DIVISION DE EDUCACION CONTINUA
FACULTAD DE INGENIERIA U.N.A.M.

CURSO: INGENIERIA MARITIMA. MODULO: "FUNDAMENTOS DE HIDRAULICA Y MATEMATICAS" DEL 11 DE MARZO AL 7 DE JUNIO.
MEXICO, D.F.

INTRODUCTION.

PROF. ING. JOSE RAYNAL.
JUNIO 1985.

1. Introduction

HYDROLOGY VERY broadly defined is the study of water. An Ad Hoc Panel on Hydrology of the Federal Council for Science and Technology (1962) defined hydrology as

> ... the science that treats of the waters of the Earth, their occurrence, circulation, and distribution, their chemical and physical properties, and their reaction with their environment, including their relation to living things. The domain of hydrology embraces the full life history of water on the earth.

This definition is much broader than the one that would be arrived at through the study of most books on hydrology. Traditionally, hydrologists have concerned themselves with the "rainfall-runoff process" and have devoted relatively little time to many other aspects of hydrology. This situation is changing, however, as many studies on all aspects of the hydrologic cycle have been undertaken in the past two decades. Most of the examples in this book deal with problems that arise in the study of the rainfall-runoff process. The reason for this is not that statistical methods cannot be applied profitably in other areas of hydrology but rather because most work and most statistical applications in hydrology have been in the rainfall-runoff area. As other aspects of hydrology are more fully developed, statistical tools will be applied to them as well.

The emphasis in the 1960's and 1970's on the quality of the environment and the preservation of the balances of nature has broadened the science of hydrology until it truly embraces the full life history of water on the Earth. This broadening of the science of hydrology has presented many new challenges to hydrologists and makes it even more imperative that a variety of tools be available to aid in the analysis of hydrologic problems. All hydrologic problems are visualized and analyzed through the use of a model. The model may be a mental conceptualization; an empirical relationship; a physical de-

vice, or a collection of mathematical, statistical, and/or empirical statements.

Most quantitative hydrologic models can be classified as deterministic, parametric, stochastic, or a combination of these. The division between these three basic types of models is not at all distinct. One can think of models as being made up of some combination of components each of which represents a point on a continuous spectrum of model "types" ranging from completely deterministic on the one hand to completely stochastic on the other.

A completely deterministic model would be one arrived at through consideration of the underlying physical relationships and would require no experimental data for its application. A parametric model may be thought of as deterministic in the sense that once model parameters are determined, the model always produces the same output from a given input. On the other hand, a parametric model is stochastic in the sense that parameter estimates depend on observed data and will change as the observed data changes. A stochastic model is one whose outputs are predictable only in a statistical sense. With a stochastic model, repeated use of a given set of model inputs produces outputs that are not the same but follow certain statistical patterns.

Most hydrology books are based on models that fall near the center of the model spectrum in that parametric, empirical models are emphasized. Eagleson's (1970) treatment of hydrology falls toward the deterministic end of the model spectrum in that emphasis is placed on the underlying theory of the phenomena being described. The treatment found in this book lies toward the stochastic end of the model spectrum in that the statistical and apparently random behavior of hydrologic quantities is emphasized.

The digital computer has made possible great advances in all types of hydrologic models. These advancements are noteworthy for both stochastic and deterministic models and have led some hydrologists to religiously adopt the philosophy that all hydrologic problems should be attacked stochastically and some the philosophy that they should be attacked deterministically. The purpose of this book is not to promote the stochastic approach but to present some basic statistical concepts that have been found useful as aids for the solution of hydrologic problems.

Many hydrologic problems can best be solved through the joint application of stochastic and deterministic methods. For instance, in the future it may be possible to deterministically predict the runoff hydrograph from a simple watershed given the rainfall input. It is unlikely, however, that rainfalls that will occur during the life of a water resources project will be deterministically predictable. Thus, one approach to project evaluation would be a statistical simulation of rainfall, deterministic conversion of the rainfall to streamflow, and a probabilistic analysis of the resulting streamflows.

Regardless of the type of stochastic model that is used, model parameters must be estimated in some way from observed hydrologic data. The validity and applicability of a stochastic model depend directly on the characteristics of the data used to estimate model parameters. A stochastic model can be no better than the data available for parameter estimation. The data used for parameter estimation must be representative of the situation in which the model is going to be used. Obviously if one is attempting to model streamflow from an urban area, model parameters cannot be estimated from forested watersheds. Similarly future hydrologic behavior of a watershed can be modeled stochastically only if available historical data are representative of future conditions. If drastic land use changes are to be made, then the model parameters must be adjusted accordingly.

In selecting data for model parameter estimation, it is important to establish that the data are homogeneous over time or can be adjusted for any nonhomogeneities that

may be present. If anything has occurred to cause a change in the characteristics of analyzed, the data must either be adjusted to account for the change or analyzed in t sections — one before the change and one after.

Some common causes of nonhomogeneity are relocating gages, especially raingag diverting streamflows, construction of dams, watershed changes such as urbanization deforestation, stream channel modifications and possibly weather modification as w as natural events of a catastrophic nature such as earthquakes, hurricane floods, etc. some instances the data can be corrected for changes. One possible adjustment would by reverse reservoir routing to determine what streamflows would have been had a res voir not been constructed. Some changes such as gradual urbanization of a watershed difficult to correct for.

The statement that the data must be representative means, for example, that d from only unusually wet or dry periods should not be used alone as this will bias the sults of the analysis. If there are only a few years of record available for analysis, t chances are good that the data are not representative of the long term variability th actually exists. Throughout this book it is assumed that the data being considered homogeneous and representative.

The concept of the return period of hydrologic events needs to be understo before proceeding. The return period is defined as the average elapsed time betwe occurrences of an event with a certain magnitude or greater. For example, a 25-y peak discharge is a discharge that is equaled or exceeded on the average once every years over a long period of time. It does not mean that an exceedance occurs every years, but that the average time between exceedances is 25 years. An exceedance is event with a magnitude equal to or greater than a certain value.

Sometimes the actual time between exceedances is called the recurrence interv With this definition for recurrence interval, the average recurrence interval for a cert event is equal to the return period of that event. In this book, recurrence interval is us in the same sense as return period.

Of course, the concept of return period can also be applied to low flows, drough shortages, etc. In this case the return period would be the average time between ever with a certain magnitude or less. Such an event might still be called an exceedance in t sense that the severity of a drought exceeds some preset level.

Regardless of whether the return period is referring to an event greater than so value or to an event less than some value, the return period can be related to a probab ity of an exceedance. If an exceedance occurs on the average once every 25 years, th the probability or chance that the event occurs in any given year is 1/25 = .04 or 4 p cent. Probability, p, and return period, T, are thus related by

$$T = \frac{1}{p} \qquad (1.$$

This is a fundamental definition in statistical hydrology.

The concept of a random sample is used throughout this book. A sample might thought of as a collection of objects selected from a larger collection of these same o jects. The larger collection of objects, if it contains all of the objects possible, is call the population. For example, 20 years of peak flow data from a certain river is a sam of the possible peak flows on the river. A random sample is one that is selected in su a fashion that any other sample could have resulted with equal likelihood. If the years of peak flow data are considered a random sample, then one is assuming that the 20 years of data are just as likely as any other possible 20 years of data and vise ver

In some types of analyses it is assumed that the order of occurrence of the data is not important, only the data values are important. The traditional hydrologic frequency analysis is an example of this. If a sample contains elements that are independent of each other, then the order of occurrence of the data is not important. This is the same as saying that the magnitude of an element in the sample is not affected by the magnitude of the other elements in the sample. Each element in the sample might be thought of as a random sample of size 1.

On the other hand there are situations where the order of occurrence of the events is important. In designing a storage reservoir to meet projected water demands, the fact that low flows tend to follow low flows makes it necessary to have a larger reservoir than would be required if the low flows occurred randomly throughout time. This is known as persistence and indicates the elements of the sample are not independent of each other. In this case the entire sequence of data values must be considered the random sample. That is, the sequence contained in the sample is assumed to be as likely as any other sequence. The individual events in the sample are not independent.

If one wanted a random sample consisting of 7 observations of daily flows on a river during a particular year, the daily flows in a particular week of that year could not be used. This is because the flow on the second, third, etc. day of the week would be dependent on the flows on the preceding days. The flow on day 2, for example, would not represent all possible daily flows but would be highly dependent on the flow during day 1. To get a random sample of daily flows, each of the 365 daily flows would have to have an equal chance of being selected. The sample of flows during the 7 consecutive days could be considered as a random sample of size 1 of weekly flows (if the week was randomly selected) but not a random sample of size 7 of daily flows.

In any hydrologic data there are errors of various kinds. The errors include measurement errors, data transmittal errors, processing errors and others. The errors may be systematic errors and show up as a bias in the data or they may be random errors. In most error analyses it is assumed that the errors are random errors and follow the normal distribution. The treatment of hydrologic data contained in this book is not concerned so much with these types of errors as it is with sampling errors.

Sampling error is a misnomer in that there are no errors in the usual sense involved. Sampling errors should more properly be called sampling variability, sampling fluctuation, or sample uncertainty. What is meant by sampling error is simply that a random sample has statistical properties that are similar to the population parameters but only equal to the population properties as the sample size gets very large (or the entire population is sampled). If two samples are selected from the same population, their statistical properties will again be similar but equal to each other only as the sample size gets very large. This variability or uncertainty in the statistical properties of a population based on estimates of the properties from samples is called "sampling error". It is clear that errors in the sense of mistakes or faulty data or carelessness are not involved in sampling errors. Sampling error is simply an inherent property of random samples. If it weren't for sampling errors, this book or hundreds of others on statistics would not be needed since populations would then be completely specified by any sample from that population.

Example 1.1. The mean annual suspended sediment load for the Green River near Munfordville, Kentucky, can be estimated from the data contained in Appendix C. This data and the resulting estimated mean annual suspended sediment load may contain many types of errors.

Systematic errors could result if the flow was sampled for sediment only when the depth of flow exceeded a preset stage. This is because low flows would not be sampled. Generally the sediment concentration in low flows is less than that in higher flows. Thus a built in bias or systematic error is produced.

Measurement errors could result from plugged samplers, samplers not properly aligned with the direction of flow, allowing the sampler to pick up some bed load, and a number of other reasons.

Data transmittal errors and processing errors can result from mistakes in transcribing data from data forms, punching data in the wrong columns on computer cards, illegibly written data, and other sources.

Sampling error can be illustrated by assuming that the tabulated data are exactly correct (contain no systematic, measurement, transmittal or processing errors). If the mean annual suspended sediment load is calculated for each successive five-year period, the results are 640,827, 484,739, 497,604 and 460,392 tons per year. Under the no-error assumption four different values of the mean annual suspended sediment discharge have been calculated each of which contains no errors yet none of which are the same. The difference in the four estimates is caused by natural variability in the phenomena (sediment) being sampled. This difference is called sampling error. If conditions on the watershed contributing to the Green River near Munfordville never changed and if the climatic conditions do not change, then the sampling error can be made as small as desired by an increase in the sample size above the five years used in this illustration.

Thus much of the statistical machinery discussed in this book is concerned with sampling errors and the estimation of population characteristics from samples of data. The fact that sampling errors are inherent in random data does not mean, however, that statistical manipulations and sophistication can in any way overcome faulty data. The quality of any statistical analysis is no better than the quality of the data used — it can be worse but no better. Furthermore, statistical considerations should not be used to replace judgement and careful thought in analyzing hydrologic data. In many instances some intelligent thought is worth reams of computer output based on a statistical analysis of some data. Statistics should be regarded as a tool, an aid to understanding, but never as a replacement for useful thought.

Rarely will one find a hydrologic problem that exactly fulfills all of the requirements for the application of one statistical technique or another. Two choices are thus available. One can redefine the problem so that it meets the requirement of the statistical theory and thus produce an "exact" answer to the artificial problem. The second approach is to alter the statistical technique where possible and then apply it to the real problem realizing that the results will be an approximate answer to the real problem. In this case the degree of the approximation depends on the severity of the violated assumptions. This latter approach is preferable and requires knowledge of available statistical techniques, of assumptions and theory underlying the techniques, and of the consequences of violating the assumptions. It is toward this latter approach that this book is oriented.

Most of the examples and exercises used in this book were selected for pedagogical reasons — not to promote a particular technique. Thus when a problem involves fitting a normal distribution to annual peak flow, the purpose of the problem revolves around learning about the normal distribution and is not to demonstrate that a normal distribution is applicable to peak flows. Similarly many examples and problems had to be simplified so that they could be realistically solved with attention being focused on the statistical technique and not the many fascinating intricacies of most real problems. That is not to say the techniques do not apply to real problems — quite the contrary. However,

most real problems involve multiple aspects, lots of data, and many considerations other than statistical. Rather than get involved in these other important aspects, many of the examples and problems are idealizations of real situations.

Since the exercises were selected as a learning aid, it will be instructive to at least read the problems at the end of each chapter. Many of the problems present useful results that supplement the material in that chapter.

Many actual problems in hydrology require considerable computation. Digital computers are used for this purpose. Special statistical-numerical procedures have been developed to simplify the computations involved and improve the accuracy of the results obtained from many of the analyses presented in this book. These procedures are not presented here. Rather the emphasis is on the principles involved. The multivariate techniques of chapters 10, 12, 14 and 15 often require extensive calculation and considerable efficiency is gained by using special purpose programs incorporating numerical shortcuts and safeguards against roundoff errors.

Finally there are many important areas of statistical analysis applicable to hydrology that are not included in this book. These omitted techniques for the most part require knowledge of the material contained in this book before they can be applied. For example Bayesian Decison Theory, which is not covered, requires a thorough knowledge of much of the material that is covered. Thus this book is an introduction to statistical methods in hydrology. Furthermore the book is not intended as a handbook or statistical "cookbook" for hydrologists. The purpose of this book is to enable the reader to better apply statistical methods to hydrologic problems through a knowledge of the methods, their foundations and limitations.

2. Probability and Probability Distributions—Basic Concepts

HYDROLOGIC PROCESSES may be thought of as stochastic processes. Stochastic in this sense means involving a variate at each instant of time where the variate is a variable that may take on any of the values of a specified set with a certain probability. An example of a stochastic hydrologic process is the annual maximum daily rainfall over a period of several years. Here the instant of time would be one year, the variable would be the maximum daily rainfall for each year and the specified set would be the set of positive numbers.

The instantaneous maximum peak flow observed during a year would be another example of a stochastic hydrologic process. Table 2.1 contains such a listing for the Kentucky River near Salvisa, Kentucky. By examining this table it can be seen that there is some order to the values yet a great deal of randomness exists as well. Even though the peak flow for each of the 66 years is listed, one cannot estimate with certainty what the peak flow for 1961 was. From the tabulated data one could surmise that the 1961 peak flow was "probably" between 20,600 cfs and 115,000 cfs. We would like to be able to estimate the magnitude of this "probably". The stochastic nature of the process, however, means that one can never estimate with certainty the exact value for the process (peak discharge) based solely on past observations.

The definition of stochastic given above has some theoretical drawbacks as we shall see. Hydrologic processes are continuous processes. The probability of realizing a given value from a continuous probability distribution is zero. Thus the probability that a variable will take on a certain value from a specified set is zero, if the variable is continuous. Practically this presents no problem since we are generally interested in the probabilities that the variate will be in some range of values. For instance we are generally not interested in the probability that the flow rate in a stream will be exactly 1000 cfs but may desire to estimate the probability that the flow will exceed 1000 cfs or be less than 1000 cfs or between 950 and 1200 cfs.

With this introduction several concepts such as probability, continuous and probability distribution have been introduced. We will now define these concepts and others

Table 2.1 Peak discharges (cfs), Kentucky River, near Salvisa, Kentucky (McCabe 1962)

1895	47,300	1917	111,000	1939	84,300
96	54,400	18	71,700	40	45,000
97	87,200	19	96,100	41	28,400
98	65,700	20	92,500	42	46,000
99	91,500	21	34,100	43	80,400
1900	53,500	22	69,000	44	55,000
01	67,800	23	73,400	45	72,900
02	70,000	24	99,100	46	71,200
03	66,900	25	79,200	47	46,800
04	34,700	26	62,600	48	84,100
05	58,000	27	93,700	49	61,300
06	47,000	28	68,700	50	87,100
07	66,300	29	80,100	51	70,500
08	80,900	30	32,300	52	77,700
09	80,000	31	43,100	53	44,200
10	52,300	32	77,000	54	20,600
11	58,000	33	53,600	55	85,000
12	67,200	34	70,800	56	82,900
13	115,000	35	89,400	57	88,700
14	46,100	36	62,600	58	60,200
15	52,400	37	112,000	59	40,300
16	94,300	38	44,000	60	50,500

as a basis for considering statistical methods in hydrology.

PROBABILITY

In the mathematical development of probability theory, the concern has been not so much how to assign probability to events, but what can be done with probability once these assignments are made. In most applied problems in water resources, one of the most important and difficult tasks is the initial assignment of probability. Considerable time is devoted in this book to determining the probability that should be assigned to events.

The definition of "probability" has been labored over for many years. One definition that is easy to grasp is the classical or a priori definition:

> If a random event can occur in n equally likely and mutually exclusive ways, and if n_a of these ways have an attribute A, then the probability of the occurrence of the event having attribute A is n_a/n written as
>
> $$\text{prob}(A) = n_a/n \qquad (2.1)$$

This definition is an a priori definition because it assumes that one can determine before the fact all of the equally likely and mutually exclusive ways that an event can occur and all of the ways that an event with attribute A can occur. The definition is somewhat circular in that equally likely is another way of saying equally probable and

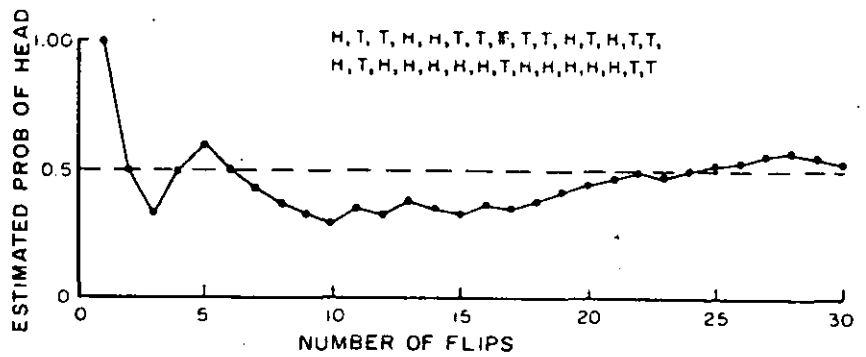


Fig. 2.1 Coin flipping experiment.

we end up using the word probable to define probability.

The classical definition of probability takes on more utility in hydrology in terms of relative frequencies and limits.

> If a random event occurs a large number of times n and the event has attribute A in n_a of these occurrences, then the probability of the occurrence of the event having attribute A is
>
> $$\text{prob}(A) = \lim_{n \to \infty} n_a/n \qquad (2.2)$$

The relative frequency concept of probability is the source of the relationship given in Chapter 1 between the return period T of an event and its probability of occurrence.

These two definitions of probability can be illustrated by considering the probability of getting a "head" in a single flip of a coin. If we know a priori that the coin is balanced and not biased toward "heads" or "tails", we can apply the first definition. There are 2 possible outcomes - heads or tails - so n is 2. There is one outcome with a head so n_a is 1. Thus the probability of a head is ½. If the coin is not balanced so that the two outcomes are not equally likely, we could not use the a priori definition. We had to know the answer to our question before we could apply the a priori definition.

This is not the case when the relative frequency definition is used. Obviously we cannot flip the coin an infinite number of times. We have to resort to a finite sample of flips. Figure 2.1 shows how the estimate of the probability of a head changes as the number of trials (flips) changes. A trend toward ½ is noted. This is called stochastic convergence towards ½. One question that might be asked is, "is the coin unbiased?" One's initial reaction is that more trials are needed. This is the plight of the hydrologist. He many times needs more trials or observations but does not have them and cannot get them. Still the data does not clearly indicate a single answer. This is where probability and statistics come into play.

Equation 2.2 allows us to estimate probabilities based on observations and does not require that outcomes be equally likely or that they all be enumerated. This advantage is somewhat offset in that estimates of probability based on observations are empirical and will only stochastically converge to the true probability as the number of observations becomes large. If two independent sets of observations are available (samples), an estimate of the probability of the event A could be determined from each set of observa-

tions. These two estimates of prob(A) would not necessarily equal each other nor would either estimate necessarily equal the true (population) prob(A) based on an infinitely large sample. This dilemma results in an important area of concern to hydrologists - how many observations are required to produce "acceptable" estimates for the probabilities of events?[1]

From either equation 2.1 or 2.2 it can be seen that the probability scale ranges from zero to one. A probability of zero means "nearly" impossible and a probability of one means "almost" certain.[2] Sometimes probability is expressed as a percent chance with a scale ranging from 0 to 100 percent. Care must be taken to insure that one does not confuse the percent chance values with true probabilities. A probability of 1 is very different from a 1 percent chance of occurrence as the former implies the event will "almost" certainly happen while the latter means it will happen only 1 time in 100.

In mathematical statistics and probability, set and measure theory are used in defining and manipulating probabilities. Consider that an experiment is any process that generates values of random variables. All possible outcomes of an experiment constitute the sample space. Any particular point in the sample is a sample point or element. An event is a collection of elements known as a set.

Example 2.1. The discharge of the Kentucky River near Salvisa over the period of a year can be considered as an experiment. The sample space representing all possible outcomes of the experiment would be the positive numbers. The possible annual maximum discharge during the year could be considered as elements or sample points in the sample space. All annual discharges above 100,000 cfs could represent an event.

To each element in the sample space of an experiment a non-negative weight is assigned such that the sum of the weights on all of the elements is one. The magnitude of the weight is proportional to the likelihood that the experiment will result in a particular element. If an element is quite likely to occur, that element would have a weight of near one. If an element was quite unlikely to occur, that element would have a weight of near zero. For elements outside the sample space, a weight of zero is assigned.

The weights assigned to the elements of the sample space are known as probabilities. Letting S represent the sample space, E_i for $i = 1, 2, \ldots$, represent elements in S, A and B represent events in S, and prob(E_i) represent the probability of E_i, it follows that

$$0 \leq \text{prob}(E_i) \leq 1 \tag{2.3}$$

Since the sample space is made up of the totality of elements in S, we have

$$S = \cup_i E_i$$

1. The term "acceptable" is left purposely vague at this point.

2. As shall be shown later, the probability of a particular value of a random variable being drawn from a continuous probability distribution is zero. Yet in sampling some value must be selected. Therefore even though the probability is zero, the event does happen. Similarly the probability of not getting that value is 1, yet the event of not getting the value does not happen. This is why the adjectives "nearly" and "almost" are used with impossible and certain.

and

$$\text{prob}(S) = \Sigma_i \text{ prob}(E_i) = 1 \tag{2.4}$$

An event is made up of a subset of elements in S so that

$$A = \cup_{i=m}^{n} E_i$$

and

$$0 \leq \text{prob}(A) = \Sigma_{i=m}^{n} \text{ prob}(E_i) \leq 1 \tag{2.5}$$

These concepts are illustrated in figure 2.2 as a Venn diagram.

Using notation from set theory and Venn diagrams, several probabilistic relationships can be illustrated. If A and B are two events in S, then the probability of A or B shown as the shaded area of figure 2.3 is given by

$$\text{prob}(A \cup B) = \text{prob}(A) + \text{prob}(B) - \text{prob}(A \cap B) \tag{2.6}$$

Note that in probability the word "or" means "either or both". The notation $\cup$ represents a union so that $A \cup B$ represents all elements in A or B or both. The notation $\cap$ rep-

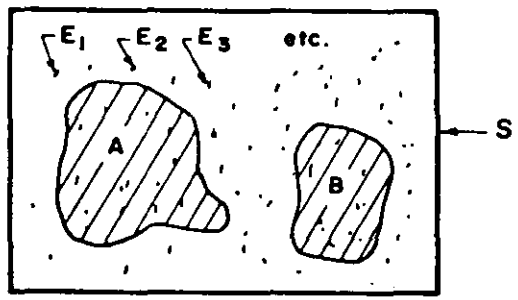


Fig. 2.2. Venn diagram illustrating a sample space, elements and events.

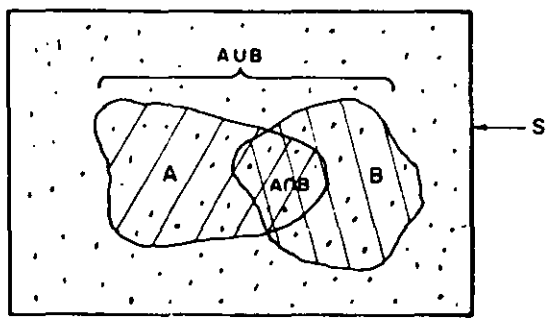


Fig. 2.3. Venn diagram showing $A \cup B$ and $A \cap B$.

resents an intersection so that $A \cap B$ represents all elements in both A and B. The last term of equation 2.6 is needed since prob(A) and prob(B) both include prob $(A \cap B)$. Thus prob$(A \cap B)$ must be subtracted once so the net result is only one inclusion of prob $(A \cap B)$ on the right hand side of the equation. If A and B are mutually exclusive, then both cannot occur and prob$(A \cap B) = 0$. In this case

$$\text{prob}(A \cup B) = \text{prob}(A) + \text{prob}(B)$$

Figure 2.2 illustrates the case where event A and B are mutually exclusive while figure 2.3 shows A and B when they are not mutually exclusive.

If A^c represents all elements in the sample space S that are not in A, then

$$A \cup A^c = S$$

and from equation 2.4 we have that

$$\text{prob}(A \cup A^c) = 1$$

This statement says that the probability of A or A^c is certainty since one or the other must occur. All of the possibilities have been exhausted. Since A and A^c are mutually exclusive

$$\text{prob}(A \cup A^c) = \text{prob}(A) + \text{prob}(A^c) = 1$$

or we have the very useful result that the probability of an event A occurring in a random experiment is

$$\text{prob}(A) = 1 - \text{prob}(A^c) \tag{2.7}$$

Equation 2.7 many times makes it easy to evaluate probability by first evaluating the probability an outcome will not occur. A^c is known as the complement of A.

If the probability of an event B depends on the occurrence of an event A, then we write prob$(B|A)$ and say the probability of B given that A has occurred. The prob$(B|A)$ is a conditional probability. The prob(B) is conditioned on the fact that A has occurred. Referring to figure 2.3 it is apparent that conditioning on the occurrence of A restricts consideration to A. Our total sample space is now A. The occurrence of B given that A has occurred is represented by $A \cap B$. Thus the prob$(B|A)$ is given by

$$\text{prob}(B|A) = \text{prob}(A \cap B)/\text{prob}(A) \tag{2.8}$$

assuming of course that prob$(A) \neq 0$. Equation 2.8 can be rearranged to give the probability of A and B as

$$\text{prob}(A \cap B) = \text{prob}(A)\,\text{prob}(B|A)$$

Now if prob$(B|A)$ = prob(B), we say that B is independent of A. Thus for independent events

$$\text{prob}(A \cap B) = \text{prob}(A)\,\text{prob}(B) \tag{2.9}$$

Example 2.2. Using the data shown in table 2.1 estimate the probability that a peak flow in excess of 100,000 cfs will occur in 2 successive years on the Kentucky River near Salvisa, Kentucky.

Solution: From table 2.1 it can be seen that a peak flow of 100,000 cfs was exceeded 3 times in the 66-year record. If it is assumed that the peak flows from year to year are independent, then the probability of exceeding 100,000 cfs in any one year is approximately 3/66 or 0.0455. Applying equation 2.9 the probability of exceeding 100,000 cfs in two successive years is found to be 0.0455 x 0.0455 or 0.00207.

Example 2.3. A study of daily rainfall at Ashland, Kentucky, has shown that in July the probability of a rainy day following a rainy day is 0.444, a dry day following a dry day is 0.724, a rainy day following a dry day is 0.276 and a dry day following a rainy day is 0.556. If it is observed that a certain July day is rainy, what is the probability that the next two days will also be rainy?

Solution: Let A be a rainy day 1 and B be a rainy day 2 following the intial rainy day. The probability of A is 0.444 since this is the probability of a rainy day following a rainy day.

$$\text{prob}(A \cap B) = \text{prob}(A)\,\text{prob}(B|A)$$

Now the prob$(B|A)$ is also 0.444 since this is the probability of a rainy day following a rainy day. Therefore

$$\text{prob}(A \cap B) = 0.444 \times 0.444 = 0.197$$

TOTAL PROBABILITY THEOREM

If $B_1, B_2, \ldots, B_n$ represents a set of mutually exclusive and collectively exhaustive events, one can determine the probability of another event A from

$$\text{prob}(A) = \sum_{i=1}^{n} \text{prob}(A|B_i)\,\text{prob}(B_i) \tag{2.10}$$

This is called the theorem of total probability. Equation 2.10 is illustrated by figure 2.4.

Example 2.4. It is known that the probability that the solar radiation intensity will reach a threshold value is 0.25 for rainy days and 0.80 for nonrainy days. It is also

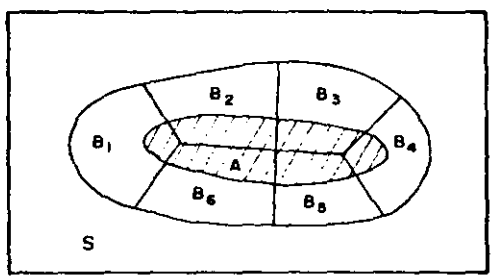


Fig. 2.4. Venn diagram for theorem of total probability.

known that for this particular location the probability of a rainy day is 0.36. What is the probability the threshold intensity of solar radiation will be reached?

Solution. Let A represent the threshold solar radiation intensity, B_1 represent a rainy day and B_2 a nonrainy day. From equation 2.10

$$\text{prob}(A) = \text{prob}(A|B_1)\,\text{prob}(B_1) + \text{prob}(A|B_2)\,\text{prob}(B_2)$$

$$= 0.25 \times 0.36 + 0.80(1 - 0.36)$$

$$= 0.602$$

BAYES THEOREM

By rewriting equation 2.8 in the form

$$\text{prob}(A)\,\text{prob}(B_j|A) = \text{prob}(B_j)\,\text{prob}(A|B_j)$$

and then substituting from equation 2.10 for prob(A), we get what is called Bayes Theorem or Bayes Rule

$$\text{prob}(B_j)\,\text{prob}(A|B_j)/\textstyle\sum_{i=1}^{n} \text{prob}(A|B_i)\,\text{prob}(B_i) = \text{prob}(B_j|A) \qquad (2.11)$$

As pointed out by Benjamin and Cornell (1970) this simple derivation of Bayes Theorem belies its importance. It provides a method for incorporating new information with previous or so-called prior probability assessments to yield new values for the relative likelihood of events of interest. These new (conditional) probabilities are called posterior probabilities. Equation 2.11 is the basis of Bayesian Decision Theory. Bayes theorem provides a means of estimating probabilities of one event by observing a second event. Such an application is illustrated in example 2.5.

Example 2.5. The manager of a recreational facility has determined that the probability he will have 1000 or more visitors on any Sunday in July depends upon the maximum temperature for that Sunday as shown in the following table. The table also gives the probabilities that the maximum temperature will fall in the indicated ranges. On a certain Sunday in July, the facility has more than 1000 visitors. What is the probability that the maximum temperature was in the various temperature classes?

temp °F T_j	prob of 1000 or more visitors	prob of being in temp class	prob of T_j | 1000 or more visitors
<60	0.05	0.05	0.005
60-70	0.20	0.15	0.059
70-80	0.50	0.20	0.197
80-90	0.75	0.35	0.517
90-100	0.50	0.20	0.197
>100	0.25	0.05	0.025

Solution: Let T_j for j = 1, 2, ..., 6 represent the 6 intervals of temperature. Then from equation 2.11

$$\text{prob}(T_j|1000 \text{ or more}) = \frac{\text{prob}(1000 \text{ or more}|T_j)\,\text{prob}(T_j)}{\sum_{i=1}^{6} \text{prob}(1000 \text{ or more}|T_i)\,\text{prob}(T_i)}$$

$$\text{prob}(T_j|1000 \text{ or more}) = \frac{\text{prob}(1000 \text{ or more}|T_j)\,\text{prob}(T_j)}{.05(.05) + .20(.15) + \ldots + .25(.05)}$$

For example

$$\text{prob}(<60F|1000 \text{ or more}) = 0.05(0.05)/0.507 = 0.005$$

Similar calculations yield the last column in the above table. Note that $\sum_{j=1}^{6}$ prob(T_j | 1000 or more) is equal to one.

COUNTING

In applying equation 2.1 to determine probabilities, one often encounters situations where it is impractical to actually enumerate all of the possible ways that an event can occur. To assist in this matter certain general mathematical formulas have been developed.

If $E_1, E_2, \ldots, E_n$ are events such that $E_i \cap E_j = 0$ for all $i \neq j$ where 0 represents an empty set and E_i can occur in n_i ways, then the compound event E made up of outcomes $E_1, E_2, \ldots, E_n$ can occur in $n_1 n_2 \ldots n_n$ ways.

The problem of sampling or selecting a sample of r items from n items is commonly encountered. Sampling can be done with replacement so that the item selected is immediately returned to the population or without replacement so that the item is not returned. The order of sampling may be important in some situations and not in others. Thus we may have four types of samples - ordered with replacement, ordered without replacement, unordered with replacement and unordered without replacement.

In case of an ordered sample with replacement, the number of ways of selecting item 1 is n since there are n items in the population. Similarly item 2 can be selected in n ways since the first item selected is returned to the population. Extending this argument to r selections, the number of ways of selecting r items from n with replacement and order being important is simply n^r.

If the items are not replaced after being selected, then the first item can be selected in n ways, the second in n-1 ways and so on until the r^{th} item can be selected in n-r+1 ways. Thus r ordered items can be selected from n without replacement in n(n-1)(n-2) ...(n-r+1) ways. This is commonly written as $(n)_r$ and called the number of permutations of n items taken r at a time.[3]

$$(n)_r = n(n-1)(n-2)\ldots(n-r+1) = n!/(n-r)! \qquad (2.12)$$

Unordered sampling without replacement is similar to ordered sampling without replacement except in the case of ordered sampling the r items selected can be arranged in r! ways. That is, an ordered sample of r items can be selected from r items in $(r)_r$ or r! ways. Thus r! of the ordered samples will contain the same elements. The number of

3. n! = n(n-1)(n-2) ... (2)(1) and is called n factorial. By definition $0! = 1$.

different unordered samples is therefore $(n)_r/r!$ commonly written as $\binom{n}{r}$ and called the binomial coefficient. The binomial coefficient gives the number of combinations possible when selecting r items from n without replacement.

$$\binom{n}{r} = (n)_r/r! = n!/(n-r)!r! \qquad (2.13)$$

Table E.1 contains values for the binomial coefficient.

Finally, in unordered sampling with replacement, selecting r items from n with replacement is equivalent to selecting r items from n+r-1 items without replacement. That is we can consider the population as containing r-1 items more than it really does since the items selected will be replaced. The number of ways of selecting r unordered items from n items is then $\binom{n+r-1}{r} = (n+r-1)!/(n-1)!r!$.

The number of ways of selecting samples under the four above conditions is summarized in the following table.

	With Replacement	Without Replacement
Ordered	n^r	$(n)_r = n!/(n-r)!$
Unordered	$\binom{n+r-1}{r} = (n+r-1)!/(n-1)!r!$	$\binom{n}{r} = n!/(n-r)!r!$

Factorials can be approximated by using Stirling's formula $n! \approx \sqrt{2\pi}\, e^{-n} n^{n+1/2}$. The error in this approximation is less than one percent for n equal to 10 and the percentage error decreases as n increases.

Example 2.6. For a particular watershed, records from 10 raingages are available. Records from 3 of the gages are known to be bad. If 4 records are selected at random from the 10 records, (a) What is the probability that 1 bad record will be selected? (b) What is the probability that 3 bad records will be selected? (c) What is the probability that at least 1 bad record will be selected?

Solution: The total number of ways of selecting 4 records from the 10 available records (order is not important) is

$$\binom{10}{4} = 10!/6!4! = 210$$

(a) The number of ways of selecting 1 bad record from 3 bad records and 3 good records from 7 good records is

$$\binom{3}{1}\binom{7}{3} = 3!/2!1! \times 7!/4!3! = 105$$

Applying equation 2.1 and letting a be 1 bad and 3 good records, the probability of a is 105/210 or 0.500.

(b) The number of ways of selecting 3 bad records and 1 good record is

$$\binom{3}{3}\binom{7}{1} = 1 \times 7 = 7$$

Thus the probability of selecting 3 bad records is 7/210 or 0.033.

(c) The probability of at least 1 bad record is equal to the probability of 1 or 2 or 3 bad records. The probability of 1 and 3 bad records is known to be 0.500 and 0.033 respectively. The probability of 2 bad records is

$$\binom{3}{2}\binom{7}{2}/210 = 3 \times 21/210 = 0.300$$

Thus the probability of at least 1 bad record is 0.500 + 0.300 + 0.033 = 0.833.

This latter result could also be determined from the fact that the probability of 0 or 1 or 2 or 3 bad records must equal one. The probability of at least 1 bad record thus equals

$$1 - \text{prob(0 bad records)} =$$

$$1 - \binom{3}{0}\binom{7}{4}/210 = 1 - 35/210 = 0.833$$

Example 2.7. For the situation described in Example 2.6, what is the probability of selecting at least 2 bad records given that one of the records selected is bad?

$$\text{prob(at least 2 bad out of 4} \mid \text{1 is bad)} =$$

$$\text{prob(at least 2 bad out of 4)/prob(1 bad in 4)}$$

$$= (0.300 + 0.033)/0.500$$

$$= 0.667$$

GRAPHICAL PRESENTATION

Hydrologists are often faced with large quantities of data requiring analysis. Since it is difficult to grasp the total data picture from tabulations such as table 2.1, a useful first step in data analysis is to plot the data as a frequency histogram. This is done by grouping the data into classes and then plotting a bar graph with the number or relative frequency of observations in a class versus the midpoint of the class interval. The midpoint of a class is called the class mark. The class interval is the difference between the upper and lower class boundaries.

In assigning data to classes, the question frequently arises as to what to do with data that falls on a class boundary. One thing that can be done is to define the boundary to more significant figures than the actual data. For instance the data of table 2.1 are to the nearest 100 cfs. By selecting the class boundaries to end in 50 cfs (i.e., 19,950, 29,950, ...), one is assured that a data value will not fall on the boundary. Another possibility is to always assign data falling on a boundary to the next higher (lower) class or to alternately assign the data to the next higher and then the next lower class.

The selection of the class interval and the location of the first class mark can appreciably affect the appearance of a frequency histogram. The appropriate width for a class interval depends upon the range of the data, the number of observations, and the behavior of the data. Several suggestions have been put forth for forming frequency histograms. Spiegel (1961) suggests that there should be 5 to 20 classes. Steel and Torrie (1960) state that the class interval should not exceed one-fourth to one-half of the standard deviation of the data. Sturges (1926) recommends that the number of classes be determined from

$$m = 1 + 3.3 \log n \qquad (2.14)$$

where m is the number of classes, n is the number of data values and the logarithm to the base 10 is used.

Whatever criteria is used, it should be kept in mind that sensitivity is lost if too few or too many classes are used. Too few classes will eliminate detail and obscure the basic pattern of the data. Too many classes result in erratic patterns of alternating high and low frequencies. If possible the class intervals and class marks should be round figures. This is not a computational or theoretical consideration, but one aimed at making it easier for those viewing the histogram to grasp its full meaning.

In some situations it may be desirable to use nonuniform class intervals. In Chapter 8 a situation is presented where the intervals are such that the expected relative frequencies are the same in each class.

Table 2.2 presents a tabulation of the data from table 2.1. The data are plotted in figure 2.5. In this case a class interval of 10,000 cfs was selected.

The influence of the number of class intervals can be seen by comparing figures 2.5 and 2.6. Figure 2.6 is a plot of the Kentucky River data using a class interval of 15,000 cfs and 7 classes. Equation 2.14 results in a recommendation of 7 classes for this data. Figure 2.6 is much smoother and much more suggestive of a probability distribution (as we shall see in chapters 5 and 6) than is figure 2.5.

Another common method of presenting data is in the form of a cumulative frequency distribution. Cumulative frequency distributions show the frequency of events less than (greater than) some given value. They are formed by summing the relative frequencies and plotting the accumulated sum against the corresponding data value. If the ordinates are summed from the smaller (larger) data values to the larger (smaller) values, the resulting cumulative frequency refers to the frequency of observations less (more) than the corresponding data value. Points on the cumulative curve should be plotted on the right (left) boundary. Figure 2.7 is a cumulative frequency plot based on table 2.2 and figure 2.5.

If the relative frequency of the i^{th} class of a frequency histogram is denoted by f_{x_i}, then one property of the histogram is

$$\sum_{i=1}^{m} f_{x_i} = 1 \qquad (2.15)$$

Table 2.2. Tabulation of peak flows, Kentucky River, near Salvisa, Kentucky

Class Mark	Number	Relative Frequency	Cumulative Relative Frequency
25,000	2	0.030	0.030
35,000	3	.045	.075
45,000	10	.152	.227
55,000	9	.136	.363
65,000	11	.167	.530
75,000	10	.152	.682
85,000	12	.182	.864
95,000	6	.091	.955
105,000	0	.000	.955
115,000	3	.045	1.000

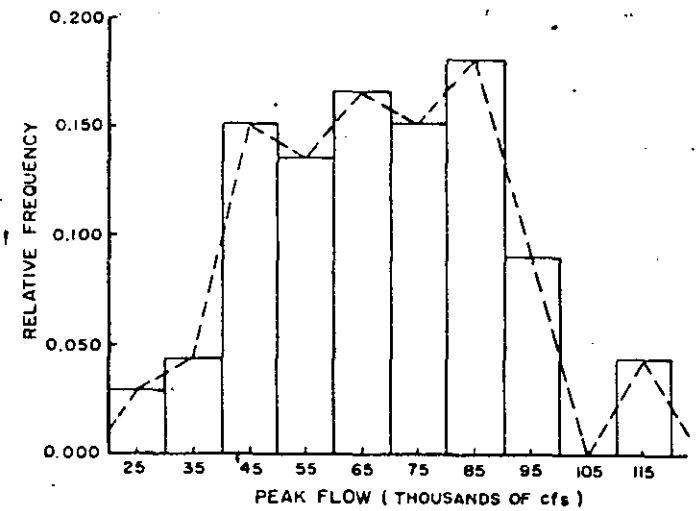


Fig. 2.5. Frequency histogram for peak flows, Kentucky River.

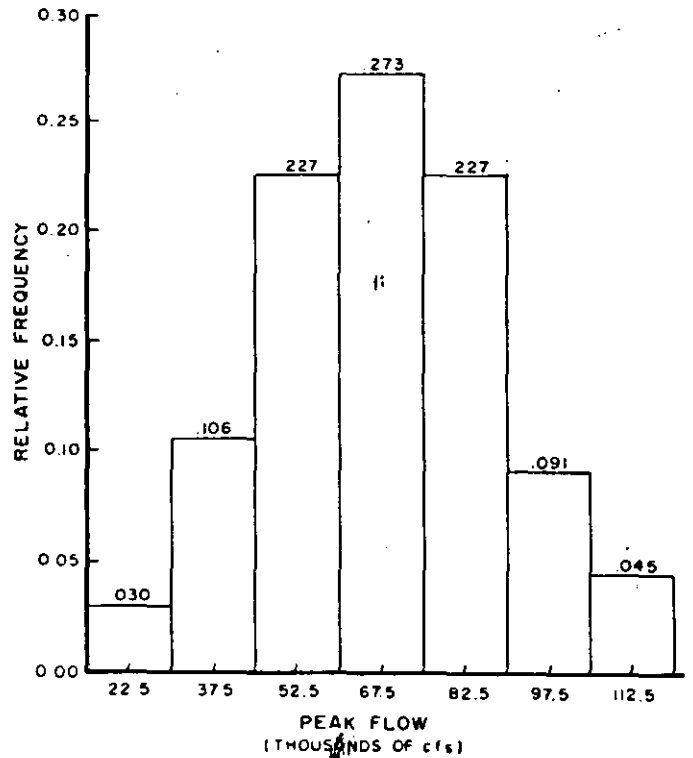


Fig. 2.6. Frequency histogram for peak flows, Kentucky River data, using a class interval of 15,000 cfs.

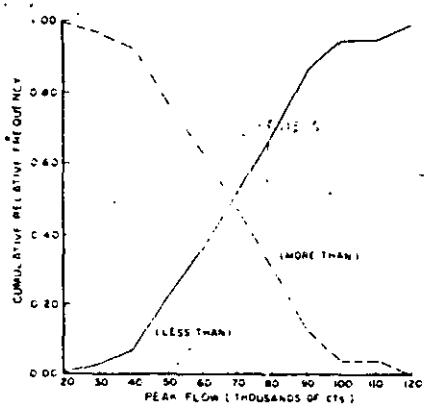


Fig. 2.7. Cumulative frequency of peak flows, Kentucky River.

where m is the number of classes. Similarly if the cumulative relative frequency at the end of the i^{th} class interval is F_{x_i}, then

$$f_{x_i} = F_{x_i} - F_{x_{i-1}} \tag{2.16}$$

and

$$F_{x_i} = \sum_{j=1}^{i} f_{x_j} \tag{2.17}$$

RANDOM VARIABLES

Simply stated a random variable is a function defined on a sample space. Random variables may be discrete or continuous. If the set of values a random variable can assume is finite (or countably infinite), the random variable is said to be a discrete random variable. If the set of values a random variable can assume is infinite, the random variable is said to be a continuous random variable. An example of a discrete random variable would be the number of rainy days experienced at a particular location over a period of 1 year. The amount of rain received over the year would be a continuous random variable. Capital letters will be used to denote random variables and the corresponding lower case letter will represent values of the random variable.

Any function of a random variable is also a random variable. If X is a random variable then Z = g(X) is a random variable as well.

Example 2.8. Every hydrologic variable can be taken as a random variable. Rainfall for any duration, streamflow, soil hydraulic properties, time between hydrologic events such as flows above a certain base or daily rainfalls in excess of some amount, the number of times a streamflow rate exceeds a given base over a period of a year and daily pan evaporation are all random variables. Quantities derived from random hydrologic variables are also random variables. The storage required in a water supply reservoir to meet a given demand is a function of the demand and the inflow to the reservoir. Since reservoir inflow is a random variable, required storage is also a random variable. As a matter of fact, the demand that is placed on the reservoir would be a random variable as well.

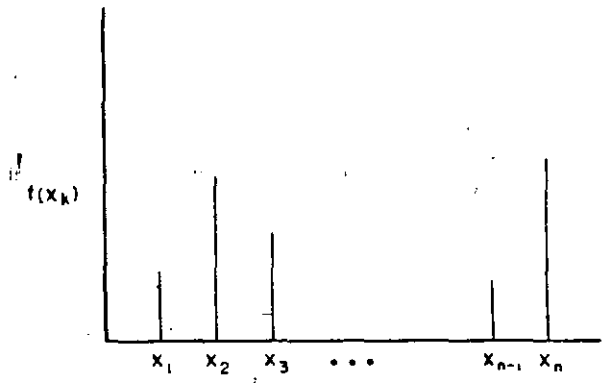


Fig. 2.8. A discrete probability distribution.

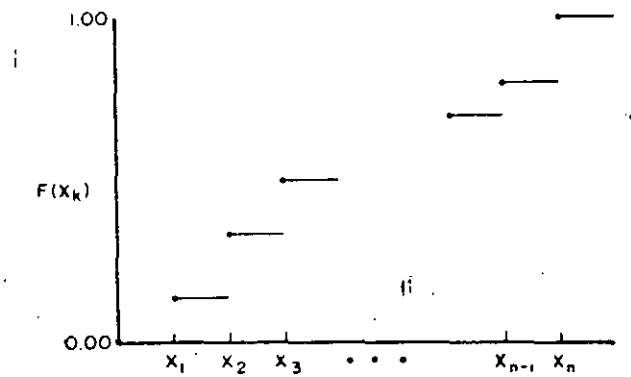


Fig. 2.9. A discrete cumulative probability distribution.

UNIVARIATE PROBABILITY DISTRIBUTIONS

If X can take on values $x_1, x_2, \ldots, x_n$ with probabilities $f_X(x_1), f_X(x_2), \ldots, f_X(x_n)$ and $\sum_i f_X(x_i) = 1$, then X is a discrete random variable. With a discrete random variable there are "lumps" or spikes of probability associated with the values that the random variable can assume. Figure 2.8 is a typical plot of the distribution of probability associated with the values that a discrete random variable can assume. This information would constitute a probability distribution or probability function for a discrete random variable.

The cumulative probability distribution, $F_X(x_k)$, for this discrete random variable is shown in figure 2.9. The cumulative distribution represents the probability that X is less than or equal to x_k.

$$F_X(x) = \sum_{x_i \le x} f_X(x_i) \tag{2.18}$$

The cumulative distribution has jumps in it at each x_i equal in magnitude to $f_X(x_i)$ or the probability that $X = x_i$. The probability that $X = x_i$ can be determined from

$$f_X(x_i) = F_X(x_i) - F_X(x_{i-1}) \tag{2.19}$$

The notation $f_X(x)$ and $F_X(x)$ denote the probability distribution and cumulative probability distribution of the random variable X evaluated at X = x.

Many times continuous data is treated as though it were discrete. The reason for this is apparent from the similarities of equations 2.17 and 2.18. Looking again at the Kentucky River peak flow data, we can define the event A as having a peak flow in the i^{th} class. Then using the notation of equation 2.2, n_a would be the number of observed peak flows in the i^{th} interval, n_i, and n would be the total number of observed peak flows. The prob(A) equals the probability that a peak flow is in the i^{th} class.

$$\text{prob}(A) = n_a/n = n_i/n = f_{x_i} \tag{2.20}$$

Thus the relative frequency, f_{x_i}, can be interpreted as a probability estimate, the frequency histogram can be interpreted as an approximation for a probability distribution, and the cumulative frequency can be interpreted as an approximation for a cumulative probability distribution.

Many times it is desirable to treat continuous random variables directly. Probability distributions of continuous random variables are smooth curves. The probability density function of a continuous random variable X is denoted by $p_X(x)$. The cumulative probability distribution function is denoted by $P_X(x)$. $P_X(x)$ represents the probability that X is less than or equal to x.[4]

$$P_X(x) = \text{prob}(X \leq x) \tag{2.21}$$

The probability density function and the cumulative probability distribution function are related by

$$d\,P_X(x) = p_X(x)\,dx \tag{2.22}$$

or

$$P_X(x) = \int_{-\infty}^{x} p_X(t)\,dt \tag{2.23}$$

The notation $p_X(x)$ and $P_X(x)$ denote the probability density function (pdf) and cumulative distribution function (cdf), respectively, of the random variable X evaluated at X = x. Thus $p_Y(a)$ represents the probability density function of the random variable Y evaluated at Y = a. $P_Y(a)$ represents the cumulative distribution function of the random variable Y and gives prob(Y ≤ a).

Any function, $p_X(x)$, defined on the real line can be a probability density function if and only if

$$\begin{aligned} &1.\ p_X(x) \geq 0 \text{ for all } x \\ &2.\ \int_{-\infty}^{\infty} p_X(x)\,dx = 1 \end{aligned} \tag{2.24}$$

4. Recall that the return period T of an event was defined as 1/p where p was the probability the event would be exceeded. We can define the return period as a function of the magnitude of the event by

$$T_X(x) = 1/[1-P_X(x)] = 1/\text{prob}(X > x)$$

From equation 2.23 we have $P_X(-\infty) = 0$ and from equation 2.24 we have $P_X(\infty) = 1$. It is also apparent that the probability that X takes on a value between a and b is given by

$$\text{prob}(a \leq X \leq b) = \int_a^b p_X(t)\,dt \tag{2.25}$$

$$= P_X(b) - P_X(a) \tag{2.26}$$

The prob(a ≤ X ≤ b) is the area under the probability density function between a and b. Earlier in the chapter it was stated that the probability of any particular value from a continuous distribution is zero. This can be seen from

$$\begin{aligned} \text{prob}(X=d) &= \int_d^d p_X(t)\,dt \\ &= P_X(d) - P_X(d) \\ &= 0 \end{aligned}$$

Since the probability that a continuous random variable takes on a specified value is zero, the expressions prob(a ≤ X ≤ b), prob(a < X ≤ b), prob(a ≤ X < b) and prob(a < X < b) are all equivalent. It is also apparent that $P_X(x)$ can be interpreted as the probability that X is strictly less than x since prob(X=x) = 0. This means as well that return periods can be defined in terms of events that equal or exceed a given value or in terms of events that strictly exceed the given value since the probability of equaling the value in the case of a continuous random variable is zero.

Figures 2.10 and 2.11 illustrate a possible probability density function and its cor-

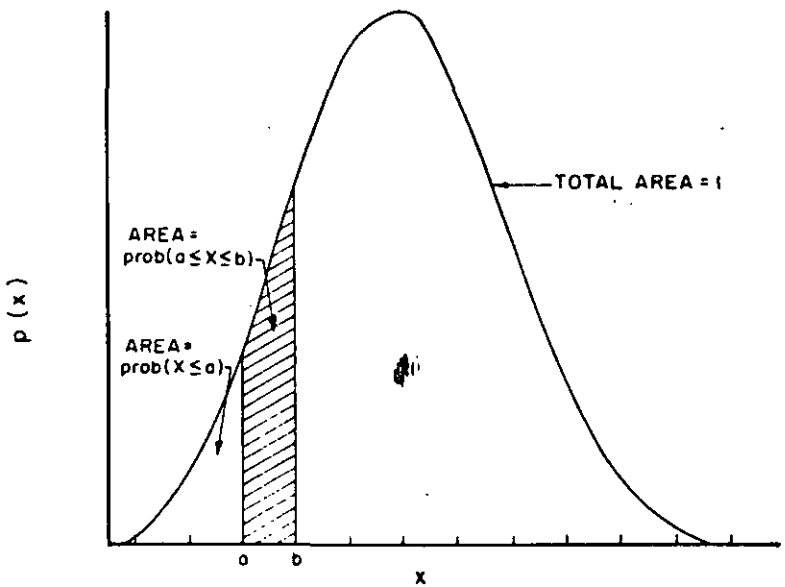


Fig. 2.10. Probability density function.

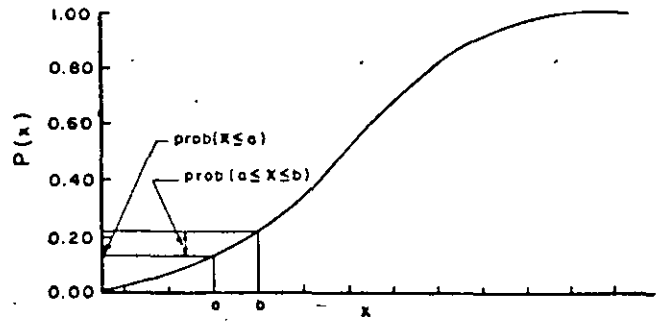


Fig. 2.11. Cumulative probability distribution function.

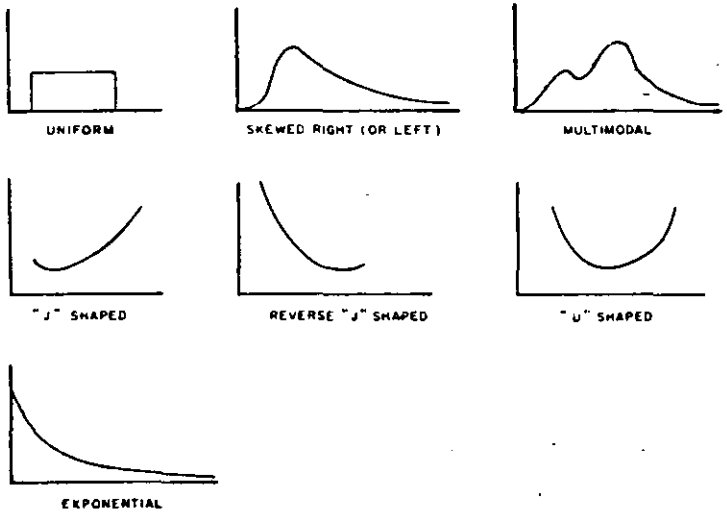


Fig. 2.12. Some possible shapes for density functions.

responding cumulative distribution function. In addition to density functions that are symmetrical and "bell-shaped", densities may take on a number of different shapes. Some of the more common shapes are shown in figure 2.12.

At this point a cautionary note is added to the effect that the probability density function, $p_X(x)$, is not a probability and can have values exceeding one. The cumulative probability distribution, $P_X(x)$, is a probability (prob$(X \leq x) = P_X(x)$) and must have values ranging from 0 to 1. Of course, $p_X(x)$ and $P_X(x)$ are related as indicated by equation 2.22 and knowledge of one specifies the other.

Example 2.9. Evaluate the constant a in the expression $p_X(x) = ax^2$ so that the following can be considered a probability density function:

$p_X(x) = ax^2, 0 \leq X \leq 5 \qquad p_X(x) = 0$, elsewhere[5]

5. Throughout this book the value of $p_X(x)$ will be taken as zero outside the range of definition of the random variable(s). This will eliminate the need for always adding the statement that $p_X(x) = 0$ elsewhere.

What is the probability that a value selected at random from this distribution will (a) be less than 2? (b) fall between 1 and 3? (c) be larger than 4? (d) be larger than or equal to 4? (e) exceed 6?

Solution: From equation 2.24 we must have

$$\int_{-\infty}^{\infty} p_X(t)\,dt = 1$$

This can be broken into

$$\int_{-\infty}^{0} p_X(t)\,dt + \int_{0}^{5} p_X(t)\,dt + \int_{5}^{\infty} p_X(t)\,dt = 1$$

The first and last term on the left hand side of this expression are both zero.

$$\int_0^5 at^2\,dt = 1$$

$$at^3/3\,\Big|_0^5 = 1$$

$$a = 3/125$$

so

$$p_X(x) = 3x^2/125$$

and

$$P_X(x) = x^3/125$$

for

$$0 \leq X \leq 5$$

a) prob$(X \leq 2) = P_X(2) = 8/125$

b) prob$(1 \leq X \leq 3) = P_X(3) - P_X(1) = 26/125$

c) prob$(X > 4) = 1 - P_X(4) = 1 - 64/125 = 61/125$

d) prob$(X \geq 4) = 61/125$ since prob$(X = 4) = 0$

e) prob$(X > 6) = 1 - P_X(6)$

$P_X(x)$ for $X > 5 = \int_{-\infty}^{5} p_X(t)\,dt$

$= \int_{-\infty}^{0} 0\,dt + \int_0^5 3t^2/125\,dt + \int_5^{\infty} 0\,dt$

$= 0 + 1 + 0 = 1$

prob$(X \geq 6) = 1 - P_X(6) = 1 - 1 = 0$

Piecewise continuous distributions satisfying the requirements for a probability distribution in which the prob(X=d) is not zero are possible. Such a distribution could be defined by

$$P_X(x) = P_1(x) \text{ for } X < d$$
$$= P_2(x) \text{ for } X \geq d \qquad (2.27)$$

Where $P_2(d) > P_1(d)$, $P_1(-\infty) = 0$, $P_2(\infty) = 1$ and $P_1(x)$ and $P_2(x)$ are nondecreasing functions of X. Figure 2.13 is a plot of such a distribution. For this situation the prob (X=d) equals the magnitude of the jump ΔP at X=d or is equal to $P_2(d) - P_1(d)$. Any finite number of discontinuities of this type are possible.

An example of a distribution as shown in figue 2.13 is the distribution of daily rainfall amounts. The probability that no rainfall is received, prob(X=0), is finite while the probability distribution of rain on rainy days would form a continuous distribution. A second example would be the probability distribution of the water level in some reservoir. The water level may be maintained at a constant level d as much as possible but may fluctuate below or above d at times. The distribution shown in figure 2.13 could represent this situation.

The relationship between relative frequency and probability can be envisioned by considering an experiment whose outcome is a value of the random variable X. Let

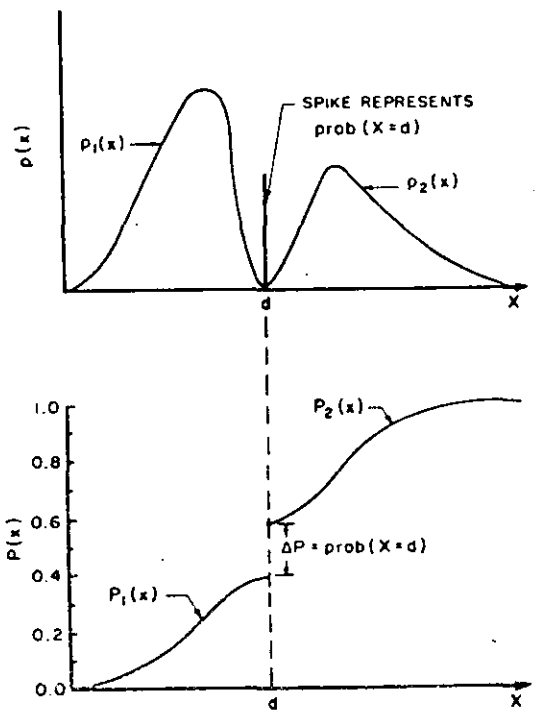


Fig. 2.13. A possible piecewise continuous probability distribution for the case prob(X=d) ≠ 0.

$p_X(x)$ be the probability density function of X. The probability that a single trial of the experiment will result in an outcome between X=a and X=b is given by

$$\text{prob}(a < X < b) = \int_a^b p_X(x)\,dx = P_X(b) - P_X(a)$$

In N independent trials of the experiment, the expected number of outcomes in the interval a to b would be

$${}_an_b = N[P_X(b) - P_X(a)]$$

and the expected relative frequency of outcomes in the interval a to b is

$${}_af_b = {}_an_b/N = P_X(b) - P_X(a)$$

In general if x_i represents the midpoint of an interval of X given by $x_i - \Delta x_i/2$ to $x_i + \Delta x_i/2$, then the expected relative frequency of outcomes of repeated, independent trials of the experiment in this interval is given by

$$f_{x_i} = P_X(x_i + \Delta x_i/2) - P_X(x_i - \Delta x_i/2) \qquad (2.28a)$$

Since the right hand side of this equation represents the area under $p_X(x)$ between $x_i - \Delta x_i/2$ and $x_i + \Delta x_i/2$, it can be approximated by

$$f_{x_i} = \Delta x_i\, p_X(x_i) \qquad (2.28b)$$

Equation 2.28 can be used to determine the expected relative frequency of repeated, independent outcomes of a random experiment whose outcome is a value of the random variable X.

If N independent observations of X are available, the actual relative frequency of outcomes in an interval of width Δx_i centered on x_i may not equal f_{x_i} as given by equation 2.28 since X is a random variable whose behavior can only be described probabilistically. The most probable outcome or the expected outcome will equal the observed outcome only if $p_X(x)$ is truly the probability density function for X and for an infinitely large number of observations. That is even if the true probability density function is being used, the actual frequency of outcomes in the interval Δx_i approaches the expected number only as the number of trials or observations becomes very large.

Example 2.10. Plot the expected frequency histogram using the probability density function of example 2.9 and a class interval of ½.

Solution: $f_{x_i} = \Delta x_i\, p_X(x_i)$

$$= p_X(x_i)/2 = 3x_i^2/250$$

The desired plot is shown in figure 2.14.

x_i	$p_X(x_i)$	f_{x_i}
.25	.00150	.00075
.75	.01350	.00675
1.25	.03750	.01875
1.75	.07350	.03675
2.25	.12150	.06075
2.75	.18150	.09075
3.25	.25350	.12675
3.75	.33750	.16875
4.25	.43350	.21675
4.75	.54150	.27075
	Sum	.99750

BIVARIATE DISTRIBUTIONS

The situation frequently arises where one is interested in the simultaneous behavior of two or more random variables. An example might be the flow rates on two streams near their confluence. One might like to know the probability of both streams having peak flows exceeding a given value.

Example 2.11. The magnitude of peak flows from small watersheds is often estimated from the "Rational Equation" given by Q=CIA where Q is the estimated flow, C is a coefficient, I is a rainfall intensity, and A is the watershed area. The assumption is made that the return period of Q will be the same as the return period of the I that is used. To verify this assumption it is necessary to study the joint probabilities of the two random variables Q and I.

If X and Y are continous random variables, their joint probability density function is $p_{X,Y}(x,y)$ and the corresponding cumulative probability distribution is $P_{X,Y}(x,y)$. These two are related by

$$p_{X,Y}(x,y) = \frac{\partial^2}{\partial x \partial y} P_{X,Y}(x,y) \tag{2.29}$$

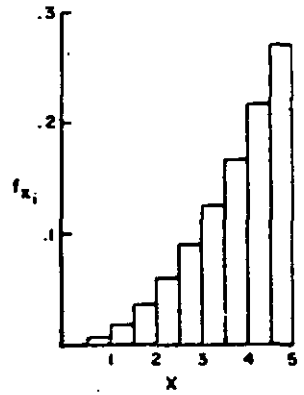


Fig. 2.14. Plot for example 2.10.

and

$$P_{X,Y}(x,y) = \text{prob}(X \leq x \text{ and } Y \leq y) = \int_{-\infty}^{x} \int_{-\infty}^{y} p_{X,Y}(t,s)\, ds\, dt \tag{2.30}$$

The corresponding relationships for X and Y being discrete random variables are

$$f_{X,Y}(x_i,y_j) = \text{prob}(X=x_i \text{ and } Y=y_j) \tag{2.31}$$

$$F_{X,Y}(x,y) = \text{prob}(X \leq x \text{ and } Y \leq y) = \sum_{x_i \leq x} \sum_{y_j \leq y} f_{X,Y}(x_i,y_j) \tag{2.32}$$

It should be noted that the bivariate analogy of equation 2.19 is

$$f_{X,Y}(x_i,y_j) = F_{X,Y}(x_i,y_j) - F_{X,Y}(x_i,y_{j-1}) - F_{X,Y}(x_{i-1},y_j) + F_{X,Y}(x_{i-1},y_{j-1}) \tag{2.33}$$

Some of the properties of continuous bivariate distributions are

1) $P_{X,Y}(x,\infty)$ is a cumulative univariate probability function of X only (the cumulative marginal distribution of X).

2) $P_{X,Y}(\infty,y)$ is a cumulative univariate probability function of Y only (the cumulative marginal distribution of Y).

3) $p_{X,Y}(x,y) \geq 0$

4) $P_{X,Y}(\infty,\infty) = 1$

5) $P_{X,Y}(-\infty,y) = P_{X,Y}(x,-\infty) = 0$

MARGINAL DISTRIBUTIONS

If one is interested in the behavior of one of a pair of random variables regardless of the value of the second random variable, the marginal distribution may be used.

For instance the marginal density of X, $p_X(x)$, is obtained by integrating $p_{X,Y}(x,y)$ over all possible values of Y.

$$p_X(x) = \int_{-\infty}^{\infty} p_{X,Y}(x,s)\, ds \tag{2.34}$$

The cumulative marginal distribution is given by

$$P_X(x) = P_{X,Y}(x,\infty) = \text{prob}(X \leq x \text{ and } Y \leq \infty) \tag{2.35a}$$

$$= \text{prob}(X \leq x) \tag{2.35b}$$

$$= \int_{-\infty}^{x} \int_{-\infty}^{\infty} p_{X,Y}(t,s)\, ds\, dt \tag{2.35c}$$

$$= \int_{-\infty}^{x} p_X(t)\, dt \tag{2.35d}$$

Similarly the marginal density and cumulative marginal distribution of Y are

$$p_Y(y) = \int_{-\infty}^{\infty} p_{X,Y}(t,y)\,dt \tag{2.36}$$

and

$$P_Y(y) = \int_{-\infty}^{y} p_Y(s)\,ds \tag{2.37}$$

The corresponding relationships for a discrete bivariate distribution are

$$f_X(x_i) = \text{prob}(X=x_i) = \Sigma_j f_{X,Y}(x_i,y_j) \tag{2.38a}$$

$$f_Y(y_j) = \text{prob}(Y=y_j) = \Sigma_i f_{X,Y}(x_i,y_j) \tag{2.38b}$$

$$F_X(x) = \text{prob}(X \le x) = \sum_{x_i \le x} f_X(x_i) = \sum_{x_i \le x} \Sigma_j f_{X,Y}(x_i,y_j) \tag{2.39a}$$

$$F_Y(y) = \text{prob}(Y \le y) = \sum_{y_j \le y} f_Y(y_j) = \sum_{y_j \le y} \Sigma_i f_{X,Y}(x_i,y_j) \tag{2.39b}$$

CONDITIONAL DISTRIBUTIONS

A marginal distribution is the distribution of one variable regardless of the value of the second variable. The distribution of one variable with restrictions or conditions placed on the second variable is called a conditional distribution. Such a distribution might be the distribution of X given that Y equals y_0 or the distribution of Y given that $x_1 \le X \le x_2$.

In general the conditional distribution of X given that Y is in some region R is arrived at using the same reasoning that was used in obtaining equation 2.8. The total sample space of Y is now the region R. Since

$$\int_{-\infty}^{\infty} \int_R p_{X,Y}(t,s)\,ds\,dt = \int_R p_Y(s)\,ds$$

then

$$\int_R p_{X,Y}(x,s)\,ds / \int_R p_Y(s)\,ds$$

represents a probability density function of X given that Y is in R. The conditional density of X given Y is in R is given by

$$p_{X|Y}(x\,|\,Y \text{ is in } R) = \int_R p_{X,Y}(x,s)\,ds / \int_R p_Y(s)\,ds \tag{2.40}$$

for X and Y continuous. Similarly the conditional distribution of $(X\,|\,Y \text{ is in } R)$ for X and Y discrete is

$$f_{X|Y}(x_i\,|\,Y \text{ is in } R) = \sum_{y_j \text{ in } R} f_{X,Y}(x_i,y_j) \Big/ \sum_{y_j \text{ in } R} f_Y(y_j) \tag{2.41}$$

The determination of conditional probabilities from equations 2.40 and 2.41 are done in the usual way.

$$\text{prob}(X \text{ is in } S\,|\,Y \text{ is in } R) = \int_S p_{X|Y}(x\,|\,Y \text{ is in } R)\,dx \tag{2.42}$$

for X and Y continuous and

$$\text{prob}(X \text{ is in } S\,|\,Y \text{ is in } R) = \sum_{x_i \text{ in } S} f_{X|Y}(x_i\,|\,Y \text{ is in } R) \tag{2.43}$$

for X and Y discrete.

For the special case where X and Y are continuous and the conditional density of X given $Y=y_0$ is desired, equation 2.40 breaks down since both the numerator and denominator become zero. In this case

$$p_{X|Y}(x\,|\,Y = y_0) = p_{X,Y}(x,y_0)/p_Y(y_0) \tag{2.44}$$

The proof of this may be found in Neuts (1973). In most books $p_{X|Y}(x\,|\,Y = y_0)$ is simply written as

$$p_{X|Y}(x\,|\,y) = p_{X,Y}(x,y)/p_Y(y) \tag{2.45}$$

and called the conditional density of X given Y. In this book when a conditional density is mentioned, an expression like equation 2.45 will be meant unless otherwise specified.

All of the above results are symmetrical with respect to X and Y. For example

$$\text{prob}(Y \text{ is in } S\,|\,X \text{ is in } R) = \int_S p_{Y|X}(y\,|\,X \text{ is in } R)\,dy \tag{2.46}$$

for X and Y continuous where

$$p_{Y|X}(y\,|\,X \text{ is in } R) = \int_R p_{X,Y}(t,y)\,dt / \int_R p_X(t)\,dt \tag{2.47}$$

If the region R of equation 2.40 is the entire region of definition with respect to Y, then

$$\int_R p_Y(s)\,ds = 1$$

and

$$\int_R p_{X,Y}(x,s)\,ds = p_X(x)$$

so that

$$p_{X|Y}(x\,|\,Y \text{ is in } R) = p_X(x)$$

This results from the fact that the condition that Y is in R when R encompasses the entire region of definition of Y is really no restriction but simply a condition stating that Y may take on any value. In this case $p_{X|Y}(x\,|\,Y \text{ is in } R)$ is identical to the marginal density of X.

INDEPENDENCE

From equation 2.40 or 2.44 it can be seen that in general the conditional density of X given Y is a function of y. If the random variables X and Y are independent, this functional relationship disappears (i.e. $p_{X|Y}(x\,|\,y)$ is not a function of y). In fact in this case

$$p_{X|Y}(x\,|\,y) = p_X(x)$$

or

$$f_{X|Y}(x_i|y_j) = f_X(x_i)$$

or the conditional density equals the marginal density. Furthermore, if X and Y are independent (continuous or discrete) random variables, their joint density is equal to the product of their marginal densities.

$$p_{X,Y}(x,y) = p_X(x)\, p_Y(y) \tag{2.48a}$$

or

$$f_{X,Y}(x_i,y_j) = f_X(x_i)\, f_Y(y_j) \tag{2.48b}$$

The random variables X and Y are independent in the probabilistic sense (stochastically independent) if and only if their joint density is equal to the product of their marginal densities.

Example 2.12. The function $p_{X,Y}(x,y) = c(5 - y/2 - x)$ for $0 \le X \le 2$ and $0 \le Y \le 2$ can serve as a bivariate continuous probability density. The value of c must be such that the prob ($X \le 2$ and $Y \le 2$) is one.

$$\text{prob}(X \le 2 \text{ and } Y \le 2) = \int_0^2 \int_0^2 c(5 - s/2 - t)\, ds\, dt = 1$$

Evaluating this equation results in $c = 1/14$. The cumulative bivariate distribution is given by

$$P_{X,Y}(x,y) = \text{prob}(X \le x, Y \le y) = \int_0^x \int_0^y (5 - s/2 - t)/14\, ds\, dt$$

or

$$P_{X,Y}(x,y) = (5xy - xy^2/4 - x^2 y/2)/14$$

The probability that $X \le 1$ and $Y \le 1$ can be evaluated from $P_{X,Y}(1,1)$ as 0.304. The marginal densities are given by

$$p_X(x) = \int_0^2 p_{X,Y}(x,s)\, ds = \int_0^2 (5 - s/2 - x)/14\, ds = (9 - 2x)/14$$

and

$$p_Y(y) = \int_0^2 p_{X,Y}(t,y)\, dt = \int_0^2 (5 - y/2 - t)/14\, dt = (8 - y)/14$$

The cumulative marginal distributions are

$$P_{X,\infty}(x,\infty) = \text{prob}(X \le x) = \int_0^x p_X(t)\, dt$$

$$= \int_0^x (9 - 2x)/14\, dt = (9x - x^2)/14$$

and

$$P_{\infty,Y}(\infty,y) = \text{prob}(Y \le y) = \int_0^y p_Y(s)\, ds$$

$$= \int_0^y (8 - y)/14\, ds = (8y - y^2/2)/14$$

The conditional densities

$$p_{X|Y}(x|y) = p_{X,Y}(x,y)/p_Y(y) = (5 - y/2 - x)/(8 - y)$$

and

$$p_{Y|X}(y|x) = p_{X,Y}(x,y)/p_X(x) = (5 - y/2 - x)/(9 - 2x)$$

The cumulative conditional distributions are

$$P_{X|Y}(x|y) = \int_0^x p_{X|Y}(t|y)\, dt = \int_0^x (5 - y/2 - t)/(8 - y)\, dt = [(5 - y/2)x - x^2/2]/(8-y)$$

and

$$P_{Y|X}(y|x) = \int_0^y p_{Y|X}(s|x)\, ds = \int_0^y (5 - s/2 - x)/(9 - 2x)\, ds = [(5 - x)y - y^2/4]/(9-2x)$$

The fact that X and Y are not independent can be seen in several ways.

1. The product of the marginal densities does not equal the bivariate distribution, i.e., $p_Y(y)\, p_X(x) \ne p_{X,Y}(x,y)$.

2. The marginal density of X given Y (or Y given X) is a function of y (or x). This is actually a restatement of 1 above.

3. The prob($X \le 1$) times prob($Y \le 1$) does not equal prob($X \le 1$, $Y \le 1$).

   $$\text{prob}(X \le 1) = \int_0^1 p_X(t)\, dt = P_{X,Y}(1,\infty) = 4/7$$

   $$\text{prob}(Y \le 1) = \int_0^1 p_Y(s)\, ds = P_{X,Y}(\infty,1) = 15/28$$

   $$\text{prob}(X \le 1)\ \text{prob}(Y \le 1) = (4/7)(15/28) = 0.306$$

   This is again a restatement of 1 above and can be used to show dependence but not independence. Independence would require that every possible point be examined whereas only one point must be found where prob($X \le x$) prob ($Y \le y$) $\ne$ prob($X \le x$, $Y \le y$) to establish dependence.

The prob($X \le \frac{1}{2} | Y \le 1$) can be determined from

$$\text{prob}(X \le \tfrac{1}{2} | Y \le 1) = \int_0^{1/2} p_{X|Y}(t|Y \le 1)\, dt$$

$$p_{X|Y}(x|Y \le 1) = \frac{\int_0^1 p_{X,Y}(x,s)\, ds}{\int_0^1 p_Y(s)\, ds} = \frac{\int_0^1 (5 - s/2 - x)/14\, ds}{\int_0^1 (8 - s)/14\, ds}$$

$$= (19 - 4x)/30$$

$$\text{Prob}(X \le \tfrac{1}{2} | Y \le 1) = \int_0^{1/2} (19 - 4t)/30\, dt = 9/30$$

Example 2.13. Show that X and Y are independent if

$p_{X,Y}(x,y) = (3y + 8x - 2xy - 12)/96$ for $2 \le x \le 5$, $0 \le y \le 4$.

Solution: Independence can be established by showing that $p_X(x)\, p_Y(y) = p_{X,Y}(x,y)$.

$$p_X(x) = \int_0^4 (3s + 8x - 2xs - 12)/96 \, ds$$

$$= (2x - 3)/12$$

$$p_Y(y) = \int_2^5 (3y + 8t - 2ty - 12)/96 \, dt$$

$$= (4 - y)/8$$

$$p_X(x)\, p_Y(y) = (2x - 3)(4 - y)/96$$

$$= (3y + 8x - 2xy - 12)/96$$

$$= p_{X,Y}(x,y)$$

X an Y are independent

It is instructive to show that the marginal density of X (or Y) is independent of Y (or X) and to select several points and show that prob($X \le x$, $Y \le y$) = prob($X \le x$) prob($Y \le y$).

Example 2.14. The tabulation below shows the occurrence of average daily temperatures (T) and average daily relative humidities (RH) on each day of a selected 8-day period for 43 consecutive years. From this data determine:

a) $f_{X,Y}(x_i,y_j)$

b) $f_X(x_i)$ and $F_X(x_i)$

c) $f_Y(y_j)$ and $F_Y(y_j)$

d) The probability that

(1) $40 \le T \le 50$ and $60 \le RH \le 80$

(2) $40 \le T \le 50$ given that $60 \le RH \le 80$

(3) $T \le 60$

(4) $RH \le 60$

(5) $T \le 40$ and $RH \le 40$

e) If T and RH are independent.

Number of occurrences

		Temperature °F					
		20-30	30-40	40-50	50-60	60-70	70-80
Relative Humidity %	0-20	2	4	6	2	2	1
	20-40	4	8	12	30	6	9
	40-60	5	15	30	60	30	20
	60-80	3	7	9	25	17	11
	80-100	1	0	2	12	8	3

Solution: (a)(b)(c) There are 6 intervals of temperature and 5 intervals of relative humidity. Let x_i be the i^{th} temperature interval for i=1 to 6 and let y_j be the j^{th} relative humidity range for j=1 to 5. Letting n_{ij} = the entry in the above tabulation corresponding to the i^{th} temperature interval and the j^{th} relative humidity interval we have the $f_{X,Y}(x_i,y_j) = n_{ij}/N$ where

$$N = \sum_{i,j} n_{ij} = 344$$

$f_X(x_i)$, $F_X(x_i)$, $f_Y(y_j)$ and $F_Y(y_j)$ are found by applying equations 2.38 and 2.39.

	i							
j	1	2	3	4	5	6	$f_Y(y_j)$	$F_Y(y_j)$
1	.0058	.0116	.0174	.0058	.0058	.0029	.0494	.0494
2	.0116	.0232	.0349	.0872	.0174	.0262	.2006	.2500
3	.0145	.0436	.0872	1744	.0872	.0581	.4651	.7151
4	.0087	.0203	.0262	.0727	.0494	.0320	.2093	.9244
5	.0029	.0000	.0058	.0349	.0233	.0087	.0756	1.0000
$f_X(x_i)$	.0436	.0988	.1715	.3750	.1831	.1279		
$F_X(x_i)$	.0436	.1424	.3139	.6889	.8720	.9999		

d) (1) $40 \le T \le 50$ and $60 \le RH \le 80 = f_{X,Y}(x_3,y_4) = 0.0262$

(2) $40 \le T \le 50$ given $60 \le RH \le 80 = f_{X|Y}(x_3 \,|\, y_4)$

$f_{X|Y}(x_3 \,|\, y_4) = f_{X,Y}(x_3,y_4)/f_Y(y_4) = 0.0262/0.2093 = 0.125$

(3) $T \le 60 = F_X(x_4) = 0.6889$

(4) $RH \le 60 = F_Y(y_3) = 0.7151$

(5) $T \le 40$ and $RH \le 40 = F_{X,Y}(x_2,y_2)$

$F_{X,Y}(x_2,y_2) = \sum_{i=1}^{2} \sum_{j=1}^{2} f_{X,Y}(x_i,y_j) = 0.0522$

e) If T and RH are independent, then $f_X(x_i)\, f_Y(y_j)$ must equal $f_{X,Y}(x_i,y_j)$. Looking at i=3 and j=3 we have $f_X(x_3) = 0.1715$ and $f_Y(y_3) = 0.4651$. Their product is 0.0798 which is not equal to $f_{X,Y}(x_3,y_3)$ of 0.0872. Therefore T and RH are dependent.

DERIVED DISTRIBUTIONS

Situations often arise where the joint probability distribution of a set of random variables is known and the distribution of some function or transformation of these variables is desired. For example the joint probability distribution of the flows in two tributaries of a stream may be known while the item of interest may be the sum of the flows in the two tributaries. Some commonly used transformations are translation and/or rotation of axes, logarithmic transformations, n^{th} root transformations for n equal 2 and 3 and certain trigonometric transformations.

Thomas (1971) presents the developments that lead to the results presented here concerning transformations and derived distributions for continuous random variables. The procedures for discrete random variables is simply an accounting procedure.

Example 2.15. Let X have the distribution function

$$f_X(x) = c/x \text{ for } X = 2,3,4,5$$

Let $Y = X^2 - 7X + 12$. The probability distribution and possible values of Y can be determined from the following table.

x	2	3	4	5
y	2	0	0	2
$f_X(x)$	c/2	c/3	c/4	c/5

Thus $f_Y(y) = c/3 + c/4 = 35c/60$ for Y=0

$= c/2 + c/5 = 42c/60$ for Y=2

= 0 elsewhere

The value for c can be evaluated from either the requirement that

$$\Sigma_{i=1}^{4} f_X(x_i) = 1 \text{ or } \Sigma_{i=1}^{2} f_Y(y_i) = 1$$

In either case the value of c will be found to be 60/77.

For a univariate continuous distribution of the random variable X, the distribution of U where

$$U = u(X) \tag{2.49}$$

is a monotonic function (u(X) is monotonically increasing if $u(x_2) > u(x_1)$ for $x_2 > x_1$ and monotonically decreasing if $u(x_2) < u(x_1)$ for $x_2 > x_1$) can be found from

$$p_U(u) = p_X(x) \left| dx/du \right| \tag{2.50}$$

Example 2.16. Find the probability of $0 < U < 10$ if $U = X^2$ and X is a continuous random variable with

$$p_X(x) = 3x^2/125 \qquad 0 < X < 5$$

Solution:

$$p_U(u) = p_X(x) \left| dx/du \right|$$

$$dx/du = 1/2x = 1/2\sqrt{u}$$

$$p_U(u) = 3x^2/250\sqrt{u}$$

or

$$p_U(u) = 3\sqrt{u}/250$$

A check to see that $p_U(u)$ is a probability density can be made by integrating $p_U(u)$ from 0 to 25

$$\int_0^{25} p_U(u)\,du = \int_0^{25} 3\sqrt{u}/250\,du = 1$$

Now $\text{prob}(0 < U < 10) = \int_0^{10} 3\sqrt{u}/250\,du = 10^{3/2}/125$

The same result could have been obtained by noting that

$$\text{prob}(0 < U < 10) = \text{prob}(0 < X < \sqrt{10})$$

$$\text{prob}(0 < X < \sqrt{10}) = \int_0^{\sqrt{10}} 3x^2/125\,dx = 10^{3/2}/125$$

In the case of a continuous bivariate density, the transformation from $p_{X,Y}(x,y)$ to $p_{U,V}(u,v)$ where $U = u(X,Y)$ and $V = v(X,Y)$ are one-to-one continuously differentiable transformations can be made by the relationship

$$p_{U,V}(u,v) = p_{X,Y}(x,y) \left| J\left(\frac{x,y}{u,v}\right) \right| \tag{2.51}$$

where $J\left(\frac{x,y}{u,v}\right)$ is the Jacobian of the transformation computed as the determinant of the matrix of partial derivatives

$$J\left(\frac{x,y}{u,v}\right) = \begin{vmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{vmatrix} \tag{2.52}$$

The limits on U and V must be determined from the individual problem at hand.

Example 2.17. Given that $p_{X,Y}(x,y) = (5 - y/2 - x)/14$ for $0 < X < 2$ and $0 < Y < 2$. If $U = X + Y$ and $V = Y/2$, what is the joint probability density function for U and V? What are the proper limits on U and V?

Solution: $p_{X,Y}(x,y) = (5 - y/2 - x)/14 = [5 + y/2 - (x+y)]/14$

$U = X + Y \qquad V = Y/2$

$$p_{U,V}(u,v) = p_{X,Y}(x,y) \left| J\left(\frac{x,y}{u,v}\right) \right|$$

$$J = \begin{vmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{vmatrix} \quad \text{or} \quad \frac{1}{J} = \begin{vmatrix} \frac{\partial u}{\partial x} & \frac{\partial v}{\partial x} \\ \frac{\partial u}{\partial y} & \frac{\partial v}{\partial y} \end{vmatrix}$$

$$\frac{1}{J} = \begin{vmatrix} 1 & 0 \\ \frac{1}{2} & \frac{1}{2} \end{vmatrix} = \frac{1}{2} \text{ or } J = 2$$

$$p_{U,V}(u,v) = 2(5 + v - u)/14 = (5 + v - u)/7$$

The limits on U and V can be determined by noting that Y = 2V and X = U - 2V. Therefore the limit of Y = 0 maps to V = 0, Y = 2 maps to V = 1, X = 0 maps to U = 2V and X = 2 maps to U = 2V + 2. These limits are shown in figure 2.15. A check can be made by integrating $p_{U,V}(u,v)$ over the region $0<V<1$, $2V<U<2V+2$.

$$\int_0^1 \int_{2v}^{2v+2} (5 + v - u)/7 \, du \, dv =$$

$$\frac{1}{7} \int_0^1 \left. 5u + vu - u^2/2 \right|_{2v}^{2v+2} dv =$$

$$\frac{1}{7} \int_0^1 (8 - 2v)\, dv = \left. (8v - v^2)/7 \right|_0^1 = 1$$

Therefore $p_{U,V}(u,v)$ over the above defined region is indeed a probability density function.

A special case of a bivariate transformation is when the distribution of U=u(X,Y) is desired. In this case one method of obtaining $p_U(u)$ is to define a dummy random variable V = v(X,Y). Equation 2.51 is then used to find the joint density of U and V, $p_{U,V}(u,v)$. The univariate density of U is now the marginal distribution of U found by integrating out V.

Example 2.18. The random variable X and Y have the joint probability density function

$$p_{X,Y}(x,y) = (5 - y/2 - x)/14 \quad 0<X<2; 0<Y<2$$

What is the distribution of the sum of X and Y?

Let U = X + Y

The desired distribution is $p_U(u)$. Define V to be Y then

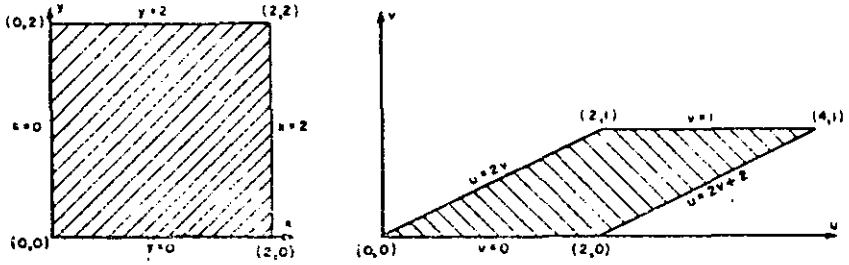


Fig. 2.15. Mapping from X,Y to U,V for Example 2.17.

$$p_{U,V}(u,v) = p_{X,Y}(x,y) \left| J\left(\frac{x,y}{u,v}\right) \right|$$

$$J\left(\frac{x,y}{u,v}\right) = 1/J\left(\frac{u,v}{x,y}\right)$$

$$J\left(\frac{u,v}{x,y}\right) = \begin{vmatrix} \frac{\partial u}{\partial x} & \frac{\partial u}{\partial y} \\ \frac{\partial v}{\partial x} & \frac{\partial v}{\partial y} \end{vmatrix} = \begin{vmatrix} 1 & 1 \\ 0 & 1 \end{vmatrix} = 1$$

$$p_{X,Y}(x,y) = [5 + y/2 - (x+y)]/14 = (5 + v/2 - u)/14$$

$$p_{U,V}(u,v) = (5 + v/2 - u)/14$$

The limits on U and V are found by noting that Y = 0 maps to V = 0, Y = 2 maps to V = 2, X = 0 maps to U = V, and X = 2 maps to U = V + 2. The limits are shown in figure 2.16.

The distribution of U now becomes the marginal distribution of U, $p_U(u)$, and is found by integrating over V.

$$p_U(u) = \int_0^u p_{U,V}(u,v)\, dv \quad \text{for } 0<U<2$$

$$= \int_{u-2}^2 p_{U,V}(u,v)\, dv \quad \text{for } 2<U<4$$

$$p_U(u) = \int_0^u (5 + v/2 - u)/14 \, dv \quad \text{for } 0<U<2$$

$$= (5u - 3u^2/4)/14$$

$$p_U(u) = \int_{u-2}^2 (5 + v/2 - u)/14 \, dv \quad \text{for } 2<U<4$$

$$= (20 - 8u + 3u^2/4)/14$$

The distribution of U is shown in figure 2.16. The fact that $p_U(u)$ is a probability distribution can be shown by integrating $p_U(u)$ over the range $0<U<4$ and showing that the integral is equal to unity. This integration will have to be done in two parts since $p_U(u)$ changes at U = 2.

In some cases the function U = u(X) may be such that it is difficult to analytically determine the distribution of U from the distribution of X. In this case it may be possible to generate a large sample of X's (chapter 13), calculate the corresponding U's and

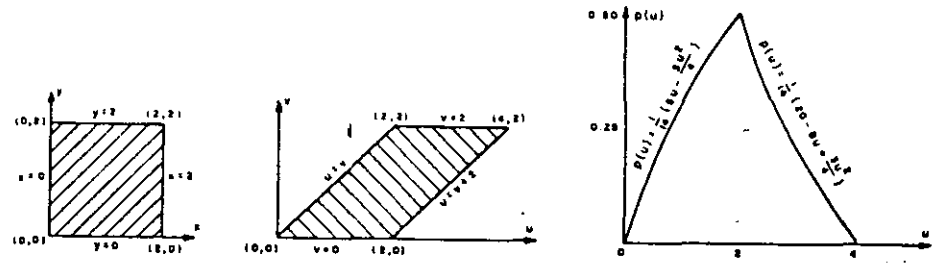


Fig. 2.16. Mapping from X,Y to U,V and plot of $p_U(u)$ for Example 2.18.

then fit a probability distribution to the U's (chapter 6). It should be noted, however, that this empirical method will not in general satisfy equation 2.50.

MIXED DISTRIBUTIONS

If $p_i(x)$ for $i = 1,2,\ldots,m$ represent probability density functions and λ_i for $i = 1,2,\ldots,m$ represent parameters satisfying $\lambda_i \geq 0$ and $\sum_{i=1}^{m} \lambda_i = 1$, then

$$p_X(x) = \sum_{i=1}^{m} \lambda_i p_i(x) \tag{2.53}$$

is a probability density function known as a mixture or mixed distribution since it is composed of a mixture of $p_i(x)$. The parameter λ_i may be thought of as the probability that a random variable is from the probability distribution $p_i(x)$ and $p_i(x)$ is the probability distribution of X given that X is from the i^{th} distribution. The cumulative distribution of X is given by

$$\begin{aligned} P_X(x) &= \int_{-\infty}^{x} \sum_{i=1}^{m} \lambda_i p_i(t)\, dt \\ &= \sum_{i=1}^{m} \lambda_i \int_{-\infty}^{x} p_i(t)\, dt \end{aligned} \tag{2.54}$$

Mixed distributions in hydrology may be applicable in situations where more than one distinct cause for an event may exist. For example flood peaks from convective storms might be described by $p_1(x)$ and from hurricane storms by $p_2(x)$. If λ_1 is the proportion of flood peaks generated by convective storms and $\lambda_2 = (1 - \lambda_1)$, is the proportion generated by hurricane storms, then equations 2.53 and 2.54 would describe the probability distribution of flood peaks.

Singh (1974) and Hawkins (1974) discuss procedures for applying mixed distributions in the form of equation 2.53 to flood frequency determinations.

Example 2.19. A certain event has probability 0.3 of being from the distribution $p_1(x) = e^{-x}$, $x>0$. The event may also be from the distribution $p_2(x) = 2e^{-2x}$, $x>0$. What is the probability that a random observation will be less than 1?

Solution:

$$\begin{aligned} P_X(x) &= 0.3\, P_1(x) + 0.7\, P_2(x) \\ &= 0.3\,(1 - e^{-x}) + 0.7\,(1 - e^{-2x}) \\ P_X(1) &= 0.3\,(1 - e^{-1}) + 0.7\,(1 - e^{-2}) \\ &= 0.3\,(1 - .368) + 0.7\,(1 - .135) \\ &= 0.795 \end{aligned}$$

Exercises

2.1 (a) Construct the theoretical relative frequency histogram for the sum of values obtained in tossing two dice. (b) Toss two dice 100 times and tabulate the frequency of occurrence of the sums of the two dice. Plot the results on the histogram of part a. (c) Why do the results of part b not equal the theoretical results of part a? What possible kinds of errors are involved? Which kind of error was the largest in your case?

2.2 Select a set of data consisting of 50 or more observations. Construct a relative frequency plot using at least two different groupings of the data. Which of the two groupings do you prefer? Why?

2.3 In a period of one week, 3 rainy days were observed. If the occurrence of a rainy day is an independent event, how many ways could the sequence consisting of 4 dry and 3 wet days be arranged?

2.4 If the occurrence of a rainy day is an independent event with probability equal to 0.3, what is the probability of (a) exactly 3 rainy days in one week? (b) the next three days will be rainy? (c) 3 rainy days in a row during any week with the other 4 days dry?

2.5 Consider a coin with the probability of a head equal to p and the probability of a tail equal to q=1-p. (a) What is the probability of the sequence HHTHTTH in 7 flips of the coin? (b) What is the probability of a specified sequence resulting in r H's and s T's? (c) How many ways can r H's and s T's be arranged? (d) What is the probability of r H's and s T's without regard to the order of the sequence?

2.6 The distribution given by $f_X(x) = 1/N$ for $X = 1,2,3,\ldots,N$ is known as the discrete uniform distribution. In the following consider $N \geq 5$. (a) What is the probability that a random value from $f_X(x)$ will be equal to 5? (b) What is the probability that a random value from $f_X(x)$ will be between 3 and 5 inclusive? (c) What is the probability that in a random sample of 3 values from $f_X(x)$ all will be less than 5? (d) What is the probability that the 3 random values from $f_X(x)$ will all be less than 5 given that 1 of the values is less than 5? (e) If 2 random values are selected from $f_X(x)$, what is the probability that one will be less than 5 and the other greater than 5? (f) For what X from $f_X(x)$ is prob$(X \leq x) = 0.5$?

2.7 Consider the continuous probability density function $p_X(x) = a \sin^2 mx$ for $0<X<\pi$. (a) What must be the value of a and m? (b) What is $P_X(x)$? (c) What is prob$(0<X<\pi/2)$? (d) What is prob$(0.5<X<2.1)$ if m = 1?

2.8 Consider the continuous probability density function given by $p_X(x) = 0.25$ for $0<X<a$. (a) What is a? (b) What is prob$(X>a/2)$? (c) What is prob$(X>a/2 \,|\, X>a/4)$? (d) What is prob$(X>a/2 \,|\, X<a/4)$?

2.9 Let $p_X(x) = 0.25$ for $0<X<a$ as in exercise 2.8. What is the distribution of $Y = \ln X$? Sketch $p_Y(y)$.

2.10 Many probability distributions can be defined simply by consulting a table of definite integrals. For example $\int_0^\infty x^{n-1} e^{-x}\, dx$ is equal to $\Gamma(n)$ where $\Gamma(n)$ is defined as the gamma function (see chapter 6). Therefore one can define $p_X(x) = x^{n-1}e^{-x}/\Gamma(n)$ to be a probability density function for $n>0$ and $0<X<\infty$. This distribution is known as the 1-parameter gamma distribution. Using a table of definite integrals, define several possible continuous probability distributions. Give the appropriate range on X and any parameters.

2.11 The annual inflow into a reservoir (acre-feet) follows a probability density given by $p_X(x) = 1/(\beta_1 - \alpha_1)$. The total annual outflow in acre-feet follows a probability distribution given by $p_Y(y) = 1/(\beta_2 - \alpha_2)$. Consider that $\alpha_1 > \beta_2$ and $\alpha_1 < \alpha_2$. (a) Calculate the expression for the probability distribution of the annual change in storage. (b) Plot the probability distribution of the annual change in storage. (c) If β_1 = 100,000, α_1 = 20,000, β_2 = 70,000 and α_2 = 50,000, what is the probability that the change in storage will be i) negative and ii) greater than 15,000 acre-feet?

2.12 The probability of receiving more than 1 inch of rain in each month is given in the following table. If a monthly rainfall record selected at random is found to have more than 1 inch of rain, what is the probability the record is for July? April?

Jan	.25	Apr	.40	Jul	.05	Oct	.05
Feb	.30	May	.20	Aug	.05	Nov	.10
Mar	.35	Jun	.10	Sep	.05	Dec	.20

2.13 It is known that the discharge from a certain plant has a probability of 0.001 of containing a fish killing pollutant. An instrument used to monitor the discharge will indicate the presence of the pollutant with probability 0.999 if the pollutant is present and with probability 0.01 if the pollutant is not present. If the instrument indicates the presence of the pollutant, what is the probability that the pollutant is really present?

2.14 A potential purchaser of a ferry across a river knows that if a flow of 100,000 cfs or more occurs, the ferry will be washed down stream, over a low dam and destroyed. He knows that the probability of a flow of this kind in any year is 0.05. He also knows that for each year that the ferry operates a net profit of $10,000 is realized. The purchase price of the ferry is $50,000. Sketch the probability distribution of the potential net profit over a period of years neglecting interest rates and other complications. Assume that if a flow of 100,000 cfs or more occurs in a year, the profit for that year is zero.

2.15 Assume that the probability density function of daily rainfall is given by

$\text{prob}(X=0) = 0.25$	$X=0$
$p_X(x) = 0.25x$	$0<X<1.732$
$p_X(x) = 0.866 - 0.25x$	$1.732<X<3.464$

(a) Is this a proper probability density function? (b) What is $\text{prob}(X>0.5)$? (c) What is $\text{prob}(X>0.5 \mid X\neq 0)$?

2.16 Consider the probability density function given by

$p_X(x) = \lambda_1/2$	$0<X<2$
$p_X(x) = (1 - \lambda_1)/4$	$2<X<6$

This is a mixture of 2 uniform distributions. (a) Sketch $p_X(x)$ for $\lambda_1 = 0.5$. (b) Sketch $p_X(x)$ for $\lambda_1 = 0.1$. (c) Sketch $p_X(x)$ for $\lambda_1 = 0.333$. (d) In a random sample from $p_X(x)$, 60% of the values were between 0 and 2. What would be an estimate for the value of λ_1?

2.17 If the joint distribution of X and Y is $p_{X,Y}(x,y)$ for X>0 and Y>0, show that in general the distribution of (a) U=X+Y is $p_U(u) = \int p_{X,Y}(x,u-x)dx$, (b) U=XY is $p_U(u) = \int p_{X,Y}(x,u/x)/x\, dx$, (c) U= X/Y is $p_U(u) = \int x p_{X,Y}(x,x/u)/u^2\, dx$, (d) U=Y/X is $p_U(u) = \int x p_{X,Y}(x,ux)dx$.

CURSO: INGENIERIA MARITIMA. MODULO: "FUNDAMENTOS DE HIDRAULICA Y MATEMATICAS" DEL 11 DE MARZO AL 7 DE JUNIO.
MEXICO, D.F.

PROPERTIES OF RANDOM VARIABLES

PROF. ING. JOSE RAYNAL.
JUNIO 1985.

3. Properties of Random Variables

IN CHAPTER 2 random variables and their probability density functions were discussed in general and somewhat abstract terms. Actually every hydrologic variable is a random variable. This includes rainfall, streamflow, infiltration rates, evaporation, reservoir storage, etc. We have also defined any process whose outcome is a random variable as an experiment. A single outcome from this experiment is called a realization of the experiment or an observation from the experiment. Thus daily rainfall values are observations generated by a set of meteorologic conditions that comprise the experiment.

The terms realization and observation can be used interchangeably; however, an observation is generally taken to be a single value of a random variable and a realization is generally taken as a time series of random variables generated by a random experiment. A 10-year record of daily rainfall might be considered as a single realization of a stochastic process (daily rainfall). A second 10-year record of daily rainfall from the same location would then be a second realization of the process.

In this chapter we will be concerned mainly with observations of random variables and with the collection of possible values that these observations may take on. The complete assemblage of all of the values representative of a particular random process is called a population. Any subset of these values would be a sample from the population. For example the pages of this book could represent a population while the pages of this chapter are a sample of that population. All of the books in a library might be taken as a population and should this book be found in the library, it would be a sample from the total population.

Generally one has at hand a sample of observations or data from which inferences about the originating population are to be made and then possibly inferences about another sample from this population. Streamflow records for the past 50 years on a particular stream would be a sample from which inferences about the behavior of the stream for all time (the population) could be made. This information could also be used to estimate the behavior of the stream during some future period of years (another but yet unrealized sample) so that a structure could be properly designed for the stream. Thus one might use information gleaned from one sample to make decisions regarding

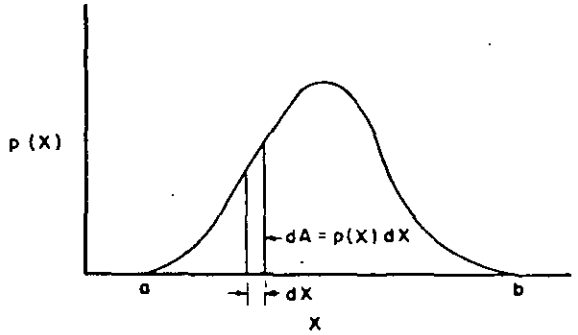


Fig. 3.2. Moment of probability distribution.

and the first moment of the total area about the origin is

$$\mu_1' = \int_A x\,dA \tag{3.1}$$

In case of a random variable and its associated probability density function such as shown in figure 3.2, the first moment about the origin is again given by

$$\mu_1' = \int_A x\,dA$$

In this case $dA = p_X(x)\,dx$ so that

$$\mu_1' = \int_{-\infty}^{\infty} x p_X(x)\,dx \tag{3.2}$$

Generalizing the situation; the i^{th} moment is

$$\mu_i' = \int_{-\infty}^{\infty} x^i\, p_X(x)\,dx \tag{3.3}$$

In the case of a discrete distribution

$$\mu_i' = \Sigma_j\, x_j^i\, f_X(x_j) \tag{3.4}$$

The i^{th} central moment is defined as the i^{th} moment about the mean, μ, of a distribution and is given by

$$\mu_i = \int_{-\infty}^{\infty} (x - \mu)^i\, p_X(x)\,dx \tag{3.5}$$

The expected value of the random variable X is defined to be

$$E(X) = \int_{-\infty}^{\infty} x\, p_X(x)\,dx \qquad X \text{ continuous} \tag{3.6}$$

$$E(X) = \Sigma_j\, x_j\, f_X(x_j) \qquad X \text{ discrete} \tag{3.7}$$

If g(X) is a function of X, then the expected value of g(X) is given by

$$E[g(X)] = \int_{-\infty}^{\infty} g(x)\, p_X(x)\,dx \qquad X \text{ continuous} \tag{3.8}$$

$$= \Sigma_j\, g(x_j)\, f_X(x_j) \qquad X \text{ discrete} \tag{3.9}$$

It is apparent that the expected value of $(x - \mu)^i$ is equal to the i^{th} central moment and $E(X) = \mu_1'$.

$$E[(X - \mu)^i] = \mu_i \tag{3.10}$$

Some rules for finding expected values are

$$E(c) = c \tag{3.11}$$

$$E[c \cdot g(X)] = c \cdot E[g(X)] \tag{3.12}$$

$$E[g_1(X) \pm g_2(X)] = E[g_1(X)] \pm E[g_2(X)] \tag{3.13}$$

MOMENT GENERATING FUNCTIONS[1]

The moment generating function of a random variable X is defined to be $E(e^{tX})$

$$M_X(t) = E(e^{tX}) \tag{3.14}$$

The k^{th} moment about the origin is then found to be the k^{th} derivative of $M_X(t)$ with respect to t and evaluated at $t = 0$.

$$\mu_k' = \left.\frac{d^k M_X(t)}{dt^k}\right|_{t=0} \tag{3.15}$$

The proof of equation 3.15 based on a power series expansion of e^{tX} may be found in Thomas (1971).

MEASURES OF CENTRAL TENDENCY

Arithmetic Mean

Generally the first property of a random variable that is of interest is its mean or average value. The mean, μ_X, of a random variable, X, is its expected value. Thus

$$\mu_X = E(X) = \mu_1' \tag{3.16}$$

A sample estimate of the population mean is the arithmetic average, $\bar{X}$, calculated from

$$\bar{x} = \Sigma_{i=1}^{n} x_i/n \tag{3.17}$$

where n is the number of observations or items in the sample. The arithmetic mean can be estimated from grouped data by

1. Some properties and useful applications of moment generating functions are given in Chapter 6.

$$\bar{x} = \frac{1}{n} \sum_{i=1}^{k} x_i n_i \qquad (3.18)$$

where k is the number of groups, n is the number of observations, n_i is the number of observations in the i^{th} group and x_i is the class mark of the i^{th} group.

Geometric Mean

The sample geometric mean, $\bar{X}_G$, is defined as[2]

$$\bar{x}_G = (\prod_{i=1}^{n} x_i)^{1/n} \qquad (3.19)$$

The logarithm of $\bar{X}_G$ is equal to the arithmetic average of the logarithms of the x_i's. The logarithm of the population geometric mean would be the expected value of logarithm of X.

Median

The sample median, X_{md}, is the observation such that half of the values lie on either side of X_{md}. The population median, μ_{md}, would be the value satisfying

$$\int_{-\infty}^{\mu_{md}} p_X(x)\,dx = 0.5 \qquad \text{X continuous} \qquad (3.20)$$

or $\mu_{md} = x_p$ where p is determined from

$$\sum_{i=1}^{p} f_X(x_i) = 0.5 \qquad \text{X discrete} \qquad (3.21)$$

The median of a sample or a population may not exist.

Mode

The mode is the most frequently occurring value. Thus the population mode, μ_{mo}, would be a value of X maximizing $p_X(x)$ and thus satisfying the equations

$$\frac{d\,p_X(x)}{dx} = 0 \text{ and } \frac{d^2\,p_X(x)}{dx^2} < 0 \qquad \text{X continuous} \qquad (3.22)$$

or the value of X associated with

$$\max_{i=1}^{n} f_X(x_i) \qquad \text{X discrete} \qquad (3.23)$$

The sample mode, X_{mo}, would simply be the most frequently occurring value in the sample. A sample or a population may have none, one or more than one mode.

Weighted Mean

The calculation of the arithmetic mean of grouped data is an example of calculating a weighted mean where n_i/n is the weighting factor. In general the weighted mean is

$$\bar{x}_w = \sum_{i=1}^{k} w_i x_i \Big/ \sum_{i=1}^{k} w_i \qquad (3.24)$$

2. $\prod_{i=1}^{n} x_i = x_1\, x_2\, x_3 \ldots x_n$

where w_i is the weight associated with the i^{th} observation or group and k is the number of observations or groups.

MEASURES OF DISPERSION

Range

The two most common measures of dispersion are the range and the variance. The range of a sample is simply the difference between the largest and smallest sample values. The range of a population is many times the interval from $-\infty$ to ∞ or from 0 to ∞. The sample range is a function of only two of the sample values but does convey some idea of the spread of the data. The population range of many continuous hydrologic variables would be 0 to ∞ and would convey little information. The range has the disadvantage of not reflecting the frequency or magnitude of values that deviate either positively or negatively from the mean since only the largest and smallest values are used in its determination. Occasionally the relative range is used which is simply the range divided by the mean.

Variance

By far the most common measure of dispersion is the variance or its positive square root the standard deviation. The variance of the random variable X is defined as the second moment about the mean and is denoted by σ^2.

$$\mathrm{Var}(X) = \sigma^2 = \mu_2 = E[(X - \mu)^2] = E(X^2) - E^2(X) \qquad (3.25)$$

Thus the variance is the average squared deviation from the mean. For a discrete population of size n, equation 3.25 becomes

$$\sigma^2 = \sum_i (x_i - \mu)^2 / n \qquad (3.26)$$

The sample estimate of σ^2 is denoted by S^2 and calculated from

$$s^2 = \sum_i (x_i - \bar{x})^2 / (n - 1) \qquad (3.27)$$

Two basic differences should be noted between equations 3.26 and 3.27. First in 3.27 $\bar{x}$ is used instead of μ. This is because in dealing with a sample, the population mean would not be known. Secondly n-1 is used in determining S^2 rather than n when calculating σ^2. The reason for this is that $\sum_i (x_i - \bar{x})^2 / n$ would result in a biased[3] estimate for σ^2. The proof that s^2 is an unbiased estimator for σ^2 is given in Appendix A.

The variance for grouped data can be estimated from

$$s^2 = \sum_{i=1}^{k} (x_i - \bar{x})^2 n_i / (n - 1) \qquad (3.28)$$

where k is the number of groups, n is the number of observations, x_i is the class mark and n_i the number of observations in the i^{th} group.

The variance of some functions of the random variable X can be determined from the following relationships:

3. Bias is discussed later in this chapter. An estimator $\hat{\theta}$ for θ is said to be unbiased if $E(\hat{\theta}) = \theta$.

$$\text{Var}(c) = 0 \tag{3.29}$$

$$\text{Var}(cX) = c^2\ \text{Var}(X) \tag{3.30}$$

$$\text{Var}(a+bX) = b^2\ \text{Var}(X) \tag{3.31}$$

The units on the variance are the same as the units on X^2. The units on the standard deviation are the same as the units on the random variable. A dimensionless measure of dispersion is the coefficient of variation defined as the standard deviation divided by the mean. The coefficient of variation is estimated from

$$c_v = s / \bar{x} \tag{3.32}$$

MEASURES OF SYMMETRY

As is apparent from figure 2.12 many distributions are not symmetrical. They may tail off to the right or to the left and as such are said to be skewed. One measure of absolute skewness would be the difference in the mean and the mode. A measure such as this would not be too meaningful, however, because it would depend on the units of measurement. A relative measure of skewness, known as Pearson's first coefficient of skewness, can be obtained by dividing the difference in the mean and the mode by the standard deviation.

$$\text{population measure of skewness} = (\mu - \mu_{mo})/\sigma \tag{3.33}$$

which could be estimated by

$$\text{sample measure of skewness} = (\bar{x} - x_{mo})/s \tag{3.34}$$

The mode of moderately skewed distributions can be estimated from (Parl 1967)

$$x_{mo} \approx \bar{x} - 3(\bar{x} - x_{md}) \tag{3.35}$$

so that

$$\text{sample measure of skewness} = 3(\bar{x} - x_{md})/s \tag{3.36}$$

If sample estimates are replaced by population values in equation 3.36, Pearson's second coefficient of skewness results.

The most commonly used measure of skewness is the coefficient of skew given by

$$\gamma = \mu_3 / \mu_2^{3/2} \tag{3.37}$$

An unbiased estimate for the coefficient of skew based on a sample of size n is

$$c_s = n^2\ M_3/(n-1)(n-2)\ s_x^3 \tag{3.38}$$

where M_3 is the sample estimate for μ_3. The sample coefficient of skew has the advantage of being a function of all of the observations in the sample. Figure 3.3 shows symmetrical, positively and negatively skewed distributions.

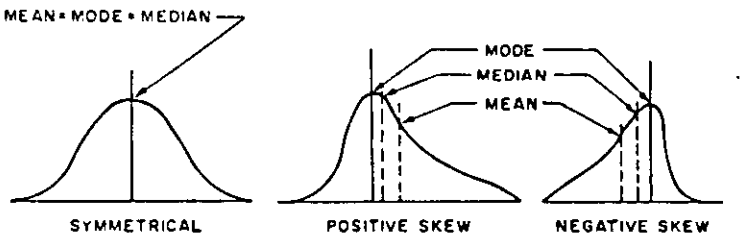


Fig. 3.3. Location of mean, median and mode.

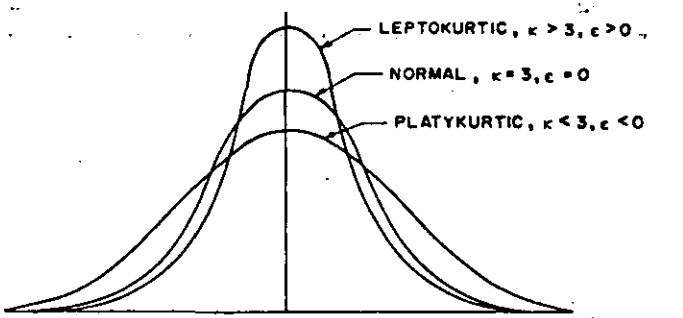


Fig. 3.4. Illustration of kurtosis.

MEASURES OF PEAKEDNESS

A fourth property of random variables based on moments is the kurtosis. Kurtosis refers to the extent of peakedness or flatness of a probability distribution in comparison with the normal probability distribution.[4] Kurtosis is defined as

$$\kappa = \mu_4 / \mu_2^2 \tag{3.39}$$

The sample estimate for the kurtosis is

$$k = M_4 / s_X^4 \tag{3.40}$$

where M_4 is the sample estimate for μ_4. According to Yevjevich (1972a), a less biased estimate for the kurtosis is obtained by multiplying equation 3.40 by $n^3/[(n-1)(n-2)(n-3)]$ where n is the sample size.

The kurtosis for a normal distribution is 3. The normal distribution is said to be mesokurtic. If a distribution has a relatively greater concentration of probability near the mean than does the normal, the kurtosis will be greater than 3 and the distribution is said to be leptokurtic. If a distribution has a relatively smaller concentration of probability near the mean than does the normal, the kurtosis will be less than 3 and the distribution is said to be platykurtic. Figure 3.4 illustrates kurtosis. The coefficient of excess, ε is defined as $\kappa - 3$. Therefore for a normal distribution ε is 0, for a leptokurtic distribution ε is positive and for a platykurtic distribution ε is negative.

4. The normal distribution is a particular probability distribution that plays a very important role in statistical theory. The normal distribution is treated in detail in Chapter 5.

MOMENTS AND EXPECTATION - JOINTLY DISTRIBUTED RANDOM VARIABLES

If X and Y are jointly distributed continuous random variables and U is some function of X and Y, U = g(X,Y), then E(U) can be found by using the methods of Chapter 2 to derive the marginal distribution of U, $p_U(u)$, so that

$$E(U) = E[g(X,Y)] = \int u\, p_U(u)\, du \tag{3.41}$$

A much simpler and more direct method of finding E[g(X,Y)] would be to use the relationship

$$E[g(X,Y)] = \int\int g(x,y)\, p_{X,Y}(x,y)\, dx\, dy \tag{3.42}$$

In either case the result is the average value of the function g(X,Y) weighted by the probability that X=x and Y=y or more simply the mean of the random variable U.

In the discrete case

$$E[g(X,Y)] = \Sigma_i \Sigma_j\, g(x_i,y_j)\, p_{X,Y}(x_i,y_j) \tag{3.43}$$

A general expression for the r,s moment of the jointly distributed random variables X and Y is

$$\mu'_{r,s} = \int\int x^r y^s\, p_{X,Y}(x,y)\, dx\, dy \tag{3.44}$$

for X and Y continuous and

$$\mu'_{r,s} = \Sigma_i \Sigma_j\, x_i^r\, y_j^s\, f_{X,Y}(x_i,y_j) \tag{3.45}$$

for X and Y discrete.

The r,s central moment is defined as

$$\mu_{r,s} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x-\mu_X)^r\, (y-\mu_Y)^s\, p_{X,Y}(x,y)\, dx\, dy \tag{3.46a}$$

for continuous random variables and as

$$\mu_{r,s} = \Sigma_i \Sigma_j\, (x_i-\mu_X)^r\, (y_j-\mu_Y)^s\, f_{X,Y}(x,y) \tag{3.46b}$$

for discrete random variables.

For most situations only moments about the origin and about the means are of interest. As in the case of univariate distributions, the r,s moment about the origin of a bivariate distribution is equal to the expected value of X^rY^s.

The cases where r = 1 and s = 0 and r = 0 and s = 1 are of special interest. For example

$$\begin{aligned} E(X^1 Y^0) &= \int\int x\, p_{X,Y}(x,y)\, dy\, dx \\ &= \int x \int p_{X,Y}(x,y)\, dy\, dx \\ &= \int x\, p_X(x)\, dx \\ &= \mu_X \end{aligned} \tag{3.47}$$

The analogous result holds for $E(X^0Y^1)$.

The most useful central moments are for (r = 2, s = 0), (r = 1, s = 1) and (r = 0, s = 2). For the case r = 2 and s = 0 we have

$$\begin{aligned} E[(X-\mu_X)^2] &= \int\int (x-\mu_X)^2\, p_{X,Y}(x,y)\, dx\, dy \\ &= \int (x-\mu_X)^2 \int p_{X,Y}(x,y)\, dy\, dx \\ &= \int (x-\mu_X)^2\, p_X(x)\, dx \\ &= \mathrm{Var}(X) \end{aligned} \tag{3.48}$$

The analogous result holds for r = 0 and s = 2. The comparable results for discrete random variables are easily obtained.

Covariance

The covariance of X and Y is defined as the 1,1 central moment

$$\begin{aligned} \mathrm{Cov}(X,Y) &= \sigma_{X,Y} = \mu_{1,1} \\ &= E[(X-\mu_X)(Y-\mu_Y)] \\ &= E(XY) - E(X)\,E(Y) \\ &= \int\int (x-\mu_X)(y-\mu_Y)\, p_{X,Y}(x,y)\, dx\, dy \end{aligned} \tag{3.49}$$

For the case where X and Y are independent, equation 3.49 can be written

$$\sigma_{X,Y} = \int (x-\mu_X)\, p_X(x)\, dx \int (y-\mu_Y)\, p_Y(y)\, dy \tag{3.50}$$

since $p_{X,Y}(x,y)$ would equal $p_X(x)\, p_Y(y)$. Furthermore both of the integrals in equation 3.50 are equal to zero so that

$$\mathrm{Cov}(X,Y) = \sigma_{X,Y} = 0 \tag{3.51}$$

if X and Y are independent. The converse of this is not necessarily true however.

The sample estimate for the population covariance $\sigma_{X,Y}$ is $S_{X,Y}$ computed from

$$s_{X,Y} = \sum_{i=1}^{n} (x_i - \bar{x})(y_i - \bar{y})/(n-1) \tag{3.52}$$

Correlation Coefficient

The covariance has units equal to the units of X times the units of Y. A normalized covariance called the correlation coefficient is obtained by dividing the covariance by the products of the standard deviations of X and Y

$$\rho_{X,Y} = \sigma_{X,Y}/\sigma_X\sigma_Y \tag{3.53}$$

It can be shown (Thomas 1971) that $-1 \le \rho_{X,Y} \le 1$. Obviously if X and Y are independent, $\rho_{X,Y} = 0$. Again the converse is not necessarily true. X and Y can be functionally

related and still have $\rho_{X,Y}$ (and $\sigma_{X,Y}$) equal to zero. Actually $\rho_{X,Y}$ is a measure of the linear dependence between X and Y. If $\rho_{X,Y} = 0$ then X and Y are linearly independent, however, they may be related by some other functional form. A value of $\rho_{X,Y}$ equal to ± 1 implies that X and Y are perfectly related by Y=a + bX. If $\rho_{X,Y} = 0$, X and Y are said to be uncorrelated. Otherwise, X and Y are correlated.

The covariance and the correlation coefficient are a measure of how the two variables X and Y vary together. If $\rho_{X,Y}$ (and $\sigma_{X,Y}$) is positive, large values of X tend to be paired with large values of Y and vice versa. If $\rho_{X,Y}$ (and $\sigma_{X,Y}$) is negative, large values of X tend to be paired with small values of Y and vice versa.

The population correlation coefficient $\rho_{X,Y}$ can be estimated by the sample correlation coefficient as

$$r_{X,Y} = s_{X,Y} / s_X s_Y \tag{3.54}$$

where s_X and s_Y are the sample estimates for σ_X and σ_Y given by equation 3.27 and $s_{X,Y}$ is the sample covariance given by equation 3.52.

Figure 3.5 demonstrates some typical values for $r_{X,Y}$. In figure 3.5a all of the points lie on the line Y=X-1 and consequently there is perfect linear dependence between X and Y and the correlation coefficient is unity. In figure 3.5b the points are either on or slightly off of the line Y=X-1, and $r_{X,Y} = 0.986$. Perfect linear dependence does not exist in this case because some of the points deviate slightly from the straight line. In measuring and relating naturally occurring hydrologic variables, a correlation coefficient of 0.986 would be considered quite good and the resulting straight line, Y=X-1 in this case, would usually be judged a good usable relationship between X and Y.

In figure 3.5c the correlation coefficient has dropped to -0.671. The points in this case are scattered about the line Y = 1.264 - 1.571X. The scatter of the points is much greater than in the previous case although the existence of some dependence (stochastic) is still in evidence.

In figure 3.5d the scatter of the points is very great, with a corresponding lack of a strong (stochastic) dependence. Generally a correlation coefficient of 0.211 is considered too small to indicate a useful stochastic dependence as knowledge about X gives very little information about Y.

In the last two paragraphs the modifier "stochastic" has appeared with the word dependence. This is because in reality there are two kinds of dependence - stochastic and functional. Generally throughout this book the word dependence alone should be taken to mean stochastic (or statistical) dependence.

Figures 3.5e and 3.5f contain examples of functionally dependent variables. In figure 3.5e the relationship is $Y = X^2/4$ for $X \geq 0$ and in figure 3.5f the relationship is $Y = -\sqrt{9 - X^2}$ for $-3 \leq X \leq 3$. The correlation coefficient for figure 3.5e is 0.963 indicating a high degree of stochastic (linear) dependence. This illustrates that even though the dependence between X and Y is nonlinear, a high correlation coefficient can result. If the plot of figure 3.5e were to cover a different range of X, the correlation coefficient would change as well.

Figure 3.5f illustrates a situation where Y and X are perfectly functionally related even though the correlation coefficient is zero. The functional relationship is not linear however. This figure demonstrates that one cannot conclude that X and Y are unrelated based on the fact that their correlation coefficients are small.

The fact that two variables have a high degree of linear correlation should not be interpreted as indicating a functional or cause and effect relationship exists between the

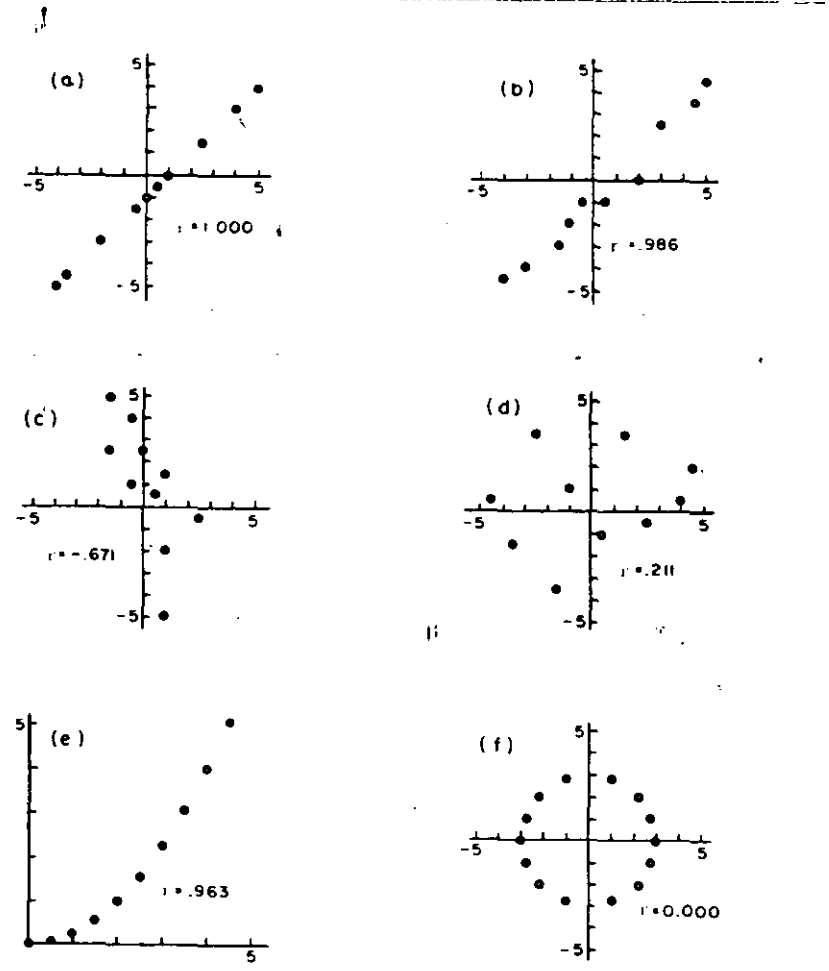


Fig. 3.5. Examples of the correlation coefficient.

two variables. The annual water yield on two adjacent watersheds may be highly positively correlated even though a high yield from one watershed does not cause a high yield from the second watershed. More likely the same climatic factors and geomorphic factors are operating on the two watersheds causing their water yields to be similar. The fact is often overlooked that high correlation does not necessarily mean a cause and effect relationship exists between the correlated variables.

Further Properties of Moments

If Z is a linear function of two random variables X and Y, then

$$Z = aX + bY$$

$$E(Z) = E(aX + bY) = aE(X) + bE(Y) \tag{3.55}$$

$$Var(Z) = Var(aX + bY) = E(aX + bY)^2 - E^2(aX + bY)$$

or

$$\mathrm{Var}(Z) = a^2\ \mathrm{Var}(X) + b^2\ \mathrm{Var}(Y) + 2ab\ \mathrm{Cov}(X,Y) \tag{3.56}$$

Equations 3.55 and 3.56 can be generalized when Y is a linear function of n random variables as follows.

$$Y = \Sigma_{i=1}^{n}\ a_i X_i$$

then

$$E(Y) = E(\Sigma_{i=1}^{n}\ a_i X_i) = \Sigma_{i=1}^{n}\ a_i E(X_i) \tag{3.57}$$

and

$$\mathrm{Var}(Y) = \Sigma_{i=1}^{n}\ a_i^2 \mathrm{Var}(X_i) + 2\ \Sigma_{i<j}\ a_i a_j\ \mathrm{Cov}(X_i,X_j) \tag{3.58}$$

A noteworthy result of equation 3.56 or 3.58 is that for uncorrelated random variables, the variance of a sum or difference is equal to the sum of the variances. This is because the variation in each of the random variables contributes to the variation of their sum or difference.

As a special case of a linear function consider the X_i to be a random sample of size n. Let the a_i all be equal to 1/n. Then Y is equal to $\bar{X}$, the mean of the sample. The Var(Y) is the Var($\bar{X}$) and can be found from equation 3.58. Since the x_i form a random sample, the $\mathrm{Cov}(X_i,X_j) = 0$ for $i \neq j$ and $\mathrm{Var}(X_i) = \mathrm{Var}(X)$. We now have

$$\mathrm{Var}(Y) = \mathrm{Var}(\bar{X}) = \Sigma_{i=1}^{n}\ \frac{1}{n^2}\ \mathrm{Var}(X) = \frac{n}{n^2}\ \mathrm{Var}(X)$$

or

$$\mathrm{Var}(\bar{X}) = \mathrm{Var}(X)/n \tag{3.59}$$

Equation 3.59 states that the variance of the mean of a random sample is equal to the variance of the sample divided by the number of observations in the sample.

If X and Y are independent random variables then the equation preceding equation 3.49 shows that the expectation of their product is equal to the product of their expectation.

$$E(XY) = E(X)E(Y) \qquad \text{if X and Y independent} \tag{3.60}$$

The variance of the product XY for X and Y independent can be obtained from

$$\mathrm{Var}(XY) = E(XY)^2 - E^2(XY)$$

and noting that

$$E(XY)^2 = \int x^2 y^2\ p_{X,Y}(x,y)\,dx\,dy$$

Since X and Y are independent $p_{X,Y}(x,y) = p_X(x)\,p_Y(y)$ and $E(XY)^2$ becomes $E(X^2)$

$E(Y^2)$ or $E(XY)^2 = (\mu_X^2 + \sigma_X^2)(\mu_Y^2 + \sigma_Y^2)$. Also from equation 3.60, $E^2(XY) = E^2(X)E^2(Y) = \mu_X^2\ \mu_Y^2$. Thus

$$\mathrm{Var}(XY) = (\mu_X^2 + \sigma_X^2)(\mu_Y^2 + \sigma_Y^2) - \mu_X^2\ \mu_Y^2$$

which reduces to

$$\mathrm{Var}(XY) = \mu_X^2 \sigma_Y^2 + \mu_Y^2 \sigma_X^2 + \sigma_X^2 \sigma_Y^2 \tag{3.61}$$

for X and Y independent.

A final word of caution on closing this section concerning the expected value of a function of random variables. The caution is that in general

$$E(g(X)) \neq g(E(X))$$

That this is true is obvious from the example of $g(X) = X^2$. From equation 3.25 it can be seen that $E(X^2) = \sigma_X^2 + \mu_X^2$.

$$E(g(X)) = E(X^2) = \sigma_X^2 + \mu_X^2$$

$$g(E(X)) = g(\mu_X) = \mu_X^2$$

thus demonstrating that in general $E(g(X)) \neq g(E(X))$.

SAMPLE MOMENTS

If x_i for i = 1 to n is a random sample, then the r^{th} sample moment about the origin is

$$M_r' = \Sigma_{i=1}^{n}\ x_i^r / n \tag{3.62a}$$

and the r^{th} sample moment about the sample mean is

$$M_r = \Sigma_{i=1}^{n}\ (x_i - \bar{X})^r / n \tag{3.62b}$$

For the bivariate case involving a random sample of x_i and y_i, the r,s sample moment about the origin is

$$M_{r,s}' = \Sigma_{i=1}^{n}\ x_i^r y_i^s / n \tag{3.63a}$$

and the r,s sample moment about $\bar{X},\bar{Y}$ is

$$M_{r,s} = \Sigma_{i=1}^{n}\ (x_i - \bar{X})^r (y_i - \bar{Y})^s / n \tag{3.63b}$$

The expected value of sample moments is equal to the population moments (Mood, et al. 1974).

Two important properties of moments worthy of repeating are:

a) The first moment about the mean is zero.

$$E(X - \mu_X) = E(X) - \mu_X = \mu_X - \mu_X = 0$$

b) The second moment about the origin is equal to the variance plus the square of the mean.

$$\sigma_X^2 = E(X - \mu_X)^2 = E(X^2) - E^2(X) = E(X^2) - \mu_X^2$$

$$E(X^2) = \sigma_X^2 + \mu_X^2$$

The moments about the mean are related to the moments about the origin by the following general equation (Thomas 1971)

$$M_r = \sum_{j=0}^{r} (-1)^j \binom{r}{j} \mu^j M'_{r-j} \qquad (3.64)$$

for the computation of sample moments on a desk calculator it is often convenient to use equation 3.64. The results of equation 3.64 for the first four sample moments are

$$M_1 = 0$$

$$M_2 = M'_2 - \bar{X}^2$$

$$M_3 = M'_3 - 3\bar{X}M'_2 + 2\bar{X}^3$$

$$M_4 = M'_4 - 4\bar{X}M'_3 + 6\bar{X}^2 M'_2 - 3\bar{X}^4$$

Sample moments can be computed from grouped data by using the equations

$$M'_r = \sum_{j=1}^{k} x_j^r n_j / n \qquad (3.65a)$$

and

$$M_r = \sum_{j=1}^{k} (x_j - \bar{X})^r n_j / n \qquad (3.65b)$$

where x_j and n_j are the class mark and number of observations respectively in the j^{th} group, n is the total number of observations and k is the number of groups.

Moments of greater than third order are generally not computed for hydrologic variables because of the small sample size. Higher order moments are very unreliable (have a high variance) for small samples. For example the variance of S^2 (the variance of the sample variance) is $[M_4 - (n-3)S^4/(n-1)]/n$ (Mood et al. 1974). Yevjevich (1972a) presents general expressions for the variance of the variance, coefficient of skew and kurtosis.

PARAMETER ESTIMATION

Thus far probability distribution functions have been written $P_X(x)$ or $F_X(x)$ depending on whether they were continuous or discrete. More correctly they should be written $P_X(x; \theta_1, \theta_2, \ldots, \theta_m)$ or $F_X(x; \theta_1, \theta_2, \ldots, \theta_m)$ indicating that in general the distributions are a function of a set of parameters as well as the random variables. To use probability distributions to estimate probabilities, values for the parameters must be

available. This section discusses methods for estimating the parameter values for probability distributions. Certain properties of these parameter estimates or statistics are also discussed. Rather than carry a dual set of relationships – one for continuous and one for discrete random variables – only the expressions for the continuous random variables will be displayed. The results are equally applicable to discrete distributions.

The usual procedure for estimating a parameter is to obtain a random sample $x_1, x_2, \ldots, x_n$ from the population X. This random sample is then used to estimate the parameters. Thus $\hat{\theta}_i$, an estimate for the parameter θ_i, is a function of the observations or random variables. Since $\hat{\theta}_i$ is a function of random variables, $\hat{\theta}_i$ is itself a random variable possessing a mean, variance and probability distribution.

Intuitively one would feel that the more observations of the random variables that were available for parameter estimation, the closer should be $\hat{\theta}$ to θ. Also if many samples were used for obtaining $\hat{\theta}$, one would feel that the average value of $\hat{\theta}$ should equal θ. These two statements deal with two properties of estimators known as consistency and unbiasedness.

Unbiasedness

An estimate $\hat{\theta}$ of a parameter θ is said to be unbiased if $E(\hat{\theta}) = \theta$. The bias, if any, is given by $E(\hat{\theta}) - \theta$.

The fact that an estimator is unbiased does not guarantee that an individual $\hat{\theta}$ is equal to θ or even close to θ, it simply means that the average of many independent estimates for θ will equal θ.

Consistency

An estimator $\hat{\theta}$ of a parameter θ is said to be consistent if the probability that $\hat{\theta}$ differs from θ by more than an arbitrary constant ε approaches 0 as the sample size approaches infinity.

Consistency is an asymptotic property since it says that by selecting n sufficiently large, the prob$(|\hat{\theta} - \theta| > \varepsilon)$ can be made as small as desired. However for small samples (as are many times used in practice) consistency does not guarantee that a small error will be made. In spite of this one feels more comfortable knowing that $\hat{\theta}$ would converge to θ if a larger sample were used.

The problem of a single estimate of θ from a small sample continues to plague us since neither unbiasedness nor consistency give us much comfort. In choosing between several methods for estimating θ in addition to being unbiased and consistent, it would be desirable if the $\mathrm{Var}(\hat{\theta})$ were as small as possible. This would mean that the probability distribution of $\hat{\theta}$ would be more concentrated about θ.

Efficiency

An estimator $\hat{\theta}$ is said to be the most efficient estimator for θ if it is unbiased and its variance is at least as small as that of any other unbiased estimator for θ. The relative efficiency of $\hat{\theta}_1$, with respect to $\hat{\theta}_2$ for estimating θ is the ratio of $\mathrm{Var}(\hat{\theta}_2)$ to $\mathrm{Var}(\hat{\theta}_1)$.

Finally it is desirable that $\hat{\theta}$ use all of the information contained in the sample relative to θ. If only a fraction of the observations in a sample are used for estimating θ, then some information about θ is lost.

Sufficiency

An estimator $\hat{\theta}$ is said to be a sufficient estimator for θ if $\hat{\theta}$ uses all of the information relevant to θ that is contained in the sample.

More formal statements of the above four properties of estimators and procedures for determining if an estimator has these properties can be found in books on mathematical statistics (Lindgren 1968; Freund 1962; Mood et al. 1974).

There are many ways for estimating population parameters from samples of data. A few of these are graphical procedures, matching selected points, method of moments, maximum likelihood, and minimum chi-square. The graphical procedure consists of drawing a line through plotted points and then using certain points on the line to calculate the parameters. This procedure is very arbitrary and dependent upon the individual doing the analysis. Frequently the method is employed when few observations are available with the thought that few observations will not produce good parameter estimates anyway. When few points are available is precisely the time when the best methods of parameter estimation should be used.

The method of matching points is not a commonly used method but can produce reasonable first approximations to the parameters. The procedure can be valuable in getting initial estimates for the parameters to be employed in iterative solutions that can arise when the method of moments or maximum liklihood are used.

Example 3.1. A certain set of data is thought to follow the distribution $p_X(x) = e^{-x/\theta}/\theta$ for $X>0$. In this particular data set, 75 percent of the values are less than 3.0. Estimate the parameter θ.

Solution: $p_X(x) = e^{-x/\theta}/\theta$

$$P_X(x) = \int_0^x e^{-t/\theta}/\theta \, dt = 1 - e^{-x/\theta}$$

$$1 - P_X(x) = e^{-x/\theta}$$

$$x/\theta = -\ln(1 - P_X(x))$$

$$\theta = -x/\ln(1 - P_X(x)) = -3.00/\ln(1 - 0.75)$$

$$\hat{\theta} = -3.00/-1.39 = 2.16$$

Comment: If sample size n is available the above procedure could be used to obtain n estimates for θ. These n estimates could then be averaged to obtain $\hat{\theta}$. If the probability distribution of interest had m parameters, then the value of $P_X(x)$ and x at m points would be used to obtain m equations in the m unknown parameters. The method of matching points is not recommended for general use in getting final parameter estimates.

Method of Moments

One of the two most commonly used methods for estimating the parameters of a probability distribution is the method of moments. For a distribution with m parameters, the procedure is to equate the first m moments of the distribution to the first m sample moments. This results in m equations which can be solved for the m unknown parameters. Moments about the origin, the mean, or any other point can be used.

Example 3.2. Estimate the parameter λ of the distribution $p_X(x) = \lambda e^{-\lambda x}$ for $X>0$ by the method of moments.

Solution: The first moment about the origin of $p_X(x)$ is

$$\mu = \lambda \int_0^\infty x e^{-\lambda x} \, dx$$

$$= \lambda \cdot 1/\lambda^2 = 1/\lambda$$

Thus the mean of $p_X(x)$ is $1/\lambda$ so that λ can be estimated by $\hat{\lambda} = 1/\bar{X}$.

Example 3.3. Use the method of moments to estimate the parameters of

$$p_X(x) = (2\pi\theta_2^2)^{-1/2} e^{-(x-\theta_1)^2/2\theta_2^2} \qquad -\infty < X < \infty$$

Solution:

$$\mu = (2\pi\theta_2^2)^{-1/2} \int_{-\infty}^{\infty} x \, e^{-(x-\theta_1)^2/2\theta_2^2} \, dx$$

let $y = (x - \theta_1)/\theta_2$ so that $dx = \theta_2 \, dy$

and $$\mu = (1/\sqrt{2\pi}) \int_{-\infty}^{\infty} (\theta_2 y + \theta_1) e^{-y^2/2} \, dy$$

$$= (\theta_2/\sqrt{2\pi}) \int_{-\infty}^{\infty} y e^{-y^2/2} \, dy + (\theta_1/\sqrt{2\pi}) \int_{-\infty}^{\infty} e^{-y^2/2} \, dy$$

The first integral has an integrand h(y) such that h(-y) = -h(y) and is therefore zero. The second integral can be written as

$$2 \int_0^\infty e^{-y^2/2} \, dy = \sqrt{2\pi}$$

Therefore $\mu = \theta_1$, or the parameter θ_1 of this distribution is equal to the mean of the distribution and can be estimated by

$$\hat{\theta}_1 = \bar{X}$$

The second moment about the mean is equal to the variance.

$$\sigma^2 = \int_{-\infty}^{\infty} (x - \mu)^2 (2\pi\theta_2^2)^{-1/2} e^{-\frac{1}{2}(x-\theta_1)^2/\theta_2^2} \, dx$$

but $\theta_1 = \mu$ so

$$\sigma^2 = (2\pi\theta_2^2)^{-1/2} \int_{-\infty}^{\infty} (x - \mu)^2 e^{-\frac{1}{2}(x-\mu)^2/\theta_2^2} \, dx$$

let $y = (x - \mu)/\sqrt{2}\theta_2$ so that $dx = \sqrt{2}\,\theta_2 \, dy$

and

$$\sigma^2 = (2\theta_2^2/\sqrt{\pi}) \int_{-\infty}^{\infty} y^2 e^{-y^2} \, dy$$

$$= (4\theta_2^2/\sqrt{\pi}) \int_0^\infty y^2 e^{-y^2} \, dy$$

$$\sigma^2 = \theta_2^2$$

Thus the parameter θ_2^2 is equal to the variance and can be estimated by S^2 the sample variance.

$\hat{\theta}_2^2 = S^2$

Maximum Likelihood:

Assume we have in hand n random observations $x_1, x_2, \ldots, x_n$. Their joint probability distribution is $p_{\underline{X}}(x_1, x_2, \ldots, x_n; \theta_1, \theta_2, \ldots, \theta_m)$. Since for a random sample the x_i's are independent their joint distribution can be written $p_X(x_1, \theta_1, \theta_2, \ldots, \theta_m) \, p_X(x_2; \theta_1, \theta_2, \ldots, \theta_m) \ldots p_X(x_n; \theta_1, \theta_2, \ldots, \theta_m)$. Now this latter expression is proportional to the probability that the particular random sample would be obtained from the population and is known as the likelihood function.

$$L(\theta_1, \theta_2, \ldots, \theta_m) = \prod_{i=1}^{n} p_X(x_i; \theta_1, \theta_2, \ldots, \theta_m) \tag{3.66}$$

The m parameters are unknown. The values of these m parameters that maximize the likelihood that the particular sample in hand is the one that would be obtained if n random observations were selected from $p_X(x; \theta_1, \theta_2, \ldots, \theta_m)$ are known as the maximum likelihood estimators. The parameter estimation procedure becomes one of finding the values of $\theta_1, \theta_2, \ldots, \theta_m$ that maximize the likelihood function. This can be done by taking the partial derivative of $L(\theta_1, \theta_2, \ldots, \theta_m)$ with respect to each of the θ_i's and setting the resulting expressions equal to zero. These m equations in m unknowns are then solved for the m unknown parameters.

Since many probability distributions involve the exponential function, it is many times easier to maximize the natural logarithm of the likelihood function. Since the logarithmic function is monotonic, the values of the θ's that maximize the logarithm of the likelihood function also maximize the likelihood function.

Example 3.4. Find the maximum likelihood estimator for the parameter λ of the distribution $p_X(x) = \lambda e^{-\lambda x}$ for $X > 0$.

Solution:

$$L(\lambda) = \prod_{i=1}^{n} \lambda e^{-\lambda x_i} = \lambda^n e^{-\lambda \Sigma x_i}$$

$$l(\lambda) = \ln L(\lambda) = n \ln(\lambda) - \lambda \Sigma_{i=1}^{n} x_i$$

$$\frac{\partial l(\lambda)}{\partial \lambda} = n/\lambda - \Sigma_{i=1}^{n} x_i = 0$$

$$\hat{\lambda} = n/\sum_{i=1}^{n} x_i = 1/\bar{X}$$

Example 3.5. Find the maximum likelihood estimators for the parameters θ_1 and θ_2^2 of the distribution

$$p_X(x) = (1/\sqrt{2\pi\theta_2^2}) \, e^{-1/2 (x-\theta_1)^2/\theta_2^2} \qquad -\infty < X < \infty$$

Solution (all summations from 1 to n):

$$L(\theta_1, \theta_2^2) = (2\pi\theta_2^2)^{-n/2} \, e^{-1/2 \Sigma (x_i - \theta_1)^2/\theta_2^2}$$

$$l(\theta_1, \theta_2^2) = \ln L(\theta_1, \theta_2^2) =$$

$$= -n \ln(2\pi)/2 - n \ln(\theta_2^2)/2 - 1/2\theta_2^2 \Sigma (x_i - \theta_1)^2/\theta_2^2$$

$$\frac{\partial l(\theta_1, \theta_2^2)}{\partial \theta_1} = \theta_2^{-2} \Sigma (x_i - \theta_1) = 0$$

or $\hat{\theta}_1 = \Sigma x_i/n = \bar{X}$

$$\frac{\partial l(\theta_1, \theta_2^2)}{\partial \theta_2^2} = -n/2\theta_2^2 + 1/2\theta_2^4 \Sigma (x_i - \bar{X})^2 = 0$$

$$\hat{\theta}_2^2 = \Sigma (x_i - \bar{X})^2/n = (n-1) S^2/n$$

Example 3.5 shows that the maximum likelihood estimators are not unbiased. It can be shown, however, that the maximum likelihood estimators are asymptotically (as $n \to \infty$) unbiased. Maximum likelihood estimators are sufficient and consistent. If an efficient estimator exists, maximum likelihood estimators, adjusted for bias, will be efficient. In addition to these four properties, maximum likelihood estimators are said to be invariant, that is, if $\hat{\theta}$ is a maximum likelihood estimator of θ and the function $h(\theta)$ is continuous, then $h(\hat{\theta})$ is a maximum likelihood estimator of $h(\theta)$.

The method of moments and the method of maximum likelihood do not always produce the same estimates for the parameters. In view of the properties of the maximum likelihood estimators, this method is generally preferred over the method of moments. Cases arise, however, where one can get maximum likelihood estimators only by iterative numerical solutions (if at all) thus leaving room for the use of more readily obtainable estimates possibly by the method of moments. The accuracy of the method of moments is severely affected if the data contains errors in the tails of the distribution where the moment arms are long (Chow 1954). This is especially troublesome with highly skewed distributions.

Finally it should be kept in mind that the properties of maximum likelihood estimators are asymptotic properties (for large n) and there well may exist better estimation procedures for small samples.

CHEBYSHEV INEQUALITY

Certain general statements about random variables can be made without placing restrictions on their distributions. More precise probabilistic statements require more restrictions on the distribution of the random variables. Exact probabilistic statements require complete knowledge of the probability distribution of the random variable.

One general result that applies to random variables is known as the Chebyshev inequality. This inequality states that a single observation selected at random from any probability distribution will deviate more than $k\sigma$ from the mean μ of the distribution with probability less than or equal to $1/k^2$.

$$\text{prob}\left[|x - \mu| \geq k\sigma\right) \leq 1/k^2 \tag{3.67}$$

For most situations this is a very conservative statement. The Chebyshev inequality produces an upper bound on the probability of a deviation of a given magnitude from

the mean.

Example 3.6. The data of table 2.1 has a mean of 67,500 cfs and a standard deviation of 21,000 cfs. Without making any distributional assumptions regarding the data, what can be said of the probability that the peak flow in a year selected at random will deviate more than 40,000 cfs from the mean?

Solution: Applying Chebyshev's inequality we have $k\sigma = 40{,}000$ cfs. Using 21,000 cfs as an estimate for σ we obtain $k = 1.905$.

$$\text{prob}(|x - \mu| \geq k\sigma) \leq 1/k^2 = 1/(1.905)^2 = 0.276$$

The probability that the peak flow in any year will deviate more than 40,000 cfs from the mean is thus less than or equal to 0.276 or roughly 1 in 4.

Comment: That this is a very conservative figure can be seen by noting that only 4 values out of 66 (4/66 = 0.061) lie outside the interval $67{,}500 \pm 40{,}000$. By not making any distributional assumptions, we are forced to accept very conservative probability estimates. In later chapters we will again look at this problem making use of selected probability distributions.

LAW OF LARGE NUMBERS

Chebyshev's inequality is sometimes written in terms of the mean $\bar{X}$ of a random sample of size n. In such a case equation 3.67 becomes

$$\text{prob}(|\bar{X} - \mu| \geq k\sigma/\sqrt{n}) \leq 1/k^2 \qquad (3.68)$$

If we now let $\delta = 1/k^2$ and choose n so that $n \geq \sigma^2/\delta\varepsilon^2$, we have the (weak) Law of Large Numbers (Mood and Graybill 1963) which states:

> Let $p_X(x)$ be a probability density function with mean μ and finite variance σ^2. Let $\bar{X}_n$ be the mean of a random sample of size n from $p_X(x)$. Let ε and δ be any two specified small numbers such that $\varepsilon > 0$, $0 < \delta < 1$. Then for n any integer greater than $\sigma^2/\varepsilon^2\delta$
>
> $$\text{prob}(|\bar{X}_n - \mu| \geq \varepsilon) \leq \delta \qquad (3.69)$$

This statement insures us that we can estimate the population mean with whatever accuracy we desire by selecting the sample large enough. The actual application of equation 3.69 requires knowledge of population parameters and is thus of limited usefulness.

Example 3.7. Assume that the standard deviation of peak flows on the Kentucky River near Salvisa, Kentucky, is 21,000 cfs. How many observations would be required to be at least 95 percent sure that the estimated mean peak flow was within 10,000 cfs of its true value if we know nothing of the distribution of peak flows?

Solution: Applying equation 3.69 we have

$\delta = 1 - .95 = .05$. $\varepsilon = 10{,}000$. $\sigma = 21{,}000$

$$n \geq \sigma^2/\varepsilon^2\delta = (21{,}000)^2/(10{,}000)^2\,(.05) = 88$$

We must have at least 88 observations to be 95 percent sure that the sample mean is within 10,000 cfs of the population mean if we know nothing of the population distribution except its standard deviation.

Comment: We will look at this problem again later making certain distributional assumptions.

Exercises

3.1 What is the expected mean and variance of the sum of values obtained by tossing two dice? What is the coefficient of skew and kurtosis?

3.2 Modular coefficients defined as $K_i = X_i/\bar{X}$ are occasionally used in hydrology. What is the mean, variance and coefficient of variation of modular coefficients in terms of the original data?

3.3 What effect does the addition of a constant to each observation from a random sample have on the mean, variance and coefficient of variation?

3.4 What effect does multiplying each observation in a random sample by a constant have on the mean, variance and coefficient of variation?

3.5 Without any knowledge of the probability distribution of peak flows on the Kentucky River (table 2.1), what can be said about the probability that $|\bar{Q} - \mu_Q|$ is greater than 10,000 cfs?

3.6 Without any knowledge of the probability distribution of peak flows on the Kentucky River (table 2.1), what can be said about the probability that a single random observation will deviate more than 10,000 cfs from μ_Q?

3.7 Using the data of exercise 2.2 calculate the mean and variance from the grouped data. How do the grouped data mean and variance compare to the ungrouped mean and variance? Which estimate do you prefer?

3.8 Calculate the covariance between the peak discharge Q in thousands of cfs and the area A in thousands of square miles for the following data.

Q	A	Q	A
15.50	1.250	18.00	1.400
8.50	0.871	8.75	0.297
85.00	5.690	8.25	0.322
105.00	8.270	3.56	0.178
24.80	1.620	1.90	0.148
3.80	0.175	16.50	0.872
1.76	0.148	2.80	0.091

3.9 Calculate the correlation coefficient between Q and A for the data in exercise 3.8.

3.10 Calculate the coefficient of skew for Q in exercise 3.8. Note that this estimate for γ is relatively unreliable because of the small sample.

3.11 Calculate the kurtosis and the coefficient of excess for Q in exercise 3.8. Note that these estimates are unreliable because of the small sample size.

3.12 Complete the steps necessary to arrive at equation 3.56 from 3.55.

3.13 Show that $\sigma_x \sigma_y \geq |\sigma_{xy}|$.

3.14 A convenient relationship for calculating the estimated variance of a sample of data is

$$s_x^2 = (\Sigma x_i^2 - n\bar{x}^2)/(n-1) = [\Sigma x_i^2 - (\Sigma x_i)^2/n]/(n-1)$$

Derive this relationship from equation 3.27.

3.15 The estimated covariance between X and Y of a bivariate random sample can be calculated from

$$s_{x,y} = (\Sigma x_i y_i - n\bar{x}\bar{y})/n = (\Sigma x_i y_i - \Sigma x_i \Sigma y_i/n)/n$$

Derive this expression from equations 3.49. Note that the above estimated covariance is biased. In practice the final divisor of n is replaced by n-1 to correct for bias.

3.16 In exercise 2.14 if the future maximum life of the ferry is 15 years, what is the expected net profit? Neglect the interest or discount rate.

3.17 What are the maximum likelihood estimates for the parameters of the two parameter exponential distribution? This distribution is given by

$$p_X(x) = \lambda \exp(-\lambda(x-\epsilon)) \qquad X \geq \epsilon,\ \lambda > 0$$

3.18 What are the moment estimates for the parameters of the exponential distribution given in exercise 3.17?

3.19 For the following data, what are the moment and maximum likelihood estimates for the parameters of the distribution given in exercise 3.17? x = 15.0, 10.5, 11.0, 12.0, 18.0, 10.5, 19.5.

3.20 Calculate the coefficient of skew for the Kentucky River data of table 2.1.

3.21 Calculate the kurtosis of the Kentucky River data of table 2.1.

3.22 Using the data of exercise 2.2, calculate the coefficient of skew from the grouped data.

3.23 Using the data of exercise 2.2, calculate the kurtosis from the grouped data.

3.24 What are the maximum likelihood estimates for α and β in the distribution

$$p(x) = (\beta - \alpha)^{-1} \qquad \alpha \leq X \leq \beta?$$

3.25 What are the mean and variance of $f_X(x) = 1/N$ for $x = 1, 2, \ldots, N$?

3.26 What are the mean and variance of $p_X(x) = a \sin^2 x$ for $0 \leq X \leq \pi$?

3.27 Use the method of moments to estimate a in $p_X(x) = a \sin^2 x$ for $0 \leq X \leq \pi$ based on the random sample given by X = 0.5, 3.0, 3.0, 2.5, 1.5, 1.8, 1.0, 0.8, 2.5, 2.2.

3.28 The r^{th} moment about x_0 can be written as $E(X - x_0)^r$. Show that the variance is the smallest possible second moment.

CURSO: INGENIERIA MARITIMA. MODULO: "FUNDAMENTOS DE HIDRAULICA Y MATEMATICAS" DEL 11 DE MARZO AL 7 DE JUNIO.
MEXICO, D.F.

SOME DISCRETE. PROBABILITY. DISTRIBUTIONS AND THEIR APPLICATIONS.

PROF. ING. JOSE RAYNAL.
JUNIO 1985.

4. Some Discrete Probability Distributions and Their Applications

THUS FAR probability distributions have been considered in general terms. This chapter is devoted to some particular discrete distributions and their applications. The next two chapters are devoted to selected continuous distributions.

HYPERGEOMETRIC DISTRIBUTION

Drawing a random sample of size n (without replacement) from a finite population of size N with the elements of the population divided into two groups with k elements belonging to one group is an example of sampling from a hypergeometric distribution. The two groups may be defective or nondefective objects, rainy or nonrainy days, success or failure of a project, etc. For discussion purposes we will consider that an element (or outcome) from the population is either a success or a failure. The probability of X=x successes in a sample of size n selected from a population of size N containing k successes can be determined by applying equation 2.1.

The total number of possible outcomes or ways of selecting a sample of size n from N objects is $\binom{N}{n}$. The number of ways of selecting x successes and n-x failures from the population containing k successes and N-k failures is $\binom{k}{x}\binom{N-k}{n-x}$. Thus the probability is

$$f_X(x; N, n, k) = \binom{k}{x}\binom{N-k}{n-x}/\binom{N}{n} \tag{4.1}$$

The distribution given by equation 4.1 is known as the hypergeometric distribution where $f_X(x; N, n, k)$ is the probability of obtaining X=x success in a sample of size n drawn from a population of size N containing k successes.

The cumulative hypergeometric distribution giving the probability of x or fewer successes is

$$F_X(x; N, n, k) = \sum_{i=0}^{x}\binom{k}{i}\binom{N-k}{n-i}/\binom{N}{n} \tag{4.2}$$

There are certain natural restrictions on this distribution. For example x cannot exceed k, x cannot exceed n, k cannot exceed N and n cannot exceed N. N, n, k, and x are all nonnegative integers. Furthermore the outcomes must be random and equally likely.

The mean of the hypergeometric distribution is

$$E(X) = nk/N \tag{4.3}$$

and the variance is

$$\mathrm{Var}(X) = nk(N-k)(N-n)/N^2(N-1) \tag{4.4}$$

Tables of the hypergeometric distribution are available (Beyer 1968; Lieberman and Owen 1961; Owen 1962). The terms of the hypergeometric can also be determined from table E.1.

Example 4.1. Example 2.6 is an example where the hypergeometric applies. In this example a success is selecting a bad record and N=10, k=3, n=4. The solutions can be written in terms of the hypergeometric as

(a) $f_X(1; 10, 4, 3) = 0.500$

(b) $f_X(3; 10, 4, 3) = 0.0333$

(c) $1-f_X(0; 10, 4, 3) = 1 - 0.1667 = 0.8333$

Example 4.2. Assume that during a certain September, 10 rainy days occurred. Also assume that at this particular location the occurrence of rain on any day is independent of whether or not it rained on any previous day. (This is many times not a good assumption).

A sample of 10 days is selected at random and their climatic data analyzed. (a) What is the probability that 4 of these days will have been rainy? (b) What is the probability that less than 4 of these days were rainy?

Solution: Use the hypergeometric distribution with

N=30, n=10, k=10

(a) $f_X(4; 30, 10, 10) = \binom{10}{4}\binom{20}{6}/\binom{30}{10} = 0.271$

(b) $F_X(3; 30, 10, 10) = [\binom{10}{0}\binom{20}{10}+\binom{10}{1}\binom{20}{9}+\binom{10}{2}\binom{20}{8}+\binom{10}{3}\binom{20}{7}]/\binom{30}{10} = 0.560$

Example 4.3. Examples of the hypergeometric distribution commonly found in the literature include card sampling problems (What is the probability of exactly 2 aces in a 5 card hand selected at random from a 52 card deck?) and acceptance sampling problems (What is the probability of selecting 5 defective items from a lot of 50 items if 20 items are selected and the lot actually contains 12 defectives?

Solution: Card problem

Prob(2 aces) = $f_X(2; 52, 5, 4) = \binom{4}{2}\binom{48}{3} / \binom{52}{5} = 0.040$

Acceptance Sampling Problem:

Prob(5 def) = $f_X(5; 50, 20, 12) = \binom{12}{5}\binom{38}{15} / \binom{50}{20} = 0.26$

BERNOULLI PROCESSES

Binomial Distribution

Consider a discrete time scale. At each point on this time scale an event may either occur or not occur. Let the probability of the event occurring be p for every point on the time scale. Thus the occurrence of the event at any point on the time scale is independent of the history of any prior occurrences or nonoccurrences. The probability of an occurrence at the i^{th} point on the time scale is p for i = 1, 2, ... A process having these properties is said to be a Bernoulli process.

As an example of a Bernoulli process consider that during any year the probability of the maximum flow exceeding 10,000 cfs on a particular stream is p. Common terminology for a flow exceeding a given value is an exceedance. Further consider that the peak flow in any year is independent from year to year (a necessary condition for the process to be a Bernoulli process). Let q = 1-p be the probability of not exceeding 10,000 cfs. We can neglect the probability of a peak of exactly 10,000 cfs since the peak flow rates would be a continuous process so the probability of a peak of exactly 10,000 cfs would be zero. In this example the time scale is discrete with the points being nominally 1 year in time apart. We can now make certain probabilistic statements about the occurrence of a peak flow in excess of 10,000 cfs (an exceedance).

For example the probability of an exceedance occurring in year 3 and not in years 1 or 2 can be evaluated from equation 2.9 as qqp since the process is independent from year to year. The probability of (exactly) one exceedance in any 3-year period is pqq + qpq + qqp since the exceedance could occur in either the first, second or third year. Thus the probability of (exactly) one exceedance in three years is $3pq^2$.

In a similar manner the probability of 2 exceedances in 5 years can be found from the summation of the terms ppqqq, pqpqq, pqqpq, ..., qqqpp. It can be seen that each of these terms is equivalent to p^2q^3 and that the number of terms is equal to the number of ways of arranging 2 items (the p's) among 5 items (the p's and q's). Therefore the total number of terms is $\binom{5}{2}$ or 10 so that the probability of exactly 2 exceedances in 5 years is $10\,p^2q^3$.

This result can be generalized so that the probability of X=x exceedances in n years is $\binom{n}{x} p^x q^{n-x}$. The result is applicable to any Bernoulli process so that the probability of X=x occurrences of an event in n independent trials if p is the probability of an occurrence in a single trial is given by

$$f_X(x; n, p) = \binom{n}{x} p^x q^{n-x} \qquad x = 0, 1, 2, \ldots, n \tag{4.5}$$

Equation 4.5 is known as the binomial distribution.

The binomial distribution and the Bernoulli process are not limited to a time scale. Any process that may occur with probability p at discrete points in time or space or in individual trials may be a Bernoulli process and follow the binomial distribution.

The cumulative binomial distribution is

$$F_X(x; n, p) = \sum_{i=0}^{x} \binom{n}{i} p^i q^{n-i} \qquad x = 0, 1, 2, \ldots, n \tag{4.6}$$

and gives the probability of x or fewer occurrences of an event in n independent trials if the probability of an occurrence in any trial is p.

Continuing the above example, the probability of less than 3 exceedances in 5 years is

$$F_X(2; 5, p) = \sum_{i=0}^{2} \binom{5}{i} p^i q^{5-i}$$

$$= f_X(0; 5, p) + f_X(1; 5, p) + f_X(2; 5, p)$$

The mean and variance of the binomial distribution are

$$E(X) = np \tag{4.7}$$

$$Var(X) = npq \tag{4.8}$$

The coefficient of skew is $(q-p)/\sqrt{npq}$ so that the distribution is symmetrical for p = q, skewed to the right for q>p and skewed to the left for q<p.

Example 4.4. On the average, how many times will a 10-year flood occur in a 40-year period? What is the probability that exactly this number of 10-year floods will occur in a 40-year period?

Solution: A 10-year flood has p = 1/10 = 0.1

$$E(X) = np = 40(0.1) = 4$$

$$f_X(4; 40, 0.1) = \binom{40}{4} (0.1)^4 (0.9)^{36} = 0.2059$$

Comment: This problem illustrates the difficulty of explaining the concept of return period to laymen. We have said that on the average a 10-year event occurs once every 10 years and that in a 40-year period we expect it to occur 4 times. Yet we have also shown that in about 80% (100(1-0.2059)) of all possible independent 40-year periods the 10-year event will not occur exactly 4 times. As a matter of fact the probability that it will occur 3 times is nearly identical to the probability it will occur 4 times (0.2003 vs. 0.2059). The number of occurrences, X, is truly a random variable (with a binomial distribution).

The individual and cumulative terms of the binomial distribution are tabled in many references (see for instance Beyer (1968) or Selby (1970)). The highest value of p given in most tables is 0.5. For values of p in excess of 0.5 the roles of p and q and x and n-x can be reversed since $f_X(x; n, p) = f_X(n-x; n, q)$.

The binomial distribution has an additive property (Gibra 1973). That is if X has a binomial distribution with parameters n_1 and p and Y has a binomial distribution with parameters n_2 and p, then Z=X+Y has a binomial distribution with parameters $n=n_1 + n_2$ and p.

The binomial distribution can be used to approximate the hypergeometric distribution if the sample selected is small in comparison to the number of items N from

which the sample is drawn. In this case the probability of a success would be about the same for each trial.

Example 4.5. Compare the hypergeometric and binomial for N=40, n=5, k=10 and X=0, 1, 2, 3, 4, and 5.

Solution:

X	hypergeometric $f_X(x;N,n,k) = f_X(x;40,5,10)$	binomial $f_X(x;n,p) = f_X(x;5,\frac{10}{40})$
0	0.2166	0.2373
1	0.4165	0.3955
2	0.2777	0.2637
3	0.0793	0.0879
4	0.0096	0.0146
5	0.0004	0.0010

Comment: This merely indicates that drawing a small sample without replacement from a large population and drawing the same sample with replacement (so probabilities in each trial are constant) are nearly equivalent.

Example 4.6. The operator of a boat dock has decided to put in a new facility along a certain river. In an economic analysis of the situation he decided to have the facility designed to withstand floods up to 75,000 cfs. Furthermore he has determined that if one flood greater than this occurs in a 5-year period, he can repair his facility and break even on its operation during the 5-year period. If more than one flow in excess of 75,000 cfs occurs, he will lose money. If the probability of exceeding 75,000 cfs is 0.15, what is the probability the operator will make money?

Solution: Money will be made if no floods exceeding 75,000 cfs occur during the 5-year period. Let X be the number of floods.

$$f_X(0;5,0.15) = \binom{5}{0}(0.15)^0(0.85)^5 = 0.4437$$

Comment: The probability that the operator will make the investment work for 5 years and then just break even is a very high

$$f_X(1;5,0.15) = \binom{5}{1}(0.15)^1(0.85)^4 = 0.3915$$

Thus even though the risk or probability of losing money is low (1 - 0.3915 - 0.4437 = 0.1648), the investment may not be an attractive one.

Example 4.7. In order to be 90 percent sure that a design storm is not exceeded in a 10-year period, what should be the return period of the design storm?

Solution: Let p be the probability of the design storm being exceeded. The probability of no exceedances is given by

$$f_X(0;10,p) = \binom{10}{0}p^0q^{10}$$

$$0.90 = (1-$$

$$p = 1 - (0.90)^{1/10} = 1 - 0.9895 = 0.0105$$

$$T = 1/p = 95 \text{ years}$$

Comment: To be 90 percent sure that a design storm is not exceeded in a 10-year period a 95-year return period storm must be used. If a 10-year return period storm is used, the chances of it being exceeded is

$$1 - f_X(0;10,0.1) = 0.6513$$

In general the chance of at least one occurrence of a T-year event in T years is $1 - f_X(0;T,1/T) = 1 - (1-1/T)^T$. Therefore, for a long design life, the chance of at least one occurrence of an event with a return period equal to the design life approaches 1 - 1/e or 0.632. Thus if the design life of a structure and its design return period are the same, the chances are very great that the capacity of the structure will be exceeded during its design life.

The procedure outlined in example 4.7 can be used to determine a design return period when the allowable risk is stated. Note that the design return period must be much greater than the life of the project to be reasonably sure that an exceedance will not occur. No matter what design return period is selected, there is still a chance that an exceedance will occur. Some may argue that there is an upper limit to the magnitude of natural events such as flood peaks. They would argue that a peak of 100,000 cfs from a 1-acre watershed would be impossible. In practice the probability that would be assigned to an event of this sort is so small that it can be neglected for most practical purposes.

Figure 4.1 shows the design return period that must be used to be a certain percent confident that the design will not be exceeded during the design life of the project. The parameters on the curves are the percent chance of no exceedance during the design life. For example to be 90 percent sure that a design condition will not be exceeded during a project whose design life is 100 years, the project would have to be designed on the basis of a 900-year event. Figure 4.1 is derived from calculations like those contained in example 4.7.

Figure 4.1 can also be used to evaluate the risk or percent chance of an event in excess of the design event during the design life. For example if a project is designed on the basis of a 50-year event and the design life of the project is 10 years, the designer is taking a 19 percent chance (100-81) that the design condition will be exceeded.

Since the probability of a success on any trial is independent of past history, the origin of the time scale of a Bernoulli process can be taken at any time point. Thus the probability of any combination of successes or failures is the same for any sequence of n points regardless of their location with respect to the origin.

Example 4.8. Three successes have occurred on the first 5 trials of a Bernoulli process with p=0.4. What is the probability of 3 successes in the next 5 trials?

Solution: $f_X(3;5,.4) = \binom{5}{3}(0.4)^3(0.6)^2 = 0.2304$

Geometric Distribution

The probability that the first exceedance (or success) of a Bernoulli trial occurs on the X^{th} trial can be found by noting that for the first exceedance to [illegible] on the X^{th} trial

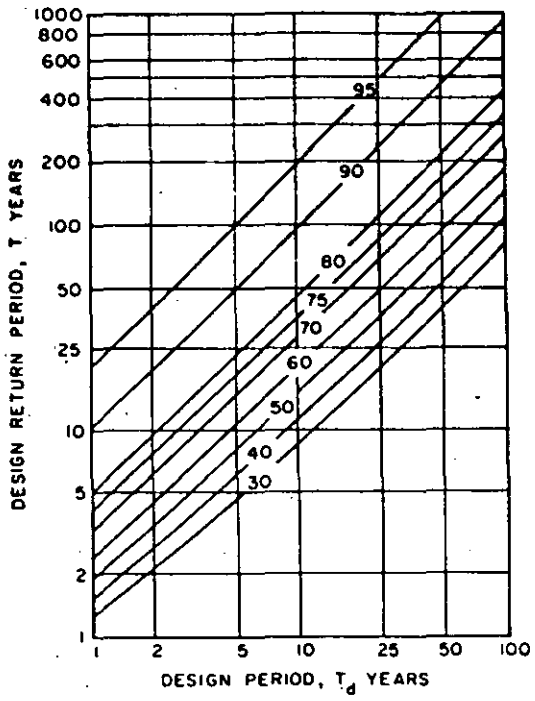


Fig. 4.1. Design return period required as a function of design life to be a given percent confident (curve parameter) that the design condition is not exceeded.

there must be X-1 preceeding trials without an exceedance followed by 1 trial with an exceedance. Thus the desired probability is pq^{x-1}. This is known as the geometric distribution

$$f_X(x;p) = pq^{x-1} \qquad x=1, 2, 3, \ldots \tag{4.9}$$

The mean and variance of the geometric distribution are

$$E(X) = 1/p \tag{4.10}$$

$$Var(X) = q/p^2 \tag{4.11}$$

Since the $E(X) = 1/p$ this means that on the average a T-year event occurs on the T^{th} year which agrees with our intuitive concept of a return period.

Example 4.9. What is the probability that a 10-year flood will occur for the first time during the fifth year after the completion of a project? What is the probability it will be at least the fifth year before a 10-year flood occurs?

Solution: The probability that the first exceedance is in year 5 is $f_X(5;0.1) = (0.1)(0.9)^4 = 0.0656$. The probability that it will be at least the fifth year before the first occurrence is equal to the probability of no occurrences in the first 4 years which is $(0.9)^4 = 0.6561$.

Example 4.10. What is the probability that exactly 9 years will elapse between occurrences of a 10-year event?

Solution: This is the same as the probability of the first occurrence on the tenth year or $f_X(10;0.1) = (0.1)(0.9)^9 = .0387$.

Negative Binomial Distribution

The probability that the k^{th} exceedance (success) occurs on the X^{th} trial ($X \geq k$) of a Bernoulli process can be found by noting that there must be k-1 exceedances in the X-1 trials preceeding the k^{th} exceedance on the X^{th} trial. The probability of k-1 exceedances in X-1 trials is given by the binomial distribution as $\binom{x-1}{k-1} p^{k-1} q^{x-k}$. The probability that the X^{th} trial results in an exceedance is p so the desired probability is given by the negative binomial distribution.

$$f_X(x;k,p) = \binom{x-1}{k-1} p^k q^{x-k} \qquad x=k, k+1, \ldots \tag{4.12}$$

The mean and variance of the negative binomial distribution are

$$E(X) = k/p$$

$$Var(X) = kq/p^2$$

As might be expected since the negative binomial is based on the binomial, the additive feature holds. Thus if X and Y are described by $f_X(x;k_1,p)$ and $f_Y(y;k_2,p)$ respectively, then $Z=X+Y$ follows the negative binomial $f_Z(z;k_1+k_2,p)$.

Example 4.11. What is the probability that the fourth occurrence of a 10-year flood will be on the fortieth year?

Solution: $f_X(40;4,0.1) = \binom{39}{3} (0.1)^4 (0.9)^{36}$

$= .0206$

Summary of Bernoulli Process

In a Bernoulli process at each instant of time (or location or trial) an event may either occur with probability p or not occur with probability q = 1-p. The probability of the event occurring is independent of the time and independent of the past history of occurrences. The number of occurrences in a given time interval (or distance or number of trials) follows the binomial distribution. The probability that the first occurrence is at the X^{th} time is described by the geometric distribution. The probability that the k^{th} occurrence was at the X^{th} time is described by the negative binomial distribution. It was also found that the probability distribution of the length of time between occurrences can be found from the geometric distribution by noting that the probability that x trials elapse between occurrences is the same as the probability that the first occurrence is at the $x+1^{st}$ time or $f_X(x+1,p) = pq^x$.

POISSON PROCESS

Poisson Distribution

Consider a Bernoulli process defined over an interval of time (or space) so that p is the probability that an event may occur during the time interval. If the time interval is allowed to become shorter and shorter so that the probability, p, of an event occurring in the interval gets smaller and the number of trials, n, increases in such a fashion that np remains constant, then the expected number of occurrences in any total time interval remains the same. It can be shown that as n gets large and p gets small so that np remains a constant, λ, the binomial distribution approaches the Poisson distribution given by

$$f_X(x;\lambda) = \lambda^x e^{-\lambda}/x! \qquad x = 0,1,2,\ldots;\ \lambda > 0 \tag{4.13}$$

The mean and variance of the Poisson distribution are

$$E(X) = \lambda \tag{4.14}$$

$$Var(X) = \lambda \tag{4.15}$$

The coefficient of skew is $\lambda^{-1/2}$ so that as λ gets large, the distribution goes from a positively skewed distribution to a nearly symmetrical distribution.

The cumulative Poisson distribution is

$$F_X(x;\lambda) = \sum_{i=0}^{x} \lambda^i e^{-\lambda}/i! \tag{4.16}$$

Table E.2 contains the cumulative terms of the Poisson distribution.

Example 4.12. What is the probability that a storm with a return period of 20 years will occur once in a 10-year period?

Solution: Using the binomial distribution the exact answer is

$$f_X(1;10,.05) = \binom{10}{1}(.05)(.95)^9 = 0.315$$

Approximating with the Poisson

$$\lambda = np = 10 \times 0.05 = 0.5$$

$$f_X(1;0.5) = 0.5e^{-0.5}/1 = 0.303$$

Thus the solutions are not identical but for most practical work the Poisson approximation is satisfactory.

Example 4.13. What is the probability of 5 occurrences of a 2-year storm in a 10-year period?

Solution: Using the binomial

$$f_X(5;10,0.5) = \binom{10}{5}(.5)^5(.5)^5 = 0.246$$

Approximating with the Poisson

$$\lambda = np = 10 \times 0.5 = 5$$

$$f_X(5;5) = 5^5 e^{-5}/5! = 0.176$$

Comment: For this situation n in not large enough and p small enough for a good approximation.

Example 4.14. What is the probability of fewer than 5 occurrences of a 20-year storm in a 100-year period?

Solution: n is relatively large and p small so the Poisson will be used.

$$\lambda = np = 100(.05) = 5$$

$$Prob(X < 5) = Prob(X \le 4) = F_X(4;5)$$

$$F_X(4;5) = \sum_{i=0}^{4} 5^i e^{-5}/i! = 0.440$$

The Poisson distribution possesses the additive property that the sum of two Poisson random variables with parameters λ_1 and λ_2 is a Poisson random variable with parameter $\lambda = \lambda_1 + \lambda_2$.

A Poisson process for a continuous time scale can be defined analogous to a Bernoulli process on a discrete time scale. The Poisson process refers to the occurrence of events along a continuous time (or location) scale. The assumptions underlying the process are:

1. The probability of an event in any short interval t to $t+\Delta t$ is $\lambda\Delta t$ (proportional to the length of the interval) for all values of t. This property is known as stationarity.

2. The probability of more than one event in any short interval t to $t+\Delta t$ is negligible in comparison to $\lambda\Delta t$.

3. The number of events in any interval of time is independent of the number of events in any other non-overlapping interval of time.

The probability distribution of the number of events X in time t for a Poisson process is given by

$$f_X(x;\lambda t) = (\lambda t)^x e^{-\lambda t}/x! \qquad \lambda > 0; t > 0; x = 0,1,2,\ldots \tag{4.17}$$

where $f_X(x;\lambda t)$ is the probability of X events in time t. Equation 4.17 is a Poisson distribution with parameter λt. The mean and variance of $f_X(x;\lambda t)$ are $E(X) = \lambda t$ and $Var(X) = \lambda t$. The parameter λ is the average rate of occurrence of the event.

Exponential Distribution

The probability distribution of the time, T, between occurrences of the event can be found by noting that the prob(T < t) is equal to 1-prob(T>t). The prob(T > t) is equal

to the probability of no occurrences in time t which is $f_X(0;\lambda t)$ or $e^{-\lambda t}$. Thus

$$\text{prob}(T \le t) = P_T(t;\lambda) = 1 - e^{-\lambda t} \tag{4.18}$$

which is a cumulative distribution known as the exponential distribution. The probability density function is

$$p_T(t;\lambda) = \frac{dP_T(t;\lambda)}{dt} = \lambda e^{-\lambda t} \tag{4.19}$$

and is the probability distribution of the length of the time interval between occurrences of the event. The mean and variance of the exponential distribution are $1/\lambda$ and $1/\lambda^2$, respectively.

Gamma Distribution

The probability distribution of the time to the n^{th} occurrence can be found by noting that the time to the n^{th} occurrence is the sum of n independent random variables, $T_1 + T_2 + \ldots + T_n$, from the exponential distribution. The method of derived distributions can be used with the result that the probability density function of the time to the n^{th} occurrence is

$$p_T(t;n,\lambda) = \lambda^n t^{n-1} e^{-\lambda t}/(n-1)! \qquad t>0;\ \lambda>0;\ n = 1,2,\ldots \tag{4.20}$$

which is the gamma distribution for integer values of the parameter n. The gamma distribution has $E(T) = n/\lambda$ and $Var(T) = n/\lambda^2$.

Example 4.15. Barges arrive at a lock at an average of 4 each hour. (a) If the arrival of barges at the lock can be considered to follow a Poisson process, what is the probability that 6 barges will arrive in 2 hours? (b) If the lock master has just locked through all of the barges at the lock, what is the probability he can take a 15 minute break without another barge arriving? (c) If the operation of the lock is such that 4 barges can be locked through at once and the lock master insists that this always be the case, what is the probability that the first barge to arrive after 4 previous barges have been locked through will have to wait at least 1 hour before being locked through?

Solution:

(a) For this problem the rate constant is 4 hours^{-1}. The probability of 6 arrivals in 2 hours can be determined from the Poisson distribution

$$f_X(x;\lambda t) = f_X(6;8) = 8^6 e^{-8}/6! = 0.1221$$

(b) The probability of no arrivals in 15 minutes is also from the Poisson

$$f_X(0;1) = 1^0 e^{-1}/0! = 0.3679$$

Note that this is not the same as the probability that it will be 15 minutes until the next arrival. The time scale is continuous so the probability that it will be exactly 15 minutes until the next arrival is zero. We can only talk of probabilities associated with time intervals, not specific times.

(c) The barge must wait for the arrival of 3 additional barges. The probability that the time T for 3 barges to arrive is greater than 1 hour

$\text{prob}(T_3 \ge 1)$ is $1-\text{prob}(T_3 \le 1)$.

The probability that $T \le 1$ for 3 arrivals comes from the gamma distribution.

$$P_3(1;n,\lambda) = \int_0^1 p_3(t;n,\lambda)dt$$

$$= \int_0^1 (4^3 t^2 e^{-4t}/2!)\,dt$$

$$= 0.762$$

The desired probability is 1 - 0.762 = 0.238.

Summary of Poisson Process

The Poisson process is a discrete process on a continuous time scale. Therefore the probability distribution of the number of events in a time T is a discrete distribution while the probability distributions for the time between events and the time to the n^{th} event are continuous distributions.

For a Poisson process the probability that an event will occur in a short time interval t to $t+\Delta t$ is $\lambda\Delta t$ for all t. The probability that more than one event occurs in Δt is negligible. The probability distribution of the number of events in a given time T is the Poisson distribution. The exponential distribution describes the time between events and the gamma distribution the time to the n^{th} event.

Example 4.16. It has been proposed that an event-based rainfall simulation model can be constructed by modeling the occurrence of rainstorms by a Poisson process and the amount of rain in each storm by some continuous probability distribution. In this way the time between rainstorms would follow an exponential distribution, the time for X rainstorms would follow a gamma distribution, and the number of rainstorms in a time interval would follow a Poisson distribution. Duckstein et al. (1975) and Fogel et al. (1974) used a modification of this approach. Part of Fogel et al's. results are shown as figure 4.2.

MULTINOMIAL DISTRIBUTION

The binomial distribution can be generalized to include the probabilities of outcomes of several types rather than the two possible outcomes of the binomial. If the probabilities associate with each of k distinct outcomes are $p_1, p_2, \ldots, p_k$, then in independent trials the probability of x_1 outcomes of type 1, x_2 outcomes of type 2, ... x_k outcomes of type k is given by the multinomial distribution as

$$f_{X_1,X_2,\ldots,X_k}(x_1,x_2,\ldots,x_k;n,p_1,p_2,\ldots,p_k) = \frac{n!}{x_1!x_2!\ldots x_k!}\, p_1^{x_1} p_2^{x_2}\ldots p_k^{x_k}$$

or

$$f_X(x;n,p) = n!\prod_{i=1}^{k} p_i^{x_i}/x_i! \tag{4.21}$$

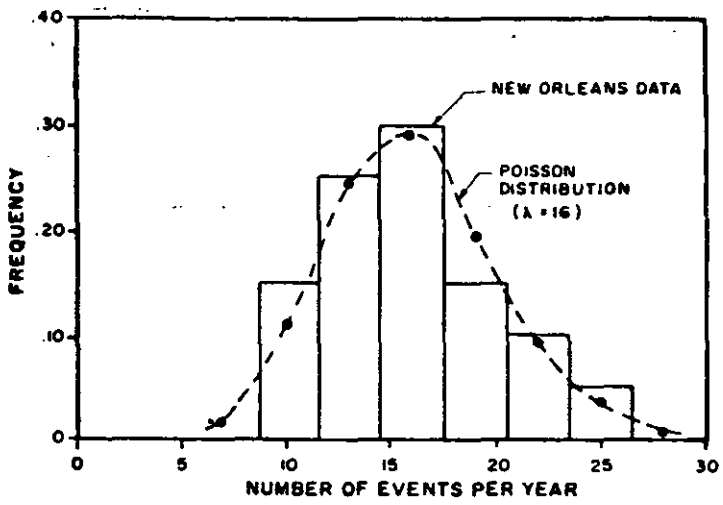


Fig. 4.2. Distribution of occurrences of warm season rainfall in which the areal mean of five gages in New Orleans, Louisiana, exceeded 0.50 inches and at least one gage recorded more than 1.0 inch. (Fogel et al. 1974).

where $\underline{X}$, $\underline{x}$ and $\underline{p}$ are $1xk$ vectors. Some restrictions on this distribution are

$$\Sigma_{i=1}^{k} p_i = 1 \text{ and } \Sigma_{i=1}^{k} x_i = n$$

The mean and variance of the multinomial distribution are

$$E(X_i) = np_i \tag{4.22}$$

$$Var(X_i) = np_i(1-p_i) \tag{4.23}$$

Example 4.17. On a certain stream the probability that the maximum peak flow during a 1-year period will be less than 5,000 cfs is 0.2 and the probability that it will be between 5,000 cfs and 10,000 cfs is 0.4. In a 20-year period, what is the probability of 4 peak flows less than 5,000 cfs and 8 peak flows between 5,000 and 10,000 cfs?

Solution: To apply the multinomial distribution we define the third event as a peak flow in excess of 10,000 cfs. This event has probability 1 - 0.2 - 0.4 = 0.4. The event of a peak flow greater than 10,000 cfs must occur 20 - 4 - 8 = 8 times. The desired probability is

$$f(4, 8, 8; 20, 0.2, 0.4, 0.4) = 20!(0.2)^4(0.4)^8(0.4)^8/4!8!8!$$

$$= 0.043$$

Comment: The expected result from 20 years of flood peak data would be

$$E(x_1) = np_1 = 20(0.2) = 4$$

$$E(x_2) = np_2 = 8$$

$$E(x_3) = np_3 = 8$$

This problem demonstrates that even though the expected results are 4, 8, 8 the probability of this happening is very low.

Exercises

4.1 Compute the terms of the binomial distribution with n = 10 and p = 0.2. Plot in the form of a histogram.

4.2 Compute the terms of the cumulative binomial with n = 10 and p = 0.2. Plot the terms.

4.3 If a project is designed on a 10-year return period, what is the probability of at least 1 exceedance during the 10-year life of the project?

4.4 What design return period should be used to insure a 95 percent chance that the design will not be exceeded in a 25-year period?

4.5 Construct a curve relating the design return period to the life of a project when a 90 percent chance of no exceedance is used.

4.6 What design return period should be used to insure a 50 percent chance of no exceedance in a 10-year period?

4.7 What design return period should be used to insure a 75 percent chance of no more than 1 exceedance in 10 years?

4.8 Construct an example where the Poisson is not a good approximation for the binomial.

4.9 In a certain locality contractors A, B and C get about 50, 25 and 25 percent, respectively, of all water resources projects. Five contracts are coming up for bid. What is the probability that contractor A will get all 5 jobs? What is the probability that A will get 2 jobs and B will get 2 jobs?

4.10 In 100 years the following number of floods were recorded at a specific location. Draw a relative frequency histogram of the data. Fit a Poisson distribution to the data and plot the relative frequencies according to the Poisson distribution on the histogram. Is the Poisson a good approximation for the data?

No. of floods	No. of occurrences
0	52
1	28
2	12
3	5
4	2
5	1
6	0

4.11 Based on a F. ..n approximation to the data of exercise 4.10, what is the probability of 5 successive years without a flood?

4.12 Based on a Poisson approximation to the data of exercise 4.10, what is the probability of exactly five years between floods?

4.13 Compute the probability of at least 1 n-year event in a k-year period using (a) n = 100, k = 20 (b) n = 500, k = 50.

4.14 Using the Poisson approximation to the binomial distribution show that the probability of at least one occurrence of a T-year event in T years is 0.632.

4.15 The Bernoulli distribution is given by

$$f(x) = p^x (1 - p)^{1-x} \qquad x = 0,1$$

What is E(x) and Var(x) for this distribution?

4.16 Use the Poisson distribution to approximate the binomial distribution of exercise 4.1. Plot the terms of this Poisson distribution on the histogram of exercise 4.1.

4.17 Two widely separated watersheds are selected for a study on peak discharges. If the occurrence of flood flows on the two basins can be considered as independent events, what is the probability of experiencing a total of 5, 20-year events on the two watersheds in a 10-year period?

4.18 A well-known scientist has predicted that during a certain three-year period a severe drought will occur on the plains east of the Rocky Mountains. He made this prediction based on his observance of sunspot activity. If the probability of a drought is 0.10 in any year, what is the probability that the scientist's prediction will come true if the occurrence of a drought is a strictly random phenomena unrelated to sunspot activity?

4.19 In a certain region there are 20 possible small watersheds suitable for a research project. Unknown to the project manager, 6 of these basins have subsurface geological features that permit large quantities of surface water to enter underground formations and leave the basin via subsurface flow. The project manager wants to select 6 watersheds from the 20 for study. (a) What is the probability that 1 of the basins having the above described geologic features will be selected? (b) What is the probability that 3 of these basins will be selected? (c) What is the probability that at least one of the basins will be selected? (d) What is the probability that all of these basins will be selected?

4.20 In the situation described in exercise 4.19 the project manager wants to pick 3 pairs of watersheds for the evaluation of an evapotranspiration suppressant. One basin in each pair will be used for a control and one will be treated with the suppressant. What is the probability that all of the control watersheds will have the geologic problem while all of the rest will not?

4.21 It is desired to model the number of rainy days in July and August as a Bernoulli process. Based on the data below and the assumption that the Bernoulli ..el is applicable: (a) What is the probability of 10 or more rainy days in each of the months of July and August? (b) What is the probability of 20 rainy days in the two-month period? (c) What assumptions concerning the Bernoulli process are likely violated by this problem? For this problem write answers in terms of summations. Do not evaluate the summations.

Year	1	2	3	4	5	6	7	8	9	10
					No. of Rainy Days					
July	10	15	17	8	9	10	17	14	20	4
August	4	9	8	3	0	10	12	2	8	6

4.22 For the binomial distribution show that $f_X(x; n, p) = f_X(x-1; n-1, p)\, f_X(1; 1, p) + f_X(x; n-1, p)\, f_X(0; 1, p)$. Write out a narrative description of the meaning of this equation.

4.23 Work exercise 4.21 using the Poisson distribution to approximate the binomial.

4.24 Pool the data of exercise 4.21 so that a single estimate is obtained for p of the binomial distribution. Compute the probability of 20 rainy days in the two month period of July-August. Compare this probability to the one computed in part b of exercise 4.21. Which answer would you prefer?

4.25 Using the data of exercise 4.21, what is the probability that the sixth wet day of August occurs on August 29, 30 or 31?

4.26 Show that for the Poisson process the time for n occurrences follows the gamma distribution. (Hint: Use the method of derived distributions to find the distribution of the time to 2 occurrences. Using the distribution of the time to 2 occurrences the method of derived distributions can be used to get the time to 3 occurrences. This process can then be repeated until a pattern emerges. Induction could also be used by showing that if the time for n-1 occurrences is given by equation 4.20 by substituting n-1 for n then the time for n occurrences is given by equation 4.20. Also the time for 1 occurrence is given by equation 4.19 which is the same as equation 4.20 with n = 1.)

5. Normal Distribution

THE MOST widely used and most important continuous probability distribution is the Gaussian or normal distribution. The normal distribution has been widely used because of its early connection with the "Theory of Errors" and because it has certain useful mathematical properties. Many statistical techniques such as analysis of variance and the testing of certain hypotheses rely on the assumption of normality. The errors involved in incorrectly assuming normality (purposely or unknowingly) depend on the use under consideration. Many statistical methods derived under the assumption of normality remain approximately valid when moderate departures from normality are present and as such are said to be robust.

The very name "normal" distribution is misleading in that it implies that random variables that are not normally distributed are abnormal in some sense. We will see from the Central Limit Theorem the conditions under which a random variable can be expected to be normally distributed. In a strict theoretical sense, most hydrologic variables cannot be normally distributed because the range on any random variable that is normally distributed is the entire real line ($-\infty$ to $+\infty$). Thus non-negative variables such as rainfall, streamflow, reservoir storage, etc., cannot be strictly normally distributed. However, it will also be shown that if the mean of a random variable is 3 or 4 times greater than its standard deviation, the probability of a normal random variable being less than zero is very small and can in many cases be neglected.

GENERAL NORMAL DISTRIBUTION

The normal distribution is a two parameter distribution whose density function is

$$p_X(x) = (2\pi\theta_2^2)^{-1/2}\, e^{-\frac{1}{2}(x-\theta_1)^2/\theta_2^2} \qquad -\infty < x < \infty$$

In examples 3.3 and 3.5 it was shown that if either the method of moments or the method of maximum likelihood is used to estimate the two parameters of this distribution, the result is $\theta_1 = \mu$ and $\theta_2^2 = \sigma^2$. For this reason the normal distribution is generally written as

$$p_X(x) = (2\pi\sigma^2)^{-1/2}\, e^{-\frac{1}{2}(x-\mu)^2/\sigma^2} \qquad -\infty < x < \infty \tag{5.1}$$

Thus the normal distribution is a two parameter distribution which is bell-shaped, continuous and symmetrical about μ (the coefficient of skew is zero). If μ is held constant and σ^2 varied, the distribution changes as in figure 5.1. If σ^2 is held constant and μ varied, the distribution does not change scale but does change location as in figure 5.2. The parameters μ and σ^2 are sometimes denoted as location and scale parameters. A common notation for indicating that a random variable is normally distributed with mean μ and variance σ^2 is $N(\mu,\sigma^2)$.

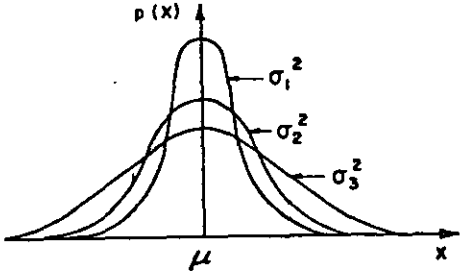


Fig. 5.1. Normal distributions with same mean and different variances.

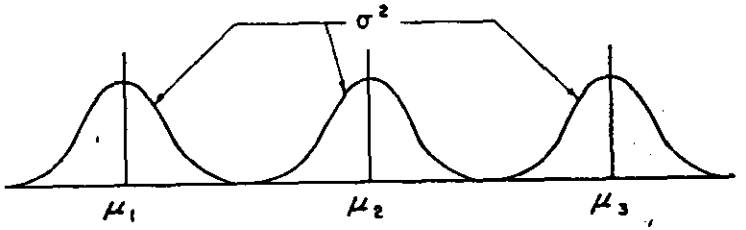


Fig. 5.2. Normal distributions with same variances and different means.

REPRODUCTIVE PROPERTIES

If a random variable X is $N(\mu,\sigma^2)$ and $Y = a + bX$, the distribution of Y can be shown to be $N(a + b\mu, b^2\sigma^2)$. This can be proven using the method of derived distributions. Furthermore if X_i for $i = 1, 2, \ldots, n$, are independently and normally distributed with mean μ_i and variance σ_i^2, then $Y = a + b_1X_1 + b_2X_2 + \ldots + b_nX_n$ is normally distributed with

$$\mu_Y = a + \Sigma_{i=1}^{n}\, b_i\mu_i \tag{5.2}$$

and

$$\sigma_Y^2 = \Sigma_{i=1}^{n}\, b_i^2\sigma_i^2 \tag{5.3}$$

Any linear function of independent normal random variables is also a normal random variable. The proof of this can be easily developed based on the moment generating function of the normal distribution. (See equation 6.99).

Example 5.1. If x_i is a random observation from the distribution $N(\mu,\sigma^2)$, what is the distribution of $\bar{X} = \sum_{i=1}^{n} x_i/n$?

Solution: $\bar{X}$ is a linear function of x_i given by $\bar{X} = (x_1 + x_2 + \ldots + x_n)/n$. From equations 5.2 and 5.3 and the reproductive properties of the normal distribution, $\bar{X}$ is normally distributed with mean

$$\mu_{\bar{X}} = a + \sum_{i=1}^{n} b_i \mu_i = 0 + \sum_{i=1}^{n} \mu/n = n\mu/n = \mu$$

and variance

$$\sigma^2_{\bar{X}} = \sum_{i=1}^{n} b_i^2 \sigma_i^2 = \sum_{i=1}^{n} \sigma^2/n^2 = n\sigma^2/n^2 = \sigma^2/n$$

Therefore $\bar{X}$ is $N(\mu,\sigma^2/n)$.

STANDARD NORMAL DISTRIBUTION

The probability that X is less than or equal to x when X is $N(\mu,\sigma^2)$ can be evaluated from

$$\text{prob}(X \le x) = P_X(x) = \int_{-\infty}^{x} (2\pi\sigma^2)^{-1/2} e^{-1/2(t-\mu)^2/\sigma^2} dt \qquad (5.4)$$

Unfortunately equation 5.4 cannot be evaluated analytically so that approximate methods of integration are required. If a tabulation of the integral was made, a separate table would be required for each value of μ and σ^2. By using the linear transformation $Z = (X - \mu)/\sigma$, the random variable Z will be N(0,1). This is a special case of $a + bx$ with $a = -\mu/\sigma$ and $b = 1/\sigma$. The random variable Z is said to be standardized (has $\mu = 0$ and $\sigma^2 = 1$) and N(0,1) is said to be the standard normal distribution. The standard normal distribution is given by

$$p_Z(z) = (2\pi)^{-1/2} e^{-z^2/2} \qquad -\infty < z < \infty \qquad (5.5)$$

and the cumulative standard normal is given by

$$\text{prob}(Z \le z) = P_Z(z) = \int_{-\infty}^{z} (2\pi)^{-1/2} e^{-t^2/2} dt \qquad (5.6)$$

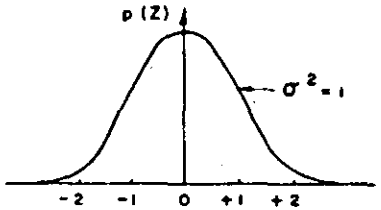


Fig. 5.3. Standard normal distribution ($\mu = 0$, $\sigma^2 = 1$).

Figure 5.3 shows the standard normal distribution which along with the transformation $Z = (X - \mu)/\sigma$ contains all of the information shown in figures 5.1 and 5.2. Both $p_Z(z)$ and $P_Z(z)$ are widely tabulated. Most tables utilize the symmetry of the normal distribution so that only positive values of Z are shown. Tables of $P_Z(z)$ may show $\text{prob}(Z \le z)$ or $\text{prob}(0 \le Z \le z)$. Care must be exercised when using normal probability tables to see what values are tabulated. Table E.3 is a table of $p_Z(z)$ for $z \ge 0$ and table E.4 contains $\text{prob}(0 \le Z \le z)$.

By studying table E.4 it can be seen that 68.26% of the normal distribution is within 1 standard deviation of the mean, 95.44% within 2 standard deviations of the mean and 99.74% within 3 standard deviations of the mean. These are called the 1, 2 and 3 sigma bounds of the normal distribution. The fact that only 0.26% of the area of the normal distribution lies outside the 3 sigma bound demonstrates that the probability of a value less than $\mu - 3\sigma$ is only 0.0013 and is the justification for using the normal distribution in some instances even though the random variable under consideration may be bounded by X = 0. If μ is greater than 3σ, the chances of an X less than zero are many times negligible (this is not always true however).

Example 5.2. Compare the 1, 2 and 3 sigma bounds under the assumption of normality and under no distributional assumptions using Chebyshev's inequality.

Solution: The 1, 2 and 3 sigma bounds of $N(\mu,\sigma^2)$ contain 68.26, 95.44 and 99.72% of the distribution. Thus the probability that X deviates more than σ, 2σ, and 3σ from μ is 0.3174, 0.0456 and 0.0028 respectively.

Chebyshev's inequality says that the $\text{prob}(|X - \mu| \ge k) \le 1/k^2$. This corresponds to a probability that X deviates more than σ, 2σ and 3σ from μ of less than 1.00, less than 0.25, and less than 0.11 respectively.

Comment: By making no distributional assumptions, we are forced to make very conservative probability statements. It is to be emphasized that Chebyshev's inequality gives an upper bound to the probability and not the probability itself.

Example 5.3. As an example of using tables of the normal distribution consider a sample drawn from a N(15, 25). What is the $\text{prob}(15.6 \le X \le 20.4)$?

Solution: The desired probability could be evaluated from

$$\text{prob}(15.6 \le X \le 20.4) = \int_{15.6}^{20.4} (50\pi)^{-1/2} e^{-(x-15)^2/50} dx$$

However this integral is difficult to evaluate. Making use of the standard normal distribution, we can transform the limits on X to limits on Z and then use standard normal tables.

$x = 15.6$ transforms to $z = (15.6 - 15.0)/5 = 0.12$

$x = 20.4$ transforms to $z = (20.4 - 15.0)/5 = 1.08$

The desired probability is

$$\text{prob}(0.12 \le Z \le 1.08) = P_Z(1.08) - P_Z(0.12)$$

From table E.4 it can be seen that $P_Z(1.08) = 0.8599$ and $P_Z(0.12) = 0.5478$. The desired probability is 0.3121.

Example 5.4. What is prob$(10.5 \le X \le 20.4)$ if X is distributed N(15, 25)?

Solution:

$x = 10.5$ transforms to $z = (10.5 - 15.0)/5 = -0.9$

$x = 20.4$ transforms to $z = 1.08$.

The desired probability is the prob$(-0.9 \le Z \le 1.08)$.

Table E.4 does not contain $P_Z(z)$ for $Z<0$. We must therefore make use of the symmetry of the normal distribution. Because of symmetry

$$\text{prob}(-0.9 \le Z \le 0) = \text{prob}(0 \le Z \le 0.9)$$

$$= 0.3159$$

$$\text{prob}(-0.9 \le Z \le 1.08) = \text{prob}(-0.9 \le Z \le 0) + \text{prob}(0 \le Z \le 1.08)$$

$$= 0.3159 + 0.3599$$

$$= 0.6758$$

Comment: Many times in solving problems of this type it is useful to sketch a normal distribution and then shade in the area corresponding to the desired probability. For this problem the sketch would be as in figure 5.4.

Example 5.5. Repeat example 3.7 assuming the Kentucky River data is $N(\mu, 21{,}000^2)$.

Solution: From example 5.1 it is known that $\bar{X}$ is $N(\mu, 21{,}000^2/n)$. Therefore $Z = (\bar{X} - \mu)\sqrt{n}/21{,}000$ is N(0,1). From the problem statement $|\bar{X} - \mu| \le 10{,}000$. n must be determined so that

$$\text{prob}(-10{,}000\sqrt{n}/21{,}000 < Z < 10{,}000\sqrt{n}/21{,}000) = 0.95$$

From table E.4 it is seen that 95% of the normal distribution is enclosed by $-1.96 < Z <$

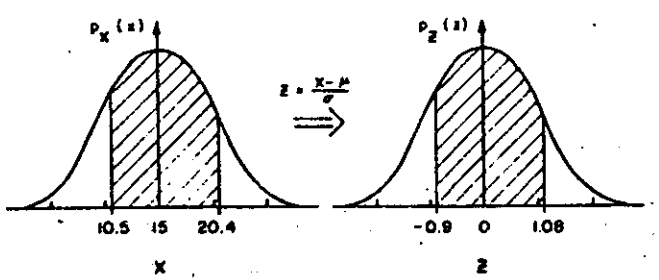


Fig. 5.4. Sketch for example 5.4.

1.96. From this n is calculated as

$$n \ge (21{,}000(1.96)/10{,}000)^2 = 16.94$$

or at least 17 observations are required to be 95% sure that $\bar{X}$ is within 10,000 cfs of μ if X is $N(\mu, 21{,}000^2)$.

Comment: By assuming normality, the required minimum number of observations has been reduced from 88 to 17. The Law of Large Numbers has placed a lower limit on n without knowledge of the distribution of X. The price for this ignorance of the distribution of X is seen to be very great if in fact X is normally distributed.

CENTRAL LIMIT THEOREM

The conditions under which a random variable might be expected to follow a normal distribution are specified by the Central Limit Theorem.

> If S_n is the sum of n independently and identically distributed random variables X_i each having a mean μ and variance σ^2, then in the limit as n approaches infinity, the distribution of S_n approaches a normal distribution with mean $n\mu$ and variance $n\sigma^2$.

In practice if the X_i are identically and independently distributed, n does not have to be very large for S_n to be approximated by a normal distribution. If interest lies in the central part of the distribution of S_n, values of n as small as 5 or 6 will result in the normal distribution producing reasonable approximations to the true distribution of S_n. If interest lies in the tails of the distribution of S_n as it often does in hydrology, larger values of n may be required.

As stated above, the Central Limit Theorem is of limited value in hydrology since most hydrologic variables are not the sum of a large number of independently and identically distributed random variables. Fortunately under some very general conditions it can be shown that if X_i for $i = 1, 2, \ldots, n$ is a random variable independent of X_j for $j \ne i$ and $E(X_i) = \mu_i$ and $\text{Var}(X_i) = \sigma_i^2$, then the sum $S_n = X_1 + X_2 + \ldots + X_n$ approaches a normal distribution with $E(S_n) = \sum_{i=1}^{n} \mu_i$ and $\text{Var}(S_n) = \sum_{i=1}^{n} \sigma_i^2$ as n approaches infinity (Thomas 1971). One condition for this generalized Central Limit Theorem is that each X_i has a negligible effect on the distribution of S_n (i.e., there cannot be one or two dominating X_i's).

This general theorem is very useful in that it says that if a hydrologic random variable is the sum of n independent effects and n is relatively large, the distribution of the variable will be approximately normal. Again how large n must be depends on the area of interest (central part or tail of the distribution) and on how good an approximation is needed.

Example 5.6. In the last chapter the gamma distribution for integer values of n was derived as the sum of n exponentially distributed random variables. The mean and variance of the exponential distribution are given as $1/\lambda$ and $1/\lambda^2$ respectively. The Central Limit Theorem gives the mean and variance of the sum of n values from the exponential distribution as n/λ and n/λ^2 for large n. This agrees with the mean and variance of the gamma distribution. In chapter 6 the coefficient of skew of the gamma distribution is given as

$2/\sqrt{n}$. Thus the sum of n random variables from an exponential distribution is a gamma distribution which approaches a normal distribution (with γ_1 approaching 0) as n gets large.

CONSTRUCTING NORMAL CURVES FOR DATA

Frequently the histogram of a set of observed data suggests that the data may be approximated by a normal distribution. One way to investigate the goodness of this approximation is by superimposing a normal curve on the frequency histogram and then visually compare the two distributions. Statistical procedures for testing the hypothesis that a set of data can be approximated by a normal (or any other) distribution are given in chapter 8.

Consider the data of table 2.1 and the frequency histogram of figure 2.5. The probability (or relative frequency) of a peak flow in any one of the class intervals assuming a normal distribution can be obtained by integrating the normal distribution over the limits of the class interval. For example the expected (according to the normal distribution) relative frequency in the first interval can be calculated from

$$f_{25,000} = \int_{19,950}^{29,950} (\sqrt{2\pi}\ 21{,}000)^{-1}\ e^{-\frac{1}{2}(x-67,500)^2/(21,000)^2}\ dx$$

since the mean of the data is 67,500 cfs and the standard deviation is 21,000 cfs. This integral is easily evaluated using standard normal tables as

$$f_{25,000} = 0.4881 - 0.4633 = 0.0248$$

An approximation to the relative frequency in a class interval can also be made by using equation 2.28b.

$$f_{x_i} = \Delta x_i\ p_X(x_i)$$

Table 5.1. Expected relative frequencies according to the normal distribution for the Kentucky River data.

Class Mark x_i	z_i	$p_Z(z_i)$		Expected Rel. Freq. f_{x_i}	Observed Rel. Freq.
25,000	-2.02	0.0519		0.0247	0.030
35,000	-1.55	0.1200		0.0571	0.045
45,000	-1.07	0.2251		0.1072	0.152
55,000	-0.60	0.3332		0.1587	0.136
65,000	-0.12	0.3961		0.1886	0.167
75,000	0.36	0.3739		0.1780	0.152
85,000	0.83	0.2827		0.1346	0.182
95,000	1.31	0.1691		0.0805	0.091
105,000	1.79	0.0804		0.0383	0.000
115,000	2.26	0.0310		0.0148	0.045
			Sum	0.9826	1.000

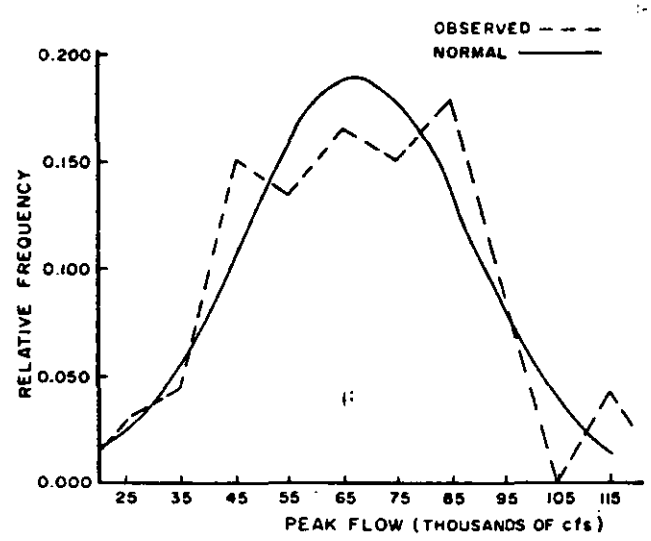


Fig. 5.5. Comparison of normal distribution with the observed distribution of peak flows in the Kentucky River.

Using the standard normal through the transformation

$$p_X(x_i) = p_Z(z_i)\ \left|\frac{dz}{dx}\right| = p_Z(z_i)/\sigma \qquad (5.7)$$

for the first class interval $\Delta x_i = 10{,}000$, $z_i = (25{,}000 - 67{,}500)/21{,}000 = -2.02$, $p_Z(z_i) = 0.0519$ (from table E.3 or equation 5.3) and σ is estimated by $s = 21{,}000$.

$$f_{25,000} = 10{,}000 \times 0.0519/21{,}000 = 0.0247$$

Similar calculations for each of the class intervals are shown in table 5.1 with the results plotted in figure 5.5. The sum of the expected relative frequencies is not one since the entire range of the normal distribution was not covered.

The procedure of integrating $p_X(x)$ over each class interval or of using equation 2.28b can be used for any continuous probability distribution to get the expected relative frequencies for that distribution.

NORMAL APPROXIMATIONS

The normal distribution can be shown to be a good approximation to several other distributions both discrete and continuous. Before using the normal to approximate some other distribution, care must be taken to see that the conditions for the approximation to be valid are met. Generally the approximations are quite good in the central part of the distribution with the accuracy dropping off in the tails of the distribution. Throughout our study of distributions, the sensitivity of the tails of distributions to distributional assumptions will be of concern. This is of particular importance in hydrology when the magnitude of a rare event is to be estimated since this estimate must come from the tail of the distribution being used.

Whenever a continuous distribution is used to approximate a discrete distribution,

corrections sometimes known as half-interval corrections must be applied to the continuous distribution. For example the probability that X is equal to some positive integer X can be evaluated for a discrete distribution. This same probability is zero if a continuous distribution is used. When a continuous distribution is used to approximate the prob(X = x), the prob(x - ½ < X < x + ½) must be evaluated. This illustrates the general rule that a ½ interval correction must be added to the upper limit and subtracted from the lower limit. The prob(X = x, x + 1, x + 2, ... y) in a discrete case is approximated by prob(X - ½< X< y + ½) in the continuous case. The prob(X≤x) in a discrete case is approximated by prob(X≤x + ½) in the continuous case. More examples of these corrections are shown in table 5.2.

The Central Limit Theorem provides the mechanism by which the normal distribution becomes an approximation for several other distributions.

Binomial Distribution

It was stated in Chapter 4 that if X is a binomial random variable with parameters n_1 and p and Y is a binomial random variable with parameters n_2 and p then Z = X + Y is a binomial random variable with parameters $n = n_1 + n_2$ and p. Extending this to the sum of several binomial random variables, the Central Limit Theorem would indicate that the normal distribution approximates the binomial distribution if n is large. Thus as n gets large the distribution of

$$Z = (X - \mu)/\sigma = (X - np)/\sqrt{np(1-p)} \qquad (5.8)$$

approaches a N(0,1). This is sometimes known as the DeMoivre-Laplace limit theorem (Mood et al. 1974).

Example 5.7. X is a binomial random variable with n = 25 and p = 0.3. Compare the binomial and normal approximation to the binomial for evaluating the prob(5< X≤8).

Solution: Using the binomial distribution this is equivalent to

$$\Sigma_{i=6}^{8} f_X(i;25,0.3) = 0.483$$

Using the normal this is approximated by prob(5.5<X<8.5).

The lower limit on Z, the standard normal variable is

$$Z = (5.5 - np)/\sqrt{np(1-p)} = (5.5 - 25(.3))/\sqrt{25(.3)(.7)} = -0.873$$

Table 5.2. Corrections for approximating a discrete random variable by a continuous random variable.

Discrete	Continuous
X=x	x-½<X<x+½
x ≤ X≤y	x-½<X<y+½
X<x	X<x+½
X≥x	X≥x-½
X> x	X≤ x-½
X≥ x	X≥x+½

The upper limit is

$$Z = (8.5 - 25(.3))/\sqrt{25(.3)(.7)} = 0.435$$

For a N(0,1) the prob(-0.873<Z<0.435) is 0.308 + 0.168 = 0.476. Therefore the exac probability of 0.483 is approximated by the normal to be 0.476 for an n of 25.

Negative Binomial Distribution

Following reasoning similar to that given for the binomial distribution, the negativ binomial distribution with large k can be approximated by a normal distribution. In th case of the negative binomial the distribution of

$$Z = (X - \mu)/\sigma = (X - k/p)/\sqrt{kq/p^2} \qquad (5.1$$

approaches N(0,1) as k gets large.

Example 5.8. Work example 4.11 using the normal approximation for the negative b nomial.

Solution: The desired probability is prob(39.5<X<40.5). Using the standard normal di tribution, the limits on Z are

$$Z = (39.5 - 4.0/0.1)/\sqrt{4(0.9)/0.01} = -0.5/18.97 = -0.026$$

$$Z = (40.5 - 4.0/0.1)/\sqrt{4(0.9)/0.01} = +0.026$$

prob(-0.026 < Z < 0.026) = 0.0208

This compares favorably with the 0.0206 computed using the negative binomial.

Poisson Distribution

The sum of two Poisson random variables with parameters λ_1 and λ_2 is also Poisson random variable with parameter $\lambda = \lambda_1 + \lambda_2$. Extending this to the sum of large number of Poisson random variables, the Central Limit Theorem indicates that f large λ, the Poisson may be approximated by a normal distribution. In this case the d tribution of

$$Z = (X - \mu)/\sigma = (X - \lambda)/\lambda^{1/2} \qquad (5.1$$

approaches a N(0,1). Since the Poisson is the limiting form of the binomial and the b nomial can be approximated by the normal, it is no surprise that the Poisson can also b approximated by the normal.

Continuous Distributions

Many continuous distributions can be approximated by the normal distribution f certain values of their parameters. For instance in example 5.4, it was shown that f large n the gamma distribution approaches the normal distribution. To make these a proximations one merely equates the mean and variance of the distribution to be a proximated to the mean and variance of the normal and then uses the fact that

$Z = (X - \mu)/\sigma$

is N(0,1) if X is $N(\mu,\sigma^2)$. Not all continuous distributions can be approximated by the normal and for those that can the approximation is only valid for certain parameter values. Things to look for are parameters that produce near zero skew, symmetry and tails that asymptotically approach $p_X(x) = 0$ as X approaches large and small values.

Exercises

5.1 Consider sampling from a normal distribution with a mean of 0 and a variance of 1. What is the probability of selecting (a) an observation between 0.5 and 1.5? (b) an observation outside the interval -0.5 to +0.5? (c) 3 observations inside and 2 observations outside the interval of 0.5 and 1.5? (d) 4 observations inside the interval 0.5 to 1.5 exactly two of which are not in the interval -0.5 to1.0?

5.2 What is the probability of selecting an observation at random from a N(100, 2500) that is (a) less than 75? (b) equal to 75?

5.3 For the Kentucky River data of table 2.1, what is the probability of a peak flow exceeding 100,000 cfs if the peaks are assumed to be normally distributed?

5.4 Construct the theoretical distribution for the data of exercise 2.2 if it is assumed that the data are normally distributed. From a visual comparison with the data histogram, would you say the data are normally distributed?

5.5 Work exercise 4.1 using the normal approximation to the binomial and plot the results on the histogram developed for exercise 4.1.

5.6 Show that if X is $N(\mu,\sigma^2)$ then $Y = a + bX$ is $N(a + b\mu, b^2\sigma^2)$.

5.7 For a particular set of data the coefficient of variation is 0.4. If the data are normally distributed, what percent of the data will be less than 0.0?

5.8 A sample of 150 observations has a mean of 10,000, a standard deviation of 2,500 and is normally distributed. Plot a frequency histogram showing the number of observations expected in each interval.

5.9 Appendix C contains a listing of the annual runoff from Cave Creek watershed near Fort Spring, Kentucky. What is the probability that the true mean annual runoff is less than 14.0 inches if one can assume the true variance is 22.56 inches^2? What other assumptions are needed?

5.10 Random digits are the numbers 0, 1, 2, ..., 9 selected in such a fashion that each is equally likely (i.e., has probability 1/10 of being selected). An experiment is performed by selecting 5 random digits, adding them together and calling their sum X. The experiment is repeated 10 times and $\bar{X}$ is calculated. What is the probability that $\bar{X}$ is less than 21.5? (Exercise 13.9 requires that this experiment be carried out.)

5.11 Plot the individual terms of the Poisson distribution for $\lambda = 2$. Approximate the Poisson by the normal and plot the normal approximations on the same graph.

5.12 Repeat exercise 4.11 for $\lambda = 9$.

5.13 Assume the data of exercise 4.21 is normally distributed. (a) Within each month what is the probability of 10 or more rainy days? (b) What is the probability of 20 or more rainy days in the July-August period? (c) What is the difference in assuming the data are normally distributed and in assuming the data are binomially distributed and approximating the binomial with the normal?

5.14 Plot the observed frequency histogram and the frequency histogram expected from the normal distribution for the annual flows for the following rivers. Discuss how well the normal approximates the data in terms of the coefficient of variation and skewness. (Note data are in Appendix C).

a) North Llano River near Junction, Texas
b) Spray River, Banff, Canada
c) Piscataquis River near Dover-Foxcroft, Maine.

5.15 The occurrence of rainstorms is sometimes considered to be a Poisson process so that the time between rainstorms is exponentially distributed. If for a certain locality the mean of this exponential distribution is 10 days, what is the probability that the elapsed time for 15 storms to occur will exceed 120 days?

5.16 Lane and Osborn (1973) present the following data for the mean number of days with more than 0.10 inches of precipitation at Tombstone, Arizona. If the occurrence of more than 0.10 inches of rain in any month can be considered as an independent Poisson process, what is the probability of fewer than 30 days with more than 0.10 inches of rain in one year at Tombstone?

Month	No. Days	Month	No. Days
Jan.	2	July	7
Feb.	2	Aug.	7
Mar.	2	Sept.	3
Apr.	1	Oct.	2
May	0	Nov.	2
June	2	Dec.	2
		Total	32

5.17 An experimenter is measuring the water level in an experimental towing channel. Because of waves and surges, a single measurement of the water level is known to be inaccurate. Past experience indicates the variance of these measurements is 0.0025 ft^2. How many independent observations are required to be 90% confident that the mean of all the measurements will be within .02 feet of the true water level?

5.18 At a certain location the annual precipitation is approximately normally distributed with a mean of 45 inches and a standard deviation of 15 inches. Annual runoff can be approximated by $R = -7.5 + 0.5P$ where R is annual runoff and P is annual precipitation. What is the mean and variance of annual runoff? What is the probability that the

annual runoff will exceed 20 inches?

5.19 Plot a frequency distribution for a mixture of two normal distributions. Use as the first distribution a N(0,1) and as the second a N(1,1). Use as values for the mixing parameter 0.2, 0.5 and 0.8.

6. Some Continuous Probability Distributions

THERE ARE many continuous probability distributions in addition to the normal distribution. This chapter covers some of these distributions, methods for estimating their parameters, properties of the distributions, and potential applications for them. Further discussion on distribution selection is contained in Chapter 7.

UNIFORM DISTRIBUTION

If a continuous random process is defined over an interval a to b and the probability of an outcome of this process being in a subinterval of a to b is proportional to the length of the subinterval, the process is said to be uniformly distributed over the interval a to b. The probability density function for the continuous uniform distribution is

$$p_X(x) = 1/(\beta - \alpha) \text{ for } \alpha \le x \le \beta \tag{6.1}$$

and the cumulative distribution function is

$$P_X(x) = (x - \alpha)/(\beta - \alpha) \text{ for } \alpha \le x \le \beta \tag{6.2}$$

The mean and variance of the uniform distribution are

$$E(X) = (\beta + \alpha)/2 \tag{6.3}$$

$$\text{Var}(X) = (\beta - \alpha)^2/12 \tag{6.4}$$

The methods of moments yields the following estimators for the parameters α and β:

$$\hat{\alpha} = \bar{X} - \sqrt{3}\,S$$
$$\hat{\beta} = \bar{X} + \sqrt{3}\,S \tag{6.5}$$

The method of maximum likelihood when applied to the uniform distribution results in the estimators for α and β being the smallest and largest sample values respectively. That this is the case can be seen by writing out the likelihood function and then selecting those values of α and β (within the constraints that $\alpha \le X \le \beta$ for all X) that maximize the function.

The uniform distribution finds its greatest application as the distribution of $P_X(x)$ for all probability density functions. That is the prob$(P_X(x) \le y)$ is uniformly distributed over the interval $0 \le y \le 1$ for any continuous probability distribution. We will use this fact in Chapter 13 for generating random observations from probability distributions.

Example 6.1. Use the method of moments to estimate the parameters of the uniform distribution based on the following sample: 1,4,3,4,5,6,7,6,9,5. What are the maximum likelihood estimators for this sample?

Solution: By method of moments

$\bar{x} = 5.00$ and $s = 2.26$

$\hat{\beta} = \bar{x} + \sqrt{3}\, s = 8.83$

$\hat{\alpha} = \bar{x} - \sqrt{3}\, s = 1.17$

By maximum likelihood

$\hat{\alpha} = 1.00$ (smallest sample value)

$\hat{\beta} = 9.00$ (largest sample value)

Comment: This problem illustrates that the method of moments and the method of maximum likelihood do not always produce the same parameter estimates. In this case the parameters estimated by moments are not reasonable since values of X outside the limits of $\hat{\alpha}$ and $\hat{\beta}$ are present in the sample. This is a common problem when the method of moments is used to estimate the parameters of the uniform distribution for small samples. Of course for large samples neither the moment nor the maximum likelihood estimates will be "good" if the sample is not truly a random sample from a uniform distribution.

EXPONENTIAL DISTRIBUTION

In chapter 4 it was shown that the exponential distribution arises as the probability distribution of the time between occurrences of events of a Poisson process. The exponential distribution has also been used as the time-to-failure distribution for machine components. The exponential density function is given by

$$p_X(x) = \lambda e^{-\lambda x} \qquad x > 0,\ \lambda > 0 \tag{6.6}$$

and the cumulative exponential by

$$P_X(x) = \int_0^x \lambda e^{-\lambda t}\, dt = 1 - e^{-\lambda x} \qquad x > 0 \tag{6.7}$$

The mean and variance of the exponential distribution are

$$E(X) = 1/\lambda \tag{6.8}$$

$$\mathrm{Var}(X) = 1/\lambda^2 \tag{6.9}$$

The coefficient of skew is a constant 2 indicating the exponential is skewed to the right for all values of λ. The curve labeled $\alpha_1 = 1$ in figure 6.2 is an exponential distribution with $\lambda = 1$. Examples 3.2 and 3.4 demonstrated that when either the method of moments or maximum likelihood is used for parameter estimation, the result is

$$\hat{\lambda} = 1/\bar{X} \tag{6.10}$$

or the parameter λ may be estimated by the reciprocal of the sample mean.

Example 6.2. Haan and Johnson (1967) studied the physical characteristics of depressions in north-central Iowa. The data tabulated below shows the number of depressions falling into various classes based on the surface area of the depression. Plot a relative frequency histogram of the data. Superimpose on the histogram the best fitting exponential distribution. Estimate the probability that a depression selected at random will have an area greater than 2.25 acres.

Area (acres)	No. depressions
0 - ½	106
½ - 1	36
1 - 1½	18
1½ - 2	9
2 - 2½	12
2½ - 3	2
3 - 3½	5
3½ - 4	1
4 - 4½	4
4½ - 5	5
5 - 5½	2
5½ - 6	6
6 - 6½	3
6½ - 7	1
7 - 7½	1
7½ - 8	1

Solution: The relative frequencies are computed by dividing the number of depressions in each class by the total number of depressions. The best fitting exponential is estimated by using equation 6.10 to estimate the exponential parameter λ. $\bar{X}$ is calculated from equation 3.18 as 1.27 acres. Then $\hat{\lambda} = 1/\bar{x} = 0.787$. The expected relative frequency in each class is then calculated from equation 2.28b as

$$f_{x_i} = \Delta x_i\, p_A(x_i)$$

where x_i is the midpoint of the class interval, $\Delta x_i = ½$, and $p_A(x_i)$ is the exponential distribution of area given by

$$p_A(x_i) = \hat{\lambda} e^{-\hat{\lambda} x_i}$$

Therefore

$$f_{x_i} = ½ \times 0.787\, e^{-0.787 x_i}$$

For example for the second class interval

$$f_{0.75} = 0.393\, e^{-0.787(.75)} = 0.22$$

compared to an observed value of 0.17.

The estimated probability that a depression will have an area in excess of 2.25 acres is

$$\text{prob}(A > 2.25) = 1 - \text{prob}(A \leq 2.25) = 1 - P_A(2.25)$$

$$= 1 - (1 - e^{-.787(2.25)}) = 0.170$$

The observed fraction of depressions with areas in excess of 2.25 acres is 0.197.

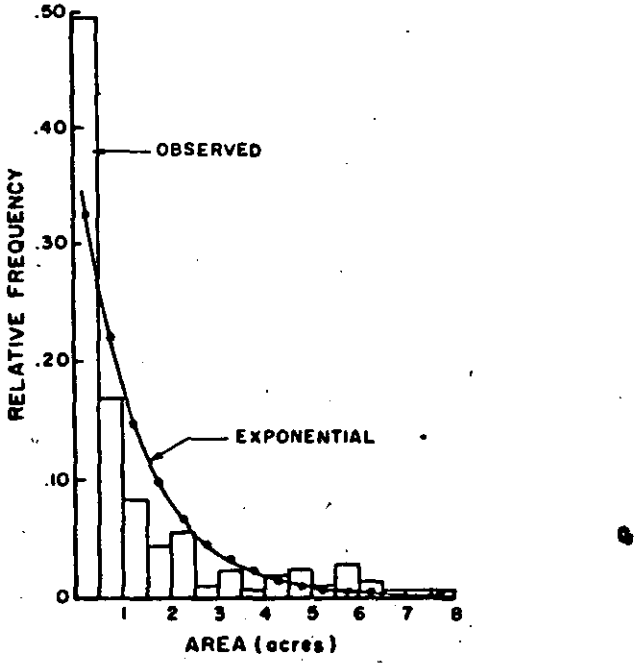


Fig. 6.1. Observed and expected (according to the exponential distribution) number of depressions in various size categories for example 6.2.

GAMMA DISTRIBUTION

The distribution of the sum of n exponentially distributed random variables each with parameter λ was given in chapter 4 as a gamma distribution with parameters n and λ. In general n does not have to be an integer. The gamma density function is given by

$$p_X(x) = \lambda^\eta x^{\eta-1} e^{-\lambda x} / \Gamma(\eta) \qquad x, \lambda, \eta > 0 \tag{6.11}$$

$\Gamma(\eta)$ is the gamma function having the properties

$$\Gamma(\eta) = (\eta - 1)! \text{ for } \eta = 1,2,3\ldots.$$

$$\Gamma(\eta + 1) = \eta\Gamma(\eta) \text{ for } \eta > 0 \tag{6.12}$$

$$\Gamma(\eta) = \int_0^\infty t^{\eta-1} e^{-t}\, dt \text{ for } \eta > 0$$

$$\Gamma(1) = \Gamma(2) = 1;\ \Gamma(½) = \sqrt{\pi}$$

Table E.12 contains $\Gamma(\eta)$ for $1 \leq \eta \leq 2$. For other values of η, one of the above equations can be used.

The mean, variance and coefficient of skew for the gamma distribution are

$$E(X) = \eta/\lambda \tag{6.13}$$

$$\text{Var}(X) = \eta/\lambda^2 \tag{6.14}$$

$$\gamma = 2/\sqrt{\eta} \tag{6.15}$$

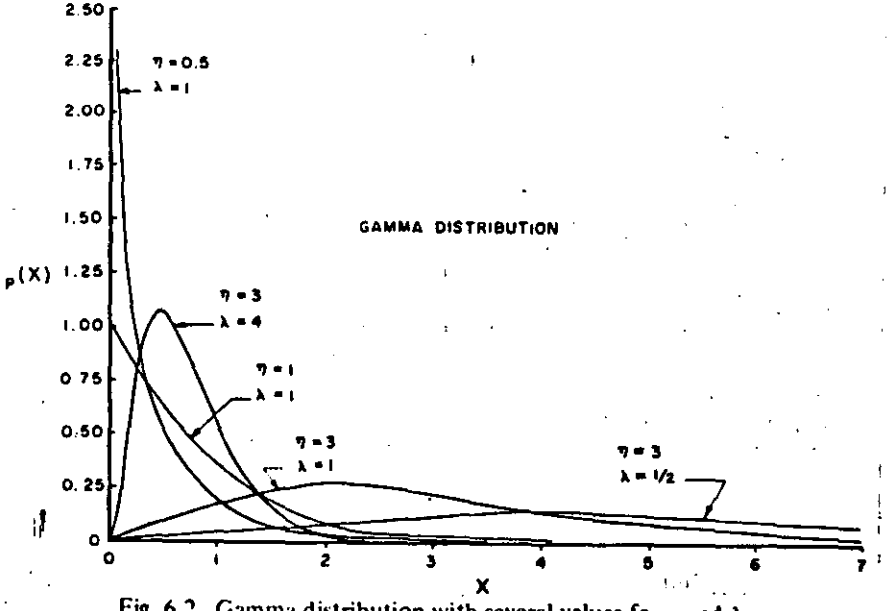


Fig. 6.2. Gamma distribution with several values fc id λ.

The gamma distribution is positively skewed with γ decreasing as η increases. Plots of the distribution for various values of η and λ are shown in figure 6.2. A wide variety of shapes ranging from reverse J-shaped for $\eta \leq 1$ to single peaked with the peak (mode) at $x = (\eta - 1)/\lambda$ for $\eta > 1$ can be produced by the gamma density function. Changing λ and holding η constant changes the scale of the distribution while changing η and holding λ constant changes the shape of the distribution. Thus λ and η are sometimes known as scale and shape parameters.

The cumulative gamma distribution is

$$P_X(x) = \int_0^x \lambda^\eta t^{\eta-1} e^{-\lambda t}/\Gamma(\eta)\, dt \qquad (6.16)$$

which can be evaluated using a table of the incomplete gamma function such as table E.8. If η is an integer, the cumulative gamma distribution is given by (Mood et al. 1974)

$$P_X(x) = 1 - e^{-\lambda x} \sum_{j=0}^{\eta-1} (\lambda x)^j/j!$$

The exponential distribution is a special case of the gamma distribution with $\eta = 1$. If X and Y are independent gamma random variables with parameters η_1, λ and η_2, λ respectively then $Z = X + Y$ is a gamma variable with parameters $\eta = \eta_1 + \eta_2$ and $\lambda = \lambda$. This can be extended to the sum of any number of independent gamma random variables having a common parameter λ. It is an expected result since in Chapter 4 the gamma distribution was shown to arise as the distribution of the sum of n independent exponential random variables.

The moment estimators for the parameters of the gamma distribution result from equations 6.13 and 6.14 as

$$\hat{\lambda} = \bar{X}/S^2$$
$$\hat{\eta} = \bar{X}^2/S^2 \qquad (6.17)$$

The maximum likelihood estimators for λ and η are given by

$$\ln \hat{\eta} - \Psi(\hat{\eta}) = \ln(\bar{X}/\bar{X}_G)$$
$$\hat{\eta}/\hat{\lambda} = \bar{X} \qquad (6.18)$$

Table 6.1. Correction factor for the maximum likelihood estimator for the parameter η of the gamma distribution.

$\hat{\eta}$	$\Delta\hat{\eta}$	$\hat{\eta}$	$\Delta\hat{\eta}$	$\hat{\eta}$	$\Delta\hat{\eta}$
0.2	0.034	1.0	0.009	1.8	0.004
0.3	0.029	1.1	0.008	1.9	0.003
0.4	0.025	1.2	0.007	2.2	0.003
0.5	0.021	1.3	0.006	2.3	0.002
0.6	0.017	1.4	0.006	3.1	0.002
0.7	0.014	1.5	0.005	3.2	0.001
0.8	0.012	1.6	0.005	5.5	0.001
0.9	0.011	1.7	0.004	5.6	0.000

where $\bar{X}_G$ is the sample geometric mean and $\Psi(x) = d \ln\Gamma(x)/dx$ is the psi-function. Thom (1958) has proposed an approximate relationship based on the truncation of a series expansion of the maximum likelihood estimator for η given by

$$\hat{\eta} = (1 + \sqrt{1+4y}/3)/4y - \Delta\hat{\eta} \qquad (6.19)$$

where y is $\ln\bar{X} - \overline{\ln X}$, $\Delta\hat{\eta}$ is a correction term arising because of the truncation and $\overline{\ln X}$ is the mean natural logarithm of the observations. Table 6.1 contains the values for $\Delta\hat{\eta}$ for $\hat{\eta}$ ranging from 0.2 to 5.6. For $\hat{\eta} > 5.6$ the correction is negligible (as it is anyway for many practical situations regardless of the value of $\hat{\eta}$). The procedure for finding the correction factor is to assume that $\hat{\eta}$ is equal to the first term of equation 6.19 and use the $\Delta\hat{\eta}$ from table 6.1 corresponding to this initial estimate for $\hat{\eta}$.

The parameter λ is then estimated by

$$\hat{\lambda} = \hat{\eta}/\bar{X} \qquad (6.20)$$

Thom (1958) states that for $\eta < 10$ the method of moments produces unacceptable estimates for both λ and η. For η near 1 the method of moments uses only 50 percent of the sample information for estimating λ and only 40 percent for η. This means the maximum likelihood estimators would do as well with one half the number of observations.

Greenwood and Durand (1960) present the following rational fraction approximations for the maximum likelihood estimators

$$\hat{\eta} = (0.5000876 + 0.1648852y - 0.0544274y^2)/y \qquad (6.21)$$

for

$$0 \leq y \leq 0.5772$$

and

$$\hat{\eta} = \frac{8.898919 + 9.05995y + 0.9775373y^2}{y(17.79728 + 11.968477y + y^2)} \qquad (6.22)$$

for

$$0.5772 < y \leq 17.0$$

where

$$y = \ln\bar{X} - \overline{\ln X} \qquad (6.23)$$

λ is then estimated from equation 6.20. Greenwood and Durand (1960) state that the maximum error in equation 6.21 is 0.0088% and in equation 6.22 0.0054%.

Equations 6.19, 6.20 and 6.21 produce estimates for η and λ that have a slight asymptotic bias. For small samples the bias may be appreciable (Shenton and Bowman 1970). Bowman and Shenton (1968) present the following approximate relationship for estimating the bias in the parameter η when equation 6.19, 6.21 or 6.22 is used.

$$E(\hat{\eta} - \eta) = (3\eta - 0.677 + 0.111/\eta + 0.032/\eta^3)/(n - 3) \text{ for } \begin{matrix} n \geq 4 \\ \eta \geq 1 \end{matrix} \qquad (6.24)$$

where $E(\hat{\eta} - \eta)$ is the bias in η with error of less than 1.4%. The result of using this relationship for estimating the bias in $\hat{\eta}$ for a sample size n from a gamma distribution having a population parameter of $\eta = 2$ is shown in figure 6.3. In practice equation 6.24 can be used to correct $\hat{\eta}$ for bias. If the population η were known, there would of course be no need for estimating η.

Bowman and Shenton (1968) suggest that the bias in $\hat{\eta}$ can be approximated from

$$E(\hat{\eta} - \eta) = 3\hat{\eta}/n \qquad (6.25)$$

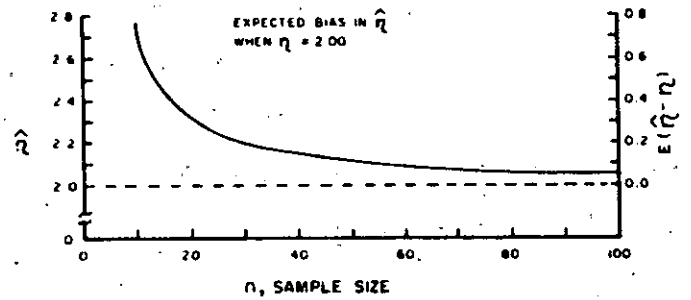
which yields

$$E(\eta) = (n - 3)\hat{\eta}/n \qquad n \geq 4$$

The gamma distribution has been widely used in hydrology. Rainfall probabilities for durations of days, weeks, months and years have been estimated by the gamma distribution (Barger and Thom 1949; Barger, Shaw and Dale 1959; Friedman and Janes 1957; Mooley and Crutcher 1968). Annual runoff (Markovic 1965) has been described by the gamma distribution.

Example 6.3. The annual water yield for Cave Creek (near Fort Spring, Kentucky) is shown in the following table. Estimate the parameters of the gamma distribution for this data using both the method of moments and the method of maximum likelihood. Assuming the data follows a gamma distribution, estimate the probability of an annual water yield exceeding 20.00 inches.

Year	Annual RO (inches)	Year	Annual RO (inches)
1953	13.26	1962	18.89
1954	3.31	1963	12.82
1955	15.17	1964	11.58
1956	15.50	1965	15.17
1957	14.22	1966	10.40
1958	21.20	1967	18.02
1959	7.70	1968	16.25
1960	17.64	1969	11.77
1961	22.91	1970	17.92



6.3. Expected bias in $\hat{\eta}$ for the gamma distribution with $\eta = 2$.

Solution: Method of Moments

$$\bar{x} = \Sigma x/n = 263.73/18 = 14.65$$

$$s = \sqrt{(\Sigma x^2 - n\bar{x}^2)/(n - 1)} = \sqrt{(4246.15 - 3863.20)/17} = 4.75$$

$$\hat{\lambda} = \bar{x}/s^2 = 14.65/22.56 = 0.649$$

$$\hat{\eta} = \bar{x}^2/s^2 = 214.62/22.56 = 9.513$$

Method of Maximum Likelihood (Thom procedure)

$$\overline{\ln x} = \Sigma_{i=1}^{18} \ln x_i/18 = 2.612$$

$$\ln\bar{x} = \ln 14.65 = 2.684$$

$$y = \ln\bar{x} - \overline{\ln x} = 2.684 - 2.612 = 0.072$$

$$\hat{\eta} = (1 + \sqrt{1 + 1.333(0.072)})/4(0.072) = 7.107$$

$$\hat{\lambda} = \hat{\eta}/\bar{x} = 7.107/14.65 = 0.485$$

Method of Maximum Likelihood (Greenwood and Durand Procedure)

$$y = \ln\bar{X} - \overline{\ln X} = 0.072$$

$$\hat{\eta} = (0.5000876 + 0.1648852y - 0.0544274y^2)/y$$

$$= 7.107$$

$$\hat{\lambda} = \hat{\eta}/X = 0.485$$

Thus the maximum likelihood estimators are $\hat{\lambda} = 0.485$ and $\hat{\eta} = 7.107$. These estimates may be corrected for bias using either equation 6.24 or 6.25. If 6.25 is used

$$E(\hat{\eta} - \eta) = 3\hat{\eta}/n = 3(7.107)/18 = 1.185$$

Therefore

$$E(\eta) = \hat{\eta} - E(\hat{\eta} - \eta) = 7.107 - 1.185 = 5.922$$

If $\eta = 5.922$ is substituted into equation 6.24, the result is $E(\hat{\eta} - \eta) = 1.141$ which is in good agreement with the 1.185 produced by equation 6.25. The final estimated for η is now $\hat{\eta} = 5.922$ and $\hat{\lambda} = \hat{\eta}/X = 0.404$. Using the method of moments the parameter estimates are $\hat{\eta} = 9.513$ and $\hat{\lambda} = 0.649$ while the maximum likelihood estimates are $\hat{\eta} = 5.922$ and $\hat{\lambda} = 0.404$. Following the recommendation of Thom (1958), the latter estimates will be used in estimating the probability of an annual water yield in excess of 20.00 inches.

$$\text{prob(yield} > 20.00) = 1 - \text{prob(yield} \leq 20.00) = 1 - P_x(20.00)$$

$$P_X(20.00) = \int_0^{20.00} \lambda^\eta x^{\eta-1} e^{-\lambda x}/\Gamma(\eta)\, dx$$

which can be evaluated using table E.8. This table contains $1 - P_X(x)$. The table is entered with

$$\chi^2 = 2\lambda x \text{ and } \nu = 2\eta$$

which are parameters required in using the tables. In our case

$$\chi^2 = 2(0.404)20 = 16.17$$

$$\nu = 2(5.922) = 11.844$$

by interpolating in the table $1 - P_X(20.00)$ is found to be 0.176 which is the desired probability. Thus the prob(yield>20.00) = 0.176 if the annual water yield follows a gamma distribution with parameters $\eta = 5.922$ and $\lambda = 0.404$.

Comment: If the moment parameter estimates had been used, the resulting probability would have been 0.130 which is reasonably close to 0.176. This is because η is reasonably close to the 10.00 that Thom (1958) suggested is the smallest value of η for which the method of moments results in good parameter estimates. For this data $C_s = 2/\sqrt{\hat{\eta}} = 0.82$ so that the distribution is moderately skewed to the right. If the normal distribution had been used to estimate prob(X>20.00) the result would have been 0.126 which again is a reasonable approximation. However if the annual water yield with a return period of 100 years or a 1 percent chance of being exceeded is evaluated by the gamma with $\eta = 5.922$ and $\lambda = 0.404$ and by the normal with $\mu = 14.56$ and $\sigma = 4.75$ the results are 32.2 inches and 25.6 inches again showing the sensitivity of estimates of rare events to the distributional assumption even though in the main body of distribution the agreement is good.

The magnitude for the 100-year event from the gamma distribution was determined by noting that a 100-year event corresponds to $1 - P_X(x) = 0.01$. Table E.8 was entered with $\nu = 2\hat{\eta} = 2(5.922) = 12$. The value of χ^2 corresponding to $1 - P_X(x) = 0.01$ is about 26. Therefore $x = \chi^2/2\hat{\lambda} = 26/2(0.404) = 32.2$.

Generally 18 observations are not enough to make reliable probability estimates or to determine the proper probability distribution to use. It is a small enough number that one can follow through all of the needed calculations for this example in a short time on a desk calculator, however. The fact that the gamma and normal estimates differ greatly for this data at large return periods does not mean the gamma (or the normal) is a better approximation for the data. This question will be taken up later. Exercise 6.21 should be consulted for another approximate solution to this example.

LOGNORMAL DISTRIBUTION

The Central Limit Theorem was used in deriving the general result that if a random variable X is made up of the sum of many small effects, then X might be expected to be normally distributed. Similarly if X is equal to the product of many small effects, that is if $X = X_1 X_2 \ldots X_n$, then the lnX can be expected to be normally distributed. This can be seen by letting $Y = \ln X$ so that $Y = \ln(X_1 X_2 \ldots X_n) = \ln X_1 + \ln X_2 + \ldots + \ln X_n$. Since the X_i are random variables, the $\ln X_i$ are also random variables and Y = lnX is a random variable made up from the sum of many other random variables. From the Central Limit Theorem, Y can be expected to be normally distributed with mean μ_y and variance σ_y^2.

$$p_Y(y) = (2\pi\sigma_y^2)^{-1/2} \exp[-\tfrac{1}{2}(y - \mu_y)^2/\sigma_y^2] \qquad -\infty < y < \infty \tag{6.26}$$

The distribution of X can be found from

$$p_X(x) = p_Y(y)\, |dy/dx|$$

Since $Y = \ln X$

$$|dy/dx| = 1/x \qquad x>0$$

and

$$p_X(x) = (2\pi x^2 \sigma_y^2)^{-1/2} \exp[-\tfrac{1}{2}(\ln x - \mu_y)^2/\sigma_y^2] \qquad x>0 \tag{6.27}$$

Note that equation 6.26 gives the distribution of Y as a normal distribution with mean μ_y and variance σ_y^2. Equation 6.27 gives the distribution of X as the lognormal distribution with parameters μ_y and σ_y^2. Y = lnX is normally distributed while X is lognormally distributed.

The parameters μ_y and σ_y^2 can be estimated by $\bar{Y}$ and S_y^2 in the usual manner by first transforming all of the X_i's to Y_i's by

$$y_i = \ln x_i$$

then

$$\bar{y} = \sum y_i/n \tag{6.28}$$

and

$$s_y^2 = (\Sigma y_i^2 - n\bar{y}^2)/(n-1) \tag{6.29}$$

with all of the summations from 1 to n. If a digital computer is used the above equations are easily applied. Chow (1954), in a comprehensive investigation of the lognormal distribution, presents the following relationships for calculating $\bar{Y}$ and S_y^2 without taking the logarithms of all of the data.

$$\bar{Y} = \tfrac{1}{2}\ln[\bar{X}^2/(C_v^2 + 1)] \tag{6.30}$$

$$S_y^2 = \ln(C_v^2 + 1) \tag{6.31}$$

where C_v is the coefficient of variation of the original data ($C_v = S_x/\bar{X}$). Brakensiek (1958) discusses a least squares procedure for estimating the parameters of the lognormal distribution.

The mean and variance of the lognormal distribution are

$$E(X) = \exp(\mu_y + \sigma_y^2/2) \tag{6.32}$$

$$\text{Var}(X) = \mu_x^2 [\exp(\sigma_y^2) - 1] \tag{6.33}$$

The coefficient of variation of the X's is

$$C_v = [\exp(\sigma_y^2) - 1]^{1/2} \tag{6.34}$$

The coefficient of skew of the X's is

$$\gamma = 3C_v + C_v^3 \tag{6.35}$$

Thus the lognormal distribution is positively skewed with the skew decreasing as the coefficient of variation decreases.

Tables of the standard normal distribution can be used to evaluate the lognormal distribution. From equation 6.27 we have $p_X(x) = p_Y(y)/x$. But $p_Y(y)$ is a normal density function. From equation 5.7, $p_Y(y) = p_Z(z)/\sigma_y$ or

$$p_X(x) = p_Z(z)/x\sigma_y \tag{6.36}$$

The prob($X \leq x$) is equal to the prob($Y \leq y$) since $Y = \ln X$ is a monotonic, single valued function. Since Y is normally distributed, prob($Y \leq y$) = prob($Z \leq z$) where

$$Z = (Y - \mu_y)/\sigma_y \tag{6.37}$$

Therefore the standard normal tables can be used to evaluate $p_X(x)$ and $P_X(x)$.

Certain reproductive properties of the lognormal follow directly from the reproductive properties of the normal distribution. For example if X is lognormally distributed then $Y = aX^b$ is lognormally distributed with

$$\mu_{\ln Y} = \ln a + b\mu_{\ln X}$$

and

$$\sigma^2_{\ln Y} = b^2\sigma^2_{\ln X}$$

This follows from the fact that $\ln Y = \ln a + b \ln X$, $\ln X$ is normally distributed, and $\ln Y$ is a linear function of $\ln X$ so is also normally distributed. Thus Y is lognormally distributed. This can be extended so that if $X_1, X_2, \ldots X_n$ are independent and lognormally distributed, then $Y = aX_1^{b_1} X_2^{b_2} \ldots X_n^{b_n}$ is lognormally distributed with

$$\mu_{\ln Y} = \ln a + \Sigma_{i=1}^{n} b_i \mu_{\ln X_i} \tag{6.38}$$

and

$$\sigma^2_{\ln Y} = \Sigma_{i=1}^{n} b_i^2 \sigma^2_{\ln X_i} \tag{6.39}$$

Two special cases of the above are if $Z = XY$ and $Z = X/Y$ with X and Y being independently and lognormally distributed. Then Z is lognormally distributed with its mean and variance easily determined from equations 6.38 and 6.39.

Because of its simplicity, readily available tables for its evaluation, and the fact that many hydrologic variables are bounded by zero on the left and positively skewed, the lognormal distribution has received wide usage in hydrology.

Example 6.4. Use the lognormal distribution and calculate the expected relative frequency for the third class interval of the data in table 5.1.

Solution: The expected relative frequency according to the lognormal distribution is

$$f_{45,000} = \Delta x \, p_X(45{,}000)$$

The evaluation of $p_X(x)$ from equation 6.36 requires an estimate for μ_y and σ_y. These are estimated from equations 6.30 and 6.31.

$$C_v = s_x/\bar{x} = 21{,}000/67{,}500 = 0.311$$

$$\bar{y} = \tfrac{1}{2} \ln[\bar{x}^2/(C_v^2 + 1)] = \tfrac{1}{2} \ln[67{,}500^2/(0.311^2 + 1)] = 11.0737$$

$$s_y = \sqrt{\ln(C_v^2 + 1)} = \sqrt{0.311^2 + 1} = 0.30395$$

Now

$$z = (\ln x - \bar{y})/s_y = (\ln 45{,}000 - 11.0737)/0.30395 = -1.182$$

so from the standard normal table we get

$$p_Z(z) = 0.198$$

and

$$p_X(x) = p_Z(z)/x\, s_y = 0.198/45{,}000(0.30395) = 1.4476 \times 10^{-5}$$

$$f_{45,000} = 10{,}000(1.4476 \times 10^{-5}) = 0.145$$

or the expected relative frequency according to the lognormal distribution is 0.145.

Example 6.5. Assume the data of table 5.1 follow the lognormal distribution. Calculate the magnitude of the 100-year peak flow.

Solution: The 100-year peak flow corresponds to a prob($X > x$) of 0.01. X must be evaluated such that $P_X(x) = 0.99$. This can be accomplished by evaluating Z such that $P_Z(z) = 0.99$ and then transforming to X. From the standard normal tables the value of Z corresponding to $P_Z(z)$ of 0.99 is 2.326. From equation 6.37

$$y = s_y z + \bar{y}$$

The values of s_y and $\bar{y}$ are given in example 6.4.

$$y = .30395(2.326) + 11.0737 = 11.781$$

$$x = \exp(y) = 130{,}700 \text{ cfs}$$

The 100-year peak flow according to the lognormal distribution is about 131,000 cfs.

EXTREME VALUE DISTRIBUTIONS

Many times interest exists in extreme events such as the maximum peak discharge of a stream or minimum daily flows. The extreme value of a set of random variables is also a random variable. The probability distribution of this extreme value random variable will in general depend on the sample size and the parent distribution from which the sample was obtained. Hahn and Shapiro (1967) and Benjamin and Cornell (1970) contain very readable treatments of some of the extreme value distributions.

Consider a random sample of size n consisting of $x_1, x_2, \ldots, x_n$. Let Y be the largest of the sample values. Let $P_Y(y)$ be the prob($Y \leq y$) and $P_{X_i}(x)$ be the prob($X_i \leq x$). Let $p_Y(y)$ and $p_{X_i}(x)$ be the corresponding probability density functions.

$P_Y(y)$ = prob($Y \leq y$) = prob(all of the x's $\leq$ y). If the x's are independently and identically distributed we have

$$P_Y(y) = P_{X_1}(y)\, P_{X_2}(y) \ldots P_{X_n}(y) = [P_X(y)]^n \qquad (6.40)$$

$$p_Y(y) = d\,P_Y(y)/dy = n[P_X(y)]^{n-1} p_X(y) \qquad (6.41)$$

Therefore the probability distribution of the maximum of n independently and identically distributed random variables depends on the sample size n and the parent distribution $P_X(x)$ of the sample. A similar result can be derived for the distribution of the smallest of n independently and identically distributed random variables.

Example 6.6. Assume that the time between rains follows an exponential distribution with a mean of 4 days. Also assume that the time between rains is independent from one rain to the next. Irrigators may be interested in the maximum time between rains. Over a period of 10 rains, what is the probability that the maximum time between rains exceeds 8 days?

Solution: 10 rains means 9 interrain periods or n = 9. From equation 6.40 the probability that the maximum interrain time is less than 8 days is

$$P_Y(8) = [P_X(8)]^9$$

In this example $P_X(y)$ is the cumulative exponential with parameter $\lambda = 1/X = 1/4$.

$$P_Y(8) = [1 - e^{-8/4}]^9 = (1 - e^{-2})^9 = 0.271$$

Therefore the probability that the maximum interrain time will be greater than 8 is 1 - 0.271 = 0.729.

Comment: The probability density function for the maximum interrain time is from equation 6.41

$$p_Y(y) = n(1 - e^{-y/4})^{n-1}\, e^{-y/4}/4$$

This distribution is plotted in figure 6.4 for various values of n. Note that for even moderately large n, the probability is very high that the extreme value (longest interrain time) will be from the tail of the parent (exponential) distribution.

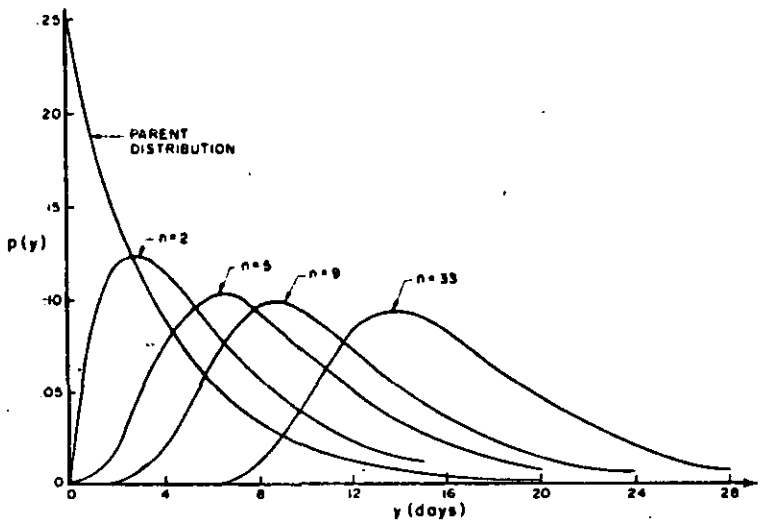


Fig. 6.4. Distribution of the largest sample value from a sample of size n from an exponential distribution.

Frequently the parent distribution from which the extreme is an observation is not known and cannot be determined. If the sample size is large, use can be made of certain general asymptotic results that depend on limited assumptions concerning the parent distribution to find the distribution of extreme values. Much of the work on extreme value distributions is due to Gumbel (1954, 1958).

Three types of asymptotic distributions have been developed based on different (but not all) parent distributions. The types are:

a. Type I - parent distribution unbounded in direction of the desired extreme and all moments of the distribution exist (exponential type distributions).
b. Type II - parent distribution unbounded in direction of the desired extreme and all moments of the distribution do not exist (Cauchy type distributions).
c. Type III - parent distribution bounded in the direction of the desired extreme (limited distributions).

Interest may exist in either the distribution of the largest or smallest extreme values. Examples of parent distributions falling under the various types are:

a. Type I - extreme value largest - normal, lognormal, exponential, gamma
b. Type I - extreme value smallest - normal
c. Type II - extreme value largest or smallest - Cauchy distribution (Hahn and Shapiro 1967; Thomas 1971)
d. Type III - extreme value largest - beta distribution (Hahn and Shapiro 1967; Gibra 1973; Benjamin and Cornell 1970)
e. Type III - extreme value smallest - beta, lognormal, gamma, exponential

The type II or Cauchy type extreme value distributions have found little application in hydrology. The distribution of the largest extreme value in hydrology generally arises as a type I extreme value largest distribution since most hydrologic variables are unbounded on the right. (See Van Montfort (1970) for a test to determine whether a type I or type II extreme value largest best fits the observed data.) The distribution of extreme value smallest commonly found in hydrologic work is the type III extreme value smallest since many hydrologic variables are bounded on the left by zero. The following is a treatment of these two (type I largest and type III smallest) extreme value distributions plus the type I smallest because of its symmetry with the type I largest.

Extreme Value Type I

The type I extreme value has been referred to as Gumbel's extreme value distribution, the extreme value distribution, the Fisher-Tippett type I distribution, and the double exponential distribution. The type I asymptotic distribution for maximum (minimum) values is the limiting model as n approaches infinity for the distribution of the maximum (minimum) of n independent values from an initial distribution whose right (left) tail is unbounded and which is an exponential type; that is, the initial cumulative distribution approaches unity (zero) with increasing (decreasing) values of the random variable at least as fast as the exponential distribution approaches unity. The normal, lognormal, exponential and gamma distributions all meet this requirement for maximum values while the normal distribution satisfies the requirement for minimum values.

The type I extreme value distribution has been used for rainfall depth-duration-frequency studies (Hershfield 1961) and as the distribution of the yearly maximum of daily river flows. Gumbel (1958) states that this latter application assumes (1) the distribution of daily discharges (the parent distribution) is of the exponential type, (2) n=365 is a sufficiently large sample and (3) the daily discharges are independent. Gumbel states that the first and second assumptions cannot be checked since the analytical form of the distribution of discharges is unknown and that the third assumption is clearly not true so that the number of independent observations is something less than 365. In spite of violating the last assumption, experience with the type I for the maximum of daily discharges has been reasonably good. Maximum annual flood peaks would more nearly fulfill assumption three although the sample size would be much less than 365.

The probability density function for the type I extreme value distribution is

$$p_X(x) = \exp\{\mp(x-\beta)/\alpha - \exp[\mp(x-\beta)/\alpha]\}/\alpha \qquad (6.42)$$

$$-\infty < x < \infty; -\infty < \beta < \infty; \alpha > 0$$

where the - applies for maximum values and the + for minimum values. The parameters α and β are scale and location parameters with β being the mode of the distribution. The type I for maximum and minimum values are symmetrical with each other about β. Figure 6.5 is a plot of the distributions for $\alpha = 3{,}897$ and $\beta = 7{,}750$.

The mean and variance of the extreme value type I distribution are

$$E(X) = \beta + 0.577\,\alpha \quad \text{(maximum)}$$
$$= \beta - 0.577\,\alpha \quad \text{(minimum)} \qquad (6.43)$$

$$\text{Var}(X) = 45\,\alpha^2 \quad \text{(both)} \qquad (6.44)$$

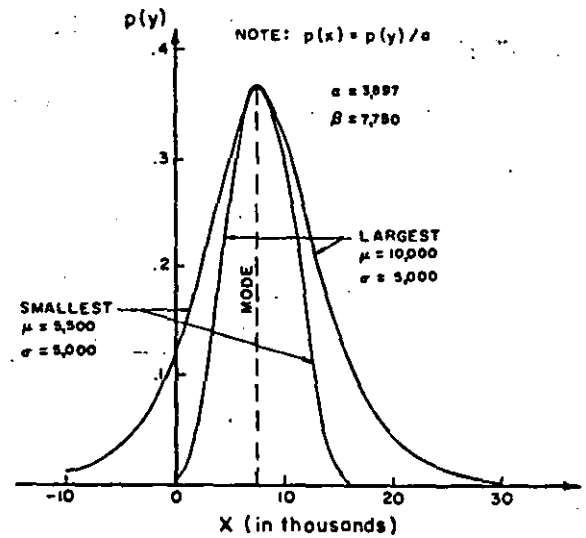


Fig. 6.5. Example of extreme value type I density curves.

The skewness coefficient is

$$\gamma = 1.1396 \quad \text{(maximum)}$$
$$= -1.1396 \quad \text{(minimum)} \qquad (6.45)$$

Thus the type I has a constant coefficient of skewness.

If the transformation

$$Y = (X - \beta)/\alpha \qquad (6.46)$$

is used, the type I extreme value density function becomes

$$p_Y(y) = \exp[\mp y - \exp(\mp y)] \qquad (6.47)$$

where the - applies for the maximum values and the + for the minimum values. The cumulative distribution is

$$P_Y(y) = \int_{-\infty}^{y} \exp[\mp t - \exp(\mp t)]\,dt \qquad -\infty < y < \infty$$

$$= \exp[-\exp(-y)] \quad \text{(maximum)} \qquad (6.48)$$

$$= 1 - \exp[-\exp(y)] \quad \text{(minimum)} \qquad (6.49)$$

The cumulative distribution for maximum values has been tabulated (U.S. National Bureau of Standards 1953). This same table can be used for minimum values by noting that

$$P_{min}(y) = 1 - P_{max}(-y) \tag{6.51}$$

The parameters of the type I extreme value distribution can be estimated in a number of ways. Lowery and Nash (1970) compared several methods for estimating α and β and concluded that the method of moments was as satisfactory as other methods. If the method of moments is used, the estimators are

$$\hat{\alpha} = S/1.283 \tag{6.52}$$

and

$$\hat{\beta} = \bar{X} - 0.45S \quad \text{(maximum)}$$
$$= \bar{X} + 0.45S \quad \text{(minimum)} \tag{6.53}$$

The maximum likelihood estimators (Lowery and Nash 1970) can be determined by a simultaneous solution to the equations

$$\hat{\alpha} = \bar{X} - (\Sigma x_i e^{-x_i/\hat{\alpha}})/\Sigma e^{-x_i/\hat{\alpha}} \tag{6.54}$$

$$\hat{\beta} = -\hat{\alpha} \ln(\Sigma e^{-x_i/\hat{\alpha}}/n) \tag{6.65}$$

Unfortunately these equations cannot be easily solved explicitly for $\hat{\alpha}$ and $\hat{\beta}$ so that a numerical solution is required.

The type I extreme value distribution for maximums has been used to define the "mean annual flood". The probability that an observation from this distribution will exceed the mean of the distribution is $1 - P_Y(y)$ where $P_Y(y)$ is evaluated from equation 6.48 for $y = (\mu - \beta)/\alpha$. Since $\mu = E(X) = \beta + 0.577\alpha$ (equation 6.43), we simply have that $y = 0.577$ and $P_Y(y) = 0.5703$. The probability of a value in excess of the mean is $1 - P_Y(y) = 0.4297$. The return period of a flood equal in magnitude to the mean is $T = [1 - P_Y(y)]^{-1} = 2.33$ years. Thus many times the "mean annual flood" refers to a flood with a return period of 2.33 years.

Extreme Value Type III Minimum (Weibull)

The extreme value type III distribution arises when the extreme is from a parent distribution that is limited in the direction of interest. This distribution has found its greatest use in hydrology as the distribution of low stream flows. Naturally low flows are bounded by zero on the left. The type III for minimum values is also known as the Weibull distribution and is defined as

$$p_X(x) = \alpha x^{\alpha-1} \beta^{-\alpha} \exp[-(x/\beta)^{\alpha}] \qquad x \geq 0;\ \alpha, \beta > 0 \tag{6.56}$$

The cumulative Weibull is given by

$$P_X(x) = 1 - \exp[-(x/\beta)^{\alpha}] \tag{6.57}$$

The mean and variance of the distribution are

$$E(X) = \beta\Gamma(1 + 1/\alpha) \tag{6.58}$$

$$Var(X) = \beta^2[\Gamma(1 + 2/\alpha) - \Gamma^2(1 + 1/\alpha)] \tag{6.59}$$

Hahn and Shapiro (1967) give the coefficient of skew as

$$\gamma = \frac{\Gamma(1 + 3/\alpha) - 3\Gamma(1 + 2/\alpha)\,\Gamma(1 + 1/\alpha) + 2\Gamma^3(1 + 1/\alpha)}{[\Gamma(1 + 2/\alpha) - \Gamma^2(1 + 1/\alpha)]^{3/2}} \tag{6.60}$$

The parameters of the Weibull distribution can be estimated by the method of moments by substituting the sample mean and variance for the population mean and variance respectively in equations 6.58 and 6.59 and then solving the two equations simultaneously for $\hat{\alpha}$ and $\hat{\beta}$.

The maximum likelihood estimates can be determined by letting

$$\lambda = \beta^{-\alpha}$$

and then solving the equations

$$\hat{\lambda} = n/\Sigma_{i=1}^{n} x_i^{\hat{\alpha}} \tag{6.61}$$

and

$$\hat{\alpha} = n/(\hat{\lambda} \Sigma_{i=1}^{n} x_i^{\hat{\alpha}} \ln x_i - \Sigma_{i=1}^{n} \ln x_i) \tag{6.62}$$

simultaneously for $\hat{\alpha}$ and $\hat{\lambda}$. $\hat{\beta}$ is then given by

$$\hat{\beta} = (\hat{\lambda})^{-1/\hat{\alpha}} \tag{6.63}$$

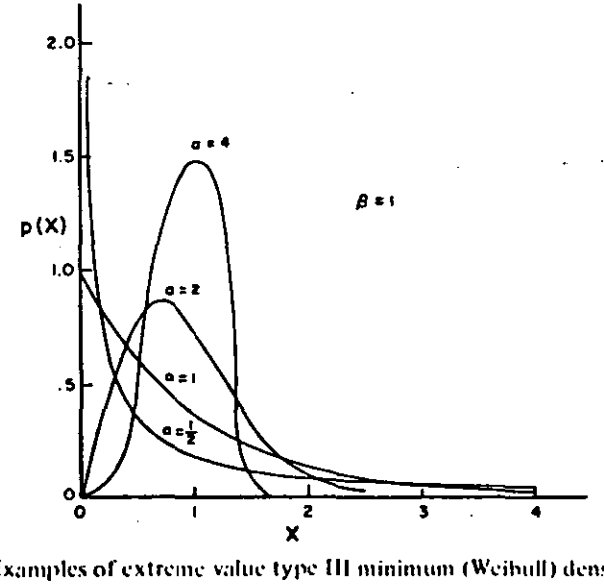


Fig. 6.6. Examples of extreme value type III minimum (Weibull) density curves.

Either method of parameter estimation is difficult. Exercise 6.18 provides a method for simplifying the solution by the moment equations.

The Weibull probability density function can range from a reverse-J with $\alpha < 1$, to an exponential with $\alpha = 1$ and to a nearly symmetrical distribution (figure 6.6) as α increases.

If the lower bound on the parent distribution is not zero, a displacement parameter must be added to the type III extreme value distribution for minimums so that the density function becomes

$$p_X(x) = \alpha(x - \varepsilon)^{\alpha-1}(\beta - \varepsilon)^{-\alpha} \exp[-\{(x - \varepsilon)/(\beta - \varepsilon)\}^{\alpha}] \tag{6.64}$$

and the cumulative distribution function becomes

$$P_X(x) = 1 - \exp[-\{(x - \varepsilon)/(\beta - \varepsilon)\}^{\alpha}] \tag{6.65}$$

By using the transformation

$$y = \{(x - \varepsilon)/(\beta - \varepsilon)\}^{\alpha}$$

Table 6.2. Table for solution of equations 6.60, 6.70 and 6.71.

α	$1/\alpha$	$A(\alpha)$	$B(\alpha)$
-1.000	0.02	0.446	40.005
-0.971	0.03	0.444	26.987
-0.917	0.04	0.442	20.481
-0.867	0.05	0.439	16.576
-0.638	0.10	0.425	8.737
-0.254	0.20	0.389	4.755
0.069	0.30	0.346	3.370
0.359	0.40	0.297	2.634
0.631	0.50	0.246	2.159
0.896	0.60	0.193	1.815
1.160	0.70	0.142	1.549
1.430	0.80	0.092	1.334
1.708	0.90	0.044	1.154
2.000	1.00	0.000	1.000
2.309	1.10	-0.040	0.867
2.640	1.20	-0.077	0.752
2.996	1.30	-0.109	0.652
3.382	1.40	-0.136	0.563
3.802	1.50	-0.160	0.486
4.262	1.60	-0.180	0.418
4.767	1.70	-0.196	0.359
5.323	1.80	-0.208	0.308
5.938	1.90	-0.217	0.263
6.619	2.00	-0.224	0.224
7.374	2.10	-0.227	0.190
8	2.20	-0.229	0.161

tables of e^{-y} can be used to determine $P_X(x)$. Equation 6.64 is sometimes known as the 3-parameter Weibull distribution or as the bounded exponential distribution.

The mean and variance of the three parameter Weibull distribution are

$$E(X) = \varepsilon + (\beta - \varepsilon)\,\Gamma(1 + 1/\alpha) \tag{6.66}$$

and

$$\mathrm{Var}(X) = (\beta - \varepsilon)^2\,[\Gamma(1 + 2/\alpha) - \Gamma^2(1 + 1/\alpha)] \tag{6.67}$$

The coefficient of skew is again given by equation 6.60. Through algebraic manipulation, equations 6.66 and 6.67 can be put in the form (Gumbel 1958)

$$\beta = \mu + \sigma A(\alpha) \tag{6.68}$$

$$\varepsilon = \beta - \sigma B(\alpha) \tag{6.69}$$

where

$$A(\alpha) = [1 - \Gamma(1 + 1/\alpha)]\,B(\alpha) \tag{6.70}$$

$$B(\alpha) = [\Gamma(1 + 2/\alpha) - \Gamma^2(1 + 1/\alpha)]^{-1/2} \tag{6.71}$$

The moment estimates for α, β, and ε can now be obtained by (1) solving equation 6.60 for $\hat{\alpha}$, (2) solving 6.70 and 6.71 for $A(\alpha)$ and $B(\alpha)$, (3) solving 6.68 for $\hat{\beta}$ and (4) solving 6.69 for $\hat{\varepsilon}$. Table 6.2 can be used to simplify the calculations.

Example 6.7. The minimum annual daily discharges on a stream are found to have an average of 125 cfs, a standard deviation of 50 cfs and a coefficient of skew of 1.4. Using both the type III minimum and the type I minimum extreme value distributions, evaluate the probability of an annual minimum flow being less than 100 cfs.

Solution: Type III minimum using interpolation in table 6.2.

$1/\alpha = 0.79$

$A(\alpha) = 0.098$

$B(\alpha) = 1.36$

$\hat{\alpha} = 1/0.79 = 1.266$

$\hat{\beta} = 125 + 50(0.098) = 129.9$ (eq. 6.68)

$\hat{\varepsilon} = 129.9 - 50(1.36) = 61.9$ (eq. 6.69)

$\mathrm{prob}(X \le 100) = P_X(100) = 1 - e^{-y}$ (eq. 6.65) where

$y = \{(x - \hat{\varepsilon})/(\hat{\beta} - \hat{\varepsilon})\}^{\hat{\alpha}} = \{(100 - 61.9)/(129.9 - 61.9)\}^{1.266} =$ [illegible]

$P_X(100) = 0.3$

Type I minimum

$\hat{\alpha} = s/1.283 = 50/1.283 = 38.97$ (eq.6.52)

$\hat{\beta} = \bar{x} + 0.45s$ (eq. 6.53)

$= 125 + .45(50) = 147.5$

$\text{prob}(X \leq 100) = P_X(100) = 1 - \exp(-e^Y)$ (eq. 6.49)

where $Y = (X - \hat{\beta})/\hat{\alpha} = (100 - 147.5)/38.97 = 1.219$

$P_X(100) = 1 - 0.745 = 0.255$

Comment: The results of applying these two distributions to this problem are very different. This should be expected since it is a situation where the type I for minimums would not be expected to apply since there would be a lower bound and since the coefficient of skew was given as 1.4 while the coefficient of skew for the type I minimum is -1.1396.

Discussion

The theory on which the extreme value distributions depend is not as strong as the Central Limit Theorem for the normal distribution. More assumptions concerning the underlying or parent distribution must be made and the rate of convergence to an asymptotic extreme value distribution may be rather slow. However, the extreme value distributions do provide a connection between observed extreme events and models which may be used to evaluate the probabilities of future extreme events.

The conditions under which the various extreme value distributions arise are such that for many parent distributions (lognormal, gamma) the distribution of maximum values and the distribution of minimum values are not of the same type. The minimum values from a lognormal would be expected to follow the type III distribution while the maximum values would follow a type I distribution.

Various types of extreme value distributions are related. The logarithms of a random variable that follows a type III minimum are distributed as the type I minimum extreme value distribution. Chow (1954) has shown that if the coefficient of variation of the type I maximum extreme value distribution is 0.364, the distribution is practically the same as the lognormal distribution with the same coefficient of variation and coefficient of skew (1.139).

BETA DISTRIBUTION

A distribution that has both an upper and lower bound is the beta distribution. Generally the beta distribution is defined over the interval 0 to 1. It can, however, be transformed to any interval a to b. If the limits of the distribution are unknown, they become parameters of the distribution making it a four parameter rather than a two parameter distribution. The beta density function is given by

$$p_X(x) = x^{\alpha-1}(1-x)^{\beta-1}/B(\alpha,\beta) \quad 0<x<1;\ \alpha,\beta>0 \tag{6.72}$$

The function $B(\alpha,\beta) = \int_0^1 x^{\alpha-1}(1-x)^{\beta-1}\,dx$ is called the beta function. The beta function is related to the gamma function by

$$B(\alpha,\beta) = \Gamma(\alpha)\Gamma(\beta)/\Gamma(\alpha+\beta) \tag{6.73}$$

The beta function is tabulated.

The mean and variance of the beta distribution are

$$E(X) = \alpha/(\alpha+\beta) \tag{6.74}$$

$$\text{Var}(X) = \alpha\beta/[(\alpha+\beta+1)(\alpha+\beta)^2] \tag{6.75}$$

The mean and variance can be used to get the moment estimators for α and β.

PEARSON DISTRIBUTIONS

Karl Pearson (Elderton 1953) has proposed that frequency distributions can be represented by

$$p_X(x) = \exp \int_{-\infty}^{x} (t+\alpha)/(\beta_0 + \beta_1 t + \beta_2 t^2 + \ldots)\,dt \tag{6.76}$$

By choosing appropriate values for the parameters, equation 6.76 becomes a large number of families of distributions including the normal, beta and gamma distributions.

The Pearson type III has found application in hydrology especially as the distribution of flood peaks. This distribution can be written

$$p_X(x) = p_0(1 + x/\alpha)^{\alpha/\delta}\, e^{-x/\delta} \tag{6.77}$$

with the mode at $X = 0$. The lower bound of the distribution is $X = -\alpha$. The difference in the mean and mode is δ and the value of $p_X(x)$ at the mode is p_0. It can be shown that the Pearson type III is the same as the three parameter gamma distribution. By shifting equation 6.77 so that the mode is at $X = \alpha$ and the lower bound is at $X = 0$, we have

$$p_X(x) = p_0\, e^{-(x-\alpha)/\delta}\,(x/\alpha)^{\alpha/\delta} \tag{6.78}$$

The gamma distribution has the mode at $(\eta-1)/\lambda$ and the mean at η/λ. Thus $\alpha = (\eta-1)/\lambda$ and $\delta = \eta/\lambda - (\eta-1)/\lambda = 1/\lambda$. The value of $p_X(x)$ at the mode for the gamma distribution is

$$p_X(x_{mo}) = p_X((\eta-1)/\lambda) = \lambda^\eta[(\eta-1)/\lambda]^{\eta-1}\exp[-\lambda(\eta-1)/\lambda]/\Gamma(\eta)$$

$$p_0 = \lambda(\eta-1)^{\eta-1}\exp[-(\eta-1)]/\Gamma(\eta)$$

Substituting these quantities into 6.78 results in

$$p_X(x) = \lambda^\eta x^{\eta-1} e^{-\lambda x}/\Gamma(\eta)$$

which is the gamma distribution (eq. 6.11).

SOME IMPORTANT DISTRIBUTIONS OF SAMPLE STATISTICS

We have already seen that sample statistics as functions of random variables are themselves random variables. Statistical tests depend on the probability distribution of test statistics which are merely sample statistics. In this section three of the important distributions of sample statistics are briefly discussed.

Chi-square Distribution

If Z is a standardized, normally distributed random variable, $Z = (X - \mu)/\sigma$, then

$$Y = \sum_{i=1}^{n} Z_i^2 \tag{6.79}$$

where Y is the sum of squares of n random values of Z and has a chi-square distribution with n degrees of freedom. The chi-square distribution is a special case of the gamma distribution when $\lambda = \frac{1}{2}$ and η is a multiple of $\frac{1}{2}$. The distribution thus has a single parameter $\nu = 2\eta$ known as the degrees of freedom. The expression for the distribution is

$$p_{\chi^2}(x) = x^{(\nu-2)/2}\, e^{-x/2} / [2^{\nu/2}\, \Gamma(\nu/2)] \qquad x, \nu > 0 \tag{6.80}$$

The mean and variance of the distribution are

$$E(\chi^2) = \nu \tag{6.81}$$

$$\mathrm{Var}(\chi^2) = 2\nu \tag{6.82}$$

The parameter ν is usually known in any application of the chi-square to statistical testing. Equation 6.81 produces the moment estimator for ν as $\hat{\nu} = \bar{X}$. In figure 6.2 the curve labeled $\lambda = \frac{1}{2}$ is a chi-square distribution with $\nu = 6$. The coefficient of skew for the chi-square distribution is $2/\sqrt{\nu/2}$. The cumulative chi-square distribution is contained in table E.6 in the form

$$\alpha = P_{\chi^2}(\chi_\alpha^2) = \int_0^{\chi_\alpha^2} p_{\chi^2}(x)\, dx \tag{6.83}$$

for various values of ν and α.

In table E.6 the notation $\chi^2_{\alpha,\nu}$ refers to the value of χ^2 such that 100α percent of the distribution lies to the left of $\chi^2_{\alpha,\nu}$. Thus for $\nu = 10$, 95% of the χ^2 distribution lies to the left of 18.3 or the probability that a random observation drawn from a χ^2 distribution with $\nu = 10$ lies between 0 and 18.3 is 0.95.

The fact that the sum of squares of n random standard normal variates is a chi-square distribution with $\nu = n$ makes it apparent that if X_i is a chi-square random variable with parameter ν_i, then $X = \sum X_i$ is a chi-square random variable with parameter $\nu = \sum \nu_i$ if all the X_i are independent.

If $Z_1, Z_2, \ldots, Z_n$ is a random sample from a standard normal distribution, then $Y = \sum Z_i^2 = \sum (X_i - \bar{X})^2/\sigma^2$ has a chi-square distribution with $\nu = n - 1$. Furthermore since $s^2 = \sum (X_i - \bar{X})^2/(n-1)$, the quantity $(n-1)s^2/\sigma^2$ has a chi-square distribution with $\nu = n - 1$ (Mood et al. 1974).

The t Distribution

If Y is a standardized normal variate and U is a chi-square variate with ν degrees of freedom and Y and U are independent, then

$$X = Y\sqrt{\nu} \,/\, \sqrt{U} \tag{6.84}$$

has a t distribution with ν degrees of freedom. The t distribution is given by

$$p_T(t) = \Gamma[(\nu+1)/2]\, (1 + t^2/\nu)^{-(\nu+1)/2} / [\sqrt{\pi\nu}\, \Gamma(\nu/2)] \qquad -\infty < t < \infty;\ \nu > 0 \tag{6.85}$$

The mean and variance of the t distribution are

$$E(T) = 0 \tag{6.86}$$

$$\mathrm{Var}(T) = \nu/(\nu - 2) \qquad \text{for } \nu > 2 \tag{6.87}$$

The cumulative t distribution is contained in table E.5 in the form

$$\alpha = P_T(t) = \int_{-\infty}^{t} p_T(x)\, dx \tag{6.88}$$

for various values of ν and α.

The interpretation of table E.5 for the t distribution is similar to that of table E.6 for the χ^2 distribution. If one wants the value of T such that 100α % of the t distribution lies to the left of t, the value of $t_{\alpha,\nu}$ would be read from the row containing the correct value of ν. If one wanted the value of T such that $\mathrm{prob}(|T| > t) = \alpha$, the value of $t_{1-\alpha/2,\nu}$ would be read from the row containing the correct value of ν.

One use of the t distribution is as the sampling distribution of the mean from a normal distribution with unknown variance. If we write

$$T = (\bar{X} - \mu)/\sqrt{S^2/n}$$

and then divide the numerator and denominator by σ, we get

$$T = [(\bar{X} - \mu)/(\sigma/\sqrt{n})]/\sqrt{S^2/\sigma^2} = Y/\sqrt{U}$$

where $Y = (\bar{X} - \mu)/(\sigma/\sqrt{n})$ has a standard normal distribution and $U = (n-1)S^2/\sigma^2$ has a χ^2 distribution. Thus from equation 6.84, T has a t distribution with n - 1 degrees of freedom.

As ν of the t distribution gets large, the t distribution approaches the standard normal distribution. This can be seen by comparing the row labeled $\nu = \infty$ of table E.5 with table E.4. Thus for large samples, the sampling distribution of the mean of a normal distribution with unknown variance approaches a normal distribution. We have already seen that the distribution of the sample mean from a normal distribution with a known variance is exactly a normal distribution. One can reason that as the sample size increases, the estimate for the variance improves to the point where the sampling distribution of the mean of a normal distribution with unknown variance can be approximated by the sampling distribution of the mean of a normal distribution with a known variance which is itself a normal distribution. In practice one rarely knows the variance of the distribution from which a sample is obtained.

Example 6.8. A sample of size 8 from a normal distribution results in $\bar{x} = 12.7$ and $s^2 =$

9.8. What is the probability that $\bar{x}$ is in error by more than 1.0?

Solution: $(\bar{X} - \mu)/\sqrt{S^2/n}$ has a t distribution with n - 1 degrees of freedom. To be in error by more than 1.0 units we must have $\bar{X} - \mu > 1.0$

$$\text{prob}(|\bar{X} - \mu| > 1.0) = \text{prob}(|T| > (\bar{x} - \mu)/\sqrt{s^2/n})$$

$$= \text{prob}(|T| > 1.0/\sqrt{9.8/8}) = \text{prob}(|T| > .904)$$

The desired probability is the shaded area of figure 6.7. The total area to the left of T = .904 is .802 as estimated by linear interpolation from table E.5. Therefore the area to the right of T = .904 is 1 - 0.802 = 0.198. By symmetry the area to the left of -0.904 is 0.198. The desired probability is 0.198 + 0.198 = 0.396.

If a standard normal distribution had been used rather than a t distribution, it would have been necessary to find prob(|Z| > 0.904). This probability can be found from table E.4 to be 0.366. Thus even for a sample as small as 8, the normal is a reasonable approximation.

The F Distribution

If U is a chi-square variate with γ = m degrees of freedom and V is a chi-square variate with γ = n degrees of freedom and U and V are independent, then

$$X = (U/m)/(V/n) \tag{6.89}$$

has an F distribution with γ_1 = m and γ_2 = n degrees of freedom (m and n are known as the numerator and denominator degrees of freedom respectively). The F distribution is given by

$$p_f(f) = \Gamma[(\gamma_1+\gamma_2)/2]\ \gamma_1^{\gamma_1/2}\gamma_2^{\gamma_2/2} f^{(\gamma_1-2)/2}(\gamma_2+\gamma_1 f)^{-(\gamma_1+\gamma_2)/2}/[\Gamma(\gamma_1/2)\Gamma(\gamma_2/2)] \tag{6.90}$$

$\gamma_1, \gamma_2, f > 0$

The mean and variance of the F distribution are

$$E(F) = \gamma_1/(\gamma_2 - 2) \tag{6.91}$$

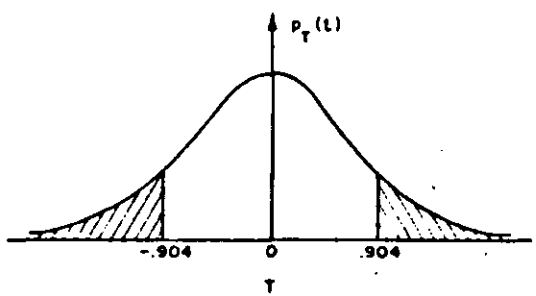


Fig. 6.7. Sketch for example 6.8.

$$\text{Var}(F) = \gamma_2^2(\gamma_1 + 2)/[\gamma_1(\gamma_2 - 2)(\gamma_2 - 4)] \tag{6.92}$$

The cumulative F distribution is contained in table E.7 as a function of m and n for values of $P_F(f) = 0.90, 0.95, 0.975, 0.99$ and 0.995.

Table E.7 contains values of $F_{\alpha,m,n}$ such that 100α % of the distribution with m and n degrees of freedom lies to the left of $F_{\alpha,m,n}$. For example, the probability that a random observation from an F distribution with 5 numerator and 10 denominator degrees of freedom exceeds 3.33 is 1.00 - 0.95 = 0.05.

TRANSFORMATIONS

Many times a transformation can be made in an attempt to arrive at a probability distribution that will describe the data. Common transformations are logarithmic transformations, translations along the X axis, and n^{th} power transformations for n = 1/3, 1/2, 2 and 3.

We have already made one application of the logarithmic transformation to get the lognormal distribution from the normal distribution. Other distributions can be transformed by means of this transformation as well. Benson (1968) discusses the use of the log-Pearson type III distribution for flood frequencies.

Translations are especially useful in the case of bounded distributions. We made use of a translation in deriving the three parameter extreme value type III for minimums from the corresponding two parameter distribution. In general a translation is accomplished by subtracting a location parameter, ε, from the random variable. For example

$$p_X(x) = \lambda e^{-\lambda(x-\varepsilon)} \qquad \lambda > 0;\ x > \varepsilon \tag{6.93}$$

could be considered a two parameter exponential distribution with the lower bound at $X = \varepsilon$. Exercises 3.17 and 3.18 deal with estimating the two parameters of this distribution. In general the addition of a displacement parameter, if the displacement parameter is unknown, makes parameter estimation via maximum likelihood much more difficult.

Moment estimators are relatively simple in that the addition of a displacement parameter affects the mean by $\mu = \mu_X + \varepsilon$ and has no effect on the variance or skewness. Thus a three parameter gamma distribution might be given by

$$p_X(x) = \lambda^\eta(x - \varepsilon)^{\eta-1}e^{-\lambda(x-\varepsilon)}/\Gamma(\eta) \qquad x > \varepsilon \tag{6.94}$$

with the moment estimators for λ, η and ε determined from

$$E(X) = \mu = \varepsilon + \eta/\lambda$$

$$\text{Var}(X) = \sigma^2 = \eta/\lambda^2$$

$$\gamma = 2/\sqrt{\eta}$$

The fact that γ must be now used to estimate η means that for small samples accuracy is lost since γ is based on the third sample moment. As shown earlier, the three parameter gamma is the same as the Pearson type III distribution.

Sangal and Biswas (1970) have used the three parameter lognormal distribution ob-

tained by fitting a normal distribution to the logarithms of $(X - \epsilon)$ where ϵ is a parameter that must be estimated from the data. They found for 10 Canadian rivers that the three parameter lognormal distribution fit the observed distribution of peak flows. They also state that the Gumbel extreme value distribution is a special case of the three parameter lognormal distribution.

The three parameter lognormal is given by

$$p_X(x) = ((x - \epsilon)^2\ 2\pi\sigma_y^2)^{-1/2} \exp\{-\tfrac{1}{2}[\ln(x - \epsilon) - \mu_y]^2/\sigma_y^2\} \quad x > \epsilon \tag{6.95}$$

where

$$E(X) = \epsilon + \exp(\mu_y + \sigma_y^2/2) \tag{6.96}$$

$$\mathrm{Var}(X) = \exp(2\mu_y + \sigma_y^2)\exp(\sigma_y^2) - 1 \tag{6.97}$$

$$\gamma = \exp[(\sigma_y^2) - 1]^{3/2} + 3[\exp(\sigma_y^2) - 1]^{1/2} \tag{6.98}$$

Stidd (1953) and Kendall (1967) discuss transforming variables by $Y = X^{1/3}$ and then fitting a normal distribution to Y. They discuss this transformation in terms of precipitation probabilities.

MORE ON MOMENT GENERATING FUNCTIONS

In chapter 3 the moment generating function of the random variable X, $M_X(t)$, was defined as $E(e^{tX})$. In this section we will give some properties of moment generating functions and show how they can be used for purposes other than generating moments.

One important property of moment generating functions is the uniqueness property. That is, if X_1 and X_2 are two random variables with moment generating functions $M_{X_1}(t)$ and $M_{X_2}(t)$, respectively, and if $M_{X_1}(t) = M_{X_2}(t)$, then the probability distributions of X_1 and X_2 are identical. This property means that there is a one-to-one correspondence between a density function and its moment generating function. Appendix B contains a listing of the moment generating functions for most of the distributions discussed in this book.

A second important property of moment generating functions is that if $X_1, X_2, \ldots X_n$ are independent random variables and $Y = a_1X_1 + a_2X_2 + \ldots + a_nX_n$, then the moment generating function of Y is given by

$$M_Y(t) = M_{X_1}(a_1 t)\, M_{X_2}(a_2 t) \ldots M_{X_n}(a_n t) \tag{6.99}$$

This property is very useful in deriving the distribution of the means of a random sample or the distribution of the sum of n independent random variables.

A third property of moment generating function is that if $Y = a + bX$ then

$$M_Y(t) = e^{at} M_X(bt) \tag{6.100}$$

This property is useful in determining the distribution of a linear function of a random variable.

Example 6.9. Show that the sum of n independent exponential random variables with parameter λ is a gamma random variable.

Solution: From Appendix B we have the moment generating function of the exponential is $(1 - t/\lambda)^{-1}$. From 6.99 we have

$$M_Y(t) = \prod_{i=1}^{n} M_{X_i}(t)$$

$$= \prod_{i=1}^{n} (1 - t/\lambda)^{-1} = (1 - t/\lambda)^{-n}$$

From Appendix B this can be seen to be the moment generating function of the gamma distribution with parameters n and λ. From the uniqueness property of moment generating function, Y must be a gamma random variable with parameters n and λ.

Example 6.10. If X is a normal random variable with mean μ and variance σ^2, show that $Y = a + bX$ is also normal. What is the mean and variance of Y?

Solution: From Appendix B

$$M_X(t) = e^{\mu t + \sigma^2 t^2/2}$$

From equation 6.100

$$M_Y(t) = e^{at}\, M_X(bt) = e^{at}\, e^{\mu bt + \sigma^2 b^2 t^2/2}$$

$$= e^{(a + b\mu)t + b^2\sigma^2 t^2/2}$$

which can be seen to be the moment generating function for a normal distribution with mean $a + b\mu$ and variance $b^2\sigma^2$.

Exercises

6.1 Show that the mean of the uniform distribution is $(\beta + \alpha)/2$ and the variance is $(\beta - \alpha)^2/12$.

6.2 What is the skewness and kurtosis of the uniform distribution?

6.3 What is the skewness and coefficient of variation for the exponential distribution?

6.4 Fit the gamma distribution to the data of exercise 2.2. Plot the expected relative frequency according to the gamma distribution on the plot of exercise 2.2.

6.5 Repeat exercise 6.4 using the lognormal distribution.

6.6 Fit the lognormal distribution to the Kentucky River data of table 2.1. Is this a good approximation for the data?

6.7 Work exercise 5.8 using the lognormal distribution.

6.8 A set of data having a mean of 4.5 and a standard deviation of 2.0 is thought to fit

low the type I extreme value distribution for maximums. What proportion of the observations from this distribution exceed 6.0? Plot the probability density function.

6.9 Repeat exercise 6.8 using the type I extreme value distribution for minimums.

6.10 Repeat exercise 6.8 using the Weibull distribution.

6.11 Repeat exercise 6.8 using the lognormal distribution.

6.12 Show that the exponential distribution is memoryless (i.e. show that prob($X \geq t+\tau | X > t$) = prob($X > t$)).

6.13 Plot the probability density function and the cumulative probability distribution for the lognormal distribution with $\mu_x = 50{,}000$ and $\sigma_x = 25{,}000$.

6.14 Plot the theoretical distribution of the largest value selected from a normal distribution with $\mu = 2$ and $\sigma = 4$ for sample sizes of n = 2, 5, 9 and 33. Compare the results with those of example 6.5.

6.15 Derive expressions analogous to equations 6.40 and 6.41 for the smallest of n independently and identically distributed random variables.

6.16 Verify equation 6.47 from equations 6.42 and 6.46.

6.17 Assume that during month 1 the mean and standard deviation of the monthly rainfall are 0.750 and 0.433 inches respectively. Similarly during month 2 the mean and standard deviations of monthly rainfall are 3.000 and 0.866 inches respectively. Assume monthly rainfall amounts can be approximated by the gamma distribution and that rainfall in month 2 is independent of rainfall in month 1. What is the probability of receiving more than 3 inches of rain during the two month period?

6.18 Show that for the two parameter Weibull distribution the parameter α is a function only of the coefficient of variation. Using this fact, describe a procedure for estimating α and β of the distribution.

6.19 If peak discharge, q, is lognormally distributed with mean μ_q and variance σ_q^2, what is the probability distribution of stage, S? Assume stage and discharge are related by $q = aS^b$.

6.20 Work exercise 6.19 assuming the peak discharges are distributed as the type I extreme value distribution.

6.21 In example 6.3 let $\hat{\eta}$ be approximated by 6.0. Calculate $\hat{\lambda}$ from equation 6.20 and then evaluate the prob(yield > 20.0) by using the equation following equation 6.16. Compare the results with those of example 6.3.

6.22 Use the method of moments to estimate the parameters of the three parameter lognormal distribution for the North Llano River near Junction, Texas. What is the return period of a mean annual flow of 273 cfs or more?

6.23 Using moment generating functions show that if X is $N(\mu,\sigma^2)$ then $\bar{X}$ is $N(\mu,\sigma^2/n)$ where n is the sample size.

6.24 Show that if X is exponentially distributed with parameter λ then $\bar{X}$ is distributed gamma with parameters $n\lambda$ and n where n is the sample size.

6.25 Show that if X_i is binomially distributed with parameters n_i and p then $Y = \Sigma X_i$ is binomial with parameters Σn_i and p.

6.26 Show that if X_i is geometrically distributed with parameter p then $Y = \Sigma_{i=1}^{k} X_i$ is negative binomially distributed with parameters p and k.

2.27 Show that if X_i is distributed Poisson with parameter λ_i then $Y = \Sigma X_i$ is Poisson with parameter $\Sigma \lambda_i$.

6.28 Show that if X_i is distributed chi-square with parameter γ_i then $Y = \Sigma X_i$ is chi-square with parameter $\Sigma \gamma_i$.

6.29 Calculate the return period associated with an annual runoff of 0.500 inches for Walnut Gulch near Tombstone, Arizona. (Data in Appendix C) Assume (a) lognormal distribution (b) gamma distribution (c) extreme value type I (d) normal distribution.

6.30 Assume the data of exercise 4.10 are distributed as a two-parameter exponential distribution. Estimate the parameters of this distribution and prepare a table comparing the observed and expected number of floods over the 100-year period.

7. Probability Plotting and Frequency Analysis

ONE OF the earliest and most frequent uses of statistics in hydrology has been that of frequency analysis. Early applications of frequency analysis were largely in the area of flood flow estimation. Today nearly every phase of hydrology is subjected to frequency analyses. Over the years and continuing today there have been volumes of material written on the best probability distribution to use in various situations. One cannot, in most instances, analytically determine which probability distribution should be used. Certain limit theorems such as the Central Limit Theorem and Extreme Value Theorems might provide guidance. One should also evaluate the experience that has been accumulated with the various distributions and how well they describe the phenomena of interest. Certain properties of the distributions can be used in screening distributions for possible application in a particular situation. For example the range of the distribution, the general shape of the distribution and the skewness of the distribution many times indicate that a particular distribution may or may not be applicable in a given situation. When two or more distributions appear to describe a given set of data equally well, the distribution that has been traditionally used should be selected unless there are contrary overriding reasons for selecting another distribution. However, if a traditionally used distribution is inferior, its use should not be continued just because "that's the way it's always been done".

The first part of this chapter discusses plotting data in the form of a cumulative probability distribution. Various types of graph paper known as probability paper have been developed for this purpose and are applicable to hydrologic data.

The second topic covered is analytical frequency analysis at a point. In this discussion a simplified technique based on frequency factors is presented for determining the magnitude of an event with a given return period. In general the frequency factor is a function of the distributional assumption that is made and of the mean, variance and for some distributions the coefficient of skew of the data.

Regional frequency analysis is then discussed. Regional frequency analysis attempts to use data from several locations in a "homogeneous" region to determine the frequency relationship for a point. The chapter closes with a discussion of the frequency

analysis of precipitation data.

GRAPHICAL CONSTRUCTION OF PROBABILITY PAPER

In chapter 2 plotting probability density functions and cumulative probability distributions on arithmetic paper was discussed. In general when the cumulative distribution function $P_X(x)$ is plotted on arithmetic paper versus the value of X, a straight line does not result. To get a straight line on arithmetic paper, $P_X(x)$ would have to be given by the expression $P_X(x) = ax + b$ or $p_X(x) = a$ which can be shown to be the uniform distribution given by equations 6.1 and 6.2. Thus if the cumulative distribution of a set of data plots as a straight line on arithmetic paper, the data follows a uniform distribution. The first part of this chapter is devoted to a discussion of special paper called probability paper. Probability paper can be developed so that any cumulative distribution can be plotted as a straight line. Generally a separate type of probability paper is required for each of the different probability distributions to plot as a straight line. The scaling of the probability paper may even have to change as the parameters of a particular distribution change.

Constructing probability paper is a process of transforming the probability scale so that the resulting cumulative curve is a straight line. The transformation technique will be illustrated with the normal distribution. The coordinates shown in table 7.1 were obtained from a table of the standardized normal distribution. If these data are plotted on arithmetic paper, the curved line of figure 7.1 results. This curved line can be linearized graphically by drawing a straight line between two points on the curve line covering the desired range. The probability scale is then derived as indicated by the dashed lines. As an example consider the 0.3 point on the arithmetic scale. The dashed line is followed vertically upward to the curved line, horizontally toward the straight line and then vertically upward to the probability scale. For $P_Z(z)$ greater than 0.5 the same procedure is followed. As can be seen from figure 7.1 the probability scale is compressed near the center ($P_Z(z) = 0.5$) and expanded in the two tails.

If all of the labeling in figure 7.1 is removed except the probability scale, the result is normal probability paper. Figure 7.2 shows a normal probability paper with the probability scale expressed as a percent. This figure has two scales so that one can represent the percent greater than and one the percent less than a certain value. Any normally distributed data will plot as a straight line on normal probability paper with the mean corresponding to the 50 percent point and one standard deviation from the mean corre-

Table 7.1. Coordinates for cumulative standardized normal distribution.

Z	P(z)	Z	P(z)
-3.0	.0013	0.5	.6915
-2.5	.0062	1.0	.8413
-2.0	.0227	1.5	.9332
-1.5	.0668	2.0	.9772
-1.0	1587	2.5	.9938
-0.5	.3085	3.0	.9987
0	.500		

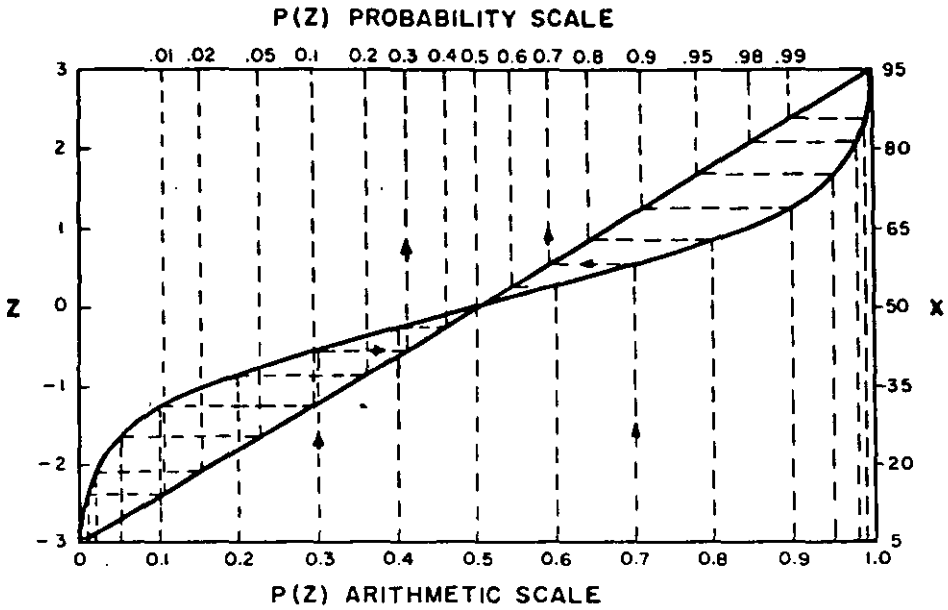


Fig. 7.1. Graphical construction of probability paper.

sponding to the 15.87 and 84.13 percent points. The right scale in figure 7.1 shows how a set of data with $\mu = 50$ and $\sigma = 15$ would appear. The standard normal transformation $Z = (X - \mu)/\sigma$ is used to transform the Z scale to the X scale.

MATHEMATICAL CONSTRUCTION OF PROBABILITY PAPER

For many probability distributions, probability paper can be constructed analytically so that the cumulative distribution function plots as a straight line. This can be done by transforming the cumulative probability function to the form

$$Y = a Z + b \tag{7.1}$$

where Y is a function of the parameters and $P_X(x)$, Z is a function of parameters and x, and a and b are functions of parameters.

The procedure will be illustrated with the exponential distribution

$$P_X(x) = 1 - e^{-\lambda x} \tag{7.2}$$

which can be written

$$-\ln(1 - P_X(x)) = \lambda x \tag{7.3}$$

Comparing equations 7.1 and 7.3 it can be seen that

$$Y = -\ln(1 - P_X(x)) \tag{7.4}$$

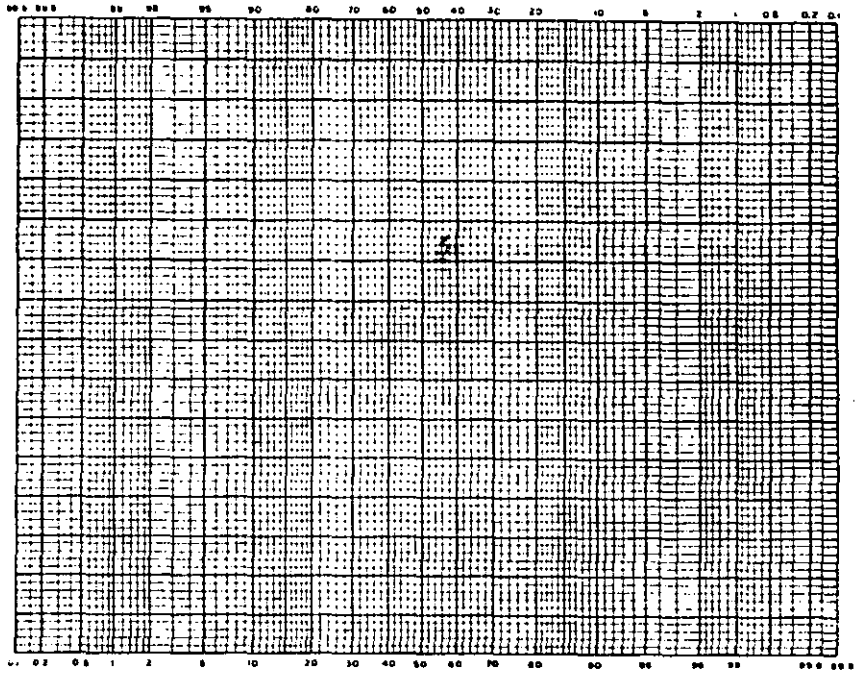


Fig. 7.2. Normal probability paper.

$$Z = x \tag{7.5}$$

$$a = \lambda$$

$$b = 0$$

Next a table is prepared containing x, $P_X(x)$, Y and Z. Y is plotted against Z and the corresponding values of $P_X(x)$ and x used to label the axes. These steps are illustrated in table 7.2 and figure 7.3 for the exponential distribution with $\lambda = 1/3$. Values of $P_X(x)$ were assumed and x, Y and Z calculated. For some distributions it may be easier to assume x and calculate $P_X(x)$. Finally Y is plotted against Z and the Y axis is labeled with the corresponding value of $P_X(x)$ and the Z axis with the corresponding x.

For many distributions, the exponential included, the same graph paper may be used for all values of the parameters of the distribution. Thus any exponentially distributed data will plot as a straight line on the paper shown at the right of figure 7.3. The slope of the line will change as the parameter λ changes. For some distributions such as

Table 7.2. Coordinates for exponential probability paper.

λ Equation 7.2	$P_X(x)$ Assumed	Y Equation 7.4	Z Equation 7.5
.033	.01	.011	.033
.153	.05	.051	.153
.318	.10	.106	.318
.669	.20	.223	.669
1.171	.30	.357	1.171
1.533	.40	.511	1.533
2.085	.50	.695	2.085
2.751	.60	.917	2.751
3.615	.70	1.205	3.615
4.833	.80	1.611	4.833
6.906	.90	2.302	6.906
9.018	.95	3.006	9.018
13.830	.99	4.610	13.830

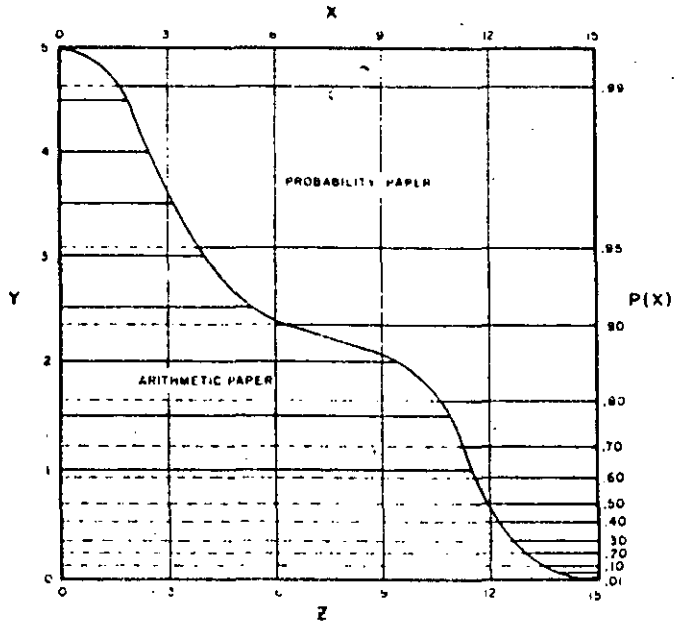


Fig. 7.3. Construction of exponential paper.

the gamma distribution a separate graph paper is required for different values of the parameters if the distribution is to plot as a straight line. As is apparent from the exponential distribution, the probability scale is not symmetrical for many distributions.

Equation 7.3 shows that standard semilogarithmic paper can be used as exponential

probability paper. If $1 - P_X(x)$ is plotted on the logarithmic scale and x is plotted on t arithmetic scale, a straight line with slope $-\lambda$ will result. The data of table 7.2 are plott in this fashion in figure 7.4. If $\ln(1 - P_X(x))$ is plotted versus λx, all exponential distri tions will plot as a single line. Note that it is not possible to plot the point $P_X(x) = 1.($ This is as expected since the exponential distribution asymptotically approaches, b never reaches, $P_X(x) = 1.00$.

Many types of probability paper are commercially available including paper for t normal, lognormal, exponential, certain cases of the gamma, extreme value (type Weibull and chi-square distributions.

PROBABILITY PLOTTING

A probability plot is a plot of a magnitude versus a probability. Determining probability to assign a data point is commonly referred to as determining the pl ting position. If one is dealing with a population, determining the plotting positi is merely a matter of determining the fraction of the data values less (greater) than equal to the value in question. Thus the smallest (largest) population value would plot 0 and the largest (smallest) population value would plot at 1.00. Assigning plotting p tions to sample data is not as straightforward because one can never be sure that a sa ple contains the smallest and largest values of the unknown population. Thus plott positions of 0 and 1 should be avoided for sample data unless one has additional inf mation on the population limits.

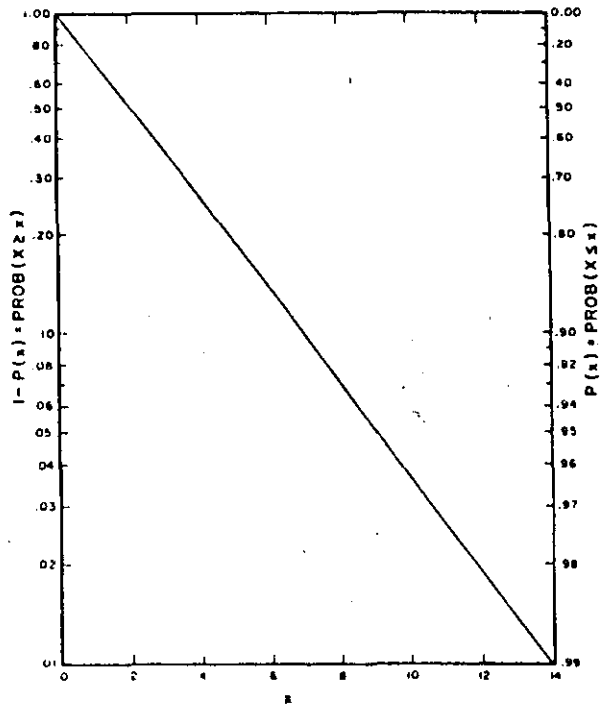


Fig. 7.4. Demonstration of case of semi-logarithmic paper for plotting exponentially tributed data.

Plotting positions may be expressed as a probability from 0 to 1 or a a percent from 0 to 100. Which method is being used should be clear from the context. In some discussions of probability plotting, especially in hydrologic literature, the probability scale is used to denote prob($X \geq x$) or $1 - P_X(x)$. In this book we will adopt this convention. The reason for it is the return period, $T_X(x)$, is $1/\text{prob}(X \geq x) = 1/(1 - P_X(x))$ or the reciprocal of the probability scale. One can always transform the probability scale from $1 - P_X(x)$ to $P_X(x)$ or even $T_X(x)$ if desired.

Probability plotting of hydrologic data requires that individual observations or data points be independent of each other and that the sample data be representative of the population (unbiased). There are four common types of sample data – complete duration series, annual series, partial duration series, and extreme value series.

The complete duration series consists of all available data. An example would be all the available daily flow data for a stream. (This particular data set would most likely not have independent observations.)

The annual series consists of one value per year such as the maximum peak flow each year. The data in table 2.1 is an annual series.

The partial duration series consists of all values above (below) a certain base. All peak flows above 40,000 cfs in the Kentucky River, Salvisa, Kentucky, would represent a partial duration series. This series may have more or less values in it than the annual series. For example there would be 5 years (1904, 1921, 1930, 1941, 1954) that would not have contributed any data to a partial duration series with a base of 40,000 cfs for the data in table 2.1; however, some years may have more than one peak above the base.

The annual series and the partial duration series have been shown by Langbein (1949) to be closely related and to approach one another for long return periods. Table 7.3 shows the relationship between the two series for return periods up to 100 years. Beard (1974) has shown that the relationship between annual series and partial duration series flood peaks varies throughout the U.S. and recommends the use of empirically derived, regionalized relationships. Frequently the annual series and the partial duration series are combined so that the largest (smallest) annual value plus all independent values above (below) some base are used.

The extreme value series consists of the largest (smallest) observation in a given time interval. The annual series is a special case of the extreme value series with the time interval being one year.

Regardless of the type of sample data used, the plotting position can be determined in the same manner. Gumbel (1958) states the following criteria for plotting position relationships:

1. The plotting position must be such that all observations can be plotted.

Table 7.3. Comparison of return periods for annual and partial duration series.

Partial Duration Series	Annual Series	Partial Duration Series	Annual Series
0.50	1.16	5	5.52
1.00	1.58	10	10.50
1.45	2.00	50	50.50
2.00	2.54	100	100.50

2. The plotting position should lie between the observed frequencies of $(m - 1)/n$ and m/n where m is the rank of the observation beginning with $m = 1$ for the largest (smallest) value and n is the number of years of record (if applicable) or the number of observations.
3. The return period of a value equal to or larger than the largest observation and the return period of a value equal to or smaller than the smallest observation should converge toward n.
4. The observations should be equally spaced on the frequency scale.
5. The plotting position should have an intuitive meaning, be analytically simple, and be easy to use.

Several plotting position relationships are presented in Chow (1964). Three of the most common relationships for plotting positions are shown in table 7.4. Unless specifically stated to the contrary, the Weibull relationship is used in the remainder of this book. Benson (1962a) in a comparative study of several plotting position relationships found on the basis of theoretical sampling from extreme value and the normal distributions that the Weibull relationship provided estimates that were consistent with experience.

The Weibull plotting position formula meets all 5 of the above criteria. (1) All of the observations can be plotted since the plotting positions range from $1/(n + 1)$ which is greater than zero to $n/(n + 1)$ which is less than one. Probability paper for many distributions does not contain the points zero and one. (2) The relationship $m/(n + 1)$ lies between $(m - 1)/n$ and m/n for all values of m and n. (3) The return period of the largest value is $(n + 1)/1$ which approaches n as n gets large and the return period of the smallest value is $(n + 1)/n = 1 + 1/n$ which approaches 1 as n gets large. (4) The difference between the plotting position of the $(m + 1)^{st}$ and m^{th} value is $1/(n + 1)$ for all values of m and n. (5) The fact that condition 3 is met plus the simplicity of the Weibull relationship fulfills condition 5.

One objection to the Hazen plotting position is that the return period for the largest $(m = 1)$ event is 2n or twice the record length. An objection to the California plotting position is the smallest value $(m = n)$ has a plotting position of 1 which implies the smallest sample value is the smallest possible value. A value of 1 cannot be plotted on many types of probability paper.

It should be noted that all of the relationships give similar values near the center of the distribution but may vary considerably in the tails. Predicting extreme events depends on the tails of the distribution so care must be exercised. The quantity $1 - P_X(x)$ represents the probability of an event with a magnitude equal to or greater than the event in question. When the data are ranked from the largest $(m = 1)$ to the smallest $(m = n)$ the plotting positions determined from equations 7.6 through 7.8 correspond to $1 - P_X(x)$. If the data are ranked from the smallest $(m = 1)$ to the largest $(m = n)$, the plotting position formulas are still valid; however, the plotting position now corresponds

Table 7.4. Plotting position relationships.

Name	Source	Relationship	Eq. No.
California	California (1923)	m/n	7.6
Hazen	Hazen (1930)	$(2m - 1)/2n$	7.7
Weibull	Weibull (1939)	$m/(n + 1)$	7.8

to the probability of an event equal to or smaller than the event in question which is $P_X(x)$. Probability paper may contain scales of $P_X(x)$, $1 - P_X(x)$, $T_X(x)$ or a combination of these.

As an example of probability plotting consider the data in table 2.1. The steps in plotting this data are:

1. Rank the data from the largest (smallest) to the smallest (largest) value. If two or more observations have the same value, several procedures can be used for assigning a plotting position. The procedure adopted here is to assume they have different values and assign each a unique rank. For example, in the following data, the value of 3 should be ranked both third and fourth; 1, 3, 4, 3, 2, 5.
2. Calculate the plotting position from equation 7.8.
3. Select the type of probability paper to be used. Normal probability paper is used in this example.
4. Plot the observations on the probability paper.

Table 7.5. Determination of plotting position for Kentucky River data.

Rank	Value	Plotting Position	Rank	Value	Plotting Position
1	115,000	1.49	34	68,600	50.7
2	112,000	2.99	35	67,800	52.2
3	111,000	4.48	36	67,200	53.7
4	99,100	5.97	37	66,900	55.2
5	96,100	7.46	38	66,300	56.7
6	94,300	8.95	39	65,700	58.2
7	93,700	10.45	40	62,600	59.7
8	92,500	11.90	41	61,300	61.2
9	91,500	13.40	42	60,200	62.7
10	89,400	14.90	43	58,000	64.2
11	88,700	16.40	44	58,000	65.7
12	87,200	17.90	45	55,000	67.1
13	87,100	19.40	46	54,400	68.6
14	85,000	20.90	47	53,600	70.2
15	84,300	22.40	48	53,500	71.6
16	84,100	23.90	49	52,400	73.2
17	82,900	25.40	50	52,300	74.6
18	80,900	26.90	51	50,500	76.1
19	80,400	28.30	52	47,300	77.6
20	80,100	29.90	53	47,000	79.1
21	80,000	31.30	54	46,800	80.6
22	79,200	32.80	55	46,100	82.1
23	77,700	34.30	56	46,000	83.5
24	77,000	35.80	57	45,000	85.1
25	73,400	37.30	58	44,200	86.5
26	72,900	38.80	59	44,000	88.0
27	71,700	40.30	60	43,000	89.5
28	71,200	41.80	61	40,300	91.0
29	70,800	43.30	62	34,700	92.5
30	70,500	44.80	63	34,100	94.0
31	70,000	46.20	64	32,300	95.5
32	69,000	47.80	65	28,400	97.0
33	,700	49.20	66	20,600	98.5

The data of table 2.1 are ranked and the plotting positions calculated in table 7.5. Figure 7.5 presents the plotted data. The theoretical normal line is drawn through the mean plus one standard deviation at 84.1%, and the mean minus one standard deviation at 15.9%. The normal distribution is generally not used for flood frequency analysis. It is used here because the data fit the normal reasonably well and because of the ease of constructing the best fitting straight line.

Many times data on one or two historical events which occurred prior to the initiation of any systematic data collection program is available or can be estimated. This data can be incorporated into the probability plot if the rank of the event is known. For ex-

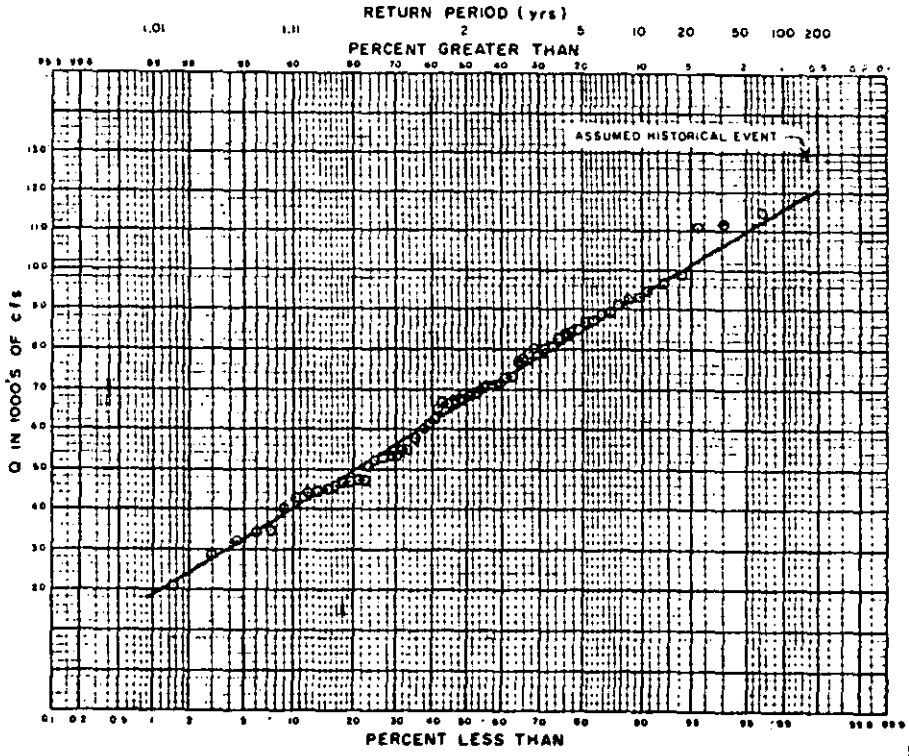


Fig. 7.5. Normal probability plot Kentucky River

ample if historical records indicate that in 1800 there was a peak flow on the Kentucky River near Salvisa (table 2.1) of 130,000 cfs, and this flow exceeded all flows since that time and nothing is known about flows before 1800, the event could be given a return period of 162 years (m = 1, n = 161) and a probability of 0.68 percent. The remaining data would then be analyzed as above. The assumed historical event is plotted with an asterisk in figure 7.5.

When probability plots of hydrologic data are made, frequently one or two extreme events are present that appear to be from a different population because they plot far off of the line defined by the other points. The treatment of these "outliers" is an unresolved and controversial question. The fact that this occurs frequently in hydrologic data should not be surprising. Using methods discussed in chapter 4, the probability of at least one occurrence of a n-year event in a k-year record can be calculated as $1 - (1 - 1/n)^k$. For example, the probability of at least one occurrence of a 100-year event in a 32-year record is $1 - 0.99^{32}$ or 0.275. If we have four independent 32-year records, we expect one to contain at least one 100-year event. This is the case even though the 100-year event is from the same population as the other 31 events in the 32-year record.

Benson (1962c) has stated that these extreme events can be treated if some historical information is available. If it is known that the largest event in a k-year record is actually the largest event to have occurred during the past n years, this large event can be assigned a return period of n + 1 years in accordance with equation 7.8. The second largest value would then be assigned a return period of (k + 1)/2 and so on for the remaining events. Dalrymple (1960) and Benson (1950) discuss in some detail the use of historical data in graphical hydrologic frequency analyses.

In the absence of historical data, the investigator has little information on which to base the return period of an extreme event. Hazen (1930) discusses a method of utilizing flow information on nearby streams. Other types of "regional" analysis are discussed by Beard (1962), Dalrymple (1960), Benson (1950) and later in this chapter. Beard (1962), Dalrymple (1960) and Benson (1968) and others state that a frequency analysis of a single short record is relatively unreliable and they recommend a regional analysis to produce more information that can be applied to the record in question. This would especially be the case if one were trying to estimate the magnitude of an event with a long return period from a short record.

When probability plots are made and a line drawn through the data, the tendency to extrapolate the data to high return periods is great. The distance on the probability paper from a return period of 20 years to a return period of 200 years is not very much; however, if the data do not truly follow the assumed distribution with population parameters equal to the sample statistics (i.e., $\mu = \bar{X}$ and $\sigma^2 = s^2$ for the normal), the error in this extrapolation can be quite large. This fact has already been referred to when it was stated that the estimation of probabilities in the tails of distributions is very sensitive to distributional assumptions. Since one of the usual purposes of probability plotting is to estimate events with longer return periods, Blench (1959) and Dalrymple (1960) have criticized the blind use of analytical flood frequency methods because of this tendency toward extrapolation.

If a set of data plots as a straight line on probability paper, the data can be said to be distributed as the distribution corresponding to the probability paper. Since it would be rare for a set of data to plot exactly on a line, a decision must be made as to whether or not the deviations from the line are random deviations or represent true deviations indicating the data does not follow the given probability distribution. Examining figure 7.5, it is apparent that the deviations from a straight line are small so it can be assumed

that the data can be approximated by the normal distribution.

So far two tests, both based on judgement, have been described for determining if a set of data follows a certain distribution. The first method was to visually compare observed and theoretical frequency histograms and the second to visually compare observed and theoretical cumulative frequency curves in the form of probability plots. In Chapter 8 statistical tests based on these two visual tests will be presented.

ANALYTICAL HYDROLOGIC FREQUENCY ANALYSIS

Much of the material that has been discussed thus far could properly fall under the heading of frequency analysis. Frequency analysis is merely a procedure for estimating the frequency of occurrence or probability of occurrence of past and/or future events. Thus probability plotting with or without any distributional assumptions is a method of frequency analysis.

Frequency analysis of hydrologic data requires that the data be homogeneous and independent. The restriction of homogeneity assures that all the observations are from the same population (i.e., a stream gaging station has not been moved, a watershed has not become urbanized, or no structures have been placed on the stream or its major tributaries). The restriction of independence assures that a hydrologic event such as a single large storm does not enter the data set more than once. For example a single storm system may produce two or more large runoff peaks only one of which (the largest) should enter the data set. Further for the prediction of the frequency of future events the restriction of homogeneity requires that the data on hand be representative of future flows (i.e. there will be no new structures, diversions, land use changes, etc. in the case of streamflow data).

Hydrologic frequency analyses can be made with or without making any distributional assumptions. The procedure to be followed in either case is much the same. If no distributional assumptions are made, the investigator merely plots the observed data on any kind of paper (not necessarily probability paper) and uses his best judgement to determine the magnitude of past or future events for various return periods. The previous section discussed flood frequency analysis without the use of an analytical frequency distribution. If a distributional assumption is made, the magnitude of events for various return periods is selected from the theoretical "best-fit" line according to the assumed distribution. If an analytical technique is used, it is recommended that the data still be plotted so that one can get an idea of how well the data fit the assumed analytical form and to spot potential problems.

Chow (1951) has shown that many frequency analyses can be reduced to the form

$$X_T = \bar{X}(1 + C_v K_T) \tag{7.9}$$

where X_T is the magnitude of the event having a return period T and K_T is a frequency factor. This relationship comes about by writing any X as

$$X = \bar{X} + \Delta X \tag{7.10}$$

and then stating that ΔX, the deviation from the mean, is the product of the standard deviation s and a frequency factor K.

$$X = \bar{X} + sK \tag{7.11}$$

Recalling that $C_v = s/\bar{X}$; equation 7.11 takes on the form of equation 7.9. Chow (1951, 1964) presents the frequency factors for many different types of frequency distributions. Some of these are presented here.

Equation 7.9 can also be used to construct the probability scale on plotting paper so that the distribution corresponding to K_T plots as a straight line.

Normal Distribution

For the normal distribution it can easily be shown that K_T is the standardized normal variate Z given by table E.4. A standard normal table can be used along with equation 7.9 to determine the magnitude of normally distributed events corresponding to various probabilities. For example the magnitude of a 20-year peak flow for the data of table 2.1 can be determined by calculating

$$c_v = s_x / x = 21{,}000/67{,}500 = .311$$

and

$$1 - P_X(x) = 1/T_X(x) = 1/20 = .05$$

The 20-year event corresponds to a prob($X \geq x$) of .05 so the probability of an event less than the 20-year event is 0.95. The value of Z corresponding to a probability of 0.95 is found from standard normal tables to be 1.645. Thus

$$X_{20} = \bar{x}\,(1 + c_v K_{20})$$

$$= 67{,}500\,(1 + 0.311 \times 1.645)$$

$$= 102{,}000 \text{ cfs}$$

which agrees with the value given by figure 7.5.

Lognormal Distribution

For the lognormal distribution, the frequency factor is given by Chow (1964) as

$$K_T = [\exp(\sigma_Y K_Y - \sigma_Y^2/2) - 1]/[\exp(\sigma_Y^2) - 1]^{1/2} \qquad (7.12)$$

where $Y = \ln X$ and

$$K_Y = (Y_T - \mu_Y)/\sigma_Y$$

Frequency factors for the lognormal distribution are given in table 7.6. The numbers in columns 2, 3, 4, and 5 are negative.

Log Pearson Type III Distribution

Benson (1968) reported on a method of flood frequency analysis based on the log Pearson type III distribution which is obtained when the logs of observed data are used along with the Pearson type III distribution (equation 6.77). This method is applied as follows:

1. Transform the n annual flood magnitudes, X_i, to their logarithmic values, Y_i (i.e. $Y_i = \log X_i$ for $i = 1, 2, \ldots, n$).
2. Compute the mean logarithm, Y.
3. Compute the standard deviation of the logarithms, S_Y.

Table 7.6. Frequency factors for lognormal distribution (Chow, 1964).

	Probability in percent equal to or greater than the given variate								
	99	95	80	50	20	5	1	0.1	c_v
	−	−	−		+	+	+	+	
0	2.33	1.65	0.84	0	0.84	1.64	2.33	3.09	0
0.1	2.25	1.62	0.85	0.02	0.84	1.67	2.40	3.22	0.033
0.2	2.18	1.59	0.85	0.04	0.83	1.70	2.47	3.39	0.067
0.3	2.11	1.56	0.85	0.06	0.82	1.72	2.55	3.56	0.100
0.4	2.04	1.53	0.85	0.07	0.81	1.75	2.62	3.72	0.136
0.5	1.98	1.49	0.86	0.09	0.80	1.77	2.70	3.88	0.166
0.6	1.91	1.46	0.85	0.10	0.79	1.79	2.77	4.05	0.197
0.7	1.85	1.43	0.85	0.11	0.78	1.81	2.84	4.21	0.230
0.8	1.79	1.40	0.84	0.13	0.77	1.82	2.90	4.37	0.262
0.9	1.74	1.37	0.84	0.14	0.76	1.84	2.97	4.55	0.292
1.0	1.68	1.34	0.84	0.15	0.75	1.85	3.03	4.72	0.324
1.1	1.63	1.31	0.83	0.16	0.73	1.86	3.09	4.87	0.351
1.2	1.58	1.29	0.82	0.17	0.72	1.87	3.15	5.04	0.381
1.3	1.54	1.26	0.82	0.18	0.71	1.88	3.21	5.19	0.409
1.4	1.49	1.23	0.81	0.19	0.69	1.88	3.26	5.35	0.436
1.5	1.45	1.21	0.81	0.20	0.68	1.89	3.31	5.51	0.462
1.6	1.41	1.18	0.80	0.21	0.67	1.89	3.36	5.66	0.490
1.7	1.38	1.16	0.79	0.22	0.65	1.89	3.40	5.80	0.517
1.8	1.34	1.14	0.78	0.22	0.64	1.89	3.44	5.96	0.544
1.9	1.31	1.12	0.78	0.23	0.63	1.89	3.48	6.10	0.570
2.0	1.28	1.10	0.77	0.24	0.61	1.89	3.52	6.25	0.596
2.1	1.25	1.08	0.76	0.24	0.60	1.89	3.55	6.39	0.620
2.2	1.22	1.06	0.76	0.25	0.59	1.89	3.59	6.51	0.643
2.3	1.20	1.04	0.75	0.25	0.58	1.88	3.62	6.65	0.667
2.4	1.17	1.02	0.74	0.26	0.57	1.88	3.65	6.77	0.691
2.5	1.15	1.00	0.74	0.26	0.56	1.88	3.67	6.90	0.713
2.6	1.12	0.99	0.73	0.26	0.55	1.87	3.70	7.02	0.734
2.7	1.10	0.97	0.72	0.27	0.54	1.87	3.72	7.13	0.755
2.8	1.08	0.96	0.72	0.27	0.53	1.86	3.74	7.25	0.776
2.9	1.06	0.95	0.71	0.27	0.52	1.86	3.76	7.36	0.796
3.0	1.04	0.93	0.71	0.28	0.51	1.85	3.78	7.47	0.818
3.2	1.01	0.90	0.69	0.28	0.49	1.84	3.81	7.65	0.857
3.4	0.98	0.88	0.68	0.29	0.47	1.83	3.84	7.84	0.895
3.6	0.95	0.86	0.67	0.29	0.46	1.81	3.87	8.00	0.930
3.8	0.92	0.84	0.66	0.29	0.44	1.80	3.89	8.16	0.966
4.0	0.90	0.82	0.65	0.29	0.42	1.78	3.91	8.30	1.000
4.5	0.84	0.78	0.63	0.30	0.39	1.75	3.93	8.60	1.081
5.0	0.80	0.74	0.62	0.30	0.37	1.71	3.95	8.86	1.155

Table 7.7. K_T values for positive skew coefficients Pearson type III distribution.[1]

Skew Coef.	Recurrence Interval in Years							
	1.0101	2	5	10	25	50	100	200
γ	Percent Chance (≥)							
	99	50	20	10	4	2	1	0.5
3.0	-0.667	-0.396	0.420	1.180	2.278	3.152	4.051	4.970
2.9	-0.690	-0.390	0.440	1.195	2.277	3.134	4.013	4.904
2.8	-0.714	-0.384	0.460	1.210	2.275	3.114	3.973	4.847
2.7	-0.740	-0.376	0.479	1.224	2.272	3.093	3.932	4.783
2.6	-0.769	-0.368	0.499	1.238	2.267	3.071	3.889	4.718
2.5	-0.799	-0.360	0.518	1.250	2.262	3.048	3.845	4.652
2.4	-0.832	-0.351	0.537	1.262	2.256	3.023	3.800	4.584
2.3	-0.867	-0.341	0.555	1.274	2.248	2.997	3.753	4.515
2.2	-0.905	-0.330	0.574	1.284	2.240	2.970	3.705	4.444
2.1	-0.946	-0.319	0.592	1.294	2.230	2.942	3.656	4.372
2.0	-0.990	-0.307	0.609	1.302	2.219	2.912	3.605	4.298
1.9	-1.037	-0.294	0.627	1.310	2.207	2.881	3.553	4.223
1.8	-1.087	-0.282	0.643	1.318	2.193	2.848	3.499	4.147
1.7	-1.140	-0.268	0.660	1.324	2.179	2.815	3.444	4.069
1.6	-1.197	-0.254	0.675	1.329	2.163	2.780	3.388	3.990
1.5	-1.256	-0.240	0.690	1.333	2.146	2.743	3.330	3.910
1.4	-1.318	-0.225	0.705	1.337	2.128	2.706	3.271	3.828
1.3	-1.383	-0.210	0.719	1.339	2.108	2.666	3.211	3.745
1.2	-1.449	-0.195	0.732	1.340	2.087	2.626	3.149	3.661
1.1	-1.518	-0.180	0.745	1.341	2.066	2.585	3.087	3.575
1.0	-1.588	-0.164	0.758	1.340	2.043	2.542	3.022	3.489
.9	-1.660	-0.148	0.769	1.339	2.018	2.498	2.957	3.401
.8	-1.733	-0.132	0.780	1.336	1.993	2.453	2.891	3.312
.7	-1.806	-0.116	0.790	1.333	1.967	2.407	2.824	3.223
.6	-1.880	-0.099	0.800	1.328	1.939	2.359	2.755	3.132
.5	-1.955	-0.083	0.808	1.323	1.910	2.311	2.686	3.041
.4	-2.029	-0.066	0.816	1.317	1.880	2.261	2.615	2.949
.3	-2.104	-0.050	0.824	1.309	1.849	2.211	2.544	2.856
.2	-2.178	-0.033	0.830	1.301	1.818	2.159	2.472	2.763
.1	-2.252	-0.017	0.836	1.292	1.785	2.107	2.400	2.670
0	-2.326	0	0.842	1.282	1.751	2.054	2.326	2.576

contd.

4. Compute the coefficient of skewness, C_s.

$$C_s = [n^2\Sigma y^3 - 3n\Sigma y\Sigma y^2 + 2(\Sigma y)^3]/[(n-1)(n-2)s_y^3] \quad (7.13)$$

$$= n\Sigma(y - \bar{y})^3/[(n-1)(n-2)s_y^3]$$

5. Compute

Table 7.7 (contd.). K_T values for negative skew coefficients Pearson type III distribution.[1]

Skew Coef.	Recurrence Interval in Years							
	1.0101	2	5	10	25	50	100	200
γ	Percent Chance (≥)							
	99	50	20	10	4	2	1	0.5
.0	-2.326	0	0.842	1.282	1.751	2.054	2.326	2.576
-.1	-2.400	0.017	0.846	1.270	1.716	2.000	2.252	2.482
-.2	-2.472	0.033	0.850	1.258	1.680	1.945	2.178	2.388
-.3	-2.544	0.050	0.853	1.245	1.643	1.890	2.104	2.294
-.4	-2.615	0.066	0.855	1.231	1.606	1.834	2.029	2.201
-.5	-2.686	0.083	0.856	1.216	1.567	1.777	1.955	2.108
-.6	-2.755	0.099	0.857	1.200	1.528	1.720	1.880	2.016
-.7	-2.824	0.116	0.857	1.183	1.488	1.663	1.806	1.926
-.8	-2.891	0.132	0.856	1.166	1.448	1.606	1.733	1.837
-.9	-2.957	0.148	0.854	1.147	1.407	1.549	1.660	1.749
-1.0	-3.022	0.164	0.852	1.128	1.366	1.492	1.588	1.664
-1.1	-3.087	0.180	0.848	1.107	1.324	1.435	1.518	1.581
-1.2	-3.149	0.195	0.844	1.086	1.282	1.379	1.449	1.501
-1.3	-3.211	0.210	0.838	1.064	1.240	1.324	1.383	1.424
-1.4	-3.271	0.225	0.832	1.041	1.198	1.270	1.318	1.351
-1.5	-3.330	0.240	0.825	1.018	1.157	1.217	1.256	1.282
-1.6	-3.388	0.254	0.817	0.994	1.116	1.166	1.197	1.216
-1.7	-3.444	0.268	0.808	0.970	1.075	1.116	1.140	1.155
-1.8	-3.499	0.282	0.799	0.945	1.035	1.069	1.087	1.097
-1.9	-3.553	0.294	0.788	0.920	0.996	1.023	1.037	1.044
-2.0	-3.605	0.307	0.777	0.895	0.959	0.980	0.990	0.995
-2.1	-3.656	0.319	0.765	0.869	0.923	0.939	0.946	0.949
-2.2	-3.705	0.330	0.752	0.844	0.888	0.900	0.905	0.907
-2.3	-3.753	0.341	0.739	0.819	0.855	0.864	0.867	0.869
-2.4	-3.800	0.351	0.725	0.795	0.823	0.830	0.832	0.833
-2.5	-3.845	0.360	0.711	0.771	0.793	0.798	0.799	0.800
-2.6	-3.889	0.368	0.696	0.747	0.764	0.768	0.769	0.769
-2.7	-3.932	0.376	0.681	0.724	0.738	0.740	0.740	0.741
-2.8	-3.973	0.384	0.666	0.702	0.712	0.714	0.714	0.714
-2.9	-4.013	0.390	0.651	0.681	0.683	0.689	0.690	0.690
-3.0	-4.051	0.396	0.636	0.660	0.666	0.666	0.667	0.667

1. Water Resources Council (1967).

$$Y_T = \bar{y} + s_Y K_T$$

where K_T is obtained from table 7.7. Note that this relationship is identical to equation 7.13 except the logarithms are used.

6. Compute X_T = antilog Y_T.

This method has as a special case the lognormal distribution when $C_s = 0$. For short

periods of record the skew coefficient calculated from equation 7.13 may not be a reliable estimate of the population skew coefficient and then it may be desirable to replace it with a regionalized coefficient (Beard 1962, 1974; Benson 1968). Figure 7.6 contains regionalized skew coefficients of annual streamflow maximum logarithms computed by the U. S. Geological Survey.

The frequency factors of table 7.7 can be used for the Pearson type III distribution in the same manner as for the log Pearson type III. The actual data values rather than their logarithms would then be used.

Extreme Value Type I Distribution

Chow (1951) presents the following relationship for the frequency factor for the extreme value type I maximum distribution

$$K_T = -0.7797\{0.5772 + \ln[\ln T_X(x)/(T_X(x) - 1)]\} \qquad (7.14)$$

where $T_X(x)$ is the desired return period of the quantity being calculated. Potter (1949) presented some curves that simplified the application of the extreme value type I. Kendall (1959) presents the frequency factors shown in table 7.8 for the extreme value type I distribution. The values computed from equation 7.14 are equivalent to an infinite sample size in table 7.8.

Other Distributions

Any of the distributions discussed in Chapter 6 can be fit to data by using the methods discussed in that chapter. Frequency factors for some of the other distributions are given by Chow (1951, 1964).

General Considerations

Many proponents (and opponents) of one analytical form for flood flow frequencies or another have come to the fore over the past few decades. The proponents claim

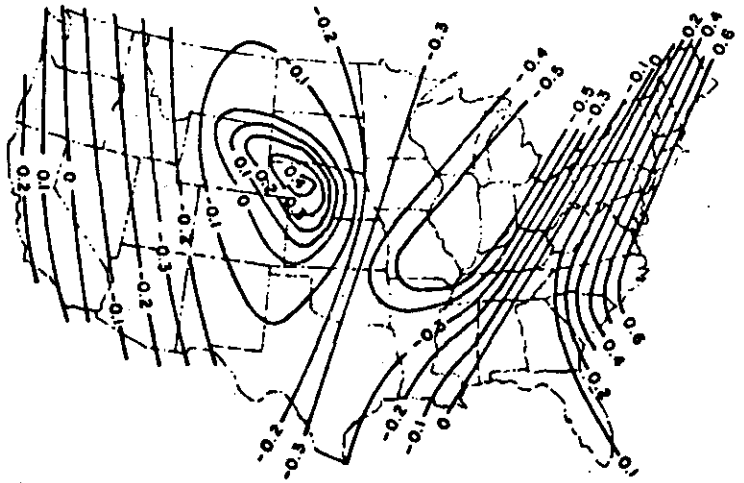


Fig. 7.6. Generalized skew coefficients of annual maximum streamflow logarithms.

Table 7.8. Frequency factors for extreme value type I distribution.

Sample size n	Return Period 5	10	15	20	25	50	75	100	1000
15	0.967	1.703	2.117	2.410	2.632	3.321	3.721	4.005	6.265
20	0.919	1.625	2.023	2.302	2.517	3.179	3.563	3.836	6.006
25	0.888	1.575	1.963	2.235	2.444	3.088	3.463	3.729	5.842
30	0.866	1.541	1.922	2.188	2.393	3.026	3.393	3.653	5.727
35	0.851	1.516	1.891	2.152	2.354	2.979	3.341	3.598	
40	0.838	1.495	1.866	2.126	2.326	2.943	3.301	3.554	5.576
45	0.829	1.478	1.847	2.104	2.303	2.913	3.268	3.520	
50	0.820	1.466	1.831	2.086	2.283	2.889	3.241	3.491	5.478
55	0.813	1.455	1.818	2.071	2.267	2.869	3.219	3.467	
60	0.807	1.446	1.806	2.059	2.253	2.852	3.200	3.446	
65	0.801	1.437	1.796	2.048	2.241	2.837	3.183	3.429	
70	0.797	1.430	1.788	2.038	2.230	2.824	3.169	3.413	5.359
75	0.792	1.423	1.780	2.029	2.220	2.812	3.155	3.400	
80	0.788	1.417	1.773	2.020	2.212	2.802	3.145	3.387	
85	0.785	1.413	1.767	2.013	2.205	2.793	3.135	3.376	
90	0.782	1.409	1.762	2.007	2.198	2.785	3.125	3.367	
95	0.780	1.405	1.757	2.002	2.193	2.777	3.116	3.357	
100	0.779	1.401	1.752	1.998	2.187	2.770	3.109	3.349	5.261
∞	0.719	1.305	1.635	1.866	2.044	2.592	2.911	3.137	4.936

that some particular method is superior to some other method and "prove" their claim by a few rationalizations and some case studies. The fact remains that these rationalizations involve questionable assumptions. There is no direct theoretical connection between any analytical form of the frequency distribution and the underlying mechanisms governing flood flows except through the limit theorems. The primary consideration in selecting a particular analytical form for the frequency distribution is that the distribution "fit" the observed data (D. V. Anderson 1967; Benson 1968).

Benson (1968) reported on the results of a study by a work group consisting of 18 representatives from 12 federal agencies of the United States government. This group studied six methods of flood frequency analysis on 10 streams located throughout the United States. The records on these streams ranged in length from 40 to 97 years with an average of 55 years. The drainage areas ranged from 16.4 to 36,800 square miles. The six methods of analysis consisted of (1) the gamma distribution, (2) Gumbel distribution, (3) Gumbel distribution using the logarithms of the data, (4) lognormal distribution, (5) log Pearson type III distribution and (6) Hazen's method. The computational procedures used were much like those presented in this book. The Hazen method consists of using an equation like equation 7.9 along with a table of empirically derived frequency factors

that are a function of the return period and the coefficient of skew (Hazen 1930).

Large differences were produced by the six different methods especially at long return periods. The results showed that the lognormal, log Pearson type III and Hazen methods were about equally good. The group suggested that the log Pearson type III be used unless there was a good reason to use some other method. This recommendation was made even though the group realized that "there are no rigorous statistical criteria on which to base a choice of method". Benson's (1968) report states that the study showed that "the range of uncertainty in flood analysis, regardless of the method used, is still quite large" and that many questions concerning it remain unresolved.

In a follow-up study, Beard (1974) examined flood peaks from 300 stations scattered throughout the U. S. Several probability distributions were tried including the log Pearson type III, lognormal, Gumbel's extreme value distribution, and the two- and three-parameter gamma distributions. Beard concluded that only the lognormal and log Pearson type III with a regionalized skew coefficient were not greatly biased in estimating future flood frequencies. He stated that the latter distribution produced somewhat more consistent results but that "... regardless of the methodology employed, substantial uncertainty in frequency estimates from station data will exist...".

In selecting a particular analytical form for a frequency curve, one may be tempted to select a distribution with a large number of parameters. Generally the more parameters a distribution has, the better it will adapt to a set of data. However, for the sample size usually available in hydrology, the reliability in estimating more than 2 or 3 parameters may be quite low. Thus a compromise must be made between flexibility of the distribution and reliability of the parameters.

Recognizing the short record lengths available for frequency analyses, methods of augmenting natural data by synthetic data are being developed. In some cases the rainfall record pertaining to a watershed is much longer than its streamflow record. In this event it may be possible to calibrate a deterministic streamflow model to the watershed and then use the long rainfall record to generate a long synthetic streamflow hydrograph. This synthetic hydrograph can then be combined with existing data into a single frequency analysis. In the absence of rainfall records, it may be possible to transfer records from a nearby station or to stochastically generate a series of rainfall data. This data could then be used with the calibrated deterministic model to augment natural streamflow data. One might consider weighting the natural data more than the augmented data in the final frequency analysis. Carey and Haan (1975) present a method for determining the useability of a deterministic model for extending streamflow records for subsequent analysis via a statistical model.

Treatment of Zeros

Most hydrologic variables are bounded on the left by zero. A zero in a set of data that is being logarithmically transformed requires special handling. One solution is to add a small constant to all of the observations. Another method is to analyze the non-zero values and then adjust the relation to the full period of record. This method biases the results as the zero values are essentially ignored. A third and theoretically more sound method would be to use the theorem of total probability (equation 2.10).

$$\text{prob}(X \le x) = \text{prob}(X \le x \mid X = 0)\,\text{prob}(X = 0) + \text{prob}(X \le x \mid X \neq 0)\,\text{prob}(X \neq 0)$$

Since $\text{prob}(X \le x \mid X = 0)$ is zero, the relationship reduces to

$$\text{prob}(X \le x) = \text{prob}(X \neq 0)\,\text{prob}(X \le x \mid X \neq 0)$$

In this relationship $\text{prob}(X \neq 0)$ would be estimated by the fraction of non-zero values and $\text{prob}(X \ge x \mid X \neq 0)$ would be estimated by a standard analysis of the non-zero values with the sample size taken to be equal to the number of non-zero values. This relation can be written as a function of cumulative probability distributions.

$$1 - P_X(x) = k[1 - P_X^*(x)]$$

or

$$P_X(x) = 1 - k + kP_X^*(x) \tag{7.15}$$

where $P_X(x)$ is the cumulative probability distribution of all X ($\text{prob}(X \le x \mid X \ge 0)$), k is the probability that X is not zero, and $P_X^*(x)$ is the cumulative probability distribution of the non-zero values of X (i.e. $\text{prob}(X \le x \mid X \neq 0)$). This type of mixed distribution with a finite probability that X = 0 and a continuous distribution of probability for X>0 was discussed in Chapter 2. Jennings and Benson (1969) have demonstrated the applicability of this approach to analyzing flood flow frequencies with zeros present.

Equation 7.15 can be used to estimate the magnitude of an event with return period $T_X(x)$ by solving first for $P_X^*(x)$ and then using the inverse transformation of $P_X^*(x)$ to get the value of X. For example the 10-year event with k = 0.95 is found to be the value of X satisfying

$$P_X^*(x) = [P_X(x) - 1 + k]/k = (0.9 - 1 + 0.95)/0.95 = 0.89$$

Note that it is possible to generate negative estimates for $P_X^*(x)$ from equation 7.15. For example if k = 0.50 and $P_X(x) = 0.05$, the estimated $P_X^*(x)$ is

$$P_X^*(x) = [P_X(x) - 1 + k]/k = (0.05 - 1 + 0.50)/0.5 = -0.9$$

This merely means that the value of X corresponding to $P_X(x) = 0.05$ is zero.

Example 7.1. Seventy-five years of peak flow data are available from an annual series; 20 of the values are zero; and the remainin 55 values have a mean of 100 cfs, a standard deviation of 35.1 cfs, and are lognormally distributed. (a) Estimate the probability of a peak exceeding 125 cfs. (b) Estimate the magnitude of the 25-year peak flow.

Solution:

(a) $\text{prob}(X \ge 125) = 1 - \text{prob}(X < 125) = 1 - P_X(125)$

Applying equation 7.15

$$P_X(125) = 1 - k + k\,P_X^*(125)$$

$$k = 55/75 = 0.733$$

$P_X^*(125)$ can be evaluated by solving equation 7.9 for K_T and then using table 7.6 to get the probability

$$X_T = \bar{X}(1 + C_v K_T)$$

$$125 = 100(1 + 35.1\ K_T/100)$$

$$K_T = 0.712$$

By interpolation from table 7.6, this K_T for a C_v of 0.351 corresponds to a probability of 0.21 for $X \geq x$ or 0.79 for $X \leq x$.

$$P_X(125) = 1 - 0.733 + 0.733(0.79) = 0.846$$

$$\text{prob}(X>125) = 1 - P_X(125) = 0.154$$

The probability of a peak flow in any year exceeding 125 cfs is 0.154. The conditional probability of a peak exceeding 125 cfs given that the peak is not zero is 0.21.

(b) $$P_X^*(x) = [P_X(x) - 1 + k]/k = [1 - (1/T) - 1 + k]/k$$

$$= (1 - 0.04 - 1 + 0.733)/0.733 = 0.945$$

The value of X corresponding to $P_X^*(x) = 0.945$ can be obtained by interpolation from table 7.6. However, since this table is nonlinear, the equation following equation 7.12 will be used. From equations 6.30 and 6.31

$$\bar{y} = \ln [\bar{x}^2/(c_v^2 + 1)]/2 = 4.547$$

$$s_Y^2 = \ln (c_v^2 + 1) = 0.116$$

$$K_Y = (Y_T - \bar{y})/s_Y$$

K_Y corresponding to $P_X^*(x) = 0.945$ is obtained from table E.4 as 1.60. Therefore

$$Y_T = 1.60(0.116)^{1/2} + 4.547 = 5.092$$

$$X_T = \exp(Y_T) = 163 \text{ cfs}$$

Example 7.2. Appendix B contains a listing of the peak discharges on the Cumberland River at Cumberland Falls, Kentucky. Beginning in 1934 the data consists of a partial duration series. In this problem only the data from 1934 through 1970 are used.

a) Plot the annual series and partial duration series on lognormal probability paper.
b) Plot the "best" fitting lognormal distribution of the annual series on the plot of part a.
c) Plot the expected partial duration series frequency distribution on the plot of part a assuming the annual series follows the lognormal distribution.
d) Plot the annual series on normal probability paper. Draw in the "best" fitting normal distribution.
e) Plot the annual series on extreme value type I probability paper. Draw in the "best" fitting extreme value type I distribution.
f) Compute the 100-year peak flow based on the annual series assuming the data follows the (1) normal, (2) lognormal, (3) extreme value type I, (4) Pearson type III and (5) log-Pearson type III distributions.

Solution:

The following table contains the data that was used in the solution to this problem. The annual series data represents the largest peak flow in each of the 37 years from 1934 through 1970. The partial duration series data is the largest 37 values observed during the 37-year period. The plotting position was calculated from the Weibull relationship m/(n+1) where m is the rank and n is 37.

(a) Figures 7.7, 7.8 and 7.9 are the lognormal, normal and extreme value type plots. The lognormal plot contains both the annual series and partial duration series data. The other two plots contain only the annual series data.

(b) Several points on the best fitting lognormal line were calculated from $X_T = \bar{X}(1 + C_v K_T)$ where K_T was interpolated from table 7.6 for a C_v of 0.3165. For e

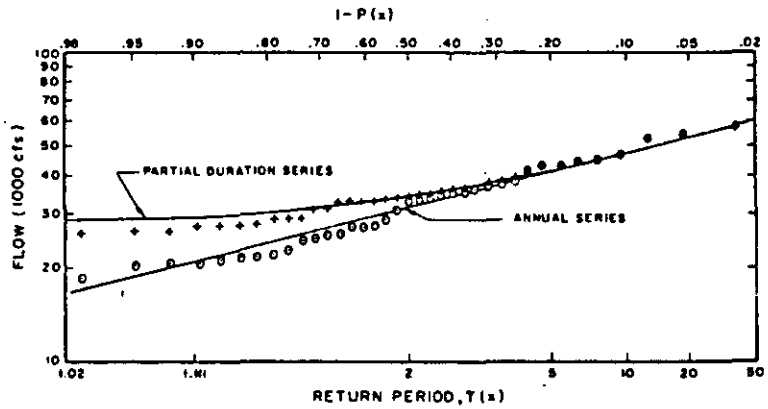


Fig. 7.7. Lognormal probability plot for data of example 7.2.

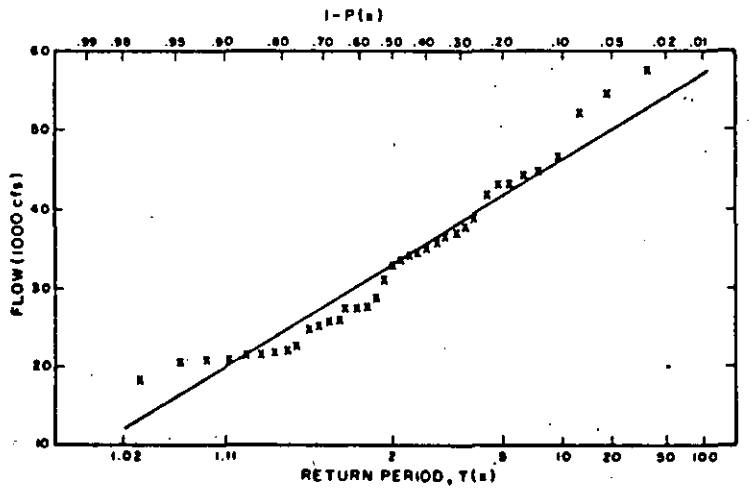


Fig. 7.8. Normal probability plot for data of exar 7.2.

Data used in solution of example 7.2.

Rank	Plotting Position Rank/38	Flow (1000 cfs) Annual Series	Flow (1000 cfs) Partial Duration Series
1	.0263	57.4	57.4
2	.0526	54.2	54.2
3	.0789	52.3	52.3
4	.1050	46.5	46.5
5	.1320	44.8	44.8
6	.1580	44.1	44.1
7	.1840	43.3	43.3
8	.2110	43.1	43.1
9	.2370	41.6	41.6
10	.2630	38.8	40.0
11	.2890	37.6	38.8
12	.3160	37.0	37.6
13	.3420	36.3	37.6
14	.3680	35.8	37.0
15	.3950	35.1	36.3
16	.4210	34.8	35.8
17	.4470	34.2	35.1
18	.4740	33.7	34.8
19	.5000	33.0	34.2
20	.5260	31.0	33.7
21	.5530	28.9	33.7
22	.5790	27.5	33.1
23	.6050	27.4	33.0
24	.6320	27.3	33.0
25	.6580	25.9	32.5
26	.6840	25.8	31.0
27	.7110	25.0	30.8
28	.7370	24.9	29.0
29	.7630	22.5	28.9
30	.7890	21.8	28.7
31	.8160	21.7	27.6
32	.8420	21.4	27.5
33	.8680	21.4	27.4
34	.8950	20.9	27.3
35	.9210	20.7	26.2
36	.9470	20.4	26.1
37	.9740	18.2	25.9

Annual Series.	
X = Flow (cfs)	Y = lnX
$\bar{x}$ = 32.873	$\bar{y}$ = 10.353
s_x = 10.405	s_y = 0.3147
C_{vx} = 0.3165	C_{vy} = 0.0304
C_{sx} = 0.600	C_{sy} = 0.031

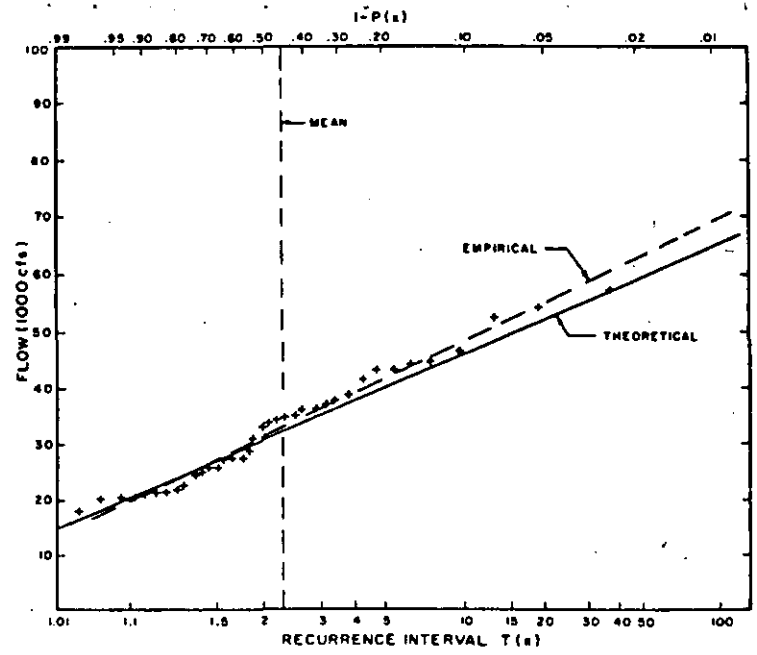


Fig. 7.9. Extreme value type I probability plot for data of example 7.2.

ample for $1 - P_X(x) = .05$, K_T is 1.85. Thus $X_T = 32.873\ (1 + 0.3165(1.85))$ or 52,122 cfs. This same value can be obtained by noting the $Y = \ln X$ is normally distributed. Therefore Y corresponding to $1 - P_X(x) = .05$ can be obtained from $Y_T = \bar{Y}(1 + C_v K_T)$ where K_T is found to be 1.645 from a standard table. Thus $Y_T = 10.353(1 + .03049\ (1.645)) = 10.87070$. Now $X_T = \exp(Y_T) = \exp(10.871) = 52{,}600$. A straight line was drawn through the calculated points.

(c) The expected partial duration series was computed using table 7.3 and the straight line lognormal frequency calculated in part b. For example, the partial duration return period corresponding to an annual series return period of two years is 1.45 years. The 2-year lognormal flow is 31,500 cfs (from graph). This point is plotted at a return period of 1.45. This procedure was repeated for several return periods and a smooth curve drawn through the resulting points.

(d) The normal distribution plot is shown in figure 7.8. The straight line was drawn through the points:

$\bar{x}$ or 32,873 and 50%

$\bar{x} + s$ or 32,873 + 10,405 = 43,278 and 15.9%

$\bar{x} - s$ or 32,873 - 10,405 = 27,468 and 84.1%

(e) The extreme value type I distribution is shown in figure 7.9. The straight line was again calculated from $X_T = \bar{X}(1 + C_v K_T)$ where K_T is from table 7.8. The "theo-

retical" line was determined using a K_T corresponding to equation 7.14 while the "empirical" line was determined using a K_T corresponding to 40 observations.

(f) In all cases except one, the 100-year flow is calculated from $X_{100} = \bar{X}(1+C_v K_{100})$ using the appropriate K_{100}. For the log-Pearson type III the value is from $X_{100} = \exp(Y_{100})$ where $Y_{100} = \bar{Y}(1+C_v K_{100})$. The results are summarized in the following table.

Distribution	100-Year Flow		Table No.
	K_{100}	Q_{100}	For K_{100}
normal	2.33	57,115	E.4
lognormal	3.02	64,300	7.6
ex. value I	3.14	65,542	eq. 7.14
ex. value I (n=40)	3.55	69,825	7.8
Pearson III	2.76	61,588	7.7
log-Pearson III	2.33	65,268	7.7

Comment: This problem illustrates the application of several different probability distributions in making a hydrologic frequency analysis at a single station. The data used in this example is well "behaved" in that there are no outliers and the data plots reasonably well on probability paper.

The predicted 100-year flow varied from 57,115 cfs for the normal distribution to 69,825 cfs for the extreme value type I distribution using an n of 40 to determine the frequency factor. For many streams, the variation between methods is much greater than this (Benson 1968).

This example illustrates the difficulty of selecting a "best" frequency distribution for peak flows. All of the distributions used in this example fit the data about equally well.

The calculated skewness of the logarithms was 0.031. This stream gaging station is located in southeastern Kentucky near the line labeled 0 in figure 7.6 indicating good agreement between the calculated skewness of logarithms and the regional value. If the regional value of zero is used, the log Pearson III is the same as the lognormal and the estimated 100-year flow becomes 64,300 cfs.

REGIONAL FREQUENCY ANALYSIS

Methods of frequency analysis discussed up to this point have been concerned with data from a single location. The data from a single location is a random sample (usually of relatively few observations), from an unknown population. As indicated in the previous sections of this chapter, not only are the parameters of this population unknown, but the frequency distribution(s) that generated the sample are unknown as well. Recognizing the variability inherent in a random sample of this nature, many hydrologic frequency studies are done on a regional basis.

In a regional frequency analysis, attempts are made to define a region that is hydrologically homogeneous in terms of the characteristic being studied. Then hydrologic data from several locations within this region are combined into a single regional frequency analysis. There are many ways that regional studies can be made. A few of the more common ones are described here.

One of the first steps in a regional study is to define the region itself. Many methods have been used for doing this. Some are based on the location of watershed divides, political boundaries, land resource regions and physiographic regions. Regional boundaries are also defined in terms of similarity of flood frequency curves or flow duration curves. Dalrymple (1960) discusses a test for determining if flood frequency curves in a region can be considered as homogeneous. DeCoursey (1973) discusses the use of discriminant analysis and canonical correlation for regionalization of hydrologic data.

One common method of regional frequency analysis postulates that a particular analytical frequency function is applicable. The parameters of this distribution are then related to physiographic and meteorologic factors within the region. By measuring these factors at some point in the region, the parameters for the frequency distribution can then be estimated and probabilities calculated.

Haan and Read (1970) studied the annual water yield from small watersheds in Kentucky. They found that the frequency distribution of annual water yield in Kentucky could be described by the normal distribution. The mean annual water yield was then related to annual rainfall and watershed area, perimeter and relief ratio by multiple regression. The standard deviation of annual water yield was found to be relatively constant throughout the state. Thus the frequency distribution of annual water yield for an ungaged basin could be estimated by first estimating the mean of the normal distribution from the multiple regression equation and using the constant standard deviation. Since the mean and standard deviation completely define the normal distribution, water yield for various fequencies could then be determined.

A method of regional frequency analysis that is widely used for flood frequencies is described in detail by Dalrymple (1960). The method consists of computing a base flood frequency relationship in terms of the return period and the ratio of the peak flow for a given return period to an index flood (usually the mean annual flood) for several stations in the region. The median value of this ratio is then plotted versus the return period. Figure 7.10 is such a plot based on data for 18 stations in Alberta and Saskatchewan, Canada, as reported by Durant and Blackwell (1959).

The next step is to relate the index flood to watershed characteristics. The area of the watershed is generally used along with other geomorphic and meteorologic factors. Again this step is usually accomplished by using multiple regression.

The flood frequency curve for an ungaged location can now be estimated by first estimating the magnitude of the index flood and then multiplying by the ratio read from the base frequency curve for various return periods.

A third method of regional frequency analysis is to determine the magnitude, Q_T, of a hydrologic variable, say flood peaks, for various return periods, T, at several points in a region. Q_T is then related to basin characteristics. A different relationship will result for every value of T. Benson (1962b) found that for flood frequencies in New England a relationship of the form

$$Q_1 = aA^b S^c St^d I^e t^f 0^g$$

results. In this expression Q_1 is the 1-year annual peak discharge; A is the drainage area; S is the main channel slope; St is the percent of surface storage area plus 0.5 percent; I is the T-year, 24-hour rainfall intensity; t is the average January degrees below freezing; and 0 is an orographic factor. a, b, c, d, e, f and g are regression coefficients.

In a regional analysis all of the events from the various stations cannot be considered as independent. For example Benson (1962c) reports that in a regional flood-frequency study in New England (Benson 1962b) records from 164 basins were analyzed. Five or six major storms were responsible for the 3 or 4 highest flood peaks at most of

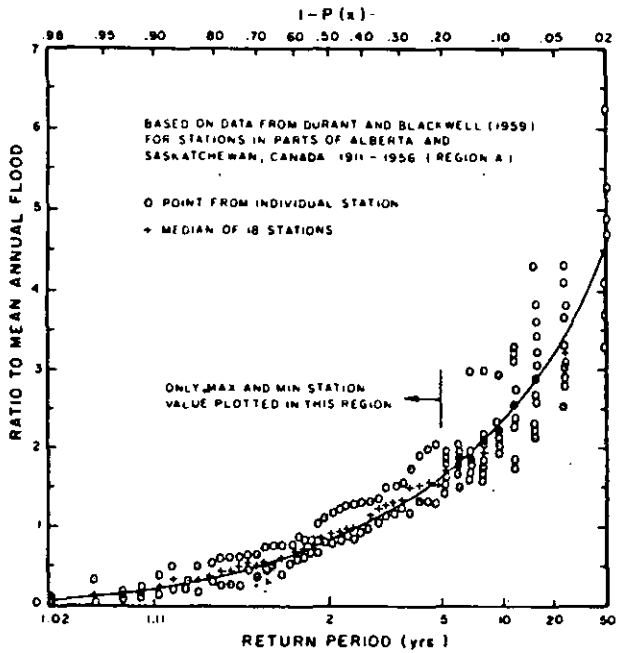


Fig. 7.10. Example relationship between Q_1/Q_{2-3} and the return period.

the 164 stations. If a single storm produced a major peak at each of the 164 stations, this cannot be considered as 164 independent events but more likely 164 measurements of the same event.

Alexander (1954) states that the number of independent stations, n', required to give the same variance as n stations having an average correlation of $\bar{r}$ is given by $n' = n/[1 + \bar{r}(n-1)]$. Thus if the $\bar{r}$ is 0.30 and n is 164, n' is only about 3.3. This means 3.3 independent stations would provide as much information (in terms of the variance) as 164 stations having an average correlation coefficient of 0.30. In the limit as n gets large, the number of equivalent independent stations approaches $1/\bar{r}$. Correlation is discussed more fully in Chapter 11.

FREQUENCY ANALYSIS OF PRECIPITATION DATA

The amount of rainfall (depth) that can be expected to occur in a given period of time (duration) on the average once every so many years (frequency) is an important design variable for many hydraulic structures. Depth-duration-frequency relationships have been developed for the United States (Hershfield 1961) for durations of 30 minutes to 24 hours and return periods of 1 to 100 years and published as U. S. Weather Bureau TP 40.

The procedure used in developing TP 40 (Hershfield 1961) was to prepare four key base maps showing the 2-year, 1-hour; 2-year, 24-hour; 100-year, 1-hour and 100-year, 24-hour rainfalls for the United States. Annual series data were used consisting of the

maximum 60-minute and 24-hour rainfall depths converted to a partial duration series by using the factors shown in table 7.9. For example if the 5-year partial duration series value estimated from the maps is 2.00 inches the corresponding annual series depth would be 0.96(2.00) or 1.92 inches. For return periods greater than 10 years, the conversion factor is essentially unity.

The 2-year rainfall amounts were determined by plotting on log-log paper the return period versus the rainfall depth using the California plotting position formula (equation 7.6), drawing a smooth curve through the points, and reading the 2-year value (figure 7.11).

The 100-year rainfall amounts were determined by using the type I Extreme Value distribution for selected stations with long rainfall records. The ratio of the 100-year to the 2-year rainfall amount was then determined for these stations and a map prepared showing the value of this ratio. The 100-year rainfall amounts for the stations with short

Table 7.9. Empirical factors for converting partial duration series to annual series (Hershfield 1961).

Return period	Conversion factor
2	.88
5	.96
10	.99

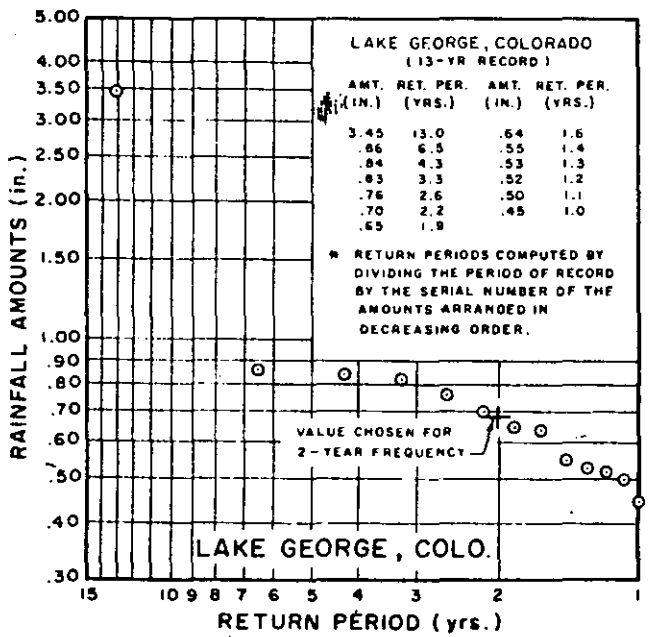


Fig. 7.11. Determination of 2-year 1-hour rainfall depth (U.S. Weather Bureau TP 24, Rainfall Intensities for local Drainage Design in the United States. U.S. Department of Commerce, Washington, D. C.).

records was estimated by the 100-year to 2-year ratio.

The rainfall depths for other return periods were determined by plotting the 2-year and 100-year depths on special paper (figure 7.12), connecting the points by a straight line, and reading off the desired rainfall depths. The spacing of the return periods along the abscissa of this special paper was empirical from 1 to 10 years based on free-hand plotting of partial duration series data and theoretical according to the type I extreme value distribution from 20 to 100 years. The transition between 10 and 20 years is smoothed by hand from the type I values.

The rainfall depths for durations other than 1 hour or 24 hours were obtained by plotting the 1-hour and 24-hour values on a second special paper (figure 7.13) and connecting the points with a straight line. This diagram was obtained empirically from an analysis of records from 200 first-order U. S. Weather Bureau stations. The depth of rainfall for the 30-minute duration is obtained by multiplying the 1-hour value by 0.79.

From these analyses, curves called depth (or intensity)-duration-frequency curves can be prepared. Data from TP 40 were used to construct figure 7.14. As more rainfall data becomes available it will be possible to perform frequency analyses for various durations directly without having to resort to diagrams such as figures 7.12 and 7.13.

Rainfall data for longer durations such as weeks or months can be analyzed by using the gamma distribution. Barger and Thom (1949) have shown the gamma distribution applicable to rainfall data. Barger, Shaw and Dale (1959), Friedman and Janes (1957), Strommen and Horsfield (1969) and Mooley and Crutcher (1968) are among those who have used the gamma distribution for rainfall.

By using equation 7.15, it can be seen that the probability of a rainfall R exceeding X is given by

$$\text{prob}\,(R>x) = k[1 - P_R{}^*(x)] \qquad (7.16)$$

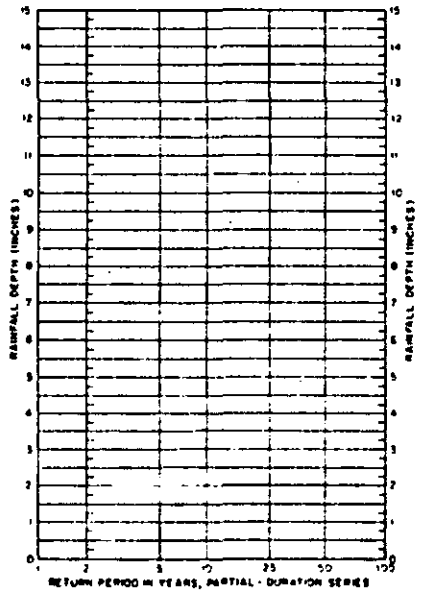


Fig. 7.12. Rainfall depth versus return period.

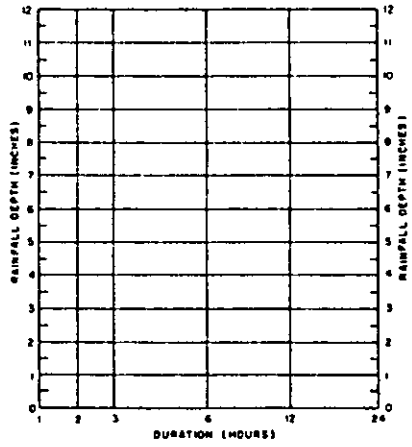


Fig. 7.13. Rainfall depth-duration diagram.

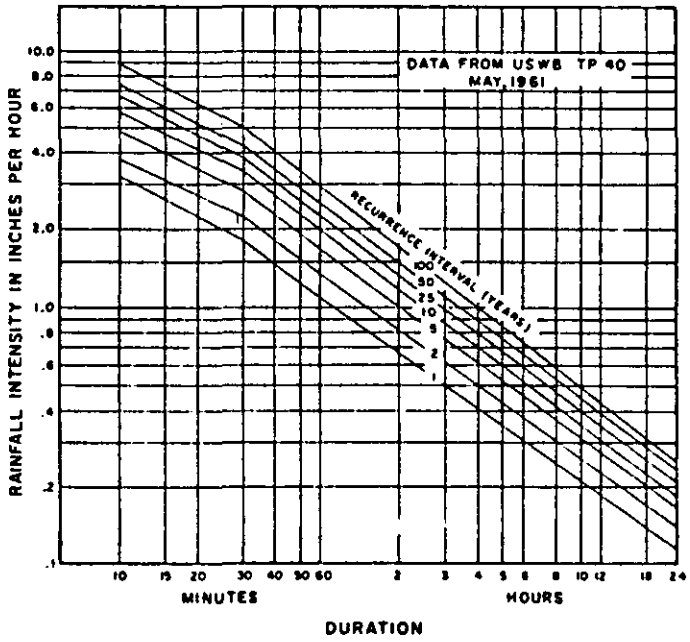


Fig. 7.14. Rainfall intensity-duration-frequency relationship Lexi[illegible]n, Kentucky.

and the probability of R being less than x is given by

$$\text{prob}\,(R<x) = 1 - \text{prob}\,(R>x)$$
$$= 1 - k + kP_R^*(x) \qquad (7.17)$$

where k is the probability of rain or the proportion of time intervals with rainfall and $P_R^*(x)$ is the cumulative probability distribution of rain given that $R \neq 0$. Many times the gamma distribution is used for rainfall data. The parameters of the gamma distribution generally are determined by using equations 6.19 and 6.20. Bridges and Haan (1972) have presented a technique for determining the reliability of rainfall estimates from the gamma distribution based on simulation studies.

FREQUENCY ANALYSES OF OTHER HYDROLOGIC VARIABLES

The principles set forth on flood frequencies and rainfall frequencies also apply to frequencies of other hydrologic variables. Basically the quantity to be analyzed must be defined, the data tabulated and then a frequency analysis made. For instance in the case of flow volume-frequency studies, the duration(s) of interest must be specified and then the maximum or minimum flow volumes for each year having the specified duration are tabulated. The maximum flow volumes would be used in the case of flood-flow volumes and the minimum volumes would be used in the case of low-flow studies.

Frequency analyses can be applied on water quality parameters such as dissolved oxygen, biological oxygen demand, sediment loads and many other quantities. Care must be taken to see that the data used meet the necessary requirements of homogeneity, independence and representativeness. For example, if sediment concentration frequencies are being studied and part of the data are collected during low flows and part during high flows, the data may not be homogeneous because of the relationship between sediment concentration and flow rate.

Exercises

7.1 Assume that daily rainfall on rainy days follows an exponential distribution. The average daily rainfall on rainy days is 0.3 inches. If 30% of all days are rainy, what is the probability that on some future day, the amount of rainfall received will exceed 1.00 inch? Assume daily rainfalls are independent.

7.2 Derive a table of frequency factors for the exponential distribution corresponding to T = 2, 5, 10, 20, 50 and 100 years.

7.3 Select several streams in a single locality and prepare a plot of the ratio of the T-year flood to the mean annual flood (as in figure 7.10).

7.4 An analysis of 50 years of data showed that the probability of a flood peak exceeding 90,000 cfs on a certain river was .02. During a 10-year period 2 such peaks occurred. If the original estimate of the probability of this exceedance was correct, what is the probability of getting 2 such exceedances in 10 years?

7.5 Forty years of peak streamflow data are available. All but one of the data points indicate that a lognormal distribution with $\bar{X}$ = 125,000 cfs and s_X = 50,000 describes the data very nicely. The one outlier is equal to 285,000 cfs. What is the probability that an event of 285,000 cfs or greater could occur in the 40-year period if the flood peaks truly follow the lognormal distribution with $\bar{X}$ and s_X as given?

7.6 Select a set of data consisting of 20 or more independent observations. Plot these data on normal probability paper using all three of the plotting position relationships contained in table 7.4.

7.7 Compute the 100-year peak flow for the annual series data of example 7.2 assuming the data follow the gamma distribution.

7.8 Prepare a plot on log-log paper of low flow frequency-volume-duration for Cave Creek near Fort Spring, Kentucky. Plot volume in inches as the ordinate, duration in months (use 1, 2, 3, 6 and 12 months) as the abscissa and use as curve parameters frequency (use 2, 5, 10 and 25 years).

7.9 Work exercise 7.8 for maximum flow frequency-volume-duration on Cave Creek.

7.10 Plot the annual runoff data for Walnut Gulch near Tombstone, Arizona, on normal and lognormal probability paper. Does either of these distributions appear to "fit" the data?

7.11 Plot on normal probability paper the annual runoff data for (a) Piscataquis River near Dover-Foxcroft, Maine (b) North Llano River near Junction, Texas and (c) Spray River, Banff, Canada. Is there any apparent relationship between the curvature (or lack of it) and the skewness?

7.12 Work exercise 7.11 only plot the data on lognormal probability paper.

7.13 For the Piscataquis River near Dover-Foxcroft, Maine, estimate the 100-year annual flow assuming the data follow the (a) normal distribution (b) lognormal distribution (c) Pearson type III distribution (d) log Pearson type III distribution (e) extreme value distribution.

7.14 Work exercise 7.13 for the 100-year annual flow on the North Llano River near Junction, Texas.

7.15 Work exercise 7.13 for the 100-year annual flow on the Spray River, Banff, Canada.

7.16 In reference to exercises 7.13, 7.14 and 7.15, which distribution would you expect to give the "best" estimate for the 100-year flow on each of the three rivers. Discuss in terms of the means, variances, coefficient of variation and skewness.

7.17 Plot the annual peak discharge of Walnut Gulch near Tombstone, Arizona, on lognormal probability paper. Draw in what you consider the best fitting straight line. Estimate the mean and variance of the data from this plot.

7.18 Plot the suspended sediment load data for the Green River at Munfordville,

Kentucky on normal and lognormal probability paper. Draw in the best fitting straight line.

7.19 Use the lognormal distribution to estimate the 25-year runoff volume for July on Walnut Gulch near Tombstone, Arizona. Plot the data on lognormal probability paper and draw in the theoretical best fitting straight line.

CURSO: INGENIERIA MARITIMA. MODULO: "FUNDAMENTOS DE HIDRAULICA Y MATEMATICAS" DEL 11 DE MARZO AL 7 DE JUNIO.
MEXICO, D.F.

CONFIDENCE INTERVALS AND HYPOTHESIS TESTING.

PROF. ING. JOSE RAYNAL.
JUNIO 1985.

8. Confidence Intervals and Hypothesis Testing

IN CHAPTER 3 parameter estimation was discussed in general terms. In chapters 4, 5 and 6 specific methods for estimating the parameters of certain probability distributions were discussed. Again it should be recalled that parameter estimates are called statistics, are functions of the sample (random) values, and are themselves random variables. Parameter estimates have associated with them probability distributions.

Thus far we have discussed methods of getting point estimates for parameters and certain properties of these point estimates. The possible errors in these point estimates due to inherent variability in random samples of data have not been discussed. The first part of this chapter is devoted to the reliability of parameter estimates. The second part of the chapter is devoted to testing hypotheses regarding population parameters.

CONFIDENCE INTERVALS

A parameter θ is estimated by $\hat{\theta}$. The statistic $\hat{\theta}$ is a random variable having a probability distribution. If $\hat{\theta}$ can take on any value in some continuous range, then prob $(\hat{\theta} = \theta)$ is zero. Rather than a point estimate for θ, it may be more desirable to get an interval estimate such that the probability that this interval contains θ can be specified. Such an interval is known as a confidence interval. This statement may be written

$$\text{prob}(L < \theta < U) = 1 - \alpha \tag{8.1}$$

where L and U are the lower and upper confidence limits so that the interval from L to U is the confidence interval and $1 - \alpha$ is the confidence level or confidence coefficient. Note that in equation 8.1, θ is not a random variable. One does not say that the probability that θ is between L and U is $1 - \alpha$ but that the probability is $1 - \alpha$ that the interval L to U contains θ. The difference in these two interpretations is subtle but based on the fact that θ is a constant while L and U are random variables.

Mood et al. (1974) discuss a general method for determining confidence intervals.

Ostle (1963) presents expressions for the confidence intervals for many different statistics. In the discussion to follow, a procedure known as the method of pivotal quantities for determining confidence limits will be illustrated. This method consists of finding a random variable V that is a function of the parameter θ but whose distribution does not involve any other unknown parameters. Then v_1 and v_2 are determined such that

$$\text{prob}(v_1 < V < v_2) = 1 - \alpha \,. \tag{8.2}$$

This inequality is then manipulated so that it is in the form of equation 8.1 where U and L are random variables depending on V but not θ.

Mean of a Normal Distribution

As an example of using equation 8.2 the confidence intervals on the mean of a normal distribution will be determined. We have shown that the quantity $V = (\bar{X} - \mu)/S_{\bar{X}}$ has a t distribution with n - 1 degrees of freedom where n is the number of observations used to estimate $\bar{X}$. Using equation 8.2 we have

$$\text{prob}(v_1 < (\bar{X} - \mu)/S_{\bar{X}} < v_2) = 1 - \alpha \tag{8.3}$$

If it is desired that the confidence interval be symmetrical in probability, v_1 and v_2 can be chosen so that the probability that a random t is less than v_1 equals the probability that a random t exceeds v_2. Since the $100(1 - \alpha)$ percent confidence interval is being sought, both of these probabilities must be $\alpha/2$. The probability that the confidence intervals do not contain θ has been divided equally between the upper and lower bounds. In the following the notation $t_{\gamma,n}$ corresponds to the value of t such that the probability of a random t with n degrees of freedom being less than $t_{\gamma,n}$ is γ (see figure 8.1).

Equation 8.3 is equivalent to

$$\text{prob}(t_{\alpha/2,n-1} < (\bar{X} - \mu)/S_{\bar{X}} < t_{1-\alpha/2,n-1}) = 1 - \alpha$$

Since the t distribution is symmetrical, $t_{\alpha/2,n-1} = -t_{1-\alpha/2,n-1}$. Therefore

$$\text{prob}(-t_{1-\alpha/2,n-1} < (\bar{X} - \mu)/S_{\bar{X}} < t_{1-\alpha/2,n-1}) = 1 - \alpha$$

or

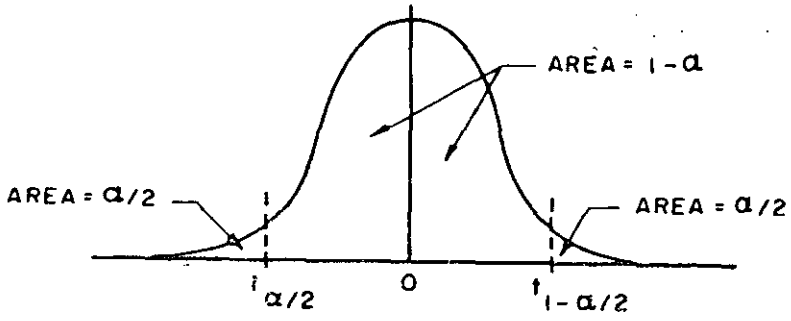


Fig. 8.1 Illustration of confidence intervals using the t distribution.

$$\text{prob}(\bar{X} - t_{1-\alpha/2,n-1} S_{\bar{X}} < \mu < \bar{X} + t_{1-\alpha/2,n-1} S_{\bar{X}}) = 1 - \alpha$$

This latter equation is in the form of equation 8.1 so the confidence limits are

$$\begin{aligned} l &= \bar{x} - t_{1-\alpha/2,n-1} s_{\bar{X}} \\ u &= \bar{x} + t_{1-\alpha/2,n-1} s_{\bar{X}} \end{aligned} \tag{8.4}$$

Since $\bar{X}$ and $S_{\bar{X}}$ are both random variables, L and U are random variables as well with estimates l and u given by equation 8.4.

Example 8.1. The sample mean and variance of the Kentucky River data contained in table 2.1 have been calculated as $\bar{X} = 67{,}500$ and $S_{\bar{X}} = 21{,}000$. What are the 95% confidence limits on the mean assuming the sample is from a normal population?

Solution:

$$s_{\bar{X}} = s_X/\sqrt{n} = 21{,}000/\sqrt{66} = 2{,}585$$

From table E.5

$$t_{1-\alpha/2,n-1} = t_{0.975,65} = 2.00.$$

From equation 8.4

$$l = \bar{x} - t_{1-\alpha/2,n-1} s_{\bar{X}} = 67{,}500 - 2.00(2{,}585) = 62{,}330$$

$$u = \bar{x} + t_{1-\alpha/2,n-1} s_{\bar{X}} = 67{,}500 + 2.00(2{,}585) = 72{,}670$$

Thus we can say that we are 95% confident that the interval 62,330 to 72,670 contains the true population mean.

Comment: If a 90% confidence interval is calculated, it is found to be 63,183 to 71,817. Thus the 90% confidence interval is shorter than the 95% confidence interval but our degree of confidence that the interval contains μ_X has decreased from 95% to 90%.

If a second independent sample of peak flows on the Kentucky River near Salvisa were available, this sample would have a different mean and variance. Thus the 95% confidence intervals would be different as well. If many samples were available and the 95% confidence limits were calculated for each, 95% of the confidence limits would contain the true population mean while 5% would not if the data were actually from a normal distribution.

The $100(1 - \alpha)\%$ confidence interval on the mean can be made as small as desired by increasing the sample size. This is because $S_{\bar{X}}$ decreases as the sample size is increased. An increase in the reliability of the sample mean comes at the expense of an increase in the sample size. Unfortunately in many hydrologic problems the sample size is fixed. For a normal distribution, equations 8.4 provides a means for determining the sample size required in order to estimate μ_X within a given reliability.

If the population variance of the normal distribution is known, then the pivotal

quantity in equation 8.3 becomes $(\bar{X} - \mu)/\sigma_{\bar{X}}$ which has a standard normal distribution. The confidence limits then become

$$l = \bar{x} - z_{1-\alpha/2}\,\sigma_{\bar{X}}$$

$$u = \bar{x} + z_{1-\alpha/2}\,\sigma_{\bar{X}} \qquad (8.5)$$

where $z_{1-\alpha/2}$ is the value of Z from the standard normal distribution such that the area to the right of Z is $\alpha/2$.

Equations 8.4 and 8.5 are based on the assumption that the underlying population of the random variable X has a normal distribution. Only through the Central Limit Theorem can these relations be applied to non-normal distributions. Confidence limits calculated by these relationships for the means of random samples from non-normal populations are only approximate with the approximation improving as the sample size increases. If these approximations are not satisfactory, other methods are available (Ostle 1963, Mood et al. 1974).

Variance of a Normal Distribution

The quantity $(n-1)S^2/\sigma^2$ has a chi-square distribution with n - 1 degrees of freedom. Letting this quantity equal V in equation 8.2 results in

$$\text{prob}(v_1 < (n-1)S_X^2/\sigma_X^2 < v_2) = 1 - \alpha$$

Choose v_1 equal to $\chi^2_{\alpha/2,n-1}$ and v_2 as $\chi^2_{1-\alpha/2,n-1}$. Then

$$\text{prob}(\chi^2_{\alpha/2,n-1} < (n-1)S_X^2/\sigma_X^2 < \chi^2_{1-\alpha/2,n-1}) = 1 - \alpha$$

or

$$\text{prob}\left((n-1)S_X^2/\chi^2_{1-\alpha/2,n-1} < \sigma_X^2 < (n-1)S_X^2/\chi^2_{\alpha/2,n-1}\right) = 1 - \alpha$$

which is in the form of equation 8.1. Thus the confidence limits on σ^2 are

$$l = (n-1)s_X^2/\chi^2_{1-\alpha/2,n-1}$$

$$u = (n-1)s_X^2/\chi^2_{\alpha/2,n-1} \qquad (8.6)$$

Again equations 8.6 are strictly valid only if X is from a normal distribution and approximate for X from a non-normal distribution with the approximation improving as the sample size increases.

The chi-square distribution is not symmetrical so that $s_X^2 - l$ is not equal to $u - s_X^2$. As the sample size and thus the degrees of freedom on the chi-square distribution increases, the distribution approaches a symmetrical distribution so that the upper and lower confidence limits are nearly the same distance from s_X^2. This is illustrated in figure 8.2.

Example 8.2. Determine the 90% confidence limits on the variance for the situation described in example 8.1.

Solution:

$$\chi^2_{\alpha/2,n-1} = \chi^2_{0.05,65} = 47.4$$

$$\chi^2_{1-\alpha/2,n-1} = \chi^2_{0.95,65} = 84.8$$

$$l = (n-1)s_X^2/\chi^2_{1-\alpha/2,n-1} = 65(21{,}000)^2/84.8 = 338.0 \times 10^6$$

$$u = (n-1)s_X^2/\chi^2_{\alpha/2,n-1} = 65(21{,}000)^2/47.4 = 604.7 \times 10^6$$

The 90% confidence intervals on the standard deviation are found by taking the square roots of the above limits to be 18,400 to 24,590 cfs.

Comment: In the preceding two examples the confidence limits on the mean and variance of a normal distribution were calculated. If the joint confidence limits on $\bar{X}$ and S_X^2 are desired, they cannot be computed separately as was done in these examples. Mood et al. (1974) discuss the estimation of joint confidence intervals.

One-Sided Confidence Intervals

Situations may arise where one is only interested in an interval estimate on one side of a parameter. For instance it may be desired to find only a lower confidence limit. In this situation equation 8.1 becomes

$$\text{prob}(L < \theta) = 1 - \alpha \qquad (8.7)$$

The same procedure for finding L would be followed as was used in the two-sided case except now all of the probability α will be in one tail. For instance the one-sided lower limit on the mean of a normal distribution with an unknown variance would be

$$l = \bar{x} - t_{1-\alpha,n-1}\,s_{\bar{X}} \qquad (8.8)$$

The analogous results would hold for any one-sided, lower or upper confidence limit.

Parameters of Probability Distributions

For a wide class of distributions for large samples the maximum likelihood estimators for the parameters of the distribution are asymptotically normally distributed with the true parameter value as the mean and a variance of $\{n\,E[\frac{\partial}{\partial\theta}\ln p_X(x,\theta)]^2\}^{-1}$.

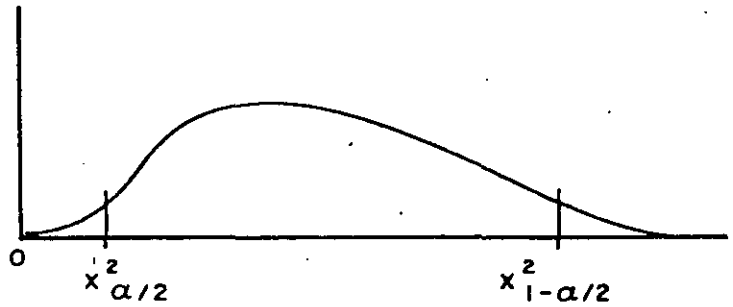


Fig. 8.2. Illustration of confidence limits on a chi-square distribution.

Using this information it is possible to construct confidence intervals and joint confidence intervals for the parameters of these distributions. The book by Mood et al. (1974) should be consulted for the procedures to be used.

HYPOTHESIS TESTING

Many times the acceptability of statistical models can be judged without actually making any statistical tests. This would be the case when observed data is predicted very closely by the model or when observed data deviates very greatly from the model. On the other hand a common occurrence is for the observed data to deviate some from the model but not enough for one to state that the model is obviously inadequate. In this latter situation one must determine whether the deviations represent true inadequacies in the model or whether the deviations are chance variations from the true model.

The general procedure to be followed in making statistical tests is

1. Formulate the hypothesis to be tested.
2. Formulate an alternative hypothesis.
3. Determine a test statistic.
4. Determine the distribution of the test statistic.
5. Define the rejection region or critical region of the test statistic.
6. Collect the data needed to calculate the test statistic.
7. Determine if the calculated value of the test statistic falls in the rejection region of the distribution of the test statistic.

For many statistical tests, steps 2-4 have been completed and may be found in a wide variety of statistics books. For many of the tests that a hydrologist might like to make, adequate test statistics and their distributions have not been determined.

It is not possible to develop tests that are absolutely conclusive. All of the tests have a possiblity of two kinds of error – rejecting a true hypothesis or accepting a false hypothesis. These errors are called Type I and Type II errors, respectively. Table 8.1 depicts the two types of errors. The probability of a Type I error is denoted by α and the probability of a Type II error by β. The significance level is defined as 100α (in percent). In testing hypotheses, the probability of a Type I error can be specified; however, the probability of a Type II error is not known unless the true parameter values being tested are known. In general as the value of α decreases, the magnitude of β increases.

As an example, assume we select an observation x_0 at random from a normal distribution with variance σ_0^2 and hypothesize that the distribution has a mean μ_0. The test statistic could be x_0 itself which has a normal distribution with unknown mean and variance σ_0^2. If the hypothesis is true (something that is not known or the

Table 8.1. Errors in hypothesis testing.

decision	true situation	
	hypothesis true	hypothesis false
accept hypothesis	No error	Type II error
reject hypothesis	Type I error	No error

test would not be made), the distribution of the test statistic would be a normal distribution with mean μ_0 and variance σ_0^2 and would appear as in Figure 8.3. If it is decided to accept the hypothesis if x_0 is within 2 standard deviations of μ_0 and reject the hypothesis otherwise, the critical region or rejection region would be the shaded area in Figure 8.3. From the properties of the normal distribution, it is known that 95.44 percent of the area of the normal curve is within 2 standard deviations of the mean so the critical region occupies 4.56 percent of the area. It is also apparent that there is a 4.56 percent chance that x_0 will be in the critical region and the hypothesis rejected even though it is true. Thus by definition $\alpha = 0.0456$ or there is a 4.56% chance of making a Type I error due to random variation in the x_0 selected. It is more common to specify α and from this information determine the critical region. For example if one wanted α to be 0.10 then the critical region would be $|(x_0 - \mu_0)/\sigma_0| > 1.645$ which is the value of the standard normal distribution such that the area outside the limits -1.645 to 1.645 is 0.10.

In order to evaluate β, the true parameter values must be known. Again consider selecting a single value x_0 from a normal population with variance σ_0^2 and an unknown mean. Let the hypothesis be that $\mu = \mu_0$ and the alternative be $\mu \neq \mu_0$. If μ actually equals μ_1, then the situation depicted in figure 8.4 would exist and there is a 100β percent chance that x_0 will fall in the acceptance region of $N(\mu_0, \sigma_0^2)$ and thus a Type II error committed. From figure 8.4 it can be seen that as α is increased, β will decrease. It can also be seen that the nearer μ_1 is to μ_0, the greater will be β. This is because it

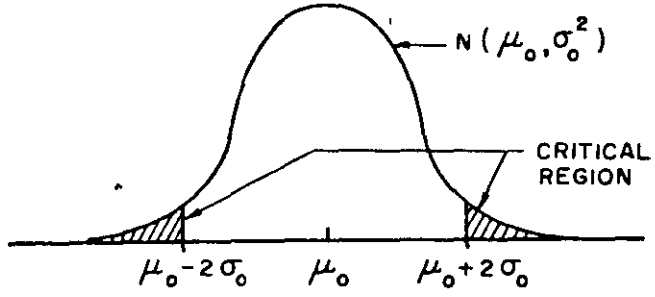


Fig. 8.3. Critical Region.

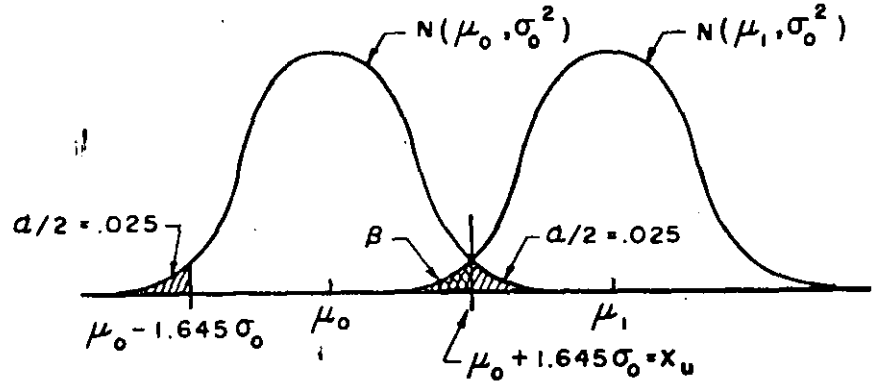


Fig. 8.4. Illustration of α and β.

is increasingly difficult to tell the difference between the two distributions. It is not possible to determine the magnitude of β since it is a function of the unknown population mean μ_1. Example 8.3 shows how β can be evaluated if μ_1 is known. Of course, μ_1 would not be known or else one would not hypothesize $\mu = \mu_0$.

Example 8.3. Assume a single observation is selected from a normal distribution with mean $\mu_1 = 7$ and variance $\sigma_0^2 = 9$. It is hypothesized that $\mu = \mu_0 = 5$. If the test is conducted at the 10 percent significance level, what is β?

Solution:

Reference should be made to figure 8.5.

$\alpha = 0.10$

$\alpha/2 = 0.05$ which corresponds to $z_{1-\alpha/2} = 1.645$ from the standard normal distribution

$(X_u - \mu_0)/\sigma_0 = Z_u$ where X_u is the boundary of the upper critical-region

$(x_u - 5)/3 = 1.645$

$x_u = 9.935$

A_u = the area of a normal distribution with mean of 7 and variance of 9 to the left of 9.935.

The standardized variate corresponding to $x_u = 9.935$ is

$z_u = (9.935 - 7)/3 = 0.978$

The area to the left of $z_u = 0.978$ from a standard normal distribution is 0.8365.

Similarly if X_l is the boundary of the lower critical region, we have $(x_l - 5)/3 = -1.645$ or $x_l = 0.0645$. A_l is the area of a normal distribution with mean 7 and variance 9 to the left of 0.0645. $z_l = (0.0645 - 7)/3$ or $z_l = -2.31$. $A_l = 0.0104$. Now $\beta = A_u - A_l$ or $\beta = 0.8365 - 0.0104 = 0.8261$. Thus the probability of accepting the hypothesis that $\mu = 5$ when in fact $\mu = 7$ is 0.8261 when α is 0.10. The probability of a type II error is 0.8261.

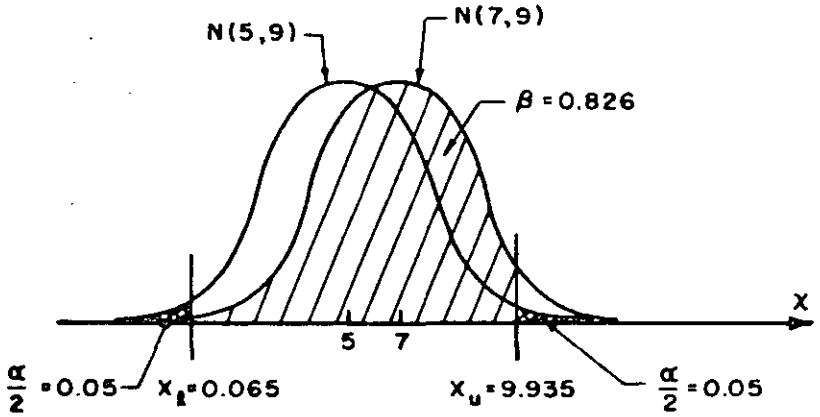


Fig. 8.5. Illustration for example 8.3.

If calculations such as contained in example 8.3 are carried out for various values of μ_1, a curve relating β to μ_1 can be constructed. Such a curve is shown in figure 8.6. Figure 8.6 shows the β curve for $\alpha = 0.05$ and $\alpha = 0.10$. Curves such as shown in figure 8.6 are often called operating characteristic (OC) curves.

Figure 8.6 verifies the earlier statements that β increases as α decreases and β increases as the true mean, μ_1, approaches the hypothesized mean, μ_0. In fact as μ_1 gets close to μ_0, the probability of accepting $\mu = \mu_0$ when $\mu = \mu_1$ is true gets very large. This may not be a serious problem in practice since we may not care for instance whether μ is 5 or 5.5.

The quantity $1 - \beta$ is called the power of a test. Ideally we would like the power to be large for all values of μ_1. In fact in testing a hypothesis, we would like α to be small and the power to be large. Figure 8.7 shows that power of a test is a function of α and true parameter values. The power of a test is also a function of the test itself. For instance we could have chosen as our test statistic $X_0 + 3$ and then rejected the hypothesis if $x_0 + 3$ fell in the critical region. Figure 8.7 compares the power of this test with the test that rejected the hypothesis if x_0 fell in the critical region.

Figure 8.7 shows that for certain values of μ_1, the $X_0 + 3$ test is more powerful than the X_0 test. Ideally we would like to use the test that was the most powerful over the entire range of the unknown parameter. Such a test is known as a uniformly most powerful test. Unfortunately uniformly most powerful tests do not exist in many situations.

Selecting which test to use comes down to the purposes of the test and the consequences of making an error. In our example, if accepting the hypothesis $\mu = 5$ when in fact $\mu > 5$ is a very serious error whereas accepting it if $\mu < 5$ is of little consequence, we might prefer the $X_0 + 3$ test since it is more powerful in the region $\mu > 5$. If the consequence of an error depended only on the magnitude of the error, the X_0 test might be preferred.

From the above discussion, it should be apparent that the selection of α and the

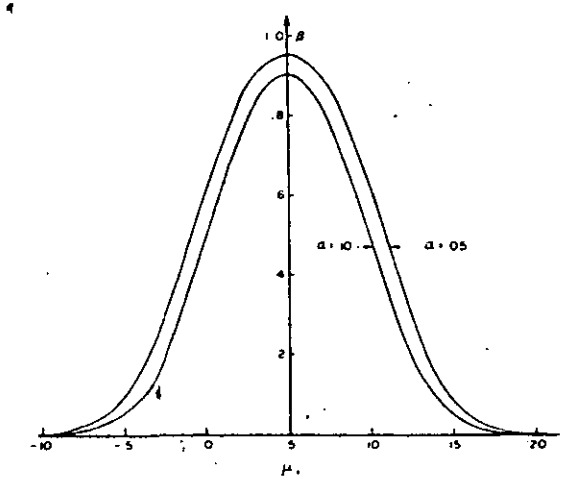


Fig. 8.6. Probability of a Type II error as a function of the true mean, μ_1, for the situation described in example 8.3.

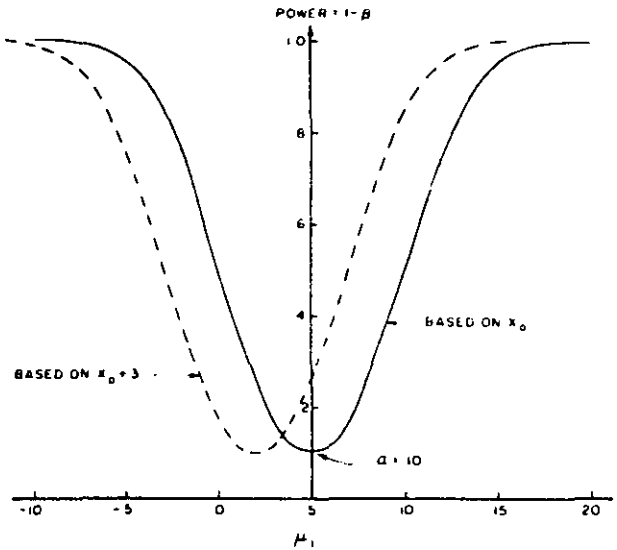


Fig. 8.7. Example of a power curve.

type of test to be used depends on the problem at hand. Mood et al. (1974) discuss these concepts in more theoretical terms. The level of significance, α, is usually chosen to be 0.10, 0.05 or 0.01. In theory α should be based on the problem at hand. In practice α is generally arbitrarily selected.

Many tests of hypothesis are of the type $\theta = \theta_1$ versus the alternative $\theta \neq \theta_1$. Accepting such a hypothesis as true does not mean that one strictly feels that $\theta = \theta_1$ but rather that θ is not significantly different from θ_1. For example, if we calculate the mean of a random sample and then accept the hypothesis that the true mean is 5, we may not believe that the true mean is exactly 5 but rather the true mean is not significantly different from 5. What constitutes a significant difference has been defined by the type of test used and the level of significance. Furthermore a statistically significant difference and a physically significant difference are not the same. For example if $\hat{\theta} = 4.0$ is an estimate for θ and a test of hypothesis shows $\hat{\theta}$ is not significantly different from zero, it does not mean $\theta = 0$ should be used in some physical analysis if this physical analysis is sensitive to differences in θ of this order of magnitude. A physically significant difference depends on the problem being studied.

The following is a discussion of several common tests of hypotheses. The hypothesis to be tested is denoted by H_o and the alternative hypothesis by H_a. For the tests that follow to be correct statistical tests, the assumptions involved in developing the test statistic must not be violated. In practice at least some of the assumptions are generally violated with the result that the tests are only approximate tests. This approximation is manifest in the fact that the actual level of significance will not be equal to $100\alpha\%$.

H_o: $\mu = \mu_1$, H_a: $\mu = \mu_2$, normal distribution, known variance

In this case H_o is a simple hypothesis and H_a is a simple alternative hypothesis.

The test statistic is developed by considering that

$$Z = \sqrt{n}\,(\bar{X} - \mu_1)/\sigma_X$$

has a standard normal distribution. If $\mu_1 > \mu_2$ then H_o is rejected if

$$\bar{x} \leq \mu_1 - z_{1-\alpha}\,\sigma_X/\sqrt{n} \tag{8.9}$$

If $\mu_1 < \mu_2$ then H_o is rejected if

$$\bar{x} \geq \mu_1 + z_{1-\alpha}\,\sigma_X/\sqrt{n} \tag{8.10}$$

In the preceding expressions, $z_{1-\alpha}$ represents the point on the standard normal distribution such that $\text{prob}(Z \geq z_{1-\alpha}) = \alpha$.

H_o: $\mu = \mu_1$, H_a: $\mu = \mu_2$, normal distribution, unknown variance

The test statistic for this situation is

$$T = \sqrt{n}\,(\bar{X} - \mu_1)/S_X$$

Then H_o is rejected if

$$\bar{x} \leq \mu_1 - t_{1-\alpha,n-1}\,s_X/\sqrt{n} \quad \text{for } \mu_1 > \mu_2 \tag{8.11}$$

and

$$\bar{x} \geq \mu_1 + t_{1-\alpha,n-1}\,s_X/\sqrt{n} \quad \text{for } \mu_1 < \mu_2 \tag{8.12}$$

H_o: $\mu = \mu_0$, H_a: $\mu \neq \mu_0$, normal distribution, known variance

This hypothesis is a simple hypothesis with a compound alternative hypothesis. Again the test statistic is

$$Z = \sqrt{n}\,(\bar{X} - \mu_o)/\sigma_X$$

H_o is rejected if

$$|z| = |\sqrt{n}\,(\bar{x} - \mu_0)/\sigma_X| > z_{1-\alpha/2} \tag{8.13}$$

H_o: $\mu = \mu_0$, H_a: $\mu \neq \mu_0$, normal distribution, unknown variance

Generally a population variance is not known and must be estimated. In that case H_o is tested by using

$$T = \sqrt{n}\,(\bar{X} - \mu_0)/S_X$$

H_o is rejected if

$$|t| = |\sqrt{n}\,(\bar{x} - \mu_0)/s_X| > t_{1-\alpha/2,n-1} \tag{8.14}$$

This test cannot be applied to every set of data. The assumption has been made that the observations are from a normal distribution.

Example 8.4. Appendix B lists the annual runoff for Cave Creek near Fort Spring, Kentucky, for the period 1953 to 1970. Test the hypothesis that the mean annual runoff is 16.5 inches.

Solution: The testing procedures we have available to us all are based on the assumption of normality. If we assume the annual runoff is normally distributed, we can use equation 8.14 to test H_o: $\mu = 16.5$ versus H_a: $\mu \neq 16.5$.

For this data $\bar{X}$ is 14.65 inches and S_X is 4.75 inches. There are 18 observations. The test statistic is

$$t = \sqrt{n}\,(\bar{x} - \mu_0)/s_X = \sqrt{18}\,(14.65 - 16.5)/4.75 = -1.65$$

Using a level of significance of $\alpha = 0.05$ we find $t_{1-\alpha/2,n-1} = t_{0.975,17} = 2.11$. Since $|t| = 1.65 < 2.11$, we accept the hypothesis that the mean is 16.5.

Comment: Some statisticians do not like to "accept" H_o. Their reasoning is we have not proven H_o, only found strong evidence to support it. As a result of a statistical test, their conclusions would be either reject H_o or fail to reject H_o. In this book we will consider failing to reject H_o equivalent to accepting H_o. It should be kept in mind however that we have not proven H_o.

For instance, in this example, we have calculated the sample mean to be 14.65 and accepted the hypothesis that the population mean is 16.5. This illustrates two points. First, the data and the test obviously do not prove that $\mu = 16.5$. Second, what we really have accepted is not that the mean is 16.5 but that when sampling from this distribution using a sample of size 18, the difference between the sample mean of 14.65 and the hypothesized mean of 16.5 can reasonably be ascribed to chance variations due to the random sample. Our conclusion is that based on this sample, we cannot say that the population mean is not 16.5 or based on this sample the population mean is not significantly different from 16.5.

Test for Differences in Means of Two Normal Distributions

If the variances of the two normal distributions are known, then the H_o: $\mu_1 - \mu_2 = \delta$ versus H_a: $\mu_1 - \mu_2 \neq \delta$ can be tested by calculating the test statistic

$$Z = (\bar{X}_1 - \bar{X}_2 - \delta)/(\sigma_1^2/n_1 + \sigma_2^2/n_2)^{1/2} \qquad (8.15)$$

In this case, Z has a standard normal distribution so the rejection region is $|z| > z_{1-\alpha/2}$.

If the variance of the two normal distributions are equal but unknown, the H_o: $\mu_1 - \mu_2 = \delta$ versus H_a: $\mu_1 - \mu_2 \neq \delta$ is tested by calculating the statistic

$$T = (\bar{X}_1 - \bar{X}_2 - \delta)/\{[(n_1+n_2)][(n_1-1)S_1^2 + (n_2-1)S_2^2]/[n_1 n_2 (n_1+n_2-2)]\}^{1/2} \qquad (8.16)$$

which has a t distribution with $n_1 + n_2 - 2$ degrees of freedom. Thus H_o is rejected if $|t| > t_{1-\alpha/2,n_1+n_2-2}$.

Again note that these two tests are based on sample normality. For large samples, the Central Limit Theorem may enable us to use these tests as approximate tests for non-

normal samples.

Gibra (1973), Ostle (1963) and others discuss testing the H_o: $\mu_1 - \mu_2 = \delta$ versus H_a: $\mu_1 - \mu_2 \neq \delta$ when sampling from two normal populations with unknown and unequal variances. Ostle recommends the following approximate procedure. Compute the test statistic

$$T' = (\bar{X}_1 - \bar{X}_2 - \delta)/(S_1^2/n_1 + S_2^2/n_2)^{1/2} \qquad (8.17)$$

The hypothesis is rejected if

$$|t'| > (w_1 t_1 + w_2 t_2)/(w_1 + w_2)$$

where

$$w_1 = s_1^2/n_1$$

$$w_2 = s_2^2/n_2$$

$$t_1 = t_{1-\alpha/2,n_1-1}$$

$$t_2 = t_{1-\alpha/2,n_2-1}$$

Test of H_o: $\sigma^2 = \sigma_o^2$ Versus H_a: $\sigma^2 \neq \sigma_o^2$ Normal Population

A test of H_o: $\sigma^2 = \sigma_o^2$ versus H_a: $\sigma^2 \neq \sigma_o^2$ when sampling from a normal distribution with sample size n can be made by calculating the test stastic

$$\chi_c^2 = \Sigma_{i=1}^{n} (X_i - \bar{X})^2/\sigma_o^2$$

and then accepting H_o if

$$\chi^2_{\alpha/2,n-1} < \chi_c^2 < \chi^2_{1-\alpha/2,n-1} \qquad (8.18)$$

Otherwise H_o is rejected.

Test of H_o: $\sigma_1^2 = \sigma_2^2$ Versus H_a: $\sigma_1^2 \neq \sigma_2^2$ for Two Normal Populations

To test the hypothesis that the sample variances of two normal populations are equal, the sample test statistic is

$$F_c = S_1^2/S_2^2$$

where S_1^2 is the larger sample variance. F is distributed as an F distribution with $n_1 - 1$ and $n_2 - 1$ degrees of freedom where n_1 is the sample size for the sample having the larger variance and n_2 is the sample size for the sample with the smaller variance. H_o is rejected if

$$F_c > F_{1-\alpha,n_1-1,n_2-1} \qquad (8.19)$$

Test for Equality of Variances from Several Normal Distributions

To test the H_o: $\sigma_1^2 = \sigma_2^2 = \ldots = \sigma_k^2$ for k independent samples each from a normal

population with mean μ_i and variance σ_i^2, it is first necessary to calculate the k sample variances S_i^2. The quantity Q/h is approximately distributed as a chi-square distribution with k - 1 degrees of freedom where

$$Q = \Sigma_{i=1}^{k}(n_i-1)\ln[\Sigma_{i=1}^{k}(n_i-1)S_i^2/(N-k)] - \Sigma_{i=1}^{k}(n_i-1)\ln S_i^2$$

$$h = 1 + \{\Sigma_{i=1}^{k}[1/(n_i-1)] - 1/(N-k)\}/3(k-1)$$

and

$$N = \Sigma_{i=1}^{k} n_i$$

H_o is rejected if

$$Q/h > \chi^2_{1-\alpha,k-1} \qquad (8.20)$$

In this test, H_a is that σ_i^2 are not all equal. This means that at least one σ_i^2 is different from the other σ_i^2. The test is known as Bartlett's test for homogeneity of variances. Homogeneity of variance is also known as homoscedasticity.

Example 8.5. For the preceding example, test the hypothesis that the variance is 36.00.

Solution: The assumption of normality is used. The test is based on equation 8.18 using $\alpha = 0.05$

$$\chi_c^2 = \Sigma_{i=1}^{n}(x_i - \bar{x})^2/\sigma_o^2 = (n-1)s^2/\sigma_o^2$$

$$= 17(4.75)^2/36 = 10.65$$

From a chi-square table

$$\chi^2_{\alpha/2,n-1} = \chi^2_{0.025,17} = 7.56$$

$$\chi^2_{1-\alpha/2,n-1} = \chi^2_{0.0975,17} = 30.2$$

Since 10.65 is between 7.56 and 30.2, H_o is accepted.

TESTING THE GOODNESS OF FIT OF DATA TO PROBABILITY DISTRIBUTIONS

Two ways of judging whether or not a particular distribution adequately describes a set of observations have already been discussed. Both of these methods required a visual judgement of goodness of fit. One method was to compare the observed relative frequency curve with the theoretical relative frequency curve. The second method was to plot the data on appropriate probability paper and judge as to whether or not the resulting plot is a straight line. Statistical tests corresponding to these visual tests will be discussed. In the following discussion, the hypothesis being tested is that the data are from a specified probability distribution.

Chi-Square Goodness of Fit Test

One of the most commonly used tests for goodness of fit of empirical data to specified theoretical frequency distributions is the chi-square test. This test makes a comparison between the actual number of observations and the expected number of observations (expected according to the distribution under test) that fall in the class intervals. The expected numbers are calculated by multiplying the expected relative frequency by the total number of observations. The test statistic is calculated from the relationship

$$\chi_c^2 = \Sigma_{i=1}^{k}(O_i - E_i)^2/E_i \qquad (8.21)$$

where k is the number of class intervals, O_i is the observed and E_i the expected (according to the distribution being tested) number of observations in the i^{th} class interval. The distribution of χ_c^2 is a chi-square distribution with k-p-1 degrees of freedom where p is the number of parameters estimated from the data. The hypothesis that the data are from the specified distribution is rejected if

$$\chi_c^2 > \chi^2_{1-\alpha,k-p-1} \qquad (8.22)$$

Example 8.6. As an example of using the chi-square test, consider the data of table 2.2 and test the hypothesis that the data are from a normal distribution. The observed and expected numbers in each class interval are obtained by multiplying the relative frequency by 66 which is the number of observations. Table 8.2 shows the calculation χ_c^2. The degrees of freedom is k-3 or 7 since two parameters (μ_X and σ_X^2) were estimated for the normal distribution. Comparing χ_c^2 of 10.31 with $\chi^2_{0.90,7} = 12.02$, it is concluded that the normal distribution adequately describes the data for $\alpha = 0.10$. If χ_c^2 had exceeded $\chi^2_{1-\alpha,k-p-1}$, the hypothesis that the normal distribution describes the data would be rejected.

Comment: By examining table 8.2 and equation 8.21, it is apparent that the chi-square goodness of fit test is quite sensitive in the tails of the assumed distribution. Because of this many statisticians recommend that classes be combined if the expected number in a class is less than 3 (or 5). If the 3 criteria is used, the first two classes and the last two classes must be combined. This makes the calculation of χ_c^2 as shown in table 8.3 and χ_c^2 value is reduced to 3.55. The degrees of freedom are reduced to 5.

Table 8.2. Chi-square test on Kentucky River data.

Class Mark	Observed Number	Expected Number	O-E	$\frac{(O-E)^2}{E}$
25.000	2	1.65	0.35	0.0742
35.000	3	3.76	-0.76	0.1536
45.000	10	7.07	2.93	1.2142
55.000	9	10.49	-1.49	0.2116
65.000	11	12.41	-1.41	0.1602
75.000	10	11.75	-1.75	0.2606
85.000	12	8.44	3.56	1.5016
95.000	6	5.35	0.65	0.0789
105.000	0	2.57	-2.57	2.5700
115.000	3	0.99	2.01	4.0809
Total	66	64.48	+1.52	10.3058

Perhaps a better way of conducting the chi-square goodness of fit test is to define the class intervals so that under the hypothesis being tested the expected number of observations in each class interval is the same. This means that the class intervals will be of unequal width and that the interval widths will be a function of the distribution being tested.

Example 8.7. A chi-square test for normality of Kentucky River data using 10 class intervals each having the same expected frequency can be conducted as follows.

Ten class intervals means that the expected relative frequency or probability in each interval is 0.1. The class boundaries can be determined by using a standard normal table. For instance the boundaries of the 4th class interval on a standard normal distribution are $z_{0.3} = -0.52$ to $z_{0.4} = -0.25$. The actual class boundaries are

$$x_{0.3} = s_X z_{0.3} + \bar{x}$$

$$= 21{,}000(-0.52) + 67{,}500 = 56{,}580$$

$$x_{0.4} = 21{,}000(-0.25) + 67{,}500 = 62{,}250$$

Table 8.4 contains the data for conducting the chi-square test based on 10 class intervals having equal expected numbers of observations (66/10 or 6.6) in each interval. In this case χ_c^2 is 4.913 which is less than $\chi^2_{0.90,7}$ of 12.02. The hypothesis is again accepted.

The Kolmogorov-Smirnov Test

An alternative to the chi-square goodness of fit test is the Kolmogorov-Smirnov test. This test is conducted as follows:

1) Let $P_X(x)$ be the completely specified theoretical cumulative distribution function under the null hypothesis.
2) Let $S_n(x)$ be the sample cumulative density function based on n observations. For any observed x, $S_n(x) = k/n$ where k is the number of observations less than or equal to x.

Table 8.3. Chi-square test on Kentucky River data (modified).

Class Mark	Observed Number	Expected Number	O-E	$\frac{(O-E)^2}{E}$
25,000	5	5.41	-0.41	0.0311
35,000				
45,000	10	7.07	2.93	1.2142
55,000	9	10.49	-1.49	0.2116
65,000	11	12.41	-1.41	0.1602
75,000	10	11.75	-1.75	0.2606
85,000	12	8.44	3.56	1.5016
95,000	6	5.35	0.65	0.0789
105,000				
115,000	3	3.56	-0.56	0.0881
Total	66	64.48	+1.52	3.5463

Table 8.4. Chi-square test based on equal expected numbers per class interval.

Class Number	Boundaries lower	Boundaries upper	Observed Number	Expected Number	$\frac{(O-E)^2}{E}$
1	$-\infty$	40,620	6	6.6	0.055
2	40,620	49,860	9	6.6	0.873
3	49,860	56,580	7	6.6	0.024
4	56,580	62,250	4	6.6	1.024
5	62,250	67,500	6	6.6	0.055
6	67,500	72,750	8	6.6	0.300
7	72,750	78,420	4	6.6	1.024
8	78,420	85,140	9	6.6	0.873
9	85,140	94,380	8	6.6	0.300
10	94,380	∞	5	6.6	0.388
		Totals	66	66.0	4.913

3) Determine the maximum deviation, D, defined by

$$D = \max |P_X(x) - S_n(x)|$$

4) If, for the chosen significance level, the observed value of D is greater than or equal to the critical tabulated value of the Kolmogorov-Smirnov statistic, the hypothesis is rejected. The Kolmogorov-Smirnov test statistic is contained in table E.9.

This test can be conducted by calculating the quantities $P_X(x)$ and $S_n(x)$ at each observed point or by plotting the data as in figure 7.5 and selecting the greatest deviation on the probability scale of a point from the theoretical line. If the latter approach is used care must be taken to select the largest deviation on the probability scale which is not necessarily linear. The data should not be grouped for this test.

Note that for the Kolmogorov-Smirnov test, $P_X(x)$ is a completely specified, cumulative probability distribution. That is no parameters for the distribution must be estimated from observed data. Crutcher (1975) points out that when parameters must be estimated to specify $P_X(x)$, the Kolmogorov-Smirnov test is conservative with respect to the Type I error. That is if the critical value is exceeded by the test statistic obtained from the observed values, the hypothesis is rejected with considerable confidence. Crutcher (1975) presents a table of critical values for sample sizes of 25 and 30 as well as infinitely large samples for the exponential, gamma, normal and extreme value distributions when parameters of these distributions must be estimated. In general these critical values are smaller than the values given in table E.9.

Example 8.8. Test the hypothesis that the Kentucky River peak flow data are normally distributed. Use the Kolmogorov-Smirnov test.

Solution: The data are plotted in figure 7.5. The maximum deviation between the best fitting line, $P_X(x)$, and the plotted points, $S_n(x)$, on the probability scale is about 0.06 at X = 47,000 cfs and again at 67,000 cfs. The critical value of the Kolmogorov-Smirnov test statistic from table E.9 for $\alpha = 0.10$ and n = 66 is 0.150. Therefore we cannot re-

ject the hypothesis of a normal distribution.

General Comments of Goodness of Fit Tests

Many hydrologists discourage the use of the chi-square and Kolmogorov-Smirnov tests when testing hydrologic frequency distributions. The reason for this is the importance of the tails of hydrologic frequency distributions and the insensitivity of these statistical tests in the tails of the distributions. In the example above a critical value of the Kolmogorov-Smirnov statistic of 0.150 was obtained. It is nearly impossible to get a deviation of this magnitude in the tails of distributions when the procedures outlined in this chapter are followed. The sensitivity of the chi-square test can be improved in the tails of the distribution if classes are not combined to get an expected frequency of 3 to 5 as recommended earlier. The disadvantage of this is that a single observation in a class with a low expectation can result in a χ_c^2 value in excess of the critical value. This single observation can lead to rejecting the hypothesis. Unfortunately, no satisfactory alternate tests are presently available for making goodness of fit tests.

Neither the chi-square test nor the Kolmogorov-Smirnov test are very powerful in the sense that the probability of accepting the hypothesis when it is in fact false is very high when these tests are used. This is especially true for small samples. These criticisms of the goodness of fit tests will be illustrated in the exercises dealing with simulation in Chapter 13.

Exercises

8.1 A sample of 20 random observations produced a mean of 145 and a variance of 30. What are the 95% confidence intervals on the mean assuming a normal distribution if (a) the true variance is estimated by 30; (b) the true variance is 30. Discuss the reason you feel that the confidence intervals computed for part (a) are wider than for part (b).

8.2 What are the 95% confidence intervals on the variance for the samples of exercise 8.1?

8.3 Test the hypothesis that the true mean of the data producing the sample whose properties are given in exercise 8.1 is 165.

8.4 Discuss any connection between hypothesis testing and confidence intervals that you can discern. What are the differences?

8.5 Assuming the data are normally distributed, test the hypothesis that the mean peak discharge on the Kentucky River near Salvisa (table 2.1) for the period 1895-1916 is different than it is for the period 1939-1960.

8.6 Repeat exercise 8.5 except test for equality of variances.

8.7 Using the data of table 2.1, test the hypothesis that the variances of the peak discharges are the same for the three periods 1895-1916, 1917-1938, 1939-1960.

8.8 Test the hypothesis that the mean monthly rainfall for September and October are the same on the Walnut Gulch watershed near Tombstone, Arizona. What assumptions did you make? Are these assumptions reasonable?

8.9 Repeat exercise 8.8 for equality for variances.

8.10 Test the hypothesis that the difference in the mean monthly rainfall on Walnut Gulch near Tombstone, Arizona, for September and October is 0.50 inches. Discuss the validity of the assumptions that are made.

8.11 Test the hypothesis that monthly rainfall in October on the Walnut Gulch watershed near Tombstone, Arizona, is normally distributed.

8.12 Test the hypothesis that annual rainfall on the Walnut Gulch watershed near Tombstone, Arizona, is normally distributed.

8.13 Comment on the results of exercises 8.11 and 8.12 in terms of the Central Limit Theorem.

8.14 Would the plotting position relationship used in exercise 7.6 have any effect on the results of a test for normality on the data set you selected?

8.15 Use the Kolmogorov-Smirnov test to answer exercise 7.10.

8.16 Use the Kolmogorov-Smirnov test to test for normality the three sets of data plotted in exercise 7.11.

8.17 Use the Kolmogorov-Smirnov test to test for lognormality the three sets of data plotted in exercise 7.12.

8.18 Work exercise 8.16 using the chi-square test.

8.19 Work exercise 8.17 using the chi-square test.

8.20 What distribution do you think would fit the data of exercise 2.2? Use the chi-square test to evaluate your assertion.

8.21 The following are experimentally determined values of Manning's n for plastic pipe as determined by Haan (1965). Test the hypothesis that the mean value of n is different from the recommended design value of 0.0090.

0.0092	0.0085	0.0083	0.0091
0.0078	0.0084	0.0091	0.0088
0.0086	0.0090	0.0089	0.0093
0.0081	0.0092	0.0085	0.0090
0.0085	0.0088	0.0088	0.0093

9. Simple Linear Regression

NOTATION: IN THIS chapter an upper case letter will represent a variable, a lower case letter will represent the difference between a variable and its mean, and a subscript will be used to denote a particular value for the variable. Thus Y represents a variable which may take on values Y_1, Y_2, Y_3, etc. $\bar{Y}$ is the mean of Y, $y = Y - \bar{Y}$ and $y_i = Y_i - \bar{Y}$. Parameters are denoted by Greek letters and a corresponding English letter is used to denote an estimate for the parameter. Thus α is a parameter estimated by a ($\hat{\alpha} = a$). The small letter e will be used to denote the difference between an observed value of Y and its predicted value $\hat{Y}$. Thus $Y - \hat{Y} = e$ and $Y_i - \hat{Y}_i = e_i$. All summations in this chapter will run from 1 to n unless otherwise specified where n is number of observations on Y and X.

SIMPLE REGRESSION[1]

Possibly the most common model used in hydrology is based on the assumption of a linear relationship between two variables. Generally the objective of such a model is to provide a means of predicting or estimating one variable, the dependent variable, from knowledge of a second variable, the independent variable.[2] Such a situation is shown in figure 9.1 where a relationship is sought between the annual runoff, Y, and the annual precipitation, X, for Cave Creek near Lexington, Kentucky. The data used in constructing figure 9.1 is contained in table 9.1.

Two questions are of immediate concern. Can a model of the form

$$Y = \alpha + \beta X \tag{9.1}$$

1. Many times the term regression is reserved for use when all of the X variables being considered are random variables. In this book liberties will be taken and the term applied whether or not the X variables are random variables.

2. Dependent and independent in this sense are not the same as dependence or independence of random variables.

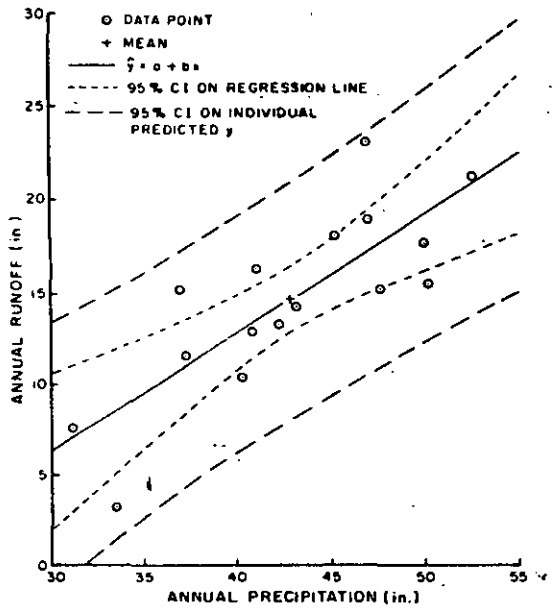


Fig. 9.1. Annual rainfall-runoff relation for Cave Creek.

Table 9.1. Annual precipitation and runoff for Cave Creek, near Lexington, Kentucky.

Year	Precipitation (inches)	Runoff (inches)	Year	Precipitation (inches)	Runoff (inches)
1953	42.39	13.26	1961	47.08	22.91
1954	33.48	3.31	1962	47.08	18.89
1955	47.67	15.17	1963	40.89	12.82
1956	50.24	15.50	1964	37.31	11.58
1957	43.28	14.22	1965	37.15	15.17
1958	52.60	21.20	1966	40.38	10.40
1959	31.06	7.70	1967	45.39	18.02
1960	50.02	17.64	1968	41.03	16.25

adequately represent the relationship between Y and X? For what values of α and β is the representation the best?

In looking at the question of the "best" straight line, a criteria for judging "best-ness" is needed. One intuitive criteria would be to estimate α and β by a and b so as to minimize the deviation e_i between the observed values of Y, Y_i, and the predicted values of Y, $\hat{Y}_i$. In this way values for a and b would be sought that minimize the sum

$$\Sigma(Y_i - \hat{Y}_i) = \Sigma e_i = \Sigma(Y_i - \hat{\alpha} - \hat{\beta} X_i) = \Sigma(Y_i - a - bX_i) \tag{9.2}$$

Closer scrutiny of equation 9.2 reveals that it is not desirable to minimize the sum in an algebraic sense because that would be equivalent to finding an a and b such that Σe_i is

minus infinity.

Another criteria might be to find an a and b such that Σe_i is zero. The fallacy with this can be seen by considering two points. If the line $Y = a + bX$ goes through the two points, then Σe_i would be zero; however, the sum is also zero for any line that overpredicts one point by the same amount that it underpredicts the second point. Thus there is an infinity of lines such that $\Sigma e_i = 0$ and an additional restriction or criterion is needed to select a single line.

The Σe_i may be positive or negative. A criterion that is not sign dependent is needed. Such a criterion might be to minimize $\Sigma |e_i|$ or to minimize Σe_i^2. Since absolute values are difficult to work with mathematically, the second criterion is generally selected. Thus it is desired to estimate α and β by a and b such Σe_i^2 is a minimum. Denoting this sum by M, we have

$$M = \Sigma e_i^2 = \Sigma (Y_i - \hat{Y}_i)^2 = \Sigma (Y_i - a - bX_i)^2 \tag{9.3}$$

This sum can be minimized with respect to a and b by taking the partial derivatives of M with respect to a and b and setting the resulting equations equal to zero.

$$\frac{\partial M}{\partial a} = -2\Sigma (Y_i - a - bX_i) = 0$$

$$\frac{\partial M}{\partial b} = -2\Sigma X_i (Y_i - a - bX_i) = 0$$

These equations can then be written in the following form known as the normal equations.

$$\Sigma (Y_i - a - bX_i) = 0 \tag{9.4}$$

$$\Sigma X_i (Y_i - a - bX_i) = 0 \tag{9.5}$$

The solution of the normal equations in terms of a and b is

$$b = [\Sigma X_i Y_i - \Sigma X_i \Sigma Y_i / n] / [\Sigma X_i^2 - (\Sigma X_i)^2 / n] = \Sigma x_i y_i / \Sigma x_i^2 \tag{9.6}$$

$$a = (\Sigma Y_i - b \Sigma X_i)/n = (\bar{Y} - b\bar{X}) \tag{9.7}$$

Equations 9.6 and 9.7 provide estimates for α and β such that Σe_i^2 is a minimum. Since the procedure is based on minimizing Σe_i^2, the estimates a and b are commonly called least squares estimates. Equation 9.4 indicates that this solution also satisfies $\Sigma e_i = 0$. Equation 9.7 indicates that the line $Y = a + bX$ goes through the point $Y = \bar{Y}$ and $X = \bar{X}$.

The line $Y = a + bX$ is commonly known as the regression line of Y on X. The procedure of determining a and b is known as simple regression. The term "simple" regression is used when only one independent variable is involved as opposed to multiple regression when several independent variables are involved. The parameter estimates, a and b, are known as regression coefficients.

Equations 9.6 and 9.7 show that a and b are functions of the sample values of Y and X. If another sample of observations were obtained and α and β were estimated from this sample, different estimates would result. We have already seen that

$$Y_i - \hat{Y}_i = Y_i - a - bX_i = e_i$$

Similarly

$$Y_i - \alpha - \beta X_i = \varepsilon_i$$

Thus e_i represents the deviation between an observed Y_i and its predicted value $\hat{Y}_i$ based on the regression equation estimated from the particular sample of data at hand. ε_i represents the deviation between an observed Y_i and the assumed true but unknown relation between Y and X given by $Y = \alpha + \beta X$.

Example 9.1. Determine the regression coefficients for the data plotted in figure 9.1.

Solution: The data required for solving equations 9.6 and 9.7 are contained in table 9.2. The equation used to calculate b would depend on the method of calculation. If a small desk calculator is used, the first of equations 9.6 might be employed. If an electronic calculator or computer is used, the latter of equations 9.6 might be employed. Generally less roundoff error will result if the latter form of equation 9.6 is used.

$$b = \Sigma x_i y_i / \Sigma x_i^2 = 369.4320/570.0559 = 0.6480$$

$$a = \bar{Y} - b\bar{X} = 14.63 - 0.6480(42.94) = -13.1951$$

Table 9.2. Calculations on data of Table 9.1.

	Y	X	y	x	$\hat{Y}$	$Y-\hat{Y}$
	13.26	42.39	- 1.37	- 0.55	14.27	-1.01
	3.31	33.48	-11.32	- 9.46	8.50	-5.19
	15.17	47.67	0.54	4.73	17.70	-2.53
	15.50	50.24	0.87	7.30	19.36	-3.86
	14.22	43.28	- 0.41	0.34	14.85	-0.63
	21.20	52.60	6.57	9.66	20.89	0.31
	7.70	31.06	- 6.93	-11.88	6.93	0.77
	17.64	50.02	3.01	7.08	19.22	-1.58
	22.91	47.08	8.28	4.14	17.31	5.60
	18.89	47.08	4.26	4.14	17.31	1.58
	12.82	40.89	- 1.81	- 2.05	13.30	-0.48
	11.58	37.31	- 3.05	- 5.63	10.98	0.60
	15.17	37.15	0.54	- 5.79	10.88	4.29
	10.40	40.38	- 4.23	- 2.56	12.97	-2.57
	18.02	45.39	3.39	2.45	16.22	1.80
	16.25	41.03	1.62	- 1.91	13.39	2.86
Total	234.04	687.05	- 0.04	0.01	234.08	-0.04
Mean	14.63	42.94	0.00	0.00	14.63	0.00

Therefore $\hat{Y} = -13.1951 + 0.6480X$. This line is plotted in figure 9.1.

Comment: The last two columns of table 9.2 contain $\hat{Y}_i$ and $Y_i - \hat{Y}_i$. Note that except for rounding errors, $\bar{Y} = \bar{\hat{Y}}$, $\Sigma(Y_i - \hat{Y}_i) = \Sigma e_i = 0$, and $\bar{e} = 0$.

EVALUATING THE REGRESSION

The second question is now considered. Can the data be adequately described by the regression line? Naturally the answer to this query depends on the definition of adequate. The question will not be answered here but methods for assessing the adequacy of the model will be explored.

One approach that does not involve any assumptions is to determine how much of the variability in the dependent variable is explained by the regression. The variability in the dependent variable is quantified as a sum of squares. Looking at figure 9.2, it can be seen that Y_i can be expressed as

$$Y_i = \bar{Y} + \hat{Y}_i - \bar{Y} + Y_i - \hat{Y}_i$$

or

$$Y_i - \hat{Y}_i = (Y_i - \bar{Y}) - (\hat{Y}_i - \bar{Y})$$

Squaring both sides of this equation and summing over all of the observations results in

$$\Sigma(Y_i - \hat{Y}_i)^2 = \Sigma(Y_i - \bar{Y})^2 - 2\Sigma(Y_i - \bar{Y})(\hat{Y}_i - \bar{Y}) + \Sigma(\hat{Y}_i - \bar{Y})^2 \qquad (9.8)$$

Since $\hat{Y}_i = a + bX_i$ and $a = \bar{Y} - b\bar{X}$, $(\hat{Y}_i - \bar{Y})$ can be written

$$(\hat{Y}_i - \bar{Y}) = b(X_i - \bar{X})$$

so that

$$-2\Sigma(Y_i - \bar{Y})(\hat{Y}_i - \bar{Y}) = -2b\Sigma(Y_i - \bar{Y})(X_i - \bar{X}) \qquad (9.9)$$

However from equation 9.6

$$\Sigma(Y_i - \bar{Y})(X_i - \bar{X}) = b\Sigma(X_i - \bar{X})^2$$

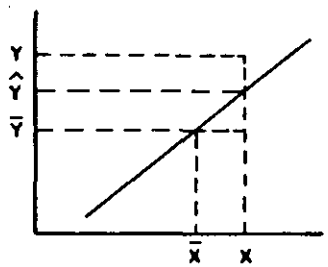


Fig. 9.2. Components of Y.

Therefore

$$-2\Sigma(Y_i - \bar{Y})(\hat{Y}_i - \bar{Y}) = -2b^2\Sigma(X_i - \bar{X})^2 = -2\Sigma[b(X_i - \bar{X})]^2$$

or

$$-2\Sigma(Y_i - \bar{Y})(\hat{Y}_i - \bar{Y}) = -2\Sigma(\hat{Y}_i - \bar{Y})^2 \qquad (9.10)$$

We can now write equation 9.8 as

$$\Sigma(Y_i - \hat{Y}_i)^2 = \Sigma(Y_i - \bar{Y})^2 - \Sigma(\hat{Y}_i - \bar{Y})^2$$

Rearranging terms results in

$$\Sigma(Y_i - \bar{Y})^2 = \Sigma(Y_i - \hat{Y}_i)^2 + \Sigma(\hat{Y}_i - \bar{Y})^2$$

However $\Sigma(Y_i - \bar{Y})^2$ is equal to $\Sigma Y_i^2 - n\bar{Y}^2$ so we have

$$\Sigma Y_i^2 = n\bar{Y}^2 + \Sigma(Y_i - \hat{Y}_i)^2 + \Sigma(\hat{Y}_i - \bar{Y})^2 \qquad (9.11)$$

The total sum of squares, ΣY_i^2, has been partitioned into three components. These three components are:

1. $n\bar{Y}^2$, the sum of squares due to the mean,
2. $\Sigma(Y_i - \hat{Y}_i)^2 = \Sigma e_i^2$, the sum of squares of deviations from regression or the residual sum of squares, and
3. $\Sigma(\hat{Y}_i - \bar{Y})^2$, the sum of squares due to regression.

The sum of squares about the mean or the sum of squares corrected for the mean is

$$\Sigma(Y_i - \bar{Y})^2 = \Sigma y_i^2 = \Sigma Y_i^2 - n\bar{Y}^2 = \Sigma(Y_i - \hat{Y}_i)^2 + \Sigma(\hat{Y}_i - \bar{Y})^2 \qquad (9.12)$$

or using equations 9.9 and 9.10

$$\Sigma y_i^2 = \Sigma e_i^2 + b\Sigma x_i y_i$$

Therefore the total sum of squares corrected for the mean is made up of two components - the sum of squares of deviation from regression (also known as the error or residual sum of squares) and the sum of squares due to regression. The larger the sum of squares due to regression is in comparison to the residual sum of squares, the more of the total sum of squares corrected for the mean is explained by the regression equation. The ratio of the sum of squares due to regression to the total sum of squares corrected for the mean can be used as a measure of ability of the regression line to explain variations in the dependent variable. This ratio is commonly denoted by r^2 and may be written in a number of ways.

$$r^2 = \text{sum of squares due to regression/sum of squares corrected for mean} = \Sigma(\hat{Y}_i - \bar{Y})^2/\Sigma y_i^2 \qquad (9.13)$$

From equations 9.9 and 9.10 this may be written as

$$r^2 = b \Sigma x_i y_i / \Sigma y_i^2 = (\Sigma x_i y_i)^2 / (\Sigma x_i^2 \Sigma y_i^2) \qquad (9.14)$$

r^2 is called the coefficient of determination. If the regression equation perfectly predicts every value of Y_i, then e_i would be zero for every i and Σe_i^2 would be zero. Under these conditions equation 9.12 states that $\Sigma y_i^2 = \Sigma(\hat{Y}_i - \bar{Y})^2$ so that from equation 9.13, r^2 is seen to be one. On the other hand, if the regression equation is explaining none of the variations in Y, then Σe_i^2 will equal Σy_i^2 and $\Sigma(\hat{Y}_i - \bar{Y})^2$ will be zero. Under this condition r^2 will be zero. Thus the range in possible values for r^2 is from zero to one. The closer r^2 is to one, the better the regression equation "fits" the data. r^2 is the fraction of the total sum of squares about the mean that is explained by the regression equation.

From equation 9.6 and 9.14 we can write

$$r^2 = b \Sigma x_i y_i / \Sigma y_i^2 = b^2 \Sigma x_i^2 / \Sigma y_i^2 = b^2 s_X^2 / s_Y^2$$

or

$$r = b s_X / s_Y \qquad (9.15)$$

Since $0 \leq r^2 \leq 1$, we have $-1 \leq r \leq 1$. The sign on r is identical to the sign on b since s_X and s_Y are always positive. From equation 9.14 it can be seen that r may also be written as

$$r = \Sigma x_i y_i / (\Sigma x_i^2 \Sigma y_i^2)^{1/2} = s_{X,Y} / s_X s_Y \qquad (9.16)$$

which would be equal to the sample correlation coefficient given by equation 3.54 if X and Y were both random variables. In fact, r is commonly called the correlation coefficient and can be shown to be equal to the correlation between Y and $\hat{Y}$. Correlation is discussed in more detail in Chapter 11.

Example 9.2. What percent of the variation in Y is accounted for by the regression of example 9.1?

Solution:

$$r^2 = b \Sigma x_i y_i / \Sigma y_i^2 = 0.6480(369.4230)/363.0714 = 0.66$$

Thus 66 percent of the variation in Y is explained by the regression equation. The remaining 34 percent of the variation is due to unexplained causes.

CONFIDENCE INTERVALS AND TESTS OF HYPOTHESES

Thus far in the discussion of simple regresssion no assumptions have been made concerning the model. In order to use some well-developed theorems concerning hypothesis testing and confidence interval estimation, it is necessary to make the assumption that the ϵ_i are identically and independently distributed as a normal distribution with a mean of zero and a variance of σ^2. (A shorthand way of writing this is ϵ_i is i.i.d. $N(0,\sigma^2)$.[3]

This assumption contains a lot of implications. The fact that the $E(\epsilon_i) = 0$ has been guaranteed by our estimation procedures. The assumption of independence means that the correlation between ϵ_i and ϵ_j for any $i \neq j$ must be zero. The assumption that the ϵ_i are identically distributed with variance σ^2 means that the variance of ϵ_i must equal the variance of ϵ_j for all i and j. That is, the variance of ϵ_i cannot change as X_i changes. This is known as homoscedasticity. Finally we must have the ϵ_i normally distributed.

The assumption of normality of the ϵ_i can be checked by the procedure of Chapter 8. A rough check would be to note that, for the normal distribution, 95% of the values of e_i should be within 2 standard deviations of the mean or only about 5% of the residuals should lie outside the interval -2σ to 2σ.[4]

Under the normality assumption, we have $E(\epsilon) = 0$. The $Var(\epsilon)$ is given by

$$Var(\epsilon) = E(\epsilon^2) - E^2(\epsilon) = E(\epsilon^2) = \sigma^2$$

The positive square root of the Var (ϵ) is known as the standard error of the regression equation.

An unbiased estimate (Graybill 1961) for Var (ϵ_i) is S^2 calculated from

$$s^2 = \Sigma e_i^2 / (n-2) = \Sigma(Y_i - \hat{Y}_i)^2 / (n-2) \qquad (9.17)$$

The least squares estimation procedure produces estimates for α and β such that the standard error of the regression equation is a minimum.

Another way to look at the coefficient of determination is to write equation 9.13 as

$$r^2 = (\Sigma y_i^2 - \Sigma e_i^2)/\Sigma y_i^2 = 1 - \Sigma e_i^2 / \Sigma y_i^2$$

or

$$r^2 = 1 - (n-2)s^2/(n-1)s_Y^2 \qquad (9.18)$$

Therefore, if the estimated standard error of the regression equation is nearly equal to the standard deviation of Y, r^2 will be close to zero and the regression equation is of little value in explaining variation in Y.

Example 9.3. Is there reason to believe the residuals of example 9.1 are not normally distributed?

Solution:

---

3. For further discussion of the assumptions involved in regression analysis, see the closing section of this chapter "General Considerations". Also see Johnston (1963) and Graybill (1961).

4. For a further discussion of examining the e_i, reference should be made to Draper and Smith (1966).

$$s^2 = \hat{\sigma}^2 = \Sigma e_i^2 / (n - 2)$$

$$= 123.7/14 = 8.83$$

$$s = 2.97$$

95% of the e_i should be between -2s and 2s, or between -5.94 and +5.94. An inspection of table 9.2 shows that none of the 16 observations are outside this interval. The number of observations is not sufficient to determine if the e_i are $N(0,\sigma^2)$, however, there is not sufficient evidence to reject this possibility.

Inferences on Regression Coefficients

In order to place confidence intervals on α and β and to test hypotheses concerning them, it is necessary to know the Var(a) and Var(b) which will be designated as σ_a^2 and σ_b^2 and estimated by S_a^2 and S_b^2.

$$b = \Sigma(X_i - \bar{X})(Y_i - \bar{Y}) / \Sigma(X_i - \bar{X})^2$$

$$= [\Sigma Y_i(X_i - \bar{X}) - \bar{Y}\Sigma(X_i - \bar{X})]/\Sigma x_i^2$$

$$= \Sigma Y_i x_i / \Sigma x_i^2$$

$$\text{Var}(b) = \text{Var}[\Sigma Y_i x_i / \Sigma x_i^2]$$

Now $Y_i = e_i + \hat{Y}_i$. If the model is correct, $Y_i = \varepsilon_i + \alpha + \beta X_i$. But $\alpha + \beta X_i$ is a constant so $\text{Var}(Y_i) = \text{Var}(\varepsilon_i) = \sigma^2$. Furthermore, the Y_i must be independent. Therefore treating the x_i as constants, we can write[5]

$$\text{Var}(b) = \text{Var}[(x_1 Y_1 + x_2 Y_2 + \ldots + x_n Y_n)/\Sigma x_i^2]$$

$$= \Sigma x_i^2 \text{Var} Y_i / (\Sigma x_i^2)^2 = \sigma^2 \Sigma x_i^2 / (\Sigma x_i^2)^2$$

$$\sigma_b^2 = \sigma^2 / \Sigma x_i^2 \qquad (9.19)$$

$$a = \bar{Y} - b\bar{X}$$

$$\text{Var}(a) = \text{Var}(\bar{Y} - b\bar{X}) = \text{Var}\,\bar{Y} + \text{Var}\,b\bar{X}^{6}$$

$$= \sigma^2/n + \bar{X}^2 \text{Var}\, b$$

$$= \sigma^2/n + \sigma^2 \bar{X}^2 / \Sigma x_i^2$$

5. Here the assumption is made that X is a fixed quantity, not a random variable. It can be shown (Graybill 1961; Johnston 1963) that the same result can be obtained by assuming that Y and X are both random variables and are jointly normally distributed or by assuming that Y and X are both random variables with the conditional distribution of Y being normal with mean $\alpha + \beta X$ and variance σ^2 and the marginal distribution of X independent of α, β, and σ^2 (Graybill 1961).

6. Cov(Y,b) = 0 (Draper and Smith 1966).

$$\sigma_a^2 = \sigma^2 [1/n + \bar{X}^2 / \Sigma x_i^2] \qquad (9.20)$$

If the model is correct, then the quantities b/S_b and a/S_a are distributed as a t distribution with n - 2 degrees of freedom. Thus the confidence limits on α can be estimated from[7]

$$L_\alpha = a - t_{1-\alpha/2,n-2}\, S_a$$

$$U_\alpha = a + t_{1-\alpha/2,n-2}\, S_a \qquad (9.21)$$

where S_a is estimated from equation 9.20 by using S^2 as an estimate for σ^2.

Similarly the confidence limits on β are

$$L_\beta = b - t_{1-\alpha/2,n-2}\, S_b$$

$$U_\beta = b + t_{1-\alpha/2,n-2}\, S_b \qquad (9.22)$$

where S_b is estimated from equation 9.19 using S^2 as an estimate for σ^2.

Test of hypotheses concerning α and β can be made by noting that $(a - \alpha_0)/S_a$ and $(b - \beta_0)/S_b$ both have t distributions with n - 2 degrees of freedom. Thus the hypothesis H_0: $\alpha = \alpha_0$ versus H_a: $\alpha \neq \alpha_0$ is tested by computing

$$t = (a - \alpha_0)/s_a \qquad (9.23)$$

H_0 is rejected if $|t| > t_{1-\alpha/2,n-2}$.

Similarly H_0: $\beta = \beta_0$ versus H_a: $\beta \neq \beta_0$ is tested by computing

$$t = (b - \beta_0)/s_b \qquad (9.24)$$

H_0 is rejected if $|t| > t_{1-\alpha/2,n-2}$.

The significance of the overall regression equation can be evaluated by testing the hypothesis that $\beta = 0$. If this hypothesis is accepted, then $\hat{Y} = \bar{Y}$ or the regression line does not explain a significant amount of the variation in Y. In this situation one would be as well off using $\bar{Y}$ as an estimator for Y regardless of the value of X. The H_0: $\beta = 0$ is equivalent to H_0: $r = 0$.

Example 9.4. Compute the 95% confidence intervals on α and β and test the hypothesis that $\alpha = 0$ and the hypothesis that $\beta = 0.500$ for the regression of example 9.1.

Solution:

$$s_a = s[1/n + \bar{X}^2/\Sigma x_i^2]^{1/2}$$

$$= 2.97\,[1/16 + 1843.8436/570.0559]^{1/2}$$

$$= 5.39$$

7. Do not confuse the parameter α in $Y = \alpha + \beta X$ and the confidence level α $t_{1-\alpha/2}$. They are not the same.

$s_b = s/(\Sigma x_i^2)^{1/2} = 2.97/(570.0559)^{1/2}$

$= 0.125$

$t_{1-\alpha/2,n-2} = t_{.975,14} = 2.14$

The 95% confidence intervals on α are

$l_\alpha = a - t_{1-\alpha/2,n-2}\, s_a = -13.1951 - 2.14(5.39) = -24.73$

$u_\alpha = a + t_{1-\alpha/2,n-2}\, s_a = -13.1951 + 2.14(5.39) = -1.66$

The 95% confidence intervals on β are

$l_\beta = b - t_{1-\alpha/2,n-2}\, s_b = .6480 - 2.14(.125) = 0.36$

$u_\beta = b + t_{1-\alpha/2,n-2}\, s_b = .6480 + 2.14(.125) = 0.92$

To test H_0: $\alpha = 0$ versus H_a: $\alpha \neq 0$, compute

$t = (a - 0.00)/s_a = -13.1951/5.39 = -2.44$

$t_{1-\alpha/2,n-2} = t_{.975,14} = 2.14$

Since $|t| > t_{.975,14}$, we reject H_0: $\alpha = 0$.

To test the H_0: $\beta = .5$ versus H_a: $\beta \neq .5$, compute

$t = (b - 0.5)/s_b = (0.6480 - 0.5)/0.125 = 1.184$

Since $|t| < t_{.975,14}$, we cannot reject H_0. The slope is not significantly different from 0.5.

Comment: The significance of the overall regression can be evaluated by testing H_0: $\beta = 0$. Under this hypothesis

$t = (b - 0.00)/s_b = 0.6480/0.125 = 5.184$

Since $|t| > t_{.975,14}$ we reject H_0. The regression equation is explaining a significant amount of the variation in Y.

Confidence Intervals on Regression Line

Confidence intervals on the regression line can be determined by first calculating the variance of $\hat{\hat{Y}}_k$ where $\hat{\hat{Y}}_k$ represents the predicted mean value of $\hat{Y}_k$ for a given X_k.

$\hat{Y}_k = a + bX_k$

From equation 3.55

$\mathrm{Var}(\hat{\hat{Y}}_k) = \mathrm{Var}(a) + X_k^2\, \mathrm{Var}(b) + 2X_k\, \mathrm{Cov}(a,b)$

Mood and Graybill (1963) give $\mathrm{Cov}(a,b) = -\sigma^2 \bar{X}/\Sigma x_i^2$. Therefore

$\mathrm{Var}(\hat{\hat{Y}}_k) = \sigma^2 [1/n + (\bar{X}^2 + X_k^2 - 2\bar{X}X_k)/\Sigma x_i^2]$

or

$$\mathrm{Var}(\hat{\hat{Y}}_k) = \sigma^2 (1/n + x_k^2/\Sigma x_i^2) \tag{9.25}$$

The standard error of $\hat{\hat{Y}}_k$ could be estimated by $S_{\hat{\hat{Y}}_k}$ calculated as

$$s_{\hat{\hat{Y}}_k} = s(1/n + x_k^2/\Sigma x_i^2)^{1/2} \tag{9.26}$$

Equation 9.25 indicates that the variance of $\hat{\hat{Y}}_k$ depends on the particular value of X at which the variance is being determined. The $\mathrm{Var}(\hat{\hat{Y}}_k)$ is a minimum when $X_k = \bar{X}$ and increases as X_k deviates from $\bar{X}$.

Confidence limits on the regression line are now given by

$$L = \hat{\hat{Y}}_k - S_{\hat{\hat{Y}}_k}\, t_{1-\alpha/2,n-2}$$
$$U = \hat{\hat{Y}}_k + S_{\hat{\hat{Y}}_k}\, t_{1-\alpha/2,n-2} \tag{9.27}$$

where $\hat{\hat{Y}}_k = a + bX_k$ and $S_{\hat{\hat{Y}}_k}$ is given by equation 9.26. Since $S_{\hat{\hat{Y}}_k}$ increases as x_k or $X_k - \bar{X}$ increases, the confidence intervals are the narrowest at $X_k = \bar{X}$ and widen as X_k deviates from $\bar{X}$.

The confidence limits on an individual predicted value of Y would be wider than the confidence interval on the regression line since for an individual Y, the $\mathrm{Var}(\varepsilon) = \sigma^2$ would have to be added to the $\mathrm{Var}(\hat{\hat{Y}}_k)$. Thus the variance of an individual predicted value of Y would be $\mathrm{Var}(\hat{\hat{Y}}_k) + \sigma^2$. Confidence intervals on an individual predicted value of Y could then be estimated from equations 9.27 where the expression

$$s(1 + 1/n + x_k^2/\Sigma x_i^2)^{1/2} \tag{9.28}$$

would be substituted for $s_{\hat{\hat{Y}}_k}$. The confidence limits on a future predicted value of Y are the same as those for an individual predicted value of Y.

Example 9.5. Calculate the 95% confidence limits for the regression line of example 9.1. Calculate the 95% confidence interval for an individual predicted value of Y for the same problem.

Solution: $s = 2.97$, $n = 16$, $\Sigma x_i^2 = 570.0559$, $t_{.975,14} = 2.14$ and $\bar{X} = 42.94$. Therefore from equations 9.27 we have for the 95% confidence intervals on the regression line

$$\begin{matrix} l \\ u \end{matrix} = -13.1951 + 0.6480\, X_k \mp 2.97\, [1/16 + (X_k - 42.94)^2/570.0559]^{1/2}\, 2.145$$

where the - applies to the lower limit, l, and the + to the upper limit, u. Similarly the 95% confidence interval on an individual predicted value of Y is given by:

$$\begin{matrix} l \\ u \end{matrix} = -13.1951 + 0.6480\, X_k \mp 2.97\, [1 + 1/16 + (X_k - 42.94)^2/570.0559]^{1/2}\, 2.145$$

By substituting various values of X_k into these equations, the desired confidence limits are obtained. These intervals are plotted in figure 9.1.

Confidence Intervals on Standard Error

Confidence intervals may be placed on σ^2 by noting that the quantity $(n-2)S^2/\sigma^2$ is distributed as a chi-square distribution with $n-2$ degrees of freedom. Thus confidence limits on σ^2 are given by

$$L = (n-2)S^2/\chi^2_{1-\alpha/2,n-2} \qquad U = (n-2)S^2/\chi^2_{\alpha/2,n-2} \tag{9.29}$$

where S^2 is determined from equation 9.17.

EXTRAPOLATION

The extrapolation of a regression equation beyond the range of X used in estimating α and β is discouraged for two reasons. First as can be seen from figure 9.1 and equation 9.27, the confidence intervals on the regression line become very wide as the distance from $\bar{X}$ is increased. Second the relation between Y and X may be nonlinear over the entire range of X and only approximately linear for the range of X investigated. A typical example of this is shown in figure 9.3.

GENERAL CONSIDERATIONS

Many authors discuss several different linear models depending on the assumptions made concerning Y, X and ϵ (Graybill 1961; Benjamin and Cornell 1970; Mood and Graybill 1963). These different models revolve around whether X (or $\underline{X}$ in multiple regression) is a random or nonrandom variable, whether measurement errors are made on Y and/or X, the distribution of X if X (or $\underline{X}$) is a random variable and the joint distribution of Y and X (or $\underline{X}$) if X is a random variable.

The most common assumptions are:

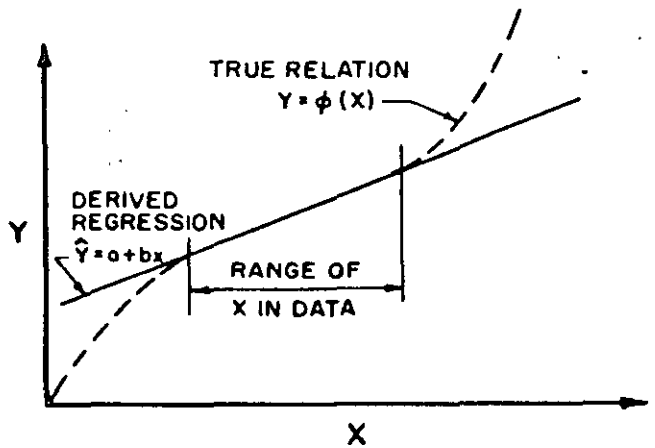


Fig. 9.3. Effect on nonlinearity and extrapolation.

(1) X is a nonrandom variable measured without error, Y is a random variable and that $E(Y_i|X)$ is normally and independently distributed with mean $\alpha + \beta X$ and variance σ^2.

(2) Y and X are both random variables having a joint distribution with the conditional distribution of Y being $N(\alpha + \beta X, \sigma^2)$ and the marginal distribution of X being independent of α, β and σ^2.

It turns out that under either of the above conditions, the procedures given in this chapter are valid for tests of hypotheses and confidence interval estimation at a specified level of significance. Graybill (1961) points out that the power of the tests are not the same for the two conditions.

If X is a fixed variable measured without error and ϵ_i is independently and identically distributed $N(0,\sigma^2)$ or Y and X are from a bivariate normal distribution and are measured without error or Y and X are from a bivariate non-normal population with the conditional distribution of Y being $N(\alpha + \beta X, \sigma^2)$ and the marginal distribution of X independent of α, β and σ^2, then the least squares estimates of α, β and σ^2 are also maximum likelihood estimators. The least squares estimates for the regression coefficients are unbiased.

If significant measurement errors are made on the X variables, then complications arise. For this situation reference can be made to Graybill (1961) or Johnston (1963). Certainly measurement errors are always present, however, if these errors are small relative to X, then the theory presented in this chapter and chapters 10, 11 and 12 may still be applied.

The reason that measurement errors on X cause problems can be seen by considering the model $Y = \alpha + \beta X + \epsilon$. If Y and X contain measurement errors, then Y and X are not observed. What is observed is Y^* and X^* where $Y^* = Y + e_Y$ and $X^* = X + e_X$ where e_Y and e_X are the measurement errors on Y and X. Thus the normal equations are solved in terms of $Y^* = \alpha + \beta X^* + \epsilon$ or $Y + e_Y = \alpha + \beta(X + e_X) + \epsilon = \alpha + \beta X + \beta e_X + \epsilon$. Now if e_X is small in comparison to X, this latter equation becomes $Y = \alpha + \beta X + \epsilon - e_Y$ or $Y = \alpha + \beta X + e_i$ which can be handled by the methods outlined in this chapter.

Recall that no distributional assumptions are required to get the least squares estimates for α and β. The assumptions are involved when confidence intervals and tests of hypotheses are of concern or when it is desired to state that the least squares estimates for α and β are also maximum likelihood estimates. Johnston (1963) points out that the least squares estimates for α and β are biased if significant measurement errors are present on X.

One of the assumptions used in developing confidence intervals and tests of hypotheses was that the e_k are independent. If e_i is correlated with e_{i+1}, the least squares estimates of α and β are unbiased, however, the sampling variance of α and β will be unduly large and will be underestimated by the least squares formulas for variances. Also the sampling variances on predictions made with the resulting equation will be needlessly large. Correlation between e_i and e_{i+1} frequently arise when time series data are being analyzed. This type of correlation is known as autocorrelation or serial correlation. Johnston (1963) discusses least squares estimation procedures in the presence of autocorrelation.

In some situations the assumption of homoscedasticity ($Var(\epsilon_i) = \sigma^2$ for all i) is violated. Quite commonly $Var(\epsilon_i)$ increases as X increases. Such a situation is depicted in figure 9.4. Draper and Smith (1966) and Johnston (1963) discuss least squares esti-

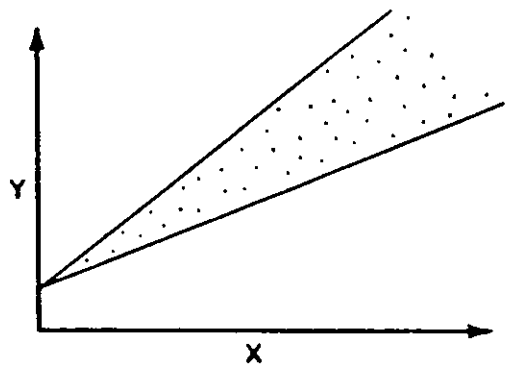


Fig. 9.4. Illustration of situation where $\mathrm{Var}(\varepsilon_i) \neq \sigma^2$ for all i.

mation under this condition.

Another point to be made concerning hypothesis testing in general is that a statistically significant difference and a physically significant difference are two entirely different quantities. For example when the H_o: $\beta = 0$ was tested in example 9.4, the conclusion was that the regression line explained a significant amount of the variation in Y. This refers to a statistically significant amount of the variation at the chosen level of significance. It means that recognizing an α% chance of an error, the relationship $Y = a + bX$ cannot be attributed to chance. It does not imply a cause and effect relationship between Y and X.

Looking at the confidence limits on the regression as plotted in figure 9.1 and the scatter of the data, it can be seen that this simple relationship $Y = a + bX$ leaves a lot to be desired in terms of predicting annual runoff. Whether or not the derived relationship is usable depends on the use to be made of the predicted values of Y and not on the fact that the H_o: $\beta = 0$ is rejected.

Exercises

9.1 The following data are the maximum air and soil temperature (bare soil at 2 inch depth) recorded for the first 30 days of July, 1973, at Lexington, Kentucky. Derive a linear relationship via simple regression for predicting the maximum soil temperature from the maximum air temperature. Estimate σ and r^2 for

Max Temp		Max Temp		Max Temp	
Air	Soil	Air	Soil	Air	Soil
80	83*	81	87	89	90*
87	87	78	93	87	90*
89	92*	87	95	91	91
86	87*	88	92*	83	84*
82	89	75	80	87	90
83	88	82	84*	84	85*
88	92	84	92	82	86
88	97	87	96	85	87
90	95	85	91*	82	93
84	84*	87	92*	85	93

the resulting regression. Test the hypothesis that (a) the intercept is zero (b) the slope is one (c) the regression is explaining a significant amount of the variation in the maximum soil temperature. Would you recommend using this relationship for predicting maximum soil temperature?

9.2 The asterisks following the soil data in exercise 9.1 indicate days on which rainfall occurred. Using only these rainfall days, work exercise 9.1.

9.3 Calculate the regression coefficients in the relationship $Q_s = a + bQ$ where Q_s is the annual suspended sediment load and Q is the annual water discharge for the Green River at Munfordville, Kentucky. Calculate the standard error of the regression equation and the correlation coefficient. Plot the data along with the 95% confidence intervals on the regression line. Is this a usable prediction equation?

9.4 Show that the correlation coefficient in simple regression is equivalent to the correlation between Y and $\hat{Y}$.

9.5 Calculate the regression equation for the data of table 9.1 considering the runoff as the independent variable and the precipitation as the dependent variable. Does the resulting regression equation agree with the regression equation in example 9.1? Should it agree? Why? Which equation should be used?

9.6 A technique used by hydrologists to detect changes in the hydrologic response of a watershed is to examine mass curves for changes in slope. A mass curve is a plot of the accumulation over time of one variable versus the accumulation over time of a second variable. The data below are the annual runoff (RO) and precipitation (Prec) for Thorn Creek experimental watershed in Pulaski County, Virginia. It is thought that there was a change in the hydrologic characteristics of this watershed during the 11-year period of study. Plot the accumulated precipitation as the abscissa and the accumulated runoff as the ordinate. Does there appear to be a change in the rainfall-runoff relationship? During what year? Calculate the slope of the regression lines describing the data both before and after the apparent change. Test the hypothesis that these slopes are not significantly different.

Year	Prec	RO	Year	Prec	RO
1958	37.8	10.9	1964	31.1	1.1
1959	33.3	3.9	1965	28.8	0.8
1960	31.5	7.9	1966	39.4	1.4
1961	36.8	2.7	1967	35.2	2.5
1962	43.1	4.4	1968	33.0	2.4
1963	24.0	3.1			

9.7. Occassionally it is desirable to restrict the intercept of a simple regression to zero thus requiring the regression line to pass through the origin. Derive the normal equation for the slope in this case. Use the resulting equation to calculate the slope of the line describing the data plotted for exercise 9.6. Neglect the apparent change in the slope for this problem (i.e., use all of the data to estimate β in the equation accumulated runoff β [accumulated precipitation]).

9.8 Hydrologists frequently use watershed physical characteristics as an aid in studying watershed hydrology. The data below are the area (square miles) and length (miles) of several Colorado mountain watersheds (Julian et al. 1967). Derive a linear regression equation for predicting the area of similar watersheds as a function of the watershed length. Plot the data and the derived regression line. Plot the 95% confidence intervals on the regression line.

Area	Length	Area	Length
47.1	6.7	338.0	35.0
89.1	11.5	86.9	13.6
178.0	20.7	87.9	12.9
129.0	17.5	58.0	10.2
82.6	13.5	69.8	12.0
109.0	17.5	55.9	9.0
41.2	9.0	172.0	20.0
295.0	26.5	96.0	14.9
106.0	15.1	161.0	21.0

CURSO: INGENIERIA MARITIMA. MODULO: "FUNDAMENTOS DE HIDRAULICA Y MATEMATICAS" DEL 11 DE MARZO AL 7 DE JUNIO.
MEXICO, D.F.

MULTIPLE LINEAR REGRESSION.

PROF. ING. JOSE RAYNAL.
JUNIO 1985.

10. Multiple Linear Regression

NOTATION

The notation set forth in Chapter 9 will be followed in this chapter unless otherwise noted. Additionally, vectors and matrices will be denoted by underlined letters such as $\underline{y}$, $\underline{X}$ or $\underline{\beta}$. The inverse of a matrix $\underline{X}$ will be denoted by $\underline{X}^{-1}$. The transpose of $\underline{X}$ will be denoted by $\underline{X}^t$. The number of rows and columns in a matrix will be shown as $\underset{n \times p}{\underline{X}}$ if $\underline{X}$ has n rows and p columns. Thus $\underset{n \times 1}{\underline{Y}}$ represents a column vector with n elements. The element of $\underline{X}$ corresponding to the i^{th} row and the j^{th} column will be denoted by $X_{i,j}$. The expression $\underline{X} = [X_{i,j}]$ indicates that $\underline{X}$ is made up of elements $X_{i,j}$. A matrix made up of elements which are deviations from a mean will be denoted by a lower case, underlined letter $\underline{y}$. The ij^{th} element of $\underline{y}$ will be given as $y_{i,j}$. The i^{th} element of a vector $\underset{n \times 1}{\underline{Y}}$ will be given by Y_i. A brief review of some matrix operations is presented in Appendix D.

The concepts of Chapter 9 must be understood before proceeding to this chapter. Calculations would normally be done on a digital computer for problems dealing with multiple regression. Standard programs are available so the emphasis in this chapter is not on computing but on the principles involved in multiple regression.

GENERAL LINEAR MODEL

Quite often a dependent variable is dependent on several other quantities. For example the peak rate of runoff from watersheds in a given region may be related to the watershed area, slope of the mainstream, rainfall, etc. A model for predicting peak runoff would have to contain all of these variables. In this chapter the linear model discussed in Chapter 9 is extended to include several independent variables.

A general linear model of the form

$$Y = \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_p X_p \qquad (10.1)$$

is discussed where Y is a dependent variable, X_1, X_2, ..., X_p are independent variables

and $\beta_1, \beta_2, \ldots, \beta_p$ are unknown parameters. This model is linear in the parameters, β_j. The model

$$Y = \beta_1 + \beta_2 X_2 + \beta_3 X_3^2 + \beta_4 \ln(X_4)$$

is also linear in the parameters, β_j, whereas the models

$$Y = \beta_1 + \beta_2 X^{\beta_3}$$

$$Y = \beta_1 + \beta_2 X_2 + \beta_3 \exp(\beta_4 X_3)$$

and

$$Y = \beta_1 + \beta_2 X_2 + \beta_2^2 X_2$$

are not linear in the parameters.

In practice n observations would be available on Y with the corresponding n observations on each of the p independent variables. Thus n equations like equation 10.1 can be written, one for each observation. Essentially we will be solving n equations for the p unknown parameters. Thus n must be equal to or greater than p. In practice n should be at least 3 or 4 times as large as p. The n equations are

$$\begin{aligned} Y_1 &= \beta_1 X_{1,1} + \beta_2 X_{1,2} + \ldots + \beta_p X_{1,p} \\ Y_2 &= \beta_1 X_{2,1} + \beta_2 X_{2,2} + \ldots + \beta_p X_{2,p} \\ &\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ Y_n &= \beta_1 X_{n,1} + \beta_2 X_{n,2} + \ldots + \beta_p X_{n,p} \end{aligned} \tag{10.2}$$

where Y_i is the i^{th} observation on Y and $X_{i,j}$ is the i^{th} observation on the j^{th} independent variable. Equations 10.2 can be written

$$Y_i = \sum_{j=1}^{p} \beta_j X_{i,j} \tag{10.3}$$

for i = 1 to n. In matrix notation the equations become

$$\underset{n\times 1}{\underline{Y}} = \underset{n\times p}{\underline{X}}\ \underset{p\times 1}{\underline{\beta}} \tag{10.4}$$

where $\underline{Y}$ is an nx1 vector of observations, $\underline{X}$ is an nxp matrix made up of n observations on each of p independent variables, and $\underline{\beta}$ is a px1 vector of unknown parameters. If the matrices in equation 10.4 are written out, we get

$$\begin{bmatrix} Y_1 \\ Y_2 \\ Y_3 \\ \cdot \\ \cdot \\ Y_n \end{bmatrix} = \begin{bmatrix} X_{1,1} & X_{1,2} & X_{1,3} & \ldots & X_{1,p} \\ X_{2,1} & X_{2,2} & X_{2,3} & \ldots & X_{2,p} \\ X_{3,1} & X_{3,2} & X_{3,3} & \ldots & X_{3,p} \\ \cdot & \cdot & \cdot & & \cdot \\ \cdot & \cdot & \cdot & & \cdot \\ X_{n,1} & X_{n,2} & X_{n,3} & \ldots & X_{n,p} \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \\ \cdot \\ \cdot \\ \beta_p \end{bmatrix} \tag{10.5}$$

When the model is written in the form of equation 10.5, it is easy to see that $\underline{Y}$ is an nx1 vector of observations on the dependent variable, $\underline{X}$ is an nxp matrix made up of n observations on each of p independent variables, and $\underline{\beta}$ is a px1 vector of unknown parameters. For equation 10.4 to have an intercept term, it is necessary that $X_{i,1} = 1$. β_1 is the intercept. In the following development, it is assumed that $X_{i,1} = 1$ for i = 1 to n.

The model discussed in Chapter 9

$$Y = \alpha + \beta X$$

is a special case of equation 10.5 with $X_{i,1} = 1$, $X_{i,2} = X$, $\beta_1 = \alpha$ and $\beta_2 = \beta$.

Following the pattern of Chapter 9, the unknown parameters, $\underline{\beta}$, can be estimated by minimizing $\sum e_i^2$ where $e_i = Y_i - \hat{Y}_i = Y_i - \sum_{j=1}^{p} \hat{\beta}_j X_{i,j}$. In matrix notation

$$\sum e_i^2 = \underline{e}' \underline{e} = (\underline{Y} - \underline{X}\hat{\underline{\beta}})' (\underline{Y} - \underline{X}\hat{\underline{\beta}}) \tag{10.6}$$

Differentiating this expression with respect to $\hat{\underline{\beta}}$ and setting the partial derivative equal to zero results in

$$\underline{0} = -2\underline{X}' (\underline{Y} - \underline{X}\hat{\underline{\beta}})$$

or

$$\underline{X}' \underline{Y} = \underline{X}' \underline{X}\hat{\underline{\beta}} \tag{10.7}$$

which represents the normal equations. The solution of equation 10.7 is obtained by premultiplying by $(\underline{X}' \underline{X})^{-1}$

$$(\underline{X}' \underline{X})^{-1} \underline{X}' \underline{Y} = (\underline{X}' \underline{X})^{-1} (\underline{X}' \underline{X})\hat{\underline{\beta}} = \hat{\underline{\beta}}$$

or we have the result that $\underline{\beta}$ can be estimated by

$$\hat{\underline{\beta}} = (\underline{X}' \underline{X})^{-1} \underline{X}' \underline{Y} \tag{10.8}$$

The $\underline{X}' \underline{X}$ matrix plays an important role in estimating $\underline{\beta}$ and in the variance of the $\hat{\beta}_i$'s. The $\underline{X}' \underline{X}$ matrix is made up of the sum of squares and cross products of the independent variables. For the pxp matrix $\underline{X}' \underline{X}$ to be inverted, its rank must be p.

If we define $z_{i,j}$ to be $(X_{i,j} - \bar{X}_j)/S_j$ and let $\underline{z} = [z_{i,j}]$, then $\underline{z}' \underline{z}/(n-1)$ is a pxp correlation matrix, $\underline{R} = [R_{i,j}]$, where $R_{i,j}$ is the correlation coefficient between the i^{th} and j^{th} independent variables. By definition $R_{i,j} = 1$ for i = j. If $|R_{i,j}| = 1$ for some $i \neq j$, then the i^{th} independent variable is a linear function of the j^{th} independent variable and the rank of $\underline{X}' \underline{X}$ will be less than p. Therefore, if equation 10.8 is to be used to estimate the p unknown parameters $\underline{\beta}$, there can be no linear dependence in $\underline{X}$. This means that an independent variable cannot be a (perfect) linear function of any other independent variable. Furthermore, for the rank of $\underline{X}' \underline{X}$ to be p, an independent variable cannot be linearly dependent on any linear function of the remaining independent variables. For example, if p is 4 and $X_2 = aX_1 + bX_3 + c$, then X_2 is a linear function of X_1 and X_3 so that the rank of $\underline{X}' \underline{X}$ would be at most 3. If there is near linear dependence in $\underline{X}$, the calculation of $(\underline{X}' \underline{X})^{-1}$ may involve roundoff errors and loss of signifi-

cance leading to nonsensical estimates for $\underline{\beta}$ (Draper and Smith 1966).

As in the case of simple regression, the total sum of squares can be partitioned into 3 parts. Draper and Smith (1966) demonstrate that equation 9.11 can be written in matrix notation as

$$\underline{Y}' \underline{Y} = n\bar{Y}^2 + (\hat{\underline{\beta}}' \underline{X}' \underline{Y} - n\bar{Y}^2) + (\underline{Y}' \underline{Y} - \hat{\underline{\beta}}' \underline{X}' \underline{Y}) \tag{10.9}$$

so that the three components of the total sum of squares, $\underline{Y}' \underline{Y}$ or ΣY_i^2, are:

1. $n\bar{Y}^2$, the sum of squares due to the mean,
2. $\underline{Y}' \underline{Y} - \hat{\underline{\beta}}' \underline{X}' \underline{Y} = (\underline{Y} - \underline{X}\hat{\underline{\beta}})' (\underline{Y} - \underline{X}\hat{\underline{\beta}}) = \underline{e}' \underline{e} = \Sigma e_i^2 = \Sigma (Y_i - \hat{Y}_i)^2$, the sum of squares of deviations from regression or the residual sum of squares, and
3. $\hat{\underline{\beta}}' \underline{X}' \underline{Y} - n\bar{Y}^2 = \Sigma (\hat{Y}_i - \bar{Y})^2$, the sum of squares due to regression.

A multiple coefficient of determination, R^2, can now be defined from equation 9.13 as

$$R^2 = \frac{\text{Sum of squares due to regression}}{\text{Sum of squares about the mean}}$$

$$R^2 = (\hat{\underline{\beta}}' \underline{X}' \underline{Y} - n\bar{Y}^2)/(\underline{Y}' \underline{Y} - n\bar{Y}^2) \tag{10.10}$$

As in the case of r^2, the range of R^2 is from 0 to 1. The multiple correlation coefficient is defined as the positive square root of R^2. Again R^2 is the fraction of the total sum of squares corrected for the mean that is explained by the regression equation $\hat{\underline{Y}} = \underline{X}\hat{\underline{\beta}}$.

Quite frequently the partitioning of the sum of squares is shown in the form of an analysis of variance (ANOV) table. A mean square in the ANOV is simply a sum of squares divided by its degrees of freedom.

Continuing the analogy with simple regression, we define $\underline{\varepsilon}$ as $\underline{Y} - \underline{X}\underline{\beta}$. Our estimation procedure guarantees that $E(\underline{\varepsilon}) = 0$. An unbiased estimate for the $\text{Var}(\underline{\varepsilon})$ or σ^2 is S^2 where

$$s^2 = \Sigma e_i^2/(n-p) = \Sigma (Y_i - \hat{Y}_i)^2/(n-p)$$

$$= \underline{e}' \underline{e}/(n-p) = (\underline{Y} - \underline{X}\hat{\underline{\beta}})' (\underline{Y} - \underline{X}\hat{\underline{\beta}})/(n-p) = (\underline{Y}' \underline{Y} - \hat{\underline{\beta}}' \underline{X}' \underline{Y})/(n-p) \tag{10.11}$$

The standard error of the regression equation σ is estimated by S. An expression for R^2 that is analogous to equation 9.18 is

Table 10.1. ANOV for multiple regression.

Source	Degrees of freedom	Sum of squares	Expected mean square
Mean	1	$n\bar{Y}^2$	
Regression	p-1	$\hat{\underline{\beta}}' \underline{X}' \underline{Y} - n\bar{Y}^2$	
Residual	n-p	$\underline{Y}' \underline{Y} - \hat{\underline{\beta}}' \underline{X}' \underline{Y}$	σ^2
Total	n	$\underline{Y}' \underline{Y}$	

$$R^2 = 1 - (n-p)s^2/(n-1)s_Y^2 \tag{10.12}$$

Again this shows that if the regression equation is explaining a large part of the variation in Y, the standard error of the equation will be significantly less than the standard deviation of Y.

Example 10.1. Benson (1962) studied flood frequencies on many streams in the northeastern United States. The following table contains a partial listing of some of Benson's data. Using this data: (1) Estimate the regression coefficients for the model

$$Q = \beta_1 + \beta_2 A + \beta_3 I$$

where Q is the mean annual flood in thousands of cfs, A is the watershed area in thousands of square miles and I is the average annual maximum 24-hour rainfall depth in inches. (2) Calculate R^2. (3) Calculate $\hat{Q}_i$ for each observation on the independent variables. (4) Calculate e_i for each $\hat{Q}_i$.

Station No.	Q	A	I	$\hat{Q}$	e
2	15.50	1.250	1.7	18.115	-2.615
6	8.50	0.871	2.1	13.135	-4.635
7	85.00	5.690	1.9	76.508	8.492
9	105.00	8.270	1.9	110.437	-5.437
12	24.80	1.620	2.1	22.985	1.815
14	3.80	0.175	2.4	3.985	-0.185
22	1.76	0.148	3.2	3.639	-1.879
33	18.00	1.400	2.7	20.099	-2.099
35	8.75	0.297	2.9	5.595	3.155
39	8.25	0.322	2.9	5.924	2.326
45	3.56	0.178	2.8	4.029	-0.469
48	1.90	0.148	2.7	3.634	-1.734
61	16.50	0.872	2.1	13.148	3.352
63	2.80	0.091	2.9	2.886	-0.086

Solution: To maintain consistency in notation, let $Y_i = Q_i$, $X_{i,1} = 1$, $X_{i,2} = A_i$ and $X_{i,3} = I_i$. For this problem n is 14 and p is 3. The column of data under Q is the 14x1 vector Y, a column of 1's along with the data under A and I is the 14x3 matrix $\underline{X}$, and the 3x1 vector $\underline{\beta}$ is made up of β_1, β_2 and β_3.

From equation 10.8, we have

$$\hat{\underline{\beta}} = (\underline{X}' \underline{X})^{-1} \underline{X}' \underline{Y}$$

$$(\underline{X}' \underline{X}) = \begin{vmatrix} \Sigma X_{i,1}^2 & \Sigma X_{i,2} X_{i,1} & \Sigma X_{i,3} X_{i,1} \\ \Sigma X_{i,1} X_{i,2} & \Sigma X_{i,2}^2 & \Sigma X_{i,3} X_{i,2} \\ \Sigma X_{i,1} X_{i,3} & \Sigma X_{i,2} X_{i,3} & \Sigma X_{i,3}^2 \end{vmatrix}$$

$$(\underline{X}' \underline{X}) = \begin{bmatrix} 14.00 & 21.33 & 34.30 \\ 21.33 & 108.741 & 43.34 \\ 34.30 & 43.34 & 86.99 \end{bmatrix}$$

Using the relationship given in Appendix D, $(\underline{X}' \underline{X})^{-1}$ is found to be

$$(\underline{X}' \underline{X})^{-1} = \begin{bmatrix} 3.71678 & -0.18094 & -1.37537 \\ -0.18094 & 0.02028 & 0.06124 \\ -1.37537 & 0.06124 & 0.52329 \end{bmatrix}$$

$$(\underline{X}' \underline{Y}) = \begin{bmatrix} \Sigma Y_i \\ \Sigma X_{i,2}\, Y_i \\ \Sigma X_{i,3}\, Y_i \end{bmatrix} = \begin{bmatrix} 304.1200 \\ 1465.8927 \\ 627.8000 \end{bmatrix}$$

$$= (\underline{X}' \underline{X})^{-1} \underline{X}' \underline{Y} = \begin{bmatrix} 1.6570 \\ 13.1510 \\ 0.0112 \end{bmatrix}$$

The parameter estimates are $\hat{\beta}_1 = 1.6570$, $\hat{\beta}_2 = 13.1510$ and $\hat{\beta}_3 = 0.0112$.

From equation 10.10, we get

$$R^2 = (\underline{\beta}' \underline{X}' \underline{Y} - n\bar{Y}^2) / (\underline{Y}'\underline{Y} - n\bar{Y}^2)$$

$$= [19{,}788.911 - 14(21.7229)^2] / [19{,}960.066 - 14(21.7229)^2]$$

$$= 0.9872$$

This means that 99% of the variation in Y is explained by the regression equation

$$\hat{Q} = 1.6570 + 13.1510\, A + 0.0112\, I$$

Values for $\hat{Q}$ contained in the above table were calculated from this relationship. Values for e_i were computed from

$$e_i = Y_i - \hat{Y}_i = Q_i - \hat{Q}_i$$

and are also contained in the above tabulation.

The ANOV table for this example would be

Source	d.f.	SS	MS
Mean	1	6,606.381	
Regression	2	13,182.600	6591.300
Residual	11	171.090	15.554
Total	14	19,960.066	

From the ANOV table $R^2 = 13{,}182.600/(19{,}960.066 - 6{,}606.381) = 0.99$ and $s^2 = 15.554$ or the estimated standard error of the regression equation is $s = 3.94$.

Comment: The purpose of this example is to demonstrate the meaning of the various matrices and to provide practice in their calculation. Hydrologic significance should not be attached to the high R^2 since the watersheds are all close to one another (Maine) and the units on Q are cfs and the watershed area is contained in the equation. Many of the gaging stations are located at various points along the same stream.

The number of significant figures that are carried in the calculations should be as large as practical. In reporting the results, the number of significant figures should be reduced. Thus the reported results on the above regression might be

$$Q = 1.66 + 13.15\, A + 0.011\, I$$

If a large number of significant figures are not carried in computing the $(\underline{X}' \underline{X})^{-1}$ matrix, significant errors can result. To demonstrate this, the elements of the $\underline{X}' \underline{X}$ and $\underline{X}' \underline{Y}$ matrices were rounded to two decimal places resulting in estimates for $\underline{\beta}$ of $\hat{\beta}_1 = 1.10$, $\hat{\beta}_2 = 12.24$ and $\hat{\beta}_3 = 5.28$.

CONFIDENCE INTERVALS AND TESTS OF HYPOTHESES

As was the case in simple regression, in order to use some well-developed theorems on confidence intervals and tests of hypotheses in multiple regression, some assumptions must be made. All of the comments of Chapter 9 regarding the assumptions in simple regression remain valid in multiple regression. The assumption will now be made that the ε_i are identically, independently and normally distributed with mean 0 and variance σ^2. (See "General Comments" section of Chapter 9.)

Confidence Intervals on Standard Error

Confidence intervals can be placed on σ^2 by noting that the quantity $(n - p)S^2/\sigma^2$ has a chi-square distribution. Thus the confidence limits on σ^2 are

$$L = \frac{(n-p)S^2}{\chi^2_{1-\alpha/2,\,n-p}} \qquad U = \frac{(n-p)S^2}{\chi^2_{\alpha/2,\,n-p}} \tag{10.13}$$

Inferences on the Regression Coefficients

To make inferences concerning β, the variance of $\hat{\beta}$ must be estimated. The variance-covariance matrix of $\hat{\beta}$ is given by

$$\mathrm{Cov}(\hat{\underline{\beta}}) = E\{[\hat{\underline{\beta}} - E(\hat{\underline{\beta}})][\hat{\underline{\beta}} - E(\hat{\underline{\beta}})]'\}$$

$E(\hat{\underline{\beta}}) = \underline{\beta}$, $\hat{\underline{\beta}} = (\underline{X}' \underline{X})^{-1} \underline{X}' \underline{Y}$ and $\underline{Y} = \underline{X}\underline{\beta} + \underline{\varepsilon}$

So that

$$\hat{\underline{\beta}} = (\underline{X}'\underline{X})^{-1}\underline{X}'(\underline{X}\underline{\beta} + \underline{\varepsilon})$$

$$= (\underline{X}'\underline{X})^{-1}(\underline{X}'\underline{X})\underline{\beta} + (\underline{X}'\underline{X})^{-1}\underline{X}'\underline{\varepsilon}$$

$$= \underline{\beta} + (\underline{X}'\underline{X})^{-1}\underline{X}'\underline{\varepsilon}$$

or

$$\hat{\underline{\beta}} - \underline{\beta} = (\underline{X}'\underline{X})^{-1}\underline{X}'\underline{\varepsilon}$$

and

$$\mathrm{Cov}(\hat{\underline{\beta}}) = E[(\hat{\underline{\beta}} - \underline{\beta})(\hat{\underline{\beta}} - \underline{\beta})']$$

$$= E[(\underline{X}'\underline{X})^{-1}\underline{X}'\underline{\varepsilon}\,\underline{\varepsilon}'\underline{X}(\underline{X}'\underline{X})^{-1}]$$ [1]

Again treating the $\underline{X}$ as fixed (see "Inferences on Regression Coefficients", Chapter 9), we have

$$\mathrm{Cov}(\hat{\underline{\beta}}) = (\underline{X}'\underline{X})^{-1}\underline{X}'E(\underline{\varepsilon}\,\underline{\varepsilon}')\underline{X}(\underline{X}'\underline{X})^{-1}$$

Since $\underline{\varepsilon}$ is an nx1 matrix, $\underline{\varepsilon}\,\underline{\varepsilon}'$ is an nxn matrix.

$$\underline{\varepsilon}\,\underline{\varepsilon}' = |\varepsilon_i\varepsilon_j|$$

$$E(\underline{\varepsilon}\,\underline{\varepsilon}') = |E(\varepsilon_i\varepsilon_j)|$$

Since ε_i is i.i.d. $N(0,\sigma^2)$, we have

$$E(\varepsilon_i\varepsilon_j) = \sigma^2 \quad \text{for } i = j$$

$$= 0 \quad \text{for } i \neq j$$

Therefore $E(\underline{\varepsilon}\,\underline{\varepsilon}')$ is a symmetric matrix with all of the diagonal elements (i = j on the diagonal) equal to σ^2 and all of the off diagonal elements equal to 0.

$$E(\underline{\varepsilon}\,\underline{\varepsilon}') = \begin{bmatrix} \sigma^2 & 0 & \dots & 0 \\ 0 & \sigma^2 & \dots & 0 \\ . & . & \dots & . \\ 0 & 0 & \dots & \sigma^2 \end{bmatrix} = \sigma^2\,\underline{I}$$

where I is the nxn identity matrix. We now have

$$\mathrm{Cov}(\hat{\underline{\beta}}) = (\underline{X}'\underline{X})^{-1}\underline{X}'\sigma^2\underline{I}\,\underline{X}(\underline{X}'\underline{X})^{-1}$$

1. $(\underline{X}'\underline{X})^{-1}$ is a symmetric matrix. A symmetric matrix is equal to its transpose.

or

$$\mathrm{Cov}(\hat{\underline{\beta}}) = \sigma^2(\underline{X}'\underline{X})^{-1} \tag{10.14}$$

The variance of $\hat{\beta}_i$ is equal to the covariance of $\hat{\beta}_i$ with itself and is therefore σ^2 times the i^{th} diagonal element of $(\underline{X}'\underline{X})^{-1}$. The covariance of $\hat{\beta}_i$ with $\hat{\beta}_j$ is σ^2 times the i,jth element of $(\underline{X}'\underline{X})^{-1}$. If we let $\underline{C} = \underline{X}'\underline{X}$ then $\underline{C}^{-1} = (\underline{X}'\underline{X})^{-1}$ and

$$\mathrm{Var}(\hat{\beta}_i) = \sigma^2_{\hat{\beta}_i} = C^{-1}_{ii}\sigma^2 \tag{10.15}$$

where C^{-1}_{ii} is the i^{th} diagonal element of $(\underline{X}'\underline{X})^{-1}$.

If the model is correct, then the quantity $\hat{\beta}_i/S_{\hat{\beta}_i}$ is distributed as a t distribution with n - p degrees of freedom where $S_{\hat{\beta}_i}$ is an estimate for $\sigma_{\hat{\beta}_i}$ and is calculated as the positive square root of $C^{-1}_{ii}S^2$.

Confidence intervals on β_i are given by

$$L_{\beta_i} = \hat{\beta}_i - t_{1-\alpha/2,n-p}\,S_{\hat{\beta}_i}$$

$$U_{\beta_i} = \hat{\beta}_i + t_{1-\alpha/2,n-p}\,S_{\hat{\beta}_i} \tag{10.16}$$

A test of the hypothesis $\beta_i = \beta_0$ where β_0 is a known constant can be made by noting that $(\hat{\beta}_i - \beta_0)/S_{\hat{\beta}_i}$ has a t distribution. Thus, to test H_o: $\beta_i = \beta_0$ versus H_a: $\beta_i \neq \beta_0$, the test statistic

$$t = \frac{\hat{\beta}_i - \beta_0}{S_{\hat{\beta}_i}} \tag{10.17}$$

is computed. H_o is reject if $|t| > t_{1-\alpha/2,n-p}$.

Since in general $\hat{\beta}_i$ is not independent of $\hat{\beta}_j$ (their covariance is given by $C^{-1}_{ij}\sigma^2$), repeated application of equation 10.17 to test H_o: $\beta_i = \beta_{0_i}$ and H_o: $\beta_j = \beta_{0_j}$ are not independent tests.

A test of H_o: $\beta_i = 0$ versus H_a: $\beta_i \neq 0$ is equivalent to testing the hypothesis that the i^{th} independent variable is not contributing significantly to explaining the variation in the dependent variable. If H_o: $\beta_i = 0$ is accepted, many times it is advisable to delete the i^{th} independent variable from the model.

A test of the hypothesis that the entire regression equation is not explaining a significant amount of the variation in Y is equivalent to H_o: $\beta_2 = \beta_3 = \dots = \beta_p = 0$ versus H_a: at least one of these β's is not zero. Since β_i is not independent of β_j, repeated application of equation 10.17 is not a valid way to test this hypothesis. Use can be made of the fact that the ratio of the mean square due to regression to the residual mean square has an F distribution with p - 1 and n - p degees of freedom. To test H_o: $\beta_2 = \beta_3 = \dots = \beta_p = 0$, calculate the test statistic

$$F = \frac{(\hat{\underline{\beta}}'\underline{X}'\underline{Y}) - n\bar{Y}^2)/(p-1)}{(\underline{Y}'\underline{Y} - \hat{\underline{\beta}}'\underline{X}'\underline{Y})/(n-p)} \tag{10.18}$$

and reject H_o if F exceeds $F_{1-\alpha,p-1,n-p}$.

A test of the hypothesis that k of the independent variables are not contributing significantly to explaining the linear variation in the dependent variable can be made.

To do this we arrange the model so that the last k variables are the ones to be tested. The hypothesis is that the last k independent variables are not contributing significantly to explaining the linear variation in Y.[2] This is equivalent to H_o: $\beta_{p-k+1} = \beta_{p-k+2} = \ldots = \beta_p = 0$ versus H_a: at least one of these β's is not zero. To test H_o, denote the full model as the model containing all p of the independent variables. Denote as the reduced model the model obtained by deleting the last k independent variables. The reduced model contains p - k independent variables. Now let

Q_2 = sum of squares due to regression on the full model with p - 1 degrees of freedom.
Q_1 = residual sum of squares on the full model with n - p degrees of freedom.
Q_2^* = sum of squares due to regression on the reduced model with p - k - 1 degrees of freedom.

The quantity

$$F = \frac{(Q_2 - Q_2^*)/k}{Q_1/(n-p)} \tag{10.19}$$

will have an F distribution with k and n - p degrees of freedom. H_o is rejected if F exceeds $F_{1-\alpha,k,n-p}$.

Note that $Q_2 - Q_2^*$ is the reduction in the sum of squares due to regression brought about by deleting k independent variables. If Q_2^* nearly equals Q_2, then the deletion of the k variables has not greatly changed the ability of the model to explain the linear variation in Y. Under these conditions F will be small and H_o will not be rejected indicating that one might eliminate the last k variables from further consideration. Rejection of H_o does not imply that all of the last k variables are important - it only implies that at least one of these variables is explaining a significant amount of the variation in Y.

Confidence Intervals on the Regression Line

To place confidence limits on $\hat{Y}_h$ where $Y_h = \underline{X}_h \underline{\beta}$ it is necessary to have an estimate for the variance of $\hat{Y}_h$. In this discussion $\hat{Y}_h$ is an estimate of Y (a scalar) at the point $\underline{X}_h$ (a 1xp vector) in p dimensional space. $\hat{\underline{\beta}}$ is a px1 vector consisting of the estimates for $\underline{\beta}$. The $\mathrm{Var}(\hat{Y}_h)$ is given by (Draper and Smith 1966)

$$\mathrm{Var}(\hat{Y}_h) = \sigma^2 \underline{X}_h (\underline{X}' \underline{X})^{-1} \underline{X}_h' \tag{10.20}$$

which can be estimated by replacing σ^2 with s^2. The confidence limits on Y_h are given by

$$L = \underline{X}_h \hat{\underline{\beta}} - t_{1-\alpha/2,n-p} \sqrt{\mathrm{Var}(\hat{Y}_h)}$$
$$U = \underline{X}_h \hat{\underline{\beta}} + t_{1-\alpha/2,n-p} \sqrt{\mathrm{Var}(\hat{Y}_h)} \tag{10.21}$$

The confidence intervals on an individual predicted value of Y_h are given by equations

2. In practice the model does not have to be so arranged. The order of the X's makes no difference. The assumption here is the last k variables are under test. This makes the notation easier.

10.21 where $\mathrm{Var}(\hat{Y}_h)$ is replaced by the variance of an individual predicted value of Y at $\underline{X}_h$ which is given by $\sigma^2(1 + \underline{X}_h(\underline{X}'\underline{X})^{-1}\underline{X}_h')$.

Other Inferences in Regression

Many other tests of hypotheses can be made and confidence intervals constructed relative to multiple regression. For example, one might make tests concerning linear relationships among the β's or that the β's obtained from one situation are equal to those obtained from another situation. Reference can be made to Graybill (1961), Johnston (1963) or Draper and Smith (1966) for these and other tests.

Example 10.2. For the regression equation of example 10.1: (1) Test the hypothesis that the regression equation is not explaining a significant amount of the variation of Y. (2) Test the H_o: $\beta_2 = 0$. (3) Test the H_o: $\beta_3 = 0$. (4) Calculate the 95% confidence limits on β_2. (5) Calculate the 95% confidence limits on the regression line at the point A = 4,000 square miles and I = 2.0 inches. (6) Calculate the 95% confidence intervals on σ^2.

Solution:

(1) This H_o is equivalent to H_o: $\beta_2 = \beta_3 = 0$ versus H_a: at least one of β_2 or $\beta_3 \neq 0$. The test is conducted by calculating the test statistic from equation 10.18. The quantities in equation 10.18 are contained in the ANOV table with the numerator being the mean square due to regression and the denominator being the residual mean square.

$$F = 6591.300/15.554 = 424$$

The tabulated $F_{.95,2,11}$ is 3.98. Therefore, H_o is rejected. The regression equation is explaining a significant amount of the variation in Y.

(2) H_o: $\beta_2 = 0$ $\qquad$ H_a: $\beta_2 \neq 0$

The test statistic is from equation 10.17.

$$t = 13.1510/s_{\hat{\beta}_2}$$
$$s_{\hat{\beta}_2} = (C_{22}^{-1} s^2)^{1/2} = [.02028(15.554)]^{1/2} = 0.562$$
$$t = 13.1510/0.562 = 23.4$$
$$t_{.975,11} = 2.201$$

Reject H_o. Area is explaining a significant amount of the variation in Y.

(3) H_o: $\beta_3 = 0$ $\qquad$ H_a: $\beta_3 \neq 0$

The test statistic is again from equation 10.17

$$s_{\hat{\beta}_3} = (C_{33}^{-1} s^2)^{1/2} = [1.5233(15.554)]^{1/2} = 2.85$$

$t = 0.0112/2.85 = 0.004$

Since $|t| < t_{.975,11}$, we cannot reject H_o. The mean annual maximum 24-hour rainfall depth is not explaining a significant amount of the variation in the mean annual peak flow.

(4) The 95% confidence limits on β_2 are calculated from equations 10.15 as

$l = 13.1510 - 2.201(0.562) = 11.91$

$u = 13.1510 + 2.201(0.562) = 14.39$

(5) The 95% confidence limits on the regression line at $X_{2,h} = 4.00$ and $X_{3,h} = 2.0$ are determined from equation 10.21. The $Var(\hat{Y}_h)$ is from equation 10.20.

$$Var(\hat{Y}_h) = 15.554\, \underline{X}_h\, (\underline{X}'\underline{X})^{-1}\, \underline{X}_h'$$

$(\underline{X}'\underline{X})^{-1}$ is given in example 10.1.

$\underline{X}_h = (1.0, 4.0, 2.0)$

$\underline{X}_h (\underline{X}'\underline{X})^{-1} \underline{X}_h' = 0.16529$

$Var(\hat{Y}_h) = 2.571$

$l = \underline{X}_h \hat{\underline{\beta}} - t_{1-\alpha/2,n-p}\, [Var(\hat{Y}_h)]^{1/2}$

$\underline{X}_h \hat{\underline{\beta}} = 1.6570 + 13.1510(4) + 0.0112(2) = 54.28$

$l = 54.28 - 2.201(1.60) = 50.76$

$u = 54.28 + 2.201(1.60) = 57.80$

(6) The 95% confidence intervals on σ^2 are calculated from equation 10.13

$l = (n-p)s^2/\chi^2_{1-\alpha/2,n-p} = 11(15.554)/21.9 = 7.81$

$u = (n-p)s^2/\chi^2_{\alpha/2,n-p} = 11(15.554)/3.82 = 44.79$

The 95% confidence intervals on σ would be 2.80 to 6.69.

Comment: The hypothesis H_o: $\beta_2 = 0$ and H_o: $\beta_3 = 0$ were both tested in this example as though the tests were independent. In fact β_2 and β_3 are not independent. The Cov $(\hat{\beta}_2, \hat{\beta}_3)$ can be determined from $C^1_{23}\, s^2$ as $.0612(15.554) = 0.9519$. The correlation between $\hat{\beta}_2$ and $\hat{\beta}_3$ can be estimated from $Cov(\hat{\beta}_2, \hat{\beta}_3)/\sigma_{\hat{\beta}_2}\, \sigma_{\hat{\beta}_3}$ as $.9519/.562(2.85) = 0.59$. The test of H_o: $\beta_3 = 0$ is made relative to the full model which includes all of the β's. The acceptance of H_o implies that $\beta_3 = 0$ given that β_1 and β_2 are in the model. In general, if there are p β's and H_o: $\beta_i = 0$ is tested for each of them with the result that k of the hypotheses can be accepted, one cannot eliminate these k variables from the model on the basis of this test alone since each of the individual H_o: $\beta_i = 0$ assumes

all of the other p - 1 β's are still in the model. To eliminate k variables at once, the test must be based on equation 10.19.

As an example of the application of equation 10.19, the H_o: $\beta_3 = 0$ will be tested. The ANOV for the full model is contained in example 10.1. The reduced model is simply $Y = \beta_1 + \beta_2 X$ where X is the watershed area in 1000's of square miles. Since this is a simple regression situation, we can compute the sum of squares due to regression from $b\Sigma x_i y_i$ where $b = \Sigma x_i y_i / \Sigma x_i^2$. The result of this calculation is the sum of squares due to regression for the reduced model which is 13,182.60.

The test statistic from equation 10.19 is

$$F = \frac{(Q_2 - Q_2^*)/k}{Q_1/(n-p)} = \frac{(13{,}182.60 - 13182.60)/1}{171.090/11} = 0.00$$

$F_{.95,1,11} = 4.84$ so we accept H_o: $\beta_3 = 0$. Note that this test is identical to the test conducted in part 3 of this example. From tables E.5 and E.7 it can be seen that $F_{1-\alpha,1,n} = t^2_{1-\alpha/2,n}$ so for the special case where k = 1 variable is being tested against zero, equations 10.17 and 10.19 produce identical results.

Since H_o: $\beta_3 = 0$ was accepted, the next logical step is to eliminate 1 from the model and consider only A. In so doing the resulting regression equation is

$Q = 1.69 + 13.15A$

The dependence of β's again is evident since the intercept is not the same as was obtained when rainfall depth was included in the model. This is a somewhat special example in that β_2 accounts for nearly all of the variation in Y leaving virtually none of the variation to be explained by β_3. Again one reason for this unusual situation is the units on Y and A and the proximity of all of the watershed.

WHICH LINE IS BEST

A common situation in which multiple regression is used is when one dependent variable and several independent variables are available and it is desired to find a linear model for predicting unobserved values for the dependent variable. The model that is developed does not necessarily have to contain all of the independent variables. Thus the points of concern are: (1) can a linear model be used and (2) what independent variables should be included?

A factor complicating the selection of the model is that in most cases the independent variables are not statistically independent at all but are correlated. One of the first steps that should be done in a regression analysis is to compute the correlation matrix, $\underline{R}$, of the independent variables. The correlation matrix can be computed as follows. Let

$$z_{i,j} = (X_{i,j} - \bar{X}_j)/S_j \qquad (10.22)$$

where $\bar{X}_j$ and S_j are the mean and standard deviation of the j^{th} independent variable. Then define $\underline{z} = [z_{i,j}]$ so that the correlation matrix is

$$\underline{R} = \underline{z}'\,\underline{z}/(n-1) = [R_{i,j}] \qquad (10.23)$$

where $R_{i,j}$ is the correlation between the i^{th} and j^{th} independent variables. $\underline{R}$ is a sym-

metric matrix since $R_{i,j} = R_{j,i}$. We have already seen that if $R_{i,j} = 1$ for $i \neq j$, then either variable i or variable j must be omitted from the model or else the $\underline{X}' \underline{X}$ matrix cannot be inverted. If $R_{i,j}$ is close to unity (but not equal to unity), then $\underline{X}' \underline{X}$ can be inverted and $\underline{\beta}$ estimated. If $R_{i,j}$ is close to unity then the $\text{Var}(\hat{\beta}_i)$ or $\text{Var}(\hat{\beta}_j)$ may be very large. Tests of hypothesis on β_i and β_j may indicate that neither is significantly different from zero when in fact either β_i or β_j when used alone may be significantly different from zero. The problem here is that since X_i and X_j are nearly linearly related, they both are attempting to explain the same thing in the linear model. By having both X_i and X_j in the model, the part of the variation in Y that either would explain if used alone may be split between them in such a fashion that neither is significant. In other words the effect of one causative factor (which may be reflected in either X_i or X_j) is being divided between two correlated variables.

Retaining variables in a regression equation that are highly correlated makes the interpretation of the regression coefficients difficult. Many times the sign of the regression coefficient may be the opposite of what is expected if the corresponding variable is highly correlated with another independent variable in the equation.

A common practice in selecting a multiple regression model (and one that is not necessarily being advocated) is to perform several regressions on a given set of data using different combinations of the independent variables. The regression that "best" fits the data is then selected. A commonly used criterion for the "best" fit is to select the equation yielding the largest value of R^2.

Looking at equations 10.21 and 10.22, another and perhaps better criterion is apparent. The confidence intervals on the regression line are a function of S, the estimated standard error. The line with the smallest standard error will have the narrowest confidence intervals.

Many times the two criteria of the largest R^2 and the smallest S give the same results but not always. As more variables are added to a regression equation, the R^2 value can never decrease. Thus from the standpoint of the R^2 criterion, one should use all of the variables he can lay his hands on. This, however, makes a clumsy equation and one in which it is extremely difficult to place a meaningful interpretation on the coefficients.

As more variables are added to a regression equation, the standard error may get larger. This can be seen from equation 10.11. Every time a variable is added, n - p gets smaller as does $\underline{Y}' \underline{Y} - \hat{\underline{\beta}}' \underline{X}' \underline{Y}$. However, the numerator may not and many times does not decrease proportionally to n - p so that as variables are added S may actually increase. This is a tip-off that the added variables are not contributing significantly to the regression and can just as well be left out.

All of the variables retained in a regression should make a significant contribution to the regression unless there is an overriding reason (theoretical or intuitive) for retaining a non-significant variable. The variables retained should have physical significance. If two variables are equally significant when used alone but are not both needed, the one that is easiest to obtain should be used.

The number of coefficients estimated should not exceed 25 to 35 percent of the number of observations. This is a rule of thumb used to avoid "over-fitting" whereby oscillations in the equation may occur between observations on the independent variables.

Thus far all decisions on which regression equation to use have been made by the investigator. In many cases this is the most reliable method of selecting a regression equation. Since the advent of the digital computer, it has been possible to perform many regressions on large sets of data. This has led to several formal procedures for selecting

a regression equation. Two methods will be discussed here – all-possible-regressions and stepwise regression. For a discussion of some other techniques, reference should be made to Draper and Smith (1966).

All-possible-regressions involves calculating regression equations having every possible combination of the X variables. If all of the equations are required to have an intercept term, 2^{p-1} regression equations would have to be calculated where p is the number of independent variables one of which is always equal to one to produce the intercept term. Thus if p = 4, 8 regression equations would be calculated (not an impossible task or a bad procedure); however, if p = 11, 1024 regressions would have to be calculated and examined. Thus as p gets even moderately large, the number of regressions required becomes prohibitive and intelligent thought could eliminate many of them. When this many regressions are calculated, the probability of getting a significant regression by chance becomes large.

One of the most commonly used procedures for selecting the "best" regression equation is stepwise regression. This procedure consists of building the regression equation one variable at a time by adding at each step the variable that explains the largest amount of the remaining unexplained variation. After each step all the variables in the equation are examined for significance and discarded if they are no longer explaining a significant amount of the variation. Thus the first variable added is the one with the highest simple correlation with the dependent variable. The second variable added is the one explaining the largest variation in the dependent variable that remains unexplained by the first variable added. At this point the first variable is tested for significance and retained or discarded depending on the results of this test. The third variable added is the one that explains the largest portion of the variation that is not explained by the variables already in the equation. The variables in the equation are then tested for significance. This procedure is continued until all of the variables not in the equation are found to be insignificant and all of the variables in the equation are significant. This is a very good procedure to use but care must be exercised to see that the resulting equation is rational.

Of course, the real test of how good the resulting regression model is depends on the ability of the model to predict the dependent variable for observations on the independent variables that were not used in estimating the regression coefficients. To make a comparison of this nature, it is necessary to randomly divide the data into two parts. One part of the data is then used to develop the model and the other part to test the model. Unfortunately, many times in hydrologic applications, there are not enough observations to carry out this procedure.

EXTRAPOLATION

The comments on extrapolation contained in chapter 9 relative to simple regression are equally applicable to multiple regression. In multiple regression an additional problem arises. It is sometimes difficult to tell the range of the data. For example in example 10.1, A ranges from 0.091 to 8.27 and I ranges from 1.7 to 3.2. Is the point A = 6.0 and I = 2.7 in the range of the data?

A plot of A and I for example 10.1 is shown in figure 10.1. From this plot it is apparent that A and I do not cover the entire range defined by $0.091 < A < 8.27$ and $1.7 < I < 3.2$. The point A = 6.0 and I = 2.7 does not appear to be in the range of the data. In more than 2 dimensions it is much more difficult to visualize the range of the data.

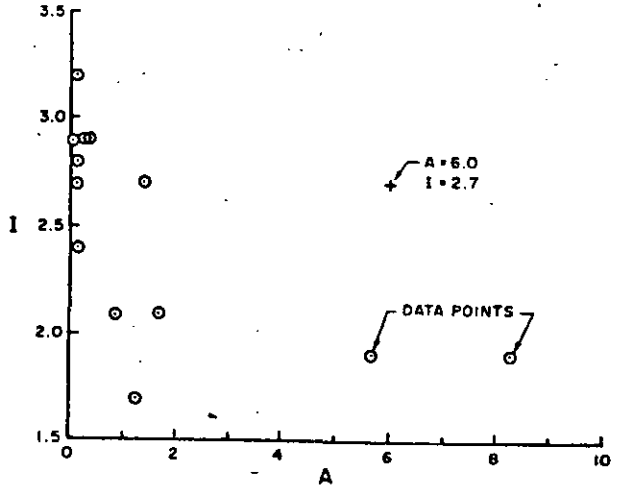


Fig. 10.1. Range of data used in example 10.1.

AN APPLICATION OF MULTIPLE REGRESSION

The following illustration of using multiple regression is adapted from Haan and Read (1970). A part of their study was devoted to developing a prediction equation for the mean annual runoff for small watersheds in Kentucky. The data for the example is contained in table 10.2. The number of observations (13) is very small and does not permit splitting the sample and using a portion of the data for testing the resulting model. Table 10.3 contains definitions of the symbols used in table 10.2. The correlation matrix for the independent variables is contained in table 10.4. Since the correlation matrix is symmetrical, it is customary to show only the diagonal elements and the elements either above or below the diagonal.

Table 10.2. Data from Haan and Read (1970).

Watershed No.	RO	Prec.	A	S	L	P	d_i	R_s	F	R_r
1	17.38	44.37	2.21	50	2.38	7.93	0.91	0.38	1.36	332
2	14.62	44.09	2.53	7	2.55	7.65	1.23	0.48	2.37	55
3	15.48	41.25	5.63	19	3.11	11.61	2.11	0.57	2.31	77
4	14.72	45.50	1.55	6	1.84	5.31	0.94	0.49	3.87	68
5	18.37	46.09	5.15	16	4.14	11.35	1.63	0.39	3.30	68
6	17.01	49.12	2.14	26	1.92	5.89	1.41	0.71	1.87	230
7	18.20	44.03	5.34	7	4.73	12.59	1.30	0.27	0.94	44
8	18.95	48.71	7.47	11	4.24	12.33	2.35	0.52	1.20	72
9	13.94	44.43	2.10	5	2.00	6.81	1.19	0.53	4.76	40
10	18.64	47.72	3.89	18	2.10	9.87	1.65	0.60	3.08	115
11	17.25	48.38	0.67	21	1.15	3.93	0.62	0.48	2.99	352
12	17.48	49.00	0.85	23	1.27	3.79	0.83	0.61	3.53	300
13	13.16	47.03	1.72	5	1.93	5.19	0.99	0.52	2.33	39

Table 10.3. Definition of symbols used by Haan and Read (1970).

RO	Mean annual runoff (inches)
Prec	Mean annual precipitation (inches)
A	Area (square miles)
S	Average land slope (%)
L	Axial length (miles)
P	Perimeter (miles)
d_i	Diameter of largest circle that can be drawn entirely within the basin (miles)
R_s	Shape factor – ratio of d_i to d_o where d_o is the diameter of the smallest circle that can be drawn which entirely encloses the basin (-)
F	Stream frequency – ratio of number of streams in basin to total area of basin (square miles)$^{-1}$
R_r	Relief ratio – ratio of total relief to largest dimension of basin generally parallel to main stream (feet per mile)

Table 10.4. Correlation matrix for data of Haan and Read (1970).

	A	S	L	P	d_i	R_s	F	R_r
A	1.00							
S	-.17	1.00						
L	.90	-.21	1.00					
P	.96	-.10	.92	1.00				
d_i	.91	-.16	.67	.81	1.00			
R_s	-.25	.05	-.58	-.41	.15	1.00		
F	-.48	-.30	-.53	-.48	-.32	.29	1.00	
R_r	-.52	.80	-.54	-.51	-.50	.18	-.08	1.00

The mean and standard deviation of RO are 16.55 and 1.93 inches respectively. Table 10.5 contains the results of the multiple regression of RO on all 9 of the independent variables. Since an intercept term was included p is equal to 10. In the ANOV table, the sum of squares for the mean and the total sum of squares are not shown. Instead the total sum of squares corrected for the mean is given. The F that is given is the calculated F for the overall regression equation (from equation 10.18) used in testing the hypothesis that the regression is not explaining a significant amount of the variation in Y. Since $F_{.95,9,3}$ is 8.81, this hypothesis is rejected.

The lower part of table 10.5 contains the estimated regression coefficients, the standard errors of the regression coefficients and the calculated t (equation 10.17) used in testing H_o: $\beta_i = 0$. The only β's with calculated t's greater than 2.0 are those based on Prec, P, and R_r. If all of the variables except these three and the intercept are eliminated at one time, the regression shown in table 10.6 results. In going to the second regression R^2 has been reduced from 0.97 to 0.91, the F increased to 28.7, and the standard error has remained unchanged. All of the regression coefficients with the exception of the intercept are now significantly different from zero at the one percent level of significance since $t_{.99,5,9}$ is 3.25.

Table 10.5. Regression analysis of data of Haan and Read (1970). (10 independent variables).

Analysis of Variance

Source	Degrees of Freedom	Sum of Squares	Mean Square
Regression	9	43.45	4.83
Residual	3	1.44	0.48
Total corrected for mean	12	44.89	

$R^2 = 0.97$ R = 0.98
F = 10.0 Std. Error = 0.69

Variable	$\hat{\beta}$	$s_{\hat{\beta}}$	t
Constant	-14.770	6.810	-2.17
Prec	0.450	.150	3.04
A	0.170	1.430	0.12
S	-0.020	0.050	-0.37
L	0.290	0.780	0.38
P	0.990	0.370	2.68
d_i	-3.020	4.810	-0.63
R_s	5.640	9.060	0.62
F	0.370	0.270	1.39
R_r	0.013	0.006	2.13

The t test used to test the hypothesis that $\beta_i = 0$ makes the test assuming that all of the other β's are still in the equation. Thus when a decision is made to eliminate more than one variable, the t's are unreliable and the F test using equation 10.9 should be used. This test determines if several variables are simultaneously making a significant contribution to explaining the variation in the dependent variable. As an illustration of the use of equation 10.14, the hypothesis that $\beta_A = \beta_S = \beta_L = \beta_{d_i} = \beta_{R_s} = \beta_F = 0$ will be tested. For this example n = 13, p = 10, k = 6, $Q_2 = 43.45$, $Q_2^* = 40.64$ and $Q_1 = 1.44$. The F calculated from equation 10.19 is 0.98. Since $F_{.95,6,3} = 8.94$, it is concluded that the variables A, S, L, d_i, R_s and F are not significant.

The resulting prediction model is

$$RO = -9.65 + 0.43\ Prec + 0.62\ P + 0.010\ R_r$$

The observed values of runoff and values predicted from the above equation are shown in the lower half of table 10.6.

To demonstrate the behavior of s, R^2 and F, several regressions were run using various combinations of the data in table 10.2. The results of these regressions are summarized in table 10.7 and figure 10.2. This table illustrates that R^2 never increases as variables are removed from the equation while S may decrease as some variables are removed and then increase as more variables are removed. R^2 is approaching unity as the number of variables is increasing. If the number of variables were increased to 12, then p would be 13 (since the model has an intercept) and R^2 would be unity. In figure 10.2 the lines connect the best values of the quantities S, R^2 and F contained in table 10.7. This is because it is possible, for example, to have many combinations of 3 variables in the regression equation and each combination would produce a different S, R^2 and F.

Table 10.6. Regression analysis of data of Haan and Read (1970) (4 independent variables).

Analysis of Variance

Source	Degrees of Freedom	Sum of Squares	Mean Square
Regression	3	40.64	13.55
Residual	9	4.25	.47
Total	12	44.89	

$R^2 = 0.91$ R = 0.95
F = 28.7 Std. Error = 0.69

Variable	$\hat{\beta}$	$s_{\hat{\beta}}$	t
Constant	-9.650	4.440	-2.17
Prec	0.430	0.093	4.62
P	0.620	0.075	8.25
R_r	0.010	0.002	5.19

Observed and predicted mean runoff values

WATERSHED	MEAN RUNOFF (IN.) OBSERVED	PREDICTED	DIFFERENCE (IN.)
Bear Branch	17.38	17.76	0.38
Cave Creek	14.63	14.59	-0.04
Flat Creek	15.48	16.04	0.56
Plum Creek No. 4	14.72	13.89	-0.83
Little Plum Creek	18.37	17.86	-0.51
McGills Creek	17.01	17.47	+0.46
McDougal Creek	18.20	17.49	-0.71
West Bays Fork	18.95	19.64	+0.69
Rose Creek	13.94	14.06	+0.12
Wood Creek	18.64	18.14	-0.50
Cane Branch	17.24	17.23	-0.01
Helton Branch	17.48	16.87	-0.61
Perry Creek	13.16	14.17	+1.01

Table 10.7. Some results of several regressions on S, R^2 and F.

Eq. No.	Variables Included[1]									s	R^2	F
	Prec	A	S	L	P	d_i	R_s	F	R_r			
1	*	x	x	x	*	x	x	x	*	0.69	0.97	10.0
2	*	x	x		*	x	x	x	*	0.62	0.97	14.3
3	*	x	x		*		x	x	*	0.59	0.96	17.5
4	*	x			*		x	x	*	0.59	0.95	20.3
5	*				*		x	x	*	0.60	0.94	23.7
6	*				*			x	*	0.64	0.93	25.2
7	*				*				*	0.69	0.91	28.6
8	*		*	*						1.13	0.74	8.7
9	*			*					*	1.06	0.78	10.4
10	*		*		*					0.98	0.81	12.5
11	*				*					1.30	0.62	8.2
12	x									1.86	0.15	1.92
13					*				*	1.20	0.68	10.7

1. * indicates variable included in regression and was significant.
x indicates variable included in regression and was not significant.

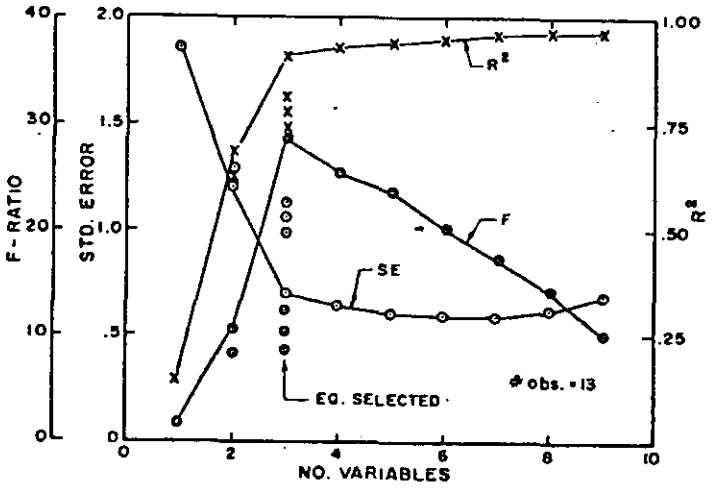


Fig. 10.2. Behavior of s, R^2, and F as a function of n for data continued in table 10.2.

TRANSFORMING NONLINEAR MODELS

Many models are not naturally linear models but can be transformed to linear models. For example

$$Y = \alpha X^{\beta} \tag{10.24}$$

is not a linear model. It can be linearized by using a logarithmic transformation

$$\ln Y = \ln \alpha + \beta \ln X \tag{10.25}$$

or

$$Y' = \alpha' + \beta' X' \tag{10.26}$$

where

$$\begin{aligned} Y' &= \ln Y \\ \alpha' &= \ln \alpha \\ \beta' &= \beta \\ X' &= \ln X \end{aligned} \tag{10.27}$$

Standard regression techniques can now be used to estimate α' and β' for equation 10.26 and α and β estimated from equations 10.27. Two important points should be noted. First the estimates of α and β obtained in this way will be such that $\Sigma(Y_i' - \hat{Y}_i')^2$ is a minimum and not such that $\Sigma(Y_i - \hat{Y}_i)^2$ is a minimum. Second the error term on equation 10.26 is additive ($Y' = \alpha' + \beta' X' + \epsilon'$) implying that it is multiplicative on equation 10.24 ($Y = \alpha X^{\beta} \epsilon$). These errors are related by $\epsilon' = \ln \epsilon$. The assumptions used in hypothesis testing and confidence intervals must now be valid for ϵ' and the tests and confidence intervals made relative to the transformed model.

In some situations the logarithmic transformation makes the data conform more closely to the regression assumptions. For example if the data plot as in figure 10.3, a logarithmic transformation may make the assumption of constant variance on the error more realistic.

The normal equations for a logarithmic transformation are based on a constant percentage error along the regression line while the standard regression is based on a constant absolute error along the regression line. For example the difference between $Y_i =$

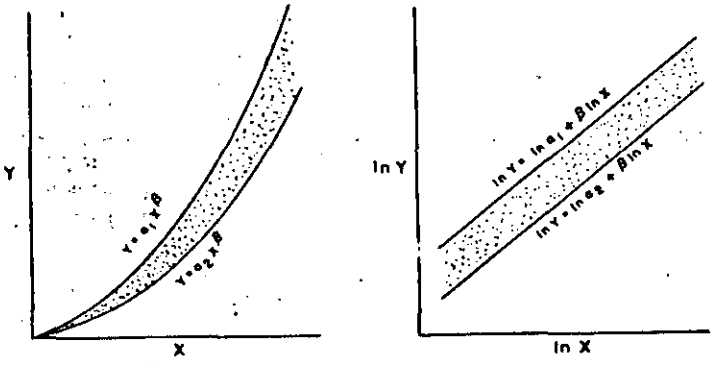


Fig. 10.3 Example of the effect of a logarithmic transformation on the error variance.

200 and $Y_i = 100$ on an arithmetic scale is 100 times as large as the difference between $Y_i = 2$ and $Y_i = 1$. However on a logarithmic scale ln 200 - ln 100 = 5.29832 - 4.60517 = .69315 which is the same as ln 2 - ln 1 = .69315 - .000 = .69315. In a situation of this type, the standard regression procedure would attempt to fit the point at $Y_i = 100$ in order to minimize $\Sigma(Y_i - \hat{Y}_i)^2$ at the expense of the point $Y_i = 1$ since its contribution to $\Sigma(Y_i - \hat{Y}_i)^2$ is small. The logarithmically transformed model would give equal percentage weight to both points.

The above discussion can be extended to the model

$$Y = \beta_o X_1^{\beta_1} X_2^{\beta_2} \dots X_p^{\beta_p}$$

through the transformation

$$\ln Y = \beta_o + \beta_1 \ln X_1 + \beta_2 \ln X_2 + \dots + \beta_p \ln X_p$$

Other models and transformations are available. For example

$$Y = a \exp(bX)$$

can be transformed to

$$\ln Y = \ln a + bX$$

Yevjevich (1972a) lists several possible transformations. Whatever the transformation, it must be remembered that the principles of and assumptions regarding least squares apply to the tranformed model, not the original model.

GENERAL COMMENTS

Regression analysis should be regarded simply as a tool for exploiting linear tendencies that may exist between a dependent variable and a set of independent variables. It is also a useful devise for estimating the parameters of a model that is linear or can be transformed to a linear model.

Any regression analysis should be preceded by a great deal of thought devoted to what variables should be included in the analysis, how these variables might influence the dependent variable, the correlations among the independent variables, and the ease of using a predictive model based on the selected independent variables. The ready availability of digital computers and library regression programs has led many to collect data with little thought, throw it into the computer, and hope for a model. This temptation must be avoided.

Not infrequently an investigator finds he cannot develop a satisfactory regression equation from the data at hand. This should not be surprising since if a relationship exists, it may be much more complex than is indicated by a linear model. Commonly factors that are important in determining the behavior of a dependent variable are omitted from a regression equation. In this case a good predictive model cannot be expected.

In some regression problems, it is possible to improve the model by including cross-product terms (called interactions) by multiplying together two independent variables to form a new variable. Thus a variable X_q may be defined as $X_r X_s$. Ratios may be used

such as $X_q = X_r / X_s$. Powers of variables may improve the model $X_q = X_r^n$ where n is a known constant. If any of these procedures are used, care must be exercised to see that large correlations (see chapter 11) are not built into the $\underline{X}' \underline{X}$ matrix.

The hydrologist frequently knows (or thinks he does anyway) the factors affecting a particular phenomena. He cannot, however, easily measure these factors and is forced to use either another related factor or a rough measure of the important factor. For instance flood peaks depend among other things on how rapidly the surface runoff reaches a particular point on a stream. This in turn depends on surface flow characteristics such as the steepness and roughness of the flow surface and the distance the flow must travel to a stream channel plus the stream flow characteristics such as roughness, hydraulic radius, slope, tortuosity, length, etc. All of the factors are not linearly related to flood peaks and could not be included in the model if they were. Indices or summaries are used such as the average land slope and the average channel slope. The hydrologist hopes that the real causative factors are correlated with these indices sufficiently to reflect this true importance and that the dependent variable is linearly related to the indices. These are large and important assumptions. They point out that there is a limit to how well one can predict a dependent variable with a regression model.

If it is at all possible, the first step in a regression analysis should be the development of the form of the predictive model based on a rational analysis of the problem. Regression analysis can then be used to develop the parameters of the model, test the importance of the variables included, and develop confidence intervals for the predictions.

Exercises

10.1 Use the matrix methods of this chapter to work example 9.1.

10.2 Compute R for example 10.1.

10.3 Use the matrix methods of this chapter to work example 9.4.

10.4 Use the matrix methods of this chapter to work example 9.5. Calculate the confidence interval for the point X equals 50.0 inches of rainfall.

10.5 If $\underline{X}$ is an nxp matrix of n observations on p variables, and $\underline{Z}$ is the nxp matrix of deviations of the variables from their means, what is contained in the matrix $\underline{Z}' \underline{Z}/(n-1)$? ($\underline{Z} = [z_{ij}] = [x_{ij} - \bar{x}_j]$)

10.6 Use the data in table 10.2 to develop a prediction equation for annual runoff using the model $Y = \beta_o X_1^{\beta_1} X_2^{\beta_2} \dots X_p^{\beta_p}$. Would you prefer this equation over the one contained in table 10.6? Compare the equations in terms of the confidence interval on the regression lines at the mean values of the variables contained in the respective equations.

10.7 Show that $\Sigma(Y - \hat{Y})^2 = \underline{Y}' \underline{Y} - \hat{\beta}' \underline{X}' \underline{Y}$.

10.8 Derive the normal equations that minimize $\Sigma(Y - \hat{Y})^2$ for the model $Y = a X^{\beta}$. Suggest a method for solving these equations.

10.9 The relationship between stage and discharge (rating curve) for many streams has

been found to follow an equation of the type $Q = \alpha S^{\beta}$ where Q is the discharge and S is the stage. Using the following data from the Cumberland River at Cumberland Falls, Kentucky, derive such a rating curve. Test the hypothesis that $\beta = 1.5$.

S(ft)	Q(cfs)	S(ft)	Q(cfs)
15.50	59,600	8.82	22,600
14.10	51,900	8.30	20,400
14.90	56,100	11.70	38,600
14.15	52,300	11.40	37,200
14.10	54,200	11.27	36,300
13.50	50,100	10.30	30,900
14.55	57,400	10.90	34,200
12.96	46,500	10.70	33,100
12.60	43,500	10.07	28,900
12.21	41,600	9.48	25,800
12.47	43,300	8.80	22,600
12.58	44,100	7.77	18,200
12.67	44,800		

10.10 The data in table 10.8 is a partial listing of the data used by Benson (1964) in a study of floods in the Southwest. Derive a prediction equation for Q_p, the mean annual flood, in terms of the remaining variables. Consider both the models given by equation 10.1 and by the multiple regression extension of equation 10.24.

Table 10.8. Independent variables, by station, in rain-flood area.

A, contributing drainage area in square miles.
S, main-channel slope (85 to 10 percent points), in feet per mile.
St, percentage of area in lakes and ponds, increased by 1 percent.
E, altitude index (mean of 85 and 10 percent points), in feet above mean sea level.
L, basin length (total length of main channel), in miles.
H, basin rise (elevation difference between 85 and 10 percent points), in feet.
P, mean annual precipitation, in inches.
I, 10-year, 24-hour rainfall intensity in inches.
R, ratio of runoff to precipitation during months when annual peak discharges occur.
R_a, mean annual runoff, in inches.
Q_p, mean annual flood, cfs.

A	S	St	E	L	H	P	I	R	R_a	Q_p
131	1.61	1.66	40	30.7	37	60.0	7.49	0.86	26.40	4500
527	2.16	2.39	44	48.8	79	60.5	7.34	.96	20.40	7080
499	3.36	1.07	182	49.6	125	58.5	7.01	.81	21.10	20600
753	2.52	1.06	142	82.9	157	59.5	7.10	.73	22.40	18700
510	5.11	1.02	156	48.2	185	60.0	7.19	.68	23.00	15200
238	4.58	1.03	144	41.7	143	59.0	7.36	.68	23.00	12200
1,700	2.43	1.04	118	99.5	181	60.0	7.19	.78	22.20	34400
148	4.67	1.04	88	34.6	121	58.0	7.54	.91	19.30	6700
82	6.27	1.18	96	22.6	106	58.5	7.75	.74	21.00	4870
1,445	3.17	1.19	438	80.8	192	42.0	6.01	.69	9.88	32300
586	4.18	1.02	429	41.2	129	43.0	6.10	.77	10.20	16900
236	6.57	1.11	392	36.5	180	44.5	6.26	.64	11.60	4290
2,846	2.20	1.21	389	127.0	209	43.0	6.14	.62	9.48	22000
4,858	1.25	1.27	300	235.0	221	44.5	6.33	.65	9.00	22500
113	6.17	1.12	225	21.6	100	51.0	6.94	1.00	12.20	4210
6,543	1.13	1.20	265	292.0	248	46.5	6.45	.68	11.50	26600
114	7.22	1.03	262	21.6	117	56.0	7.21	.66	18.50	7500
9,440	.96	1.18	203	427.0	308	49.0	6.78	.69	12.70	43500
1,143	2.29	1.21	357	88.6	152	43.5	6.30	.65	9.09	9620
1,943	1.57	1.17	306	146.0	172	44.0	6.37	.62	8.17	11100
2,714	1.25	1.16	266	214.0	201	44.5	6.50	.62	8.59	14400
3,623	1.29	1.13	232	253.0	245	45.5	6.61	.63	8.95	16700
376	4.27	2.16	344	36.6	117	43.0	6.37	.56	10.10	5890
1,604	1.80	1.33	258	114.0	153	45.0	6.62	.64	10.60	10700
501	2.76	1.19	246	65.9	136	48.0	6.84	.54	13.10	9000

11. Correlation

IN CHAPTER 3 the population correlation coefficient between two random variables X and Y was defined in terms of the covariance of X and Y and the variances of X and Y as

$$\rho_{X,Y} = \sigma_{X,Y} / \sigma_X \sigma_Y \tag{11.1}$$

The sample estimate $r_{X,Y}$ for $\rho_{X,Y}$ is similarly given by

$$r_{X,Y} = s_{X,Y} / s_X s_Y \tag{11.2}$$

where $s_{X,Y}$ is the sample covariance between X and Y and s_X and s_Y are the sample standard deviations of X and Y respectively. Figure 3.5 and the accompanying description discussed some typical values for $r_{X,Y}$ and their meaning. Here it was emphasized that (1) $\rho_{X,Y}$ can range from -1 to 1, (2) $\rho_{X,Y} = \pm 1$ implies a perfect linear relationship between X and Y, (3) $\rho_{X,Y} = 0$ implies linear independence but leaves room for other types of dependence, and (4) if X and Y are independent then $\rho_{X,Y} = 0$.

In chapters 9 and 10 the concept of correlation was extended to give a measure of the strength of the linear relationship between a random variable Y and a second variable which was a linear function of one or more X variables each of which may or may not be a random variable.

Throughout the text many of the results that have been developed have included the assumption that the random variables were independent or that the sample being analyzed was composed of random observations. A random observation simply means that every possible element in the sample space has an equal chance of being selected during any trial.

Random variables may be either uncorrelated ($\rho_{X,Y} = 0$) or correlated ($\rho_{X,Y} \neq 0$). Even when sampling from uncorrelated populations, it would be rare for the sample correlation coefficient to be exactly zero. More likely it will deviate from zero due to chance. Thus statistical tests are needed to evaluate whether the deviation of the sample

correlation coefficient from zero may be ascribed to chance or whether the deviation is too large to attribute to chance.

If successive observations in a time series of hydrologic data are correlated, this must be taken into account in any inferences made about the data or in attempts to model the process that produced the data. Again a procedure is required for determining if the sampled elements from a time series can be considered as random. These and other properties of correlation are the subject of this chapter.

INFERENCES ABOUT POPULATION CORRELATION COEFFICIENTS

Situations frequently arise where it is desired to test H_o: $\rho_{X,Y} = 0$ or H_o: $\rho_{X,Y} = \rho^*$ where ρ^* is known. These and other tests about the population correlation coefficient will be discussed in this section. For a more detailed treatment, reference can be made to Graybill (1961).

As in the case of all hypothesis tests, certain assumptions are needed. In this section the assumption is made that X and Y are random variables from a bivariate normal distribution. The population correlation coefficient is given by ρ and the sample estimate of ρ given by r is based on a random sample.

If $\rho = 0$, then the quantity

$$t = r[(n - 2)/(1 - r^2)]^{1/2} \tag{11.3}$$

has a t distribution with n - 2 degrees of freedom where n is the sample size. Thus to test H_o: $\rho = 0$, the test statistic is calculated from equation 11.3 and H_o is rejected if $|t| > t_{1-\alpha/2,n-2}$.

If n is moderately large ($n \geq 25$) then the quantity W is approximately normally distributed with mean ω and variance $(n - 3)^{-1}$ where

$$W = \tfrac{1}{2} \ln[(1 + r)/(1 - r)] = \operatorname{arctanh} r \tag{11.4}$$

and

$$\omega = \tfrac{1}{2} \ln[(1 + \rho)/(1 - \rho)] = \operatorname{arctanh} \rho \tag{11.5}$$

To test the hypothesis H_o: $\rho = \rho^*$ against the alternative H_a: $\rho \neq \rho^*$ for ρ^* a known constant, the quantity

$$z = (W - \omega)(n - 3)^{1/2} \tag{11.6}$$

can be considered to be normally distributed with a mean of zero and a variance of one. If $|z| > z_{1-\alpha/2}$ (z is the standard normal variable), H_o is rejected.

Confidence limits on ρ can be estimated from

$$l = \tanh[W - z_{1-\alpha/2} / (n - 3)^{1/2}]$$
$$u = \tanh[W + z_{1-\alpha/2} / (n - 3)^{1/2}] \tag{11.7}$$

Consider k bivariate normal populations having population correlation coefficients of $\rho_1, \rho_2, \ldots, \rho_k$ and sample correlation coefficients of $r_1, r_2, \ldots, r_k$ based on samples

of size $n_1, n_2, \ldots, n_k$. Then the hypothesis $H_o: \rho_1 = \rho_2 = \ldots = \rho_k = \rho^*$ for ρ^* a known constant is tested by noting that

$$\chi^2 = \Sigma_{i=1}^{k} (\text{arctanh } r_i - \text{arctanh } \rho^*)^2 (n_i - 3) \tag{11.8}$$

has a chi-square distribution with k degrees of freedom. H_o is rejected if $\chi^2 > \chi^2_{1-\alpha,k}$.

The hypothesis $H_o: \rho_1 = \rho_2 = \ldots = \rho_k$ (all correlation coefficients are equal) is tested by noting that

$$\chi^2 = \Sigma_{i=1}^{k} (n_i - 3)(W_i - \overline{W})^2 \tag{11.9}$$

has a chi-square distribution with k - 1 degrees of freedom. In equation 11.9, W_i is given by equation 11.4 as $W_i = \text{arctanh } r_i$ and

$$\overline{W} = \Sigma_{i=1}^{k} (n_i - 3)W_i / \Sigma_{i=1}^{k} (n_i - 3) \tag{11.10}$$

H_o is rejected if $\chi^2 > \chi^2_{1-\alpha,k-1}$.

If the hypothesis that all of the correlation coefficients are equal is not rejected, it may be desirable to calculate a "best" combined estimate $\bar{r}$ of the common correlation ρ ("best" means weighted with inverse variance). Such an estimate is given by

$$\bar{r} = \tanh (\overline{W} - m\rho^*/2) \tag{11.11}$$

where

$\overline{W}$ is given by equation 11.10

$$\rho^* = \Sigma_{i=1}^{k} r_i / k$$

and

$$m = \Sigma_{i=1}^{k} [(n_i - 3)/(n_i - 1)] / \Sigma_{i=1}^{k} (n_i - 3)$$

Example 11.1. Burges and Johnson (1973) presents the following sample correlation coefficients for monthly flow volumes for the Sauk River in Washington and Arroyo Seco in California. In the following table r_j represents the sample correlation coefficient between the monthly flow volumes in months j and j - 1. Assume the coefficients are based on 30 observations each and that the parent populations are all bivariate normal (Burges and Johnson actually used the lognormal distribution in their study). (1) Test the hypothesis that ρ_s for the Sauk River is equal to 0.50. (2) Compute the 95% confidence limits for ρ_s of the Sauk River. (3) Test the hypothesis that ρ_s on Arroyo Seco is zero. (4) Test the hypothesis that on each of the streams all of the monthly correlation coefficients are equal. (5) Assume the hypothesis in part 4 is accepted for the Sauk River and estimate an average correlation coefficient for the Sauk River.

Solution:

(1) $H_o: \rho_s = 0.5$ for Sauk River
From equation 11.6

Month	j	Sauk River	r_j	Arroyo Seco
October	1	0.61		0.00
November	2	.58		.00
December	3	.50		.00
January	4	.31		.45
February	5	.38		.21
March	6	.37		.70
April	7	.44		.60
May	8	.34		.75
June	9	.17		.98
July	10	.65		.97
August	11	.93		.96
September	12	.51		.00

$$z = (W - \omega)(n - 3)^{1/2}$$

where

$W = \text{arctanh } r = \text{arctanh } (.34) = .35409$

$\omega = \text{arctanh } \rho = \text{arctanh } (.50) = .54931$

$z = (.35409 - .54931)(27)^{1/2} = -1.015$

$z_{1-\alpha/2} = z_{.975} = 1.96$

Since $|z| < 1.96$ we cannot reject $H_o: \rho_s = 0.5$ for the Sauk River.

(2) The 95% confidence limits on ρ_s for the Sauk River are calculated from equation 11.7 as

$$l = \tanh[W - z_{1-\alpha/2} / (n - 3)^{1/2}]$$

$$= \tanh[.35409 - 1.96/5.2] = -0.0228$$

$$u = \tanh[W + z_{1-\alpha/2} / (n - 3)^{1/2}]$$

$$= \tanh[.35409 + 1.96/5.2] = 0.624$$

(3) $H_o: \rho_s = 0$ for Arroyo Seco is tested by using equation 11.3.

$$t = r[(n - 2)/(1 - r^2)]^{1/2} = 0.21 (28/0.9559)^{1/2} = 1.1386$$

$$t_{1-\alpha/2,n-2} = t_{0.975,28} = 2.05$$

Therefore we cannot reject $H_o: \rho_s = 0$.

(4) The test, H_o: all ρ_j are equal, is tested by using equation 11.9.

$$\chi^2 = \Sigma_{i=1}^{k} (n_i - 3)(W_i - \overline{W})^2$$

i	Sauk River r_i	Sauk River W_i=arctanh r_i	Arroyo Seco r_i	Arroyo Seco W_i=arctanh r_i
1	0.61	0.71	0.00	0.00
2	.58	0.66	.00	0.00
3	.50	0.55	.00	0.00
4	.31	0.32	.45	0.49
5	.38	0.40	.21	0.21
6	.37	0.39	.70	0.87
7	.44	0.47	.60	0.69
8	.34	0.35	.75	0.97
9	.17	0.17	.98	2.30
10	.65	0.78	.97	2.09
11	.93	1.66	.96	1.95
12	.51	0.56	.00	0.00

$$\overline{W} = \Sigma(n_i - 3)\, W_i / \Sigma(n_i - 3) = (n - 3)\Sigma W_i/(n - 3)k = \Sigma W_i/k$$

Sauk River $\overline{W} = 0.585$

Arroyo Seco $\overline{W} = 0.798$

$$\chi^2 = \Sigma(n_i - 3)\,(W_i - \overline{W})^2 = 27\,\Sigma(W_i - \overline{W})^2 = 27[\Sigma W_i^2 - 12\,\overline{W}^2]$$

Sauk River $\chi^2 = 27[5.707 - 12(.585)^2] = 43.208$

Arroyo Seco $\chi^2 = 27[15.919 - 12(.798)^2] = 223.489$

$$\chi^2_{1-\alpha,k-1} = \chi^2_{.95,11} = 19.67$$

Therefore H_o is rejected for both the rivers.

(5) An average correlation coefficient for the Sauk River is calculated from equation 11.11.

$$\bar{r} = \tanh(\overline{W} - m\rho^*/2)$$

where

$$\overline{W} = 0.585$$

$$m = \Sigma[(n_i - 3)/(n_i - 1)]/\Sigma(n_i - 3)$$

$$= (27/29)/27 = 1/29$$

$$\rho^* = \Sigma r_i/k$$

$$\bar{r} = \tanh(.585 - 0.483/2(29)) = \tanh(.577) = .52$$

Comment: In parts (4) and (5) of this problem several simplifications were made in the summations since n_i was equal to 30 for all i. In general this cannot be done.

Graybill (1961) presents the exact probability distribution of r and states that for small samples, the exact distribution should be used in hypothesis testing. References to tables that aid in hypothesis testing for small samples and examples of their use are also given.

Again it is emphasized that the above tests are based on a random sample from multivariate normal distributions. Even under these conditions only the test of H_o: $\rho = 0$ conducted using equation 11.3 is "exact". The other tests are approximate with the approximation improving as the sample size increases.

For non-normal populations, it may be possible to transform the variables to a normal situation and then apply the above tests to the transformed data. If a transformation of a non-normal random variable is not possible or not desired, then the above tests must be considered as approximate with the approximation becoming poorer as the coefficient of skew of the random variables increase.

SERIAL CORRELATION

It is not uncommon to find in a time series of hydrologic data that an observation at one time period is correlated with the observation in the preceding time period. Such correlation is termed serial correlation or autocorrelation. By definition the elements of a sample of data possessing serial correlation are not random elements. A serially correlated sample of size n contains less information about a process than a completely random sample of size n. In a serially correlated sample, part of the information contained in each observation is already known through its correlation with the preceding observation.

Serial correlation can also exist between an observation at one time period and an observation k time periods earlier for k = 1, 2, ... In this discussion of serial correlation, it is assumed that observations are equally spaced in time and that the statistical properties of the process do not change with time (stationary process). The population serial correlation coefficient is denoted by $\rho(k)$ (and frequently called the autocorrelation coefficient) where k is the lag or number of time intervals between the observations being considered. The sample serial correlation coefficient will be given by r(k). The sample serial correlation coefficient for a sample of size n is given by

$$r(k) = \frac{\Sigma_{i=1}^{n-k} x_i x_{i+k} - \Sigma_{i=1}^{n-k} x_i \Sigma_{i=1}^{n-k} x_{i+k}/(n-k)}{[\Sigma_{i=1}^{n-k} x_i^2 - (\Sigma_{i=1}^{n-k} x_i)^2/(n-k)]^{1/2}\,[\Sigma_{i=1}^{n-k} x_{i+k}^2 - (\Sigma_{j=1}^{n-k} x_{i+k})^2/(n-k)]^{1/2}} \qquad (11.12)$$

From equation 11.12 is is seen that r(0) is unity. That is, the correlation of an observation with itself is one. Equation 11.12 also demonstrates that as k increases, the number of pairs of observations used in estimating $\rho(k)$ decreases since all of the summations contain n - k terms. Serial correlation should only be estimated for k considerably less than n.

If $\rho(k) = 0$ for all $k \neq 0$, the process is said to be a purely random process. This indicates that all of the observations in a sample will be independent of each other. Chapter 14, Yevjevich (1972b), Matalas (1966, 1972b), Julian (1967) and others treat hydrologic time series in more detail.

Anderson (1942) has proposed a test of significance for the serial correlation coef-

ficient for a circular, normal, stationary time series. A circular series is one that closes on itself so that x_n is followed by x_1. Under these assumptions

$$r(k) = \frac{\sum_{i=1}^{n} x_i x_{i+k} - n\bar{x}^2}{(n-1)s_x^2} \qquad (11.13)$$

Although the assumption of a circular series is unrealistic, values of r(k) from equation 11.13 will not differ greatly from those calculated from equation 11.12 if n is large in comparison to k. Under these conditions r(k) will be approximately normally distributed with mean $-1/(n-1)$ and variance $(n-2)/(n-1)^2$ if $\rho(k) = 0$. The confidence limits on $\rho(k)$ are then estimated by

$$l = (-1 - z_{1-\alpha/2}\sqrt{n-2})/(n-1)$$
$$u = (-1 + z_{1-\alpha/2}\sqrt{n-2})/(n-1) \qquad (11.14)$$

If the calculated r(k) falls outside these confidence limits, the hypothesis that $\rho(k)$ is zero (H_o: $\rho(k) = 0$ versus H_a: $\rho(k) \neq 0$) is rejected.

Example 11.2. Frequently in the analysis of runoff volumes, one finds there is significant serial correlation caused by storages on the watershed. Appendix C contains a listing of the monthly and annual runoff volumes for Cave Creek near Lexington, Kentucky. Test the hypothesis that $\rho(1) = 0$ for the annual runoff volumes.

Solution: This solution assumes $\alpha = 0.05$ and is based on equation 11.14 and therefore assumes that the annual runoff is normally distributed and is a stationary time series. Furthermore $\rho(1)$ is estimated from equation 11.13 assuming that the series is circular (in this case this is equivalent to assuming $x_{n+1} = x_1$ in calculating r(1)).

$$r(1) = [\Sigma x_i x_{i+1} - (\Sigma x_i)^2/18]/[\Sigma x_i^2 - (\Sigma x_i)^2/18]$$

$$r(1) = (3831.1777 - 3864.08)/(4246.15 - 3864.08) = -0.086$$

$$l = (-1 - z_{.975}\sqrt{16})/17$$

$$= (-1 - 1.96\sqrt{16})/17 = -0.520$$

$$u = (-1 + 1.96\sqrt{16})/17 = 0.402$$

Since $-0.520 < r(1) < 0.402$, H_o: $\rho(1) = 0$ is not rejected.

Comment: From the width of the confidence interval, it is apparent that the above test is not very powerful for small samples. A sample of around 400 observations would be required to reject H_o: $\rho(k) = 0$ if $r(k) = \pm 0.1$.

Matalas (1967b) has suggested that for hydrologic data $\rho(1)$ tends to be greater than zero due to persistence caused by storage. If $\rho(1)$ is found to be less than zero, it is in many cases difficult to explain hydrologically. In this case one might take $\rho(1)$ as

equal to zero.

Matalas and Langbein (1962) state that in an autocorrelated series, each observation represents part of the information contained in the previous observation. They discuss stationary time series having $\rho(1) \neq 0$ and $\rho(i) = 0$ for $i = 2,3,...$ They state that n observations of a nonrandom series having $\rho(1)>0$ give only as much information (measured in terms of a variance) about the mean as some lesser number, n_e, of observations in a purely random time series.

This lesser number of observations is called the effective number of observations and is given by

$$n_e = n\{[1 + \rho(1)]/[1 - \rho(1)] - 2\rho(1)[1 - \rho^n(1)]/n[1 - \rho(1)]^2\}^{-1} \qquad (11.15)$$

If $\rho(1) = 0$, then $n_e = n$. If $\rho(1)>0$, then $n_e<n$. Equation 11.15 is expressed graphically as figure 11.1. As an example, a 50-year record for which $\rho(1) = 0.2$ contains only as much information about the mean as a 33-year record with $\rho(1) = 0$. Note that if n is large or $\rho(1)$ small, the second term in the brackets of equation 11.15 can be neglected with little loss in accuracy.

CORRELATION AND REGIONAL ANALYSIS

Matalas and Langbein (1962), Yevjevich (1972a), Alexander (1954) and others demonstrate that the information relative to estimating the regional mean contained in data from n stations in a region having an average interstation correlation of $\bar{\rho}$ is equivalent to the information contained in n' uncorrelated stations in the region where n' is given by

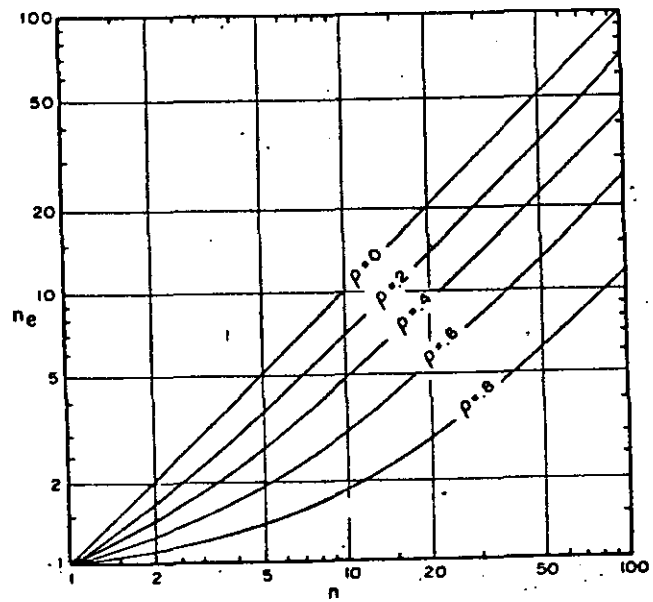


Fig. 11.1. Relation between n and n_e for various values of $\rho(1)$ (after Matalas an Langbein 1962).

$$n' = n[1 + \bar{\rho}(n - 1)]^{-1} \tag{11.16}$$

As n gets large, n' approaches $1/\bar{\rho}$. For a $\bar{\rho}$ of 0.2, the maximum information about the regional mean contained in n stations could not exceed the information contained in 5 uncorrelated stations.

From a consideration of equation 11.16, it seems it would be logical to establish relatively few independent hydrologic stations in a region rather than several correlated stations. However, by the very concept of a hydrologic region, the hydrologic characteristics are going to be correlated.

Correlation within a region can be exploited to yield improved estimates of a particular hydrologic variable at a point through correlation with another hydrologic variable at that point or a similar characteristic at another point. For instance let Y and X represent two random hydrologic variables having no serial correlation for which n_1 and $n_1 + n_2$ observations, respectively are available. Also consider that Y and X are correlated with a correlation coefficient of $\rho_{Y,X}$. Now the record on Y can be extended by using the correlation between Y and X. This relation is merely a simple regression considering Y as the dependent and X the independent variable. The relation is developed based on the n_1 common observations. From equation 9.16, it can be shown that the regression between Y and X is given by

$$y = r_{Y,X}\, s_Y\, x / s_x \tag{11.17}$$

where $r_{Y,X}$ is the estimate for $\rho_{Y,X}$ and y and x are deviations from their respective means. Now n_2 estimates of Y can be computed from equation 11.17 based on the n_2 observations on X not common to the observations on Y. Let $\bar{Y}_1$ and $\bar{Y}_2$ represent the mean of Y based on the original n_1 observations and the n_2 estimated observations respectively. A new weighted mean for Y based on $n_1 + n_2$ observations can now be computed from

$$\bar{Y} = (n_1 \bar{Y}_1 + n_2 \bar{Y}_2)/(n_1 + n_2) \tag{11.18}$$

For the n_2 additional observations to improve the estimate of $\bar{Y}$, it is necessary that $r_{Y,X}$ be greater than $1/(n_1 - 2)$ (Matalas and Langbein 1962).

If the random variables Y and X contain significant serial correlation, the situation is somewhat more complex. Matalas and Langbein (1962), Matalas and Rosenblatt (1962) and Yevjevich (1972a) contain treatments of this case. In general serial correlation serves to decrease the information relative to the mean while cross-correlation tends to improve information relative to the mean.

CORRELATION AND CAUSE AND EFFECT

At this point it should be apparent that a high correlation between two variables does not necessarily imply that there is a cause and effect relation between the variables. The fact that the monthly flows on adjacent small streams are correlated does not mean that changes in the monthly flow of one stream causes a corresponding change in the other stream. More likely both changes are caused by the same external factors operating on both watersheds.

Again it is emphasized that independent variables are uncorrelated but uncorrelated variables are not necessarily independent. The dependence in correlated variables is a

stochastic dependence and not a physical or cause and effect dependence.

SPURIOUS CORRELATION

Spurious correlation is any apparent correlation between variables that are in fact uncorrelated. Spurious correlation can arise due to clustering of data. For example, in figure 11.2, the correlation of Y with X within either of the data clusters is near zero. When the data from both clusters are used to calculate a single correlation coefficient, this correlation is found to be quite high. This is spurious correlation.

Figure 11.3 shows a plot of Y versus X where both Y_i and X_i are random variables obtained by adding 11 to a random observation from a standard normal distribution. For a sufficiently large sample $r_{X,Y}$ would be zero. If both Y_i and X_i are divided by yet a third random observation Z_i obtained in the same manner as X_i and Y_i and the correlation between Y_i/Z_i and X_i/Z_i computed, for a sufficiently large sample the correlation will be near 0.5. Figure 11.4 is a plot of Y_i/Z_i versus X_i/Z_i. Figure 11.4 indicates that X_i furnishes information useful in estimating Y_i when in fact Y_i and X_i are uncorrelated.

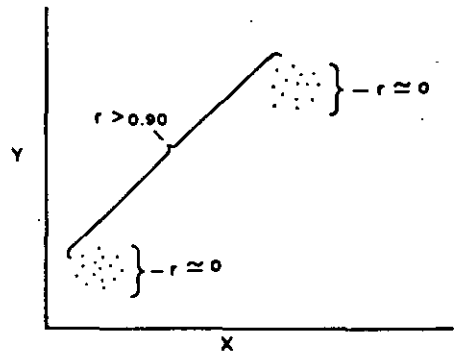


Fig. 11.2. Spurious correlation due to data clustering.

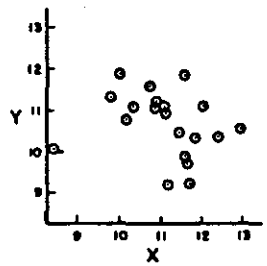


Fig. 11.3. Absence of correlation between two random variables.

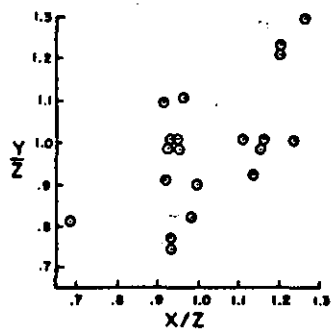


Fig. 11.4. Spurious correlation introduced by dividing 2 random variables by a common third random variable.

Table 11.1. Special Cases of Spurious Correlation*

SPURIOUS CORRELATION: RATIOS

$$\left[\text{Mean of ratio } x_1/x_2 \cong (\bar{X}_1/\bar{X}_2)(1 + C_2^2 - r_{12}C_1C_2)\text{ ; standard deviation of } x_1/x_2 \cong \bar{X}_1/\bar{X}_2\,(C_1^2 + C_2^2 - 2r_{12}C_1C_2)^{1/2}\right]$$

Model	(1)	(2) Graph	(3) r_{yz} (approx.)	Special cases (4) Given	(5) r_{yz} (approx.)	(6) Relative values of C	(7) r_{yz} (approx.)
I	General case of ratio correlation	$y = x_1/x_2$; $z = x_3/x_4$	$\dfrac{r_{13}C_1C_3 - r_{14}C_1C_4 - r_{23}C_2C_3 + r_{24}C_2C_4}{(C_1^2 + C_2^2 - 2r_{12}C_1C_2)^{1/2}\,(C_3^2 + C_4^2 - 2r_{34}C_3C_4)^{1/2}}$				
II	Correlation of a ratio with a third variable	$y = x_1/x_2$; $z = x_3$	$\dfrac{r_{13}C_1 - r_{23}C_2}{(C_1^2 + C_2^2 - 2r_{12}C_1C_2)^{1/2}}$	$r_{13} = 0$	$\dfrac{-r_{23}C_2}{(C_1^2 + C_2^2 - 2r_{12}C_1C_2)^{1/2}}$		
III	Correlation of ratios with a common numerator	$y = x_1/x_2$; $z = x_1/x_4$	$\dfrac{C_1^2 - r_{14}C_1C_4 - r_{12}C_1C_2 + r_{24}C_2C_4}{(C_1^2 + C_2^2 - 2r_{12}C_1C_2)^{1/2}\,(C_1^2 + C_4^2 - 2r_{14}C_1C_4)^{1/2}}$	$r_{12} = r_{14} = r_{24} = 0$	$\dfrac{C_1^2}{(C_1^2 + C_2^2)^{1/2}\,(C_1^2 + C_4^2)^{1/2}}$	$C_1 = C_2 = C_4$ $C_1 > C_2, C_4$ $C_1 = 2C_2 = 2C_4$ $C_1 = 3C_2 = 3C_4$	0.5 >0.5 0.8 0.9
IV	Correlation of ratios with a common denominator	$y = x_1/x_2$; $z = x_3/x_2$	$\dfrac{r_{13}C_1C_3 - r_{12}C_1C_2 - r_{23}C_2C_3 + C_2^2}{(C_1^2 + C_2^2 - 2r_{12}C_1C_2)^{1/2}\,(C_2^2 + C_3^2 - 2r_{23}C_2C_3)^{1/2}}$	$r_{12} = r_{13} = r_{23} = 0$	$\dfrac{C_2^2}{(C_1^2 + C_2^2)^{1/2}\,(C_2^2 + C_3^2)^{1/2}}$	$C_1 = C_2 = C_3$ $C_1 > C_2, C_3$ $C_2 = 2C_1 = 2C_3$ $C_2 = 3C_1 = 3C_3$	0.5 >0.5 0.8 0.9
V	Correlation when denominator of one ratio equals numerator of other	$y = x_1/x_2$; $z = x_2/x_4$	$\dfrac{r_{12}C_1C_2 - r_{14}C_1C_4 + r_{24}C_2C_4 - C_2^2}{(C_1^2 + C_2^2 - 2r_{12}C_1C_2)^{1/2}\,(C_2^2 + C_4^2 - 2r_{24}C_2C_4)^{1/2}}$	$r_{12} = r_{14} = r_{24} = 0$	$\dfrac{-C_2^2}{(C_1^2 + C_2^2)^{1/2}\,(C_2^2 + C_4^2)^{1/2}}$	$C_1 = C_2 = C_4$ $C_2 > C_1, C_4$ $C_2 = 2C_1 = 2C_4$ $C_2 = 3C_1 = 3C_4$	−0.5 <−0.5 −0.8 −0.9
VI	Correlation of a ratio with its own numerator	$y = x_1/x_2$; $z = x_1$	$\dfrac{C_1 - r_{12}C_2}{(C_1^2 + C_2^2 - 2r_{12}C_1C_2)^{1/2}}$	$r_{12} = 0$	$\dfrac{C_1}{(C_1^2 + C_2^2)^{1/2}}$	$C_1 = C_2$ $C_1 > C_2$ $C_1 = 2C_2$ $C_1 = 3C_2$	0.71 >0.71 0.89 0.95
VII	Correlation of a ratio with its own denominator	$y = x_1/x_2$; $z = x_2$	$\dfrac{r_{12}C_1 - C_2}{(C_1^2 + C_2^2 - 2r_{12}C_1C_2)^{1/2}}$	$r_{12} = 0$	$\dfrac{-C_2}{(C_1^2 + C_2^2)^{1/2}}$	$C_2 = C_1$ $C_2 > C_1$ $C_2 = 2C_1$ $C_2 = 3C_1$	−0.71 < 0.71 −0.89 −0.95
VIII	Correlation of a ratio with the reciprocal of its numerator	$y = 1/x_2$; $z = x_2/x_4$	$\dfrac{r_{24}C_2 - C_2}{(C_2^2 + C_4^2 - 2r_{24}C_2C_4)^{1/2}}$	$r_{24} = 0$	$\dfrac{-C_2}{(C_2^2 + C_4^2)^{1/2}}$	$C_2 = C_4$ $C_2 > C_4$ $C_2 = 2C_4$ $C_2 = 3C_4$	−0.71 <−0.71 −0.89 −0.95

Table 11.1. (continued)

SPURIOUS CORRELATION: PRODUCTS

$$\left[\text{Mean of product } x_1x_2 \cong \bar{X}_1\bar{X}_2(1 + r_{12}C_1C_2)\text{ ; standard deviation of } x_1x_2 \cong \bar{X}_1\bar{X}_2(C_1^2 + C_2^2 + 2r_{12}C_1C_2)^{1/2}\right]$$

Model	(1)	(2) Graph	(3) r_{yz} (approx.)	Special cases (4) Given	(5) r_{yz} (approx.)	(6) Relative values of C	(7) r_{yz} (approx.)
I	General case of product correlation	$y = x_1x_2$; $z = x_3x_4$	$\dfrac{r_{13}C_1C_3 + r_{14}C_1C_4 + r_{23}C_2C_3 + r_{24}C_2C_4}{(C_1^2 + C_2^2 + 2r_{12}C_1C_2)^{1/2}\,(C_3^2 + C_4^2 + 2r_{34}C_3C_4)^{1/2}}$				
II	Products with one common factor	$y = x_1x_2$; $z = x_1x_3$	$\dfrac{r_{13}C_1C_3 + r_{12}C_1C_2 + r_{23}C_2C_3 + C_1^2}{(C_1^2 + C_2^2 + 2r_{12}C_1C_2)^{1/2}\,(C_1^2 + C_3^2 + 2r_{13}C_1C_3)^{1/2}}$	$r_{12} = r_{13} = r_{23} = 0$	$\dfrac{C_1^2}{(C_1^2 + C_2^2)^{1/2}\,(C_1^2 + C_3^2)^{1/2}}$	$C_1 = C_2 = C_3$ $C_1 > C_2, C_3$ $C_1 = 2C_2 = 2C_3$ $C_1 = 2C_2 = 3C_3$ $C_1 = 3C_2 = 3C_3$	0.5 >0.5 0.80 0.85 0.9
III	Correlation of a product with a third variable	$y = x_1x_2$; $z = x_3$	$\dfrac{r_{13}C_1 + r_{23}C_2}{(C_1^2 + C_2^2 + 2r_{12}C_1C_2)^{1/2}}$	$r_{13} = 0$	$\dfrac{r_{23}C_2}{(C_1^2 + C_2^2 + 2r_{12}C_1C_2)^{1/2}}$		
IV	Correlation of a product with one of the variables	$y = x_1x_2$; $z = x_2$	$\dfrac{r_{12}C_1 + C_2}{(C_1^2 + C_2^2 + 2r_{12}C_1C_2)^{1/2}}$	$r_{13} = 0$	$\dfrac{C_2}{(C_1^2 + C_2^2)^{1/2}}$	$C_1 = C_2$ $C_2 > C_1$	0.71 >0.71

* From Benson (1965) in this table r is the population correlation coefficient.

The correlation between Y_i/Z_i and X_i/Z_i is spurious.

Pearson (1896-1897) investigated the spurious correlation that can arise between ratios. Let $Y = X_1/X_2$ and $Z = X_3/X_4$. The correlation between Y and Z, ρ_{YZ}, was found to be a function of the variances, covariances and means of the X's. Pearson's derivation assumed that the X's were normally distributed and that the coefficient of variation of each X was small enough so that its third and higher powers could be neglected. Reed (1921) arrived at the same results without specifying the parent distribution of the X's. Pearson's general formula is

$$\rho_{YZ} = \frac{\rho_{13}C_1C_3 - \rho_{14}C_1C_4 - \rho_{23}C_2C_3 + \rho_{24}C_2C_4}{(C_1^2 + C_2^2 - 2\rho_{12}C_1C_2)^{1/2}(C_3^2 + C_4^2 - 2\rho_{34}C_3C_4)^{1/2}} \qquad (11.19)$$

where ρ_{ij} is the correlation between X_i and X_j and C_i is the coefficient of variation of X_i.

Chayes (1949) and Benson (1965) considered many special cases of equation 11.19. For example if $X_2 = X_4$, $\rho_{12} = \rho_{13} = \rho_{34} = 0$, $\rho_{24} = 1$, and $C_1 = C_2 = C_3 = C_4$, equation 11.19 reduces to $\rho_{XY} = 0.5$ which is the case shown in figure 11.4. Benson (1965) produced a table (Table 11.1) showing many special cases of ratio and product correlations.

Spurious correlation can arise in hydrology when dimensionless terms or standardized variables are used. Benson (1965) presents several examples of possible spurious correlation in hydrology.

Exercises

11.1 Calculate the first order serial correlation coefficients for the sediment load and annual discharge data for the Green River at Munfordville, Kentucky. Test the hypothesis that these two correlations are equal. Discuss the assumptions you have made and how they affect the validity of the tests you have made.

11.2 Calculate the correlation between the sediment load and annual discharge for the Green River at Munfordville, Kentucky. Test the hypothesis that this correlation is equal to 0.50.

11.3 Verify the "comment" of example 11.2.

11.4 Calculate the first order serial correlation coefficient for the Spray River, Banff, Canada. Test the hypothesis that the first order serial correlation is zero.

11.5 Work exercise 11.4 for the Piscataquis River near Dover-Foxcroft, Maine.

11.6 If the annual runoff from the Spray River, Banff, Canada, is normally distributed, how many independent observations would provide as much information relative to estimating the mean annual runoff as does the 45 years of actual record?

11.7 Work exercise 11.6 for the Piscataquis River, near Dover-Foxcroft, Maine and its 54 years of record.

11.8 The following data were collected on two streams in southeastern Kentucky. Use the data to extend the peak flow record of Cave Branch through 1972. Estimate the average peak flow for the entire record plus estimated record for Cave Branch. Is this estimated average an improvement over an estimate based only on the actual observed record of Cave Branch?

Peak Flow Data (cfs)

Year	Cave Br.	Helton Br.	Year	Helton Br.
1956	98	76	1967	69
1957	198	136	1968	36
1958	154	54	1969	10
1959	30	18	1970	122
1960	71	65	1971	67
1961	44	32	1972	54
1962	184	182		
1963	127	130		
1964	27	21		
1965	54	46		
1966	24	26		

CURSO: INGENIERIA MARITIMA. MODULO: "FUNDAMENTOS DE HIDRAULICA Y MATEMATICAS" DEL 11 DE MARZO AL 7 DE JUNIO.
MEXICO, D.F.

MULTIVARIATE ANALYSIS.

PROF. ING. JOSE RAYNAL.
JUNIO 1985.

12. Multivariate Analysis

THIS CHAPTER is devoted to a discussion of several multivariate techniques that may prove useful in the analysis of hydrologic data. The treatment is largely concerned with the principles and techniques rather than inferential aspects of multivariate analysis. Topics considered are principal components analysis, factor analysis, canonical correlation and multivariate regression analysis. These multivariate techniques have been around for some time but are only recently finding their way into hydrologic literature and hydrologic applications. For a more complete treatment of these topics, especially the inferential aspects of multivariate analysis, reference should be made to the books by Morrison (1967), Press (1972), Cooley and Lohnes (1971), Harman (1967), and Anderson (1958). Synder (1962) and Wong (1963) were among the first to apply multivariate analysis to hydrology.

NOTATION

In this chapter an upper case underlined letter will denote a matrix and a lower case underlined letter will denote a column vector. Thus $\underline{Z}$ could be an nxp matrix made up of p nxl column vectors $\underline{z}_j$ for $j = 1,2,...,p$.

$$\underline{Z} = (\underline{z}_1, \underline{z}_2, ..., \underline{z}_p)$$

$$\underline{Z} = [z_{ij}] \quad i = 1, 2, ... n; j = 1, 2, ..., p$$

$$\underline{z}_j = [z_{ij}] \quad i = 1, 2, ..., n$$

PRINCIPAL COMPONENTS

In chapter 10 it was shown that when data are collected on p variables, these p variables are many times correlated. This correlation indicates that some of the information contained in one variable is also contained in some of the other p - 1 variables. The objective of principal components analysis is to transform the p original correlated variables into p uncorrelated or orthogonal components. These components are linear functions of the original variables. Such a transformation can be written

$$\underline{Z} = \underline{X}\,\underline{A} \tag{12.1}$$

where $\underline{X}$ is an nxp matrix of n observations on p variables. Since we are dealing with variances and covariances, all X's will be assumed to be deviations from their respective means so that $\underline{X}$ is a matrix of deviations from means. $\underline{Z}$ is an nxp matrix of n values for each of p components, and $\underline{A}$ is a pxp matrix of coefficients defining the linear transformation.

Since the original p-variate set of observations contained in $\underline{X}$ contains correlation, it might be possible to characterize the variance of $\underline{X}$ with $q<p$ orthogonal components. Thus it is desired to construct $\underline{Z}$ so that each component, $\underline{z}_j$, (an nxl column vector) explains the maximum amount of the variance of $\underline{X}$ left unexplained by the first j - 1 components. In this way it may be found that the first q components explain most of the system variance and that the last p - q components explain little of the system variance. The total variance of $\underline{X}$ is defined to be the sum of the variances of the p variables contained in $\underline{X}$. The variance-covariance matrix of $\underline{X}$ is defined to be the pxp matrix $\underline{\Sigma}$ where $\underline{\Sigma} = [\sigma_{ij}]$ and σ_{ij} is the covariance of the i^{th} and j^{th} variables in $\underline{X}$ for $i \neq j$ and σ_{ii} is the variance of the i^{th} variable. $\underline{\Sigma}$ is estimated by $\underline{S}$ whose elements are given by

$$s_{ij} = \frac{1}{n-1} \sum_{k=1}^{n} X_{kj} X_{ki} \tag{12.2a}$$

so

$$\underline{S} = \underline{X}' \underline{X} /(n - 1) \tag{12.2b}$$

The total system variance, V, is defined as the sum of the variances of the original variables and can be estimated as

$$V = \text{Trace } \underline{S} = \sum_{i=1}^{p} s_{ii} \tag{12.3}$$

The j^{th} principal component, $\underline{z}_j$, is the linear function

$$\underline{z}_j = \underline{X}\,\underline{a}_j \tag{12.4}$$

where $\underline{z}_j$ is an nxl and $\underline{a}_j$ a pxl column vector. $\underline{z}_j$ can also be written $\underline{z}_j = [z_{ij}]$ where

$$z_{ij} = \sum_{k=1}^{p} X_{ik}\, a_{kj} \tag{12.5}$$

The variance of $\underline{z}_j$ is found from

$$\text{Var}(\underline{z}_j) = \text{Var}(\underline{X}\underline{a}_j) = \underline{a}_j' \text{ Var}(\underline{X})\, \underline{a}_j \tag{12.6}$$

and may be estimated by $\underline{a}_j'\ \underline{S}\ \underline{a}_j$. Note that this is simply a matrix equivalent of equation 3.58 for the variance of a linear function.

The variances of the first principal component z_1 is estimated by

$$\mathrm{Var}(\underline{z}_1) = \underline{a}_1' \underline{S}\, \underline{a}_1 \tag{12.7}$$

$\underline{z}_1$ is thus defined by the vector $\underline{a}_1$ that maximizes the variance of $\underline{z}_1$ subject to the constraint that $\underline{a}_1' \underline{a}_1 = 1$. This is a normalizing constraint without which there would be no unique solution.

Equation 12.7 can be maximized by using the Langrangian multiplier λ_1 to introduce the constraint $\underline{a}_1' \underline{a}_1 = 1$. Let

$$Q = \underline{a}_1' \underline{S}\, \underline{a}_1 + \lambda_1 (1 - \underline{a}_1' \underline{a}_1)$$

Q is maximized by differentiation.

$$\frac{\partial Q}{\partial \underline{a}_1} = 2(\underline{S} - \lambda_1 \underline{I})\, \underline{a}_1 = \underline{0}$$

or

$$(\underline{S} - \lambda_1 \underline{I})\, \underline{a}_1 = \underline{0} \tag{12.8}$$

For the solution of equation 12.8 to be other than the trivial solution $\underline{a}_1 = \underline{0}$, we must have

$$|\underline{S} - \lambda_1 \underline{I}| = 0 \tag{12.9}$$

This is a classical characteristic value problem. λ_1 is called the characteristic root and $\underline{a}_1$ the characteristic vector of $\underline{S}$. Equation 12.9 has p solutions for λ_1. This is easily seen by considering the special case of $\underline{S}$ to be a 2x2 matrix in which case equation 12.9 becomes

$$\begin{vmatrix} s_{11} - \lambda_1 & s_{12} \\ s_{21} & s_{22} - \lambda_1 \end{vmatrix} = 0$$

or $(s_{11} - \lambda_1)(s_{22} - \lambda_1) - s_{21} s_{12} = 0$. This is a quadratic equation in λ_1 having 2 solutions. Special properties[1] of $\underline{S}$ guarantee that the p solutions will be real.

Multiplying equation 12.8 by $\underline{a}_1'$ results in

$$\underline{a}_1' \underline{S} \underline{a}_1 = \lambda_1 \tag{12.10}$$

Since our objective was to maximize $\underline{a}_1' \underline{S} \underline{a}_1$, the desired solution to equation 12.9 is the largest characteristic root (the largest value) for λ.

Equations 12.7 and 12.10 demonstrate the important point that

$$\mathrm{Var}(\underline{z}_1) = \lambda_1 \tag{12.11}$$

Having found the characteristic root, λ_1, of $\underline{S}$, the characteristic vector, $\underline{a}_1$, is found from equation 12.8 using the constraint that $\underline{a}_1' \underline{a}_1 = 1$ which is equivalent to $\sum_{k=1}^{p} a_{k1}^2 = 1$.

The second principal component is found in a similar manner. Now it is desired to find $\underline{a}_2$ such that $\mathrm{Var}(\underline{z}_2) = \underline{a}_2' \underline{S}\, \underline{a}_2$ is maximized subject to the constraints that $\underline{a}_2' \underline{a}_2 = 1$ and $\underline{a}_1' \underline{a}_2 = \underline{a}_2' \underline{a}_1 = 0$. This latter constraint guarantees that $\underline{z}_1$ and $\underline{z}_2$ are orthogonal (uncorrelated). Using a procedure similar to the above for $\underline{a}_1$, let Q be

$$Q = \underline{a}_2' \underline{S}\, \underline{a}_2 + \lambda_2 (1 - \underline{a}_2' \underline{a}_2) + \gamma \underline{a}_1' \underline{a}_2$$

$$\frac{\partial Q}{\partial \underline{a}_2} = 2(\underline{S} - \lambda_2 \underline{I})\, \underline{a}_2 + \gamma \underline{a}_1 = \underline{0} \tag{12.12}$$

Premultiplication by $\underline{a}_1'$ results in

$$2\underline{a}_1' \underline{S}\, \underline{a}_2 + \gamma = 0 \tag{12.13}$$

Premultiplication of equation 12.8 by $\underline{a}_2'$ results in $\underline{a}_2' \underline{S}\, \underline{a}_1 = 0$. Substituting this into equation 12.13 results in $\gamma = 0$. Therefore from equation 12.12 we have

$$(\underline{S} - \lambda_2 \underline{I})\, \underline{a}_2 = \underline{0} \tag{12.14}$$

from which it follows that $\underline{a}_2$, the coefficients of the second largest principal component, are the coefficients of the characteristic vector associated with the second largest characteristic root of $\underline{S}$. Premultiplying equation 12.14 by $\underline{a}_2'$ also results in $\lambda_2 = \underline{a}_2' \underline{S}\, \underline{a}_2 = \mathrm{Var}(\underline{z}_2)$. In general the j^{th} principal component of the p-variate sample $\underline{X}$ is the linear function $\underline{z}_j = \underline{X}\, \underline{a}_j$ where $\underline{a}_j$ are the elements of the characteristic vector associated with the j^{th} largest characteristic root of $\underline{S}$. From 12.1 we can find $\underline{Z}'\underline{Z}$ as $\underline{Z}'\underline{Z} = (\underline{XA})'(\underline{XA}) = \underline{A}'\underline{X}'\underline{XA} = (n-1)\,\underline{A}'\underline{SA}$. It can be easily shown (see equations 12.24 thru 12.28) that $\underline{A}'\underline{SA}$ is a diagonal pxp matrix with the i^{th} diagonal element equal to λ_i. This matrix may also be written as $\underline{A}'\underline{SA} = \underline{D}_\lambda$ where $\underline{D}_\lambda$ is the diagonal matrix whose diagonal elements are the characteristic roots of $\underline{S}$.

One property of matrices is that if $\underline{E}$ is an orthogonal matrix, then the trace of $\underline{E}'\underline{FE}$ equals the trace of $\underline{F}$. Therefore

$$\mathrm{Trace}(\underline{D}_\lambda) = \mathrm{Trace}(\underline{A}'\underline{SA}) = \mathrm{Trace}(\underline{S}) = V$$

However

$$\mathrm{Trace}(\underline{D}_\lambda) = \sum_{i=1}^{p} \lambda_i$$

so the sum of the characteristic roots which equals the sum of the variances of the principal components also equals the total system variance.

The covariance between $\underline{z}_i$ and $\underline{z}_j$ is $\mathrm{Cov}(\underline{z}_i, \underline{z}_j) = \mathrm{Cov}(\underline{X}\underline{a}_i, \underline{X}\underline{a}_j) = \underline{a}_i' \underline{S}\, \underline{a}_j = 0$ for $i \neq j$. Therefore $\underline{z}_i$ and $\underline{z}_j$ are uncorrelated.

Some important properties obtained thus far are:

(1) $\underline{z}_i$ and $\underline{z}_j$ are uncorrelated for $i \neq j$
(2) $\mathrm{Var}(\underline{z}_i) = \underline{a}_i' \underline{S}\, \underline{a}_i = \lambda_i$
(3) $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_p$
(4) $\sum_{k=1}^{p} \mathrm{Var}(z_i) = \sum_{k=1}^{p} \lambda_i = \mathrm{Trace}\, \underline{S} = V$

1. The number of real, nonzero roots of a symmetric matrix is equal to the rank of the matrix. If $(\underline{X}'\underline{X})^{-1}$ exists, then the rank of $\underline{S}$ is p. $\underline{S}$ is symmetric since $s_{ij} = s_{ji}$.

(5) $\underline{Z} = \underline{X}\underline{A}$ where
$\underline{Z} = (\underline{z}_1, \underline{z}_2, \ldots, \underline{z}_p)$ and
$\underline{A} = (\underline{a}_1, \underline{a}_2, \ldots, \underline{a}_p)$

From item (4) above, it can be seen that the fraction of the total variance accounted for by the j^{th} principal component is λ_j/Trace $\underline{S}$. In many situations the first q components account for a large fraction (say 90% or more) of the system variance indicating that the last p-q components are not needed in terms of explaining variance. Many times these last p-q components are discarded with the effect that the problem has been reduced from one of dealing with an nxp $\underline{X}$ matrix containing correlation to dealing with an nxq (q<p) $\underline{Z}$ matrix that is orthogonal.

The questions of how many components are needed to satisfactorily explain the system variance or what part of the total system variance should be explained is an unresolved one. Morrison (1967) suggests that only the first four or five components should be extracted since later components will be difficult to physically interpret in terms of the problem at hand. Unfortunately there are no statistical tests that can be used to determine the significance of a component. The sampling theory of principal components is not well developed especially when the components are extracted from the correlation matrix rather than the covariance matrix as in later examples.

The covariance between the original variables, $\underline{X}$, and the principal components, $\underline{Z}$, is given by

$$\text{Cov}(\underline{X}, \underline{Z}) = \text{Cov}(\underline{X}, \underline{X}\underline{A}) = \underline{S}\underline{A} \tag{12.15}$$

The covariance between the variables and the j^{th} component is given by

$$\text{Cov}(\underline{X}, \underline{z}_j) = \text{Cov}(\underline{X}, \underline{X}\underline{a}_j) = \underline{S}\underline{a}_j \tag{12.16}$$

From $(\underline{S} - \lambda_j \underline{I})\,\underline{a}_j = \underline{0}$ we have $\underline{S}\underline{a}_j = \lambda_j \underline{a}_j$. Therefore $\text{Cov}(\underline{X}, \underline{z}_j) = \lambda_j \underline{a}_j$. The covariance between the i^{th} variable and the j^{th} component is given by $\lambda_j a_{ij}$. The correlation between the i^{th} variable and the j^{th} component is

$$\text{Cor}(\underline{x}_i, \underline{z}_j) = \text{Cov}(\underline{x}_i, \underline{z}_j)/[\text{Var}(\underline{x}_i)\,\text{Var}(\underline{z}_j)]^{1/2}$$

The $\text{Var}(\underline{x}_i) = s_{ii} = s_i^2$ and $\text{Var}(\underline{z}_j) = \lambda_j$. Therefore

$$\text{Cor}(\underline{x}_i, \underline{z}_j) = \lambda_j^{1/2} a_{ij}/s_i \tag{12.17}$$

This equation can be used to transform $\underline{A}$ into a pxp matrix of correlations between the i^{th} observed variable and the j^{th} computed component. These correlations can then be used in an attempt to assess the physical meaning of the components.

In some situations some of the p variables can be eliminated from further consideration by examining the correlations defined by equation 12.17. If a variable has no significant correlation with a component, then that variable is not contributing much to the variance of the component. By eliminating the variable from the component, the fraction of the system variance explained by the component would be changed very little. The difficulty here is that this particular variable may be correlated with a second component in which case its elimination would decrease the variance explained by the second component. For these reasons, variables are generally eliminated only if they are not correlated with any of the q components retained for analysis.

Example 12.1. Consider the data in table 10.2. Let $\underline{X}$ be a 13x3 matrix made up of 13 observations on A, S and L. Compute the principal components of $\underline{X}$ based on the covariance matrix. Compute the correlation between the variables and the components.

Solution:

$\underline{S}$ is computed from equation 12.2.

$$\underline{S} = \begin{bmatrix} 4.465 & -4.519 & 2.177 \\ -4.519 & 155.769 & -2.955 \\ 2.177 & -2.955 & 1.322 \end{bmatrix}$$

$$(\underline{S} - \lambda\underline{I}) = \begin{bmatrix} 4.465 - \lambda & -4.519 & 2.177 \\ -4.519 & 155.769 - \lambda & -2.955 \\ 2.177 & -2.955 & 1.322 - \lambda \end{bmatrix}$$

$|\underline{S} - \lambda\underline{I}|$ is computed from the relationship given in Appendix D,

$$|\underline{S} - \lambda\underline{I}| = (4.465 - \lambda)(155.769 - \lambda)(1.322 - \lambda) + (-4.519)$$
$$(-2.955)(2.177) + (2.177)(-4.519)(-2.955) -$$
$$(2.177)(155.769 - \lambda)(+2.177) - (-4.519)(-4.519)$$
$$(1.322 - \lambda) - (4.465 - \lambda)(-2.955)(-2.955) = 0$$

which simplifies to

$$\lambda^3 - 161.548\lambda^2 + 872.130\lambda - 171.154 = 0$$

The solutions to this cubic equation are

$$\lambda_1 = 155.963$$

$$\lambda_2 = 5.387$$

$$\lambda_3 = 0.207$$

Note that $\Sigma\lambda_i = \text{Trace } \underline{S} = 161.557$.

The first principal component accounts for $100\lambda_1/\text{Trace } \underline{S} = 100(155.963)/161.557 = 96.54$ percent of the total system variance. The coefficients of the characteristic vectors can be computed from equation 12.7. For example for the first principle component we have

$$(\underline{S} - \lambda_1 \underline{I})\,\underline{a}_1 = \underline{0}$$

$$\begin{bmatrix} 4.465 - 155.963 & -4.519 & 2.177 \\ -4.519 & 155.769 - 155.963 & -2.955 \\ 2.177 & -2.955 & 1.322 - 155.933 \end{bmatrix} \begin{bmatrix} a_{11} \\ a_{21} \\ a_{31} \end{bmatrix} = 0$$

or

$$\begin{array}{lll} -151.498\, a_{11} & -4.519\, a_{21} & +2.177\, a_{31} = 0 \\ -4.519\, a_{11} & -.194\, a_{21} & -2.955\, a_{31} = 0 \\ 2.177\, a_{11} & -2.955\, a_{21} & -154.641\, a_{31} = 0 \end{array}$$

Solving these three equations simultaneously for a_{11}, a_{21}, and a_{31} results in

$$a_{21} = -51.43\, a_{31} \text{ and } a_{11} = 1.5503\, a_{31}.$$

Using the constraint that $a_{11}^2 + a_{21}^2 + a_{31}^2 = 1$ the solution is $a_{11} = .030$, $a_{21} = -.999$ and $a_{31} = .020$. Similarly for λ_2 and λ_3 we get

$$\begin{array}{lll} a_{12} = .892 & a_{22} = .036 & a_{32} = .451 \\ a_{13} = -.452 & a_{23} = .004 & a_{33} = .892 \end{array}$$

Thus

$$\underline{A} = \begin{bmatrix} .030 & .892 & -.452 \\ -.999 & .036 & .004 \\ .020 & .451 & .892 \end{bmatrix}$$

The values for the principal components can now be calculated from

$$\underline{Z} = \underline{X}\underline{A}.$$

$$\begin{bmatrix} -0.96 & 33.54 & -0.19 \\ -0.64 & -9.46 & -0.02 \\ 2.46 & 2.54 & 0.54 \\ -1.62 & -10.46 & -0.73 \\ 1.98 & -0.46 & 1.57 \\ -1.03 & 9.54 & -0.65 \\ 2.17 & -9.46 & 2.16 \\ 4.30 & -5.46 & 1.67 \\ -1.07 & -11.46 & -0.57 \\ 0.72 & 1.54 & -0.47 \\ -2.50 & 4.54 & -1.42 \\ -2.32 & 6.54 & -1.30 \\ -1.45 & -11.46 & -0.64 \end{bmatrix} \begin{bmatrix} .030 & .892 & -.452 \\ -.999 & .036 & .004 \\ .020 & .451 & .892 \end{bmatrix} =$$

$$\begin{bmatrix} -33.55 & 0.25 & 0.40 \\ 9.44 & -0.92 & 0.24 \\ -2.45 & 2.53 & -0.62 \\ 10.39 & -2.15 & 0.05 \\ 0.55 & 2.46 & 0.51 \\ -9.58 & -0.87 & -0.07 \\ 9.56 & 2.57 & 0.92 \\ 5.62 & 4.39 & -0.47 \\ 11.41 & -1.62 & -0.06 \\ -1.53 & 0.48 & -0.73 \\ -4.64 & -2.71 & -0.11 \\ -6.63 & -2.42 & -0.08 \\ 11.40 & -1.99 & 0.05 \end{bmatrix}$$

The correlation matrix between the variables and the components can be computed from equation 12.17. For example the correlation between $\underline{x}_2$ and $\underline{z}_1$ is

$$\text{Cor}(\underline{x}_2, \underline{z}_1) = \lambda_1^{1/2}\, a_{21}/s_2 = 155.963^{1/2}(-0.999)/155.769^{1/2} = -0.9995$$

The resulting correlation matrix is

$$\begin{bmatrix} 0.178 & 0.979 & -0.097 \\ -1.000 & 0.007 & 0.000 \\ 0.212 & 0.911 & 0.353 \end{bmatrix}$$

Example 12.1 illustrates that using the $\underline{S}$ matrix in a principal component analysis presents some problems if the units of the X variables differ greatly. In example 12.1, the magnitude of the observations associated with the second variable were much greater than those associated with the other two variables. Consequently the variance of $\underline{x}_2$ was much greater than either $\text{Var}(\underline{x}_1)$ or $\text{Var}(\underline{x}_3)$. $\underline{x}_2$ accounted for 100 $\text{Var}(\underline{x}_2)$/ Trace $\underline{S}$ or 96.4% of the system variance. This means that the first principal component is merely a restatement of $\underline{x}_2$. This can also be seen from the fact that the correlation between $\underline{x}_2$ and $\underline{z}_1$ is 1.000.

In most hydrologic studies the problem of noncommensurate units on the X's has been handled by standardizing the X's through the transformation $(x_{ij} - \bar{x}_j)/s_j$. The covariance matrix of the standardized variables becomes the correlation matrix, $S = R$ as can be seen from equation 12.2. The principal components analysis is then done on R. The total system "variance" now becomes Trace $\underline{R} = p$ since $\underline{R}$ has ones on the diagonal.

The characteristic roots and vectors are determined from

$$(\underline{R} - \lambda_j \underline{I})\, \underline{a}_j = \underline{0} \qquad (12.18)$$

and the numerical value of the components is computed from

$$z_{ij} = \sum_{k=1}^{p} \left[\frac{x_{ik} - \bar{x}_k}{s_k} \right] a_{kj} \qquad (12.19)$$

The correlation between the i^{th} standardized variable and the j^{th} component (equation

12.17) reduces to

$$\text{Cor}(\underline{x}_i, \underline{z}_j) = \lambda_j^{1/2} a_{ij} \qquad (12.20)$$

These correlations are sometimes called factor loadings. The factor loadings can be used to attach physical significance to the components. If a particular component is highly correlated with 1, 2 or 3 variables, then the component is a reflection of these variables. For example in a study of watershed geomorphic factors, it might be found that a component is highly correlated with the average stream slope and the basin relief ratio. This being the case, that particular component might be termed a measure of watershed steepness.

Example 12.2. Repeat example 12.1 using $\underline{R}$ instead of $\underline{S}$

Solution:

$$\underline{R} = \begin{bmatrix} 1.0000 & -.1713 & .8958 \\ -.1713 & 1.0000 & -.2059 \\ .8958 & -.2059 & 1.0000 \end{bmatrix}$$

$$|\underline{R} - \lambda \underline{I}| = (1-\lambda)^3 - (1-\lambda)(.8768435) + .06343748 = 0$$

which has solutions

$\lambda_1 = 1.9692$

$\lambda_2 = 0.9273$

$\lambda_3 = 0.1035$

In this formulation $\underline{z}_1$ accounts for 100(1.9692)/3 or 65.64% of the system "variance" while $\underline{z}_2$ and $\underline{z}_3$ account for 30.91% and 3.45% respectively.

The corresponding characteristic vectors are

$$\underline{A} = (\underline{a}_1, \underline{a}_2, \underline{a}_3) = \begin{bmatrix} .679 & .208 & -.704 \\ -.265 & .964 & .029 \\ .684 & .167 & .710 \end{bmatrix}$$

The factor loadings computed from $\lambda_j^{1/2} a_{ij}$ are

$$\begin{bmatrix} .953 & .200 & -.226 \\ -.372 & .928 & .009 \\ .960 & .161 & .228 \end{bmatrix}$$

Since component 1 is highly correlated with both area and length, this component might be called a "size" component. Likewise component 2 might be called a slope component. In terms of explaining the "variance" of $\underline{R}$, component 3 could be eliminated since it explains only 3.40% of the variance and is not correlated with any of the variables. We cannot eliminate any variables, however, since component 1 is strongly dependent on X_1 and X_3 while component 2 depends on X_2.

In terms of explaining the variance of $\underline{R}$, we have reduced our problem from one of considering a 13x3 $\underline{X}$ matrix with correlations to a 13x2 $\underline{Z}$ matrix without correlations (assuming Z_3 is discarded).

The values for the components are computed from

$$\underline{Z} = \underline{X}\underline{A}$$

where $\underline{X} = [(x_{ij} - \bar{x}_j)/s_j] =$

$$\begin{bmatrix} -0.46 & 2.69 & -0.16 \\ -0.30 & -0.76 & -0.01 \\ 1.16 & 0.20 & 0.47 \\ -0.77 & -0.84 & -0.63 \\ 0.94 & -0.04 & 1.37 \\ -0.49 & 0.76 & -0.56 \\ 1.03 & -0.76 & 1.88 \\ 2.03 & -0.44 & 1.46 \\ -0.51 & -0.92 & -0.49 \\ 0.34 & 0.12 & -0.41 \\ -1.18 & 0.36 & -1.23 \\ -1.10 & 0.52 & -1.13 \\ -0.69 & -0.92 & -0.55 \end{bmatrix}$$

$\underline{Z} = \underline{X}\underline{A} =$

$$\begin{bmatrix} -1.13 & 2.47 & 0.28 \\ -0.02 & -0.80 & 0.18 \\ 1.06 & 0.52 & -0.48 \\ -0.73 & -1.07 & -0.07 \\ 1.58 & 0.39 & 0.31 \\ -0.92 & 0.54 & -0.03 \\ 2.19 & -0.20 & 0.59 \\ 2.49 & 0.25 & -0.41 \\ -0.44 & -1.07 & -0.02 \\ -0.08 & 0.12 & -0.52 \\ -1.74 & -0.10 & -0.03 \\ -1.66 & -0.09 & -0.01 \\ -0.60 & -1.12 & -0.07 \end{bmatrix}$$

REGRESSION ON PRINCIPAL COMPONENTS

Many times a principal components analysis is the first step in the development of a prediction model for some dependent variable, Y. Once the principal components are derived, they are used as the independent variables in a multiple regression analysis with the dependent variable, Y. Because of the differing units usually present in the original independent variables, the principal components are generally abstracted from the correlation matrix.

The steps in performing a multiple regression on principal components are outlined here. First the independent variables are standardized and the dependent variable centered so that $\underline{X} = [x_{ij}]$ and $\underline{Y} = [y_i]$ where

$$x_{ij} = (X_{ij} - \bar{X}_j)/s_j \text{ and } y_i = Y_i - \bar{Y} \qquad (12.21)$$

where Y_i is the i^{th} observation on Y, $\bar{Y}$ is the mean of Y, X_{ij} is the i^{th} observation on the j^{th} variable and $\bar{X}_j$ and s_j are the mean and standard deviation of the j^{th} variable. Centering Y is not necessary. It eliminates the need for an intercept and simplifies notation. The matrix of principal components, $\underline{Z}$, is determined from $\underline{Z} = \underline{X}\underline{A}$ with $\underline{A}$ being a pxp matrix whose j^{th} column is $\underline{a}_j$, the characteristic vector computed from equation 12.18 with $\underline{R} = \underline{X}'\underline{X}/(n-1)$.

The regression model is

$$\underline{Y} = \underline{Z}\underline{\beta} \text{ or } Y_i = \Sigma_{j=1}^{p} \beta_j z_{ij} \tag{12.22}$$

where $\underline{Y}$ is an nxl vector whose elements are the n observations of the centered dependent variable, $\underline{Z}$ is an nxp matrix whose elements, z_{ij}, represent the i^{th} value of the j^{th} principal component.

$\underline{\beta}$ is estimated from equation 10.8 as

$$\hat{\underline{\beta}} = (\underline{Z}'\underline{Z})^{-1}\underline{Z}'\underline{Y} \tag{12.23}$$

The expression for $\hat{\underline{\beta}}$ can be simplified by writing $\underline{Z}$ as

$$\underline{Z} = (\underline{z}_1, \underline{z}_2, \ldots, \underline{z}_p) \tag{12.24}$$

where $\underline{z}_j$ is an nxl vector whose elements are the n values of the j^{th} principal component.

$$\underline{Z}' = \begin{bmatrix} \underline{z}_1' \\ \underline{z}_2' \\ \cdot \\ \cdot \\ \underline{z}_p' \end{bmatrix} \tag{12.25}$$

so that

$$\underline{Z}'\underline{Z} = [\underline{z}_i'\ \underline{z}_j] \tag{12.26}$$

From equation 12.4 we have

$$\underline{Z}'\underline{Z} = [\underline{a}_i'\underline{X}'\underline{X}\,\underline{a}_j] = (n-1)[\underline{a}_i'\underline{R}\,\underline{a}_j] \tag{12.27}$$

Now $\underline{a}_i'\underline{R}\,\underline{a}_j$ is 0 for $i \neq j$ and is λ_j for $i = j$. Thus $\underline{Z}'\underline{Z}$ is a pxp matrix whose off-diagonal elements ($i \neq j$) are all zero and whose j^{th} diagonal element ($i = j$) is $(n-1)\lambda_j$.

$$\underline{Z}'\underline{Z} = (n-1)\begin{bmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_j & & \cdot \\ \cdot & & & \cdot \\ \cdot & & & \cdot \\ 0 & & & \lambda_p \end{bmatrix} = (n-1)\,\underline{D}_\lambda \tag{12.28}$$

$(\underline{Z}'\underline{Z})^{-1}$ is therefore

$$(\underline{Z}'\underline{Z})^{-1} = \frac{1}{n-1}\begin{bmatrix} 1/\lambda_1 & & & 0 \\ & 1/\lambda_2 & & \\ & & \cdot & \\ 0 & & & 1/\lambda_p \end{bmatrix} = \underline{D}_\lambda^{-1}/(n-1) \tag{12.29}$$

Equation 12.22 can now be written as

$$\begin{bmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \\ \cdot \\ \cdot \\ \hat{\beta}_p \end{bmatrix} = \frac{1}{n-1}\begin{bmatrix} 1/\lambda_1 & & & 0 \\ & 1/\lambda_2 & & \\ & & \cdot & \\ 0 & & & 1/\lambda_p \end{bmatrix}\begin{bmatrix} \underline{z}_1' \\ \underline{z}_2' \\ \cdot \\ \cdot \\ \underline{z}_p' \end{bmatrix}\begin{bmatrix} Y_1 \\ Y_2 \\ \cdot \\ \cdot \\ Y_n \end{bmatrix} \tag{12.30}$$

or

$$\hat{\beta}_j = \underline{z}_j'\,\underline{Y}/(n-1)\lambda_j = \underline{a}_j'\,\underline{X}'\underline{Y}'/(n-1)\lambda_j \tag{12.31}$$

From equation 10.14 and the above results it is apparent that

$$\mathrm{Cov}(\hat{\beta}_i, \hat{\beta}_j) \begin{cases} = 0 & \text{for } i \neq j \\ = \mathrm{Var}(\hat{\beta}_j) = \sigma^2/(n-1)\lambda_j & \text{for } i = j \end{cases} \tag{12.32}$$

where σ is the standard error of the regression equation.

Thus $\hat{\beta}_i$ is independent of $\hat{\beta}_j$ for $i \neq j$. The independence of the $\hat{\beta}$'s is a result of the orthogonality of the principal components. Since the $\hat{\beta}$'s are independent, the t-test given by equation 10.17 can be repeatedly applied to test hypotheses on the β's from a single regression equation. Furthermore, the numerical value for the β's retained in the regression will not be altered by eliminating any number of the other β's. This is the distinct advantage of having an orthogonal matrix of independent variables.

A second advantage of having independent β's is that the interpretation of the β's in terms of the independent variables is greatly simplified. Thus if some hydrologic meaning can be attached to a component through an examination of the factor loadings, hydrologic significance can also be attached to the β's. Unfortunately in most hydrologic applications of principal components analysis, a clear and distinct interpretation of the principal components has not been possible. This in turn means the hydrologic significance of the β's is unclear as well.

Some authors (DeCoursey and Deal 1974) state that yet another advantage for using regression on principal components as compared to normal multiple regression is that the resulting regression coefficients are more stable when applied to a new set of data because the coefficients are fitted on the basis of only statistically significant orthogonal components. This could imply that using an equation based on regression on principal components for prediction on a sample not included in the equation development would have a smaller standard error on this sample than would a normal multiple regression equation. If this is the case, it would be an important advantage for the regression on principal components technique. An adequate demonstration of this hypothesis needs to be developed, however.

A disadvantage of using principal components in a regression sis is that even

if all but one of the components is eliminated, all of the original variables (the X's) must still be measured since each component is a function of all of the X's (equation 12.4).

In reporting the results of a regression on principal components, it is generally desirable to transform the resulting regression equation into an equation in terms of the original X variables. This can be done since $y_i - \bar{Y}_i = \bar{Y}$, the $\hat{\beta}_j$'s are known constants, $z_{ij} = \Sigma_{k=1}^{p} a_{kj} x_{ik}$ and $x_{ij} = (X_{ij} - \bar{X}_j)/s_j$. Thus equation 12.22 becomes

$$Y_i = \bar{Y} + \Sigma_{j=1}^{p} \hat{\beta}_j \{\Sigma_{k=1}^{p} a_{kj} [(X_{ik} - \bar{X}_k)/s_k]\} \quad (12.33)$$

Equation 12.33 can then be simplified by collecting terms to be of the form

$$Y_i = \beta_o^* + \Sigma_{j=1}^{p} \beta_j^* X_{ij} \quad (12.34)$$

where the β^*'s are constants. If only q (q<p) components are retained in the final regression equation and the components are rearranged so that the first q components are retained, the first summation in equation 12.33 would run from 1 to q; however, the second summation would still run from 1 to p. This means the summation in equation 12.34 would run from 1 to p. It also means that even though the equation contains only q components, all p of the original variables must be measured to predict Y.

Some of the original X variables can be eliminated from the analysis before any regressions are performed by examining the factor loadings and eliminating variables that are not highly correlated with any of the components. The remaining X variables are then resubmitted to a principal components analysis with the multiple regression being performed on the new components. This procedure has the advantage of reducing the number of variables that must be measured to use the resulting regression equation. It has the disadvantage of eliminating X variables rather arbitrarily (there is no statistical test for the significance of the factor loadings) without ever having them in a position to determine their usefulness in explaining the variation in the dependent variable, Y.

In many applications of regression on principal components, the last p-q components are discarded before the regression is performed. The number of retained components, q, is selected so that a large proportion of the variance of $\underline{X}$ is accounted for. This procedure reduces the number of coefficients that must be estimated but runs the risk of eliminating a component that may explain a significant amount of the variation in $\underline{Y}$ even though it explains little of the variance of $\underline{X}$.

Equation 12.31 gives $\hat{\beta}_j = \underline{a}_j' \underline{X}' \underline{Y}/(n-1) \lambda_j$ while equation 12.32 gives $\text{Var}(\hat{\beta}_j) = \sigma^2/(n-1)\lambda_j$. The statistical significance of β_j is tested using equation 10.17 with $\beta_o = 0$. Thus the test statistic is

$$t = \underline{a}_j' \underline{X}' \underline{Y}/\sqrt{(n-1)\lambda_j}\,\sigma \quad (12.35)$$

There is no reason to believe before the regression is performed that this test statistic will be nonsignificant for small values of λ_j (i.e. for the last p-q components). Therefore the regression should be performed on all of the components and then the components that prove to be nonsignificant can be eliminated.

The value of the test statistic given by equation 12.35 can be shown to be proportional to the correlation between $\underline{Y}$ and $\underline{z}_j$ as follows:

$$\text{Cov}(\underline{z}_j, \underline{Y}) = \underline{z}_j' \underline{Y}/(n-1) = \underline{a}_j' \underline{X}' \underline{Y}/(n-1) \quad (12.36)$$

$$\text{Cor}(\underline{z}_j, \underline{Y}) = \text{Cov}(\underline{z}_j, \underline{Y})/\sqrt{\text{Var}(\underline{z}_j)\text{Var}(\underline{Y})} = \underline{a}_j' \underline{X}' \underline{Y}/(n-1)\, s_Y \sqrt{\lambda_j} \quad (12.37)$$

Therefore

$$t = \sqrt{n-1}\ \text{Cor}(\underline{z}_j, \underline{Y})\, s_Y/\sigma \quad (12.38)$$

or the significance of the j^{th} component is directly proportional to its correlation with the dependent variable. Equation 12.38 can be used to test the significance of the j^{th} component.

At this point it should be noted that if a dependent variable $\underline{Y}$ is regressed on p principal components extracted from a pxp correlation matrix and then transformed via equation 12.33, the results are identical to those that would be obtained by a direct regression of $\underline{Y}$ on the original p variables. This is because multiple regression is a linear operation and the principal components are independent linear functions of the original variables that explain all of the variance of the variables.

ROTATION OF PRINCIPAL COMPONENTS – I

For principal components to be of maximum usefulness, it is necessary to be able to attach physical significance to the components. That is it must be possible to interpret the significance of a component in terms of the problem being analyzed. The interpretation of principal components is done through the factor loading matrix which shows the correlations between the components and the original variables.

For maximum ease of interpretation of the components, it would be desirable for the factor loadings of an individual component to be high (> .90 ?) on some of the variables, low on some of the other variables, and to have a minimum number of intermediate loadings. The high loadings would indicate a high correlation with the variable or a strong linear similarity between the component and the variable indicating that both are measures of the hydrologic attribute. A low loading would indicate that the component and variable were independent or that the variable was not influential with the component. The intermediate loadings would indicate some relationship between the component and the variable but its real significance would be difficult to assess.

To aid in the interpretation of the principal components, it is sometimes desirable to make an orthogonal transformation of the factor loading matrix in hopes of producing a factor loading matrix having the properties described in the preceding paragraph. The mechanics of transformations of this type are described in the section – Rotation of Principal Components – II.

FACTOR ANALYSIS

The purpose of factor analysis is to partition a p-variate observed vector into some factors common to all of the p variables and some factors unique to each of the p variables. A factor model might be written

$$\underset{pxl}{\underline{X}} = \underset{pxm}{\underline{G}}\ \underset{mxl}{\underline{F}} + \underset{pxp}{\underline{H}}\ \underset{pxl}{\underline{U}} \quad (12.39)$$

where $\underline{X}$ is a p-variate vector of observed variables, $\underline{G}$ is a matrix of coefficients, $\underline{F}$ is a vector of common factors, $\underline{H}$ is a diagonal matrix of coefficients, and $\underline{U}$ is a vector of unique factors.

Matalas and Reiher (1967) and Wallis (1967) discuss some applications of factor analysis in hydrology. A common application of factor analysis in hydrology has been conducted by ignoring the unique factors. The common factors, $\underline{F}$, are then taken as a linear function of the observed variables.

$$\underset{n\times p}{\underline{F}} = \underset{n\times p}{\underline{X}}\ \underset{p\times p}{\underline{B}} \tag{12.40}$$

$\underline{F}$ is a matrix of factor scores consisting of n values for each of the p factors, $\underline{X}$ is a matrix of n observations on each of p standardized variables, and $\underline{B}$ is pxp matrix of coefficients known as factor score coefficients.

One way estimating $\underline{B}$ is to define $\underline{B}$ as

$$\underline{B} = \underline{A}\ \underline{D}_\lambda^{-1/2} \tag{12.41}$$

where $\underline{A}$ is the pxp matrix whose columns are the characteristic vectors of the correlation matrix $\underline{R}$, $\underline{D}_\lambda$ is a pxp diagonal matrix whose diagonal elements are the characteristic roots of $\underline{R}$, and $\underline{D}_\lambda^{-1/2}$ is a diagonal matrix formed by inverting $\underline{D}_\lambda$ and then taking the square root of its elements.

The factor score matrix $\underline{F}$ is "orthogonal" since

$$\underline{F}'\underline{F} = \underline{B}'\underline{X}'\underline{X}\ \underline{B} = \underline{D}_\lambda^{-1/2\prime}\ \underline{A}'(n-1)\ \underline{RAD}_\lambda^{-1/2} = (n-1)\ \underline{D}_\lambda^{-1/2}\underline{D}_\lambda\underline{D}_\lambda^{-1/2} = (n-1)\ \underline{I} \tag{12.42}$$

The correlation between the j^{th} observed variable and the k^{th} factor is called the factor loading. The pxp matrix $\underline{L}$ containing the loadings is called the factor structure or factor loading matrix.

$\underline{F}$ and the principal components $\underline{Z}$ defined in the previous section are related by

$$\underline{Z} = \underline{X}\ \underline{A} = \underline{X}\ \underline{B}\ \underline{D}_\lambda^{1/2} = \underline{F}\ \underline{D}_\lambda^{1/2} \tag{12.43}$$

Since multiplication of a variable by a constant does not change its correlation with another variable, $\mathrm{Cor}(\underline{X}_i, \underline{F}) = \mathrm{Cor}(\underline{X}_i, \underline{Z})$ or from equation 12.20 we have

$$\underline{L} = \underline{A}\ \underline{D}_\lambda^{1/2} \tag{12.44}$$

In the section on principal components it was shown that $\underline{A}'\underline{R}\ \underline{A} = \underline{D}_\lambda$. From this result it is apparent that $\underline{R}\ \underline{A} = \underline{A}\ \underline{D}_\lambda$. Writing equation 12.44 as $\underline{L} = \underline{A}\ \underline{D}_\lambda\ \underline{D}_\lambda^{-1/2}$ it can be seen that $\underline{L} = \underline{R}\ \underline{A}\ \underline{D}_\lambda^{-1/2}$ or the factor loading matrix is given by

$$\underline{L} = \underline{R}\ \underline{B} \tag{12.45}$$

The factor loading matrix has the property that

$$\underline{L}'\underline{L} = \underline{D}_\lambda^{1/2\prime}\ \underline{A}'\underline{A}\ \underline{D}_\lambda^{1/2} = \underline{D}_\lambda \tag{12.46}$$

and

$$\underline{L}\ \underline{L}' = \underline{A}\ \underline{D}_\lambda^{1/2}\ \underline{D}_\lambda^{1/2\prime}\ \underline{A}' = \underline{A}\ \underline{D}_\lambda\ \underline{A}' = \underline{R} \tag{12.47}$$

Equation 12.47 shows that the factor loading matrix has the ability to reproduce $\underline{R}$. $\underline{L}$ is not unique in this respect. Any orthogonal rotation of $\underline{L}$ will also reproduce $\underline{R}$. For example if $\underline{L}^* = \underline{LT}$ where $\underline{T}$ is orthogonal then

$$\underline{L}^*\underline{L}^{*\prime} = \underline{LT}\ \underline{T}'\underline{L}' = \underline{L}\ \underline{L}' = \underline{R} \tag{12.48}$$

It should be noted however, that

$$\underline{L}^{*\prime}\underline{L}^* = \underline{T}'\underline{L}'\underline{L}\ \underline{T} = \underline{T}'\underline{D}_\lambda\ \underline{T} \tag{12.49}$$

which is not a diagonal matrix.

The factors $\underline{F}$ defined by equation 12.40 are thus unique only up to an orthogonal rotation. $\underline{B}$ can be determined from equation 12.45 as

$$\underline{B} = \underline{R}^{-1}\ \underline{L} \tag{12.50}$$

Any $\underline{B}^*$ given by

$$\underline{B}^* = \underline{R}^{-1}\ \underline{L}^* \tag{12.51}$$

will also produce a set of "orthogonal" factors $\underline{F}^* = \underline{XB}^*$. This is apparent from

$$\underline{F}^{*\prime}\underline{F}^* = \underline{B}^{*\prime}\underline{X}'\underline{X}\ \underline{B}^* = \underline{L}^{*\prime}\underline{R}^{-1}(n-1)\ \underline{R}\ \underline{R}^{-1}\ \underline{L}^*$$

$$= (n-1)\ \underline{T}'\underline{L}'\underline{R}^{-1}\ \underline{R}\ \underline{R}^{-1}\ \underline{LT}$$

$$= (n-1)\ \underline{T}'\underline{L}'\underline{R}^{-1}\ \underline{LT}$$

but

$$\underline{L}'\underline{R}^{-1}\ \underline{L} = \underline{D}_\lambda^{1/2\prime}\ \underline{A}'\underline{B} = \underline{D}_\lambda^{1/2\prime}\ \underline{A}'\underline{AD}_\lambda^{-1/2} = \underline{I}$$

so

$$\underline{F}^{*\prime}\underline{F}^* = (n-1)\ \underline{T}'\underline{T} = (n-1)\ \underline{I}$$

ROTATION OF PRINCIPAL COMPONENTS – II

Equation 12.43 can be taken as the defining relationship for principal components. The matrix $\underline{L}$ contains the correlations between the components and the original variables. $\underline{L}$ can be rotated by an orthogonal matrix $\underline{T}$ so that the correlations between the variables and the components are changed. The most common method of rotation is called "varimax" rotation and was first defined by Kaiser (1958). Harman (1967) presents a discussion of the actual computations involved in a varimax rotation.

Basically a varimax rotation involves computing an orthogonal matrix of rotation $\underline{T}$ that produces an $\underline{L}^* = \underline{LT}$ such that the correlations between some of the variables and components are high, some are low, and a minimum of the correlations are intermediate. This makes interpretation of the components or placing physical significance on the components easier. Most computer centers having principal components pro-

grams will also have programs for varimax rotation.

Once the rotated factor loading matrix $\underline{L}^*$ is computed, the rotated principal components can be computed from

$$\underline{Z}^* = \underline{X}\,\underline{B}^*\,\underline{D}_\lambda^{-1/2} = \underline{X}\,\underline{R}^{-1}\,\underline{L}^*\,\underline{D}_\lambda^{-1/2} = \underline{X}\,\underline{R}^{-1}\,\underline{L}\,\underline{T}\,\underline{D}_\lambda^{-1/2} \qquad (12.52)$$

Many of the varimax programs do not actually provide the user with $\underline{T}$, but produce $\underline{L}^*$.

$\underline{Z}^*$ can be shown to be "orthogonal" in the sense the $\underline{Z}^{*\prime}\,\underline{Z}^* = (n-1)\,\underline{D}_\lambda$ as follows

$$\underline{Z}^{*\prime}\,\underline{Z}^* = (\underline{X}\,\underline{B}^*\,\underline{D}_\lambda^{-1/2})'\,\underline{X}\,\underline{B}^*\,\underline{D}_\lambda^{-1/2}$$

$$= (\underline{F}^*\,\underline{D}_\lambda^{-1/2})'\,\underline{F}^*\,\underline{D}_\lambda^{-1/2}$$

$$= \underline{D}_\lambda^{-1/2\prime}\,\underline{F}^{*\prime}\,\underline{F}^*\,\underline{D}_\lambda^{-1/2}$$

$$= (n-1)\,\underline{D}_\lambda$$

Therefore the rotated principal components are uncorrelated.

Since generally the rotated components are easier to interpret, they are many times more meaningful in a regression analysis than the unrotated components. One procedure for using regression on principal components is to

(1) Compute the factor loading matrix $\underline{L}$.
(2) Compute $\underline{L}^* = \underline{L}\underline{T}$ by a varimax rotation of $\underline{L}$.
(3) Eliminate X variables that have no high (>0.90?) correlations with any of the rotated components.
(4) Repeat steps (1) and (2) with the remaining variables.
(5) Perform a multiple regression of $\underline{Y}$ on $\underline{Z}^*$ where $\underline{Z}^*$ is computed from equation 12.52 based on the results of step (4).
(6) Eliminate the nonsignificant components from the regression.
(7) Put the resulting equation in the form of equation 12.34.

Wallis (1965) suggests a procedure similar to this. The comments made in the section "Regression on Principal Components", made relative to the model $\underline{Y} = \underline{Z}\,\underline{\beta}$ using the principal components apply as well to the model $\underline{Y} = \underline{Z}^*\underline{\beta}$ using the rotated principal components. Dawdy and Feth (1967) present situations where rotated factor loading matrices may be used in studying inter-relationships among indpendent variables. Shelton and Sewell (1969) and Overton (1969) discuss what is one of the most effective uses of principal components analysis – screening of parameters. They demonstrate how principal components can be used to eliminate redundant variables and variables that do not contribute to explaining the variation present in the original data matrix.

Example 12.3. Compute the components obtained by rotating the factor loading matrix of example 12.2 using the rotation defined by

$$\underline{T} = \begin{bmatrix} .866 & .500 & 0 \\ -.500 & .866 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Solution:

The fact that $\underline{T}$ is orthogonal is seen from

$$\underline{T}'\,\underline{T} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \underline{I}$$

The rotated factor loadings are

$$\underline{L}^* = \underline{L}\,\underline{T} = \begin{bmatrix} 0.9532 & .2002 & -.2264 \\ -.3725 & .9280 & .0092 \\ .9601 & .1612 & .2284 \end{bmatrix} \begin{bmatrix} .866 & .500 & 0 \\ -.500 & .866 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$\underline{L}^* = \begin{bmatrix} .725 & .650 & -.226 \\ -.787 & .617 & .009 \\ .751 & .620 & .228 \end{bmatrix}$$

$$\underline{Z}^* = \underline{X}\,\underline{R}^{-1}\,\underline{L}^*\,\underline{D}_\lambda^{-1/2}$$

$$\underline{R}^{-1} = \begin{bmatrix} 1.0000 & -.1713 & .8958 \\ -.1713 & 1.0000 & -.2059 \\ .8958 & -.2059 & 1.0000 \end{bmatrix}^{-1} = \begin{bmatrix} 5.066 & -.069 & -4.553 \\ -.069 & 1.045 & .277 \\ -4.553 & .277 & 5.135 \end{bmatrix}$$

$$\underline{R}^{-1}\,\underline{L}^* = \begin{bmatrix} .311 & .429 & -2.188 \\ -.664 & .772 & .089 \\ .355 & .394 & 2.206 \end{bmatrix}$$

$$\underline{R}^{-1}\,\underline{L}^*\,\underline{D}_\lambda^{-1/2} = \begin{bmatrix} .437 & .413 & -.704 \\ -.932 & .743 & .029 \\ .471 & .380 & .710 \end{bmatrix}$$

$\underline{X} = [(x_{ij} - \bar{x}_j)/s_j]$ (see example 12-2).

$$\underline{Z}^* = \underline{X}\,\underline{R}^{-1}\,\underline{L}^*\,\underline{D}_\lambda^{-1/2} = \begin{bmatrix} -2.78 & 1.75 & .28 \\ .57 & -.69 & .18 \\ .54 & .81 & -.48 \\ .15 & -1.18 & .07 \\ 1.09 & .88 & .31 \\ -1.19 & .15 & -.03 \\ 2.04 & .57 & .59 \\ 1.98 & 1.07 & -.41 \\ .40 & -1.08 & -.02 \\ -.16 & .08 & -.52 \\ -1.44 & -.69 & -.03 \\ -1.50 & -.49 & -.01 \\ .30 & -1.18 & -.07 \end{bmatrix}$$

Comment: The purpose of this example is to demonstate with numbers the meanings

of some of the matrices that have been used. The rotation is not a varimax rotation. As a matter of fact the original components are much easier to interpret than the rotated ones since $\underline{z}_1$ is highly correlated with $\underline{x}_1$ and $\underline{x}_3$, $\underline{z}_2$ is highly correlated with $\underline{x}_2$, and $\underline{z}_3$ is not correlated with any of the variables. The correlations involved are contained in $\underline{L}$ and are

$$\text{Cor}(\underline{x}_1, \underline{z}_1) = 0.9532$$

$$\text{Cor}(\underline{x}_3, \underline{z}_1) = 0.9601$$

$$\text{Cor}(\underline{x}_2, \underline{z}_2) = 0.9280$$

The rotated components are correlated about equally with all of the variables and would be difficult to interpret as seen by the matrix $\underline{L}^*$.

AN APPLICATION OF REGRESSION ON PRINCIPAL COMPONENTS

In Chapter 10 the procedures involved in multiple regression were illustrated using a set of data from Haan and Read (1970). In this section that problem will be re-analyzed using regression on principal components following the steps outlined earlier. The data for the problem are contained in table 10.2. The data consist of three classes of variables: (1) runoff, the dependent variable, (2) precipitation and (3) eight geomorphic variables. The principle components will be determined for the eight geomorphic variables. These components along with precipitation will then be the independent variables in the regression with runoff as the dependent variable.

The first step in the analysis was to compute the characteristic vectors and characteristic values for the correlation matrix shown in table 10.4. The results of this computation yielded the characteristic vectors:

$$\underline{A} = \begin{bmatrix} .463 & .058 & -.176 & .131 & -.288 & -.166 & .272 & -.742 \\ -.139 & .653 & -.183 & .390 & .544 & .190 & .149 & -.123 \\ .456 & .091 & .182 & .059 & -.204 & .824 & -.119 & .110 \\ .461 & .103 & -.033 & 2.71 & .141 & -.364 & -.734 & .109 \\ .394 & -.042 & -.051 & .107 & -.040 & -.134 & .451 & .593 \\ -.191 & -.176 & -.796 & -.223 & .009 & .291 & -.361 & -.184 \\ -.239 & -.499 & -.018 & .819 & -.101 & .112 & -.004 & -.024 \\ -.311 & .519 & -.103 & .151 & -.741 & -.097 & -.132 & .158 \end{bmatrix}$$

and the characteristic values:

$$\underline{D}_\lambda = \begin{bmatrix} 4.359 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1.802 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1.193 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0.447 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0.124 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0.045 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0.025 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.003 \end{bmatrix}$$

The factor loading matrix was then calculated from equation 12.44 as

$$\underline{L} = \begin{bmatrix} .967 & .077 & -.192 & .087 & -.101 & -.035 & .043 & -.043 \\ -.291 & .877 & -.200 & .261 & .192 & .040 & .024 & -.071 \\ .953 & .123 & .198 & .040 & -.072 & .175 & -.019 & .064 \\ .962 & .139 & -.036 & .181 & .050 & -.077 & -.116 & .063 \\ .824 & -.056 & -.553 & .072 & -.014 & -.028 & .072 & .034 \\ -.398 & -.236 & -.870 & -.149 & .003 & .062 & -.057 & -.011 \\ -.500 & -.670 & -.020 & .547 & -.036 & .024 & .001 & -.001 \\ -.650 & .697 & -.112 & .101 & -.261 & -.021 & -.021 & .009 \end{bmatrix} \begin{matrix} \text{Var.} \\ A \\ S \\ L \\ P \\ d_i \\ R_s \\ F \\ R_r \end{matrix}$$

At this point an attempt can be made to attach physical significance to the components. Component 1 is highly correlated with A, L, and P and moderately correlated with d_i. Thus component 1 might be termed a size component. Component 2 is moderately correlated with S, F, and R_r and might be called a slope or steepness variable although this interpretation is not really clear. Component 3 has moderate correlation with R_s and might be named a measure of watershed shape. The remaining components are not highly correlated with any of the variables.

The factor loading matrix was subjected to a varimax rotation resulting in

$$\underline{L}^* = \begin{bmatrix} \underline{0.955} & -0.162 & 0.135 & -0.199 & -0.021 & -0.003 & 0.042 & -0.042 \\ -0.038 & \underline{0.979} & -0.009 & -0.137 & 0.146 & -0.011 & 0.012 & 0.000 \\ 0.782 & -0.212 & 0.465 & -0.264 & 0.018 & 0.238 & 0.003 & -0.001 \\ \underline{0.915} & -0.100 & 0.302 & -0.164 & 0.106 & -0.017 & -0.150 & 0.003 \\ \underline{0.938} & -0.157 & -0.261 & -0.110 & 0.068 & -0.044 & 0.077 & 0.042 \\ -0.095 & 0.050 & \underline{-0.988} & 0.110 & -0.002 & -0.007 & 0.006 & -0.010 \\ -0.322 & -0.196 & -0.175 & \underline{0.909} & 0.007 & -0.011 & 0.008 & -0.001 \\ -0.407 & 0.839 & -0.103 & -0.064 & -0.340 & -0.011 & 0.009 & -0.002 \end{bmatrix} \begin{matrix} \text{Var.} \\ A \\ S \\ L \\ P \\ d_i \\ R_s \\ F \\ R_r \end{matrix}$$

Again we can attempt to place physical significance on the components. The correlations exceeding 0.90 are underlined. Component 1 is clearly a size or "bigness" component in that it is correlated with A, P, and d_i. The correlation with L has diminished while that on d_i has increased. Component 2 is now more clearly a steepness component in that it is highly correlated with S and moderately correlated with R_r. Component 3 is now highly correlated with R_s and is clearly a shape component. Component 4 now has a high correlation with F and is a measure of stream density.

The rotated factors are somewhat easier to interpret although the rotation has not been so helpful in this example as it is in some situations. It can be seen that the rotated factor loading matrix has fewer intermediate correlations (.45 to .75). It is these intermediate correlations that make the interpretation of the components difficult.

Those variables having a rotated factor loading of less than 0.9 were discarded. This left the six variables A, S, P, d_i, R_s and F. The characteristic vectors and characteristic values of the correlation matrix of these six variables were computed and the factor loading matrix subjected to a varimax rotation. The correlation matrix is the same as table 10.4 omitting the rows and columns corresponding to L and R_r. This new correlation matrix is $\underline{R}$. The characteristic values for this 6x6 correlation matrix are 3.135, 1.322, 1.073, 0.433, 0.029 and 0.008.

If the practice of extracting components until 90 percent of the system variance is explained is adopted, only the first three of these components would be retained. The

last three components account for only 7.8 percent of the system variance. All six of the components will be retained in this example.

The rotated factor loading matrix for the remaining six variables is

$$\underline{L}^* = \begin{bmatrix} 0.953 & -0.091 & -0.157 & -0.220 & -0.062 & 0.074 \\ -0.119 & 0.980 & 0.027 & -0.161 & 0.002 & 0.000 \\ 0.916 & -0.011 & -0.324 & -0.180 & 0.152 & -0.007 \\ 0.957 & -0.074 & 0.237 & -0.117 & -0.079 & -0.058 \\ -0.071 & 0.031 & 0.990 & 0.125 & -0.007 & -0.005 \\ -0.282 & -0.195 & 0.167 & 0.925 & -0.006 & 0.000 \end{bmatrix}$$

The rotated principal components are now determined from equation 12.52 ($\underline{Z}^* = \underline{X}\,\underline{R}^{-1}\,\underline{L}^*\,\underline{D}_\lambda^{-1/2}$) as

$$\underline{Z}^* = \begin{bmatrix} -0.888 & 3.011 & -1.201 & -0.397 & -0.032 & -0.080 \\ -0.650 & -1.031 & -0.162 & -0.329 & -0.065 & -0.176 \\ 2.597 & 0.445 & 0.691 & 0.080 & -0.015 & -0.077 \\ -1.321 & -0.915 & -0.277 & 0.572 & -0.117 & -0.026 \\ 1.898 & 0.381 & -1.154 & 0.832 & -0.079 & 0.021 \\ -0.604 & 0.587 & 2.012 & -0.696 & 0.103 & -0.009 \\ 0.675 & -1.035 & -1.921 & -0.883 & 0.265 & 0.107 \\ 3.094 & -0.466 & 0.413 & -0.657 & -0.345 & 0.058 \\ -0.161 & -0.730 & 0.006 & 1.264 & -0.048 & -0.046 \\ 1.325 & 0.332 & 0.875 & 0.398 & 0.317 & 0.039 \\ -2.414 & 0.259 & -0.339 & -0.036 & -0.132 & 0.059 \\ -1.781 & 0.550 & 0.833 & 0.354 & -0.027 & 0.167 \\ -1.768 & -1.388 & 0.223 & -0.572 & -0.081 & -0.037 \end{bmatrix}$$

Next a multiple regression was performed using the model

$$RO = \beta_o + \beta_1\,Prec + \beta_2 z_1^* + \beta_3 z_2^* + \ldots + \beta_7 z_6^*$$

This model is slightly different than the one discussed earlier in that it contains the Prec variable in addition to the principal components. Prec could have been incorporated into the correlation matrix with the eight geomorphic variables. This was not done since Prec is not correlated with any of the geomorphic variables and would thus require a component for its characterization. Rather than go through the mechanics of extracting a component to explain Prec, the Prec variable was used directly. In some situations, Prec could be correlated with a geomorphic variable such as elevation.

The results of the regression on this model are contained in table 12.1. The regression coefficient on z_4^* was not significantly different from zero while the coefficients on z_5^* and z_6^* were significant only at the 90% level. Thus two more regressions were run — one by eliminating z_4^* only and one by eliminating z_4^*, z_5^* and z_6^*. When the latter three rotated components were eliminated, the R^2 dropped to 0.805 and the standard error climbed to 1.046. Elimination of z_4^* only resulted in an R^2 of 0.905 and a standard error of 0.844. Thus the equation containing Prec and all of the rotated principal components except z_4^* was accepted as the "best" equation. Table 12.2 presents more details on this equation. In comparing table 12.1 and 12.2 it is apparent that the regression coefficients and their standard errors have changed. This is because the inclusion of

Table 12.1: Results of first regression using rotated principal components.

Correlation matrix

	RO	Prec	z_1^*	z_2^*	z_3^*	z_4^*	z_5^*	z_6^*
RO	1.000							
Prec	.387	1.000						
z_1^*	.431	- .271	1.000					
z_2^*	.442	.018	- .000	1.000				
z_3^*	- .111	.450	.001	- .000	1.000			
z_4^*	- .186	- .118	.000	- .001	.000	1.000		
z_5^*	.063	- .248	.000	.000	- .000	.000	1.000	
z_6^*	.615	.578	- .000	.000	- .000	.000	- .001	1.000

Analysis of Variance

Source	Degrees of Freedom	Sum of Squares	Mean Square
Regression	7	41.256	5.894
Residual	5	3.690	0.738
Total corrected for mean	12	44.947	

$R^2 = 0.9179$ $R = 0.9581$
$F = 7.99$ Std. error = 0.859

Variable	$\hat{\beta}$	$s_{\hat{\beta}}$	t
Constant	-4.933	8.389	-0.588
Prec	0.466	0.182	2.562
z_1^*	0.645	0.156	4.135
z_2^*	0.727	0.216	3.368
z_3^*	-0.699	0.307	-2.280
z_4^*	-0.343	0.385	-0.890
z_5^*	2.347	1.581	1.485
z_6^*	5.994	3.987	1.503

Prec as an independent variable has resulted in some correlation among the independent variables even though the rotated principle components are orthogonal (table 12.1).

It should be noted at this point that had the last three components been eliminated because they collectively only account for 7.2 percent of the system variance, the resulting equation would not be nearly as good as the final equation that included both the fifth and sixth rotated components.

The final step was to convert the equation into the form of equation 12.34. Let $\underline{R}^{-1}\,\underline{L}^*\,\underline{D}_\lambda^{-1/2} = \underline{E}$. Then equation 12.33 becomes

$$\hat{Y}_i = \beta_o + \beta_1\,Prec_i + \Sigma_{j=2}^{6}\,\hat{\beta}_j\,(\Sigma_{k=1}^{6}\,E_{kj}\,[(X_{ik} - \bar{X}_k)/s_k])$$

Table 12.2. Results of "best" regression using rotated principal components.

Analysis of Variance

Source	Degrees of Freedom	Sum of Squares	Mean Square
Regression	6	40.671	6.779
Residual	6	4.275	0.713
Total corrected for mean	12	44.947	

$R^2 = 0.9049$ $R = 0.9513$
$F = 9.51$ Std. error = 0.8441

Variable	$\hat{\beta}$	$s_{\hat{\beta}}$	t
Constant	-6.471	8.066	-0.802
Prec	0.499	0.175	2.856
z_1^*	0.657	0.152	4.315
z_2^*	0.725	0.212	3.423
z_3^*	-0.735	0.299	-2.458
z_5^*	2.464	1.547	1.592
z_6^*	5.469	3.874	1.412

which after substitution of the correct quantities results in

$$\hat{Y} = -16.050 + 0.499\,\text{Prec} + 1.730\,A + 0.081\,S + 0.695\,P$$

$$- 8.104\,d_1 + 12.057\,R_s + 0.692\,F$$

At this point the reader should compare this last equation with the equation in Table 10.6. This comparison should include the ease of obtaining the equations, the ease of applying the equations, the R^2 and the standard error.

DeCoursey and Deal (1974) present another example of developing a prediction equation for mean annual discharge based on 90 stream locations in Oklahoma. They compare standard multiple regression with several alternate methods of regression using principal components. They conclude that "...any one of the methods is probably as good as the next...". They point out advantages and disadvantages of the methods they investigated.

Wang and Huber (1967) investigated water yield prediction via regression on principal components for watersheds in Utah. In their study they found that a component could be important as a predictor for water yield even though it explained little of the "variance" of the system of independent variables.

CANONICAL CORRELATION

Canonical correlation is used to study the correlation between two sets of variables. Let $\underline{X}$ be a p-variate random vector with $E(\underline{X}) = \underline{\mu}$ and $\text{Var}(\underline{X}) = \underline{\Sigma}$. $\underline{X}$ and $\underline{\Sigma}$ can

be partitioned so that

$$\underline{X} = \begin{bmatrix} \underline{Y} \\ \underline{Z} \end{bmatrix} \qquad \underline{\Sigma} = \begin{bmatrix} \underline{\Sigma}_{11} & \underline{\Sigma}_{12} \\ \underline{\Sigma}_{21} & \underline{\Sigma}_{22} \end{bmatrix} \tag{12.53}$$

where $\underline{Y}$ is $p_1 x 1$, $\underline{Z}$ is $p_2 x 1$, $\underline{\Sigma}_{11} = \text{Var}(\underline{Y})$ is $p_1 x p_1$, $\underline{\Sigma}_{22} = \text{Var}(\underline{Z})$ is $p_2 x p_2$, $\underline{\Sigma}_{12} = \underline{\Sigma}_{21}' = \text{Cov}(\underline{Y}, \underline{Z})$ is $p_1 x p_2$, $p_1 + p_2 = p$, and without loss of generality $p_1 \le p_2$.

Canonical correlation can be used to investigate the correlation structure between the two sets of variables $\underline{Y}$ and $\underline{Z}$. Specifically the correlation between linear functions of $\underline{Y}$ and $\underline{Z}$ will be studied. We can define two new variables by $U_1 = \underline{\alpha}_1' \underline{Y}$ and $V_1 = \underline{\gamma}_1' \underline{Z}$ where $\underline{\alpha}_1$ is a $p_1 x 1$ and $\underline{\gamma}_1$ is a $p_2 x 1$ vector of coefficients. The covariance between U_1 and V_1 is

$$\text{Cov}(U_1, V_1) = \text{Cov}(\underline{\alpha}_1' \underline{Y}, \underline{\gamma}_1' \underline{Z}) = \underline{\alpha}_1' \underline{\Sigma}_{12} \underline{\gamma}_1 \tag{12.54}$$

The variances of U_1 and V_1 are

$$\text{Var}(U_1) = \text{Var}(\underline{\alpha}_1' \underline{Y}) = \underline{\alpha}_1' \underline{\Sigma}_{11} \underline{\alpha}_1$$

$$\text{Var}(V_1) = \text{Var}(\underline{\gamma}_1' \underline{Z}) = \underline{\gamma}_1' \underline{\Sigma}_{22} \underline{\gamma}_1 \tag{12.55}$$

Thus the correlation between U_1 and V_1 is

$$\text{Cor}(U_1, V_1) = \underline{\alpha}_1' \underline{\Sigma}_{12} \underline{\gamma}_1 / \sqrt{\underline{\alpha}_1' \underline{\Sigma}_{11} \underline{\alpha}_1 \; \underline{\gamma}_1' \underline{\Sigma}_{22} \underline{\gamma}_1} \tag{12.56}$$

In canonical correlation it is desired to find the $\underline{\alpha}_1$ and $\underline{\gamma}_1$ that maximize the correlation between U_1 and V_1. Since correlations between pairs of variables are not changed by linear transformations of the variables, a unique solution can be achieved only by requiring that

$$\text{Var}(U_1) = \underline{\alpha}_1' \underline{\Sigma}_{11} \underline{\alpha}_1 = 1$$

$$\text{Var}(V_1) = \underline{\gamma}_1' \underline{\Sigma}_{22} \underline{\gamma}_1 = 1 \tag{12.57}$$

so that

$$\text{Cor}(U_1, V_1) = \underline{\alpha}_1' \underline{\Sigma}_{12} \underline{\gamma}_1$$

Thus it is desired to find the $\underline{\alpha}_1$ and $\underline{\gamma}_1$ that maximize $\underline{\alpha}_1' \underline{\Sigma}_{12} \underline{\gamma}_1$ subject to equations 12.57. The variables U_1 and V_1 are called the first canonical variates and the correlation between U_1 and V_1 is called the first canonical correlation.

Next $U_2 = \underline{\alpha}_2' \underline{Y}$ and $V_2 = \underline{\gamma}_2' \underline{Z}$ are sought so that $\text{Var}(U_2) = \text{Var}(V_2) = 1$, U_2 and V_2 are uncorrelated with U_1 and V_1 respectively, and $\text{Cor}(U_2, V_2)$ is a maximum. U_2 and V_2 thus defined are the second canonical variates and $\text{Cor}(U_2, V_2)$ is the second canonical correlation.

This process is continued until p_1 canonical variates and canonical correlations are found. The equations needed for calculating the $\underline{\alpha}$'s and $\underline{\gamma}$'s may be found in Press (1972) and Cooley and Lohnes (1971).

The hope in using canonical correlation is that the first canonical correlation (or at

least the first few canonical correlations) account for a large part of the correlation structure between $\underline{Y}$ and $\underline{Z}$. This being the case, attention can be focused on the two variates U_1 and V_1 rather than the p variates contained in $\underline{Y}$ and $\underline{Z}$.

The variates $\underline{Y}$ and $\underline{Z}$ used in a canonical correlation analysis are defined prior to the analysis and not as a result of the analysis. DeCoursey (1973) used canonical correlation in a regional flood frequency study. In this study the variables in $\underline{Y}$ where the flood peaks associated with 4 different return periods and the variables in $\underline{Z}$ were 12 watershed characteristics. The purpose of DeCoursey's canonical correlation analysis was to select the variables in $\underline{Z}$ to be used in a regression analysis for predicting $\underline{Y}$. He found that the first two canonical variates were needed to describe his system. The variables selected for use in the regression analysis were those that had relatively large coefficients in either $\underline{Y}_1$ or $\underline{Y}_2$.

MULTIVARIATE REGRESSION ANALYSIS

The situation occasionally arises where it is desired to predict several dependent quantities from the same set of independent variables. Such a situation might be predicting the mean annual flood, 10-year peak flow and 25-year peak flow from a given set of watershed and climatic variables. It may be desirable to use the same independent variables to predict each of the three dependent variables since this would limit the data required.

Such a model can be written

$$\underset{nxq}{\underline{Y}} = \underset{nxp}{\underline{X}} \; \underset{pxq}{\underline{\beta}} \qquad (12.58)$$

where $\underline{Y}$ contains n observations on each of q dependent variables, $\underline{X}$ contains n observations on each of p independent variables, and $\underline{\beta}$ is a pxq matrix of coefficients. Press (1972) discusses this model in more detail.

The coefficients, $\underline{\beta}$, can be estimated in a manner similar to that employed in multiple regression as

$$\hat{\underline{\beta}} = (\underline{X}'\underline{X})^{-1}\underline{X}'\underline{Y} \qquad (12.59)$$

This equation can be written as

$$(\underline{b}_1, \underline{b}_2, \ldots, \underline{b}_q) = (\underline{X}'\underline{X})^{-1}\underline{X}'(\underline{y}_1, \underline{y}_2, \ldots, \underline{y}_q) \qquad (12.60)$$

where $\hat{\underline{\beta}} = (\underline{b}_1, \underline{b}_2, \ldots, \underline{b}_q)$ and $\underline{Y} = (\underline{y}_1, \underline{y}_2, \ldots, \underline{y}_q)$ are partitioned into q pxl vectors. Furthermore

$$\underline{b}_j = (\underline{X}'\underline{X})^{-1}\underline{X}'\underline{y}_j \qquad (12.61)$$

demonstating that the solution to equations 12.59 is equivalent to q multiple regressions each involving the same $\underline{X}$ but a different vector of independent variables. Test of hypothesis concerning $\underline{\beta}_j$ can be made using the procedures set forth in Chapter 10.

In multivariate regression as in multiple regression one commonly has a large number of independent variables all of which are not important in predicting the q dependent variables. If q separate multiple regressions are performed and independent variables eliminated using the procedures of chapter 10, it would be unlikely that the resulting q

equations would contain the same set of independent variables.

If the multivariate regression model is used, all q of the equations given by equation 12.61 will contain the same set of independent variables. Press (1972) presents a procedure for testing the hypothesis that $\underline{\beta}_i = \underline{\beta}_i^*$ where $\underline{\beta}_i$ is a lxq vector made up of the coefficients associated with the i^{th} independent variable for each of the q dependent variables and $\underline{\beta}_i^*$ is a lxq vector of constants. To test that the i^{th} independent variable was not significant would be equivalent to the testing that $\underline{\beta}_i = \underline{0}$. Thus a procedure is available for eliminating variables from equation 12.58 to produce a usable model.

One distinct advantage in using the same independent variables for estimating several dependent variables is that the correlation structure of the dependent variables is preserved. DeCoursey (1973) used such an approach to derive prediction equations for the 2-, 5-, 10-, and 25-year peak flows on watersheds in Oklahoma. In situations like this, it is highly desirable to retain the observed correlations among the dependent variables in the resulting prediction equations. If this is not done, the observed correlations among the various dependent variables are not necessarily retained. In the case of flood peak estimation, this retention of correlation results in stable and consistent estimates for the various peak flows.

Another example where the retention of the correlation structure among a set of dependent variables is important is in estimating the parameters descriptive of runoff hydrographs. Rice (1967) discusses this application of multivariate, multiple regression in simultaneously estimating the runoff volume, peak discharge, and a base time parameter for runoff hydrographs based on data presented by Reich (1962). Rice states that even though three separate regressions produce slightly better fits to the original pool of data, the multivariate solution might be more effective in predicting storm hydrographs for storms on watersheds not included in the original data sample.

Exercises

12.1 Calculate the correlation matrix for the first two variables contained in the table of exercise 10.8.

12.2 Calculate the characteristic values and characteristic vectors associated with the correlation matrix of exercise 12.1.

12.3 Compute the numerical values of the principal components of the data in the first two columns of the table in exercise 10.8 (based on the correlation matrix).

12.4 (a) Work exercise 12.1 using the first three variables. (b) Work exercise 12.2. basec on the first three variables. (c) Work exercise 12.3 based on the first three variables

12.5 (a) Work exercises 12.1, 12.2 and 12.3 based on the covariance matrix. (b) Worl exercise 12.4 based on the covariance matrix.

12.6 Work exercise 12.4 using all of the variables in the table of exercise 10.8 excep Q_p. (Note: Don't try this without a computer – life is too short!)

12.7 Calculate the factor loadings for the data of (a) exercise 12.2, (b) exercise 12.4 c (c) exercise 12.5.

12.8 Show that $\underline{Z}'\underline{Z} = (n-1)\,\underline{D}_\lambda$ by using as an example the data of exercise (a) 12.2, (b) 12.4 or (c) 12.5.

12.9 Investigate the availability of a computer program to perform a varimax rotation of the factor loading matrix. Use the program on the data of exercise 12.7.

12.10 Compute the numerical values for the rotated components based on the varimax rotation requested in exercise 12.9.

12.11 (a) From an examination of the rotated factor loading matrix of exercise 12.5, can any of the variables be eliminated? (b) If so, eliminate them and compute a new rotated factor loading matrix.

12.12 Perform a multiple regression on the rotated principal components of exercise 12.5 or 12.11. Eliminate the nonsignificant components. Put the regression equation in the form of equation 12.34. Compare the results with those of exercise 10.8. Which regression equation do you prefer? Why?

12.13 Perform a multivariate multiple regression using the data of exercise 10.8. Treat both R_s and Q_p as dependent variables and the remaining variables as independent variables.

13. Data Generation

CHAPTER 15 discusses several stochastic models that have been found useful in hydrology. Stochastic models contain random components. These random components contain random elements. If a stochastic model is to be used to generate hydrologic data, methods must be available for generating the random elements of the models.

A random element is usually thought of as an element selected in a fashion such that each element in the population has an equal chance of being selected. If the sample consists of selecting a number at random from a large selection of possible numbers in such a fashion that each number has an equal chance of being selected, the procedure is equivalent to sampling from a uniform distribution. More generally a random element can be selected from any probability distribution as long as the elements are independent of each other. This chapter first sets forth techniques for generating random samples from probability distributions. Next a method for generating a multivariate random sample that preserves the correlations between the variates is presented. Finally several possible areas of application for data generation methods are discussed.

In any application of data generation methods, it must be kept firmly in mind that data generation cannot improve or overcome faulty data. At best one can generate a set of data having statistical properties equal to the properties of the sample used in estimating the population parameters. In addition to this, data stochastically generated is subject to the same sampling errors as natural data. As a matter of fact, data generation has been widely used to study sampling errors. This application is discussed later in the chapter.

UNIVARIATE DATA GENERATION

A random number is defined as a number selected at random from a population of numbers in such a fashion that every number in the population has an equal chance of being selected. A random digit would be one of the numbers 0, 1, ..., 9 selected in such a fashion that any one of these numbers would have an equal ability of being selected. In a sample of 100 random digits, the expected result (b[illegible]with a very low

probability of occurrence) would be ten each of the digits 0, 1, ..., 9.

A random number table is made up of random digits. Many of these tables have the digits grouped into sequences of 5 random digits such as 81425, 13607, etc. A random number of any size can be constructed from these sequences in many ways. If a random number in the interval 0.000 to 0.999 inclusive is desired, any 3 digits of a 5 number sequence could be used. The first digit from 3 successive 5 digit sequences or any other combination of 3 digits could be selected as long as the procedure used did not permit an individual digit to be used in any two of the resulting random numbers. A nine digit random number could similarly be selected by using 4 digits from one sequence and 5 from another. Many large tables of random numbers are available (Rand Corporation 1965). Table E.10 is a table of random numbers. Computer routines for generating random numbers are included as a part of the program libraries for most computers. Care must be exercised when using computer routines in that some generate biased samples.

Many computer routines generate random numbers in the interval (0,1). A random number, Y, in the interval (a,b) can be generated from random numbers in the interval (0,1), R_u, by the relationship $Y = (b - a)R_u + a$.

Random observations may be generated from probability distributions by making use of the fact that the cumulative probability function for any continuous variate is uniformly distributed over the interval 0 to 1. Thus for any random variable Y with probability density function $p_Y(y)$, the variate

$$P_Y(y) = \int_{-\infty}^{y} p_Y(x)\,dx \tag{13.1}$$

is uniformly distributed over (0,1).

A procedure, illustrated in figure 13.1, for generating a random value y from $p_Y(x)$ is

1) Select a random number R_u from a uniform distribution in the interval (0,1).
2) Set $P_Y(y) = R_u$ in equation 13.1.
3) Solve for y.

Step 3 in this procedure is known as obtaining the inverse transform of the probability distribution.

As an example consider the Weibull distribution with

$$p_Y(x) = \alpha(x - \gamma)^{\alpha-1}(\beta - \gamma)^{-\alpha} \exp\{-[(x - \gamma)/(\beta - \gamma)]^{\alpha}\} \tag{13.2}$$

and

$$P_Y(y) = 1 - \exp\{-[(y - \gamma)/(\beta - \gamma)]^{\alpha}\} \tag{13.3}$$

Solving for y results in the inverse transform

$$y = (\beta - \gamma)[-\ln(1 - P_Y(y))]^{1/\alpha} + \gamma \tag{13.4}$$

By substituting R_u for $P_Y(y)$, random values of Y from the 3-parameter Weibull distribution can be generated from

$$y = (\beta - \gamma)[-\ln(1 - R_u)]^{1/\alpha} + \gamma \tag{13.5}$$

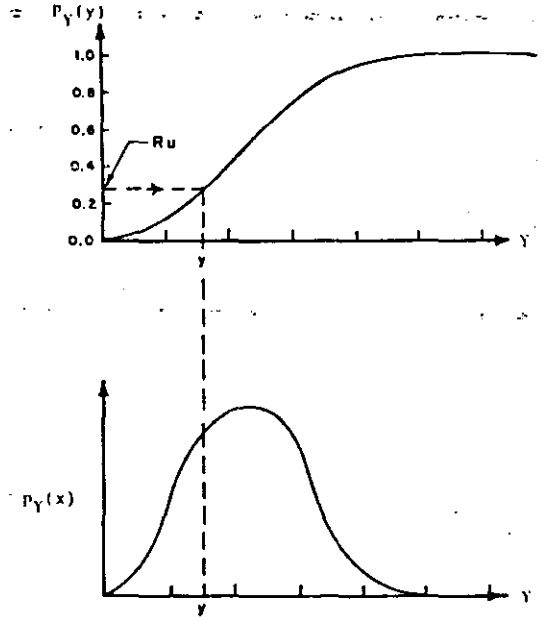


Fig. 13.1. Procedure for generating a random observation from a probability distribution.

For some distributions it is not possible to solve equation 13.1 explicitly for y. That is, an analytic inverse transform cannot be found. The normal and gamma distributions are examples of this. Fortunately in the case of the normal distribution, numerically generated tables of standard random normal deviates are available (Rand Corporation 1965, table E.11). A standard random normal deviate is a random observation from a standard normal distribution. Random observations for any normal distribution can be generated from the relationship

$$y = \sigma R_N + \mu \tag{13.6}$$

where R_N is a standard random normal deviate and μ and σ are the parameters of the desired normal distribution of Y. Computer routines are available for generating standard random normal deviates.

For some distributions relationships with other distributions can be used in the generating process. For example a gamma variate with integer values for η has been shown to be the sum of η exponential variates each with parameter λ. Therefore gamma variates with integer values for η can be generated by summing η values generated from an exponential distribution.

Whittaker (1973) discusses a method for generating random gamma variates with any shape parameter η. Since the gamma distribution is closed under addition, a gamma random variable with any shape parameter can be constructed if one with a shape parameter in the interval $0<\eta<1$ can be constructed. Let R_{u1}, R_{u2} and R_{u3} be independent uniform random variables on (0,1). Define S_1 and S_2 by

$$S_1 = R_{u1}^{1/\eta}, \quad S_2 = R_{u2}^{1/(1-\eta)}$$

then if $S_1 + S_2 \leq 1$, define Y and Z as $Z = S_1/(S_1 + S_2)$ and

$$Y = -Z \ln(R_{u3})/\lambda \tag{13.7}$$

Then Y has a gamma distribution with shape parameter η and scale parameter λ.

This procedure requires the generation of at least 3 uniform random variables. If $S_1 + S_2 > 1$, the R_{u1} and R_{u2} are rejected and new values generated. The probability that $S_1 + S_2 \leq 1$ is given as $\pi\eta(1 - \eta)\,\mathrm{cosec}\,(\pi\eta)$ and has a minimum of $\pi/4$ at $\eta = \frac{1}{2}$ and is symmetric about this value.

To generate a gamma variate with $\eta > 1$, a gamma variate with an integer shape parameter and a shape parameter less than one can be added as long as the scale parameter, λ, is held constant. For example to generate a gamma random variate with $\eta = 3.6$ and any λ, a gamma variate with $\eta = 3$ and λ can be added to a gamma variate with $\eta = 0.6$ and λ.

Table 13.1 presents a summary of some analytical methods for generating observations from selected common probability distributions. The table is modified from Hahn and Shapiro (1967).

Where analytical inverse transforms cannot be found, numerical procedures can be employed. One numerical method is to select a random number between 0 and 1 and then numerically integrate equation 13.1 along the X-axis until the accumulated integral equals the selected random number. At this point y would be equal to the value of X that had been reached.

A second numerical method and one that would be faster if a large number of random observations were needed would be to numerically integrate equation 13.1 starting at the extreme left of the distribution. The integration would proceed to the right in small increments along the X-axis until the accumulated integral was sufficiently close to 1. At each step of the integration, the value of X and the accumulated integral would be saved in the form of a table. The generation process would then consist of selecting a random number in the interval (0,1), entering the table with this random number considered as an accumulated integral, and finding the corresponding value of X. This value of X would then be set equal to the desired random variate Y.

Example 13.1. Generate 22 observations from an exponential distribution with $\lambda = 2$. Plot the observations on semilogarithmic (probability) paper. Estimate $\hat{\lambda}$ from the observations.

Solution: The 22 observations are generated from the relationship $y = -\ln(R_u)/\lambda$ where R_u is a randomly selected number in the interval 0 to 1. The values of Y so generated are shown below. Figure 13.2 is a plot of the resulting numbers along with the lines describing the exponential distribution with parameter $\lambda = 2$ and with parameter $\hat{\lambda} = 1.718$ calculated as $\hat{\lambda} = 1/\bar{y}$.

Comment: This problem illustrates the random variations possible when sampling from a probability distribution. As the sample size increases, $\hat{\lambda}$ should approach λ and the plotted points will lie more nearly on the line describing the exponential distribution with $\lambda = 2$.

In hydrologic frequency analysis the data represent a sample from an unknown population. Thus uncertainty as to the proper frequency distribution exists as well as uncertainty in the values for population parameters.

R_u	y	rank	plotting position
0.329	0.556	9	0.391
.494	.353	12	.522
.051	1.488	3	.130
.187	.838	6	.261
.166	.898	5	.217
.658	.209	15	.652
.608	.249	14	.609
.033	1.706	2	.087
.394	.466	11	.478
.959	.021	22	.957
.735	.154	18	.783
.352	.522	10	.435
.310	.586	8	.348
.754	.141	20	.870
.670	.200	16	.696
.027	1.806	1	.043
.803	.109	21	.913
.585	.268	13	.565
.750	.144	19	.826
.081	1.257	4	.174
.708	.173	17	.739
.266	0.662	7	.304
	Σ 12.806		
	$\bar{y} = .582$	$\hat{\lambda} = 1.718$	

Several exercises at the end of this chapter are designed to help develop a "feel" for the scatter that can be expected when sampling from various frequency distributions. Problems dealing with testing distributional assumptions are also included. These problems demonstrate that for small samples or for similar frequency distributions, a single set of data is not a reliable indicator of the distribution that generated the sample.

MULTIVARIATE DATA GENERATION

Consider the vector $\underline{Y}$ made up of random variables Y_i that are normally distributed.

Random values for this p-variate, normally distributed, 1xp vector $\underline{Y}$ that preserve the means, variances, covariances and correlations between the variables can be generated by using principal components. Recall from Chapter 12 that $\underline{Z} = \underline{X}\,\underline{A}$ where $\underline{Z}$ is an nxp matrix of n values for each of p components, $\underline{X}$ is an nxp matrix of n observations on each of p standardized variables, and $\underline{A}$ is a pxp orthogonal matrix of characteristic vectors of the correlation matrix $\underline{R}$.

Since $\underline{A}$ is orthogonal we have

$$\underline{X} = \underline{Z}\,\underline{A}'$$

or

$$\underline{x} = \underline{z}\,\underline{A}' \tag{13.8}$$

Table 13.1. Simulation of random observations from frequency distributions[1,2]

Distribution	Expression	Relationship for Random Observation y'
Exponential	$p_x(x) = \lambda e^{-\lambda(x-t)}$	$y' = -\ln(R_u)/\lambda + t$
Gamma [integer values of η][3]	$p_x(x) = \lambda^\eta x^{\eta-1} e^{-\lambda x}/\Gamma(\eta)$	$y' = -\Sigma_{i=1}^{\eta} \ln(R_{u_i})/\lambda$
Chi-Square	$p_{\chi^2}(x) = x^{(\nu-2)/2} e^{-x/2}/[2^{\nu/2}\,\Gamma(\nu/2)]$	$y' = \Sigma_{i=1}^{\nu} R_{N_i}^2$
Log-Normal	$p_x(x) = (2\pi x^2 \sigma_{\ln x}^2)^{-\frac{1}{2}} \exp[-\frac{1}{2}(\ln x - \mu_{\ln x})^2/\sigma_{\ln x}^2]$	$y' = \exp(\sigma_{\ln x} R_N + \mu_{\ln x})$
Beta [integer values of α and β]	$p_x(x) = x^{\alpha-1}(1-x)^{\beta-1}\Gamma(\alpha+\beta)/\Gamma(\alpha)\Gamma(\beta)$	$y' = \Sigma_{i=1}^{2\alpha} R_{N_i}^2 / \Sigma_{i=1}^{2\alpha} R_{N_i}^2 + \Sigma_{i=2\alpha+1}^{2\alpha+2\eta} R_{N_i}^2$
Weibull	$p_x(x) = \alpha(x-\varepsilon)^{\alpha-1}(\beta-\varepsilon)^{-\alpha} \exp[-\{(x-\varepsilon)/(\beta-\varepsilon)\}^\alpha]$	$y' = (\beta-\varepsilon)[-\ln(R_u)]^{1/\alpha} + \varepsilon$
Poisson	$f_x(x) = \lambda^x e^{-\lambda}/x!$	$y' = k$ where k is lowest integer such that $\Sigma_{i=1}^{k+1} -\ln(R_{u_i})/\lambda > 1$
Binomial	$f_x(x) = \binom{n}{x} p^x (1-p)^{n-x}$	$y' = \Sigma_{i=1}^{k} k_i$ where $k_i = \begin{cases} 0 \text{ if } R_{u_i} < p \\ 1 \text{ if } R_{u_i} \geq p \end{cases}$

1. R_N is a random observation from a standard normal distribution. R_u is a random number from the interval 0 to 1.
2. Modified from Hahn and Shapiro (c1967).
3. For noninteger values of η use equation 13.7.

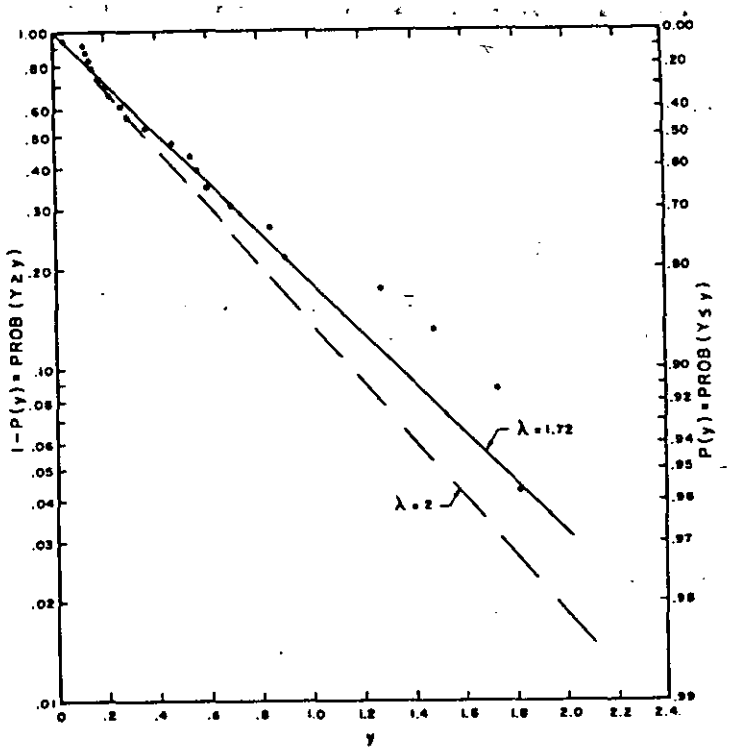


Fig. 13.2. Exponentially generated data for example 13.1.

where $\underline{z}$ is a 1xp vector consisting of a single value for each of the p uncorrelated components. The mean and variance of the j^{th} principal component are 0 and λ_j respectively.

Equation 13.8 can be used to generate standardized normally distributed random variables that preserve their correlation. The components are uncorrelated so that a random value of $\underline{z}_j$ is generated as $\underline{z}_j = (z_1, z_2, \ldots, z_p)$ where z_k is a random observation from a normal distribution with mean zero and variance λ_k. Post multiplying $\underline{z}_j$ by $\underline{A}'$ then produces $\underline{x}_j$. n values for $\underline{x}_j$ can be generated by repeating this process n times. Then

$$\underline{X} = \begin{bmatrix} \underline{x}_1 \\ \underline{x}_2 \\ \cdot \\ \cdot \\ \cdot \\ \underline{x}_n \end{bmatrix} \qquad (13.9$$

is an nxp matrix of standardized normally distributed random variables having the property that $\underline{X}'\underline{X} = (n-1)\underline{R}$.

Letting $\underline{X} = [x_{ij}]$, the matrix $\underline{Y} = [y_{ij}]$ can be computed from $y_{ij} = \sigma_j x_{ij} + \mu_j$ where μ_j and σ_j are the desired mean and standard deviation of the j^{th} variable. The re-

sulting nxp matrix $\underline{Y}$ is made up of n random observations on p variables with the mean and variance of the j^{th} variable being μ_j and σ_j^2 respectively and the correlation between the i^{th} and j^{th} variable being contained in the matrix $\underline{R} = [r_{ij}]$. Since the desired correlations and the variances are produced, the correct covariances are also produced.

Example 13.2. Generate a sample of 20 observations from the three variate normal distribution having the properties $\mu_1 = 3.173$, $\mu_2 = 16.462$, $\mu_3 = 2.566$, $\sigma_1 = 2.113$, $\sigma_2 = 12.481$, $\sigma_3 = 1.150$, $\rho_{12} = -0.1713$, $\rho_{13} = 0.8958$ and $\rho_{23} = -0.2059$.

Solution: This correlation structure corresponds to the correlation matrix in example 12.2. The procedure is to first generate 20 observations from a 3 variate standard normal distribution having the desired correlation structure by 20 applications of equation 13.8. The matrix $\underline{A}$ is contained in example 12.2. A 1 x 3 vector $\underline{z}$ is generated as (z_1, z_2, z_3) where z_i is a random observation from a normal distribution with mean 0 and variance λ_i. The λ_i are obtained from example 12.2 as 1.9692, 0.9273 and 0.1035. The 1 x 3 vector $\underline{x} = (x_1, x_2, x_3)$ is then computed from equation 13.8. Finally a 1 x 3 vector $\underline{y} = (y_1, y_2, y_3)$ is computed as $y_i = x_i \sigma_i + \mu_i$. This process is repeated 20 times generating the required 20 values for $\underline{y}$. The following table contains the resulting 20 observations on $\underline{Y}$.

$$\underline{Y} = \begin{bmatrix} 4.55 & -5.56 & 2.89 \\ 3.57 & -1.94 & 2.37 \\ 5.95 & 1.64 & 4.18 \\ 3.02 & -8.45 & 2.61 \\ 1.42 & 17.40 & 2.07 \\ 3.04 & 14.43 & 1.60 \\ 4.27 & 14.63 & 3.56 \\ 2.24 & 44.56 & 2.42 \\ 0.87 & 30.09 & 1.71 \\ 6.03 & 25.67 & 4.29 \\ 3.46 & 11.00 & 2.33 \\ 4.73 & -0.00 & 4.13 \\ 5.64 & 14.85 & 3.53 \\ 2.73 & 17.57 & 2.56 \\ 4.97 & 11.10 & 3.98 \\ 5.37 & 23.91 & 3.81 \\ 3.00 & 3.58 & 1.78 \\ 3.64 & 11.74 & 2.53 \\ 1.47 & 4.38 & 1.68 \\ 5.31 & 26.00 & 4.03 \end{bmatrix}$$

The means, standard deviations and correlations of this $\underline{Y}$ are shown below. These statistics are not the same as the desired population parameters (as expected) since they are based on a random sample of size 20.

The above procedure was also carried out for samples of 200 and 999 observations with the results shown below. Again it should be kept in mind that these results are based on random samples. A second random sample of the same size would result in different estimates for the population parameters.

n	$\bar{y}_1$	$\bar{y}_2$	$\bar{y}_3$	s_1	s_2	s_3	r_{12}	r_{13}	r_{23}
population	3.17	16.46	2.57	2.11	12.48	1.15	-.17	.90	-.21
20	3.79	12.83	2.90	1.52	13.14	0.94	-.13	.89	.03
200	3.12	17.10	2.58	2.22	11.39	1.20	-.13	.91	-.15
999	3.13	15.48	2.54	2.08	12.38	1.12	-.20	.90	-.23

APPLICATIONS OF DATA GENERATION[1]

Data generation techniques or Monte Carlo simulation have been widely used in hydrology. These uses range from generating large samples of data from known probability distributions to studying the probabilistic behavior of complex water resources systems. The use of simulation in hydrology is certainly not a recent development. In 1927 Sudler (1927) generated a 1,000-year record of annual runoff values to develop probability distributions of reservoir capacities. Chow (1964) has indicated that how much risk and uncertainty are associated with a proposed investment can be estimated by the use of multiple sequences of generated data.

Fiering (1966) has discussed the stochastic simulation of water resources systems. In his paper, he makes the following points:

(1) Synthetic hydrologic traces do not provide a mechanism for overcoming biased or faulty data.
(2) Simulation is not a substitute for analytical solution.
(3) When system simulation appears necessary, it is statistically unjustifiable to rely solely on the observed sequence of hydrologic events.

McMahon et al., (1972) discuss the use of simulated streamflow in reservoir design. Burges and Linsley (1971) investigated the influence of the number of traces used in determining the frequency distribution of reservoir stage. In their study, they generated inflows from both an annual and a monthly, normal, Markov model. They found that, in general, fewer traces were required to define the storage distribution when the monthly model was used than when the annual model was used. They also found that about 1,000 traces should be used to determine the storage frequency distribution when the annual model is used.

Hahn and Shapiro (1967) discuss evaluating system performance by Monte Carlo simulation. Benjamin and Cornell (1970) discuss using simulation to derive the probability distribution of a random variable that is a function of other random variables. Smart (1973) discusses the use of simulation to determine relationships between certain parameters of random geomorphological models. Shreve (1970) used simulation to generate a sample of topologically random channel networks. Fiering (1961) discusses simulation in reservoir design. Fiering and Jackson (1971) develop models for simulating streamflow.

The stochastic nature of quantities estimated from stochastic models can be investigated using data generation techniques. The design of any water resources system is dependent upon estimates of hydrologic quantities. These estimates are based on some type of stochastic model whether it be a flood frequency curve or a comprehensive river basin simulation model. One of the first steps in developing design estimates is the selection of the stochastic model to be used.

1. Some specific stochastic models for generating hydrologic data are discussed in Chapter 15.

Regardless of what stochastic model is finally selected, the parameters of this model must be estimated from historical data. Since the parameters are functions of random variables (the historical data), the parameters themselves are random variables. Furthermore the design estimate that is arrived at using the model is a random variable since it is dependent on the model parameters.

As an example consider the design capacity of a reservoir required to meet a given criterion. This capacity might be determined based on an available historical streamflow record. If a different historical streamflow record were available and was used to determine the required capacity, the estimate based on this historical record would differ from the estimate based on the original historical record. The estimated design would be a random variable since it is a function of the available streamflow record and streamflow is a random variable. Intuitively if two extremely long streamflow records were used, one would expect less difference in the estimated reservoir capacity than if two short streamflow records were used. Furthermore one would expect the estimated capacity based on the long record to more closely approximate the "true"capacity than the estimate based on a short record.

In general the variance of a parameter estimate is a decreasing function of the sample size. The larger the sample, the smaller the variance of the parameter estimate. This in turn implies that the variance of the design estimate will decrease as the sample size increases. The difference in a design estimate and its true population value may be thought of as a prediction error.

A general procedure for determining the probability distribution of prediction errors as a function of sample size is presented in Haan (1972b). The procedure assumes that the correct stochastic model is being employed. The procedure is as follows:

(1) Estimate the parameters of the stochastic model and assume these estimates are equal to the population values.
(2) Simulate k independent sets of data of the type being studied with the model using the assumed population parameters. Each set of data consists of n observations or years of record.
(3) Re-estimate the parameters of the model being used from the n simulated observations for each of the k data sets. This results in k parameter sets.
(4) Estimate the desired quantity, Q, (mean annual runoff, 50-year peak flow, 90-day low flow, etc.) with the model using each of the k parameters sets. This will result in k estimates for Q.
(5) Look at the probability distribution, $p_Q(Q)$, of the k estimates for Q and determine the probability of an individual estimate being outside some acceptable limits. If Q^* represents the estimate of Q, and Q_l and Q_u are the lower and upper limits, then the probability that Q^* will be outside the desired interval is given by

$$1 - \int_{Q_l}^{Q_u} p_Q(t)\,dt$$

(6) Repeat steps 2 through 5 for various values of n, the record length.
(7) Select the record length that gives an acceptable probability (sufficiently low) of Q^* falling outside the interval Q_l to Q_u.

This procedure can be applied to many types of stochastic models. Haan (1972a) presents an illustration of its used in conjunction with the Thomas and Fiering (1962)

streamflow simulation model[2]. In this example a set of population parameters were assumed for the model and k = 100 sets of observations generated. Each set of observations consisted of n years of record. The process was carried out for n = 10, 15, 25 and 50 years. Using each set of observations and each record length, the parameters of the Thomas and Fiering streamflow model were estimated, and from these parameters, the mean annual flow determined by simulation. This gave 100 estimates for the mean annual flow for each of the 4 record lengths. A probability distribution (normal distribution) was fit to the 100 estimates for the mean annual flow. The probability that a single estimate of the mean annual flow would deviate more than a given amount, d, from the population mean annual runoff of 371.6 mm was evaluated. This entire process was repeated 3 times giving a total of 300 simulated traces.

The results of this anlyses, presented in table 13.2, show the expected result that as the number of years of record increase, the probability of making an error greater than a given value decreases. For example, for this particular stream, there is a probability of 0.22 of missing the true mean annual runoff by more than 50.8 mm if 10 years of record are available whereas the probability is only 0.03 if 50 years of data are available for parameter estimation.

This procedure for estimating prediction error probabilities requires that the population parameters for the stochastic model be known. Since this is rarely the case in hydrology, these parameters must be estimated from all of the available information. Obviously, these estimated parameters will not equal the population parameters, but when used as population values along with the above simulation technique, will yield estimates of error probabilities that can serve as a guide in determining how much data is needed to insure an acceptably low probability of making an unacceptable error with the stochastic model.

Exercises

13.1 Without using a table of standard normal deviates generate 20 observations from a normal distribution with a mean of 100 and a variance of 100. What is the mean and variance of the 20 observations?

13.2 Select 100 observations from a normal distribution with mean 0 and variance 1 (use a table of standard normal deviates). Plot a histogram of these observations. Test the hypothesis that these are from a normal distribution using the χ^2 test and the Kolmogorov-Smirnov test. Why do the mean and variance of the data not equal 0 and 1 respectively?

13.3 Generate 20 observations from an exponential distribution with $\lambda = 0.5$. (a) Test the hypothesis that the observations are normally distributed. (b) Test the hypothesis that the observations are exponentially distributed.

13.4 Generate independent samples of size 10, 20, 30, 50 and 100 from a N(0,1). Plot the observations on probability paper using [illegible] plot for each sample. Repeat this entire process 5 times. Study the resulting probability plots in an attempt to develop a "feel" for the scatter that one can expect when sampling from a frequency distribution. (This might be undertaken as a class project with each student working through a sample of 10, 20, 30, 50 and 100. The results can then be shared.)

2. This model is discussed in Chapter 15.

Table 13.2. Probability of error greater than d in mean annual runoff for problem described by Haan (1972a).

n, in years	d in millimeters						
	6.40	12.70	25.40	38.10	50.80	76.20	101.60
10	0.88	0.76	0.54	0.34	0.22	0.06	0.02
15	.86	.72	.46	.28	.14	.04	.00
25	.84	.68	.40	.22	.10	.02	.00
50	.78	.58	.28	.10	.03	.00	.00

13.5 Any number of variations of exercise 13.3 can be worked using different initial distributions, test distributions, parameter values and sample sizes. Some variations should be used to assist in developing a "feel" for the scatter present in random samples from frequency distributions and for the discriminatory power of the chi-square and Kolmogorov-Smirnov tests.

13.6 Write a computer program for generating random observations from a gamma distribution for integer values of η. Generate independent sets of size 10, 20, 30, 40, 50 and 100 observations using $\eta = 2$ and $\lambda = 1.5$. Test the hypothesis that these generated values are from a (a) gamma distribution, (b) normal distribution, (c) exponential distribution.

13.7 Repeat example 13.2 for samples of size 20, 200 and 999. Why are your results not identical to those of example 13.2?

13.8 Weekly rainfall during a particular week of the year at a weather station is thought to follow a gamma distribution with $\eta = 2$ and $\lambda = 1.5$. If 25 years of data are available for estimating the parameters of the gamma distribution, what is the probability that the estimated 50-year weekly rainfall based on the estimated gamma parameters will be in error by more than 0.5 inches?

14. Analysis of Hydrologic Time Series

THIS CHAPTER is an introduction into the analysis of time series of hydrologic data. The purpose of the chapter is to provide an introduction to the terminology and basic concepts used in the analysis of time series. There have been many books and articles written on the subject of time series and stochastic processes. Yevjevich (1972b) has written a book dealing entirely with stochastic processes in hydrology. Bendat and Piersol (1966; 1971) have prepared very readable books on the analysis of random data. Kiesel (1969) and Matalas (1966) have prepared comprehensive treatments of time series analysis of hydrologic data. Two shorter but excellent summaries of many aspects of analyzing stochastic time series in hydrology are presented by Matalas (1967b) and Julian (1967). These references and others contain much more information than is presented here and should be consulted by those requiring more than an introductory knowledge to time series analysis.

DEFINITIONS

A sequence of values collected over time on a particular variable is a time series.

A time series can be composed of a quantity either observed at discrete times, averaged over a time interval, or recorded continuously with time. An ensemble of time series is a set of several time series measuring the same variable. A single time series is called a realization. Thus an ensemble is made up of several realizations.

A time series may be composed of only deterministic events, only stochastic events or a combination of the two. Most generally a hydrologic time series will be composed of a stochastic component superimposed on a deterministic component. For example the series composed of average daily temperature at some point would contain seasonal variation, the deterministic component, plus random deviations from the seasonal values, the stochastic component. The deterministic components may be classified as a periodic component, a trend, a jump or a combination of these. Figure 14.1 shows typical stochastic time series with various types of deterministic components.

Trends in a hydrologic time series can result from gradual natural or man-induced

changes in the hydrologic environment producing the time series. Changes in watershed conditions over a period of several years can result in corresponding changes in streamflow characteristics that show up as trends in time series of streamflow data. Urbanization on a large scale may result in changes in precipitation amounts that show up as trends in precipitation (Huff and Changnon 1973).

Jumps in time series may result from catastrophic natural events such as earthquakes or large forest fires that may quickly and significantly alter the hydrologic regime of an area. Man-made changes such as the closure of a new dam or the beginning or cessation of pumping of ground water may also cause jumps in certain hydrologic time series.

Astronomic cycles are generally responsible for periodicities in natural hydrologic time series. Annual cycles are many times apparent in streamflow, precipitation, evapotranspiration, groundwater level, soil moisture and other types of hydrologic data. Weekly cycles may be present in water use data such as industrial, domestic or irrigation demands. Many times the latter time series will contain both annual and weekly periodicities. Salas-LaCruz and Yevjevich (1972) and Yevjevich (1972c) discuss periodicities and trends in hydrologic data in more detail.

The time scale of time series may be either discrete or continuous. A discrete time scale would result from observations at specific times with the times of the observations separated by Δt or from observations that are some function of the values that actually occurred during Δt. Most hydrologic time series fall in this latter category. Examples would be the average monthly flow in a stream (Δt = 1 month), annual peak discharge

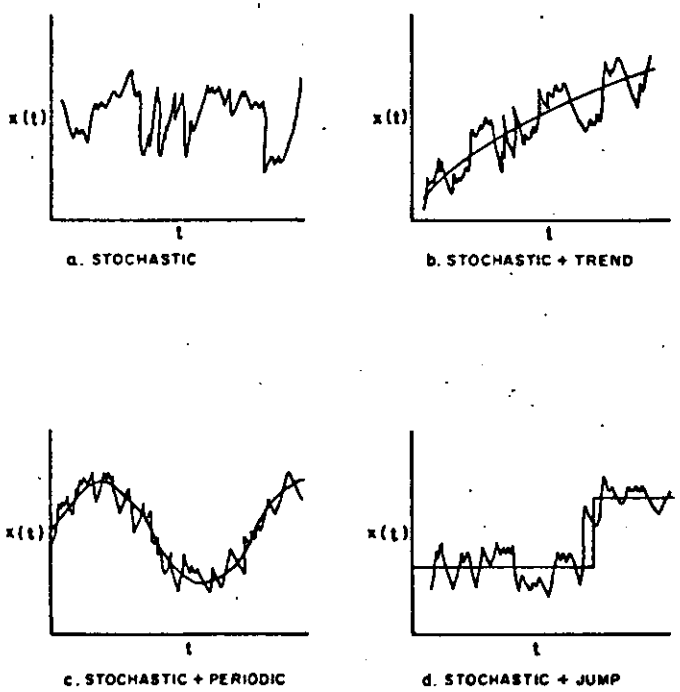


Fig. 14.1. Time series containing stochastic and several types of deterministic components.

(Δt = 1 year), and daily rainfall (Δt = 1 day).

A continuous time scale results when data is recorded continuously with time such as the stage at a stream gaging location. Even when a continuous time scale is used for collecting the data, the analysis is usually done by selecting values at specific time intervals. For example raingage charts are usually analyzed by reading the data at selected times (i.e., every 5 minutes) or at "break points" (here Δt is not a constant). In some research studies, the continuous time data is analyzed via an analog computer. Examples of this would be turbulence studies and pressure measurements.

In this chapter it will be assumed that the data is available at discrete times evenly spaced Δt time units apart. Even though the discussion centers around a time scale concept, a distance or space scale can be used as well. For example the width of a stream along a certain reach might be a stochastic process where the width would be the random variable and distance along the reach the "time".

The random variable described by the time series may be discrete or continuous. A sequence of 0's and 1's denoting rainless and rainy days would be a discrete stochastic process with a discrete time scale. The amount of daily rainfall would be a continuous stochastic process with a discrete time scale (Δt = 1 day).

A stochastic process can be represented by X(t). The probability density function of X(t) is denoted by $p(x;t)$ which describes the probabilistic behavior of X(t) at the specified time, t. If the properties of a time series do not change with time, the series is called stationary. For a stationary series $p(x;t_1)$ equals $p(x;t_2)$ where t_1 and t_2 represent any two different possible times. If $p(x;t_1)$ and $p(x;t_2)$ are not equal, the series is termed nonstationary. Of the series shown in figure 14.1, only that given in 14.1a can possibly be stationary. If the deterministic component is removed from 14.1b, c and d, they too might be stationary.

The properties of a time series can be obtained based on a single realization over a time interval or based on several realizations at a particular time. The properties based on a time interval of a single realization are known as time average properties. The properties based on several realizations at a given time are known as the ensemble properties. If the time average properties and the ensemble properties are the same, the time series is said to be ergodic.

Figure 14.2 shows several possible realizations for a continuous stochastic process on a continuous time scale. The time average over the time interval 0 to T of the i^{th} realization is given by

$$\bar{X}_i = \int_0^T X_i(t)\, dt/T \qquad (14.1)$$

If the realization had been on a discrete time scale, the time average would be determined from

$$\bar{X}_i = \sum_{j=1}^{n} X_i(t_j)/n \qquad (14.2)$$

where n is the total number of equally spaced points at which $X_i(t)$ was observed.

The ensemble average at time t is given by

$$\overline{X(t)} = \sum_{i=1}^{m} X_i(t)/m \qquad (14.3)$$

where m is the number of realizations in the ensemble. As m gets large

$$\overline{X(t)} = \int_{-\infty}^{\infty} X(t)\, p(x;t)\, dx \qquad (14.4)$$

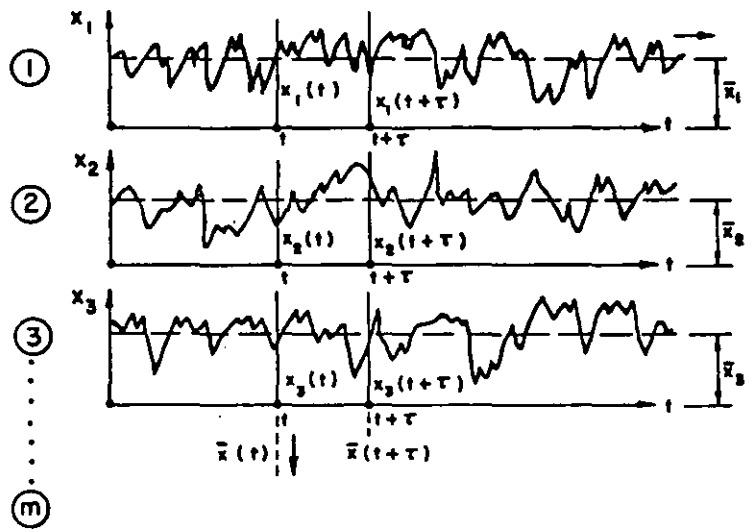


Fig. 14.2. Several realizations of a stochastic process.

If the process is such that $\overline{X(t)} = \overline{X(t+\tau)}$ for all values of t and τ, the process is said to be stationary in the mean or first-order stationary.

The ensemble covariance of X(t) and X(t+τ) is given by

$$\text{Cov}(X(t), X(t+\tau)) = \Sigma_{i=1}^{m} (X_i(t) - \overline{X(t)})(X_i(t+\tau) - \overline{X(t+\tau)})/m \qquad (14.5)$$

If the covariance given by equation 14.5 is independent of t but dependent on τ, the lag, the time series is stationary in the covariance. If $\tau = 0$, equation 14.5 gives the variance of the series. Stationarity in the covariance implies stationarity of the variance. If a series is stationary in the mean and in the covariance, the series is said to be second-order stationary or weakly staionary. If a series is stationary in the covariance but not in the mean, the term weakly stationary or second-order stationary should not be used. For many hydrologic applications, one is satisfied with second-order stationarity. If a process is second-order stationary and p(x;t) is a normal distribution, the process can be shown to be stationary.

Bendat and Piersol (1966) state that in actual practice, random data representing stationary physical phenomena are generally ergodic. For ergodic random processes, the time average mean, as well as all other time average properties, equals the ensemble averaged value. Thus the properties of a stationary random phenomena can be measured properly, in most cases, from a single observed time history record.

Generally only one realization of a stochastic process is available. More than one realization can be obtained by breaking the single realization into several shorter series. Unfortunately, most hydrologic records are so short that breaking them into even shorter series may not be practical. If the statistical properties of the parts of a time series are not significantly different from one another, the series is said to be self-stationary.

For a single realization, the mean is determined from equation 14.1 or 14.2. The covariance can be determined by

$$\text{Cov}(X_i(t), X_i(t+\tau)) = \int_0^{T-\tau} (X_i(t) - \bar{X}_i)(X_i(t+\tau) - \bar{X}_i)dt/(T-\tau) \qquad (14.6)$$

or

$$\text{Cov}(X_i(t), X_i(t+\tau)) = \Sigma_{j=1}^{n} (X_i(t_j) - \bar{X}_i)(X_i(t_j+\tau) - \bar{X}_i)/(n-1) \qquad (14.7)$$

depending on whether the series is on a continuous or discrete time scale.

In the remainder of this chapter, it is assumed that the random or stochastic component of the time series is stationary and ergodic so that the time average properties of the stochastic component can be used. This eliminates the need for more than one realization.

AUTOCORRELATION

One method of describing a stochastic process is with the autocorrelation function $\rho(\tau)$ given by

$$\rho(\tau) = \text{Cov}(X(t), X(t+\tau))/\text{Var}(X(t)) \qquad (14.8)$$

For $\tau = 0$, equation 14.8 indicates that $\rho(0) = 1$ since Cov(X(t), X(t+τ))= Var(X(t)). From figure 14.2 it can be seen that for small values of τ the covariance term would be positive since for the most part like signs are being multiplied $(X(t) - \bar{X}$ and $X(t+\tau) - \bar{X}$ have the same sign for small τ). As τ increases a point is reached where the covariance and thus $\rho(\tau)$ may become negative. Some authors call Cov(X(t), X(t+τ)) the autocorrelation function. In keeping with the terminology established earlier in this book, the Cov(X(t), X(t+τ)) will be called the autocovariance.

A plot of the autocorrelation function against the lag τ is called a correlogram. For random data such as shown in figure 14.1a, the correlogram would appear as in figure 14.3a. In the case of data containing a cyclic and stochastic component such as shown in figure 14.1c, the correlogram would appear as in figure 14.3b where p is the period of the cycle.

Correlograms are useful in determining if successive observations are independent. If the correlogram indicates a high correlation between X(t) and X(t+τ), the observations cannot be assumed to be independent. The autocorrelation function may thus be said to indicate the "memory" of a stochastic process. When $\rho(\tau)$ becomes zero, the process is said to have no memory for what occurred prior to time t-τ. In practice $\rho(\tau)$ should be zero for large τ for most random processes. If $\rho(\tau)$ for large τ exhibits a pattern that is

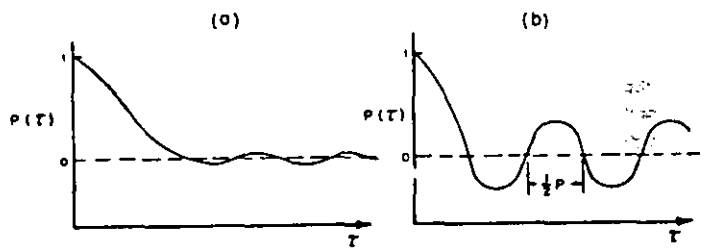


Fig. 14.3. Typical correlograms. (a) Random process. (b) Random process superimposed on a periodic process.

not zero, it may be an indication of a deterministic component. For example, if the correlogram appears as in figure 14.3b, it indicates the data contains a periodic component.

For a discrete time scale, the autocorrelation function becomes $\rho(k)$ where k is the number of time intervals separating $X(t)$ and $X(t+\tau)$. The relationship between τ and k is given by

$$\tau = k\Delta t \tag{14.9}$$

where Δt is the length of the time interval (eg. 1 day, 1 month, 1 year, etc.)

If $\rho(k) = 0$ for all $k \neq 0$, the process is said to be a purely random one. This indicates that the observations are linearly independent of each other. If $\rho(k) \neq 0$ for some $k \neq 0$, the observations k time increments apart are dependent in the statistical sense and the process is referred to as simply a random one. If a time series is nonstationary, $\rho(k)$ will not be zero for all $k \neq 0$ because of the deterministic element even if the random element is itself a purely random time series (Matalas 1967b). Unless the deterministic element is removed, one cannot determine to what extent nonzero values of $\rho(k)$ are affected by the deterministic element.

The population autocorrelation function, $\rho(k)$, may be estimated by $r(k)$ which is given by equation 11.12 with $X_i = X(t_i)$, $X_{i+k} = X(t_i + k\Delta t)$, and n as the total number of observations. Some authors use the terminology autocorrelation function for $\rho(k)$ or $\rho(\tau)$ and serial correlation function for $r(k)$ or $r(\tau)$. This distinction is not made in this text.

For any observed series it is unlikely that $r(k)$ will be exactly zero. If $r(k)$ differs from zero by more than is expected by chance, then the observations k time periods apart cannot be assumed independent. In chapter 11 procedures are given for testing the hypothesis that $\rho(k) = 0$ and for placing confidence intervals on $\rho(k)$.

SPECTRAL ANALYSIS

The autocorrelation function is useful in analyzing a time series in the time domain. Quite often periodicities in data can best be determined by analyzing the time series in the frequency domain.

The theory of variance spectrum analysis postulates that a time series is a sample from a population characterized by a variation over a continuous spectrum of frequencies. The observed time series is a random sample of a process in time (or space) that is made up of oscillations of all possible frequencies. A variance spectrum partitions the variance into a number of intervals or bands of frequency. The quantity usually shown is the spectral density which is the amount of variance per interval of frequency.

The angular frequency ω in radians per unit of time, the period p in units of time and the frequency f in cycles per unit of time are related by $\omega = 2\pi/p = 2\pi f$. Thus one can talk of the spectral density as a function of ω, p or f. In this treatment the spectral density as a function of f, $S(f)$, will be used. $S(f)$ and $S(\omega)$ are related by

$$S(\omega) = 2\pi\, S(f) \tag{14.10}$$

For a completely random series of uncorrelated numbers, the spectral density function $S(f)$ is a constant and is termed white noise. This indicates that no frequency interval contains any more variance than any other frequency interval. In the case of an autocorrelation function, this independence would be shown by $\rho(k) = 0$ for $k \neq 0$.

The mathematical development of the spectral density function indicates the relation between the autocorrelation function and the spectral density function.

$$S(f) = \int_{-\infty}^{\infty} \rho(\tau) \exp(-i\, 2\pi f \tau)\, d\tau = 2\int_{0}^{\infty} \rho(\tau) \cos(2\pi f \tau)\, d\tau \tag{14.11}$$

From a Fourier transformation

$$\rho(\tau) = \int_{-\infty}^{\infty} S(f) \cos(2\pi f \tau)\, df \tag{14.12}$$

For $\tau = 0$ we have $\rho(0) = 1$ and $\cos(0) = 1$ indicating that

$$\int_{-\infty}^{\infty} S(f)\, df = 1 \tag{14.13}$$

which is the normalized variance of the series. Thus $S(f)$ can be viewed as a probability density function that gives the contribution to the total normalized variance contained in the frequency range f_1 to f_2. This contribution can be determined from $\int_{f_1}^{f_2} S(f)df$. If the $\mathrm{Cov}(X(t), X(t+\tau))$ is taken as the autocorrelation function, then $\rho(0)$ is equal to the variance of $X(t)$ (see equation 14.8). In this case $\int_{-\infty}^{\infty} S(f)df = \mathrm{Var}\, X(t)$. We will continue to use as the autocorrelation function $\rho(\tau)$ defined by equation 14.8. The spectral density function is useful in determining which frequencies explain the variance of the time series.

Sometimes $S(f)$ is called a two-sided spectral density function and a one-sided spectral density function is defined as $G(f) = 2\, S(f)$. In reading the literature it is important to note whether $S(f)$, $G(f)$, $S(\omega)$ or $G(\omega)$ is being discussed as one can easily be confused by a factor of 2, π, or 2π.

In nearly all tabulations of hydrologic time series, the time scale is discrete rather than continuous. Spectral theory must then be modified to allow for a discrete set of frequencies over which the variance is to be distributed. For equally spaced observations taken Δt time units apart, the oscillation with the highest frequency for which any information can be obtained is one with a frequency

$$f_N = 1/2\Delta t \tag{14.14}$$

referred to as the Nyquist frequency.

A sample spectral density function $\hat{S}'(f)$ can be computed based on $r(k)$ and numerical integration of equation 14.11.

$$\hat{S}'(f) = \Delta t[r(0) + 2\,\Sigma_{k=1}^{m-1} r(k) \cos(2\pi k f \Delta t) + r(m) \cos(2\pi m f \Delta t)] \tag{14.15}$$

where m is the maximum number of correlation lags. Some authors recommend that m should not exceed 10 to 25 percent of n. The reason for keeping m relatively small in comparison to n is that as m approaches n, the number of observations available for calculating $r(m)$ gets small. If $m = n/2$, then $n/2$ observations are used to calculate $r(m)$. If $m = 0.75n$ then only $0.25n$ observations are used in calculating $r(m)$. Values for $\hat{S}'$ should be computed from equation 14.15 only for frequencies given by

$$f = k f_N / m \tag{14.16}$$

This being the case, equation 14.13 can be evaluated numerically as

$$\int_{-\infty}^{\infty} S(f)\,df = 2\int_{0}^{\infty} S(f)\,df \approx 2\Delta f\{0.5[\hat{S}'(0) + \hat{S}'(f_N)] + \sum_{k=1}^{m-1} \hat{S}'(kf_N/m)\} \tag{14.17}$$

Equation 14.16 gives Δf as f_N/m or $1/(2m\Delta t)$. The term in brackets of the right hand side of equation 14.17 thus must equal $m\Delta t$ since the integral must be 1. A useful check on the computations of $\hat{S}'(f)$ is to see if the terms in brackets actually sum to $m\Delta t$.

It is recommended that the values of $\hat{S}'(f)$ calculated from equation 14.15 be smoothed so that the final estimates for S(f) are given by

$$\begin{aligned}
&\hat{S}(0) = 0.5(\hat{S}'(0) + \hat{S}'(f_N/m)) \\
&\hat{S}(kf_N/m) = 0.25\,\hat{S}'((k-1)f_N/m)) + 0.5\,\hat{S}'(kf_N/m) + 0.25\,\hat{S}'((k+1)f_N/m) \\
&\text{for } k = 1, 2, \ldots, m-1 \\
&\hat{S}(f_N) = 0.5(\hat{S}'((m-1)f_N/m) + \hat{S}'(f_N))
\end{aligned} \tag{14.18}$$

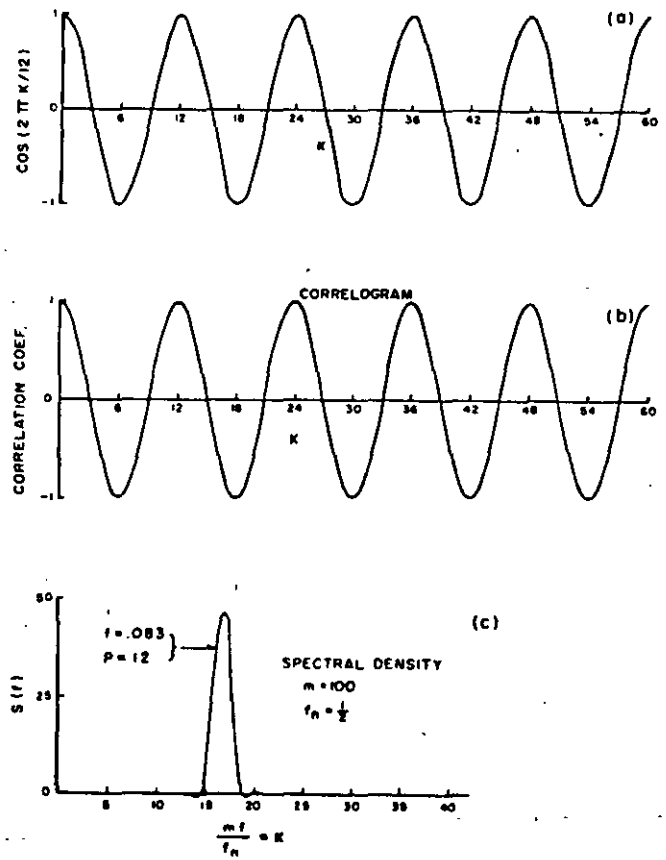


Fig. 14.4. (a) $\cos(2\pi k)/12$ and its correlogram (b) and spectral density (c).

Bendat and Piersol (1966) present the necessary steps for sampling a time series, determining Δt, n and m, and computing r(k) and $\hat{S}(f)$ to attain certain accuracy goals. Since for most hydrologic applications Δt and n are fixed, these procedures will not be covered here.

EXAMPLES OF AUTOCORRELATION AND SPECTRAL DENSITY FUNCTIONS

If X(t) is an analytic function of time, it is theoretically possible to calculate $\rho(\tau)$ and S(f) directly from equations 14.8 and 14.11 respectively. For example if $X(t) = \sin(2\pi f_o t)$ or $X(t) = \cos(2\pi f_o t)$, then $\rho(\tau) = \cos(2\pi f_o t)$ and S(f) is a Dirac delta function with a spike at $f = f_o$. Figure 14.4 shows a function $X(k) = \cos(2\pi k/12)$ and the corresponding r(k) and $\hat{S}(f)$. r(k) and $\hat{S}(f)$ where computed from equations 11.12, 14.5, and 14.18 based on m = 100 and $\Delta t = 1$. As can be seen from figure 14.4c, the conversion of a continuous function $X(t) = \sin(2\pi t/12)$ to a discrete function $X(k) = \sin(2\pi k/12)$ results in both r(k) and $\hat{S}(f)$ being approximations for $\rho(\tau)$ and S(f). The periodicity of $f_o = 1/12$ is apparent in both the correlogram and the spectral density.

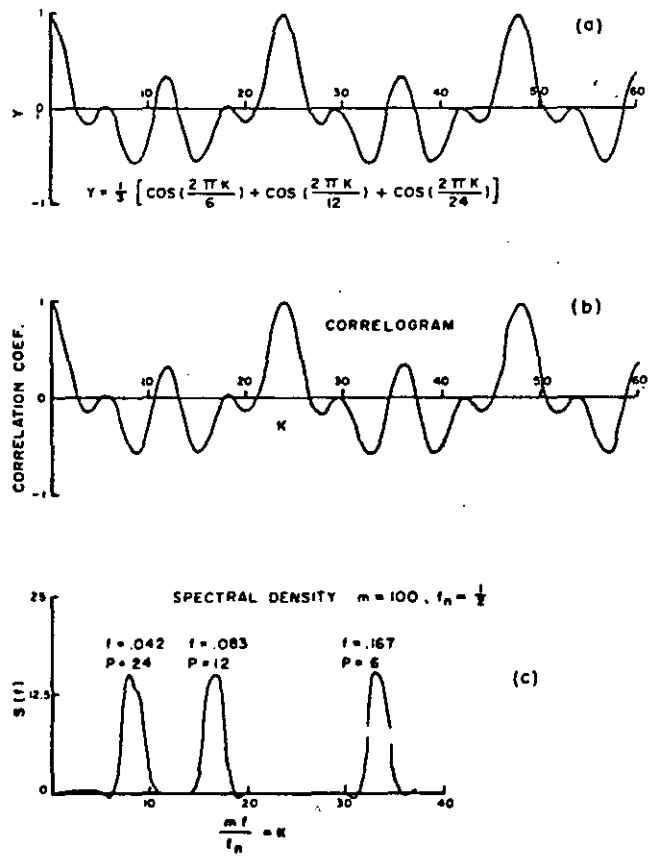


Fig. 14.5. (a) Sum of 3 cosines and its correlogram (b) and spectral density (c).

In figure 14.4c, $\hat{S}(f)$ is plotted against k. According to equation 14.7, the area under the S(f) versus f curve from f = 0 to f = ∞ should be ½. Since $k = mf/k_N$, $f_N = 1/2\Delta t$, and $\Delta t = 1$, the area under the S(f) versus k curve should be $m\Delta t$ or 100 in the case of figure 14.4.

Figure 14.5 presents the correlogram and spectral density for $X(t) = [\cos(2\pi t/6) + \cos(2\pi t/12) + \cos(2\pi t/24)]/3$. Again the correlogram is exactly equal to X(k) for this deterministic function. The spectral density now has 3 spikes corresponding to the 3 frequencies 1/6, 1/12, and 1/24. If a random component ε had been added to X(t), the correlogram would lose some of its regularity making it more difficult to spot the periodicities present in the data. The spectral density would, however, retain its 3 spikes.

Figure 14.6 presents the peak flow data for the Kentucky River near Salvisa, Kentucky (Table 2.1). Here the correlogram and spectral density contain no apparent patterns. The correlogram in this case indicates that the annual peaks are independent of each other. The 95% confidence limit for r(k) is shown as calculated based on equation 11.14 as follows.

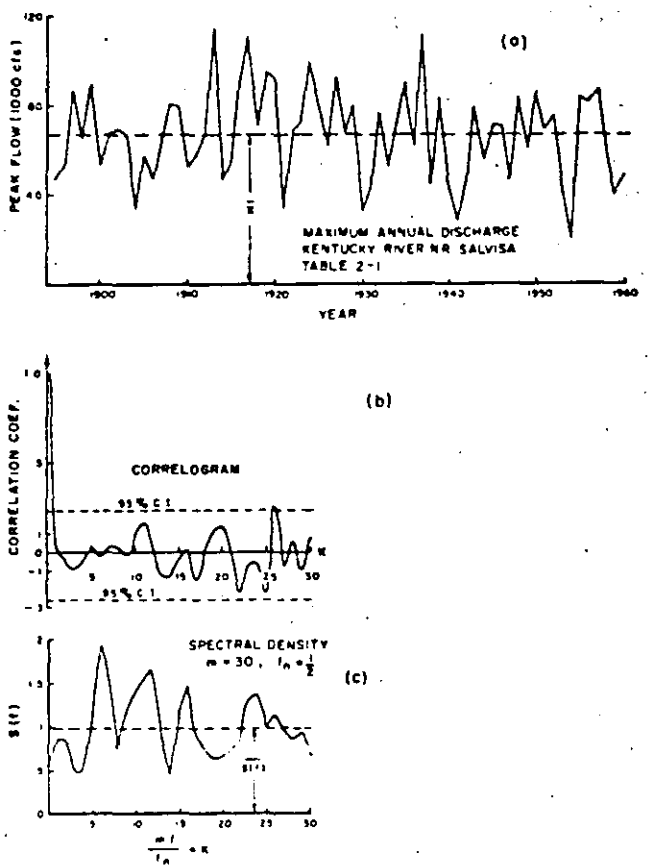


Fig. 14.6. (a) Kentucky River peak flows and its correlogram (b) and spectral density (c).

$$\begin{matrix} U \\ L \end{matrix} = (-1 \pm Z_{.975}\sqrt{64})/65 = (-1 \pm 1.96 \times 8)/65$$

$$U = 0.226 \qquad L = -0.257$$

The values for r(k) for k near n are not reliable and the confidence intervals calculated from equation 11.14 are not valid in that region.

The spectral density shown in figure 14.6c oscillates about the mean value of $\hat{S}(f)$ which is $m\Delta t/n = 1$. This is a typical "white noise" spectrum indicating that the variance is equally distributed among all frequencies. Both the correlogram and spectral density indicate that the peak flows on the Kentucky River near Salvisa can be considered a purely random stochastic process. It is this independence that makes the reordering of the peak values by ranking them from largest to smallest, a valid procedure to use in flood frequency studies.

Figures 14.7 shows the monthly runoff data for Cave Creek watershed located near

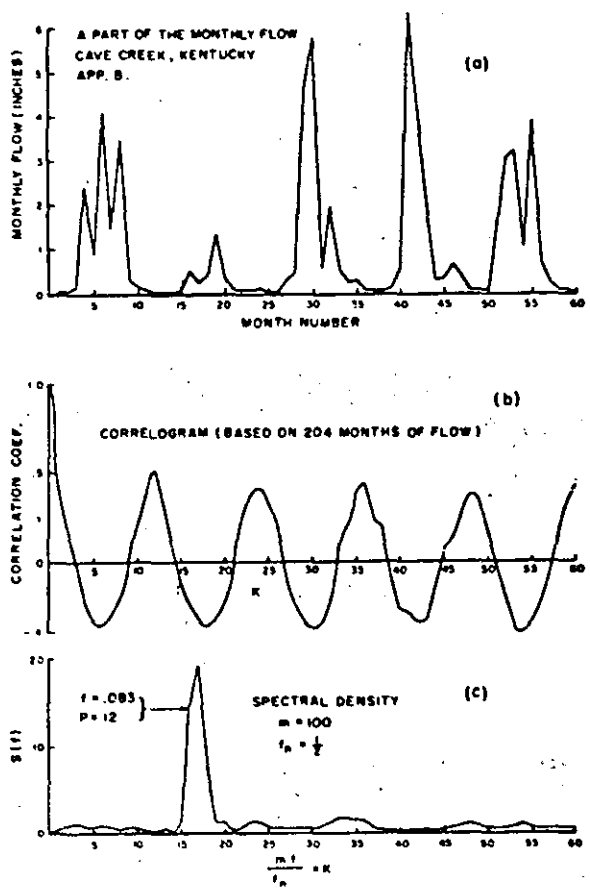


Fig. 14.7. (a) Cave Creek monthly runoff and its correlogram (b) and spectral density (c).

Lexington, Kentucky. Only a part of the record is plotted in figure 14.7a. The complete record as shown in Appendix C was used in determining the correlogram and spectral density. Both the correlogram and the spectral density indicate that monthly runoff for this watershed has a definite annual cycle. The smoothness of both the correlogram and the spectral density in comparison to the data should be noted. A comparison of figures 14.3a and 14.7b indicates that the monthly runoff is a periodic (deterministic) process with a random component. The spectral density clearly indicates that the period is 12 months.

SUMMARY

This chapter has presented an introduction to time series analysis. The tools available for studying time series are much broader and varied than has been covered in this chapter. Time series in general are composed of deterministic and stochastic components. The deterministic component may be in the form of trends, jumps or periodicities. The stochastic component may be an independent stochastic process (purely random) or a dependent stochastic process (random). Any stochastic hydrologic model must contain adequate representations of both the deterministic and stochastic components of the time series.

A correlogram is a useful device for determining the degree of dependence present in successive values of a time series. If $r(k)$ for k sufficiently large (with $k < n/4$) is significantly different from zero, the possible presence of a deterministic component should be investigated. The correlogram of the stochastic component can be used to determine the degree of dependence in the stochastic component. For the stochastic component, if $r(k) = 0$ for all $k \neq 0$, the process is a purely random process. A periodic correlogram is indicative of a periodic time series.

The spectral density is most useful for isolating periodicities in a time series. The presence of a cycle and its frequency are shown by the spectral density function. For a purely random process, the spectral density oscillates randomly about a constant value indicating that no particular frequency explains any more of the variance of the time series than any other frequency. Such a case is termed "white noise".

Exercises

14.1 Let $X(k) = \cos(2\pi k/12) + \epsilon$ where ϵ is an independent random observation from a normal distribution with a mean of zero and a variance of 0.5. Compute and plot the correlogram and spectral density function for this process by generating values for $X(k)$ for $k = 0$ to 200. Compare the results with figure 14.4.

14.2 Let $X(k) = [\cos(2\pi k/6) + \cos(2\pi k/12) + \cos(2\pi k/24)]/3 + \epsilon$ where ϵ is an independent random observation from a normal distribution with a mean of zero and a variance of 0.5. Compute and plot the correlogram and spectral density function for this process by generating values for $X(k)$ for $k = 0$ to 200. Compare the results with figure 14.5.

14.3 The following data represent the years of eruptions of the volcano Aso for the period 1229 to 1962. Let k equal the eruption number beginning in 1239 so that $k = 1$ for 1239, $k = 2$ for 1240, $k = 3$ for 1265, etc. Let $X(k)$ be the number of years since the last eruption so that $X(1) = 10$, $X(2) = 1$, $X(3) = 25$, etc. Compute the correlogram and

spectral density function for $X(k)$. Is there any apparent pattern to the eruptions?

Years of Eruptions of the Volcano Aso for the Period 1229-1962

1229	1340	1533	1828	1931
1239	1346	1542	1829	1932
1240	1369	1558	1830	1933
1265	1375	1562	1854	1934
1269	1376	1563	1872	1935
1270	1377	1564	1874	1938
1272	1387	1576	1884	1949
1273	1388	1582	1894	1950
1274	1434	1780	1897	1951
1281	1438	1804	1906	1953
1286	1473	1806	1916	1954
1305	1485	1814	1920	1955
1324	1505	1815	1927	1956
1331	1506	1826	1928	1957
1335	1522	1827	1929	1958
				1962

Data from Davis (1973).

14.4 The following data represent an unusual phenomenon in that they are observations of a true time series from the geologic past. The Ecocene Lake deposits of the Rocky Mountains consist of thinly laminated dolomitic oil shales hundreds of feet thick. It has been established that the laminations are varves, or layered deposits caused by seasonal climatic changes in the lake basins. By measuring the thickness of these laminations, a record of the annual change in the rate of deposition through the lakes history is obtained (Davis 1973). Compute and plot the correlogram and spectral density function for this data. Discuss any apparent patterns.

Thickness of Successive Varves of a Section through the Green River Oil Shale

Thickness (mm)

(Top of Section)									
6.0	12.0	7.1	14.0	8.4	11.0	3.8	8.1	7.3	13.2
7.2	11.0	7.2	13.6	7.9	10.8	3.4	7.8	6.7	12.4
7.1	9.6	7.1	12.1	7.0	9.5	4.2	6.4	6.0	9.7
7.1	8.7	7.0	12.9	6.7	8.1	4.8	4.6	5.8	9.2
7.2	7.6	7.0	12.8	6.8	7.2	4.5	3.7	5.7	9.3
7.4	7.2	7.7	11.1	7.3	7.1	3.6	4.0	6.5	8.3
8.0	7.2	8.6	9.0	7.3	6.8	3.0	4.2	8.2	6.0
8.6	7.8	9.0	7.5	7.2	7.0	2.8	4.5	10.2	5.7
10.0	8.1	12.0	7.5	8.1	7.1	4.1	5.9	12.3	6.1
11.4	7.8	13.7	8.4	9.8	5.6	6.8	7.3	13.2	6.3
									6.3
									(Bottom of Section)

Data from Davis (1973).

14.5 Compute and plot the correlogram and spectral density function for the weekly precipitation at Ashland, Kentucky, for the week of March 1-7. Discuss any apparent patterns. (Data in Appendix C).

14.6 Work exercise 14.5 using the monthly precipitation data for Walnut Gulch near Tombstone, Arizona. (Data in Appendix C).

15. Some Stochastic Hydrologic Models

EVERY DESIGN decision requiring hydrologic knowledge is based on a hydrologic model of some type. This model might be one that gives the peak discharge from a small watershed as some function of the watershed area, it might be a flood frequency curve, it might be a comprehensive "deterministic" model capable of generating synthetic streamflow records or it might be a stochastic model for generating a time series of hydrologic data. Deterministic is used here in the sense that once the model parameters are known, the same inputs to the model always produce the same outputs. Thus parametric models are included under "deterministic" even though the model parameters may be functions of observed hydrologic records and thus be random variables.

Ultimately design decisions must be based on a stochastic model or a combination of stochastic and deterministic models. This is because any system must be designed to operate in the future. Deterministic models are not available for generating future watershed inputs in the form of precipitation, solar radiation, etc., nor is it likely that deterministic models for these inputs will be available in the near future. Stochastic models must be used for these inputs.

If a design is based solely on the basis of a historical record of rainfall or streamflow, the stochastic model employed is simply the historical record itself. It should be kept in mind that any historical record is but one realization of a stochastic time series and that future realizations will resemble the historical record only in a statistical sense even if the process is stationary.

Designers of water resources systems have realized for years that evaluating their designs using past or historical records provided no guarantee that the design would perform satisfactorily in the future because future flow sequences will not be the same as past flow sequences. Typically historical flow sequences are quite short – generally less than 25 years in length. Even during the 100-year life of a project, an observed historical flow sequence of 10 to 25 years in length will not repeat itself. In all cases the designer would agree that the worst flood (or drought) on record is not the worst possible flood (or drought).

The use of historical records alone gives no idea of the risk involved. That is, if a

design is made based on an historical record, chances are that the design would be adequate if the historical record repeated itself. However, we know that the historical record will not repeat itself. There is thus a certain risk that the design will be inadequate for the unknown flow sequence that the system will actually experience.

This latter point can be illustrated by considering the design of a facility that might have a 5-year life – say a small, temporary boat dock. For this design assume that 100 years of flow records are available. The 100 years of record would provide 20 independent 5-year flow sequences. A proposed design could then be evaluated on 20 independent flow sequences equal in length to the design life of the facility. If it is found that in 15 of the sequences, the design is adequate and inadequate in the remaining 5, then one would estimate that for some future 5-year period there would be a probability of 0.25 (5/20) that the design would be inadequate. If the design is adopted, a risk of 0.25 exists. A risk of 0.25 may be unacceptable. For this case consider that an acceptable risk is 0.05. This means the design should be increased and reevaluated until it proves inadequate in only 1 of the 20 5-year observed sequences. Generally the design life of a water resources project exceeds the length of available record so that a risk evaluation using this procedure is not possible.

In other cases it may be desirable to know the severity of a shortage. For instance, one might be looking at a system to control the thermal pollution from a power plant. It may be that the design requirement is to affect the natural water temperature by less than 5 °C. During low flows the ratio of the volume of heated water discharge to the volume of natural flow may be such that it is difficult to keep the overall temperature rise to less than 5 °C. In this case it would be desirable to know the magnitude as well as the frequency of failing to meet the design standard since a 6 °C temperature rise would be less damaging than a 15 °C temperature rise.

The approach outlined above assumes that there is some probabilistic mechanism underlying the generation of streamflows and that this mechanism is sufficiently stable that it can be considered stationary. It is also assumed that the sample in hand is a representative one.

An even better approach to determining risk probabilities would be through operations (analytic or Monte-Carlo) on the underlying, exact probability distribution or distributions that the natural hydrologic process follows. Of course, this type of information is never available and in practice must be approximated. Even if the exact distribution was known, its parameters would have to be estimated from an observed record (sample) and would not equal the population parameters. Thus to overcome the objections of design evaluation based on a single (and many times short) flow trace, a data generation scheme or stochastic model is needed.

A stochastic model is a probabilistic model having parameters that must be obtained from observed data. Stochastic streamflow models for example do not convert rainfall to runoff through theoretical or empirical relationships as do deterministic models but utilize the information in past or historical streamflows. Stochastic streamflows are neither historical flows nor predictions of future flows, but they are representative of possible future flows in a statistical sense. Stochastically generated data can be used in evaluating risk probabilities providing a "satisfactory" stochastic model is available.

It should be noted that a stochastic model depends heavily on the assumptions of stationarity and representativeness. The effects of watershed changes for example cannot be evaluated. On the other hand deterministic models might be able to simulated a changing hydrology as the basin changes – but remember the future rainfall problem.

Thus one approach to watershed modeling is to stochastically generate rainfall and use a deterministic model to convert the rainfall to streamflow.

In developing a stochastic model it is assumed that the data is the result of a random process or one that involves chance. One cannot precisely state what the data values will be at any particular future time, but will be able to make statements of probability concerning future data values. In looking over past data records it is apparent that streamflows are not completely random with no constraints, but do possess certain recognizable features. For instance, if the average annual flow has been around 15 inches for a long period of time, it is unlikely that it will suddenly change to 25 inches unless the watershed is altered in some fashion. If the flows have tended to be between 10 and 20 inches per year with only an occasional yearly total outside these limits, the model should not produce a large number of flows outside these limits. Thus the model should preserve the overall mean and spread or variance of the data.

Further it may be noted that there is some degree of persistence in that low flows tend to follow low flows and high flows tend to follow high flows. A streamflow model should retain this property. From this it can be seen that historical records certainly guide us in model development.

It is not the purpose of this chapter to promote stochastic hydrologic models. Rather some of the most prominent models are discussed. There is a very rapidly expanding literature on stochastic hydrologic models. No attempt is made to cover all of the models currently in the literature or to discuss all of the features of the models that are covered in the chapter.

In selecting a stochastic model it is important to be able to state what characteristics of the phenomena being modeled are important and what characteristics are unimportant. For example, if streamflow is being modeled, the following is a partial listing of the questions that must be considered.

1. Is it necessary to model the peak flows?
2. Are annual peaks sufficient or will other peaks occurring during the year be important?
3. Is the time during the year when the peak occurs important?
4. Is the sequence or order of occurrence of the peaks important?
5. Is the simultaneous occurrence of a peak flow and some other event important?
6. Is the volume of flow important?
7. Is daily, weekly, monthly or annual volume of flow required?
8. Is the simultaneous volume and peak during some interval required?
9. Is the seasonality in volumes for durations of less than a year important?
10. Is the dependence of the flow in one time period on the flow in previous time periods important?
11. Is it sufficient to model the mean flow in a period? Is the variance important too? What about the skewness?
12. Is the relationship of the flow on one stream and that on nearby streams of concern?
13. Is the time series of flows stationary?
14. Is there evidence of trends or jumps in the flow record? Will there be trends or jumps in the future? Is it important to model these features?
15. How well do the above properties have to be modeled?
16. What is the quality and quantity of data available for model selection and parameter estimation?

17. Given the available historic data, can the model parameters be estimated with sufficient accuracy?

The answers to these and many other questions must be obtained before a model can be applied to a particular problem. Of course, the answers to these questions depend heavily on the use that is to be made of the model. Generally it is desirable to select the simplest model that will provide the information needed.

If one is considering the design of a reservoir to be located on a single stream to provide irrigation water, quite likely a model that is capable of producing synthetic monthly streamflows could be used. If a more complex model that considers daily flows or peak flows is used, it may be necessary to sacrifice accuracy of monthly flow simulation to obtain accuracy of daily flows or peaks. In any case the more complex model would be more expensive to develop and test and certainly more expensive to use in developing long synthetic streamflow traces. On the other hand if the proposed reservoir is to also provide flood control benefits, then estimates of flood peaks and possibly shorter duration flow volumes would be required.

The design of an irrigation reservoir illustrates the case where joint probabilities or joint hydrologic time series may be required. It may be that the land to be irrigated and the watershed supplying waters to the reservoir are subjected to similar climatic conditions. Thus the irrigation demand may be the greatest at precisely the same time that flows into the reservoir are the lowest and vice versa. Neglecting this possible correlation between supply and demand can result in an underdesigned reservoir. If the demand and supply are independent, designing for the highest demand at a time of the lowest supply may be overdesigning since the joint occurrence of these events may have a very low probability.

There is no substitute for a thorough knowledge of the problem to be solved and the features of the problem that must be reproduced by the simulation model. It is relatively easy to develop a simulation model for a problem by making unrealistic simplifying assumptions. It is difficult to develop a model for use in solving a problem as it really exists. It is generally better to develop an approximate solution to the real problem than an exact solution to an unreal problem.

In chapter 14 it was stated that a hydrologic time series may contain trends and/or jumps. If an historical record contains trends and/or jumps and it is desired to use this record for the estimation of the parameters of a stochastic model, it is necessary to be able to separate the deterministic and stochastic components of the historical record. Once the deterministic component is removed, the stochastic component can be used for parameter estimation. The model that is developed may have to incorporate both deterministic and stochastic components. This is especially apparent in cases where trends are present. If the trend is expected to continue into the period being modeled, the trend component must be present in the model that is developed. Trends can generally be modeled by a polynomial equation of the type

$$T_p(t) = \beta_0 + \beta_1 t + \beta_2 t^2 + \ldots$$

where $T_p(t)$ represents the trend in the parameter p as a function of time t and β_0, β_1, β_2, ... are coefficients that may be estimated by multiple regression. The order of the polynomial can also be tested by determining the highest order term having a regression coefficient that is significantly different from zero.

Jumps in a hydrologic time series may be identified by computing the mean value of the parameter of interest during the two time periods on either side of the jump.

These two means are then tested to see if they are significantly different from each other. The exact time at which a jump occurs cannot be easily identified from the data alone because of the presence of stochastic variation. In this case a review of the data gathering procedure and factors affecting the variable under study should be undertaken in an attempt to identify possible causes for the jump and the time that these factors became important.

Many stochastic models require the estimation of a large number of parameters. Again the limited hydrologic data that is available at a point may be inadequate to estimate these parameters. A regional approach to parameter estimation may help in this situation provided regional data are available (Benson and Matalas 1967; also chapters 7 and 10).

Two classical stochastic models, the Bernoulli process and the Poisson process, and some of their potential applications in hydrology were discussed in chapter 4. The remainder of this chapter is devoted to other selected models that appear frequently in the hydrologic literature.

PURELY RANDOM STOCHASTIC MODELS

Possibly the simplest stochastic process to model is where the events can be assumed to occur at discrete times with the time between events constant, the events at any one time are independent of the events at any other time, and the probability distribution of the event is known. Stochastic generation from a model of this type merely amounts to generating a sample of random observations from a univariate probability distribution as shown in chapter 13.

This type of model might be appropriate for generating a synthetic record of flood peaks. Problems with the method are the uncertainty as to the proper probability distribution to use and the uncertainty in the parameter values of the probability distribution. These two types of uncertainty exist in all stochastic models to some extent. The larger the sample for estimating the model parameters and testing the derived model, the less will be these uncertainties. Regional data can also be used in some situations to assist in distribution selection and parameter estimation.

Another slightly more advanced application of a purely random model might be in generating sequences of point storm rainfall amounts. The time between storms might be modeled as an independent Poisson or Bernoulli process (Lane and Osborn 1973) and the amount of rain as a gamma variable. The model could be made more complex by assuming the distribution parameters are a function of the time of year or that the parameters of the gamma distribution depend on the generated time since the last storm.

As indicated in chapter 14, whether or not a process can be considered as a purely random process may be indicated by its correlogram or spectral density. If r(k) is not significantly different from zero for k greater than zero or if $\hat{S}(f)$ oscillates randomly about $\bar{S}(f)$, the process may be a purely random process. The difficulties in selecting the proper probability density function and in parameter estimation remain, however.

FIRST ORDER MARKOV PROCESS

Many hydrologic time series exhibit significant serial correlation. That is, the value of the random variable under consideration at one time period is correlated with the values of the random variable at earlier time periods. The correlation of a random variable X at one time period with its value k time periods earlier is denoted by

$\rho_X(k)$ and is called the k^{th} order serial correlation. If $\rho_X(k)$ can be approximated by $\rho_X(k) = \rho_X{}^k(1)$, then the time series of the random variable X might be modeled by a first order Markov process. The first order Markov process might also be used for a model if serial correlation for lags greater than one are not important.

A first order Markov process is defined by the equation

$$X_{i+1} = \mu_X + \rho_X(1)(X_i - \mu_X) + \epsilon_{i+1} \tag{15.1}$$

where X_i is the value of the process at time i, μ_X is the mean of X, $\rho_X(1)$ is the first order serial correlation and ϵ_{i+1} is a random component with $E(\epsilon) = 0$ and $Var(\epsilon) = \sigma_\epsilon^2$. This model states that the value of X in one time period is dependent only on the value of X in the preceding time period plus a random component. It is also assumed that ϵ_{i+1} is independent of X_i. The variance of X is given by σ_X^2 and can be shown to be related to σ_ϵ^2 by

$$Var(X_{i+1}) = \sigma_X^2 = E[\mu_X + \rho_X(1)(X_i - \mu_X) + \epsilon_{i+1}]^2 - E^2(X_{i+1})$$

$$= \rho_X^2(1)\,\sigma_X^2 + \sigma_\epsilon^2$$

or

$$\sigma_\epsilon^2 = \sigma_X^2\,[1-\rho_X^2(1)] \tag{15.2}$$

If the distribution of X is $N(\mu_X, \sigma_X^2)$ then the distribution of ϵ is $N(0, \sigma_\epsilon^2)$. Random values of X_{i+1} can now be generated by selecting ϵ_{i+1} randomly from a $N(0, \sigma_\epsilon^2)$ distribution. If t is N(0,1) then $t\sigma_\epsilon$ or $t\sigma_X\sqrt{1-\rho_X^2(1)}$ is $N(0, \sigma_\epsilon^2)$. Thus a model for generating X's that are $N(\mu_X, \sigma_X^2)$ and follow the first order Markov model is

$$X_{i+1} = \mu_X + \rho_X(1)(X_i - \mu_X) + t_{i+1}\,\sigma_X\sqrt{1-\rho_X^2(1)} \tag{15.3}$$

The procedure for generating a value for X_{i+1} is to estimate μ_X, σ_X, and $\rho_X(1)$ by $\bar{X}$, S_X, and $r_X(1)$ respectively and then select a t_{i+1} at random from a N(0,1) distribution and calculate X_{i+1} based on $\bar{X}$, S_X, $r_X(1)$ and X_i. The first value of X_i, X_1, might be selected at random from a $N(\mu_X, \sigma_X^2)$. To eliminate the effect of X_1 on the generated sequence, the first 50 or 100 generated values might be discarded.

Equation 15.3 has been widely used for generating annual runoff from watersheds (Fiering and Jackson 1971). Since t is N(0,1), it is possible to generate values of X that are less than zero. If this occurs it is generally recommended that the negative X be used to generate the next value for X and then discarded. This procedure will result in a slight bias. If the occurrence of negative X's is common in the generation process, it may indicate that X is not normally distributed. In this event some other distribution of ϵ must be used.

Equation 15.3 generates normally distributed X's with a mean of μ_X, variance of σ_X^2 and first order serial correlation of $\rho_X(1)$. Serial correlation is common in hydrology and depending on the use to be made of the model may be quite important. Note that if $\rho_X(1) = 0$, equation 15.3 reduces to the independent process of selecting a random observation from $N(\mu_X, \sigma_X^2)$. On the other hand if $\rho_X(1) = 1$, equation 15.3 is completely deterministic in that X_{i+1} is completely specified by X_i ($X_{i+1} = X_i$).

For a first order Markov process, the lag k serial correlation, $\rho_X(k)$, is given by

$$\rho_X(k) = \rho_X^k(1) \tag{15.4}$$

Thus the correlogram exponentially decays from $\rho_X(0) = 1$ to $\rho_X(\infty) = 0$ according to equation 15.4. If an observed correlogram has this property, the Markov model may be an appropriate generating model.

Equation 15.3 can be applied to the logarithms of data through the transformation $Y_i = \ln(X_i)$. The generation model is given by

$$Y_{i+1} = \mu_Y + \rho_Y(1)(Y_i - \mu_Y) + t_{i+1}\,\sigma_Y\sqrt{1-\rho_Y^2(1)} \tag{15.5}$$

where μ_Y, σ_Y and $\rho_Y(1)$ refer to the mean, standard deviation and first order serial correlation of the logarithms of the original data. Generation by equation 15.5 preserves the mean, variance, coefficient of skew and first order serial correlation of the logarithms of the original data but not of the data itself.

Matalas (1967) suggests a procedure for using a first order Markov model on the logarithms that preserves the mean, variance, skewness and first order serial correlation of the original data. The procedure is based on the transformation $Y_i = \ln(X_i - \alpha)$ with the parameters of equation 15.5 related to the parameters of X through the following equations:

$$\mu_X = \alpha + \exp(\sigma_Y^2/2 + \mu_Y) \tag{15.6}$$

$$\sigma_X^2 = \exp[2(\sigma_Y^2 + \mu_Y)] - \exp(\sigma_Y^2 + 2\mu_Y) \tag{15.7}$$

$$\gamma_X = [\exp(3\sigma_Y^2) - 3\exp(\sigma_Y^2) + 2] \,/\, [\exp(\sigma_Y^2) - 1]^{3/2} \tag{15.8}$$

$$\rho_X(1) = [\exp(\sigma_Y^2\,\rho_Y(1)) - 1] \,/\, [\exp(\sigma_Y^2) - 1] \tag{15.9}$$

In these equations μ_X, σ_X^2, γ_X and $\rho_X(1)$ refer to the mean, variance, coefficient of skew and first order serial correlation of the original data and are estimated by $\bar{X}$, S_X^2, C_{sX}, and $r_X(1)$ respectively. The quantities μ_Y, σ_Y, $\rho_Y(1)$ and α are estimated from equations 15.6 thru 15.9 and then used in equation 15.5 to generate values for Y_{i+1}. X_{i+1} is then calculated from

$$X_{i+1} = \exp(Y_{i+1}) + \hat{\alpha} \tag{15.10}$$

The X's generated in this fashion have the same mean, variance, skewness and first order serial correlation as the sample used to estimate μ_X, σ_X^2, γ_X and $\rho_X(1)$. The procedure that is recommended for estimating μ_Y, σ_Y, ρ_Y and α is to solve equation 15.8 for s_Y. Equation 15.9 then yields $r_Y(1)$, equation 15.7 yields $\bar{Y}$, and equation 15.6 yields $\hat{\alpha}$.

Equation 15.1 can be used to generate X's that are distributed approximately gamma with mean $\bar{X}$, variance s_X^2 and skewness c_{sX} (Thomas and Fiering 1963). The procedure is to define γ_ϵ as the skewness of the random component, ϵ. γ_ϵ is estimated by

$$c_{s\epsilon} = [1 - r_X^3(1)]\; c_{sX} / [1 - r_X^2(1)]^{1.5} \tag{15.11}$$

Then a random element ϵ_{i+1} is defined by

$$\epsilon_{i+1} = \frac{2}{c_{s_\epsilon}} \left[1 + \frac{c_{s_\epsilon} t_{i+1}}{6} - \frac{c_{s_\epsilon}^2}{36} \right]^3 - \frac{2}{c_{s_\epsilon}} \qquad (15.12)$$

where t_{i+1} is a random value from a N(0,1). X_{i+1} is then generated by

$$X_{i+1} = \mu_X + r_X(1)(X_i - \bar{X}) + \epsilon_{i+1} s_X \sqrt{1-r_X^2(1)} \qquad (15.13)$$

with the resulting generated X's being approximately gamma distributed with mean $\bar{X}$, variance s_X^2, first order serial correlation $r_X(1)$ and skewness c_{s_X}.

The first order Markov model (equation 15.1) is also known as the first order autoregressive model since $\rho_X(1)$ is equal to the regression coefficient β that could be obtained if the regression model of chapter 9 was used taking Y as X_{i+1} and X as X_i. That this is the case is apparent from equation 9.15 with $s_Y = s_X$.

FIRST ORDER MARKOV PROCESS WITH PERIODICITY

The first order Markov model of the previous section assumes that the process is stationary in its first three moments. It is possible to generalize the model so that the periodicity in hydrologic data is accounted for to some extent. The main application of this generalization has been in generating monthly streamflow where pronounced seasonality in the monthly flows exist. Looking at figure 14.7, it is apparent that the monthly flows in Cave Creek are much greater (on the average) in some months than they are in others. This pronounced annual cycle is prevalent in many types of hydrologic and climatic data. The periodicity may affect not only the mean, but all of the moments of the data as well as the first order serial correlation.

To generalize the Markov model, we adopt the notation that the subscript i refers to the year under consideration and the subscript j refers to the season within the year. Thus j may run from 1 to 4 if 4 seasons are being considered, 1 to 12 if monthly data is being considered, 1 to 52 for weekly data, etc. In general we will take j to run from 1 to m is the number of seasons in the year.

With this notation $\mu_{x,j}$ refers to the mean of X in the j^{th} season. $\mu_{x,j}$ is estimated by $\bar{X}_j$ where

$$\bar{X}_j = \Sigma_{i=1}^n X_{i,j}/n \qquad (15.14)$$

with n equal to the number of years of data and $X_{i,j}$ the data value in the j^{th} season of the i^{th} year. Similarly $\sigma_{X,j}^2$ is estimated by $S_{X,j}^2$, $\gamma_{X,j}$ is estimated by $C_{s_X,j}$ and $\rho_{X,j}(1)$ is estimated by $r_{X,j}(1)$. Note that $\rho_{X,j}(1)$ is the first order serial correlation between values in successive seasons. If monthly streamflow is being considered, $\rho_{X,4}(1)$ would be the first order serial correlation between flows in months 4 and 5. $\rho_{X,j}(1)$ would be estimated by

$$r_{X,j}(1) = (\Sigma_{i=1}^n X_{i,j} X_{i,j+1} - n \bar{X}_j \bar{X}_{j+1}) / (n-1) s_{X,j} s_{X,j+1} \qquad (15.15)$$

where

$$s_{X,j}^2 = (\Sigma_{i=1}^n X_{i,j}^2 - n\bar{X}_j^2)/(n-1) \qquad (15.16)$$

In equation 15.15 there are some notational problems when j = m. In this case j+1 should be taken as 1 since the first season follows the m^{th} season (January follows December for example).

With this notation, the multiseason, first order Markov model for normally distributed flows becomes

$$X_{i,j+1} = \mu_{X,j+1} + (\rho_{X,j}(1)\, \sigma_{X,j+1}/\sigma_{X,j})(X_{i,j} - \mu_{X,j}) + t_{i,j+1}\, \sigma_{X,j+1} \sqrt{1-\rho_{X,j}^2(1)} \qquad (15.17)$$

In any application the population parameters are estimated by the corresponding sample statistics. The subscript notation of equation 15.17 again has problems in that $X_{i,j+1}$ is really equal to $X_{i+1,1}$ when j = m. For instance if a monthly model is considered then $X_{i,13}$ (or the 13^{th} monthly value in the i^{th} year) is actually $X_{i+1,1}$ (or the first monthly value in the next year). $t_{i,j+1}$ is again a random observation from a N(0,1). Values generated by equation 15.17 are thus the sum of the mean for the season plus a regression coefficient times the deviation from its mean of the value in the previous period plus a random component that is normally distributed with mean zero and variance $\sigma_{X,j+1}^2 [1-\rho_{X,j}^2(1)]$.

The first order Markov model can also be generalized to a seasonal model for gamma variates (Fiering and Jackson 1971) by generalizing equations 15.11, 15.12 and 15.13. Equation 15.11 becomes

$$\gamma_{\epsilon,j} = [\gamma_{X,j} - \rho_{X,j-1}^3(1)\, \gamma_{X,j-1}] / [1 - \rho_{X,j}^2(1)]^{1.5} \qquad (15.18)$$

Equation 15.12 becomes

$$\epsilon_{i,j} = \frac{2}{\gamma_{\epsilon,j}} \left[1 + \frac{\gamma_{\epsilon,j} t_{i,j}}{6} - \frac{\gamma_{\epsilon,j}^2}{36} \right]^3 - \frac{2}{\gamma_{\epsilon,j}} \qquad (15.19)$$

where $t_{i,j}$ is a random value from a N(0,1). Equation 15.13 becomes identical to equation 15.17 with population parameters replaced by their estimates except $\epsilon_{i,j+1}$ is used in place of $t_{i,j+1}$. The resulting $X_{i,j}$ will be distributed almost gamma. Because skewness varies from season to season, the representation is not statistically pure (Fiering and Jackson 1971). This is because the sum of gamma variates is not gamma unless the scale parameter, λ, is the same.

Equation 15.17 can be applied to the logarithms of the original data. In this case $X_{i,j}$ would refer to the logarithm of the value in the i^{th} year and j^{th} season. The parameters of the model would also be based on the logarithms. The model used in this way would preserve the mean, variance, skewness and first order serial correlation of the logarithms of the data but not of the data itself.

Equation 15.3 and 15.17 have been widely used in hydrology. Equation 15.17 is sometimes known as the Thomas-Fiering model because of the early work of these two men with the model (Thomas and Fiering 1962, 1963; Fiering 1967). The model in the form of equation 15.17 requires that many parameters be estimated. For each season, the mean, variance and first order serial correlation must be estimated. This results in estimating 3m parameters (a monthly model requires one to estimate 36 parameters). This large number of parameters requires considerable data. The technique based on data generation given in chapter 13 can be used for evaluating the effect of the length of record available for parameter estimation on the reliability of the Thomas-Fiering model or other stochastic models.

HIGHER ORDER AUTOREGRESSIVE MODELS

The model given by equation 15.1 can be generalized to include the effects of more than one preceding time period. Such a model has been called a multilag model, higher order Markov model or higher order autoregressive model. The model can be written as

$$X_{i+1} = \beta_o + \beta_1 X_i + \beta_2 X_{i-1} + \ldots + \beta_m X_{i-m+1} + \epsilon_{i+1} \qquad (15.20)$$

The X_i's might represent actual data values or their natural logarithms. In the case of a normal model, the random element becomes

$$\epsilon_{i+1} = \sigma_X t_{i+1} \sqrt{1-R^2} \qquad (15.21)$$

where σ_X^2 is the variance of X; R^2 is the multiple coefficient of determination between X_{i+1} and $X_i, X_{i-1}, \ldots, X_{i-m+1}$; t_{i+1} is a random observation from N(0,1) and the β's are multiple regression coefficients.

The multilag model permits one to incorporate linear influences on data in one period reflected by data in several preceding periods. The regression coefficients, β, can be estimated by the procedures of chapter 10. The question of how many lags to include can also be attacked by the methods of chapter 10 devoted to determining whether or not a particular "independent" variable is important.

One difference between this model and the multiple regression procedures is that the number of observations available for parameter estimation, n*, changes as the number of lags changes. If there are n total observations, then there are n-1 observations available for estimating $\rho_X(1)$ of the first order Markov model. If two lags are considered (m = 2), then there are only n-2 observations available for parameter estimation. In general for an m^{th} order Markov model, there are $n^* = n-m$ observations for parameter estimation. What this means is that the results of chapter 10 are not strictly applicable because the sample size and variables involved in the regressions change as the number of lags change. For instance R^2 may actually increase if the number of lags included decreases since the data set involved in the regression has changed. Generally it is recommended that if a k^{th} order lag is included, then all lags up to k also be included. For example if a third order lag is included, then the first and second order lags should be included as well.

MULTISITE MARKOV MODEL

The need for simultaneous data generation on two or more random variates or on the same random variate at two or more sites frequently arises. If a reservoir for water supply is fed by several streams, the flows from all of the streams must be considered in the design of the reservoir. The flood volume passing a city may be the sum of the volume released from an upstream reservoir and the inflows from tributaries between the reservoir and the city. The optimal operation of a system of reservoirs depends on flows in several streams. The simultaneous simulation of inflow into a water supply reservoir and outflow for irrigation purposes may be needed to design the reservoirs or to develop its operating policy. In all of these cases and more, the simultaneous behavior of the random variables (reservoir inflow, irrigation demand, etc.) is important.

If the time series for the important random variables are independent, then the generation techniques for single sites can be used to produce synthetic time series of

data on each random variable without considering any of the other random variables. For example if the inflow into a reservoir is the sum of the inflows from two streams and the flows on the two streams are stochastically independent, data can be generated using the methods discussed earlier in this chapter for each stream. In so doing the resulting generated data on the two streams will be independent. In another situation the existing dependence may not be of concern in which case independent generations can again be made.

In many situations the simultaneous behavior of two or more random variables is important and must be considered. Streamflow volumes on nearby streams are generally correlated. Correlation means that the flows in the same time period are correlated. This is known as simple correlation or as lag zero cross-correlation. The latter term is used since consideration can also be given to lag k cross-correlation. Lag k cross-correlation is the correlation between one random variable at one time point and a second random variable k time points later. A general notation for lag k cross-correlation between random variable X_j and X_h is $\rho_{j,h}(k)$ which is estimated by

$$r_{j,h}(k) = [\Sigma_{i=1}^{n-k}(X_{j,i} - \bar{X}_j)(X_{h,i+k} - \bar{X}_h)] \,/\, (n-k)\, s_{X,j}\, s_{X,h} \qquad (15.22)$$

where n is the total number of pairs of observations on X_j and X_h, $X_{j,i}$ is the i^{th} observation on X_j, $\bar{X}_j$ is the mean of the observations on X_j and $s_{X,j}^2$ is the variance of X_j. Note that for k = 0, equation 15.22 is identical to equation 11.12. The lag k cross-correlation is thus the correlation between values of X_j and values of X_h that are k time units apart.

If interest exists only in the sum of several random variables such as the total inflow into a reservoir, a sequence $Y_i = X_{1,i} + X_{2,i} + \ldots + X_{p,i}$, where $X_{k,i}$ represents the flow on stream k at time i, can be formed and then the single site models applied to the Y_i's. The multiseason, single site model could be used in this situation as well. If X_h is independent of X_j, p is large and all of the X_j contribute about equally to Y_i, the Central Limit Theorem would suggest that Y_i is approximately normally distributed. Of course if the X_j and X_h are independent, the single site model could be applied to each of the X_j's individually and then the results summed to get Y_i. If X_j and X_h are not independent, the Central Limit Theorem does not apply, however, the Y_i's may still be normally distributed.

A two site generation model that preserves the means, variances, skewness, lag one serial correlation, and lag zero cross-correlation has been presented by Fiering (1964) and discussed by DeCoursey (1971). The technique requires that one of the two sites be selected as a key site. The selection could be based on the quality of the record, length of the record, or other considerations. We will assume that site j is the key site and site h is subordinate to site j. A sequence of observations can be generated for site j using single site generation techniques. A cross-correlation model is used to generate values of site h based on generated values at site j.

$$X_{h,i} = \bar{X}_h + r_{j,h}(0)\, s_h (X_{j,i} - \bar{X}_j)/s_j + u_i\, s_h \sqrt{1 - r_{j,h}^2(0)} \qquad (15.23)$$

where u_i is a standardized random variate adjusted to incorporate the serial correlation at site h.

$$u_i = \zeta(X_{h,i-1} - \bar{X}_h)/s_h + t_i (1 - \zeta^2)^{1/2} \qquad (15.24)$$

in which t_i is a standardized random variate adjusted for skew if necessary and

$$\zeta = (r_h(1) - r_j(1)\, r^2_{j,h}(0))/\sqrt{1 - r^2_{j,h}(0)} \tag{15.25}$$

If X_h and X_j show seasonal trends, these may be accounted for by proper selection of the relation for generation at the key station and proper selection of the appropriate seasonal quantities in equations 15.23, 15.24 and 15.25.

More generally multisite generation requires simultaneous generation of data at several sites while preserving the correlation between the data at the various sites. To develop the multisite model, let $x_{j,i} = (X_{j,i} - \mu_j)/\sigma_j$ so that $x_{j,i}$ is the standardized value of the data being considered at site j during time period i. The first order Markov model (equation 15.3) for sites h and j becomes

$$x_{j,i+1} = \rho_j(1)x_{j,i} + \sqrt{1 - \rho^2_j(1)}\ \epsilon_{j,i+1} \tag{15.26}$$

and

$$x_{h,i+1} = \rho_h(1)x_{h,i} + \sqrt{1 - \rho^2_h(1)}\ \epsilon_{h,i+1} \tag{15.27}$$

The equations can be written in matrix notation as

$$\underline{X}_{i+1} = \underline{E}\,\underline{X}_i + \underline{G}\,\underline{\epsilon} \tag{15.28}$$

where $\underline{X}_i$ is a pxl vector of standardized values of the variable being generated at time i, $\underline{E}$ is a pxp diagonal matrix whose j^{th} diagonal element is $\rho_j(1)$, $\underline{G}$ is a pxp diagonal matrix whose j^{th} diagonal element is $\sqrt{1-\rho^2_j(1)}$ and $\underline{\epsilon}$ is a pxl vector of random elements defined in such a way as to preserve the first order serial correlation of the x_j's and the lag zero cross-correlation between x_j and x_h. Limiting this treatment to the normal multisite model, $\underline{\epsilon}$ will be made up of elements that are $\epsilon_{j,i+1}$ where each $\epsilon_{j,i+1}$ is independent of $x_{j,i}$, ϵ_j is N(0,1) and the correlation between ϵ_j and ϵ_h is $\rho^*_{j,h}(0)$ where $\rho^*_{j,h}(0)$ is determined in such a way as to produce the desired $\rho_{j,h}(0)$ which is the lag zero cross-correlation between x_j and x_h.

Since the x's are standardized, $\rho_{j,h}(0) = \mathrm{cov}(x_j, x_h)$. From equations 15.26 and 15.27, $\rho_{j,h}(0)$ is evaluated as

$$\begin{aligned}
&\rho_{j,h}(0) = E[x_{j,i+1}x_{h,i+1} - E(x_{j,i+1})\,E(x_{h,i+1})] = \\
&E(x_{j,i+1}\,x_{h,i+1}) \\
&= E[(\rho_j(1)x_{j,i} + \sqrt{1-\rho^2_j(1)}\ \epsilon_{j,i+1})(\rho_h(1)x_{h,i} + \\
&\sqrt{1-\rho^2_h(1)}\ \epsilon_{h,i+1})] \\
&= \rho_j(1)\rho_h(1)\,E(x_{j,i}\,x_{h,i}) + \sqrt{1-\rho^2_j(1)}\ \sqrt{1-\rho^2_h(1)}\ E(\epsilon_{j,i+1}\epsilon_{h,i+1}) \\
&= \rho_j(1)\rho_h(1)\rho_{j,h}(0) + \sqrt{1-\rho^2_j(1)}\ \sqrt{1-\rho^2_h(1)}\ \rho^*_{j,h}(0)
\end{aligned} \tag{15.29}$$

Thus $\rho^*_{j,h}(0)$ can be evaluated as

$$\rho^*_{j,h}(0) = [1 - \rho_j(1)\rho_h(1)]\ \rho_{j,h}(0) / \sqrt{(1-\rho^2_j(1))\,(1-\rho^2_h(1))} \tag{15.30}$$

Therefore, a normal, multiple site generation model that preserves the means, variances, first order serial correlation and lag zero cross-correlation is given by equation 15.28 where the elements of $\underline{\epsilon}$ are N(0,1) and have lag zero cross-correlation given by $\rho^*_{j,h}(0)$.

The procedures outlined in chapter 13 can be used to generate $\underline{\epsilon}$. In that chapter it was shown that $\underline{\epsilon}$ is given by

$$\underline{\epsilon} = \underline{A}\,\underline{D}^{1/2}_{\lambda}\,\underline{e} \tag{15.31}$$

where $\underline{e}$ represents a pxl vector of independent observations from N(0,1), $\underline{A}$ is a pxp vector of characteristic roots and $\underline{D}^{1/2}_{\lambda}$ is a pxp diagonal matrix whose j^{th} diagonal element is the square root of the j^{th} largest characteristic root of the pxp correlation matrix whose elements are $\rho^*_{j,h}(0)$. Thus equation 15.28 can be written

$$\underline{X}_{i+1} = \underline{E}\,\underline{X}_i + \underline{G}\,\underline{A}\,\underline{D}^{1/2}_{\lambda}\,\underline{e} \tag{15.32}$$

Matalas (1967) discusses a multisite normal generation model that preserves the means, variances, lag one serial correlation, lag zero cross-correlation and lag one cross-correlation. Using the notation of Matalas the model is written

$$\underline{X}_{i+1} = \underline{A}\,\underline{X}_i + \underline{B}\,\underline{\epsilon}_{i+1} \tag{15.33}$$

where $\underline{X}_i$ is a mxl vector whose p^{th} element is $X^p_i - \mu^p_X$ with X^p_i the i^{th} value of X at station p and μ^p_X the mean value at station p. The random element $\underline{\epsilon}_{i+1}$ is from a N(0,1) with $\underline{\epsilon}_{i+1}$ independent of $\underline{X}_i$. The mxm matrices $\underline{A}$ and $\underline{B}$ are defined in terms of two new mxm matrices $\underline{M}_1$ and $\underline{M}_o$.

$\underline{M}_o$ has elements equal to the expected value of $\underline{X}_i\,\underline{X}_i^t$ or $\underline{M}_o = [\rho_{h,j}(0)\sigma_h\sigma_j]$ and $\underline{M}_1$ has elements equal to the expected value of $\underline{X}_{i+1}\,\underline{X}_i^t$ or $\underline{M}_1 = [\rho_{h,j}(1)\sigma_h\sigma_j]$. The matrix $\underline{A}$ is then defined by

$$\underline{A} = \underline{M}_1\,\underline{M}_o^{-1} \tag{15.34}$$

and the matrix $\underline{B}$ by

$$\underline{B}\,\underline{B}^t = \underline{M}_o - \underline{M}_1\,\underline{M}_o^{-1}\,\underline{M}_1^t \tag{15.35}$$

Matalas (1967) suggests a solution for equation 15.35 based on principal components. Young (1968) has suggested an alternative solution by assuming that $\underline{B}$ is a lower triangular matrix (all elements above and to the right of the main diagonal are zero). He then defines the matrix $\underline{C}$ as

$$\underline{C} = \underline{B}\,\underline{B}^t = \underline{M}_o - \underline{M}_1\,\underline{M}_o^{-1}\,\underline{M}_1 \tag{15.36}$$

Since $\underline{B}$ is lower triangular, $\underline{B}\,\underline{B}^t$ is symmetric so $\underline{C}$ is also symmetric.

The case where m = 2 results in

$$\underline{C} = \begin{bmatrix} c_{11} & c_{21} \\ c_{21} & c_{22} \end{bmatrix} = \begin{bmatrix} b_{11}^2 & b_{11}b_{21} \\ b_{21}b_{11} & b_{21}^2+b_{22}^2 \end{bmatrix} = \underline{B}\,\underline{B}'$$

or

$$\begin{aligned} b_{11} &= \sqrt{c_{11}} \\ b_{21} &= c_{21}/b_{11} \\ b_{22} &= \sqrt{c_{22} - b_{21}^2} \end{aligned} \tag{15.37}$$

Similarly for the case m = 3

$$\underline{C} = \begin{bmatrix} b_{11}^2 & b_{11}b_{21} & b_{11}b_{31} \\ b_{21}b_{11} & b_{21}^2+b_{22}^2 & b_{21}b_{31}+b_{22}b_{32} \\ b_{31}b_{11} & b_{31}b_{21}+b_{32}b_{22} & b_{31}^2+b_{32}^2+b_{33}^2 \end{bmatrix}$$

or

$$\begin{aligned} b_{11} &= \sqrt{c_{11}} \\ b_{21} &= c_{21}/b_{11} \\ b_{31} &= c_{31}/b_{11} \\ b_{22} &= \sqrt{c_{22} - b_{21}^2} \\ b_{32} &= (c_{31} - b_{31}b_{21})/b_{22} \\ b_{33} &= \sqrt{c_{33} - b_{31}^2 - b_{32}^2} \end{aligned} \tag{15.38}$$

These results can easily be extended for any m. In each case the number of equations involved is m(m+1)/2.

The Matalas model can be simplified if lag one cross-correlation is not important. In this event $\underline{A}$ is a diagonal matrix with the diagonal elements equal to the lag one serial correlation. Then $\underline{M}_1$ is defined as

$$\underline{M}_1 = \underline{A}\,\underline{M}_0 \tag{15.39}$$

and $\underline{M}_0$ and $\underline{B}$ are defined as above with the solution for $\underline{B}$ as outlined above. The lag one cross-correlation is not preserved by this simplification but is set equal to $\rho_j(1)\rho_{j,h}(0)$ by it.

MARKOV CHAIN MODELS

A Markov chain is a stochastic process having the property that the value of the process at time t, X_t, depends only on its value at time t - 1, X_{t-1}, and not on the sequence of values $X_{t-2}, X_{t-3}, \ldots, X_o$ that the process passed through in arriving at X_{t-1}. This can be written

$$\begin{aligned} &\text{prob}(X_t = a_j \,|\, X_{t-1} = a_i, X_{t-2} = a_k, X_{t-3} = a_l, \ldots, X_o = a_q) = \\ &\text{prob}(X_t = a_j \,|\, X_{t-1} = a_i) \end{aligned} \tag{15.40}$$

The conditional probability, $\text{prob}(X_t = a_j \,|\, X_{t-1} = a_i)$, gives the probability that the process at time t will be in "state" j given that at time t - 1 the process was in "state" i. Equation 15.40 says that this conditional probability is independent of the "states" occupied at times prior to t - 1. A state is simply a subdivision of the process X_t into some interval. Thus if X_t represents the depth of rainfall on day t, one state might be defined as no rainfall, another as between 0.00 and 0.05 inches of rainfall, etc.

The $\text{prob}(X_t = a_j \,|\, X_{t-1} = a_i)$ is commonly called the one step transition probability. That is, it is the probability that the process makes the transition from state a_i to state a_j in one time period or one step. $\text{Prob}(X_t = a_j \,|\, X_{t-1} = a_i)$ is usually written as $p_{ij}(t)$ indicating the probability of a step from a_i to a_j at time (t). If $p_{ij}(t)$ is independent of t ($p_{ij}(t) = p_{ij}(t + \tau)$ for all t and τ), then the Markov chain is said to be homogeneous. In this event

$$\text{prob}(X_t = a_j \,|\, X_{t-1} = a_i) = p_{ij} \tag{15.41}$$

Higher order Markov chains can be defined to represent stochastic processes such that the value of the process at time t is dependent on its value in several immediately preceding time periods. Thus an n^{th} order Markov chain is one in which

$$\begin{aligned} &\text{prob}(X_t = a_j \,|\, X_{t-1} = a_i, X_{t-2} = a_k, X_{t-3} = a_e, \ldots, X_o = a_q) = \\ &\text{prob}(X_t = a_j \,|\, X_{t-1} = a_i, X_{t-2} = a_k, \ldots, X_{t-n} = a_p) \end{aligned} \tag{15.42}$$

The treatment of Markov chains in this text will be limited to first order homogeneous Markov chains.

If a process is divided into m states, then m^2 transition probabilities must be defined. However, at each step the process must either remain in state i or proceed to one of the other m - 1 states. Thus

$$\sum_{j=1}^{m} p_{ij} = 1 \tag{15.43}$$

With this restriction, an m state Markov chain requires that m(m - 1) transition probabilities (parameters) be estimated. The remaining m p_{ij}'s can be determined from equation 15.43. The m^2 transition probabilities can be represented by the mxm matrix $\underline{P}$ given by

$$\underline{P} = [p_{ij}] = \begin{bmatrix} p_{11} & p_{12} & \cdots & p_{1m} \\ p_{21} & p_{22} & & \vdots \\ \vdots & & & \vdots \\ p_{m1} & & & p_{mm} \end{bmatrix} \tag{15.44}$$

Equation 15.43 states that the elements in any row of $\underline{P}$ must sum to unity. A matrix having this property is said to be a stochastic matrix. Some authors define $\underline{P}$ as $\underline{P} = [p_{ij}]' = p_{ji}$. Under this definition the columns of $\underline{P}$ sum to unity. The definition given by equation 15.41 will be used in this treatment.

The transitional probability matrix $\underline{P}$ can be estimated from observed data by tabulating the number of times the observed data went from state i to state j, n_{ij}. Then an estimate for p_{ij} would be

$$\hat{p}_{ij} = n_{ij} / \Sigma_{j=1}^{m} n_{ij} \tag{15.45}$$

Considerable data may be required to get accurate estimates of p_{ij} if p_{ij} is low. This is because in an observed set of data, n_{ij} may be uncharacteristically high or low if p_{ij} is close to zero and the sample is small.

The more states that a process is divided into, the less accurate will be the estimates for p_{ij}. For example if a daily rainfall model is being considered, one might like to have 10 states to adequately represent the possible amounts of rainfall. However, 10 states require the estimation of 90 transition probabilities. This in turn requires a large amount of data.

Once $\underline{P}$ is known, all that is required to determine the probabilistic behavior of the Markov chain is the initial state of the chain. In the following the notation $p_j^{(n)}$ means the probability that the chain is in state j at step or time n. The 1xm vector $\underline{p}^{(n)}$ has elements $p_j^{(n)}$. Thus

$$\underline{p}^{(n)} = (p_1^{(n)}, p_2^{(n)}, \ldots, p_m^{(n)}) \tag{15.46}$$

Under this definition, $\underline{p}^{(0)}$, is the initial probability vector. $\underline{p}^{(1)}$ is then given by

$$\underline{p}^{(1)} = \underline{p}^{(0)} \underline{P} \tag{15.47}$$

and $\underline{p}^{(2)}$ is given by

$$\underline{p}^{(2)} = \underline{p}^{(1)} \underline{P} = \underline{p}^{(0)} \underline{P}^2 \tag{15.48}$$

where $\underline{P}^n$ is the n^{th} power of $\underline{P}$. In general

$$\underline{p}^{(n)} = \underline{p}^{(0)} \underline{P}^n \tag{15.49}$$

Furthermore it can be shown that

$$\underline{p}^{(n+m)} = \underline{p}^{(m)} \underline{P}^n \tag{15.50}$$

For a proof of these relationships, reference should be made to any number of books on probability or stochastic processes (see for instance Bailey 1964; Feller 1957; Brieman 1969).

As the Markov chain advances in time, $p_j^{(n)}$ becomes less and less dependent on $\underline{p}^{(0)}$. That is to say the probability of being in state j after a large number of steps becomes independent of the initial state of the chain. A point is reached where $\underline{p}^{(n)} = \underline{p}^{(n+m)}$ for a sufficiently large n. From equation 15.49 we then get for a sufficiently large n that $\underline{P}^n = \underline{P}^{n+m}$. When this occurs the chain is said to have reached a steady state.

Under steady state conditions $\underline{p}^{(n)} = \underline{p}^{(n+m)}$ and can thus be denoted simply as $\underline{p}$. The 1xm vector $\underline{p}$ can be thought of as giving the probabilities of being in the various states after a large number of steps.

Under steady state conditions

$$\underline{p} = \underline{p} \underline{P} \tag{15.51}$$

The solution of equation 15.51 thus provides $\underline{p}$.

$\underline{P}^n$ is called the n step transitional probability matrix. That is $\underline{P}^n = [p_{ij}^{(n)}]$ has elements which give the probability of going from state i to state j in n steps. Since for large n, $p_j^{(n)}$ is independent of the initial state, we must have $p_{ij}^{(n)} = p_j^{(n)}$. Thus $\underline{P}^n$ is made up of m 1xm vectors all equal to $\underline{p}$. That is for large n

$$\underline{P}^n = \begin{bmatrix} \underline{p} \\ \underline{p} \\ . \\ . \\ . \\ \underline{p} \end{bmatrix} \tag{15.52}$$

One can therefore calculate the steady state probabilities simply by computing $\underline{P}^n$ for a large enough n. In practice one would compute $\underline{P}^n$ and $\underline{P}^{2n}$. If the two differed by an acceptably small amount, $\underline{p}$ would be taken as one of the rows of $\underline{P}^{2n}$. Bailey (1964) gives a procedure for calculating $\underline{P}^n$ based on characteristic roots. On a digital computer, $\underline{P}^n$ can be easily evaluated by multiplication. This method for finding $\underline{p}$ may require n to be very large. The steady state probabilities $\underline{p}$ can also be determined directly from equation 15.51. Example 15.2 illustrates this approach.

Example 15.1. Consider a 2-state, first order Markov chain for a sequence of wet and dry days. Let state 1 be a dry day and state 2 be a wet day. Assume the transitional probability matrix to be

$$\underline{P} = \begin{bmatrix} .9 & .1 \\ .5 & .5 \end{bmatrix}$$

Thus the probability of a dry day following a wet day is given by p_{21} as 0.5. Evaluate (a) prob(day 1 wet | day 0 dry) (b) prob(day 2 wet | day 0 dry) (c) prob(day 100 wet | day 0 dry).

Solution:

(a) prob(day 1 wet | day 0 dry) = $p_{12} = p_2^{(1)} = 0.1$

(b) prob(day 2 wet | day 0 dry) = $p_2^{(2)}$

$$\underline{p}^{(2)} = \underline{p}^{(1)} \underline{P} = (.9 \; .1) \begin{bmatrix} .9 & .1 \\ .5 & .5 \end{bmatrix} = (.86 \; .14)$$

thus $p_2^{(2)} = 0.14$

(c) prob(day 100 wet | day 1 dry) = $p_2^{(100)}$

$$\underline{p}^{(100)} = \underline{p}^{(1)}\,\underline{P}^{99}$$

However, the fact that day 1 was dry would not significantly affect the probability of rain on day 100. Therefore it can be assumed that n is large and base the solution on the steady state probabilities contained in $\underline{P}^n$ for large n.

$$\underline{P}^2 = \underline{P}\,\underline{P} = \begin{bmatrix} .86 & .14 \\ .70 & .30 \end{bmatrix}$$

$$\underline{P}^4 = \underline{P}^2\,\underline{P}^2 = \begin{bmatrix} .8376 & .1624 \\ .8120 & .1880 \end{bmatrix}$$

$$\underline{P}^8 = \underline{P}^4\,\underline{P}^4 = \begin{bmatrix} .8334 & .1666 \\ .8320 & .1672 \end{bmatrix}$$

$$\underline{P}^{16} = \underline{P}^8\,\underline{P}^8 = \begin{bmatrix} .8333 & .1667 \\ .8333 & .1667 \end{bmatrix}$$

$\underline{P}^{16}$ is assumed to be the steady state n-step transitional probability matrix since the $p_{ij}^{(n)}$ are not changing much and since the two rows are identical. Thus $p_2^{(100)} = p_2^{(n)} = 0.1667$. The probability of rain on any day in the distant future is 0.1667. For this to be true an analysis of rainfall records should show that 16.67 precent of the days are wet. This serves as a check on $\underline{P}$.

Comment: Another check on the steady state probabilities is to see if equation 15.48 is valid.

$$\underline{p}\,\underline{P} = (.8333\ \ .1667)\begin{bmatrix} .9 & .1 \\ .5 & .5 \end{bmatrix} = (.8333\ \ .1667)$$

This demonstrates that $\underline{p} = (.8333\ .1667)$ is the steady state probability matrix. See example 15.2 for further comment on this.

Example 15.2. Consider a Markov chain model for the amount of water in storage in a reservoir. Let state 1 represent the nearly full condition, state 2 an intermediate condition and state 3 the nearly empty condition. Assume that the transition probability matrix is given by

$$\underline{P} = \begin{bmatrix} .4 & .6 & 0 \\ .2 & .6 & .2 \\ 0 & .7 & .3 \end{bmatrix}$$

Note that it is not possible to pass directly from state 1 to state 3 or from state 3 to state 1 without going through state 2. Over the long run, what fraction of the time is the reservoir level in each of the states?

Solution: The fraction of time spent in each state is given by $\underline{p}$. Equation 15.51 can be used to determine $\underline{p}$. Examination of equation 15.51 shows that if $\underline{p}$ is a solution so is $\lambda\underline{p}$. Therefore a solution to 15.51 is unique only up to a scalar multiplication. However, since $\underline{p}$ is a probability vector, the sum of its elements must be one. Therefore our solution technique is to find an arbitrary solution to 15.51 and then scale it so that $\Sigma p_i = 1$.

$$\underline{p}\,\underline{P} = \underline{p}$$

$$(p_1\ p_2\ p_3)\begin{bmatrix} .4 & .6 & 0 \\ .2 & .6 & .2 \\ 0 & .7 & .3 \end{bmatrix} = (p_1\ p_2\ p_3)$$

$$.4\,p_1 + .2p_2 + 0 = p_1$$

$$.6\,p_1 + .6p_2 + .7p_3 = p_2$$

$$0 \quad + .2p_2 + .3p_3 = p_3$$

If we let $p_1 = 1$, the first of these equations gives $p_2 = 3$. With $p_2 = 3$, the last of the equations give $p_3 = 6/7$. Therefore one solution of $\underline{p}\,\underline{P} = \underline{p}$ is $\underline{p} = (1\ 3\ 6/7)$. Since Σp_i must equal one, $\underline{p}$ can be scaled so that $\underline{p} = (.2059\ .6176\ .1765)$. This solution can be substituted into equation 15.51 to verify that it is in fact a solution. Another check would be to compute $\underline{P}^n$ for large n and show that $\underline{P}^n = (\underline{p}\ \underline{p}\ \underline{p})^t$. Thus over the long run the reservoir is nearly full 20.59% of the time, nearly empty 17.65% of the time and in the intermediate state 61.76% of the time.

Comment: This problem illustrates the direct solution via equation 15.51 for $\underline{p}$. The $\underline{P}^n$ approach for determining $\underline{P}$ can however, be used. For this example $\underline{P}^8$ and $\underline{P}^{16}$ can be found to be (4 significant figures):

$$\underline{P}^8 = \begin{bmatrix} .2060 & .6176 & .1763 \\ .2059 & .6176 & .1765 \\ .2057 & .6177 & .1766 \end{bmatrix} \text{ and } \underline{P}^{16} = \begin{bmatrix} .2059 & .6176 & .1765 \\ .2059 & .6176 & .1765 \\ .2059 & .6176 & .1765 \end{bmatrix}$$

Thus $\underline{P}^8$ and $\underline{P}^{16}$ are nearly identical and $\underline{P}^{16}$ has the same columns which are in fact equal to $\underline{p}$. This can be seen from

$$\underline{p}\,\underline{P} = (.2059\ \ .6176\ \ .1765)\begin{bmatrix} .4 & .6 & 0 \\ .2 & .6 & .2 \\ 0 & .7 & .3 \end{bmatrix} = (.2059\ \ .6177\ \ .1765) = \underline{p}$$

Thus $\underline{p}$ satisfies equation 15.51.

Data generation from a Markov chain requires only a knowledge of the initial state and the transitional probability matrix $\underline{P}$. To determine the state at time 2, a random number is selected between zero and one. If this random number is between $\Sigma_{j=1}^{n-1} p_{ij}$ and $\Sigma_{j=1}^{n} p_{ij}$ for n = 1, 2, ... m, the next state is taken as state n. Example 15.3 illustrates this procedure.

Example 15.3. Assume that the reservoir of example 15.2 is nearly full at t = 0. Generate a sequence of 10 possible reservoir levels corresponding to t = 1, 2, ..., 10.

Solution: The matrix $\underline{P}$ can be written in the form of a cumulative transition probability matrix $\underline{P}^*$ where

$$\underline{P}^* = [P^*_{ik}] \text{ and } P^*_{ik} = \Sigma_{j=1}^{k} p_{ij}$$

For this example

$$\underline{P}^* = \begin{bmatrix} .4 & 1.0 & 1.0 \\ .2 & .8 & 1.0 \\ 0 & .7 & 1.0 \end{bmatrix}$$

Time t	State at t	Random No.	State at t+1	Reservoir level at t
0	1	.48	2	nearly full
1	2	.52	2	intermediate
2	2	.74	2	intermediate
3	2	.15	1	intermediate
4	1	.27	1	nearly full
5	1	.03	1	nearly full
6	1	.49	2	nearly full
7	2	.02	1	intermediate
8	1	.97	2	nearly full
9	2	.96	3	intermediate
10	3			nearly empty

Markov chains have been used in hydrology for modeling rainfall (Gabriel and Neumann 1962; Pattison 1964; Bagley 1964; Grace and Eagleson 1966; Hudlow 1967). Lloyd (1967) presents a discussion of the application of Markov chains to reservoir theory.

Some of the difficulties in using Markov chains in hydrology are:

(1) Determining the number of states to use.
(2) Determining the intervals of the variable under study to associate with each state.
(3) Assigning a number to the magnitude of an event once the state is determined (i.e., how much rainfall should be assigned given that chain moved to state 3 and that state 3 encompasses all rainfalls between 1 and 2 inches).
(4) Estimating the large number of parameters involved in even a moderate size Markov chain model. A chain with 5 states has 20 parameters to estimate. If seasonality is encountered and 4 seasons are needed, 80 parameters are required.
(5) Handling situations where some transitions are dependent on several previous time periods while others are dependent on only one prior time period. Hudlow (1967) found the dry-dry transition for hourly rainfall showed a sixth order Markov dependence while a first order dependence was adequate for the other transitions.

Woolhiser, Rovey and Todorovic (1973) discuss an n-day rainfall model in which the transition from wet to dry days is based on a 2 state Markov chain and the amount of rain on rainy days is exponentially distributed. Haan et al. (1976) describe a 7 state Markov chain model of daily rainfall in which the amount of rain in each state is assumed uniformly distributed except for the last state in which a shifted exponential distribution is used.

Carey and Haan (1976) present a modified Markov chain daily rainfall simulation model in which the transitional probabilities are replaced by a continuous probability distribution. That is, given that the system is in state i on day n in season k, then the probability distribution of the amount of rain on day n+1 is given by

$$\text{prob}(X_{n+1} \le x \,|\, X_n \text{ in i, season k}) = p^k_{i1} + (1 - p^k_{i1})\, P_X(x \,|\, i,k) \qquad (15.53)$$

where X_n is the amount of rain on day n, p^k_{i1} is the probability of no rain on day n+1 given X_n was in state i and season k and $P_X(x\,|\,i,k)$ is the cumulative probability distribution of rainfall on day n+1 given that rain occurs on day n+1 and that X_n was in state i and season k. Hence, to each state in each season there is a corresponding distribution function of the form of equation 15.53. The parameters p^k_{i1} are estimated as f^k_{i1}/f^k_i where f^k_{i1} is the historical frequency of transition from state i to state 1 (no rain) in season k and f^k_i is the total number of occurrences in state i and season k. The parameters of each distribution $P_X(x\,|\,i,k)$ are determined from historical data using the set of observations $[X_{n+1}\,|\,X_{n+1} > 0, X_n \text{ in i, season k}]$.

Synthetic traces of daily rainfall are generated from equation 15.53 in the following manner:

(1) Determine the state i and season k of X_n.
(2) Generate a uniform random number, R_u, from the interval (0,1).
(3) If $R_u \le p^k_{i1}$ then $X_{n+1} = 0$.
(4) If $R_u > p^k_{i1}$, generate a random observation, x, from $P_X(x\,|\,i,k)$ and set $X_{n+1} = x$.
(5) Repeat steps 1-4 advancing in time and changing seasons as required.

For Kentucky rainfall, Carey and Haan (1976) used 3 states and 12 seasons. Gamma distributions were used for $P_X(x\,|\,i,k)$. Furthermore, for a given season, the same gamma distribution could be used for all 3 states. Thus for each season 5 parameters – 2 parameters of the gamma distribution and 3 values for p^k_{i1} – had to be estimated or a total of 60 parameters. This compares with 505 parameters when the Markov chain approach of Haan et al. (1976) was used (A 7x7 transition probability matrix for each of 12 seasons plus an exponential parameter. 12(7x6) + 1 = 505). When simulated rainfall for these two models was compared to historical rainfall at 7 Kentucky locations, the Carey-Haan model proved superior.

OTHER STOCHASTIC MODELS

The number and kind of stochastic models that might be applicable in hydrology are nearly unlimited. The purpose of this section is to provide some comments and references relative to some other stochastic models that have been suggested for use in hydrology. This discussion is by no means exhaustive in that many potentially useful

models are not covered.

Yevjevich (1972) has suggested that hydrologic time series can be modeled by a deterministic component and a stationary stochastic process. The deterministic component is composed of trends, jumps and periodicities. The stochastic component consists of an autocorrelative type dependence and an independent stochastic component. Under certain conditions this model reduces to a Markov model. This type of analysis has been applied to many kinds of hydrologic time series including water-use time series and ground water table level time series (Salas-LaCruz and Yevjevich 1972; Law 1974).

As has been mentioned many times, the selection of a model depends on the properties of the process being modeled that must be preserved. The first-order Markov or autoregressive model has been criticized for its inability to maintain a property known as the "Hurst phenomenon" (Rodriguez-Iturbe et al. 1971, 1972; Mandelbrot and Wallis 1968, 1969; Wallis and Matalas 1970). The following discussion of the Hurst phenomenon is adopted from Rodriguez-Iturbe et al. (1972).

Let $X_1, X_2, X_3, \ldots$ be a stochastic series. Define S_n to be

$$S_n = \Sigma_{i=1}^{n} X_i \qquad n = 1, 2, 3, \ldots$$

Let S_k^* be the cumulative deviation from the mean, S_n/n, after k time intervals.

$$S_k^* = S_k - kS_n/n = \Sigma_{i=1}^{k} X_i - (k/n)\Sigma_{i=1}^{n} X_i \qquad 1 \le k \le n \qquad (15.54)$$

Now let M_n^* be the maximum value of S_i^* for i = 1 to n.

$$M_n^* = \max[0, S_i^*] \text{ for } i = 1, 2, \ldots, n$$

Similarly let m_n^* be the minimum value of S_i^* for i = 1 to n.

$$m_n^* = \min[0, S_i^*] \text{ for } i = 1, 2, \ldots, n$$

Finally define the difference in M_n^* and m_n^* as R_n^*

$$R_n^* = M_n^* - m_n^*$$

R_n^* is called the adjusted range. R_n^* is the capacity that a reservoir with annual inflows X_i would have to have in order to provide a constant annual outflow equal to $\bar{X}$. The reservoir maintains the smallest acceptable reserve when m_n^* occurs and is of sufficient capacity to store the water when M_n^* occurs. This is of course an idealization; however, the statistical behavior of R_n^* is of interest.

Hurst (1951, 1956) recognized the dependence between the range and storage requirements of reservoirs. From an extensive analysis of 120 long-term geophysical records, he found that

$$R_n = R_n^*/s_n = K_n^H \qquad (15.55)$$

where H was typically around 0.72, K is constant, n the record length, s_n is the estimated standard deviation of X_i based on n years of record, and R_n is the rescaled range.

The coefficient H in equation 15.55 is known as the Hurst coefficient. For a long normal independent series, H is 0.5. Mandelbrot and Van Ness (1968) have shown that

the theoretical asymptotic proportionality of R_n to $n^{0.5}$ holds for any finite memory process but that n^H ($H \neq 0.5$) may be exhibited as an initial transient state. This transient behavior poses the question of whether Hurst's observations were of a transient or long-term nature. Computer experiments (Wallis and Matalas 1970) tend to discredit the transient hypothesis; however, the question remains unanswered.

At this time the n^H behavior of natural geophysical records is an unresolved question. The fact that the first order Markov model cannot produce long term n^H behavior for H>0.5 and that multilag models require many lags to maintain H>0.5 for a transient period has led to the development of other models for generating time series of hydrologic data. For example Mejia et al. (1972) have advanced a simulation model based on a broken line process. A second simulation method that is presently being advanced is based on Fractional Gaussian Noise (see references by Mandelbrot and by Wallis). Because of the tentative nature of these approaches, they will not be discussed here. It is expected that in the next few years, these models and others will be thoroughly tested in hydrologic applications.

Exercises

15.1 Develop a stochastic model for generating a sequence of numbers that could represent the years between eruptions of the Volcano Aso (see exercise 14.3). Use the model to generate a series of 100 possible times (years) between eruptions. Compare the correlogram and spectral density functions for the generated and observed sequences.

15.2 Assume that the time (days) between rains follows a Poisson distribution with a mean of 2 days. Further assume that the amount of rain (inches) on rainy days follows a gamma distribution with a mean of 1 inch and a variance of 0.50 inches. Simulate 1 year of rainfall using this model.

15.3 Use the first order Markov model to generate 100 years of annual runoff (inches) for Cave Creek near Fort Spring, Kentucky. (Basic data in Appendix C.)

15.4 Generate a random sample of size 100 from a gamma distribution with $\eta = 3.5$ and $\lambda = 2.5$. Plot the observed and expected relative frequencies.

15.5 The following data are presented by Burges and Johnson (1973) for the Sauk River in Washington. Based on this data and the first order, seasonal, lognormal Markov model, generate 50 years of streamflow data. Compute and plot the correlogram and spectral density function for the generated data.

Month	$\bar{X}_j$	$s_{x,j}$	$\gamma_{x,j}$	Month	$\bar{X}_j$	$s_{x,j}$	$\gamma_{x,j}$
Oct.	5.02	2.31	0.61	June	12.76	3.32	0.17
Nov.	6.50	3.38	0.58	July	9.05	3.26	0.65
Dec.	7.33	3.23	0.50	Aug.	4.44	1.47	0.93
Jan.	6.42	2.95	0.31	Sept.	3.29	1.22	0.51
Feb.	5.35	2.62	0.38				
Mar.	5.02	1.66	0.37				
Apr.	6.42	1.80	0.44				
May	10.70	2.89	0.34				

15.6 Generate 100 years of monthly streamflow data for Cave Creek near Fort Spring, Kentucky, using the seasonal first order normal Markov model. Compare the correlogram and spectral density function of the simulated and observed data. (Basic data are in Appendix C.)

15.7 Write out and explain how a model such as described by equations 15.20 and 15.21 can be used as a higher order, multiseason Markov model. Apply the model to Cave Creek near Fort Spring, Kentucky, (Appendix C), using a second order, monthly Markov model.

15.8 Use the first order, normal Markov model to generate 100 years of annual runoff for the Spray River near Banff, Canada. Compare the correlogram and spectral density functions for the observed and simulated data.

15.9 Use equation 15.29 to show the individual generating equations for a 2-site model in terms of $\rho_{1,2}(0)$, $\rho_1(1)$, $\rho_2(1)$, x_{i+1}, and x_i.

15.10 What is $\rho_{j,h}(1)$ for the model given by equation 15.25?

15.11 For the model given by equation 15.30, show for a 2-site model using Young's solution the two generating equations in terms of $\rho_{1,2}(0)$, $\rho_1(1)$, $\rho_2(1)$, x_i and x_{i+1}. Consider $\rho_{1,2}(1)$ as not important.

15.12 Generate 1 year of rainfall letting the sequence of wet and dry days be defined by the Markov chain of example 15.1 and the amount of rainfall on a rainy day by a gamma distribution with a mean of 1 and a variance of 0.50 inches.

15.13 Generate a succession of 200 water level states for the situation described in examples 15.2 and 15.3. What fraction of the time was the reservoir level in each of the three states? How does this compare to the predicted results of example 15.2?

Appendices

APPENDIX A

Proof that s^2 is Unbiased Estimator for σ^2

$$\Sigma(X_i - \mu)^2 = \Sigma(X_i - \bar{X} + \bar{X} - \mu)^2 = \Sigma[(X_i - \bar{X})^2 + 2(X_i - \bar{X})(\bar{X} - \mu) + (\bar{X} - \mu)^2]$$

$$= \Sigma(X_i - \bar{X})^2 + 2(\bar{X} - \mu)\Sigma(X_i - \bar{X}) + (\bar{X} - \mu)^2 n$$

$$\Sigma(X_i - \mu)^2 = \Sigma(X_i - \bar{X})^2 + n(\bar{X} - \mu)^2 \qquad \text{since } \Sigma(X_i - \bar{X}) = 0$$

$$= (n-1)s^2 + n(\bar{X} - \mu)^2$$

$$s^2 = \frac{1}{n-1}\Sigma(X_i - \mu)^2 - \frac{n}{n-1}(\bar{X} - \mu)^2$$

$$E(s^2) = \frac{1}{n-1}\Sigma E(X_i - \mu)^2 - \frac{n}{n-1}E(\bar{X} - \mu)^2$$

$$= \frac{1}{n-1}\Sigma \mathrm{Var}(X_i) - \frac{n}{n-1}\mathrm{Var}\,\bar{X}$$

$$= \frac{n}{n-1}\sigma^2 - \frac{n}{n-1}\sigma^2/n$$

$$= \frac{n}{n-1}\sigma^2 - \frac{1}{n-1}\sigma^2$$

$$= \frac{n-1}{n-1}\sigma^2$$

$$E(s^2) = \sigma^2$$

Therefore s^2 is an unbiased estimator for σ^2

APPENDIX B

TABLE OF COMMON DISTRIBUTIONS AND THEIR PROPERTIES

distribution	density function	mean	variance	coef. skew	mgf	text pp
hypergeometric	$\binom{k}{x}\binom{N-k}{n-x}/\binom{N}{n}$	nk/N	$nk(N-k)(N-n)/N^2(N-1)$	$\frac{(N-2k)(N-2n)(N-1)^{1/2}}{(N-2)[nk(N-k)(N-n)]^{1/2}}$		68
binomial	$\binom{n}{x} p^x q^{n-x}$	np	npq	$(q-p)/\sqrt{npq}$	$(pe^t + q)^n$	70
geometric	pq^{x-1}	$1/p$	q/p^2	$(q-p)/\sqrt{q}$	$pe^t/(1-qe^t)$	73
negative binomial	$\binom{x-1}{k-1} p^k q^{x-k}$	k/p	kq/p^2		$pe^{kt}/(1-qe^t)^k$	75
Poisson	$\lambda^x e^{-\lambda}/x!$	λ	λ	$1/\sqrt{\lambda}$	$e^{\lambda(e^t-1)}$	76
uniform	$1/(\beta - \alpha)$	$(\beta + \alpha)/2$	$(\beta - \alpha)^2/12$	0		97
exponential	$\lambda e^{-\lambda x}$	$1/\lambda$	$1/\lambda^2$	2	$(1 - t/\lambda)^{-1}$	98
normal	$\frac{e^{-\frac{1}{2}(x-\mu)^2/\sigma^2}}{(2\pi\sigma^2)^{1/2}}$	μ	σ^2	0	$e^{\mu t + \sigma^2 t^2/2}$	84
lognormal	$\frac{\exp[-\frac{1}{2}(\ln x-\mu_y)^2/\sigma_y^2]}{(2\pi x^2\sigma_y^2)^{1/2}}$	$e^{\mu_y + \sigma_y^2/2}$	$\mu_x^2(e^{\sigma_y^2} - 1)$	$3c_v + c_v^3$ $c_v = (e^{\sigma_y^2} - 1)^{1/2}$		106
gamma	$\lambda^\eta x^{\eta-1} e^{-\lambda x}/\Gamma(\eta)$	η/λ	η/λ^2	$2/\sqrt{\eta}$	$(1-t/\lambda)^{-\eta}$	101

continued

distribution	density function	mean	variance	coef. skew	mgf	text pp
extreme value type I (max)	$\exp(-y - e^{-y})$ $y = (x - \beta)/\alpha$	$\beta + 0.577\alpha$	$1.645\,\alpha^2$	1.1396	$e^{\beta t}\Gamma(1 - t/\alpha)$	112
extreme value type I (min)	$\exp(y - e^{y})$ $y = (x - \beta)/\alpha$	$\beta - 0.577\alpha$	$1.645\,\alpha^2$	-1.1396	$e^{\beta t}\Gamma(1 + t/\alpha)$	112
extreme value type III (min)	$\frac{\alpha \exp[-\{(x-\epsilon)/(\beta-\epsilon)\}^\alpha]}{(x-\epsilon)^{\alpha-1}(\beta-\epsilon)^\alpha}$	$\epsilon + (\beta-\epsilon)\Gamma(1+1/\alpha)$	$(\beta-\epsilon)^2[\Gamma(1+2/\alpha) - \Gamma^2(1+1/\alpha)]$	eq. 6.60		114
chi-square	$\frac{x^{(\nu/2)-1} e^{-x/2}}{2^{\nu/2}\,\Gamma(\nu/2)}$	ν	2ν	$2/\sqrt{\nu/2}$	$(1-2t)^{-\nu/2}$	120
t	$\frac{\Gamma[(\nu+1)/2](1 + t^2/\nu)^{-(\nu+1)/2}}{\sqrt{\pi\nu}\,\Gamma(\nu/2)}$	0	$\nu/(\nu-2)$	0	none	120
F	eq. 6.90	$\gamma_1/(\gamma_2 - 2)$	$\frac{\gamma_2^2(\gamma_1+2)}{[\gamma_1(\gamma_2-2)(\gamma_2-4)]}$		none	122
beta	$x^{\alpha-1}(1-x)^{\beta-1}/B(\alpha, \beta)$	$\alpha/(\alpha+\beta)$	$\frac{\alpha\beta}{[(\alpha+\beta+1)(\alpha+\beta)^2]}$	$\frac{2(\beta-\alpha)(\alpha+\beta+1)^{1/2}}{(\alpha\beta)^{1/2}(\alpha+\beta+2)}$		118

APPENDIX C
DATA

Monthly Runoff[1] Cave Creek near Fort Spring, Kentucky

Water Year	Oct	Nov	Dec	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Total
1953	2	5	19	240	86	416	147	354	31	18	7	1	1326
1954	0	2	4	54	22	40	139	35	8	7	6	14	331
1955	2	4	30	73	463	579	59	197	55	24	28	3	1517
1956	4	6	13	59	637	469	192	28	32	64	38	8	1550
1957	7	10	172	308	325	103	392	68	24	6	5	2	1422
1958	3	106	432	200	221	117	235	236	19	369	170	12	2120
1959	6	9	17	270	195	112	102	24	24	5	4	2	770
1960	3	36	269	219	313	291	68	19	364	138	14	30	1764
1961	12	52	79	204	295	532	476	414	159	48	18	4	2291
1962	2	6	76	346	401	508	330	79	96	30	8	7	1889
1963	39	141	124	150	146	548	52	25	14	29	11	3	1282
1964	1	4	3	87	173	788	45	21	11	8	2	16	1158
1965	15	7	347	276	230	449	146	31	8	5	1	2	1517
1966	4	2	2	48	281	79	202	332	25	14	41	11	1040
1967	7	119	357	97	161	466	50	476	33	14	15	7	1802
1968	9	38	271	135	98	425	238	199	91	29	75	16	1625
1969	14	22	112	278	216	73	237	74	40	27	66	17	1177
1970	7	25	91	130	389	291	568	206	38	14	6	27	1792

1. The entries should be divided by 100 to get monthly flow in inches. 2 = .02 inches, 240 = 2.40 inches, etc. Data from U.S.G.S. Water Supply Papers.

Peak Discharge Cumberland River at Cumberland Falls, Kentucky

Water Year	Date	Gage Height (ft)	Discharge (cfs)
1916	Dec. 18	12.90	45,200
1917	Mar. 4	13.20	47,000
1918	Jan. 28	15.50	59,600
1919	Jan. 2, 3	11.40	37,200
1920	Jan. 23	12.60	43,500
1921	Apr. 17	9.80	28,200
1922	Feb. 21	11.70	38,600
1923	Feb. 3, 4	10.30	30,900
1924	Jan. 3	11.00	35,100
1925	Feb. 16	10.30	30,900
1926	May 17	9.20	25,600
1927	Dec. 25	14.10	51,900
1928	June 30	10.40	31,600
1929	Mar. 23	14.90	56,100
1930	Nov. 18	8.80	23,600
1931	Apr. 23	9.50	26,800

1932 ?

Peak Discharge Cumberland River at Cumberland Falls, Kentucky – (continued)

Water Year	Date	Gage Height (ft)	Discharge (cfs)
1933	Dec. 29	8.80	23,600
1934	Feb. 26	8.64	22,600
	Mar. 3	10.90	34,200
	Mar. 9	8.54	22,100
	Mar. 20	8.27	20,800
	Mar. 25	9.66	27,600
1935	Mar. 13	10.70	33,100
	Mar. 21	8.80	23,300
	Mar. 28	11.20	35,800
	Apr. 2	8.20	20,400
	Apr. 7	8.57	22,200
1936	Mar. 27	9.93	29,000
	Apr. 6	10.26	31,000
1937	Jan. 2	11.27	36,300
	Jan. 17	10.68	33,000
	Jan. 25	10.82	33,700
	Feb. 10	9.20	25,300
	June 11	8.35	21,200
1938	Jan. 25	8.30	20,900
1939	Feb. 3	14.15	52,300
	Feb. 12	8.63	22,500
	Feb. 15	8.85	23,600
	Feb. 28	8.20	20,400
1940	Apr. 21	8.37	21,400
1941	July 6	9.60	27,300
1942	Mar. 18	8.51	21,700
1943	Dec. 30	12.58	44,100
	Mar. 14	8.98	23,400
	Mar. 20	9.53	25,900
	Apr. 20	9.10	23,900
1944	Feb. 18	10.72	32,500
	Feb. 29	10.77	33,000
	Mar. 20	9.13	24,000
	Apr. 12	8.80	22,500
1945	Jan. 1	8.82	22,500
	Feb. 18	8.70	22,000
1946	Jan. 8	14.10	54,200
1947	Jan. 2	9.78	27,500
	Jan. 21	9.35	25,400
1948	Feb. 14	12.67	44,800
1949	Nov. 30	8.48	21,000
	Jan. 6	10.07	28,900
	Jan. 23	8.38	20,500
	Mar. 19	8.70	22,000
	Apr. 28	8.66	21,800

Peak Discharge Cumberland River at Cumberland Falls, Kentucky – (continued)

Water Year	Date	Gage Height (ft)	Discharge (cfs)
1950	Jan. 6	9.47	25,800
	Feb. 2	13.50	50,100
	Feb. 10	8.43	20,700
1951	Feb. 1	12.96	46,500
	Feb. 21	9.55	26,200
1952	Dec. 9	9.38	25,300
	Dec. 15	9.43	25,600
	Dec. 21	9.10	23,900
	Jan. 23	9.07	23,800
	Mar. 12	8.43	20,700
	Mar. 22	11.48	37,000
1953	Jan. 10	8.31	20,200
	Feb. 22	8.26	20,000
	May 20	8.43	20,700
1954	Jan. 23	9.30	24,900
1955	Dec. 31	9.00	23,400
	Mar. 19	11.96	40,000
	Mar. 22	12.21	41,600
1956	Feb. 4	9.35	25,200
	Feb. 18	11.16	35,100
	Mar. 15	9.02	23,500
	Apr. 17	8.82	22,600
1957	Dec. 15	9.11	24,000
	Jan. 29	14.55	57,400
1958	Nov. 19	10.92	33,700
	Dec. 9	8.66	21,800
	Apr. 25	8.83	22,600
	Apr. 30	8.88	22,800
	May 7	10.04	28,700
1959	Jan. 23	9.50	25,900
1960	Dec. 20	8.67	21,800
	June 25	8.31	20,200
1961	Feb. 28	9.30	25,000
	Mar. 9	8.47	21,100
1962	Jan. 28	9.25	24,700
	Feb. 28	12.47	43,300
	Apr. 7	9.13	24,100
	Apr. 13	9.54	26,100
1963	Mar. 7	9.26	24,800
	Mar. 12	11.78	38,800
	Mar. 17	11.58	37,600
1964	Feb. 17	7.77	18,200
1965	Jan. 12	9.12	24,100
	Mar. 27	10.42	30,800
	Mar. 29	11.11	34,800

Peak Discharge Cumberland River at Cumberland Falls, Kentucky – (continued)

Water Year	Date	Gage Height (ft)	Discharge (cfs)
1966	May 2	8.54	21,400
1967	Dec. 29	8.37	20,800
	Jan. 28	8.36	20,700
	Mar. 7	11.58	37,600
	July 8	8.32	20,500
1968	Dec. 23	8.31	20,400
1969	Mar. 13	9.79	27,400
	Feb. 3	7.89	18,700
1970	Dec. 30	12.44	43,100
	Feb. 16	8.80	22,600
	Apr. 3	8.55	21,500
	Apr. 30	9.48	25,800

Data from McCabe and U.S.G.S. Water Supply Papers

Total Precipitation for Week of March 1 – March 7, Ashland, Kentucky

Mean Weekly Precipitation 1.03 inches
Period of Record 1932 – 1968
Coefficient of Variation = 0.79
Skewness = 0.59

0.56	0.17	1.35	0.38
0.48	0.22	1.05	1.87
1.34	0.52	0.21	0.98
1.93	2.99	0.26	0.14
1.26	0.62	0.11	1.47
1.18	1.98	1.94	2.54
1.05	0.38	0.04	0.58
0.26	1.78	0.56	2.32
1.89	0.97	0.47	2.20
0.05			

Data from computer tapes belonging to Kentucky Division of Water, Department of Natural Resources and Environmental Protection, Frankfort, Kentucky.

Green River at Munfordville, Kentucky, 33085

Year	Sus. Sed. Load Tons	Ann. Disch. cfs-days
1952	1,070,886	1,431,201
1953	357,517	639,808
1954	295,187	384,267
1955	744,412	975,470
1956	736,131	1,055,609
1957	539,645	949,493
1958	771,530	1,285,048
1959	303,815	630,787
1960	378,891	856,188
1961	429,815	948,601
1962	940,227	1,243,398
1963	348,653	681,589
1964	426,989	633,862
1965	499,532	996,693
1966	272,618	566,710
1967	807,457	1,483,088
1968	496,158	1,049,411
1969	112,737	492,737
1970	411,824	856,469
1971	473,785	1,146,517
1972	494,030	1,120,905

Data from U.S.G.S. Water Supply Papers.

Streamflow, Walnut Gulch near Tombstone, Arizona

Year	Runoff (inches) June	July	Aug.	Sept.	Oct.	Annual	Peak Flow inches/hour
1957	.0000	.0000	.4938	.0000	.0000	.4938	.30737
1958	.0038	.0025	.2158	.0089	.0028	.2338	.09331
1959	.0000	.1533	.0296	.0000	.0005	.1834	.07438
1960	.0000	.0000	.0012	.0002	.0000	.0014	.00145
1961	.0000	.0014	.2352	.0272	.0000	.2638	.10555
1962	.0000	.1146	.0000	.0620	.0004	.1770	.02286
1963	.0000	.0223	.3668	.0871	.0000	.4762	.07280
1964	.0000	.1277	.0136	.3143	.0000	.4556	.11526
1965	.0000	.0012	.0022	.0332	.0000	.0366	.02261
1966	.0000	.0558	.0597	.0243	.0000	.1398	.04231
1967	.0000	.0026	.0111	.1191	.0000	.1328	.12579
1968	.0000	.0003	.0556	.0000	.0000	.0559	.02196
1969	.0000	.0255	.0485	.0029	.0000	.0769	.04512
1970	.0000	.0210	.0596	.0137	.0000	.0943	.01909
1971	.0000	.0301	.2090	.0010	.0013	.2414	.09717

Flow in other months was always 0.0000

Data from Leonard Lane, U.S.D.A., A.R.S., Southwest Watershed Research Center, Tucson, Arizona.

Monthly Rainfall (inches), Walnut Gulch near Tombstone, Arizona

Raingage 63.001 Monthly Total Rainfall (55-72)

Yr.	Jan.	Feb.	Mar.	Apr.	May	June	July	Aug.	Sept.	Oct.	Nov.	Dec.	Annual
55	.62	.20	.11	.00	.10	.50	7.57	4.08	.17	.18	.00	.12	13.65
56	.61	.10	.00	.00	.00	.38	2.08	1.31	.00	.05	.05	.06	4.64
57	.87	.04	.97	.00	.04	.66	2.75	1.75	.00	.83	.00	.21	8.12
58	.00	.84	1.33	.62	.18	1.06	2.22	3.74	2.33	1.11	.29	.00	13.72
59	.00	.42	.00	.08	.00	.51	4.74	3.83	.41	1.80	.64	.96	13.39
60	1.01	.57	.22	.00	.00	.11	1.04	1.00	1.55	.69	.00	.38	6.57
61	.44	.10	.00	.00	.00	.24	2.07	4.11	.83	1.32	.37	.54	10.02
62	.74	.07	.00	.00	.00	.08	2.78	.09	.93	.36	.58	.91	6.54
63	.17	.33	.01	.05	.00	.00	2.83	3.02	.85	.62	1.19	.19	9.26
64	.32	.01	.32	.28	.00	.00	3.03	1.56	2.47	.42	96	.09	9.46
65	.28	.08	.18	.00	.00	.04	3.06	.77	.78	.00	.12	2.91	8.22
66	.48	.90	.00	.31	.00	.05	4.39	4.89	2.41	.00	.20	.14	13.77
67	.00	.09	.02	.14	.16	.19	1.75	1.31	2.15	.37	.04	3.44	9.66
68	.55	.69	.76	.11	.00	.00	2.13	5.64	.43	.00	.28	.35	10.94
69	.10	.60	.16	.00	.20	.00	4.26	3.78	2.49	.00	.26	.51	12.36
70	.00	.28	1.16	.04	.00	.18	1.48	4.06	.83	.16	.00	.42	8.61
71	.05	.20	.00	.45	.00	.07	1.67	3.11	2.16	1.66	.07	1.21	10.65
72	.00	.00	.00	.00	.04	.89	1.81	1.30	.48	2.27	1.07	.06	7.92
$\bar{X}$	.35	.31	.29	.12	.04	.52	2.87	2.74	1.18	1.44	.34	.69	10.89
S_x	.33	.29	.44	.18	.07	1.16	1.56	1.62	.92	3.21	.39	.97	
C_v	.94	.94	1.52	1.50	1.74	2.21	0.54	0.59	.78	2.24	1.15	1.40	

Data from Leonard Lane, U.S.D.A., A.R.S., Southwest Watershed Research Center, Tucson, Arizona.

Annual Discharge for Piscataquis River, Dover-Foxcroft, Maine

Basin Area 297.0 square miles
Mean Annual Discharge 578.0 cfs
Period of Record 1902-1956
Coefficient of Variation 0.216
Skewness 0.009

602.3	427.1	394.2	522.5
712.3	836.4	584.4	547.9
238.1	707.5	639.3	644.5
614.4	563.0	838.1	557.2
733.5	630.0	585.5	697.1
528.4	480.3	428.3	659.5
506.3	842.1	591.3	470.5
503.4	549.1	566.4	537.5
541.0	717.3	619.0	664.1
502.3	564.1	335.2	510.4
497.1	531.2	645.0	531.2
722.5	410.4	401.1	492.5
776.3	655.5	623.1	608.1
746.2	477.4		

Data from Yevdjevich (1963).

Annual Discharge for North Llano River, Junction, Texas

Basin Area 914.0 square miles
Mean Annual Discharge 68.3 cfs
Period of Record 1915-1957
Coefficient of Variation 1.100
Skewness 1.691

25.4	10.3	83.1	138.0
70.7	18.9	58.9	69.1
167.1	157.0	32.2	10.3
101.0	12.8	10.3	108.1
243.1	73.0	14.0	271.1
215.1	42.7	298.1	74.0
40.5	45.7	61.0	40.9
40.2	19.9	8.5	26.3
89.7	34.6	21.4	11.2
7.8	4.7	0.8	12.6
8.2	92.1		

Data from Yevdjevich (1963).

Annual Discharge for Spray River, Banff, Canada

Basin Area 289.0 square miles
Mean Annual Discharge 461.6 cfs
Period of Record 1910-1955
Coefficient of Variation 0.293
Skewness -0.985

581.2	515.1	578.8	617.2
553.9	707.2	515.1	516.1
507.8	474.1	518.8	415.0
530.8	442.2	553.0	365.1
593.2	631.0	437.1	521.1
434.8	559.9	603.8	504.1
507.8	385.9	401.1	512.8
403.9	427.0	353.1	476.8
447.8	329.1	401.1	512.8
558.1	625.0	366.0	131.1
570.1	215.1	193.9	161.1
115.9			

Data from Yevdjevich (1963).

APPENDIX D

MATRIX ALGEBRA

DEFINITION

A matrix is any rectangular array of numbers; an sxn matrix has s rows and n columns. Example:

$$\begin{bmatrix} 2 & 5 & 8 \\ 5 & 1 & 3 \end{bmatrix}$$ is a 2x3 matrix

NOTATION

A matrix is denoted by an underlined letter $\underline{C}$ or $\underline{c}$. The usual notation for the elements of a matrix, $\underline{c}$, is the corresponding letter with subscripts to indicate the position in the matrix. There are two subscripts, the first indicating the row of the matrix and the second indicating the column. A matrix is uniquely defined by the position of the elements. Thus $\underline{c} = (c_{ij})$ where c_{ij} is an element of the matrix in row i and column j is given by

$$\underset{sxn}{\underline{c}} = \begin{bmatrix} c_{11} & c_{12} & \cdots & c_{1n} \\ c_{21} & c_{22} & \cdots & c_{2n} \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ c_{s1} & c_{s2} & & c_{sn} \end{bmatrix}$$ which is an sxn matrix

A row matrix is a matrix $\underline{x}$ such that $\underset{1xn}{\underline{x}} = [x_1, x_2, \ldots, x_n]$ is an 1xn row matrix or vector.

An nx1 column or column vector $\underline{y}$ is

$$\underset{nx1}{\underline{y}} = \begin{bmatrix} y_1 \\ y_2 \\ \cdot \\ \cdot \\ \cdot \\ y_n \end{bmatrix}$$

Symmetry of Matrices

An nxn matrix $\underline{c} = (c_{ij})$ is said to be symmetric if $(c_{ji}) = (c_{ij})$ for all i and j.

Transpose

If $\underline{c} = (c_{ij})$ is a sxn matrix, then the matrix obtained from $\underline{c}$ by interchanging rows and columns is defined to be the transposed matrix and denoted as $\underline{c}'$. Thus $\underline{c}' = (c_{ji})$ is an nxs matrix. Some identities regarding transposes are:

$$(\underline{A}')' = \underline{A}$$

$$(\underline{A} + \underline{B})' = \underline{A}' + \underline{B}'$$

$$(\underline{AB})' = \underline{B}'\,\underline{A}'$$

Equality of Matrices

Two matrices $\underline{A}$ and $\underline{B}$ are considered to be equal if and only if all corresponding elements are equal. One result of this is that equal matrices must have the same number of rows and the same number of columns. In other words, $\underline{A} = \underline{B}$ when $A_{ij} = B_{ij}$ for all i and j.

Addition of Matrices

To add two matrices $\underline{A}$ and $\underline{B}$ which are of the same size, mxn, we add to the corresponding elements in the two matrices and take their sums as the corresponding element in the sum matrix. Thus if $\underline{C} = \underline{A} + \underline{B}$, then $C_{ij} = A_{ij} + B_{ij}$ for all i and j. It is readily seen that in order for matrices to be added together, they must be of the same size and the sum matrix will also be of that size. It is also seen that $\underline{A} + \underline{B} = \underline{B} + \underline{A}$.

This rule may be extended step by step so that $\underline{M} = \underline{A} + \underline{B} + \ldots + \underline{K}$ implies $M_{ij} = A_{ij} + B_{ij} + \ldots + K_{ij}$ for all i and j with matrices $\underline{M}, \underline{A}, \underline{B}, \ldots, \underline{K}$ all being the same size

Matrix Subtraction

The operation of subtraction is quite similar to addition as might be expected. The difference matrix is a matrix whose elements are the difference between the corresponding elements of the two matrices. Thus $\underline{C} = \underline{B} - \underline{A}$ implies $C_{ij} = B_{ij} - A_{ij}$ for all i and j Again, all three matrices involved in this matrix equation must be of the same size

Scalar Multiplication

To multiply a matrix $\underline{A}$ by a scalar number λ, we multiply all elements A_{ij} by λ Thus

$\underline{C} = \lambda\,\underline{A}$ implies $C_{ij} = \lambda\,A_{ij}$ for all i and j.

Matrix Multiplication

To multiply two matrices, the element in the i^{th} row and the j^{th} column of product matrix $\underline{BA}$ is obtained by multiplying the elements in the i^{th} row of $\underline{B}$ by the corresponding elements in the j^{th} column of $\underline{A}$, then summing the products so obtained. Thus

$$\underline{C} = \underline{BA} \text{ implies } C_{ij} = B_{i1}\,A_{1j} + B_{i2}\,A_{2j} + \ldots B_{in}\,A_{nj} = \Sigma_{k=1}^{n}\,B_{ik}\,A_{kj}$$

In order for the product $\underline{BA}$ to exist, it is necessary that the $\underline{B}$ matrix have the same number of columns as the $\underline{A}$ matrix has rows. Thus if $\underline{B}$ is mxn and $\underline{A}$ is nxp, the product $\underline{C} = \underline{BA}$ exists. It is also seen that in this case $\underline{C}$ will be of size mxp, that is, the product matrix will have the same number of rows as the first matrix in the product and the same number of columns as the second.

For the above example, the product $\underline{AB}$ will exist only if m = p. In general th

product $\underline{BA}$ will not equal $\underline{AB}$ even when they both exist.

Since, in general, $\underline{AB}$ is usually not equal to $\underline{BA}$, it is necessary to define two different types of multiplication. It is said that $\underline{A}$ is premultiplied by $\underline{B}$ in the product $\underline{BA}$ and that $\underline{A}$ is postmultiplied by $\underline{B}$ in the product $\underline{AB}$.

General structure for matrix multiplication is:

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} e & f \\ g & h \end{bmatrix} = \begin{bmatrix} ae + bg & af + bh \\ ce + dg & cf + dh \end{bmatrix}$$

Examples:

$$\underline{A} = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \quad \underline{B} = \begin{bmatrix} 3 & 0 \\ -1 & 4 \end{bmatrix} \quad \underline{C} = [2, 3] \quad \underline{D} = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

$$\underline{AB} = \begin{bmatrix} 1 & 8 \\ 5 & 16 \end{bmatrix} \quad \underline{BA} = \begin{bmatrix} 3 & 6 \\ 11 & 14 \end{bmatrix} \quad \underline{CD} = -1 \quad \underline{DC} = \begin{bmatrix} 2 & 3 \\ -2 & -3 \end{bmatrix}$$

$$\underline{AD} = \begin{bmatrix} -1 \\ -1 \end{bmatrix} \quad \underline{CA} = [11, 16]$$

Identity Matrix

A square nxn matrix with diagonal elements equal to 1 and off diagonal elements equal to zero is an identity matrix $\underline{I}$. Thus $I_{ij} = 1$ for $i = j$ and $I_{ij} = 0$ for $i \neq j$. For example

$$\underset{3x3}{\underline{I}} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

It is easy to see then

$$\underset{mxn}{\underline{A}} \; \underset{nxn}{\underline{I}} = \underset{mxn}{\underline{A}} \quad \text{and} \quad \underset{mxm}{\underline{I}} \; \underset{mxn}{\underline{A}} = \underset{mxn}{\underline{A}}$$

Null Matrix

An nxm matrix of zeros is called a null matrix or zero matrix. If $\underline{N}$ is a null matrix, then $N_{ij} = 0$ for all i and j. Also

$$\underset{mxn}{\underline{A}} \; \underset{nxp}{\underline{N}} = \underset{mxp}{\underline{N}}$$

and

$$\underset{nxm}{\underline{N}} \; \underset{mxp}{\underline{A}} = \underset{nxp}{\underline{N}}$$

Diagonal Matrix

A square nxn matrix with at least some nonzero diagonal elements and all off diagonal elements equal to zero is a diagonal matrix. For example

$$\underset{4x4}{\underline{D}} = \begin{bmatrix} 2 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

is a diagonal matrix. The identity matrix is a diagonal matrix. Note that

$$\underset{mxn}{\underline{A}} \; \underset{nxn}{\underline{D}} = \underline{B} \text{ where } B_{ij} = A_{ij} D_{jj}$$

and

$$\underset{nxn}{\underline{D}} \; \underset{nxm}{\underline{A}} = \underline{B} \text{ where } B_{ij} = D_{ii} A_{ij}$$

Matrix Division

Matrix division is undefined. In scalar algebra the relationship, $x\,b = y$, was solved for b by dividing both sides of the relationship by x to give $b = y/x$. This operation cannot be performed in matrix algebra.

Matrix Inversion

Suppose we let $c = 1/x$ and solve the problem by multiplication:

$$c \times b = c\,y$$

since $c = 1/x$, then $c\,x = 1$ and

$$b = c\,y$$

In matrix algebra, the reciprocal matrix, which is analogous to the reciprocal in scalar algebra, is called the inverse matrix and is denoted by adding the exponent $(^{-1})$. Thus the inverse of $\underline{X}$ is denoted by $\underline{X}^{-1}$. The inverse of $\underline{X}$, $\underline{X}^{-1}$, when multiplied by $\underline{X}$ will produce an identity matrix.

$$\underline{X}\,\underline{X}^{-1} = \underline{X}^{-1}\,\underline{X} = \underline{I}$$

A matrix must be square (have the same number of rows as columns) to be inverted. The general procedure for matrix inversion may be found in texts on matrix algebra (Hadley 1961). A procedure will be given here for inverting 2x2 and 3x3 matrices.

Some properties of inverse matrices are:

$$(\underline{A}^{-1})^{-1} = \underline{A}$$

$$(\underline{A}\,\underline{B})^{-1} = \underline{B}^{-1}\underline{A}^{-1}$$

$$(\underline{A}^{t})^{-1} = (\underline{A}^{-1})^{t}$$

Let $\underline{M}$ be a 2x2 matrix given by

$$\underline{M} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

then the inverse is given by

$$\underline{M}^{-1} = \begin{bmatrix} d/D & -b/D \\ -c/D & a/D \end{bmatrix}$$

where $D = ad - bc$

Let $\underline{M}$ be a 3x3 matrix given by

$$\underline{M} = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & k \end{bmatrix}$$

then the inverse is given by

$$\underline{M}^{-1} = \begin{bmatrix} A & B & C \\ D & E & F \\ G & H & K \end{bmatrix}$$

where

$A = (ek - fh)/Z$
$B = -(bk - ch)/Z$
$C = (bf - ce)/Z$
$D = -(dk - fg)/Z$
$E = (ak - cg)/Z$
$F = -(af - cd)/Z$
$G = (dh - eg)/Z$
$H = -(ah - bg)/Z$
$K = (ae - bd)/Z$

and $Z = aek + bfg + cdh - ahf - dbk - gec$

Orthogonal Matrices

Any matrix $\underline{A}$ for which $\underline{A}' = \underline{A}^{-1}$ is called an orthogonal matrix. Note that a matrix must be square to be orthogonal. For orthogonal matrices we have

$$\underline{A}\,\underline{A}' = \underline{A}'\underline{A} = \underline{I}$$

Determinant of a Matrix

Associated with every square matrix is a unique scalar called its determinant. The determinant of the nxn matrix $\underline{A}$ is the sum

$$\Sigma \ldots \Sigma (-1)^{\alpha} a_{1j_1} a_{2j_2} \ldots a_{nj_n}$$

of all products consisting of one element from each row and column of $\underline{A}$ where α is the number of interchanges necessary among the factors a_{ij_i} to bring the subscripts j_i in-

to their natural order, 1, 2, ..., n. The determinant of $\underline{A}$ is written $|\underline{A}|$. For the three smallest square matrices, $|\underline{A}|$ is given by:

$$\underline{A} = [a_{11}] \qquad |\underline{A}| = a_{11}$$

$$\underline{A} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \qquad |\underline{A}| = a_{11}a_{22} - a_{12}a_{21}$$

$$\underline{A} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \qquad |\underline{A}| = a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - (a_{13}a_{22}a_{31} + a_{12}a_{21}a_{33} + a_{11}a_{23}a_{32})$$

For larger matrices, books on matrix algebra should be consulted for methods of evaluating $|\underline{A}|$.

Trace

The trace of a square matrix $\underline{A}$ written as Trace $\underline{A}$ is equal to the sum of the diagonal elements of $\underline{A}$

$$\text{Trace } \underline{A} = a_{11} + a_{22} + \ldots + a_{nn}$$

APPLICATION OF MATRIX ALGEBRA TO SIMPLE REGRESSION

Dependent variable	Independent variable
Y_1	X_1
Y_2	X_2
Y_3	X_3
.	.
.	.
.	.
Y_n	X_n

$$\underset{nx1}{Y} = \begin{bmatrix} Y_1 \\ Y_2 \\ Y_3 \\ . \\ . \\ Y_n \end{bmatrix} \quad \text{and} \quad \underset{nx2}{X} = \begin{bmatrix} 1 & X_1 \\ 1 & X_2 \\ 1 & X_3 \\ . & . \\ . & . \\ 1 & X_n \end{bmatrix}$$

$$\underset{2xn}{X'} = \begin{bmatrix} 1 & 1 & 1 & \cdots & 1 \\ X_1 & X_2 & X_3 & \cdots & X_n \end{bmatrix}$$

$$\underset{2x2}{X'X} = \begin{bmatrix} n & \Sigma X \\ \Sigma X & \Sigma X^2 \end{bmatrix}$$

$$\underset{2x1}{X'Y} = \begin{bmatrix} \Sigma Y \\ \Sigma XY \end{bmatrix}$$

$$\underset{nx1}{Y} = \underset{nx1}{X\beta}$$

$$\underset{2x1}{X'Y} = \underset{2x1}{X'X\beta}$$

$$\underset{2x1}{(X'X)^{-1}X'Y} = \underset{2x1}{(X'X)^{-1}(X'X)\beta} = \underline{\beta}$$

or

$$\underline{\beta} = (X'X)^{-1}X'Y$$

$$= \begin{bmatrix} \frac{\Sigma X^2}{n\Sigma X^2 - (\Sigma X)^2} & \frac{-\Sigma X}{n\Sigma X^2 - (\Sigma X)^2} \\ \frac{-\Sigma X}{n\Sigma X^2 - (\Sigma X)^2} & \frac{n}{n\Sigma X^2 - (\Sigma X)^2} \end{bmatrix} \begin{bmatrix} \Sigma Y \\ \Sigma XY \end{bmatrix}$$

$$\underline{\beta} = \begin{bmatrix} \frac{\Sigma X^2 \Sigma Y - \Sigma X \Sigma XY}{n\Sigma X^2 - (\Sigma X)^2} \\ \frac{-\Sigma X \Sigma Y + n\Sigma XY}{n\Sigma X^2 - (\Sigma X)^2} \end{bmatrix} = \begin{bmatrix} \hat{a} \\ \hat{b} \end{bmatrix}$$

From this it can be seen that

$$\hat{b} = \frac{\Sigma XY - \Sigma X \Sigma Y/n}{\Sigma X^2 - (\Sigma X)^2/n}$$

which agrees with equation 9.6. Through algebraic manipulation it can also be shown that the expression for $\hat{a}$ agrees with equation 9.7.

APPENDIX E

STATISTICAL TABLES

Table E.1. Binomial Coefficients

n	$\binom{n}{0}$	$\binom{n}{1}$	$\binom{n}{2}$	$\binom{n}{3}$	$\binom{n}{4}$	$\binom{n}{5}$	$\binom{n}{6}$	$\binom{n}{7}$	$\binom{n}{8}$	$\binom{n}{9}$	$\binom{n}{10}$
0	1										
1	1	1									
2	1	2	1								
3	1	3	3	1							
4	1	4	6	4	1						
5	1	5	10	10	5	1					
6	1	6	15	20	15	6	1				
7	1	7	21	35	35	21	7	1			
8	1	8	28	56	70	56	28	8	1		
9	1	9	36	84	126	126	84	36	9	1	
10	1	10	45	120	210	252	210	120	45	10	1
11	1	11	55	165	330	462	462	330	165	55	11
12	1	12	66	220	495	792	924	792	495	220	66
13	1	13	78	286	715	1287	1716	1716	1287	715	286
14	1	14	91	364	1001	2002	3003	3432	3003	2002	1001
15	1	15	105	455	1365	3003	5005	6435	6435	5005	3003
16	1	16	120	560	1820	4368	8008	11440	12870	11440	8008
17	1	17	136	680	2380	6188	12376	19448	24310	24310	19448
18	1	18	153	816	3060	8568	18564	31824	43758	48620	43758
19	1	19	171	969	3876	11628	27132	50388	75582	92378	92378
20	1	20	190	1140	4845	15504	38760	77520	125970	167960	184756

Table E.2. Cumulative Poisson Distribution

x	λ = 0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
0	0.9048	0.8187	0.7408	0.6730	0.6065	0.5488	0.4966	0.4493	0.4066
1	0.9953	0.9825	0.9631	0.9384	0.9098	0.8781	0.8442	0.8088	0.7725
2	0.9998	0.9989	0.9964	0.9921	0.9856	0.9769	0.9659	0.9526	0.9371
3	1.0000	0.9999	0.9997	0.9992	0.9982	0.9966	0.9942	0.9909	0.9865
4		1.0000	1.0000	0.9999	0.9998	0.9996	0.9992	0.9986	0.9977
5				1.0000	1.0000	1.0000	0.9999	0.9998	0.9997
6							1.0000	1.0000	1.0000

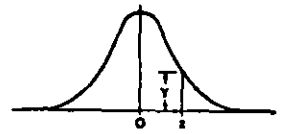
x	λ = 1.0	1.5	2.0	2.5	3.0	3.5	4.0	4.5	5.0
0	0.3679	0.2231	0.1353	0.0821	0.0498	0.0302	0.0183	0.0111	0.0067
1	0.7358	0.5578	0.4060	0.2873	0.1991	0.1359	0.0916	0.0611	0.0404
2	0.9197	0.8088	0.6767	0.5438	0.4232	0.3208	0.2381	0.1736	0.1247
3	0.9810	0.9344	0.8571	0.7576	0.6472	0.5366	0.4335	0.3423	0.2650
4	0.9963	0.9814	0.9473	0.8912	0.8153	0.7254	0.6288	0.5321	0.4405
5	0.9994	0.9955	0.9834	0.9580	0.9161	0.8576	0.7851	0.7029	0.6160
6	0.9999	0.9991	0.9955	0.9858	0.9665	0.9347	0.8893	0.8311	0.7622
7	1.0000	0.9998	0.9989	0.9958	0.9881	0.9733	0.9489	0.9134	0.8666
8		1.0000	0.9998	0.9989	0.9962	0.9901	0.9786	0.9597	0.9319
9			1.0000	0.9997	0.9989	0.9967	0.9919	0.9829	0.9682
10				0.9999	0.9997	0.9990	0.9972	0.9933	0.9863
11				1.0000	0.9999	0.9997	0.9991	0.9976	0.9945
12					1.0000	0.9999	0.9997	0.9992	0.9980
13						1.0000	0.9999	0.9997	0.9993
14							1.0000	0.9999	0.9998
15								1.0000	0.9999
16									1.0000

x	λ = 5.5	6.0	6.5	7.0	7.5	8.0	8.5	9.0	9.5
0	0.0041	0.0025	0.0015	0.0009	0.0006	0.0003	0.0002	0.0001	0.0001
1	0.0266	0.0174	0.0113	0.0073	0.0047	0.0030	0.0019	0.0012	0.0008
2	0.0884	0.0620	0.0430	0.0296	0.0203	0.0138	0.0093	0.0062	0.0042
3	0.2017	0.1512	0.1118	0.0818	0.0591	0.0424	0.0301	0.0212	0.0149
4	0.3575	0.2851	0.2237	0.1730	0.1321	0.0996	0.0744	0.0550	0.0403
5	0.5289	0.4457	0.3690	0.3007	0.2414	0.1912	0.1496	0.1157	0.0885
6	0.6860	0.6063	0.5265	0.4497	0.3782	0.3134	0.2562	0.2068	0.1649
7	0.8095	0.7440	0.6728	0.5987	0.5246	0.4530	0.3856	0.3239	0.2687
8	0.8944	0.8472	0.7916	0.7291	0.6620	0.5925	0.5231	0.4557	0.3918
9	0.9462	0.9161	0.8774	0.8305	0.7764	0.7166	0.6530	0.5874	0.5218
10	0.9747	0.9574	0.9332	0.9015	0.8622	0.8159	0.7634	0.7060	0.6453
11	0.9890	0.9799	0.9661	0.9466	0.9208	0.8881	0.8487	0.8030	0.7520
12	0.9955	0.9912	0.9840	0.9730	0.9573	0.9362	0.9091	0.8758	0.8364
13	0.9983	0.9964	0.9929	0.9872	0.9784	0.9658	0.9486	0.9261	0.8981
14	0.9994	0.9986	0.9970	0.9943	0.9897	0.9827	0.9726	0.9585	0.9400
15	0.9998	0.9995	0.9988	0.9976	0.9954	0.9918	0.9862	0.9780	0.9665
16	0.9999	0.9998	0.9996	0.9990	0.9980	0.9963	0.9934	0.9889	0.9823
17	1.0000	0.9999	0.9998	0.9996	0.9992	0.9984	0.9970	0.9947	0.9911
18		1.0000	0.9999	0.9999	0.9997	0.9994	0.9987	0.9976	0.9957
19			1.0000	1.0000	0.9999	0.9997	0.9995	0.9989	0.9980
20					1.0000	0.9999	0.9998	0.9996	0.9991
21						1.0000	0.9999	0.9998	0.9996
22							1.0000	0.9999	0.9999
23								1.0000	0.9999
24									1.0000

Table E.2. (continued). Cumulative Poisson Distribution

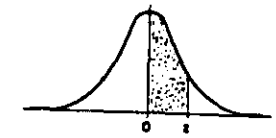
x	λ = 10.0	11.0	12.0	13.0	14.0	15.0	16.0	17.0	18.0
0	0.0000	0.0000	0.0000						
1	0.0005	0.0002	0.0001	0.0000	0.0000				
2	0.0028	0.0012	0.0005	0.0002	0.0001	0.0000	0.0000		
3	0.0103	0.0049	0.0023	0.0010	0.0005	0.0002	0.0001	0.0000	0.0000
4	0.0293	0.0151	0.0076	0.0037	0.0018	0.0009	0.0004	0.0002	0.0001
5	0.0671	0.0375	0.0203	0.0107	0.0055	0.0028	0.0014	0.0007	0.0003
6	0.1301	0.0786	0.0458	0.0259	0.0142	0.0076	0.0040	0.0021	0.0010
7	0.2202	0.1432	0.0895	0.0540	0.0316	0.0180	0.0100	0.0054	0.0029
8	0.3328	0.2320	0.1550	0.0998	0.0621	0.0374	0.0220	0.0126	0.0071
9	0.4579	0.3405	0.2424	0.1658	0.1094	0.0699	0.0433	0.0261	0.0154
10	0.5830	0.4599	0.3472	0.2517	0.1757	0.1185	0.0774	0.0491	0.0304
11	0.6968	0.5793	0.4616	0.3532	0.2600	0.1848	0.1270	0.0847	0.0549
12	0.7916	0.6887	0.5760	0.4631	0.3585	0.2676	0.1931	0.1350	0.0917
13	0.8645	0.7813	0.6815	0.5730	0.4644	0.3632	0.2745	0.2009	0.1426
14	0.9165	0.8540	0.7720	0.6751	0.5704	0.4657	0.3675	0.2808	0.2081
15	0.9513	0.9074	0.8444	0.7636	0.6694	0.5681	0.4667	0.3715	0.2867
16	0.9730	0.9441	0.8987	0.8355	0.7559	0.6641	0.5660	0.4677	0.3750
17	0.9857	0.9678	0.9370	0.8905	0.8272	0.7489	0.6593	0.5640	0.4686
18	0.9928	0.9823	0.9626	0.9302	0.8826	0.8195	0.7423	0.6550	0.5622
19	0.9965	0.9907	0.9787	0.9573	0.9235	0.8752	0.8122	0.7363	0.6509
20	0.9984	0.9953	0.9884	0.9750	0.9521	0.9170	0.8682	0.8055	0.7307
21	0.9993	0.9977	0.9939	0.9859	0.9712	0.9469	0.9108	0.8615	0.7991
22	0.9997	0.9990	0.9970	0.9924	0.9833	0.9673	0.9418	0.9047	0.8551
23	0.9999	0.9995	0.9985	0.9960	0.9907	0.9805	0.9633	0.9367	0.8989
24	1.0000	0.9998	0.9993	0.9980	0.9950	0.9888	0.9777	0.9594	0.9317
25		0.9999	0.9997	0.9990	0.9974	0.9938	0.9869	0.9748	0.9554
26		1.0000	0.9999	0.9995	0.9987	0.9967	0.9925	0.9848	0.9718
27			0.9999	0.9998	0.9994	0.9983	0.9959	0.9912	0.9827
28			1.0000	0.9999	0.9997	0.9991	0.9978	0.9950	0.9897
29				1.0000	0.9999	0.9996	0.9989	0.9973	0.9941
30					0.9999	0.9998	0.9994	0.9986	0.9967
31					1.0000	0.9999	0.9997	0.9993	0.9982
32						1.0000	0.9999	0.9996	0.9990
33							0.9999	0.9998	0.9995
34							1.0000	0.9999	0.9998
35								1.0000	0.9999
36									0.9999
37									1.0000

Table E.3. Ordinates Y of the Standard Normal Density



z	0	1	2	3	4	5	6	7	8	9
0.0	0.3989	0.3989	0.3989	0.3988	0.3986	0.3984	0.3982	0.3980	0.3977	0.3973
0.1	0.3970	0.3965	0.3961	0.3956	0.3951	0.3945	0.3939	0.3932	0.3925	0.3918
0.2	0.3910	0.3902	0.3894	0.3885	0.3876	0.3867	0.3857	0.3847	0.3836	0.3825
0.3	0.3814	0.3802	0.3790	0.3778	0.3765	0.3752	0.3739	0.3725	0.3712	0.3697
0.4	0.3683	0.3668	0.3653	0.3637	0.3621	0.3605	0.3589	0.3572	0.3555	0.3538
0.5	0.3521	0.3503	0.3485	0.3467	0.3448	0.3429	0.3410	0.3391	0.3372	0.3352
0.6	0.3332	0.3312	0.3292	0.3271	0.3251	0.3230	0.3209	0.3187	0.3166	0.3144
0.7	0.3123	0.3101	0.3079	0.3056	0.3034	0.3011	0.2989	0.2966	0.2943	0.2920
0.8	0.2897	0.2874	0.2850	0.2827	0.2803	0.2780	0.2756	0.2732	0.2709	0.2685
0.9	0.2661	0.2637	0.2613	0.2589	0.2565	0.2541	0.2516	0.2492	0.2468	0.2444
1.0	0.2420	0.2396	0.2371	0.2347	0.2323	0.2299	0.2275	0.2251	0.2227	0.2203
1.1	0.2179	0.2155	0.2131	0.2107	0.2083	0.2059	0.2036	0.2012	0.1989	0.1965
1.2	0.1942	0.1919	0.1895	0.1872	0.1849	0.1827	0.1804	0.1781	0.1759	0.1736
1.3	0.1714	0.1692	0.1669	0.1647	0.1626	0.1604	0.1582	0.1561	0.1540	0.1518
1.4	0.1497	0.1476	0.1456	0.1435	0.1415	0.1394	0.1374	0.1354	0.1334	0.1315
1.5	0.1295	0.1276	0.1257	0.1238	0.1219	0.1200	0.1182	0.1163	0.1145	0.1127
1.6	0.1109	0.1092	0.1074	0.1057	0.1040	0.1023	0.1006	0.0989	0.0973	0.0957
1.7	0.0941	0.0925	0.0909	0.0893	0.0878	0.0863	0.0848	0.0833	0.0818	0.0804
1.8	0.0790	0.0775	0.0762	0.0748	0.0734	0.0721	0.0707	0.0694	0.0682	0.0669
1.9	0.0656	0.0644	0.0632	0.0620	0.0608	0.0596	0.0584	0.0573	0.0562	0.0551
2.0	0.0540	0.0529	0.0519	0.0508	0.0498	0.0488	0.0478	0.0468	0.0459	0.0449
2.1	0.0440	0.0431	0.0422	0.0413	0.0404	0.0396	0.0387	0.0379	0.0371	0.0363
2.2	0.0355	0.0347	0.0339	0.0332	0.0325	0.0317	0.0310	0.0303	0.0297	0.0290
2.3	0.0283	0.0277	0.0271	0.0264	0.0258	0.0252	0.0246	0.0241	0.0235	0.0229
2.4	0.0224	0.0219	0.0213	0.0208	0.0203	0.0198	0.0194	0.0189	0.0184	0.0180
2.5	0.0175	0.0171	0.0167	0.0163	0.0159	0.0155	0.0151	0.0147	0.0143	0.0139
2.6	0.0136	0.0132	0.0129	0.0126	0.0122	0.0119	0.0116	0.0113	0.0110	0.0107
2.7	0.0104	0.0101	0.0099	0.0096	0.0093	0.0091	0.0088	0.0086	0.0084	0.0081
2.8	0.0079	0.0077	0.0075	0.0073	0.0071	0.0069	0.0067	0.0065	0.0063	0.0061
2.9	0.0060	0.0058	0.0056	0.0055	0.0053	0.0051	0.0050	0.0048	0.0047	0.0046
3.0	0.0044	0.0043	0.0042	0.0041	0.0039	0.0038	0.0037	0.0036	0.0035	0.0034
3.1	0.0033	0.0032	0.0031	0.0030	0.0029	0.0028	0.0027	0.0026	0.0025	0.0025
3.2	0.0024	0.0023	0.0022	0.0022	0.0021	0.0020	0.0020	0.0019	0.0018	0.0018
3.3	0.0017	0.0017	0.0016	0.0016	0.0015	0.0015	0.0014	0.0014	0.0013	0.0013
3.4	0.0012	0.0012	0.0012	0.0011	0.0011	0.0010	0.0010	0.0010	0.0009	0.0009
3.5	0.0009	0.0008	0.0008	0.0008	0.0008	0.0007	0.0007	0.0007	0.0007	0.0006
3.6	0.0006	0.0006	0.0006	0.0005	0.0005	0.0005	0.0005	0.0005	0.0005	0.0004
3.7	0.0004	0.0004	0.0004	0.0004	0.0004	0.0004	0.0003	0.0003	0.0003	0.0003
3.8	0.0003	0.0003	0.0003	0.0003	0.0003	0.0002	0.0002	0.0002	0.0002	0.0002
3.9	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0001	0.0001

Table E.4. Areas under the Standard Normal Density from 0 to z.



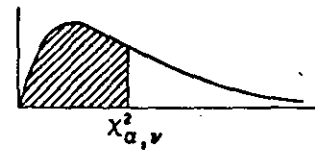
z	0	1	2	3	4	5	6	7	8	9
0.0	0.00	0.0040	0.0080	0.0120	0.0160	0.0199	0.0239	0.0279	0.0319	0.0359
0.1	0.0398	0.0438	0.0478	0.0517	0.0557	0.0596	0.0636	0.0675	0.0714	0.0753
0.2	0.0793	0.0832	0.0871	0.0910	0.0948	0.0987	0.1026	0.1064	0.1103	0.1141
0.3	0.1179	0.1217	0.1255	0.1293	0.1331	0.1368	0.1406	0.1443	0.1480	0.1517
0.4	0.1554	0.1591	0.1628	0.1664	0.1700	0.1736	0.1772	0.1808	0.1844	0.1879
0.5	0.1915	0.1950	0.1985	0.2019	0.2054	0.2088	0.2123	0.2157	0.2190	0.2224
0.6	0.2257	0.2291	0.2324	0.2357	0.2389	0.2422	0.2454	0.2486	0.2517	0.2549
0.7	0.2580	0.2611	0.2642	0.2673	0.2703	0.2734	0.2764	0.2793	0.2823	0.2852
0.8	0.2881	0.2910	0.2939	0.2967	0.2995	0.3023	0.3051	0.3078	0.3106	0.3133
0.9	0.3159	0.3186	0.3212	0.3238	0.3264	0.3289	0.3315	0.3340	0.3365	0.3389
1.0	0.3413	0.3438	0.3461	0.3485	0.3508	0.3531	0.3554	0.3577	0.3599	0.3621
1.1	0.3643	0.3665	0.3686	0.3708	0.3729	0.3749	0.3770	0.3790	0.3810	0.3830
1.2	0.3849	0.3869	0.3888	0.3906	0.3925	0.3943	0.3962	0.3980	0.3997	0.4015
1.3	0.4032	0.4049	0.4066	0.4082	0.4099	0.4115	0.4131	0.4147	0.4162	0.4177
1.4	0.4192	0.4207	0.4222	0.4236	0.4251	0.4265	0.4279	0.4292	0.4306	0.4319
1.5	0.4332	0.4345	0.4357	0.4370	0.4382	0.4394	0.4406	0.4418	0.4429	0.4441
1.6	0.4452	0.4463	0.4474	0.4484	0.4495	0.4505	0.4515	0.4525	0.4535	0.4545
1.7	0.4554	0.4564	0.4573	0.4582	0.4591	0.4599	0.4608	0.4616	0.4625	0.4633
1.8	0.4641	0.4648	0.4656	0.4664	0.4671	0.4678	0.4686	0.4693	0.4699	0.4706
1.9	0.4713	0.4719	0.4726	0.4732	0.4738	0.4744	0.4750	0.4756	0.4761	0.4767
2.0	0.4772	0.4778	0.4783	0.4788	0.4793	0.4798	0.4803	0.4808	0.4812	0.4817
2.1	0.4821	0.4826	0.4830	0.4834	0.4838	0.4842	0.4846	0.4850	0.4854	0.4857
2.2	0.4861	0.4864	0.4868	0.4871	0.4875	0.4878	0.4881	0.4884	0.4887	0.4890
2.3	0.4893	0.4896	0.4898	0.4901	0.4904	0.4906	0.4909	0.4911	0.4913	0.4916
2.4	0.4918	0.4920	0.4922	0.4924	0.4927	0.4929	0.4931	0.4932	0.4934	0.4936
2.5	0.4938	0.4940	0.4941	0.4943	0.4945	0.4946	0.4948	0.4949	0.4951	0.4952
2.6	0.4953	0.4955	0.4956	0.4957	0.4959	0.4960	0.4961	0.4962	0.4963	0.4964
2.7	0.4965	0.4966	0.4967	0.4968	0.4969	0.4970	0.4971	0.4972	0.4973	0.4974
2.8	0.4974	0.4975	0.4976	0.4977	0.4977	0.4978	0.4979	0.4979	0.4980	0.4981
2.9	0.4981	0.4982	0.4982	0.4983	0.4984	0.4984	0.4985	0.4985	0.4986	0.4986
3.0	0.4986	0.4987	0.4987	0.4988	0.4988	0.4989	0.4989	0.4989	0.4990	0.4990
3.1	0.4990	0.4991	0.4991	0.4991	0.4992	0.4992	0.4992	0.4992	0.4993	0.4993
3.2	0.4993	0.4993	0.4994	0.4994	0.4994	0.4994	0.4994	0.4995	0.4995	0.4995
3.3	0.4995	0.4995	0.4995	0.4996	0.4996	0.4996	0.4996	0.4996	0.4996	0.4997
3.4	0.4997	0.4997	0.4997	0.4997	0.4997	0.4997	0.4997	0.4997	0.4997	0.4998
3.5	0.4998	0.4998	0.4998	0.4998	0.4998	0.4998	0.4998	0.4998	0.4998	0.4998
3.6	0.4998	0.4998	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999
3.7	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999
3.8	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999
3.9	0.5000	0.5000	0.5000	0.5000	0.5000	0.5000	0.5000	0.5000	0.5000	0.5000

Table E.5. Percentile Values ($t_{\alpha,\nu}$) for the t Distribution with ν Degrees of Freedom--(shaded area = α).

ν	$t_{.995}$	$t_{.99}$	$t_{.975}$	$t_{.95}$	$t_{.90}$	$t_{.80}$	$t_{.75}$	$t_{.70}$	$t_{.60}$	$t_{.55}$
1	63.66	31.82	12.71	6.31	3.08	1.376	1.000	.727	.325	.158
2	9.92	6.96	4.30	2.92	1.89	1.061	.816	.617	.289	.142
3	5.84	4.54	3.18	2.35	1.64	.978	.765	.584	.277	.137
4	4.60	3.75	2.78	2.13	1.53	.941	.741	.569	.271	.134
5	4.03	3.36	2.57	2.02	1.48	.920	.727	.559	.267	.132
6	3.71	3.14	2.45	1.94	1.44	.906	.718	.553	.265	.131
7	3.50	3.00	2.36	1.90	1.42	.896	.711	.549	.263	.130
8	3.36	2.90	2.31	1.86	1.40	.889	.706	.546	.262	.130
9	3.25	2.82	2.26	1.83	1.38	.883	.703	.543	.261	.129
10	3.17	2.76	2.23	1.81	1.37	.879	.700	.542	.260	.129
11	3.11	2.72	2.20	1.80	1.36	.876	.697	.540	.260	.129
12	3.06	2.68	2.18	1.78	1.36	.873	.695	.539	.259	.128
13	3.01	2.65	2.16	1.77	1.35	.870	.694	.538	.259	.128
14	2.98	2.62	2.14	1.76	1.34	.868	.692	.537	.258	.128
15	2.95	2.60	2.13	1.75	1.34	.866	.691	.536	.258	.128
16	2.92	2.58	2.12	1.75	1.34	.865	.690	.535	.258	.128
17	2.90	2.57	2.11	1.74	1.33	.863	.689	.534	.257	.128
18	2.88	2.55	2.10	1.73	1.33	.862	.688	.534	.257	.127
19	2.86	2.54	2.09	1.73	1.33	.861	.688	.533	.257	.127
20	2.84	2.53	2.09	1.72	1.32	.860	.687	.533	.257	.127
21	2.83	2.52	2.08	1.72	1.32	.859	.686	.532	.257	.127
22	2.82	2.51	2.07	1.72	1.32	.858	.686	.532	.256	.127
23	2.81	2.50	2.07	1.71	1.32	.858	.685	.532	.256	.127
24	2.80	2.49	2.06	1.71	1.32	.857	.685	.531	.256	.127
25	2.79	2.48	2.06	1.71	1.32	.856	.684	.531	.256	.127
26	2.78	2.48	2.06	1.71	1.32	.856	.684	.531	.256	.127
27	2.77	2.47	2.05	1.70	1.31	.855	.684	.531	.256	.127
28	2.76	2.47	2.05	1.70	1.31	.855	.683	.530	.256	.127
29	2.76	2.46	2.04	1.70	1.31	.854	.683	.530	.256	.127
30	2.75	2.46	2.04	1.70	1.31	.854	.683	.530	.256	.127
40	2.70	2.42	2.02	1.68	1.30	.851	.681	.529	.255	.126
60	2.66	2.39	2.00	1.67	1.30	.848	.679	.527	.254	.126
120	2.62	2.36	1.98	1.66	1.29	.845	.677	.526	.254	.126
∞	2.58	2.33	1.96	1.645	1.28	.842	.674	.524	.253	.126

Taken from: Statistical Tables for Biological, Agricultural and Medical Research, published by Longman Group Ltd., London. (previously published by Oliver and Boyd, Edinburgh), and by permission of the authors and publishers.

Table E.6. Percentile Values ($\chi^2_{\alpha,\nu}$) for the Chi-Square Distribution with ν Degrees o Freedom (shaded area = α).



ν	$\chi^2_{.995}$	$\chi^2_{.99}$	$\chi^2_{.975}$	$\chi^2_{.95}$	$\chi^2_{.90}$	$\chi^2_{.75}$	$\chi^2_{.50}$	$\chi^2_{.25}$	$\chi^2_{.10}$	$\chi^2_{.05}$	$\chi^2_{.025}$	$\chi^2_{.01}$	$\chi^2_{.005}$
1	7.88	6.63	5.02	3.84	2.71	1.32	.455	.102	.0158	.0039	.0010	.0002	.0000
2	10.6	9.21	7.38	5.99	4.61	2.77	1.39	.575	.211	.103	.0506	.0201	.0100
3	12.8	11.3	9.35	7.81	6.25	4.11	2.37	1.21	.584	.352	.216	.115	.072
4	14.9	13.3	11.1	9.49	7.78	5.39	3.36	1.92	1.06	.711	.484	.297	.207
5	16.7	15.1	12.8	11.1	9.24	6.63	4.35	2.67	1.61	1.15	.831	.554	.412
6	18.5	16.8	14.4	12.6	10.6	7.84	5.35	3.45	2.20	1.64	1.24	.872	.676
7	20.3	18.5	16.0	14.1	12.0	9.04	6.35	4.25	2.83	2.17	1.69	1.24	.989
8	22.0	20.1	17.5	15.5	13.4	10.2	7.34	5.07	3.49	2.73	2.18	1.65	1.34
9	23.6	21.7	19.0	16.9	14.7	11.4	8.34	5.90	4.17	3.33	2.70	2.09	1.73
10	25.2	23.2	20.5	18.3	16.0	12.5	9.34	6.74	4.87	3.94	3.25	2.56	2.16
11	26.8	24.7	21.9	19.7	17.3	13.7	10.3	7.58	5.58	4.57	3.82	3.05	2.60
12	28.3	26.2	23.3	21.0	18.5	14.8	11.3	8.44	6.30	5.23	4.40	3.57	3.07
13	29.8	27.7	24.7	22.4	19.8	16.0	12.3	9.30	7.04	5.89	5.01	4.11	3.57
14	31.3	29.1	26.1	23.7	21.1	17.1	13.3	10.2	7.79	6.57	5.63	4.66	4.07
15	32.8	30.6	27.5	25.0	22.3	18.2	14.3	11.0	8.55	7.26	6.26	5.23	4.60
16	34.3	32.0	28.8	26.3	23.5	19.4	15.3	11.9	9.31	7.96	6.91	5.81	5.14
17	35.7	33.4	30.2	27.6	24.8	20.5	16.3	12.8	10.1	8.67	7.56	6.41	5.70
18	37.2	34.8	31.5	28.9	26.0	21.6	17.3	13.7	10.9	9.39	8.23	7.01	6.26
19	38.6	36.2	32.9	30.1	27.2	22.7	18.3	14.6	11.7	10.1	8.91	7.63	6.84
20	40.0	37.6	34.2	31.4	28.4	23.8	19.3	15.5	12.4	10.9	9.59	8.26	7.43
21	41.4	38.9	35.5	32.7	29.6	24.9	20.3	16.3	13.2	11.6	10.3	8.90	8.03
22	42.8	40.3	36.8	33.9	30.8	26.0	21.3	17.2	14.0	12.3	11.0	9.54	8.64
23	44.2	41.6	38.1	35.2	32.0	27.1	22.3	18.1	14.8	13.1	11.7	10.2	9.26
24	45.6	43.0	39.4	36.4	33.2	28.2	23.3	19.0	15.7	13.8	12.4	10.9	9.89
25	46.9	44.3	40.6	37.7	34.4	29.3	24.3	19.9	16.5	14.6	13.1	11.5	10.5
26	48.3	45.6	41.9	38.9	35.6	30.4	25.3	20.8	17.3	15.4	13.8	12.2	11.2
27	49.6	47.0	43.2	40.1	36.7	31.5	26.3	21.7	18.1	16.2	14.6	12.9	11.8
28	51.0	48.3	44.5	41.3	37.9	32.6	27.3	22.7	18.9	16.9	15.3	13.6	12.5
29	52.3	49.6	45.7	42.6	39.1	33.7	28.3	23.6	19.8	17.7	16.0	14.3	13.1
30	53.7	50.9	47.0	43.8	40.3	34.8	29.3	24.5	20.6	18.5	16.8	15.0	13.8
40	66.8	63.7	59.3	55.8	51.8	45.6	39.3	33.7	29.1	26.5	24.4	22.2	20.7
50	79.5	76.2	71.4	67.5	63.2	56.3	49.3	42.9	37.7	34.8	32.4	29.7	28.0
60	92.0	88.4	83.3	79.1	74.4	67.0	59.3	52.3	46.5	43.2	40.5	37.5	35.5
70	104.2	100.4	95.0	90.5	85.5	77.6	69.3	61.7	55.3	51.7	48.8	45.4	43.3
80	116.3	112.3	106.6	101.9	96.6	88.1	79.3	71.1	64.3	60.4	57.2	53.5	51.2
90	128.3	124.1	118.1	113.1	107.6	98.6	89.3	80.6	73.3	69.1	65.6	61.8	59.2
100	140.2	135.8	129.6	124.3	118.5	109.1	99.3	90.1	82.4	77.9	74.2	70.1	67.3

Source: Catherine M. Thompson, Table of percentage points of the χ^2 distribution, Biometrika, Vol. 32 (1941), b permission of the author and publisher.

Table E.7. Cumulative F Distribution (m Numerator and n Denominator Degrees of Freedom).

α	n \ m	1	2	3	4	5	6	7	8	9	10	12	15	20	30	60	120	∞
.90		39.9	49.5	53.6	55.8	57.2	58.2	58.9	59.4	59.9	60.2	60.7	61.2	61.7	62.3	62.8	63.1	63.3
.95		161	200	216	225	230	234	237	239	241	242	244	246	248	250	252	253	254
.975	1	648	800	864	900	922	937	948	957	963	969	977	985	993	1000	1010	1010	1020
.99		4,050	5,000	5,400	5,620	5,760	5,860	5,930	5,980	6,020	6,060	6,110	6,160	6,210	6,260	6,310	6,340	6,370
.995		16,200	20,000	21,600	22,500	23,100	23,400	23,700	23,900	24,100	24,200	24,400	24,600	24,800	25,000	25,200	25,400	25,500
.90		8.53	9.00	9.16	9.24	9.29	9.33	9.35	9.37	9.38	9.39	9.41	9.42	9.44	9.46	9.47	9.48	9.49
.95		18.5	19.0	19.2	19.2	19.3	19.3	19.4	19.4	19.4	19.4	19.4	19.4	19.5	19.5	19.5	19.5	19.5
.975	2	38.5	39.0	39.2	39.2	39.3	39.3	39.4	39.4	39.4	39.4	39.4	39.4	39.4	39.5	39.5	39.5	39.5
.99		98.5	99.0	99.2	99.2	99.3	99.3	99.4	99.4	99.4	99.4	99.4	99.4	99.4	99.5	99.5	99.5	99.5
.995		199	199	199	199	199	199	199	199	199	199	199	199	199	199	199	199	199
.90		5.54	5.46	5.39	5.34	5.31	5.28	5.27	5.25	5.24	5.23	5.22	5.20	5.18	5.17	5.15	5.14	5.13
.95		10.1	9.55	9.28	9.12	9.01	8.94	8.89	8.85	8.81	8.79	8.74	8.70	8.66	8.62	8.57	8.53	8.53
.975	3	17.4	16.0	15.4	15.1	14.9	14.7	14.6	14.5	14.5	14.4	14.3	14.3	14.2	14.1	14.0	13.9	13.9
.99		34.1	30.8	29.5	28.7	28.2	27.9	27.7	27.5	27.3	27.2	27.1	26.9	26.7	26.5	26.3	26.2	26.1
.995		55.6	49.8	47.5	46.2	45.4	44.8	44.4	44.1	43.9	43.7	43.4	43.1	42.8	42.5	42.1	42.0	41.8
.90		4.54	4.32	4.19	4.11	4.05	4.01	3.98	3.95	3.93	3.92	3.90	3.87	3.84	3.82	3.79	3.78	3.76
.95		7.71	6.94	6.59	6.39	6.26	6.16	6.09	6.04	6.00	5.96	5.91	5.86	5.80	5.75	5.69	5.66	5.63
.975	4	12.2	10.6	9.98	9.60	9.36	9.20	9.07	8.98	8.90	8.84	8.75	8.66	8.56	8.46	8.36	8.31	8.26
.99		21.2	18.0	16.7	16.0	15.5	15.2	15.0	14.8	14.7	14.5	14.4	14.2	14.0	13.8	13.7	13.6	13.5
.995		31.3	26.3	24.3	23.2	22.5	22.0	21.6	21.4	21.1	21.0	20.7	20.4	20.2	19.9	19.6	19.5	19.3
.90		4.06	3.78	3.62	3.52	3.45	3.40	3.37	3.34	3.32	3.30	3.27	3.24	3.21	3.17	3.14	3.12	3.11
.95		6.61	5.79	5.41	5.19	5.05	4.95	4.88	4.82	4.77	4.74	4.68	4.62	4.56	4.50	4.43	4.40	4.37
.975	5	10.0	8.43	7.76	7.39	7.15	6.98	6.85	6.76	6.68	6.62	6.52	6.43	6.33	6.23	6.12	6.07	6.02
.99		16.3	13.3	12.1	11.4	11.0	10.7	10.5	10.3	10.2	10.1	9.89	9.72	9.55	9.38	9.20	9.11	9.02
.995		22.8	18.3	16.5	15.6	14.9	14.5	14.2	14.0	13.8	13.6	13.4	13.1	12.9	12.7	12.4	12.3	12.1
.90		3.78	3.46	3.29	3.18	3.11	3.05	3.01	2.98	2.96	2.94	2.90	2.87	2.84	2.80	2.76	2.74	2.72
.95		5.99	5.14	4.76	4.53	4.39	4.28	4.21	4.15	4.10	4.06	4.00	3.94	3.87	3.81	3.74	3.70	3.67
.975	6	8.81	7.26	6.60	6.23	5.99	5.82	5.70	5.60	5.52	5.46	5.37	5.27	5.17	5.07	4.96	4.90	4.85
.99		13.7	10.9	9.78	9.15	8.75	8.47	8.26	8.10	7.98	7.87	7.72	7.56	7.40	7.23	7.06	6.97	6.88
.995		18.6	14.5	12.9	12.0	11.5	11.1	10.8	10.6	10.4	10.2	10.0	9.81	9.59	9.36	9.12	9.00	8.88
.90		3.59	3.26	3.07	2.96	2.88	2.83	2.78	2.75	2.72	2.70	2.67	2.63	2.59	2.56	2.51	2.49	2.47
.95		5.59	4.74	4.35	4.12	3.97	3.87	3.79	3.73	3.68	3.64	3.57	3.51	3.44	3.38	3.30	3.27	3.23
.975	7	8.07	6.54	5.89	5.52	5.29	5.12	4.99	4.90	4.82	4.76	4.67	4.57	4.47	4.36	4.25	4.20	4.14
.99		12.2	9.55	8.45	7.85	7.46	7.19	6.99	6.84	6.72	6.62	6.47	6.31	6.16	5.99	5.82	5.74	5.65
.995		16.2	12.4	10.9	10.1	9.52	9.16	8.89	8.68	8.51	8.38	8.18	7.97	7.75	7.53	7.31	7.19	7.08
.90		3.46	3.11	2.92	2.81	2.73	2.67	2.62	2.59	2.56	2.54	2.50	2.46	2.42	2.38	2.34	2.31	2.29
.95		5.32	4.46	4.07	3.84	3.69	3.58	3.50	3.44	3.39	3.35	3.28	3.22	3.15	3.08	3.01	2.97	2.93
.975	8	7.57	6.06	5.42	5.05	4.82	4.65	4.53	4.43	4.36	4.30	4.20	4.10	4.00	3.89	3.78	3.73	3.67
.99		11.3	8.65	7.59	7.01	6.63	6.37	6.18	6.03	5.91	5.81	5.67	5.52	5.36	5.20	5.03	4.95	4.86
.995		14.7	11.0	9.60	8.81	8.30	7.95	7.69	7.50	7.34	7.21	7.01	6.81	6.61	6.40	6.18	6.06	5.95

α	n \ m	1	2	3	4	5	6	7	8	9	10	12	15	20	30	60	120	∞
.90		3.36	3.01	2.81	2.69	2.61	2.55	2.51	2.47	2.44	2.42	2.38	2.34	2.30	2.25	2.21	2.18	2.16
.95		5.12	4.26	3.86	3.63	3.48	3.37	3.29	3.23	3.18	3.14	3.07	3.01	2.94	2.86	2.79	2.75	2.71
.975	9	7.21	5.71	5.08	4.72	4.48	4.32	4.20	4.10	4.03	3.96	3.87	3.77	3.67	3.56	3.45	3.39	3.33
.99		10.6	8.02	6.99	6.42	6.06	5.80	5.61	5.47	5.35	5.26	5.11	4.96	4.81	4.65	4.48	4.40	4.31
.995		13.6	10.1	8.72	7.96	7.47	7.13	6.88	6.69	6.54	6.42	6.23	6.03	5.83	5.62	5.41	5.30	5.19
.90		3.29	2.92	2.73	2.61	2.52	2.46	2.41	2.38	2.35	2.32	2.28	2.24	2.20	2.15	2.11	2.08	2.06
.95		4.96	4.10	3.71	3.48	3.33	3.22	3.14	3.07	3.02	2.98	2.91	2.84	2.77	2.70	2.62	2.58	2.54
.975	10	6.94	5.46	4.83	4.47	4.24	4.07	3.95	3.85	3.78	3.72	3.62	3.52	3.42	3.31	3.20	3.14	3.08
.99		10.0	7.56	6.55	5.99	5.64	5.39	5.20	5.06	4.94	4.85	4.71	4.56	4.41	4.25	4.08	4.00	3.91
.995		12.8	9.43	8.08	7.34	6.87	6.54	6.30	6.12	5.97	5.85	5.66	5.47	5.27	5.07	4.86	4.75	4.64
.90		3.18	2.81	2.61	2.48	2.39	2.33	2.28	2.24	2.21	2.19	2.15	2.10	2.06	2.01	1.96	1.93	1.90
.95		4.75	3.89	3.49	3.26	3.11	3.00	2.91	2.85	2.80	2.75	2.69	2.62	2.54	2.47	2.38	2.34	2.30
.975	12	6.55	5.10	4.47	4.12	3.89	3.73	3.61	3.51	3.44	3.37	3.28	3.18	3.07	2.96	2.85	2.79	2.72
.99		9.33	6.93	5.95	5.41	5.06	4.82	4.64	4.50	4.39	4.30	4.16	4.01	3.86	3.70	3.54	3.45	3.36
.995		11.8	8.51	7.23	6.52	6.07	5.76	5.52	5.35	5.20	5.09	4.91	4.72	4.53	4.33	4.12	4.01	3.90
.90		3.07	2.70	2.49	2.36	2.27	2.21	2.16	2.12	2.09	2.06	2.02	1.97	1.92	1.87	1.82	1.79	1.76
.95		4.54	3.68	3.29	3.06	2.90	2.79	2.71	2.64	2.59	2.54	2.48	2.40	2.33	2.25	2.16	2.11	2.07
.975	15	6.20	4.77	4.15	3.80	3.58	3.41	3.29	3.20	3.12	3.06	2.96	2.86	2.76	2.64	2.52	2.46	2.40
.99		8.68	6.36	5.42	4.89	4.56	4.32	4.14	4.00	3.89	3.80	3.67	3.52	3.37	3.21	3.05	2.96	2.87
.995		10.8	7.70	6.48	5.80	5.37	5.07	4.85	4.67	4.54	4.42	4.25	4.07	3.88	3.69	3.48	3.37	3.26
.90		2.97	2.59	2.38	2.25	2.16	2.09	2.04	2.00	1.96	1.94	1.89	1.84	1.79	1.74	1.68	1.64	1.61
.95		4.35	3.49	3.10	2.87	2.71	2.60	2.51	2.45	2.39	2.35	2.28	2.20	2.12	2.04	1.95	1.90	1.84
.975	20	5.87	4.46	3.86	3.51	3.29	3.13	3.01	2.91	2.84	2.77	2.68	2.57	2.46	2.35	2.22	2.16	2.09
.99		8.10	5.85	4.94	4.43	4.10	3.87	3.70	3.56	3.46	3.37	3.23	3.09	2.94	2.78	2.61	2.52	2.42
.995		9.94	6.99	5.82	5.17	4.76	4.47	4.26	4.09	3.96	3.85	3.68	3.50	3.32	3.12	2.92	2.81	2.69
.90		2.88	2.49	2.28	2.14	2.05	1.98	1.93	1.88	1.85	1.82	1.77	1.72	1.67	1.61	1.54	1.50	1.46
.95		4.17	3.32	2.92	2.69	2.53	2.42	2.33	2.27	2.21	2.16	2.09	2.01	1.93	1.84	1.74	1.68	1.62
.975	30	5.57	4.18	3.59	3.25	3.03	2.87	2.75	2.65	2.57	2.51	2.41	2.31	2.20	2.07	1.94	1.87	1.79
.99		7.56	5.39	4.51	4.02	3.70	3.47	3.30	3.17	3.07	2.98	2.84	2.70	2.55	2.39	2.21	2.11	2.01
.995		9.18	6.35	5.24	4.62	4.23	3.95	3.74	3.58	3.45	3.34	3.18	3.01	2.82	2.63	2.42	2.30	2.18
.90		2.79	2.39	2.18	2.04	1.95	1.87	1.82	1.77	1.74	1.71	1.66	1.60	1.54	1.48	1.40	1.35	1.29
.95		4.00	3.15	2.76	2.53	2.37	2.25	2.17	2.10	2.04	1.99	1.92	1.84	1.75	1.65	1.53	1.47	1.39
.975	60	5.29	3.93	3.34	3.01	2.79	2.63	2.51	2.41	2.33	2.27	2.17	2.06	1.94	1.82	1.67	1.58	1.48
.99		7.08	4.98	4.13	3.65	3.34	3.12	2.95	2.82	2.72	2.63	2.50	2.35	2.20	2.03	1.84	1.73	1.60
.995		8.49	5.80	4.73	4.14	3.76	3.49	3.29	3.13	3.01	2.90	2.74	2.57	2.39	2.19	1.96	1.83	1.69
.90		2.75	2.35	2.13	1.99	1.90	1.82	1.77	1.72	1.68	1.65	1.60	1.54	1.48	1.41	1.32	1.26	1.19
.95		3.92	3.07	2.68	2.45	2.29	2.18	2.09	2.02	1.96	1.91	1.83	1.75	1.66	1.55	1.43	1.35	1.25
.975	120	5.15	3.80	3.23	2.89	2.67	2.52	2.39	2.30	2.22	2.16	2.05	1.94	1.82	1.69	1.53	1.43	1.31
.99		6.85	4.79	3.95	3.48	3.17	2.96	2.79	2.66	2.56	2.47	2.34	2.19	2.03	1.86	1.66	1.53	1.38
.995		8.18	5.54	4.50	3.92	3.55	3.28	3.09	2.93	2.81	2.71	2.54	2.37	2.19	1.98	1.75	1.61	1.43
.90		2.71	2.30	2.08	1.94	1.85	1.77	1.72	1.67	1.63	1.60	1.55	1.49	1.42	1.34	1.24	1.17	1.00
.95		3.84	3.00	2.60	2.37	2.21	2.10	2.01	1.94	1.88	1.83	1.75	1.67	1.57	1.46	1.32	1.22	1.00
.975	∞	5.02	3.69	3.12	2.79	2.57	2.41	2.29	2.19	2.11	2.05	1.94	1.83	1.71	1.57	1.39	1.27	1.00
.99		6.63	4.61	3.78	3.32	3.02	2.80	2.64	2.51	2.41	2.32	2.18	2.04	1.88	1.70	1.47	1.32	1.00
.995		7.88	5.30	4.28	3.72	3.35	3.09	2.90	2.74	2.62	2.52	2.36	2.19	2.00	1.79	1.53	1.36	1.00

This table is abridged from "Tables of percentage points of the inverted beta distribution," Biometrika, Vol. 33 (1943).

Table E.8. Cumulative Gamma Distribution.[1]

r	$x^2 = 0.001$ $m = 0.0005$	0.002 0.0010	0.003 0.0015	0.004 0.0020	0.005 0.0025	0.006 0.0030	0.007 0.0035	0.008 0.0040	0.009 0.0045	0.010 0.0050
1	0.97477	0.96433	0.95632	0.94957	0.94363	0.93826	0.93332	0.92873	0.92442	0.92034
2	0.99950	0.99900	0.99850	0.99800	0.99750	0.99700	0.99651	0.99601	0.99551	0.99501
3	0.99999	0.99998	0.99996	0.99993	0.99991	0.99988	0.99984	0.99981	0.99977	0.99973
4							0.99999	0.99999	0.99999	0.99999
r	$x^2 = 0.01$ $m = 0.005$	0.02 0.010	0.03 0.015	0.04 0.020	0.05 0.025	0.06 0.030	0.07 0.035	0.08 0.040	0.09 0.045	0.10 0.050
1	0.92034	0.88784	0.86249	0.84148	0.82306	0.80650	0.79134	0.77730	0.76418	0.75183
2	0.99501	0.99005	0.98511	0.98020	0.97531	0.97045	0.96561	0.96079	0.95600	0.95123
3	0.99973	0.99925	0.99863	0.99790	0.99707	0.99616	0.99518	0.99412	0.99301	0.99184
4	0.99999	0.99995	0.99989	0.99980	0.99969	0.99956	0.99940	0.99922	0.99902	0.99879
5			0.99999	0.99998	0.99997	0.99995	0.99993	0.99991	0.99987	0.99984
6							0.99999	0.99999	0.99999	0.99998
r	$x^2 = 0.1$ $m = 0.05$	0.2 0.10	0.3 0.15	0.4 0.20	0.5 0.25	0.6 0.30	0.7 0.35	0.8 0.40	0.9 0.45	1.0 0.50
1	0.75183	0.65472	0.58388	0.52709	0.47950	0.43858	0.40278	0.37109	0.34278	0.31731
2	0.95123	0.90484	0.86071	0.81873	0.77880	0.74082	0.70469	0.67032	0.63763	0.60653
3	0.99184	0.97759	0.96003	0.94024	0.91889	0.89643	0.87320	0.84947	0.82543	0.80125
4	0.99879	0.99532	0.98981	0.98248	0.97350	0.96306	0.95133	0.93845	0.92456	0.90980
5	0.99984	0.99911	0.99784	0.99533	0.99212	0.98800	0.98297	0.97703	0.97022	0.96257
6	0.99998	0.99985	0.99950	0.99885	0.99784	0.99640	0.99449	0.99207	0.98912	0.98561
7		0.99997	0.99990	0.99974	0.99945	0.99899	0.99834	0.99744	0.99628	0.99483
8			0.99998	0.99994	0.99987	0.99973	0.99953	0.99922	0.99880	0.99825
9				0.99999	0.99997	0.99993	0.99987	0.99978	0.99964	0.99944
10					0.99999	0.99998	0.99997	0.99994	0.99989	0.99983
11							0.99999	0.99998	0.99997	0.99995
12									0.99999	0.99999
r	$x^2 = 1.1$ $m = 0.55$	1.2 0.60	1.3 0.65	1.4 0.70	1.5 0.75	1.6 0.80	1.7 0.85	1.8 0.90	1.9 0.95	2.0 1.00
1	0.29427	0.27332	0.25421	0.23672	0.22067	0.20590	0.19229	0.17971	0.16808	0.15730
2	0.57695	0.54881	0.52205	0.49659	0.47237	0.44933	0.42741	0.40657	0.38674	0.36788
3	0.77707	0.75300	0.72913	0.70553	0.68227	0.65939	0.63693	0.61493	0.59342	0.57241
4	0.89427	0.87810	0.86138	0.84420	0.82664	0.80879	0.79072	0.77248	0.75414	0.73576
5	0.95410	0.94488	0.93493	0.92431	0.91307	0.90125	0.88890	0.87607	0.86280	0.84915
6	0.98154	0.97689	0.97166	0.96586	0.95949	0.95258	0.94512	0.93714	0.92866	0.91970
7	0.99305	0.99093	0.98844	0.98557	0.98231	0.97864	0.97457	0.97008	0.96517	0.95984
8	0.99753	0.99664	0.99555	0.99425	0.99271	0.99092	0.98887	0.98654	0.98393	0.98101
9	0.99917	0.99882	0.99838	0.99782	0.99715	0.99633	0.99537	0.99425	0.99295	0.99147
10	0.99973	0.99961	0.99944	0.99921	0.99894	0.99859	0.99817	0.99766	0.99705	0.99634
11	0.99992	0.99987	0.99981	0.99973	0.99962	0.99948	0.99930	0.99908	0.99882	0.99850
12	0.99998	0.99996	0.99994	0.99991	0.99987	0.99982	0.99975	0.99966	0.99954	0.99941
13	0.99999	0.99999	0.99998	0.99997	0.99996	0.99994	0.99991	0.99988	0.99983	0.99977
14			0.99999	0.99999	0.99999	0.99998	0.99997	0.99996	0.99994	0.99992
15						0.99999	0.99999	0.99999	0.99998	0.99997
16									0.99999	0.99999

Table E.8. (continued)

r	$x^2 = 2.2$ $m = 1.1$	2.4 1.2	2.6 1.3	2.8 1.4	3.0 1.5	3.2 1.6	3.4 1.7	3.6 1.8	3.8 1.9	4.0 2.0
1	0.13801	0.12134	0.10686	0.09426	0.08327	0.07364	0.06520	0.05778	0.05125	0.0455
2	0.33287	0.30119	0.27253	0.24660	0.22313	0.20190	0.18268	0.16530	0.14957	0.1353
3	0.53195	0.49363	0.45749	0.42350	0.39163	0.36181	0.33397	0.30802	0.28389	0.2614
4	0.69903	0.66263	0.62682	0.59183	0.55783	0.52493	0.49325	0.46284	0.43375	0.4060
5	0.82084	0.79167	0.76137	0.73079	0.69999	0.66918	0.63857	0.60831	0.57856	0.5494
6	0.90042	0.87949	0.85711	0.83350	0.80885	0.78336	0.75722	0.73062	0.70372	0.6766
7	0.94795	0.93444	0.91938	0.90287	0.88500	0.86590	0.84570	0.82452	0.80250	0.7797
8	0.97426	0.96623	0.95691	0.94628	0.93436	0.92119	0.90681	0.89129	0.87470	0.8571
9	0.98790	0.98345	0.97807	0.97170	0.96430	0.95583	0.94631	0.93572	0.92408	0.9114
10	0.99457	0.99225	0.98934	0.98575	0.98142	0.97632	0.97039	0.96359	0.95592	0.9473
11	0.99786	0.99652	0.99503	0.99311	0.99073	0.98781	0.98431	0.98019	0.97541	0.9699
12	0.99903	0.99850	0.99777	0.99680	0.99554	0.99398	0.99200	0.98962	0.98678	0.9834
13	0.99961	0.99938	0.99903	0.99856	0.99793	0.99711	0.99606	0.99475	0.99314	0.9911
14	0.99985	0.99975	0.99960	0.99938	0.99907	0.99866	0.99813	0.99743	0.99655	0.9954
15	0.99994	0.99990	0.99984	0.99974	0.99960	0.99940	0.99913	0.99878	0.99832	0.9977
16	0.99998	0.99996	0.99994	0.99989	0.99983	0.99974	0.99961	0.99944	0.99921	0.9989
17	0.99999	0.99999	0.99998	0.99996	0.99993	0.99989	0.99983	0.99975	0.99964	0.9994
18			0.99999	0.99998	0.99997	0.99995	0.99993	0.99989	0.99984	0.9997
19				0.99999	0.99999	0.99998	0.99997	0.99995	0.99993	0.9998
20						0.99999	0.99999	0.99998	0.99997	0.9999
21								0.99999	0.99999	0.9999
22										0.9999
r	$x^2 = 4.2$ $m = 2.1$	4.4 2.2	4.6 2.3	4.8 2.4	5.0 2.5	5.2 2.6	5.4 2.7	5.6 2.8	5.8 2.9	6.0 3.0
1	0.04042	0.03594	0.03197	0.02846	0.02535	0.02259	0.02014	0.01796	0.01603	0.0143
2	0.12246	0.11080	0.10026	0.09072	0.08209	0.07427	0.06721	0.06081	0.05502	0.0497
3	0.24066	0.22139	0.20354	0.18704	0.17180	0.15772	0.14474	0.13278	0.12176	0.1116
4	0.37962	0.35457	0.33085	0.30844	0.28730	0.26739	0.24866	0.23108	0.21459	0.1991
5	0.52099	0.49337	0.46662	0.44077	0.41588	0.39196	0.36904	0.34711	0.32617	0.3062
6	0.64963	0.62271	0.59604	0.56971	0.54381	0.51843	0.49363	0.46945	0.44596	0.4231
7	0.75647	0.73272	0.70864	0.68435	0.65998	0.63557	0.61127	0.58715	0.56329	0.5397
8	0.83864	0.81935	0.79935	0.77872	0.75758	0.73600	0.71409	0.69194	0.66962	0.6472
9	0.89776	0.88317	0.86769	0.85138	0.83431	0.81654	0.79814	0.77919	0.75976	0.7399
10	0.93787	0.92750	0.91625	0.90413	0.89118	0.87742	0.86291	0.84768	0.83178	0.8152
11	0.96370	0.95672	0.94898	0.94046	0.93117	0.92109	0.91026	0.89868	0.88637	0.8733
12	0.97955	0.97509	0.97002	0.96433	0.95798	0.95096	0.94327	0.93489	0.92583	0.9160
13	0.98887	0.98614	0.98298	0.97934	0.97519	0.97052	0.96530	0.95951	0.95313	0.9461
14	0.99414	0.99254	0.99064	0.98841	0.98581	0.98283	0.97943	0.97559	0.97128	0.9664
15	0.99701	0.99610	0.99501	0.99369	0.99213	0.99029	0.98816	0.98571	0.98291	0.9797
16	0.99851	0.99802	0.99741	0.99666	0.99575	0.99467	0.99338	0.99187	0.99012	0.9881
17	0.99928	0.99902	0.99869	0.99828	0.99777	0.99715	0.99639	0.99550	0.99443	0.9931
18	0.99966	0.99953	0.99936	0.99914	0.99886	0.99851	0.99809	0.99757	0.99694	0.9962
19	0.99985	0.99978	0.99969	0.99958	0.99943	0.99924	0.99901	0.99872	0.99836	0.9979
20	0.99993	0.99990	0.99986	0.99980	0.99972	0.99962	0.99950	0.99934	0.99914	0.9989
21	0.99997	0.99995	0.99993	0.99991	0.99987	0.99982	0.99975	0.99967	0.99956	0.9994
22	0.99999	0.99998	0.99997	0.99996	0.99994	0.99991	0.99988	0.99984	0.99978	0.9997
23	0.99999	0.99999	0.99999	0.99998	0.99997	0.99996	0.99994	0.99992	0.99989	0.9998
24			0.99999	0.99999	0.99999	0.99998	0.99997	0.99996	0.99995	0.9999
25					0.99999	0.99999	0.99999	0.99998	0.99998	0.9999
26								0.99999	0.99999	0.9999
27									0.99999	0.9999

Table E.8. (continued)

r	$x^2 = 6.2$ $m = 3.1$	6.4 3.2	6.6 3.3	6.8 3.4	7.0 3.5	7.2 3.6	7.4 3.7	7.6 3.8	7.8 3.9	8.0 4.0
1	0.01278	0.01141	0.01020	0.00912	0.00815	0.00729	0.00652	0.00584	0.00522	0.00468
2	0.04505	0.04076	0.03688	0.03337	0.03020	0.02732	0.02472	0.02237	0.02024	0.01832
3	0.10226	0.09369	0.08580	0.07855	0.07190	0.06579	0.06018	0.05504	0.05033	0.04601
4	0.18470	0.17120	0.15860	0.14684	0.13589	0.12569	0.11620	0.10738	0.09919	0.09158
5	0.28724	0.26922	0.25213	0.23595	0.22064	0.20619	0.19255	0.17970	0.16761	0.15624
6	0.40116	0.37990	0.35943	0.33974	0.32085	0.30275	0.28543	0.26890	0.25313	0.23810
7	0.51660	0.49390	0.47168	0.45000	0.42888	0.40836	0.38845	0.36918	0.35056	0.33259
8	0.62484	0.60252	0.58034	0.55836	0.53663	0.51522	0.49415	0.47349	0.45325	0.43347
9	0.71975	0.69931	0.67869	0.65793	0.63712	0.61631	0.59555	0.57490	0.55442	0.53415
10	0.79819	0.78061	0.76259	0.74418	0.72544	0.70644	0.68722	0.66784	0.64837	0.62884
11	0.85069	0.84539	0.83049	0.81504	0.79908	0.78266	0.76583	0.74862	0.73110	0.71330
12	0.90367	0.89459	0.88288	0.87054	0.85761	0.84412	0.83009	0.81556	0.80056	0.78513
13	0.93657	0.93038	0.92157	0.91216	0.90215	0.89155	0.88038	0.86865	0.85638	0.84360
14	0.96120	0.95538	0.94903	0.94215	0.93471	0.92673	0.91819	0.90911	0.89948	0.88933
15	0.97619	0.97222	0.96782	0.96289	0.95765	0.95186	0.94559	0.93882	0.93155	0.92378
16	0.98579	0.98317	0.98022	0.97693	0.97326	0.96921	0.96476	0.95989	0.95460	0.94887
17	0.99174	0.99007	0.98816	0.98599	0.98355	0.98081	0.97775	0.97437	0.97064	0.96655
18	0.99532	0.99429	0.99309	0.99171	0.99013	0.98833	0.98630	0.98402	0.98147	0.97864
19	0.99741	0.99679	0.99606	0.99521	0.99421	0.99307	0.99176	0.99026	0.98857	0.98667
20	0.99860	0.99824	0.99781	0.99729	0.99669	0.99598	0.99515	0.99420	0.99311	0.99187
21	0.99926	0.99905	0.99880	0.99850	0.99814	0.99771	0.99721	0.99662	0.99594	0.99514
22	0.99962	0.99950	0.99936	0.99919	0.99898	0.99873	0.99843	0.99807	0.99765	0.99716
23	0.99981	0.99974	0.99967	0.99957	0.99945	0.99931	0.99913	0.99892	0.99867	0.99837
24	0.99990	0.99987	0.99983	0.99978	0.99971	0.99963	0.99953	0.99941	0.99926	0.99908
25	0.99995	0.99994	0.99991	0.99989	0.99985	0.99981	0.99975	0.99968	0.99960	0.99949
26	0.99998	0.99997	0.99996	0.99994	0.99992	0.99990	0.99987	0.99983	0.99978	0.99973
27	0.99999	0.99999	0.99998	0.99997	0.99996	0.99995	0.99993	0.99991	0.99989	0.99985
28		0.99999	0.99999	0.99999	0.99998	0.99998	0.99997	0.99996	0.99994	0.99992
29				0.99999	0.99999	0.99999	0.99998	0.99998	0.99997	0.99996
30						0.99999	0.99999	0.99999	0.99999	0.99998

r	$x^2 = 8.2$ $m = 4.1$	8.4 4.2	8.6 4.3	8.8 4.4	9.0 4.5	9.2 4.6	9.4 4.7	9.6 4.8	9.8 4.9	10.0 5.0
1	0.00419	0.00375	0.00336	0.00301	0.00270	0.00242	0.00217	0.00195	0.00175	0.00157
2	0.01657	0.01500	0.01357	0.01228	0.01111	0.01005	0.00910	0.00823	0.00745	0.00674
3	0.04205	0.03843	0.03511	0.03207	0.02929	0.02675	0.02442	0.02229	0.02034	0.01857
4	0.08452	0.07798	0.07191	0.06630	0.06110	0.05629	0.05184	0.04773	0.04394	0.04043
5	0.14555	0.13553	0.12612	0.11731	0.10906	0.10135	0.09413	0.08740	0.08110	0.07524
6	0.22381	0.21024	0.19736	0.18514	0.17358	0.16264	0.15230	0.14254	0.13333	0.12465
7	0.31529	0.29865	0.28266	0.26734	0.25266	0.23861	0.22520	0.21240	0.20019	0.18857
8	0.41418	0.39540	0.37715	0.35945	0.34230	0.32571	0.30968	0.29423	0.27935	0.26503
9	0.51412	0.49439	0.47499	0.45594	0.43727	0.41902	0.40120	0.38383	0.36692	0.35049
10	0.60031	0.58983	0.57044	0.55118	0.53210	0.51323	0.49461	0.47626	0.45821	0.44049
11	0.69528	0.67709	0.65876	0.64035	0.62189	0.60344	0.58502	0.56669	0.54846	0.53039
12	0.76031	0.75314	0.73666	0.71991	0.70293	0.68576	0.66844	0.65101	0.63350	0.61596
13	0.83033	0.81660	0.80244	0.78788	0.77294	0.75768	0.74211	0.72827	0.71020	0.69393
14	0.87865	0.86746	0.85579	0.84365	0.83105	0.81803	0.80461	0.79081	0.77666	0.76218
15	0.91551	0.90675	0.89749	0.88774	0.87752	0.86683	0.85569	0.84412	0.83213	0.81974
16	0.94269	0.93606	0.92897	0.92142	0.91341	0.90495	0.89603	0.88667	0.87686	0.86663
17	0.96208	0.95723	0.95198	0.94633	0.94028	0.93378	0.92687	0.91954	0.91179	0.90361
18	0.97551	0.97207	0.96830	0.96420	0.95974	0.95493	0.94974	0.94418	0.93824	0.93191
19	0.98454	0.98217	0.97955	0.97666	0.97348	0.97001	0.96623	0.96213	0.95771	0.95295
20	0.99046	0.98887	0.98709	0.98511	0.98291	0.98047	0.97779	0.97486	0.97166	0.96817
21	0.99424	0.99320	0.99203	0.99070	0.98921	0.98755	0.98570	0.98365	0.98139	0.97891
22	0.99659	0.99593	0.99518	0.99431	0.99333	0.99222	0.99098	0.98958	0.98803	0.98630
23	0.99802	0.99761	0.99714	0.99659	0.99596	0.99524	0.99442	0.99349	0.99245	0.99128
24	0.99888	0.99863	0.99833	0.99799	0.99760	0.99714	0.99661	0.99601	0.99532	0.99455
25	0.99937	0.99922	0.99905	0.99884	0.99860	0.99831	0.99798	0.99760	0.99716	0.99665
26	0.99966	0.99957	0.99947	0.99934	0.99919	0.99902	0.99882	0.99858	0.99830	0.99798
27	0.99981	0.99977	0.99971	0.99963	0.99955	0.99944	0.99933	0.99917	0.99900	0.99880
28	0.99990	0.99987	0.99984	0.99980	0.99975	0.99969	0.99962	0.99953	0.99942	0.99930
29	0.99995	0.99993	0.99991	0.99989	0.99986	0.99983	0.99979	0.99973	0.99967	0.99960

Table E.8. (continued)

r	$x^2 = 10.5$ $m = 5.25$	11.0 5.5	11.5 5.75	12.0 6.0	12.5 6.25	13.0 6.5	13.5 6.75	14.0 7.0	14.5 7.25	15.0 7.5
1	0.00119	0.00091	0.00070	0.00053	0.00041	0.00031	0.00024	0.00018	0.00014	0.00011
2	0.00525	0.00409	0.00318	0.00248	0.00193	0.00150	0.00117	0.00091	0.00071	0.00055
3	0.01476	0.01173	0.00931	0.00738	0.00585	0.00464	0.00367	0.00291	0.00230	0.00182
4	0.03280	0.02656	0.02148	0.01735	0.01400	0.01128	0.00907	0.00730	0.00586	0.00471
5	0.06225	0.05138	0.04232	0.03479	0.02854	0.02338	0.01912	0.01561	0.01273	0.01036
6	0.10511	0.08838	0.07410	0.06197	0.05170	0.04304	0.03575	0.02964	0.02452	0.02027
7	0.16196	0.13862	0.11825	0.10056	0.08527	0.07211	0.06082	0.05118	0.04297	0.03604
8	0.23167	0.20170	0.17495	0.15120	0.13025	0.11185	0.09577	0.08177	0.06963	0.05913
9	0.31154	0.27571	0.24299	0.21331	0.18657	0.16261	0.14126	0.12233	0.10562	0.09094
10	0.39777	0.35752	0.31991	0.28506	0.25299	0.22367	0.19704	0.17299	0.15138	0.13206
11	0.48605	0.44326	0.40237	0.36364	0.32726	0.29333	0.26190	0.23299	0.20655	0.18250
12	0.57218	0.52892	0.48662	0.45468	0.40640	0.36904	0.33577	0.30071	0.26992	0.24144
13	0.65263	0.61082	0.56901	0.52764	0.48713	0.44781	0.40997	0.37384	0.33960	0.30735
14	0.72479	0.68604	0.64639	0.60630	0.56622	0.52652	0.48759	0.44971	0.41316	0.37815
15	0.78717	0.75259	0.71641	0.67903	0.64086	0.60230	0.56374	0.52553	0.48800	0.45142
16	0.83925	0.80949	0.77762	0.74398	0.70890	0.67276	0.63591	0.59871	0.56152	0.52464
17	0.88135	0.85656	0.82942	0.80014	0.76596	0.73619	0.70212	0.66710	0.63145	0.59548
18	0.91436	0.89436	0.87195	0.84724	0.82038	0.79157	0.76106	0.72909	0.69596	0.66197
19	0.93952	0.92384	0.90587	0.88562	0.86316	0.83857	0.81202	0.78369	0.75380	0.72260
20	0.95817	0.94622	0.93221	0.91608	0.89779	0.87738	0.85492	0.83050	0.80427	0.77641
21	0.97166	0.96279	0.95214	0.93962	0.92513	0.90862	0.89010	0.86960	0.84718	0.82295
22	0.98118	0.97475	0.96686	0.95738	0.94618	0.93316	0.91827	0.90148	0.88279	0.86224
23	0.98773	0.98319	0.97748	0.97047	0.96201	0.95199	0.94030	0.92687	0.91165	0.89463
24	0.99216	0.98901	0.98498	0.97991	0.97367	0.96612	0.95715	0.94665	0.93454	0.92076
25	0.99507	0.99295	0.99015	0.98657	0.98206	0.97650	0.96976	0.96173	0.95230	0.94138
26	0.99696	0.99555	0.99366	0.99117	0.98798	0.98397	0.97902	0.97300	0.96581	0.95733
27	0.99815	0.99724	0.99598	0.99429	0.99208	0.98925	0.98567	0.98125	0.97556	0.96943
28	0.99890	0.99831	0.99749	0.99637	0.99487	0.99290	0.99037	0.98719	0.98324	0.97844
29	0.99935	0.99899	0.99846	0.99773	0.99672	0.99538	0.99363	0.99138	0.98854	0.98502
30	0.99963	0.99940	0.99907	0.99860	0.99794	0.99704	0.99585	0.99428	0.99227	0.98974

r	$x^2 = 15.5$ $m = 7.75$	16.0 8.0	16.5 8.25	17.0 8.5	17.5 8.75	18.0 9.0	18.5 9.25	19.0 9.5	19.5 9.75	20.0 10.0
1	0.00008	0.00006	0.00005	0.00004	0.00003	0.00002	0.00002	0.00001	0.00001	0.00001
2	0.00043	0.00034	0.00026	0.00020	0.00016	0.00012	0.00010	0.00008	0.00006	0.00005
3	0.00144	0.00113	0.00090	0.00071	0.00056	0.00044	0.00035	0.00027	0.00022	0.00017
4	0.00377	0.00302	0.00242	0.00193	0.00154	0.00123	0.00099	0.00079	0.00063	0.00050
5	0.00843	0.00684	0.00555	0.00450	0.00364	0.00295	0.00238	0.00192	0.00155	0.00125
6	0.01670	0.01375	0.01131	0.00928	0.00761	0.00623	0.00510	0.00416	0.00340	0.00277
7	0.03010	0.02512	0.02092	0.01740	0.01444	0.01197	0.00991	0.00819	0.00676	0.00557
8	0.05012	0.04238	0.03576	0.03011	0.02530	0.02123	0.01777	0.01486	0.01240	0.01034
9	0.07809	0.06688	0.05715	0.04872	0.04144	0.03517	0.02980	0.02519	0.02126	0.01791
10	0.11487	0.09963	0.08619	0.07436	0.06401	0.05496	0.04709	0.04026	0.03435	0.02925
11	0.16073	0.14113	0.12356	0.10788	0.09393	0.08158	0.07068	0.06109	0.05269	0.04534
12	0.21522	0.19124	0.16939	0.14960	0.13174	0.11569	0.10133	0.08853	0.07718	0.06709
13	0.27719	0.24913	0.22318	0.19930	0.17744	0.15752	0.13944	0.12310	0.10840	0.09521
14	0.34485	0.31337	0.28380	0.25618	0.23051	0.20678	0.18495	0.16495	0.14671	0.13014
15	0.41604	0.38205	0.34962	0.31886	0.28986	0.26267	0.23729	0.21373	0.19196	0.17193
16	0.48837	0.45296	0.41864	0.38560	0.35398	0.32390	0.29544	0.26866	0.24359	0.22022
17	0.55051	0.52383	0.48871	0.45437	0.42102	0.38884	0.35797	0.32853	0.30060	0.27423
18	0.62740	0.59255	0.55770	0.52311	0.48902	0.45565	0.42320	0.39182	0.36166	0.33282
19	0.69033	0.65728	0.62370	0.58987	0.55603	0.52244	0.48931	0.45684	0.42521	0.39458
20	0.74712	0.71662	0.68516	0.65297	0.62031	0.58741	0.55451	0.52183	0.48957	0.45793
21	0.79705	0.76965	0.74093	0.71111	0.68039	0.64900	0.61718	0.58514	0.55310	0.52126
22	0.83990	0.81589	0.79032	0.76336	0.73519	0.70599	0.67597	0.64533	0.61428	0.58304
23	0.87582	0.85527	0.83304	0.80925	0.78402	0.75749	0.72983	0.70122	0.67185	0.64191
24	0.90527	0.88808	0.86919	0.84866	0.82657	0.80301	0.77810	0.75199	0.72483	0.69678
25	0.92891	0.91483	0.89912	0.88179	0.86287	0.84239	0.82044	0.79712	0.77254	0.74683
26	0.94769	0.93820	0.92341	0.90908	0.89320	0.87577	0.85683	0.83643	0.81464	0.79156
27	0.96182	0.95295	0.94274	0.93112	0.91806	0.90352	0.88750	0.87000	0.85107	0.83076
28	0.97256	0.96582	0.95782	0.94859	0.93805	0.92615	0.91285	0.89817	0.88200	0.86446
29	0.98071	0.97554	0.96939	0.96218	0.95383	0.94427	0.93344	0.92136	0.90779	0.89293

Table E.8. (continued)

ν	$\chi^2 = 21$ $m = 10.5$	22 11.0	23 11.5	24 12.0	25 12.5	26 13.0	27 13.5	28 14.0	29 14.5	30 15.0
1	0.00001									
2	0.00003	0.00002	0.00001	0.00001						
3	0.00011	0.00007	0.00004	0.00003	0.00002	0.00001	0.00001			
4	0.00032	0.00020	0.00013	0.00008	0.00005	0.00003	0.00002	0.00001	0.00001	0.00001
5	0.00081	0.00052	0.00034	0.00022	0.00014	0.00009	0.00006	0.00004	0.00002	0.00002
6	0.00184	0.00121	0.00080	0.00052	0.00034	0.00022	0.00015	0.00009	0.00006	0.00004
7	0.00377	0.00254	0.00171	0.00114	0.00076	0.00050	0.00033	0.00022	0.00015	0.00010
8	0.00715	0.00492	0.00336	0.00229	0.00155	0.00105	0.00071	0.00047	0.00032	0.00021
9	0.01265	0.00888	0.00620	0.00430	0.00297	0.00204	0.00140	0.00095	0.00065	0.00044
10	0.02109	0.01511	0.01075	0.00760	0.00535	0.00374	0.00260	0.00181	0.00125	0.00086
11	0.03337	0.02437	0.01768	0.01273	0.00912	0.00649	0.00460	0.00324	0.00227	0.00159
12	0.05038	0.03752	0.02773	0.02034	0.01482	0.01073	0.00773	0.00553	0.00394	0.00279
13	0.07293	0.05536	0.04168	0.03113	0.02308	0.01700	0.01244	0.00905	0.00655	0.00471
14	0.10163	0.07661	0.06027	0.04582	0.03457	0.02589	0.01925	0.01423	0.01045	0.00763
15	0.13683	0.10780	0.08414	0.06509	0.04994	0.03802	0.02874	0.02157	0.01609	0.01192
16	0.17851	0.14319	0.11374	0.08950	0.06982	0.05403	0.04148	0.03162	0.02394	0.01800
17	0.22629	0.18472	0.14925	0.11944	0.09471	0.07446	0.05807	0.04494	0.03453	0.02635
18	0.27941	0.23199	0.19059	0.15503	0.12492	0.09976	0.07900	0.06206	0.04838	0.03745
19	0.33680	0.28426	0.23734	0.19615	0.16054	0.13019	0.10465	0.08343	0.06599	0.05180
20	0.39713	0.34051	0.28880	0.24239	0.20143	0.16581	0.13528	0.10940	0.08776	0.06985
21	0.45894	0.39951	0.34398	0.29306	0.24716	0.20645	0.17085	0.14015	0.11400	0.09199
22	0.52074	0.45989	0.40173	0.34723	0.29707	0.25168	0.21123	0.17568	0.14486	0.11846
23	0.58109	0.52025	0.46077	0.40381	0.35029	0.30087	0.25597	0.21578	0.18031	0.14940
24	0.63873	0.57927	0.51980	0.46160	0.40576	0.35317	0.30445	0.26004	0.22013	0.18475
25	0.69261	0.63574	0.57756	0.51937	0.46237	0.40760	0.35588	0.30785	0.26392	0.22429
26	0.74196	0.68870	0.63295	0.57597	0.51898	0.46311	0.40933	0.35846	0.31108	0.26761
27	0.78629	0.73738	0.68501	0.63032	0.57446	0.51860	0.46379	0.41097	0.36090	0.31415
28	0.82535	0.78129	0.73304	0.68154	0.62784	0.57305	0.51825	0.46445	0.41253	0.36322
29	0.85915	0.82019	0.77654	0.72893	0.67825	0.62549	0.57171	0.51791	0.46507	0.41400
30	0.88789	0.85404	0.81526	0.77203	0.72503	0.67513	0.62327	0.57044	0.51760	0.46565

ν	$\chi^2 = 31$ $m = 15.5$	32 16.0	33 16.5	34 17.0	35 17.5	36 18.0	37 18.5	38 19.0	39 19.5	40 20.0
5	0.00001	0.00001								
6	0.00003	0.00002	0.00001	0.00001						
7	0.00008	0.00004	0.00003	0.00002	0.00001	0.00001				
8	0.00014	0.00009	0.00006	0.00004	0.00003	0.00002	0.00001	0.00001		
9	0.00030	0.00020	0.00013	0.00009	0.00006	0.00004	0.00003	0.00002	0.00001	0.00001
10	0.00059	0.00040	0.00027	0.00019	0.00012	0.00008	0.00006	0.00004	0.00003	0.00002
11	0.00110	0.00076	0.00053	0.00036	0.00025	0.00017	0.00012	0.00008	0.00005	0.00004
12	0.00197	0.00138	0.00097	0.00068	0.00047	0.00032	0.00022	0.00015	0.00011	0.00007
13	0.00337	0.00240	0.00170	0.00120	0.00085	0.00059	0.00041	0.00029	0.00020	0.00014
14	0.00554	0.00401	0.00288	0.00206	0.00147	0.00104	0.00074	0.00052	0.00036	0.00026
15	0.00878	0.00644	0.00469	0.00341	0.00246	0.00177	0.00127	0.00090	0.00064	0.00045
16	0.01346	0.01000	0.00739	0.00543	0.00397	0.00289	0.00210	0.00151	0.00109	0.00078
17	0.01997	0.01505	0.01127	0.00840	0.00622	0.00459	0.00337	0.00246	0.00179	0.00129
18	0.02879	0.02199	0.01669	0.01260	0.00945	0.00706	0.00524	0.00387	0.00285	0.00209
19	0.04037	0.03125	0.02404	0.01838	0.01397	0.01056	0.00793	0.00593	0.00442	0.00327
20	0.05519	0.04330	0.03374	0.02612	0.02010	0.01538	0.01170	0.00886	0.00667	0.00500
21	0.07366	0.05855	0.04622	0.03624	0.02824	0.02187	0.01683	0.01289	0.00981	0.00744
22	0.09612	0.07740	0.06187	0.04912	0.03875	0.03037	0.02366	0.01832	0.01411	0.01081
23	0.12279	0.10014	0.08107	0.06516	0.05202	0.04125	0.03251	0.02547	0.01984	0.01537
24	0.15378	0.12699	0.10407	0.08467	0.06840	0.05489	0.04376	0.03467	0.02731	0.02139
25	0.18902	0.15801	0.13107	0.10791	0.08820	0.07160	0.05774	0.04626	0.03684	0.02916
26	0.22827	0.19312	0.16210	0.13502	0.11165	0.09167	0.07475	0.06056	0.04875	0.03901
27	0.27114	0.23208	0.19707	0.16605	0.13887	0.11530	0.09507	0.07786	0.06338	0.05124
28	0.31708	0.27451	0.23574	0.20087	0.16987	0.14260	0.11886	0.09840	0.08092	0.06613
29	0.36542	0.31987	0.27774	0.23928	0.20454	0.17356	0.14622	0.12234	0.10166	0.08394
30	0.41541	0.36753	0.32254	0.28083	0.24284	0.20808	0.17714	0.14975	0.12573	0.10486

Table E.8. (continued)

ν	$\chi^2 = 42$ $m = 21$	44 22	46 23	48 24	50 25	52 26	54 27	56 28	58 29	60 30
10	0.00001									
11	0.00002									
12	0.00003	0.00001								
13	0.00006	0.00002	0.00001							
14	0.00012	0.00003	0.00001	0.00001						
15	0.00023	0.00006	0.00003	0.00001	0.00001					
16	0.00040	0.00011	0.00005	0.00003	0.00001	0.00001				
17	0.00067	0.00020	0.00010	0.00005	0.00002	0.00001	0.00001			
18	0.00111	0.00034	0.00017	0.00009	0.00004	0.00002	0.00001	0.00001		
19	0.00177	0.00058	0.00030	0.00015	0.00008	0.00004	0.00002	0.00001	0.00001	
20	0.00277	0.00094	0.00050	0.00026	0.00013	0.00007	0.00003	0.00002	0.00001	0.00001
21	0.00421	0.00151	0.00081	0.00043	0.00022	0.00011	0.00006	0.00003	0.00001	0.00001
22	0.00625	0.00234	0.00128	0.00069	0.00036	0.00019	0.00010	0.00005	0.00003	0.00001
23	0.00908	0.00355	0.00198	0.00109	0.00059	0.00031	0.00016	0.00009	0.00004	0.00002
24	0.01291	0.00526	0.00299	0.00167	0.00092	0.00050	0.00027	0.00014	0.00007	0.00004
25	0.01797	0.00763	0.00443	0.00252	0.00142	0.00078	0.00043	0.00023	0.00012	0.00006
26	0.02455	0.01085	0.00642	0.00373	0.00213	0.00120	0.00066	0.00036	0.00020	0.00011
27	0.03292	0.01512	0.00912	0.00540	0.00314	0.00180	0.00102	0.00056	0.00031	0.00017
28	0.04336	0.02068	0.01272	0.00768	0.00455	0.00265	0.00152	0.00086	0.00048	0.00026
29	0.05616	0.02779	0.01743	0.01072	0.00647	0.00384	0.00224	0.00129	0.00073	0.00041
30	0.07157	0.03670	0.02346	0.01470	0.00903	0.00545	0.00324	0.00189	0.00109	0.00062

ν	$\chi^2 = 62$ $m = 31$	64 32	66 33	68 34	70 35	72 36	74 37	76 38
21	0.00001							
22	0.00001							
23	0.00002	0.00001						
24	0.00003	0.00001	0.00001					
25	0.00006	0.00002	0.00001	0.00001				
26	0.00009	0.00003	0.00002	0.00001	0.00001			
27	0.00014	0.00005	0.00003	0.00001	0.00001	0.00001		
28	0.00023	0.00008	0.00004	0.00002	0.00001	0.00001	0.00001	
29	0.00035	0.00012	0.00007	0.00004	0.00002	0.00001	0.00001	0.00001
30	0.00052	0.00019	0.00011	0.00006	0.00003	0.00002	0.00001	
		0.00029	0.00016	0.00009	0.00005	0.00003		

1. Compiled from E. S. Pearson and H. O. Hartley (eds.) [1954], "Biometrika Tables for Statisticians," Vol. 1, Cambridge University Press, Cambridge, England (by permission). Table yields: $1 - P(y)$, where $y = \chi^2/2\lambda$, for a gamma distribution with parameters $\eta = \nu/2$ and λ. Enter table with $\nu = 2\eta$ and $\chi^2 = 2\lambda y$.

Table E.9. Critical Values for the Kolmogorov-Smirnov Test Statistic.

Sample Size (n)	Significance Level .20	.15	.10	.05	.01
1	.900	.925	.950	.975	.995
2	.684	.726	.776	.842	.929
3	.565	.597	.642	.708	.829
4	.494	.525	.564	.624	.734
5	.446	.474	.510	.563	.669
6	.410	.436	.470	.521	.618
7	.381	.405	.438	.486	.577
8	.358	.381	.411	.457	.543
9	.339	.360	.388	.432	.514
10	.322	.342	.368	.409	.486
11	.307	.326	.352	.391	.468
12	.295	.313	.338	.375	.450
13	.284	.302	.325	.361	.433
14	.274	.292	.314	.349	.418
15	.266	.283	.304	.338	.404
16	.258	.274	.295	.328	.391
17	.250	.266	.286	.318	.380
18	.244	.259	.278	.309	.370
19	.237	.252	.272	.301	.361
20	.231	.246	.264	.294	.352
25	.21	.22	.24	.264	.32
30	.19	.20	.22	.242	.29
35	.18	.19	.21	.23	.27
40				.21	.25
50				.19	.23
60				.17	.21
70				.16	.19
80				.15	.18
90				.14	
100				.14	
Asymptotic Formula:	$\frac{1.07}{\sqrt{n}}$	$\frac{1.14}{\sqrt{n}}$	$\frac{1.22}{\sqrt{n}}$	$\frac{1.36}{\sqrt{n}}$	$\frac{1.63}{\sqrt{n}}$

Journal American Statistical Association 47:425-441, 1952, Z.W. Birnbaum.

Table E.10. Random Digits

31599	49305	11580	88139	43295	19715	75309	72806	68671	93426
00312	19818	21169	08646	85696	56004	58606	15805	94795	13297
57268	04277	09659	98845	72016	93177	68307	57609	16987	16465
56056	90234	76778	33438	13811	10103	82440	16407	29667	00096
27786	45278	81300	10551	09113	59136	82926	00859	53523	81631
44470	63805	95023	25293	46763	96300	61260	43881	44491	58327
08964	79503	70271	82519	57294	84771	54524	10469	52532	82101
01980	06265	90844	19102	32948	51414	34531	52680	43749	61007
98917	14832	13335	55865	28952	55685	27553	39667	68079	85475
82637	42270	37177	65329	64575	08586	33926	82087	43985	32307
88072	35622	27108	92753	15087	04057	96811	73998	81211	88623
22665	73151	46836	49236	08289	91204	95446	48209	47682	94388
13054	50453	12351	57302	72013	00212	20403	91276	44300	98067
25849	03692	03745	52870	22536	53563	18986	05412	03289	12278
17929	95318	25063	87279	91243	77309	23901	26018	74087	22874
91220	68801	07903	77520	36469	11213	44456	84440	85006	90197
76588	57463	15628	38730	79315	05376	10107	53352	63835	76966
04919	62581	35866	46809	24846	90581	13459	89705	14685	34049
31427	04062	11332	19407	08588	60207	22670	23384	51865	05021
69191	00800	84387	10444	96467	38636	38949	43689	73680	61060
12727	93904	95545	44729	28861	52196	71655	35996	54698	77740
72280	43015	91243	04337	91940	07452	33072	32805	88361	64313
51506	35859	34192	43621	74023	04577	02969	90704	77594	40218
31481	43264	78081	91632	35117	34759	18497	58269	77556	92349
92474	14604	81252	29582	27566	01193	65093	75772	80239	79594
20405	36228	79201	54860	20989	05035	39794	25820	43524	88085
76309	27488	85259	19418	00990	11394	38339	03207	26774	48838
81382	66556	04185	45529	79844	43009	63238	09509	78487	52443
01593	65592	74479	90553	70400	97787	36167	51098	91098	81308
99028	81310	42693	12315	12664	19047	49266	89982	70515	12949
99339	20500	21583	19901	08816	25585	57785	73865	20663	36796
57915	75318	02293	60606	62411	04016	23105	16980	10110	61703
71298	69195	68142	24458	69160	18034	11953	55314	73435	07598
03842	24220	67885	89911	95274	10157	79938	73372	62379	43138
79764	81762	66393	35091	65833	30762	69360	00798	16119	20182

Table E.11. Random Standard Normal Deviates

0.335	-1.117	0.379	-2.273	-0.668	1.378	0.559	0.231	1.015	-0.168
0.371	0.006	-0.863	-1.020	-1.213	0.507	-1.115	0.995	0.212	-1.374
-0.114	1.487	-2.508	0.692	-0.797	1.505	-1.091	-0.530	0.337	-0.467
0.182	-1.834	1.226	0.495	-1.531	-0.579	-1.753	-1.250	1.364	0.314
1.213	1.098	-0.099	-1.638	0.375	-0.228	-0.453	1.518	0.421	-1.554
0.212	1.175	-0.523	1.230	-0.761	2.161	0.108	-0.328	-0.590	-0.256
-0.687	0.758	-1.054	-0.374	-1.469	0.910	1.503	-1.050	-0.361	-1.011
2.177	0.295	-0.051	-0.708	0.499	-0.322	2.213	-1.445	-0.560	0.552
0.619	-1.050	2.406	-1.191	0.794	0.089	0.242	0.779	-0.200	-0.981
-0.769	-1.773	1.139	0.224	0.092	-0.642	-0.251	0.629	-0.018	-0.186
0.005	-1.407	0.497	-0.326	-0.277	1.915	0.667	2.454	-0.848	-1.171
-1.036	-0.892	-2.019	1.009	0.294	-0.972	-0.664	-2.436	0.003	0.043
1.780	1.044	-1.129	-1.580	1.315	0.432	0.119	0.863	0.379	1.152
0.405	0.513	1.848	-0.936	0.823	-0.050	0.035	0.467	-1.095	1.189
0.691	0.095	1.620	1.686	-1.994	-0.879	0.377	-0.668	-1.396	2.156
0.020	0.456	-0.451	0.247	0.830	1.824	-0.657	-1.120	-1.471	-0.360
-1.077	-0.022	0.273	0.547	-0.334	0.901	-0.344	0.623	-0.615	-0.553
-1.289	0.385	0.331	-1.149	0.908	-2.213	0.193	-0.213	0.154	-1.779
0.097	-2.637	0.794	0.515	-0.394	0.129	0.609	0.937	0.720	0.070
0.820	1.190	0.880	-1.392	0.084	1.416	-1.040	-0.395	0.496	0.194
0.273	-0.644	0.587	-0.459	-0.428	-1.401	0.328	-0.089	-0.956	-0.494
0.126	-0.805	-0.684	-1.313	0.825	1.722	-0.550	1.190	0.707	-0.740
-0.051	-1.499	0.494	-0.988	-1.250	-1.458	-1.660	-1.380	-0.771	-0.326
-0.157	-1.838	-0.800	-1.306	-0.968	-0.843	-1.087	-0.162	1.380	1.556
1.058	0.779	-0.209	-1.247	-0.240	0.627	0.612	-0.423	0.583	-0.070
0.625	-0.707	-0.472	-0.327	-0.951	0.606	0.033	0.280	-0.459	-0.342
1.974	-0.390	-1.804	-0.972	-1.336	-0.039	-1.453	-1.268	-0.145	0.069
0.563	-0.775	0.735	0.106	-0.349	-1.603	0.930	-1.316	0.657	-0.421
0.504	-0.151	-1.138	-1.464	-1.043	1.087	-1.015	0.473	0.229	-0.679
0.186	1.032	-1.306	-0.088	1.057	1.291	0.166	1.656	1.631	1.768
0.898	-0.210	-0.622	-0.585	-1.425	0.987	0.206	0.626	0.326	0.356
-0.542	0.228	-1.287	0.446	-0.082	0.489	-1.262	1.242	0.804	0.113
0.838	1.166	0.775	0.241	0.890	1.080	-0.076	-1.548	-1.793	0.280
0.784	-1.237	0.540	1.503	0.325	-0.371	1.079	-0.079	0.955	-0.060
1.447	-1.855	-1.484	0.025	1.710	-0.778	-1.707	0.652	-0.632	0.637

Table E.12. Gamma Function

Values of $\Gamma(n) = \int_0^\infty e^{-x} x^{n-1}\,dx$; $\Gamma(n+1) = n\Gamma(n)$

n	Γ(n)	n	Γ(n)	n	Γ(n)	n	Γ(n)
1.00	1.00000	1.25	.90640	1.50	.88623	1.75	.91906
1.01	.99433	1.26	.90440	1.51	.88659	1.76	.92137
1.02	.98884	1.27	.90250	1.52	.88704	1.77	.92376
1.03	.98355	1.28	.90072	1.53	.88757	1.78	.92623
1.04	.97844	1.29	.89904	1.54	.88818	1.79	.92877
1.05	.97350	1.30	.89747	1.55	.88887	1.80	.93138
1.06	.96874	1.31	.89600	1.56	.88964	1.81	.93408
1.07	.96415	1.32	.89464	1.57	.89049	1.82	.93685
1.08	.95973	1.33	.89338	1.58	.89142	1.83	.93969
1.09	.95546	1.34	.89222	1.59	.89243	1.84	.94261
1.10	.95135	1.35	.89115	1.60	.89352	1.85	.94561
1.11	.94739	1.36	.89018	1.61	.89468	1.86	.94869
1.12	.94359	1.37	.88931	1.62	.89592	1.87	.95184
1.13	.93993	1.38	.88854	1.63	.89724	1.88	.95507
1.14	.93642	1.39	.88785	1.64	.89864	1.89	.95838
1.15	.93304	1.40	.88726	1.65	.90012	1.90	.96177
1.16	.92980	1.41	.88676	1.66	.90167	1.91	.96523
1.17	.92670	1.42	.88636	1.67	.90330	1.92	.96878
1.18	.92373	1.43	.88604	1.68	.90500	1.93	.97240
1.19	.92088	1.44	.88580	1.69	.90678	1.94	.97610
1.20	.91817	1.45	.88565	1.70	.90864	1.95	.97988
1.21	.91558	1.46	.88560	1.71	.91057	1.96	.98374
1.22	.91311	1.47	.88563	1.72	.91258	1.97	.98768
1.23	.91075	1.48	.88575	1.73	.91466	1.98	.99171
1.24	.90852	1.49	.88595	1.74	.91683	1.99	.99581
						2.00	1.00000

* For large positive values of x, $\Gamma(x)$ approximates the asymptotic series

$$x^x e^{-x}\sqrt{\frac{2\pi}{x}}\left[1 + \frac{1}{12x} + \frac{1}{288x^2} - \frac{139}{51840x^3} - \frac{571}{2488320x^4} + \cdots\right].$$

Bibliography

Abramowitz, M. and I. A. Stegun. 1964. Handbook of Mathematical Functions. U. S. National Bureau of Standards, Applied Mathematics Series 55, U. S. Government Printing Office, Washington, D. C.

Alexander, G. N. 1954. Some Aspects of Time Series in Hydrology. Journal Institute of Engineers (Australia), September, pp 196.

Anderson, D. V. 1967. Review of Basic Statistical Concepts in Hydrology in Statistical Methods in Hydrology. Proceedings of Hydrology Symposium 5, McGill University, Feb. 24 and 25, 1966, Queen's Printer, Ottawa, Canada.

Anderson, H. W., et al. 1967. Discussions of Multivariate Methods. Volume 2, Proceedings of International Hydrology Symposium, Sept. 6-8, 1967, Fort Collins, Colorado.

Anderson, R. L. 1942. Distribution of the Serial Correlation Coefficient. Annals of Mathematical Statistics 13:1-13.

Anderson, R. L. and T. A. Bancroft. 1952. Statistical Theory in Research. McGraw-Hill Company, New York.

Anderson, T. W. 1958. An Introduction to Multivariate Statistical Analysis. John Wiley and Sons, Inc., New York.

Bagley, J. M. 1964. An Application of Stochastic Process Theory to the Rainfall-Runoff Process. Tech. Report 35, Department of Civil Engineering, Stanford University, Stanford, California.

Bailey, N. T. J. 1964. The Elements of Stochastic Processes. Wiley, New York.

Barger, G. L. and H. C. S. Thom. 1949. Evaluation of Drought Hazard. Agronomy Journal 41(11):519-526.

Barger, G. L., R. H. Shaw and R. F. Dale. 1959. Gamma Distribution Parameters from 2 and 3-week Precipitation Totals in the North Central Region of the United States. Agricultural and Home Economics Experiment Station, Iowa State University, Ames, Iowa.

Beard, L. R. 1962. Statistical Methods in Hydrology. U. S. Army Engineer District, Corps of Engineers, Sacramento, California.

Beard, L. R. 1967. Optimization Techniques for Hydrologic Engineering. Water Resources Research 3:809-815.

Beard, L. R. 1974. Flood Flow Frequency Techniques. Technical Report 119, Center for Research in Water Resources, University of Texas, Austin, Texas.

Beer, C. E. 1966. Class notes for Agricultural Engineering 532. Iowa State University, Ames, Iowa. Ditto material.

Bendat, J. S. and A. G. Piersol. 1966. Measurement and Analysis of Random Data. John Wiley and Sons, New York.

Bendat, J. S. and A. G. Piersol. 1971. Random Data: Analysis and Measurement Procedures. Wiley-Interscience, New York.

Benjamin, J. R. and C. A. Cornell. 1970. Probability, Statistics, and Decision Theory for Civil Engineers. McGraw-Hill, New York.

Benson, M. A. 1950. Use of Historical Data in Flood Frequency Analysis. Transactions American Geophysical Union 31:419-424.

Benson, M. A. 1962a. Plotting Positions and Economics of Engineering Planning. Proceedings of American Society of Civil Engineers 88(HY6) pt. 1, November, 1962.

Benson, M. A. 1962b. Factors Influencing the Occurrence of Floods in a Humid Region of Diverse Terrain. Water Supply Paper 1580-B, U. S. Geological Survey, Washington, D. C.

Benson, M. A. 1962c. Evolution of Methods for Evaluating the Occurrence of Floods. Water Supply Paper 1580-A, U. S. Geological Survey, Washington, D. C.

Benson, M. A. 1964. Factor Affecting the Occurrence of Floods in the Southwest. U. S. Geological Survey Water Supply Paper 1580-D. Washington, D. C.

Benson, M. A. 1965. Spurious Correlation in Hydraulics and Hydrology. Proceedings American Society Civil Engineers, Journal of the Hydraulics Division, HY4, July 1965.

Benson, M. A. 1968. Uniform Flood-Frequency Estimating Methods for Federal Agencies. Water Resources Research 4(5):891-908.

Benson, M. A. and N. C. Matalas. 1967. Synthetic Hydrology Based on Regional Statistical Parameters. Water Resources Research 3(4):931-935.

Beyer, W. H. (ed.). 1968. Handbook of Tables for Probability and Statistics. 2nd editio The Chemical Rubber Company, Cleveland, Ohio.

Blench, T. 1959. Empirical Methods in Spillway Design Floods. Proceedings of H drology Symposium 1, Ottawa. Sponsored by National Research Council of Canad Nov. 4 and 5, 1959.

Bowman, K. O. and L. R. Shenton. 1968. Properties of Estimators for the Gamma Di tribution. Report CTC-1, Union Carbide Corporation, Nuclear Division, Oak Ridg Tennessee.

Bowman, K. O. and L. R. Shenton. 1970. Small Sample Properties of Estimators for th Gamma Distribution. Report CTC-28, Union Carbide Corporation, Nuclear Divisio Oak Ridge, Tennessee.

Brakensiek, D. L. 1958. Fitting Generalized Lognormal Distribution to Hydrologic Dat Transactions American Geophysical Union 39(3):469-473.

Bridges, T. C. and C. T. Haan. 1972. Reliability of Precipitation Probabilities Estimate from the Gamma Distribution. Monthly Weather Review 100(8):607-611.

Brieman, L. 1969. Probability and Stochastic Processes. Houghton Mifflin Compan Boston, Massachusetts.

Burges, S. J. and A. E. Johnson. 1973. Probabilistic Short-Term River Yield Forecast Proceedings American Society of Civil Engineers (IR2):143-155.

California State Department of Public Works. 1923. Flow in California Streams. Bulleti 5. Chapter 5. (Original not seen, cited in Chow (1964)).

Carey, D. I. and C. T. Haan. 1975. Using Parametric Models of Runoff to Impro Parameter Estimates for Stochastic Models. Water Resources Research 11(6):874-87

Carey, D. I. and C. T. Haan. 1976. Supply and Demand in Water Planning: Streamflo Estimation and Conservational Water-Pricing. Water Resources Institute Report 92, Un versity of Kentucky, Lexington.

Chayes, F. 1949. On Ratio Correlation in Petrography. Journal of Geology 57(3):23 254.

Chow, V. T. 1951. A Generalized Formula for Hydrologic Frequency Analysis. Tran actions American Geophysical Union 32(2):231-237.

Chow, V. T. 1954. The Log-Probability Law and Its Engineering Applications. Proce ings American Society of Civil Engineers 80:536-1 to 536-25.

Chow, V. T. (ed.). 1964. Handbook of Applied Hydrology. McGraw-Hill, New Yor

Cooley, W. W. and P. R. Lohnes. 1971. Multivariate Data Analysis. John Wiley and Son

Inc., New York.

Crutcher, H. L. 1975. A Note on the Possible Misuse of the Kolmogorov-Smirnov Test. Journal of Applied Meteorology 14:1600-1603.

Dalrymple, T. 1960. Flood Frequency Analysis. U. S. Geological Survey Water Supply Paper 1543-A, Manual of Hydrology. Part 3, Flood Flow Techniques. U. S. Government Printing Office, Washington, D. C.

Davis, J. C. 1973. Statistics and Data Analysis in Geology. John Wiley and Sons, Inc., New York.

Dawdy, D. R. and J. H. Feth. 1967. Applications of Factor Analysis in Study of Chemistry of Groundwater Quality, Mojare River Valley, California. Water Resources Research 3(2):505-510.

DeCoursey, D. G. 1971. The Stochastic Approach to Watershed Modeling. Nordic Hydrolgy pp 186-216.

DeCoursey, D. G. 1973. Objective Regionalization of Peak Flow Rates. Floods and Droughts. Proceedings of the Second International Symposium in Hydrology. Fort Collins, Colorado. Sept. 11-13, 1972. Water Resources Publications, Fort Collins.

DeCoursey, D. G. and R. B. Deal. 1974. General Aspects of Multivariate Analysis with Application to some Problems in Hydrology. Proceedings Symposium on Statistical Hydrology. U. S. Department of Agriculture Miscellaneous Publication No. 1275, Washington, D. C. pp 47-68.

DeVries, R. N. and D. L. Sveum. 1970. A Streamflow Simulation Technique. Paper presented at 51st Annual Meeting of American Geophysical Union, April 20-24, 1970. Washington, D. C.

Draper, N. R. and H. Smith. 1966. Applied Regression Analysis. John Wiley and Sons, New York.

Duckstein, L., M. Fogel and D. Davis. 1975. Mountainous Winter Precipitation: A Stochastic Event-based Approach. Proceedings National Symposium on Precipitation Analysis for Hydrologic Modeling sponsored by the American Geophysical Union. June 26-28, 1975, Davis, California. pp 172-188.

Durant, E. F. and S. R. Blackwell. 1959. The Magnitude and Frequency of Floods on the Canadian Prairies in Spillway Design Floods, Proceedings of Hydrology Symposium 1, Ottawa, sponsored by National Research Council of Canada, Nov. 4 and 5, 1959.

Eagleson, P. S. 1970. Dynamic Hydrology. McGraw-Hill, New York.

Eiselstein, L. M. 1967. A Principal Component Analysis of Surface Runoff Data from a New Zealand Alpine Watershed. Volume 1 Proceedings of International Hydrology Symposium. Sept. 6-8, 1967. Fort Collins, Colorado.

Elderton, W. P. 1953. Frequency Curves and Correlation. Fourth Edition, Harren Pr Washington, D. C.

Federal Council for Science and Technology. Ad Hoc Panel on Hydrology. 1962. Sci tific Hydrology. Washington, D. C.

Feller, W. 1957. An Introduction to Probability Theory and its Application, Volum John Wiley and Sons, New York.

Fiering, M. B. 1961. Queueing Theory and Simulation in Reservoir Design. Proceedi American Society of Civil Engineers 87(HY6):39-69.

Fiering, M. B. 1964. Multivariate Technique for Synthetic Hydrology. Proceedi American Society of Civil Engineers 90(HY5):43-60.

Fiering, M. B. 1966. Synthetic Hydrology: An Assessment, in Water Research edited A. V. Kneese and S. C. Smith, Resources for the Future, Inc., Washington, D. C. pp. 3 341.

Fiering, M. B. 1967. Streamflow Synthesis. Harvard University Press, Cambric Massachusetts.

Fiering, M. B. and B. B. Jackson. 1971. Synthetic Streamflows. Water Resources Mo graph 1, American Geophysical Union, Washington, D. C. 98 pp.

Fogel, M. M., L. Duckstein and J. L. Sanders. 1974. An Event-based Stochastic Model Areal Rainfall and Runoff. Proceedings Symposium on Statistical Hydrology. U. S. I partment of Agriculture Miscellaneous Publication No. 1275, Washington, D. C. pp 2 261.

Freund, J. E. 1962. Mathematical Statistics. Prentice Hall, Inc., Englewood Cliffs, N

Friedman, D. G. and B. F. Janes. 1957. Estimation of Rainfall Probabilities. Connecti Agricultural Experiment Station Bulletin 332, University of Connecticut, Storrs, C necticut.

Gabriel, K. R. and J. Neumann. 1962. A Markov Chain Model for Daily Rainfall Occ rence at Tel Aviv. Quarterly Journal of the Royal Meteorological Society 88(375): 95.

Gibra, I. N. 1973. Probability and Statistical Inference for Scientists and Enginee Prentice-Hall, Englewood Cliffs, New Jersey.

Grace, R. A. and P. S. Eagleson. 1966. The Synthesis of Short-Time Increment Rainf Sequences. Report 91, Hydrodynamics Laboratory, Department of Civil Engineeri Massachusetts Institute of Technology, Cambridge, Massachusetts.

Graybill, F. A. 1961. An Introduction to Linear Statistical Models. McGraw-Hi New York.

Greenwood, J. A. and D. Durand. 1960. Aids for Fitting the Gamma Distribution by Maximum Likelihood. Technometrics 2(1):55-65.

Gumbel, E. J. 1954. Statistical Theory of Extreme Values and Some Practical Applications. U. S. National Bureau of Standards Applied Mathematics Series, No. 33, Washington, D. C.

Gumbel, E. J. 1958. Statistics of Extreme Values. Columbia University Press. New York.

Guttman, I. and S. S. Wilks. 1965. Introductory Engineering Statistics. John Wiley and Sons, New York.

Haan, C. T. 1965. Point of Impending Sediment Deposition for Open Channel Flow in a Circular Conduit. M.S. Thesis, Purdue University, West Lafayette, Indiana.

Haan, C. T. 1972a. Adequacy of Hydrologic Records for Parameter Estimation. Journal of Hydraulics Division, American Society of Civil Engineers 98(HY8):1387-1393.

Haan, C. T. 1972b. Evaluation of Prediction Error Probabilities from Stochastic Models by Simulation. Proceedings of International Symposium on Uncertainties in Hydrologic and Water Resources Systems. Dec. 11-14, 1972. Tucson, Arizona.

Haan, C. T. and D. M. Allen. 1972. Comparison of Multiple Regression and Principal Component Regression for Predicting Water Yields in Kentucky. Water Resources Research 8(6):1593-1596.

Haan, C. T., D. M. Allen and J. O. Street. 1976. A Markov Chain Model of Daily Rainfall. Water Resources Research 12(3):443-449.

Haan, C. T. and C. E. Beer. 1968. Determining the Maximum Likelihood Estimators for the Three Parameter Weibull Distribution. Iowa State Journal of Science 42(1):37-42.

Haan, C. T. and H. P. Johnson. 1967. Geometrical Properties of Depressions in North-Central Iowa. Iowa State Journal of Science 42(2):149-160.

Haan, C. T. and H. R. Read. 1970. Prediction of Monthly, Seasonal and Annual Runoff Volumes for Small Agricultural Watersheds in Kentucky. Bulletin 711, Kentucky Agricultural Experiment Station, University of Kentucky, Lexington, Kentucky.

Hadley, G. 1961. Linear Algebra. Addison-Wesley, Reading, Massachusetts.

Hahn, G. J. and S. S. Shapiro. 1967. Statistical Models in Engineering. John Wiley and Sons, New York.

Hall, W. A. and J. A. Dracup. 1970. Water Resources Systems Engineering. McGraw-Hill, New York.

Harman, H. H. 1967. Modern Factor Analysis, University of Chicago Press, Chicago, Illinois.

Harter, H. C. 1964. New Tables of the Incomplete Gamma-Function Ratio and of P centage Points of the Chi-Square and Beta Distributions. Aerospace Research Labo tories, U. S. Air Force.

Hawkins, R. H. 1974. A Note on Mixed Distributions in Hydrology. Proceedings Symp sium on Statistical Hydrology. U. S. Department of Agriculture Miscellaneous Publi tion No. 1275, Washington, D. C. pp 336-344.

Hazen, A. 1930. Flood Flows, A Study of Frequencies and Magnitudes. John Wiley a Sons, Inc. New York.

Hershfield, D. M. 1961. Rainfall Frequency Atlas of the United States. U. S. Weath Bureau Technical Paper 40. U. S. Department of Commerce, Washington, D. C.

Horton, R. E. 1933. The Role of Infiltration in the Hydrologic Cycle. Transactions the American Geophysical Union 14:446-460.

Hudlow, M. D. 1967. Streamflow Forecasting Based on Statistical Applications a Measurements Made with Rain Gage and Weather Radar. Tech. Report 7, Water R sources Institute, Texas A&M University, College Station, Texas.

Huff, F. A. and S. A. Changnon. 1973. Precipitation Modification by Major Urba Areas. Bulletin American Meteorological Society 54(12):1220-1232.

Huntsberger, D. V. 1967. Elements of Statistical Inference. Second Edition. Allyn a Bacon, Boston, Massachusetts.

Hurst, H. E. 1951. Long-term Storage Capacity of Reservoirs. Transactions Americ Society of Civil Engineers, 116:776-808.

Hurst, H. E. 1956. Methods of Using Long-term Storage in Reservoirs. Proceedings I stitute of Civil Engineers 5(5):519-590.

Inter-Agency Committee on Water Resources. 1966. Methods of Flow Frequency Anal ysis. Bulletin 13, Subcommittee on Hydrology, Inter-Agency Committee on Water R sources, U. S. Government Printing Office, Washington, D. C.

Jennings, M. E. and M. A. Benson. 1969. Frequency Curves for Annual Series with Som Zero Events or Incomplete Data. Water Resources Research 5(1):276-280.

Johnston, J. 1963. Econometric Methods. McGraw-Hill, New York.

Julian, P. R. 1967. Variance Spectrum Analysis.Water Resources Research 3:831-845.

Julian, R. W., V. Yevjevich, and H. J. Morel-Seytoux. 1967. Prediction of Water Yield i High Mountain Watersheds Based on Physiography. Hydrology Paper 22, Colorado Stat University, Fort Collins, Colorado.

Kaiser, H. F. 1958. The Varimax Criterion for Analytic Rotation in Factor Analysis.

Psychometrika 23:187-200.

Kendall, G. R. 1967. Probability Distribution of a Single Variable in Statistical Methods in Hydrology, Proceedings of Hydrology Symposium 5, McGill University, Feb. 24 and 25, 1966, Queen's Printer, Ottawa, Canada.

Kendall, G. R. 1959. Statistical Analysis of Extreme Values in Spillway Design Floods, Proceedings of Hydrology Symposium 1, Ottawa. Sponsored by National Research Council of Canada, Nov. 4 and 5, 1959.

Kisiel, C. C. 1969. Time Series Analysis of Hydrologic Data. Advances in Hydroscience, Vol. 5, Academic Press, New York, pp:1-119.

Lane, L. J. and H. B. Osborn. 1973. Hypotheses on the Seasonal Distribution of Thunderstorm Rainfall in Southeastern Arizona in Floods and Droughts. Proceedings of Second International Symposium on Hydrology, September 11-13, 1972. Fort Collins, Colorado.

Langbein, W. B. 1949. Annual Floods and the Partial Duration Flood Series. Transactions American Geophysical Union 30:879-881.

Law, A. G. 1974. Stochastic Analysis of Groundwater Level Time Series in the Western United States. Hydrology Paper, Colorado State University, Fort Collins, Colorado.

Lieberman, G. J. and D. B. Owen. 1961. Tables of the Hypergeometric Probability Distribution. Stanford University Press, Stanford, California.

Lindgren, B. W. 1968. Statistical Theory. Collier-MacMillan, Ltd., Toronto, Ontario, Canada.

Linsley, R. K., M. A. Kohler and J. C. H. Paulus. 1958. Hydrology for Engineers. McGraw-Hill, New York.

Lloyd, E. H. 1967. Stochastic Reservoir Theory in Advances in Hydroscience, Volume 4:281-335. Academic Press, New York.

Lowery, M. D. and J. E. Nash. 1970. A Comparison of Methods of Fitting the Double Exponential Distribution, Journal of Hydrology 10(3):259-275.

Mandelbrot, B. B. and J. W. Van Ness. 1968. Fractional Brownian Motions, Fractional Noises, and Applications. SIAM Rev. 10(4), 422-437.

Mandlebrot, B. B. and J. R. Wallis. 1968. Noah, Joseph, and Operational Hydrology. Water Resources Research 4(5):909-918.

Mandlebrot, B. B. and J. R. Wallis. 1969. Computer Experiments with Fractional Gaussian Noises. Water Resources Research 5(1):228-267.

Markovic, R. D. 1965. Probability Functions of Best Fit to Distributions of Annual Precipitation and Runoff. Hydrology Paper 8, Colorado State University, Fort Collins, Colorado.

Matalas, N. C. 1966. Some Aspects of Time Series Analysis in Hydrologic Studies. Proceedings Hydrology Symposium 5, Statistical Methods in Hydrology. McGill University, February 24 and 25, 1966. Queen's Printer, Ottawa, Canada.

Matalas, N. C. 1967. Mathematical Assessment of Synthetic Hydrology. Water Resources Research 3(4):937-945.

Matalas, N. C. 1967b. Time Series Analysis. Water Resources Research 3:817-8

Matalas, N. C. and W. B. Langbein. 1962. Information Content of the Mean. Journal Geophysical Research 67(9):3441-3448.

Matalas, N. C. and B. J. Reiher. 1967. Some Comments on the Use of Factor Analysis Water Resources Research 3(1):213-224.

Matalas, N. C. and J. R. Rosenblatt. 1962. Use of Correlation for Augmenting Autocorrelated Streamflow Information. U. S. Geological Survey Water Supply Paper (Original not seen, cited by Matalas and Langbein, 1962.)

McCabe, J. A. 1962. Floods in Kentucky – Magnitude and Frequency: Information Circular 9, Kentucky Geological Survey, Lexington, Kentucky.

McMahon, T. A., G. P. Lodner and C. S. Joy. 1972. Reservoir Storage Yield Estimates Based on Historical and Generated Streamflows. Civil Engineers Transactions, Institution of Engineers, Australia, CE 14(2):147-152.

Meinzer, O. E. (ed.). 1942. Hydrology. Dover Publications Inc., New York.

Mejia, J. M., I. Rodriguez-Iturbe and D. R. Dawdy. 1972. Streamflow Simulation 2. The Broken Line Process as a Potential Model for Hydrologic Simulation. Water Resources Research 8(4):931-941.

Mood, A. M. and F. A. Graybill. 1963. Introduction to the Theory of Statistics. Second Edition, McGraw-Hill Book Co., Inc., New York.

Mood, A. M., F. A. Graybill and D. C. Boes. 1974. Introduction to the Theory of Statistics. Third Edition, McGraw-Hill, New York.

Mooley, D. A. and H. L. Crutcher. 1968. An Application of the Gamma Distribution Function to Indian Rainfall. Environmental Sciences Services Administration Technical Report EDS-5, U.S. Department of Commerce, Silver Springs, Maryland.

Morrison, D. F. 1967. Multivariate Statistical Methods. McGraw-Hill, New York.

National Bureau of Standards. 1953. Probability Tables for the Analysis of Extreme Value Data. Applied Mathematics Series 22. Washington, D. C.

Neuts, M. F. 1973. Probability. Allyn and Bacon, Inc., Boston.

Ostle, B. 1963. Statistics in Research. Second Edition, Iowa State University Press, Ames, Iowa.

Overton, D. C. 1969. Eigenvector Analysis of Geomorphic Interrelations of Small Agricultural Watersheds. Transactions American Society of Agricultural Engineers 12(4): 506-508.

Owen, D. B. 1962. Handbook of Statistical Tables. Addison Wesley, Reading, Massachusetts.

Panofsky, H. A. and G. W. Brier. 1958. Some Applications of Statistics to Meteorology. The Pennsylvania State University, University Park, Pennsylvania.

Parl, B. 1967. Basic Statistics. Doubleday and Company, Inc., Garden City, New York.

Pattison, A. 1964. Synthesis of Rainfall Data. Technical Report 40, Department of Civil Engineering, Stanford University, Stanford, California.

Pearson, K. (1896-1897). On the Form of Spurious Correlation which may arise when Indices are Used in the Measurement of Organs. Royal Society of London, Proceedings 60:489-502. (Original not seen, cited by Chayes, 1949, and others.)

Pearson, K. 1930. Tables for Statisticians and Biometricians, Part I, The Biometric Laboratory, University College, London; Cambridge University Press. (Original not seen, cited in Chow (c1964).)

Pearson, K. 1957. Tables of the Incomplete Γ-Function. Biometriks Office, University College, London.

Potter, W. D. 1949. Simplification of the Gumbel Method for Computing Probability Curves. SCS-TP-78. U. S. Department of Agriculture, Soil Conservation Service, Washington, D. C., May 1949.

Press, S. J. 1972. Applied Multivariate Analysis. Holt-Rinehart and Winston, Inc., New York.

Ralston, A. 1965. A First Course in Numerical Analysis. McGraw-Hill, New York.

Rand Corporation. 1965. A Million Random Digits with 100,000 Normal Deviates. The Free Press. Glencoe, Illinois.

Reed, L. J. 1921. On the Correlation Between any Two Functions and Its Application o the General Case of Spurious Correlation. Washington Academy of Science Journal Vol. 11:449-455.

Reich, B. M. 1962. Design Hydrographs for Very Small Watersheds from Rainfall. Civil Engineering Section, Colorado State University. 57 pp.

Rice, R. M. 1967. Multivariate Methods Useful in Hydrology. Volume 1. Proceedings International Hydrology Symposium. Sept. 6-8, 1967. Fort Collins, Colorado.

Rodriguez-Iturbe, I., D. R. Dawdy and L. Garcia. 1971. Adequacy of Markovian Mod with Cyclic Components for Stochastic Streamflow Simulation. Water Resources I search 7(5):1127-1143.

Rodriguez-Iturbe, I., J. M. Mejia and D. R. Dawdy. 1972. Streamflow Simulation 1. New Look at Markovian Models, Fractional Guassian Noise, and Crossing Theory. Wa Resources Research 8(4):921-930.

Rohlf, F. J. and R. R. Sokal. 1969. Statistical Tables. W. H. Freeman and Co., S Francisco.

Salas-LaCruz, J. O. and V. Yevjevich. 1972. Stochastic Structure of Water Use Tir Series, Hydrology Paper 52, Colorado State University, Fort Collins, Colorado.

Sangal, B. P. and A. K. Biswas. 1970. The 3-Parameter Lognormal Distribution and I Applications in Hydrology. Water Resources Research 6(2):505-515.

Selby, S. M. (ed.). 1970. Handbook of Tables for Mathematics. The Chemical Rubb Publishing Company, Cleveland, Ohio.

Shelton, C. H. and J. I. Sewell. 1969. Parameter Screening for Watershed Analysi Transactions American Society of Agricultural Engineers 12(4):533-539.

Shenton, L. R. and K. O. Bowman. 1970. Remarks on Thom's Estimators for the Gan ma Distribution. Monthly Weather Review 98(2):154-160.

Shreve, R. L. 1970. Mainstream Length versus Basin Area in Topologically Randor Channel Networks. Geological Society of America, Abstracts with Programs. Vol. 2 p. 682. (Original not seen, cited by Smart (1973).)

Singh, K. P. 1974. A Two-Distribution Method for Fitting Mixed Distributions in Hy drology. Proceedings Symposium on Statistical Hydrology. U. S. Department of Agricul ture Miscellaneous Publication No. 1275, Washington, D. C. pp 371-382.

Smart, J. S. 1973. The Random Model in Fluvial Geomorphology. Proceedings Sympo sium on Fluvial Geomorphology. State University of New York, Binghampton, Sept. 28-29.

Synder, W. M. 1962. Some Possibilities for Multivariate Analysis in Hydrologic Studies. Journal of Geophysical Research 67(2):721-729.

Solomon, S. 1966. Statistical Association Between Hydrologic Variables. Proceedings Hydrology Symposium 5, Statistical Methods in Hydrology. McGill University, February 24 and 25, 1966. Queen's Printer, Ottawa, Canada.

Spiegel, M. R. 1961. Theory and Problems of Statistics, Schaum Publishing Company.

New York.

Stammers, W. M. 1967. The Application of Multivariate Techniques in Hydrology in Statistical Methods in Hydrology. Proceedings of Hydrology Symposium 5, McGill University, February 24 and 25, 1966. Queen's Printer, Ottawa, Canada. (Also discussions of this paper which immediately follow it.)

Steel, R. G. D. and J. H. Torrie. 1960. Principles and Procedures of Statistics. McGraw-Hill Book Co., New York.

Stidd, C. K. 1953. Cube Root Normal Precipitation Distributions. Transactions American Geophysical Union 34(1):31-35.

Strommen, N. O. and J. E. Horsfield. 1969. Monthly Precipitation Probabilities by Climatic Divisions. Miscellaneous Publication 1160, U. S. Department of Agriculture and U. S. Department of Commerce, Washington, D. C.

Sturges, H. A. 1926. The Choice of a Class Interval. Journal American Statistical Association 21:65-66.

Sudler, C. E. 1927. Storage Required for Regulations of Streamflow. Transactions American Society of Civil Engineers 91:622-660.

Thomas, H. A. and M. B. Fiering. 1962. Mathematical Synthesis of Streamflow Sequences for the Analysis of River Basins by Simulation, Chapter 12 of Design of Water Resources Systems. Harvard University Press, Cambridge, Massachusetts.

Thomas, H. A. and M. B. Fiering. 1963. The Nature of the Storage Yield Function, in Operations Research in Water Quality Management, Chapter 2. Harvard University Water Program, Cambridge, Massachusetts.

Thomas, J. B. 1971. An Introduction to Applied Probability and Random Processes. John Wiley and Sons, Inc., New York.

Thom, H. C. S. 1947. A Note on the Gamma Distribution, Statistical Laboratory, Iowa State College (Manuscript). (Original not seen, cited in Strommen and Horsfield, 1969.)

Thom, H. C. S. 1958. A Note on the Gamma Distribution. Monthly Weather Review 86(4):117-122.

Van Montfort, M. A. J. 1970. On Testing that the Distribution of Extremes is of Type I when Type II is the Alternative. Journal of Hydrology 11(4):421-427.

Viessman, W., T. E. Harbaugh and J. W. Knapp. 1972. Introduction to Hydrology. Intext Educational Publishers, New York.

Wallis, J. R. 1965. Multivariate Statistical Methods in Hydrology – A Comparison Using Data of Known Functional Relationship. Water Resources Research 1(4):447-462.

Wallis, J. R. 1967. When is it Safe to Extend a Prediction Equation? – An Answer Bas on Factor and Discriminant Function Analysis. Water Resources Research 3(2):375-38

Wallis, J. R. 1968. Factor Analysis in Hydrology – An Agnostic View. Water Resourc Research 4(3):521-528.

Wallis, J. R. and N. C. Matalas. 1970. Small Sample Properties of H and K – Estimato of the Hurst Coefficient h. Water Resources Research 6(6):1583-1594.

Wang, Leei-Luoh and A. C. Huber. 1967. Estimating Water Yields in Utah by Princip Components Analysis. Report PrWg 35a-1, Utah Water Research Laboratory, Logai Utah.

Water Resources Council. 1967. A Uniform Technique for Determining Flood Flo Frequencies. Bulletin 15. Water Resources Council, Washington, D. C.

Weast, R. C. (ed.). 1967. Handbook of Tables for Mathematics. The Chemical Rubbe Co., Cleveland, Ohio.

Weibull, W. 1939. "A Statistical Study of the Strength of Materials". Ing. Vetenskap Akad. Handl. (Stockholm) Vol. 151, pp 15. (Original not seen, cited in Chow (1964).

Whittaker, J. 1973. A Note on the Generation of Gamma Random Variables with Nor Integral Shape Parameter. Proceedings of the Second International Symposium on Hy drology. Water Resources Publications, Fort Collins, Colorado. pp 591-594.

Wong, Shue Tuch. 1963. A Multivariate Statistical Model for Predicting Mean Annua Flood in New England. Annals of Association of American Geographers. 53(3):298-311

Woolhiser, D. A., E. Rovey and P. Todorovic. 1973. Temporal and Spatial Variation of Parameters for the Distribution of N-Day Precipitation. Proceedings of the Second Inter national Symposium on Hydrology. Water Resources Publications, Fort Collins, Colorado. pp 605-614.

Yevdjevich, V. M. 1963. Fluctuations of Wet and Dry Years, Part I. Research Data Assembly and Mathematical Models. Hydrology Paper 1, Colorado State University, Fort Collins, Colorado.

Yevjevich, V. M. 1972. Structural Analysis of Hydrologic Time Series. Hydrology Paper 56, Colorado State University, Fort Collins, Colorado.

Yevjevich, V. M. 1972a. Probability and Statistics in Hydrology. Water Resources Publications, Fort Collins, Colorado.

Yevjevich, V. M. 1972b. Stochastic Processes in Hydrology. Water Resources Publications, Fort Collins, Colorado.

Young, G. K. 1968. Discussion of Mathematical Assessment of Synthetic Hydrology by N. C. Matalas. Water Resources Research 4(3):681-682.

CURSO INGENIERIA MARITIMA.
MODULO: "FUNDAMENTOS DE HIDRAULICA Y MATEMATICAS".
DEL 11 DE MARZO AL 7 DE JUNIO
MEXICO, D.F.

AIMS AND MEANS IN TIME SERIES ANALYSIS.

PROF. ING. JOSE RAYNAL.
JUNIO DE 1985.

1

Aims and Means in Time Series Analysis

To orient and motivate the reader, this chapter gives a summary of the most useful concepts and main objectives of time series analysis.

1.1 TIME SERIES AND STOCHASTIC PROCESSES

1.1.1 Deterministic and non-deterministic functions

It is difficult to find any branch of science which does not lead to the study of data arising in the form of *time series*. A time series is a *random* or *non-deterministic* function x of an independent variable t. In most situations the function $x(t)$ will be a function of time, but in other situations it may be a function of some other physical parameter t, for example, space.

The characteristic feature of a time series is that its future behavior cannot be predicted exactly, as would be the case for a *deterministic* function of time. In many branches of applied mathematics it is convenient to assume that certain physical processes can be described by deterministic or mathematical functions of time. For example, in the majority of electrical engineering calculations it is convenient to assume that the most important features of the supply voltage can be represented by the cosine function

$$x(t) = a \cos (2\pi f_0 t + \phi), \tag{1.1.1}$$

where f_0 is the supply frequency and a is the voltage amplitude. However, closer inspection of the amplitude reveals that it is not constant but fluctuates with time. Thus Figure 1.1 shows the voltage deviations at the terminals of a turbo-alternator as a function of time. When two records of the voltage–time curve are compared, they may not resemble each other visually. However, when their *statistical* or *average* behavior is compared, similarities begin to emerge. This observation leads to the notion of a stochastic process.

1.1.2 Stochastic processes

Since different sections of a time series resemble each other only in their average properties, it is necessary to describe these series by probability laws or models. Thus, possible values of the time series at a given time t are assumed to be described by a *random variable* $X(t)$ and its associated probability distribution. The observed value $x(t)$ of the time series at time t is then regarded as one of the infinity of values which the random variable $X(t)$ might have taken at time t.

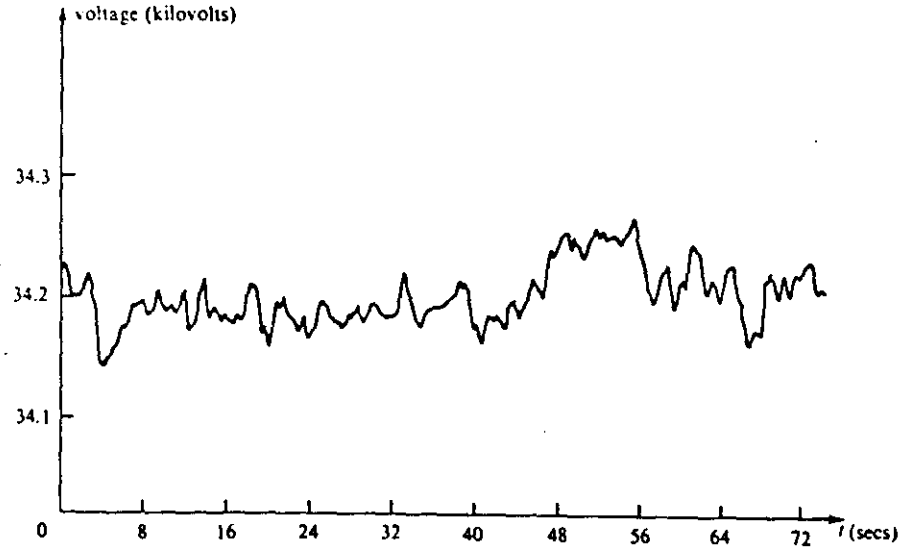


FIG. 1.1: Voltage deviations at the stator terminals of a 50-megawatt turbo-alternator

The behavior of the time series at all times can be described by a set of random variables $\{X(t)\}$ where the time variable t can take any value from $-\infty$ to $+\infty$. Thus the statistical properties of the series are described by associating probability distributions with any set of times $t_1, t_2, \ldots, t_N$. The ordered set of random variables $\{X(t)\}$ and its associated probability distributions is called a *stochastic process*. The observed time series $x(t)$ is thus regarded as one of the doubly infinite set of functions which might have been generated by the stochastic process. The set is doubly infinite because an infinite set of values is possible at any given time and there are an infinite number of time points.

Time series which occur in practice are either *discrete* or *continuous*. Examples of discrete time series are monthly figures of imports and exports or the yields on consecutive batches from a batch chemical process (Figure

5.2). Examples of continuous time series are the turbo-alternator data of Figure 1.1 or the radar return signal of Figure 5.1.

1.1.3 Experimental and non-experimental data

A more important distinction than that between discrete and continuous time series is whether the data are *non-experimental* or whether they are obtained from a planned experiment. For example, time series in economics and the social sciences are examples of non-experimental data. The economist is usually in a position where he can only *observe* the economic system and is rarely in a position to conduct planned experiments. A further difficulty associated with the analysis of economic time series is that they usually contain few observations. Therefore it is exceedingly difficult to check whether a given stochastic model provides a good fit to the data. Nevertheless, the techniques of time series analysis are of considerable importance in the analysis of economic data [1].

On the other hand, in engineering and the physical sciences, the time scale over which useful data need be collected is much smaller, so time series containing many more values can be obtained. Furthermore, it is possible to repeat experiments under similar sets of conditions so that the validity of the analysis and of different models can be checked.

1.2 TIME-DOMAIN AND FREQUENCY-DOMAIN DESCRIPTIONS OF TIME SERIES

It was stated in Section 1.1 that the stochastic process, from which the observed time series is regarded as being generated, can be described by the probability distributions associated with all possible sets of time points. To infer the nature of these probability distributions from a single or small number of series is an impossible or even meaningless exercise. In this section we discuss some of the most important simplifications which have been made in time series analysis in order to make the analysis of observed time series tractable and yet fruitful.

The most important assumptions made about a time series are that the corresponding stochastic process is *stationary*, and that a stationary stochastic process may be adequately described by the lower moments of its probability distributions. The lower moments include the *mean*, *variance*, *covariance function* and the Fourier transform of the covariance function, the power *spectrum*. An alternative approach to the above is to assume that the stochastic process can be adequately described by means of a model containing a few parameters which may be estimated from the data. These simplifying assumptions are now discussed briefly.

1.2.1 Stationarity

Examination of the output from a noise generator over a limited period of time shows that different sections of the output "look alike." In contrast, the characteristic feature of an economic time series, such as the gross national product of an industrialized country, is that its level tends to increase with time and hence different sections of the series will not be comparable. The output from the noise generator is said to be a *stationary* time series, whereas the gross national product series is said to be *non-stationary*.

Qualitatively, a stationary series is one which is in statistical equilibrium, in the sense that it contains no *trends*, whereas a non-stationary series is such that its properties change with time. Series occurring in practice are usually one of three kinds: those which exhibit stationary properties over long periods, for example, outputs from noise generators; those which are reasonably stationary over short periods, for example, measurements of atmospheric turbulence; and series which are obviously non-stationary in the sense that their visual properties are continuously changing with time.

Most methods of dealing with non-stationary time series are based on techniques for removing or filtering out the non-stationary part, leaving behind a series which can be treated as stationary. In some recent work [2], models which can describe non-stationary series have been given.

Since the statistical properties of stationary series do not change with time, these properties can be conveniently summarized by computing certain functions from the data. The function which was first used for this purpose is the autocovariance function.

1.2.2 The autocovariance function

In classical statistical work the measurements x_t $(t = 1, 2, \ldots, N)$ of some physical parameter can be assumed to be independent since the experiments which generate these observations are physically independent. If the probability distribution $f_X(x)$ associated with the measurements is *Normal* or *Gaussian*, it can be completely characterized by its mean

$$\mu = E[X] = \int_{-\infty}^{\infty} x f_X(x)\, dx \tag{1.2.1}$$

and its variance

$$\sigma^2 = E[(X - \mu)^2] = \int_{-\infty}^{\infty} (x - \mu)^2 f_X(x)\, dx. \tag{1.2.2}$$

The mean measures the location or center of gravity of the distribution and the variance its variability or spread.

If the observations x_t form part of a time series, then only for a purely random series will neighboring values be independent, that is, the size of

x_t is not influenced by the size of $x_{t-1}, x_{t-2}, \ldots$. In general, neighboring values of a time series will be *correlated*. Hence, in addition to specifying the mean μ and the variance σ^2, it is necessary in the case of a stationary Normal series to specify its *autocovariance function*

$$\gamma(u) = E[(X(t) - \mu)(X(t + u) - \mu)]. \tag{1.2.3}$$

In practice $\gamma(u)$ can be estimated by

$$c(u) = \frac{1}{N} \sum_{t=1}^{N-u} (x_t - \bar{x})(x_{t+u} - \bar{x}), \tag{1.2.4}$$

where

$$\bar{x} = \frac{1}{N} \sum_{t=1}^{N} x_t$$

is the *mean* of the observed time series. The plot of $c(u)$ versus u is called the *sample autocovariance function* of the time series. It is sometimes convenient when comparing series with different scales of measurement to normalize (1.2.4) by dividing by the variance $c(0)$. Thus the *sample autocorrelation function* is defined by

$$r(u) = \frac{c(u)}{c(0)}. \tag{1.2.5}$$

The sample autocorrelation function for the turbo-alternator data of Figure 1.1 is shown in Figure 1.2. It is seen that the voltage is highly positively correlated after one lag, which corresponds to $\frac{1}{2}$ second, is still slightly positively correlated after 1 second, but over the range $1\frac{1}{2}$ to $2\frac{1}{2}$ seconds is prominently negatively correlated. This means that if a large voltage occurs which is above the mean, there is a good chance that in about 2 seconds the voltage will be less than the mean, and vice versa. The estimates $r(u)$ for lags in the range $2\frac{1}{2}$ to 10 seconds are all very small, but persistently negative, which suggests that on the average a positive deviation from the mean tends to follow a negative one after 2 to 10 seconds. However, the values of $r(u)$ in this range are extremely small and hence inferences based on them may be suspect. The tail of the sample correlation function reveals a periodicity in the voltage waveform with a period of approximately 3 seconds. This periodicity could also account for some of the negative correlation near a lag of 2 seconds.

The autocorrelation function is useful in some situations because it gives a visual picture of the way in which the dependence in the series damps out with the lag or separation u between points in the series. However, the autocorrelation function is sometimes difficult to interpret because neighboring

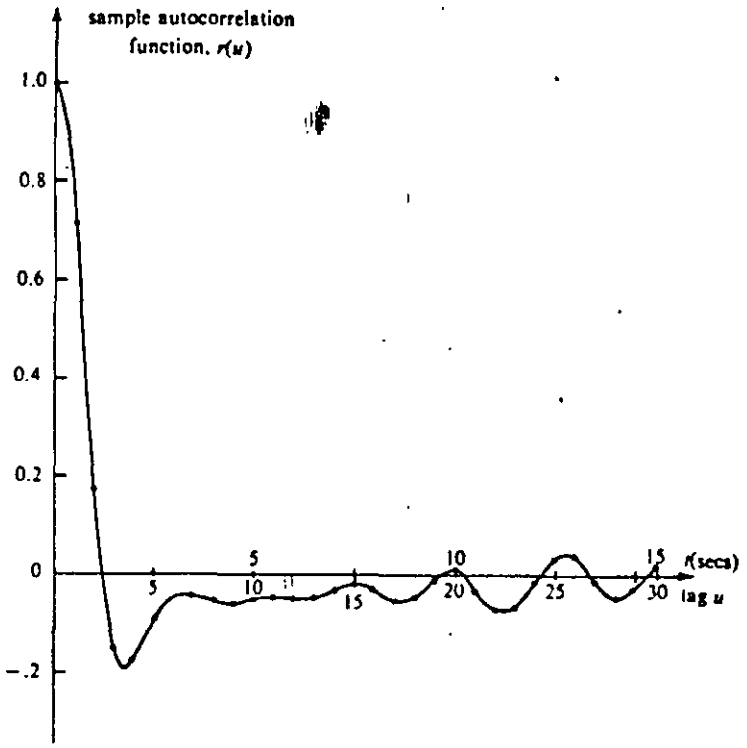


FIG. 1.2: Sample autocorrelation function for turbo-alternator data of Figure 1.1, $N = 1000$

values can be highly correlated. This means that the sample autocorrelation function may be distorted visually. A more detailed description of the properties and uses of the autocorrelation function is given in Chapter 5. Its main use in this book will be as an intermediate step in the estimation of the spectrum.

1.2.3 *The spectrum*

Suppose that the time series x_t consists of values of the cosine function (1.1.1) at discrete times. Then for frequencies f_0 which are multiples of the fundamental frequency $1/N$, it may be verified that the variance calculated according to (1.2.4) is $a^2/2$. If x_t is measured in volts, then this means that the average ac power or variance of the series is $a^2/2$ watts. More generally, if x_t consists

of a mixture of several cosine waves with frequencies f_i and amplitudes a_i, then the variance is

$$c(0) = \sum_i \tfrac{1}{2}a_i^2. \tag{1.2.6}$$

The result (1.2.6) shows that if x_t can be regarded as being made up of mixtures of cosine waves, its variance can be decomposed into components of average power or variance $\frac{1}{2}a_i^2$ at the various frequencies f_i. It will be shown in Chapter 6 that if x_t is a stationary time series, the variance of the corresponding stochastic process can be decomposed into contributions at a *continuous range of frequencies* according to

$$\sigma^2 = \gamma(0) = \int_{-\infty}^{\infty} \Gamma(f)\, df,$$

where $\Gamma(f)$ is called the power *spectrum* of the stochastic process. Thus $\Gamma(f)\,\delta f$ is an approximate measure of the average power or variance in the frequency band f to $f + \delta f$.

The *spectrum estimate* of the turbo-alternator data of Figure 1.1 is shown in Figure 1.3. The main feature of this spectrum is that there is high power at low frequencies and low power at high frequencies. This is due mainly to the high positive values of the sample autocorrelation function at lags 1 and 2. Note also that the power does not drop gradually from high to low frequencies. Instead, there is a flat region over the range 0–$\frac{1}{4}$ cps. There is also a well-defined minor peak at frequency 0.39 cps, or period 2.54 seconds, which could explain the small periodicity in the tail of the sample correlation function of Figure 1.2.

It will also be shown in Chapter 6 that the spectrum and the autocovariance function are related according to the Fourier transform relation

$$\Gamma(f) = \int_{-\infty}^{\infty} \gamma(u) \cos 2\pi f u\, du, \tag{1.2.7}$$

and hence knowledge of the autocovariance function of the process is equivalent to knowledge of the spectrum of the process.

However, in the analysis of a finite length of record, the spectrum is often preferable to the autocovariance function. First, estimates of the spectrum at neighboring frequencies are approximately independent, and hence the interpretation of the sample spectrum is usually easier than that of the sample autocovariance function. More important, in many physical problems the spectrum is of direct physical interest. Examples of the uses of spectral analysis will be given in Section 1.3.

Digital filters. Although it is necessary to assume stationarity to describe a stochastic process by its spectrum, in practice *the stationarity assumption does*

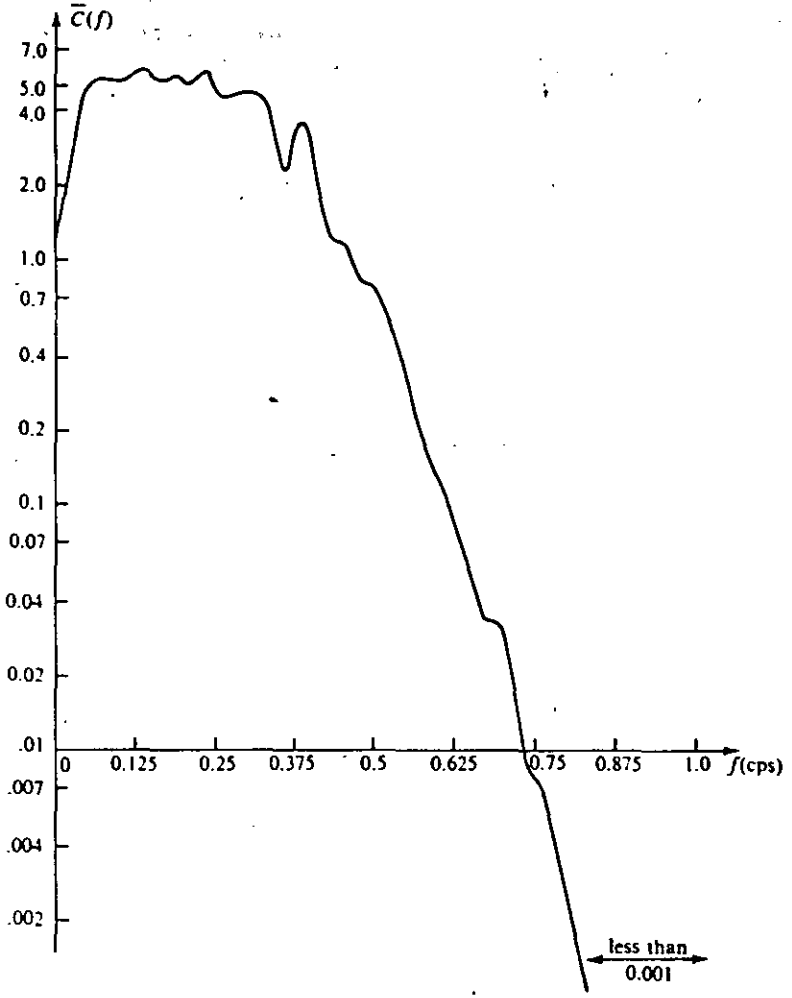


FIG. 1.3: Spectrum estimate for turbo-alternator data of Figure 1.1, $N = 1000$

not present serious problems. This is because the spectrum isolates the contributions in the series which can be attributed to different frequency bands. A non-stationary series is usually characterized by the presence of large power at low frequencies. However, in many practical applications, the information which is of interest may be at higher frequencies. In such cases, all that is necessary is to filter off the non-stationary low-frequency components and use the residual series for the spectral analysis.

A particularly simple form of digital filter for removing low-frequency components is the first difference filter

$$y_t = (x_t - x_{t-1}). \tag{1.2.8}$$

The gain $G(f)$ of this filter is shown in Figure 1.4 and measures the extent to

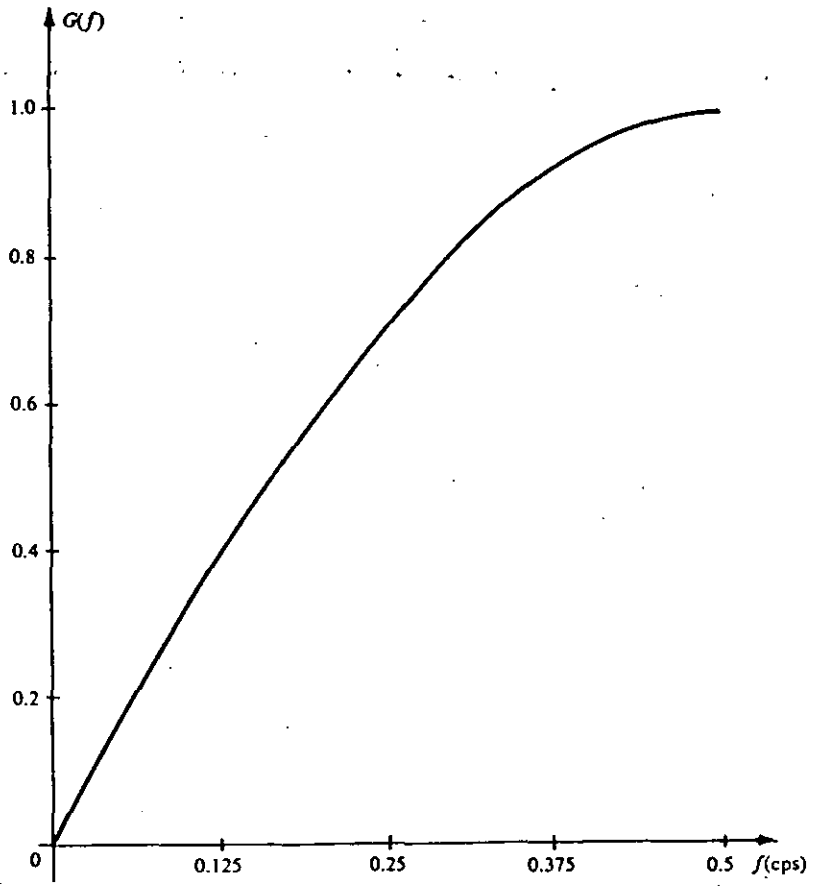


FIG. 1.4: Gain function for first difference filter

which a cosine wave of frequency f is transmitted by the filter. It is seen that low frequencies will be considerably attenuated and hence will tend not to be present in the output from the filter.

1.2.4 Parametric time series models

In many problems, such as those where it is required to *predict* future values of the series, it is necessary to construct a parametric model for the time series. To be useful, the model should be physically meaningful and involve as few parameters as possible. A powerful parametric model which has been widely used in practice for describing empirical time series is the *moving average–autoregressive process*

$$X_t - \mu = \alpha_1(X_{t-1} - \mu) + \cdots + \alpha_m(X_{t-m} - \mu) + Z_t + \beta_1 Z_{t-1} + \cdots + \beta_l Z_{t-l}, \tag{1.2.9}$$

where Z_t is a purely random series, or white noise, and μ is the mean level of X_t. The model (1.2.9) is physically meaningful since it is the discrete analog of the familiar linear differential equation used to describe linear systems. The model thus represents time series as the output from a linear system whose input is white noise. By introducing a suitable number of the parameters α and β in (1.2.9) it is possible, after suitable differencing [2], to fit *most empirical time series with a relatively small number of parameters.*

The decision as to whether to use the autocovariance function, the spectrum or a parametric model will be dictated in practice by the needs of the situation. Different needs will require different methods of approach and hence it is important to realize that *there is no one approach which should be applied to the analysis of all time series in all situations.*

1.3 OBJECTIVES OF TIME SERIES ANALYSIS

Several different applications of spectral analysis will be described in this book. Since spectral analysis is just one tool which is available for the analysis of time series, it is useful by way of introduction to discuss the nature of time series problems in greater generality.

At the risk of oversimplification, time series problems can be classified into those which require some form of *model building* and those which lead to *frequency response studies.* These categories inevitably have a certain amount of overlap.

1.3.1 Model building

One can usually distinguish between several different kinds of models, for example, *exploratory* and *sophisticated* models; *empirical* and *physical* models; and *parametric* and *non-parametric* models.

Exploratory and sophisticated models. In the early stages of an investigation, one may know very little about a particular phenomenon. The main purpose of the time series analysis at this stage may be to look at the data in several different ways to see what hypotheses are suggested. For example, a study of the spectrum of the vertical velocity field of atmospheric turbulence [3] indicated that the peak in the spectrum moved toward lower frequencies with increased solar radiation. This suggested that there were two different causes for the fluctuations in atmospheric turbulence, a high-frequency component due to frictional forces and a low-frequency component due to heat convection caused by solar radiation. As a result of this exploratory analysis, it was possible to begin to construct a more realistic model for atmospheric turbulence.

As in the above example, it often happens that a time series model which is guessed initially may be proved inadequate subsequently. The nature of the inadequacies in the exploratory model can then be used as a basis for modifying it and constructing a more sophisticated model.

Empirical and physical models. In some situations it is possible to write down detailed models for a time series based on the physics of the situation. For example, considerable effort has gone into the construction of physical models for atmospheric turbulence [3] and hydrodynamic turbulence [4]. In other situations, so little is known about the phenomenon being investigated that it is necessary to resort to the fitting of a more empirical model, such as the moving average–autoregressive model (1.2.9). A big advantage of physical models is that they usually require fewer parameters than empirical models. The decision to spend time and effort deriving a physical model or to resort to an empirical model requires judgement and insight. In general, some compromise is necessary and one should use whatever physical knowledge is available and start building on these foundations.

Parametric and non-parametric models. The moving average–autoregressive model (1.2.9) is a *parametric* model. To fit this model it is only necessary to estimate a small set of parameters from the data. On the other hand, the description of a time series provided by the autocorrelation function or the spectrum is *non-parametric* (or *multiparametric*, since an effectively infinite number of parameters is required to specify the process).

FIG. 1.5: Schematic representation of a linear system

A problem to which both of these methods may be applied is the estimation of the dynamic model of a linear system as shown in Figure 1.5. In the simplest case, where there is one input x_1 and one output x_2, it is possible to estimate the dynamic model from records $x_1(t)$, $x_2(t)$, of the input and output. For example, it may be known that a simple parametric model such as

$$T\frac{dx_2(t)}{dt} + x_2(t) = x_1(t)$$

is adequate. This model involves one parameter T, called the *time constant* of the system. In other situations it may be more convenient to describe the system by a nonparametric model requiring specification of the gain function $G(f)$ and the phase function $\phi(f)$ of the system. These are functions which describe the response of the system to cosine waves of different frequencies f.

Thus an input cosine wave $x_1(t) = a \cos 2\pi ft$ appears at the output as a cosine wave $x_2(t) = aG(f) \cos [2\pi ft + \phi(f)]$, that is, the amplitude is changed from a to $aG(f)$ and the phase is changed by $\phi(f)$. It will be shown in Chapter 10 that cross spectral analysis can be used to estimate the gain and phase of a linear system.

In some situations the gain-phase description is preferable, since the model may be needed only over a limited frequency band. In other situations the overall description provided by the parametric model may be preferable.

Since spectral analysis is a non-parametric approach, its usefulness in the area of model building is limited. However, as in the turbulence example mentioned above, it is sometimes useful as an exploratory tool for suggesting models which can then be fitted parametrically.

1.3.2 Uses of time series models

Time series models are used for a variety of purposes. Some of the most common of these are (a) prediction; (b) estimation of transfer functions; (c) filtering and control; (d) simulation and optimization; and (e) generating new physical theories.

Prediction. By prediction is meant the estimation of future values $x(t + T)$ of the time series in some future range $0 \leqslant T \leqslant l$ from values of the series up to and including time t. The prediction of economic and industrial time series is a very important application of time series and is discussed in reference [2].

Estimation of transfer functions. This application has been discussed above.

Filtering and control. A more general problem than the one of prediction described above is that of linear filtering as formulated by Wiener [5]. The linear filter is a device which operates on the input $x_1(t)$ to give an output $x_2(t)$ according to

$$x_2(t) = \int_0^\infty h(u)x_1(t - u)\, du, \tag{1.3.1}$$

where $h(u)$ is the *weighting function* or *impulse response function* of the filter. Suppose that the input $x_1(t) = s(t) + z(t)$, where $s(t)$ is the *signal* or useful information and $z(t)$ is the *noise* or unwanted information. Then, as shown in Figure 1.6, the filter is required to produce an output which is some function of the signal at a future time T. For example, the desired output might be

$$g[s(t + T)] = \frac{d}{dt}[s(t + T)].$$

The optimal filter is then defined as the weight function which minimizes the mean square of the error signal

$$\epsilon(t) = x_2(t) - g[s(t + T)]$$

between the actual and desired outputs. If models are available for the stochastic processes which describe the signal $s(t)$ and the noise $z(t)$, techniques are available [6] for calculating the gain and phase function of the optimal filter. In fact, the calculation of the optimal filter is made easier by working with the spectra of the signal $s(t)$ and the noise $z(t)$.

In control theory it is required to follow some specified signal $s(t)$ with as small an error as possible. It may be shown [6] that this reduces to a special case of the filtering problem described above.

Simulation and optimization. Many systems, for example electronic guidance systems, are much too complex to study or optimize analytically. In such a case, the system may be simulated using analog, digital or hybrid computers.

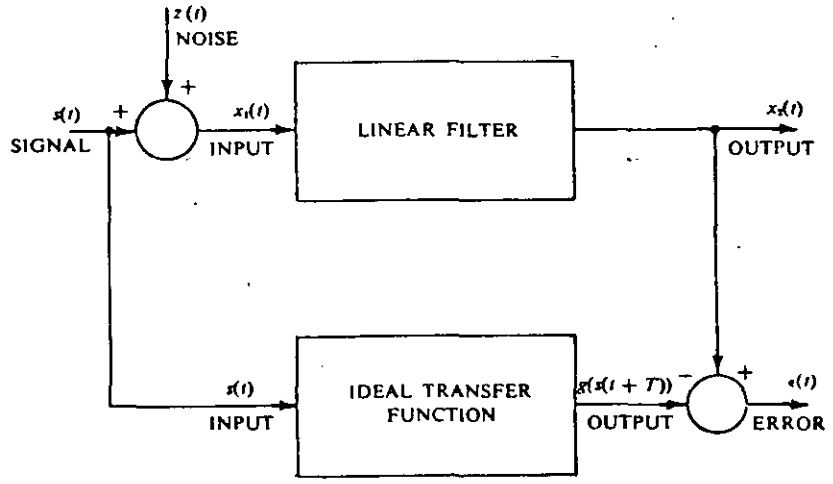


FIG. 1.6: Schematic of the general filtering problem

The disturbances entering the real system at various points can be measured, and if enough of these data are available they may be fed directly into the simulator. However, the quantity of data required for simulation studies is usually so large that it is necessary to fit models to the disturbances. The models can then be used to generate unlimited amounts of artificial data which can then be used in the simulation.

Generating new physical theories. The above applications of time series models refer to engineering problems. In applications of time series analysis in physics, the approach is somewhat different. The physicist is interested in producing theories for physical phenomena which can be used to predict as wide a range of situations as possible. Hence the time series models devised are more detailed than those required by the engineer and must be constantly

modified and enlarged as greater physical insight is obtained. The use of time series analysis in building physical models for atmospheric turbulence is well illustrated in [3].

1.3.3 Frequency response studies

Although spectral analysis has an important role to play in time series model building, it is in the area of frequency response studies that it is most relevant.

It has been mentioned above that the linear system (1.3.1) transforms a cosine wave input $x_1(t) = a \cos 2\pi ft$ into a cosine wave output $x_2(t) = a\,G(f) \cos(2\pi ft + \phi(f))$. It will be shown in Chapter 6 that if $x_1(t)$ is a stationary time series, the *spectrum* of the input is changed by multiplying it by the *square* of the gain. That is, the spectrum of $x_2(t)$ is

$$\Gamma_2(f) = G^2(f)\Gamma_1(f). \tag{1.3.2}$$

Figure 7.22 shows the effect of passing an input with the spectrum shown in (a) of that figure through three systems whose squared gains are shown in (b). In the example illustrated in Figure 7.22, the input refers to the roughness of a runway, the system is the landing gear of an aircraft and the output is a typical aircraft response such as the acceleration at the center of gravity. Using the result (1.3.2), it is seen that the combination of the input spectrum with the landing gear response marked 3 produces an output spectrum with a very sharp peak, as indicated by 3 in (c) of Figure 7.22. This indicates that large accelerations will result at this resonant frequency, producing passenger discomfort and large stresses in the landing gear. Knowing the gain plots of the landing gears of typical aircraft at typical landing speeds, it is possible to lay down norms for the roughness of runways.

In the above example, it was possible to change the input spectrum by designing suitable runways, but the characteristics of the landing gear would usually have been specified from other considerations and hence would have to be regarded as fixed. The converse holds in the design of suspension units for motorcycles or cars. In recent years certain companies have based the design of suspension systems on measurements of the roughness of the roads in a particular country. Roads differ widely in their characteristics from one country to another, and hence suspension systems may be designed accordingly.

Frequency response studies may also be applied to the design of aircraft structures and to the design of experiments to optimize the performance of industrial processes. These applications are discussed more fully in Chapter 7.

1.4 SCOPE OF THE PRESENT BOOK

There is no doubt that spectral techniques have come to play an important part in the analysis of time series. However, it is important to realize that

they do have limitations and should be applied judiciously. The pioneering work in the estimation of spectra from finite lengths of records is to be found in the books by Bartlett [7] and Blackman and Tukey [8]. Their books are essentially concerned with the estimation of the spectra of single time series. In the present book, these ideas are extended to the estimation of the spectra and cross spectra of several time series and their subsequent use in the estimation of gain and phase functions of linear systems.

Several topics in spectral analysis have not been included here. An important omission is the spectral analysis of stochastic processes defined in several dimensions, for example, the elevation of ocean waves as a function of earth coordinates. Another omission is that of higher-order spectra, for example, the bi-spectrum. These higher-order spectra are useful in the analysis of non-Normal processes and non-linear systems. Stochastic processes defined in several dimensions have been omitted because the book is already too long. However, non-linear spectra have been omitted mainly because we think that the extra complication introduced by these spectra will inhibit their widespread use. On the present evidence it is felt that parametric methods are preferable in these situations.

REFERENCES

[1] C. W. J. Granger, *Spectral Analysis of Economic Time Series*. Princeton University Press, Princeton, 1964.

[2] G. E. P. Box and G. M. Jenkins, *Statistical Models for Forecasting and Control*. To be published, Holden-Day, San Francisco.

[3] John L. Lumley and Hans A. Panofsky, *The Structure of Atmospheric Turbulence*. John Wiley, New York, 1964.

[4] G. K. Batchelor, *The Theory of Homogeneous Turbulence*. Cambridge University Press, Cambridge, 1953.

[5] N. Wiener, *The Extrapolation, Interpolation and Smoothing of Stationary Time Series with Engineering Applications*. John Wiley, New York, 1949.

[6] J. H. Laning and R. H. Battin, *Random Processes in Automatic Control*. McGraw-Hill, New York, 1956.

[7] M. S. Bartlett, *An Introduction to Stochastic Processes with Special Reference to Methods and Applications*. Cambridge University Press, Cambridge, 1953.

[8] R. B. Blackman and J. W. Tukey, *The Measurement of Power Spectra from the Point of View of Communications Engineering*. Dover, New York, 1958.

DIVISION DE EDUCACION CONTINUA
FACULTAD DE INGENIERIA U.N.A.M.

CURSO: INGENIERIA MARITIMA.
MODULO: "FUNDAMENTOS DE HIDRAULICA Y MATEMATICAS"
DEL 11 DE MARZO AL 7 DE JUNIO
MEXICO, D.F.

FOURIER ANALYSIS

PROF. ING. JOSE RAYNAL.
JUNIO DE 1985.

2

Fourier Analysis

Spectral analysis brings together two very important theoretical approaches, the statistical analysis of time series and the methods of Fourier analysis. The latter requires no detailed exposition for engineers since much of their training is deeply rooted in these methods. However, for the sake of completeness and for the benefit of other readers, this chapter describes those ideas in Fourier analysis necessary for the analysis of time series. In later chapters, it will be shown how Fourier techniques need to be modified to deal with statistical rather than deterministic functions of time.

2.1 INTRODUCTION

2.1.1 The role of Fourier analysis in applied mathematics and engineering

The analytic techniques developed by Jean-Baptiste-Joseph Fourier (1768–1830) have played an important role in the development of applied mathematics. They are particularly important in three applications: (a) for studying periodic solutions to physical problems described by differential equations, especially partial differential equations—for example, the study of wave motion of plucked strings or the transmission of electromagnetic waves in waveguides or cables; (b) as an operational device for solving differential equations—for example, ordinary differential equations with constant coefficients may be converted into algebraic equations by Fourier transformation; (c) for approximating non-periodic functions.

This book will be concerned primarily with the latter case and only incidentally with solving differential equations. It will not consider periodic solutions to physical problems. As an instance of approximating a non-periodic function, consider a deterministic function $s(t)$ of time t which will be called a *signal* and which is to be approximated by means of suitably chosen periodic functions. A deterministic signal is a function which is known exactly *for all time* and hence is a mathematical idealization. Examples of deterministic signals are

$$s(t) = e^{-|t|}, \quad -\infty \leqslant t \leqslant \infty,$$

or

$$s(t) = a \cos 2\pi f_1 t, \quad -\infty \leqslant t \leqslant \infty.$$

Many practical signals are usefully thought of as deterministic, for example, the supply voltage as a function of time, the output of a square-wave generator, the displacement of an object subjected to a suddenly applied constant force or the current through a resistor which is suddenly shunted across a charged capacitor. The dimensions of the first two of these signals are volts, of the third, feet, and of the fourth, amperes. However, the dimensions could be feet per second if the signal were a velocity or they could be units of temperature, pressure and so forth. For consistency it will always be assumed that t is measured in seconds and $s(t)$ in volts, since in most practical applications the physical quantity which is being studied will be converted into a voltage before recording.

The first deterministic example cited above is *non-periodic* while the second is *periodic*. Periodic means that there is a number T, called the *period* of the function, such that

$$s(t) = s(t + T) \tag{2.1.1}$$

for all t.

The function between time t and $t + T$ can be of any shape whatsoever. A particularly simple shape is the cosine function example mentioned above which has period $T = 1/f_1$, since

$$a \cos 2\pi f_1[t + (1/f_1)] = a \cos 2\pi f_1 t.$$

It is possible to represent non-periodic functions using any class of periodic functions. In Fourier analysis, the periodic functions used are sine and cosine functions. They have the important properties that an approximation consisting of a given number of terms achieves the minimum mean square error between the signal and the approximation, and also that they are orthogonal, so the coefficients may be determined independently of one another.

2.1.2 Finite Fourier Series

Consider a signal which is specified only at discrete times, and suppose it is required to expand it in terms of periodic functions. The discrete signal can be regarded as having been derived from a continuous signal $s(t)$ of duration T by sampling the values of the signal at spacing Δ, as shown in Figure 2.1(a). This produces $N = T/\Delta$ sample values s_r, where

$$s_r = s(t = r\Delta). \tag{2.1.2}$$

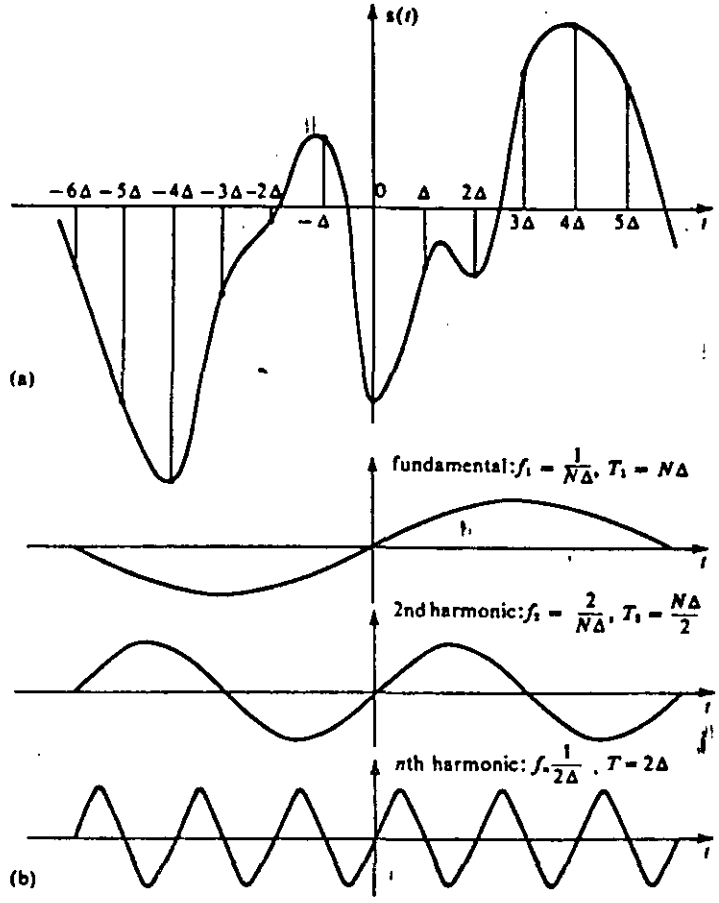


FIG. 2.1: (a) A discrete signal obtained by sampling a continuous signal
(b) The fundamental sine wave and harmonics

For convenience it is assumed that N is even and equal to $2n$ so that r may run through the integers $-n, \ldots, 0, 1, \ldots, n-1$.

Note that periodic functions which pass through the sample values may be chosen in an infinite number of ways. For example, the finite Fourier series

$$\bar{s}(t) = A_0 + 2\sum_{m=1}^{n-1}\{A_m \cos 2\pi m f_1 t + B_m \sin 2\pi m f_1 t\} + A_n \cos 2\pi n f_1 t \quad (2.1.3)$$

contains N constants, the A_m and B_m, which can be determined so that the discrete and continuous values coincide at the points $t = r\Delta$, that is, $\bar{s}(t) = s_r$. It follows that the function $\bar{s}(t)$ provides an *approximation* to the original continuous function $s(t)$ in the interval $-T/2 \leqslant t < T/2$.

On substituting $t = r\Delta$ in (2.1.3) and setting $\bar{s}(r\Delta) = s_r$, a set of N equations for the N unknown constants is obtained. The equations are

$$s_r = A_0 + 2\sum_{m=1}^{n-1}\{A_m \cos 2\pi m f_1 r\Delta + B_m \sin 2\pi m f_1 r\Delta\} + A_n \cos 2\pi n f_1 r\Delta, \quad (r = -n, \ldots, 0, 1, \ldots, n-1). \quad (2.1.4)$$

Choosing $f_1 = 1/N\Delta$ simplifies the solution of equation (2.1.4), because then the sines and cosines are *orthogonal*, that is, they satisfy the relations

$$\sum_{r=-n}^{n-1} \sin\frac{2\pi kr}{N}\cos\frac{2\pi mr}{N} = 0, \quad k, m \text{ integers};$$

$$\sum_{r=-n}^{n-1} \sin\frac{2\pi kr}{N}\sin\frac{2\pi mr}{N} = \begin{cases} 0, & k \neq m \\ \frac{N}{2}, & k = m \neq 0, n \\ 0, & k = m = 0, n; \end{cases} \quad (2.1.5)$$

$$\sum_{r=-n}^{n-1} \cos\frac{2\pi kr}{N}\cos\frac{2\pi mr}{N} = \begin{cases} 0, & k \neq m \\ \frac{N}{2}, & k = m \neq 0, n \\ N, & k = m = 0, n. \end{cases}$$

The frequency $f_1 = 1/N\Delta$ is called the *fundamental frequency* of the signal $\bar{s}(t)$, and it corresponds to a period equal to the length of the record, as shown in Figure 2.1(b). The dimensions of f_1 are cycles per second (cps) when t is measured in seconds (sec).

The function $\bar{s}(t)$ in (2.1.3) is thus composed of a sum of sine and cosine functions whose frequencies are multiples or *harmonics* of the fundamental f_1, as is illustrated in Figure 2.1(b). The highest frequency present is $n/N\Delta = 1/2\Delta$ cps, which corresponds to a period of 2 sampling intervals.

The coefficients A_m or B_m, when $f_1 = 1/N\Delta$, may be determined by multiplying both sides of (2.1.4) by $\cos(2\pi mr/N)$ or $\sin(2\pi mr/N)$ and summing over r, then making use of the orthogonality relations (2.1.5).

The final expressions for the coefficients are

$$A_m = \frac{1}{N}\sum_{r=-n}^{n-1} s_r \cos\frac{2\pi mr}{N}, \quad (2.1.6)$$

$$B_m = \frac{1}{N}\sum_{r=-n}^{n-1} s_r \sin\frac{2\pi mr}{N}, \quad (2.1.7)$$

for $m = 0, 1, \ldots, n$. A_0 is the *mean* or average value of s_r. Similar expressions may be derived when the number of points N is odd, say $2n-1$, the only difference being that the A_n term vanishes.

An example. Consider the data of Table 2.1, which gives the intensity of reflected signals from one of the *E*-layers in the ionosphere. The figures given are averages, over several months, of the intensity at fixed times during the day.

TABLE 2.1: Intensity of reflected signals from ionosphere

Time	0	1	2	3	4	5	6	7	8	9	10	11
Average intensity	−6	−20	−28	−8	−1	7	−20	−6	−7	14	19	12

Table 2.2 gives the values of the coefficients A_m and B_m computed from (2.1.6) and (2.1.7) using hour 6 as the origin of time. The coefficient A_2, for example, is obtained from

$$A_2 = \frac{1}{12}\left\{(-6)\cos(-2\pi) + (-20)\cos\left(-\frac{5\pi}{3}\right) + \cdots + (12)\cos\left(\frac{5\pi}{3}\right)\right\}$$
$$= \frac{1}{12}\{-6 - 10 + \cdots + 6\} = -2.25.$$

TABLE 2.2: Fourier decomposition of mean square for ionospheric data

Source	m	A_m	B_m	R_m	$\phi_m(°)$	Contribution to mean square
mean	0	−3.667	0	3.667	180	13.44
fundamental	1	−0.475	5.584	5.604	85	62.81
2nd harmonic	2	−2.250	−7.073	7.422	−72	110.17
3rd harmonic	3	−1.250	−0.250	1.275	−11	3.25
4th harmonic	4	−0.667	0.577	0.882	41	1.56
5th harmonic	5	−1.775	−0.334	1.806	−11	6.52
6th harmonic	6	−3.500	0	3.500	0	12.25
Total						210.00

Amplitude and phase representation. It is sometimes more convenient to write (2.1.3) in the form

$$\bar{s}(t) = R_0 + 2\sum_{m=1}^{n-1} R_m \cos(2\pi m f_1 t + \phi_m) + R_n \cos 2\pi n f_1 t, \quad (2.1.8)$$

where

$$R_m = \sqrt{A_m^2 + B_m^2}, \qquad \phi_m = \arctan -\frac{B_m}{A_m} \quad (2.1.9)$$

and

$$A_m = R_m \cos\phi_m, \qquad B_m = -R_m \sin\phi_m. \quad (2.1.10)$$

R_m is called the *amplitude* and ϕ_m the *phase* of the mth harmonic relative to an arbitrary origin of time. In the above formulae the origin of time has been taken at a point roughly halfway between the first and last value of s_r. If this origin were changed, the amplitude would remain unaltered but the phase would change accordingly. The amplitudes and phases for the ionosphere data are shown in Table 2.2.

Parseval's theorem. The mean square value or average power of the signal s_r is

$$\frac{1}{N}\sum_{r=-n}^{n-1} s_r^2.$$

Using (2.1.3) and the orthogonality property (2.1.5) it may be verified that this can be written

$$\frac{1}{N}\sum_{r=-n}^{n-1} s_r^2 = R_0^2 + 2\sum_{m=1}^{n-1} R_m^2 + R_n^2, \quad (2.1.11)$$

which is a special case of *Parseval's theorem*. In words, it states that the mean square value of s_r, or the *average power* dissipated by s_r, can be decomposed into contributions arising from each harmonic. For the zero-th and nth harmonic the contribution is R_m^2 but for the mth harmonic the average power is $2R_m^2$.

A more convenient measure is the mean square value of s_r about the mean R_0. This is simply the *variance*

$$\sigma^2 = \frac{1}{N}\sum_{r=-n}^{n-1}(s_r - R_0)^2 = 2\sum_{m=1}^{n-1} R_m^2 + R_n^2, \quad (2.1.12)$$

or, in electrical terms, the *average ac power*.

The decomposition of the mean square value of s_r for the ionospheric data is shown in Table 2.1. It is seen that the mean, fundamental and first harmonic account for about 89% of the total mean square, showing that the data are very well approximated by the model

$$s_r = -3.67 + 11.2\cos\left(\frac{\pi r}{6} + 85°\right) + 14.8\cos\left(\frac{\pi r}{3} - 72°\right).$$

The decomposition of the mean square may be displayed by plotting the average power at the harmonic versus the frequency of the harmonic. This is called a *Fourier line spectrum* and is shown in Figure 2.2 for the ionospheric data.

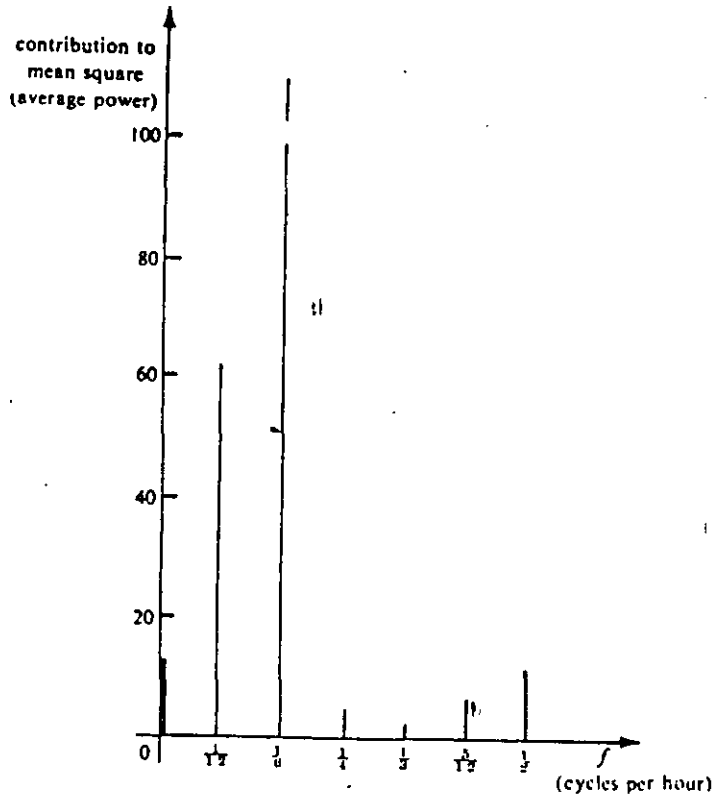


FIG. 2.2: A Fourier line spectrum (periodogram)

Complex Fourier series. The above formulae are cumbersome to manipulate, and hence for operational convenience it is preferable to express s_r in terms of *complex* amplitudes S_m, where

$$S_m = R_m e^{+j\phi_m} = A_m - jB_m, \quad j^2 = -1. \tag{2.1.13}$$

Thus (2.1.3) may be written

$$\bar{s}(t) = \sum_{m=-n}^{n-1} S_m e^{j(2\pi mt/N\Delta)}, \tag{2.1.14}$$

where $S_{-m} = S_m^*$, the asterisk denoting a complex conjugate. Similarly, formulae (2.1.6, 7) become

$$S_m = \frac{1}{N} \sum_{r=-n}^{n-1} s_r e^{-j(2\pi mr/N)}, \quad -n \leqslant m \leqslant n-1, \tag{2.1.15}$$

and Parseval's theorem (2.1.11) becomes

$$\frac{1}{N} \sum_{r=-n}^{n-1} s_r^2 = \sum_{m=-n}^{n-1} |S_m|^2. \tag{2.1.16}$$

Hence the contribution to the mean square, $2R_m^2$ in (2.1.11), is divided in (2.1.16) into two parts, each equal to $|S_m|^2 = R_m^2$, one at the frequency mf_1, the other at the frequency $-mf_1$.

Throughout this book it will be found convenient to operate with complex transforms. The resulting formulae can then be converted into real form by taking real and imaginary parts. For example, taking the real and imaginary parts of (2.1.15) gives the sine and cosine transforms (2.1.6, 7).

2.1.3 Fourier series

Suppose that a Fourier representation of a continuous signal in the interval $-T/2$ to $T/2$ is required. Note in the analysis of the previous section that if the sampling interval Δ tends to zero, then the sample points s_r will trace out the continuous signal $s(t)$. The continuous signal $\bar{s}(t)$, which is constrained to pass through the sample points s_r, must then coincide with $s(t)$, and hence in this limiting case the Fourier representation $\bar{s}(t)$ is an *exact* representation of the signal $s(t)$ over the interval $-T/2$ to $T/2$.

The Fourier coefficients S_m defined in (2.1.15) may be rewritten

$$S_m = \frac{1}{N\Delta} \sum_{r=-n}^{n-1} s_r \Delta\, e^{-j(2\pi mr\Delta/N\Delta)}, \tag{2.1.17}$$

and if Δ tends to zero and N tends to infinity in such a way that $N\Delta = T$, then $r\Delta$ tends to t, $s_r\Delta$ tends to $s(t)\,dt$ and the sum (2.1.17) tends to the integral

$$S_m = \frac{1}{T} \int_{-T/2}^{T/2} s(t)\, e^{-j(2\pi mt/T)}\, dt. \tag{2.1.18}$$

Similarly, (2.1.14) tends to

$$s(t) = \sum_{m=-\infty}^{\infty} S_m e^{j(2\pi mt/T)}. \tag{2.1.19}$$

Parseval's theorem, (2.1.16), now becomes

$$\frac{1}{T} \int_{-T/2}^{T/2} s^2(t)\, dt = \sum_{m=-\infty}^{\infty} |S_m|^2, \tag{2.1.20}$$

since (2.1.16) may be written

$$\frac{1}{N\Delta} \sum_{r=-n}^{n-1} s_r^2 \Delta = \sum_{m=-n}^{n-1} |S_m|^2,$$

and as Δ tends to zero and N tends to infinity, $s_r^2\Delta$ tends to $s^2(t)\,dt$. Equation (2.1.20) states that the mean square of the continuous periodic signal $s(t)$

can be decomposed into an infinite number of contributions at harmonics $f_m = m/T$, $m = -\infty$ to $+\infty$, of the fundamental frequency $1/T$ cps. Equation (2.1.19) is called the *Fourier series* representation of the function $s(t)$ in the interval $-T/2 \leqslant t < T/2$. Note that although the above limiting arguments are heuristic, they can be justified rigorously.

2.1.4 Fourier integrals

Up to now it has been shown that two types of signal may be represented by means of trigonometric series. The first type of signal s, consisted of a finite number (N) of equispaced ordinates, spaced Δ seconds apart. This could be represented over the given interval by a continuous signal $\bar{s}(t)$ composed of N harmonics of the fundamental frequency $1/N\Delta$ cps. The maximum frequency present is $1/2\Delta$ cps, and hence the signal $\bar{s}(t)$ is said to be *band-limited*. The second type of signal, a continuous signal $s(t)$ available over the interval $-T/2 \leqslant t \leqslant T/2$, was seen to be represented over the interval by a signal composed of an infinite number of harmonics of the fundamental frequency $1/T$ cps.

More generally, it is necessary to consider a third type of signal $s(t)$ defined over the *infinite* interval $-\infty \leqslant t \leqslant \infty$. The corresponding Fourier analysis is a limiting case of the analysis of Section 2.1.3 in which increasingly large segments T of the infinite record are analyzed. As T tends to infinity the frequency spacing $1/T$ between harmonics becomes infinitesimal, which results in a continuous amplitude distribution over frequency.

To demonstrate this limiting argument, (2.1.19) may be rewritten

$$s(t) = \sum_{m=-\infty}^{\infty} (TS_m)\, e^{j(2\pi mt/T)}\, \frac{1}{T}. \tag{2.1.21}$$

In the limit, as T tends to infinity, m/T tends to f, $1/T$ tends to df, and TS_m tends to $S(f)$. Hence (2.1.21) tends to the integral

$$s(t) = \int_{-\infty}^{\infty} S(f)\, e^{j2\pi ft}\, df. \tag{2.1.22}$$

Similarly, (2.1.18) may be written

$$TS_m = \int_{-T/2}^{T/2} s(t)\, e^{-j2\pi(m/T)t}\, dt, \tag{2.1.23}$$

which tends to

$$S(f) = \int_{-\infty}^{\infty} s(t)\, e^{-j2\pi ft}\, dt, \tag{2.1.24}$$

as T tends to infinity. The function $S(f)$ is called the *Fourier transform* of $s(t)$.

Parseval's relation (2.1.20) for the infinite interval case may be written

$$\int_{-T/2}^{T/2} s^2(t)\, dt = \sum_{m=-\infty}^{\infty} |TS_m|^2\, \frac{1}{T}, \tag{2.1.25}$$

which tends to

$$\int_{-\infty}^{\infty} s^2(t)\, dt = \int_{-\infty}^{\infty} |S(f)|^2\, df. \tag{2.1.26}$$

The limiting operation in (2.1.25) may be thought of as first distributing the power or variance $|S_m|^2$ at frequency m/T over a band of width $1/T$, giving an average power $T|S_m|^2$ over the band. This average power then tends to a continuous distribution of power over frequency as the width of the band becomes infinitesimal.

Physically, the Fourier transform $S(f)$ represents the *distribution of signal strength with frequency*, that is, it is a density function. When s is measured in volts and t in seconds, the dimensions of $S(f)$ are "volt-seconds" or "volts per unit of f" since f has the dimensions of frequency, that is, $\sec^{-1}$.

Mathematical texts on Fourier analysis give a wide variety of sufficient conditions for the existence of the integrals (2.1.22) and (2.1.24). In this book, these conditions are avoided by using the theory of distributions conceived by Dirac and made rigorous by Schwartz. An excellent account of this theory is given in [1], and [2] may also be consulted. In this theory, every generalized function has a Fourier transform which is itself a generalized function. One consequence of the theory is that it is possible to regard a Fourier series as a special case of a Fourier integral, as will be seen later. The results of Section 2.1 are summarized in Table 2.3.

2.2 FOURIER TRANSFORMS AND THEIR PROPERTIES

2.2.1 Well-behaved functions

As an example of the application of (2.1.24), consider the Fourier transform of the simple function $s(t) = e^{-|t|}$. Then

$$\begin{aligned} S(f) &= \int_{-\infty}^{\infty} e^{-|t|}\, e^{-j2\pi ft}\, dt \\ &= \frac{1}{1 - j2\pi f} + \frac{1}{1 + j2\pi f} = \frac{2}{1 + (2\pi f)^2}. \end{aligned}$$

Table 2.4 gives the Fourier transforms of some simple signals $s(t)$ which will be needed later.

These signals and their transforms are plotted in Figure 2.3. Remembering that $S(f)$ gives the distribution of signal strength with frequency, notice in Figure 2.3 that the first signal is quite smooth and hence its transform is dominated by low-frequency contributions. Also observe that sharp corners

TABLE 2.3: Summary of Fourier transforms

Description	Function	Transform	Inverse transform
Finite discrete series	s_r, $r = -n, \ldots, 0, 1, \ldots, n-1$	$S_m = \frac{1}{N}\sum_{r=-n}^{n-1} s_r\, e^{-j(2\pi mr/N)}$, $m = -n, \ldots, 0, 1, \ldots, n-1$	$\tilde{s}(t) = \sum_{m=-n}^{n-1} S_m\, e^{j(2\pi mt/N\Delta)}$, $-\infty \leqslant t \leqslant \infty$; $\tilde{s}(t=r\Delta) = s_r$, $\tilde{s}(t) = s(t+N\Delta)$
Continuous periodic function	$s(t) = s(t+T)$, $-\infty \leqslant t \leqslant \infty$	$S_m = \frac{1}{T}\int_{-T/2}^{T/2} s(t)\, e^{-j(2\pi mt/T)}\, dt$, $m = 0, \pm 1, \pm 2, \ldots$	$s(t) = \sum_{m=-\infty}^{\infty} S_m\, e^{j(2\pi mt/T)}$, $-\infty \leqslant t \leqslant \infty$; $s(t) = s(t+T)$
Continuous aperiodic function	$s(t)$, $-\infty \leqslant t \leqslant \infty$	$S(f) = \int_{-\infty}^{\infty} s(t)\, e^{-j2\pi ft}\, dt$, $-\infty \leqslant f \leqslant \infty$	$s(t) = \int_{-\infty}^{\infty} S(f)\, e^{j2\pi ft}\, df$, $-\infty \leqslant t \leqslant \infty$

in $s(t)$, as in example (b), produce ripples or side lobes in the transform and that periodicities in $s(t)$ appear as peaks in the transform, as illustrated by example (c).

FIG. 2.3: Some simple signals and their Fourier transforms

TABLE 2.4: Some simple functions and their Fourier transforms

$s(t)$	$S(f)$
$e^{-\|t\|} \quad (-\infty \leqslant t \leqslant \infty)$	$2/[1 + (2\pi f)^2]$
$\begin{cases} 0, & \|t\| > b \\ a, & \|t\| \leqslant b \end{cases}$	$2ab \dfrac{\sin 2\pi fb}{2\pi fb}$
$e^{-a\|t\|} \cos 2\pi f_0 t$	$\dfrac{a}{a^2 + [2\pi(f + f_0)]^2} + \dfrac{a}{a^2 + [2a(f - f_0)]^2}$

The signals in Table 2.4 are all even functions of t, and hence their Fourier transforms are real and even functions. In general this is not the case; for example, suppose $s(t)$ is the odd function

$$s(t) = \begin{cases} 0, & t < 0 \\ e^{-t}, & 0 \leqslant t \leqslant \infty. \end{cases}$$

Then, using (2.1.24),

$$S(f) = \frac{1}{1 + j2\pi f}.$$

This transform is complex and may be written as the sum of a real and imaginary part,

$$S(f) = \frac{1}{1 + (2\pi f)^2} - j \frac{2\pi f}{1 + (2\pi f)^2}.$$

Alternatively, it may be written in terms of an amplitude and phase function

$$S(f) = \frac{1}{1 + (2\pi f)^2} \exp(-j \arctan 2\pi f)$$

using (2.1.13), so that

$$R(f) = \frac{1}{\sqrt{1 + (2\pi f)^2}}, \qquad \phi(f) = \arctan -2\pi f.$$

Note that all these transforms damp out or "dissipate" as f tends to infinity. Situations will now be considered where the transform does not damp out.

2.2.2 *Generalized functions*

Consider two special cases of the rectangular pulse which was the second example in Table 2.4.

Unit height. If $a = 1$, then

$$S(f) = 2b \frac{\sin 2\pi fb}{2\pi fb}. \tag{2.2.1}$$

If b tends to infinity, $s(t)$ tends to a constant, equal to 1 everywhere. The behavior of $S(f)$ as b increases is illustrated in Figure 2.4, where it may be seen that $S(f)$ tends to a spike at $f = 0$. In the limit, $S(f)$ tends to a spike of infinite height at $f = 0$ and is bounded everywhere else. This is what is meant by a Dirac delta or impulse function. Hence *the Fourier transform of a constant is a delta function.*

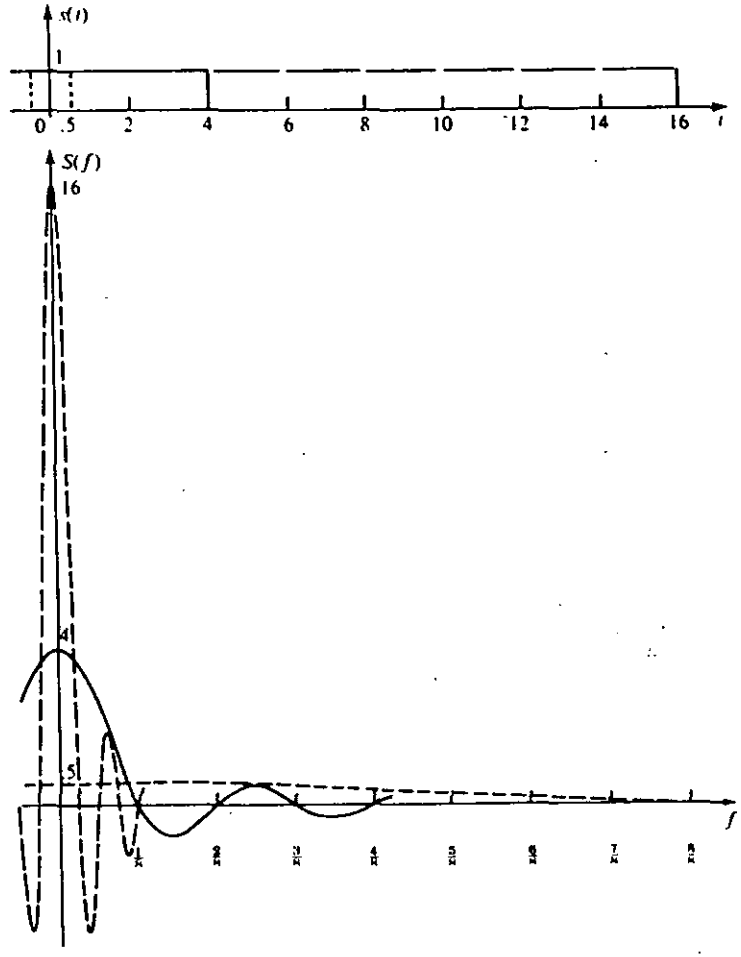


FIG. 2.4: Unit height rectangular pulses and their Fourier transforms

Unit area. If $2ab = 1$, then

$$S(f) = \frac{\sin 2\pi fb}{2\pi fb}. \tag{2.2.2}$$

As b tends to zero, $S(f)$ tends to unity everywhere. However, as b decreases $s(t)$ becomes taller and taller, as illustrated in Figure 2.5. It follows that $s(t)$ tends to a delta function at the origin.

These two cases show that the Fourier transform of a constant is a delta function and, conversely, that the Fourier transform of a delta function is a constant. This reciprocity is to be expected from the symmetry of the transform equations (2.1.22) and (2.1.24).

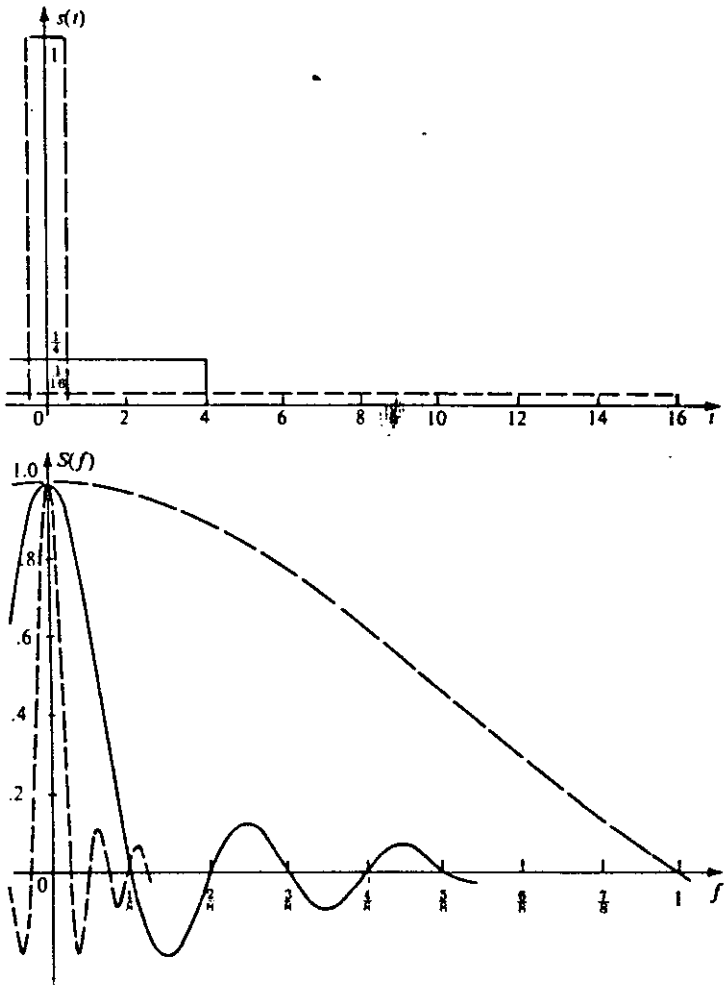


FIG. 2.5: Unit area rectangular pulses and their Fourier transforms

Delta functions. The sequence of functions (2.2.1) as b tends to infinity is not the only one which can be used to define a delta function. In general a delta function may be defined as a sequence of functions $\delta_n(t)$ such that

$$\int_{-\infty}^{\infty} \delta_n(t)\, dt = 1, \quad \text{for every } n, \tag{2.2.3}$$

and in the limit as n tends to ∞,

$$\delta(t) = \begin{cases} 0, & t \neq 0, \\ \infty, & t = 0. \end{cases} \tag{2.2.4}$$

Examples of such sequences of functions, together with their Fourier transforms, are given in Table 2.5. Note that $S_n(f)$ tends to a constant, unity, for all f as n tends to infinity.

TABLE 2.5: Sequences defining delta functions

	$\delta_n(t)$	$S_n(f)$
1.	$2n \dfrac{\sin 2\pi nt}{2\pi nt}$	$1, \quad \lvert f\rvert \leq n$ $0, \quad \lvert f\rvert > n$
2.	$\sqrt{n}\, e^{-\pi n t^2}$	$e^{-\pi f^2/n}$
3.	$\dfrac{n}{2} e^{-n\lvert t\rvert}$	$\dfrac{n^2}{n^2 + (2\pi f)^2}$
4.	$\dfrac{1}{\pi} \dfrac{n}{n^2 t^2 + 1}$	$e^{-\lvert 2\pi f/n\rvert}$
5.	$\dfrac{n \sin^2 (\pi nt)}{(\pi nt)^2}$	$1 - \dfrac{\lvert f\rvert}{n}, \quad \lvert f\rvert \leq n$ $0, \quad \lvert f\rvert > n$

One physical interpretation of the delta function is that of an input of energy into a system. Using an example from mechanics, suppose that a hard block is resting on a plane surface. If a very small but high-speed bullet is fired at the block, an exchange of energy occurs when the bullet rebounds. Assuming that the collision occurs so rapidly that the block does not have time to move, the bullet may be regarded as having imparted to the block an impulse of energy in the form of a change of momentum. An alternative interpretation drawn from electromagnetic theory is that of a unit point charge at the origin.

Unit area. If $2ab = 1$, then

$$S(f) = \frac{\sin 2\pi fb}{2\pi fb}. \qquad (2.2.2)$$

As b tends to zero, $S(f)$ tends to unity everywhere. However, as b decreases $s(t)$ becomes taller and taller, as illustrated in Figure 2.5. It follows that $s(t)$ tends to a delta function at the origin.

These two cases show that the Fourier transform of a constant is a delta function and; conversely, that the Fourier transform of a delta function is a constant. This reciprocity is to be expected from the symmetry of the transform equations (2.1.22) and (2.1.24).

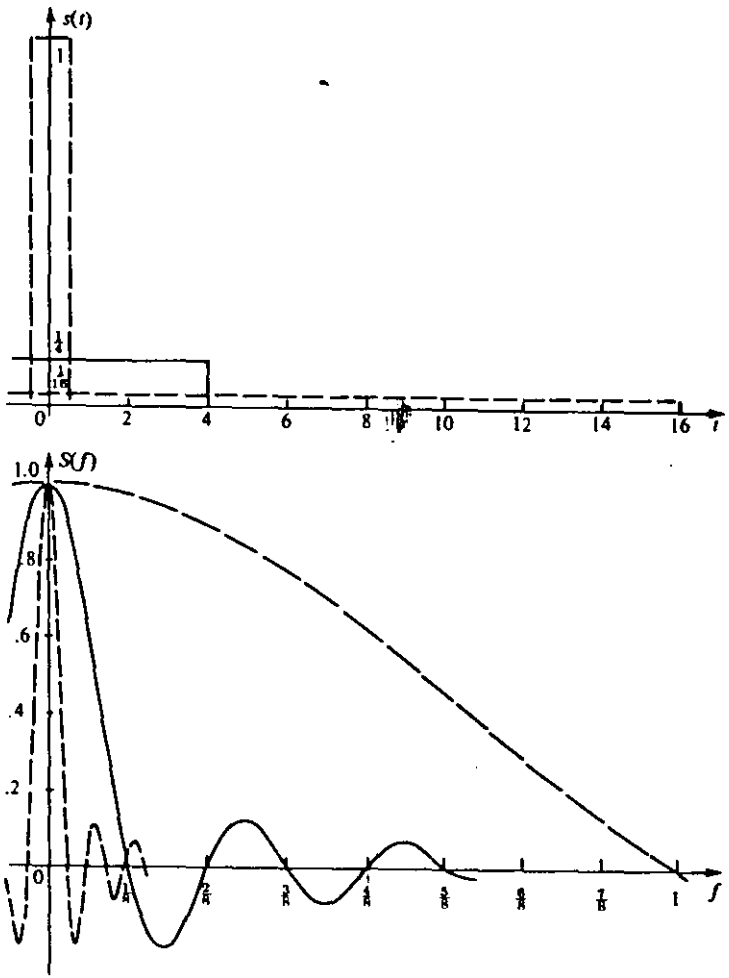


FIG. 2.5: Unit area rectangular pulses and their Fourier transforms

Delta functions. The sequence of functions (2.2.1) as b tends to infinity is not the only one which can be used to define a delta function. In general a delta function may be defined as a sequence of functions $\delta_n(t)$ such that

$$\int_{-\infty}^{\infty} \delta_n(t)\, dt = 1, \quad \text{for every } n, \qquad (2.2.3)$$

and in the limit as n tends to ∞,

$$\delta(t) = \begin{cases} 0, & t \neq 0, \\ \infty, & t = 0. \end{cases} \qquad (2.2.4)$$

Examples of such sequences of functions, together with their Fourier transforms, are given in Table 2.5. Note that $S_n(f)$ tends to a constant, unity, for all f as n tends to infinity.

TABLE 2.5: Sequences defining delta functions

	$\delta_n(t)$	$S_n(f)$
1.	$2n \dfrac{\sin 2\pi nt}{2\pi nt}$	$1, \quad \lvert f\rvert \leq n$ $0, \quad \lvert f\rvert > n$
2.	$\sqrt{n}\, e^{-\pi n t^2}$	$e^{-\pi f^2/n}$
3.	$\dfrac{n}{2} e^{-n\lvert t\rvert}$	$\dfrac{n^2}{n^2 + (2\pi f)^2}$
4.	$\dfrac{1}{\pi} \dfrac{n}{n^2 t^2 + 1}$	$e^{-\lvert 2\pi f/n\rvert}$
5.	$\dfrac{n \sin^2(\pi nt)}{(\pi nt)^2}$	$1 - \dfrac{\lvert f\rvert}{n}, \quad \lvert f\rvert \leq n$ $0, \quad \lvert f\rvert > n$

One physical interpretation of the delta function is that of an input of energy into a system. Using an example from mechanics, suppose that a hard block is resting on a plane surface. If a very small but high-speed bullet is fired at the block, an exchange of energy occurs when the bullet rebounds. Assuming that the collision occurs so rapidly that the block does not have time to move, the bullet may be regarded as having imparted to the block an impulse of energy in the form of a change of momentum. An alternative interpretation drawn from electromagnetic theory is that of a unit point charge at the origin.

The delta function may be used as an operational device for picking out the value of a signal at a given instant of time. This is contained in the result

$$\lim_{n\to\infty}\int_{-\infty}^{\infty}\delta_n(t-t_0)s(t)\,dt=\int_{-\infty}^{\infty}\delta(t-t_0)s(t)\,dt=s(t_0). \quad (2.2.5)$$

In a similar manner, by considering the limit of the sequence of mth derivatives of $\delta_n(t)$ [1], it is possible to define the mth derivative of a delta function, namely $\delta^{(m)}(t)$. This can be used to select the mth derivative of a function at a given point. This leads to a generalization of (2.2.5), namely

$$\int_{-\infty}^{\infty}\delta^{(m)}(t-t_0)s(t)\,dt=(-1)^m s^{(m)}(t_0). \quad (2.2.6)$$

Referring to the interpretation of a delta function as a unit charge at the origin, the first derivative $\delta'(t)$ corresponds to the mathematical idealization of a unit dipole. This is because the first moment of $\delta'(t)$ is

$$\int_{-\infty}^{\infty}t\,\delta'(t)\,dt=-1,$$

using (2.2.6). Hence, the moment of $\delta'(t)$ is unity, which is the standard definition for a unit dipole.

The unit step function. A function closely related to the delta function is the *unit step function*. This corresponds physically either to the application of a unit force which is then maintained or to a change in a valve position which alters the flow in a pipe by a unit amount. Mathematically this is a signal described by

$$U(t)=\begin{cases}0, & t<0,\\ \frac{1}{2}, & t=0,\\ 1, & t>0.\end{cases} \quad (2.2.7)$$

The function $U(t)$ may be regarded as the limit of a sequence of functions $U_n(t)$, for example, the limit as n tends to infinity of

$$U_n(t)=\begin{cases}\frac{1}{2}e^{nt}, & t<0,\\ 1-\frac{1}{2}e^{-nt}, & t\geqslant 0.\end{cases} \quad (2.2.8)$$

As n tends to infinity, $U_n(t)$ tends to zero for t negative and to unity for t positive. Differentiating $U_n(t)$ gives

$$\frac{d}{dt}U_n(t)=\frac{n}{2}e^{-n|t|}=\delta_n(t), \quad (2.2.9)$$

which illustrates the important result that the derivative of a step function is a delta function.

The Fourier transform of the unit step function (2.2.7) is

$$S(f)=\tfrac{1}{2}\delta(f)+\frac{1}{j2\pi f}.$$

2.2.3 Fourier series as Fourier transforms

Consider the Fourier transform of the signal

$$s_T(t)=\begin{cases}a\cos\dfrac{2\pi t}{\Delta}, & -\dfrac{T}{2}\leqslant t\leqslant\dfrac{T}{2},\\ 0, & |t|>\dfrac{T}{2},\end{cases} \quad (2.2.10)$$

which is a "periodic" signal in the interval $-T/2$ to $+T/2$. Using (2.1.24) directly, its Fourier transform is

$$S_T(f)=\frac{a}{2}\left\{T\frac{\sin\pi T[f-(1/\Delta)]}{\pi T[f-(1/\Delta)]}+T\frac{\sin\pi T[f+(1/\Delta)]}{\pi T[f+(1/\Delta)]}\right\}. \quad (2.2.11)$$

As T tends to infinity, the signal $s_T(t)$ becomes a truly periodic signal $s(t)$, periodic for all time, while the transform $S_T(f)$ tends to

$$S(f)=\frac{a}{2}\left\{\delta\left(f-\frac{1}{\Delta}\right)+\delta\left(f+\frac{1}{\Delta}\right)\right\}, \quad (2.2.12)$$

since each of the terms inside the braces is a sequence defining a delta function. Hence, the Fourier transform of a truly periodic (infinite extent) cosine wave consists of a delta function of area $a/2$ centered at $f=+(1/\Delta)$ and a delta function of area $a/2$ at $f=-(1/\Delta)$.

Similarly, the complex signal

$$s_T(t)=e^{j(2\pi mt/\Delta)},\quad -\frac{T}{2}\leqslant t\leqslant\frac{T}{2},$$

has the Fourier transform

$$S_T(f)=T\frac{\sin\pi T[f-(m/\Delta)]}{\pi T[f-(m/\Delta)]}.$$

Hence as T tends to infinity, $S_T(f)$ tends to $S(f)=\delta[f-(m/\Delta)]$. It follows that a periodic signal of period Δ with the Fourier series representation

$$s(t)=\sum_{m=-\infty}^{\infty}S_m\,e^{j(2\pi mt/\Delta)} \quad (2.2.13)$$

has the Fourier transform

$$S(f)=\sum_{m=-\infty}^{\infty}S_m\,\delta\left(f-\frac{m}{\Delta}\right), \quad (2.2.14)$$

which is a train of delta functions. Thus, by allowing generalized functions, Fourier series may be regarded as a special case of Fourier transforms.

To find the Fourier coefficients S_m corresponding to a generalized function, the classical formula (2.1.18) is no longer applicable since a generalized function cannot be integrated between finite limits. The appropriate formula to use in these circumstances is given in [1].

In particular, it may be shown that the Fourier transform of a train of delta functions

$$s(t) = \sum_{n=-\infty}^{\infty} \delta(t - n\Delta) \tag{2.2.15}$$

is

$$S(f) = \frac{1}{\Delta} \sum_{n=-\infty}^{\infty} \delta\left(f - \frac{n}{\Delta}\right). \tag{2.2.16}$$

Hence a train of delta functions transforms into a train of delta functions. Note that this result is symmetrical in the time and frequency domain.

The train of delta functions is not the only function which has a symmetric transformation. A much simpler function possessing this property is given by example 2 of Table 2.5 with $n = 1$. Thus, $s(t) = \exp(-\pi t^2)$ transforms to $S(f) = \exp(-\pi f^2)$.

At this point, the reader should satisfy himself that he is familiar with the various operational properties of Fourier transforms summarized in Appendix A2.1.

2.3 LINEAR SYSTEMS AND CONVOLUTION

2.3.1 Linear differential equations

One important reason for the usefulness of Fourier and spectral analysis is that they simplify the analysis of *time invariant linear systems*, that is, systems whose behavior can be described by linear integro-differential equations with constant coefficients. It may be shown in general [3] that the solution to such an equation may be written as a convolution integral

$$y(t) = \int_{-\infty}^{\infty} h(u)x(t - u)\,du, \tag{2.3.1}$$

where $y(t)$ is the solution and $x(t)$ is the forcing function. It will be shown in Section 2.3.4 that the solution is simplified by Fourier transforms. The transform of the solution becomes

$$Y(f) = H(f)X(f),$$

where $Y(f)$, $H(f)$ and $X(f)$ are the Fourier transforms of $y(t)$, $h(t)$ and $x(t)$ respectively. Thus, convolution in the time domain transforms to multiplication in the frequency domain.

An illustration of convolution. As an example of a convolution integral, consider a simple linear system consisting of a spring and a dashpot, shown schematically in Figure 2.6. One use of such a device is to prevent screen doors from closing too violently when released. A force applied to the spring produces an *input* displacement $x(t)$ which causes an *output* displacement $y(t)$ of the dashpot wiper. The differential equation obtained by equating forces is then

$$K[x(t) - y(t)] = D\frac{dy}{dt},$$

where K is the spring constant in pounds per foot and D is the velocity constant of the dashpot (in pounds per foot per second). Rearranging this equation gives

$$T\frac{dy}{dt} + y(t) = x(t), \tag{2.3.2}$$

where $T = D/K$ is the *time constant* of the system (in seconds).

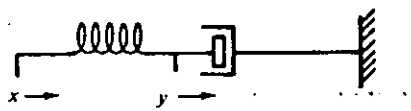


FIG. 2.6: A first-order mechanical system

Equation (2.3.2) may be used to describe the behavior of many other physical systems, for example, the temperature $y(t)$ in the outlet from a chemical reactor when the inlet temperature is $x(t)$. Equation (2.3.2) then shows that the rate of change of temperature in the outlet is directly proportional to the temperature gradient across the reactor.

The solution to equation (2.3.2) can be written as a convolution integral by introducing the integrating factor $e^{t/T}$. Thus

$$\begin{aligned} y(t) &= \int_{-\infty}^{t} x(u)\frac{e^{-(t-u)/T}}{T}\,du \\ &= \int_{-\infty}^{\infty} x(u)h(t - u)\,du, \end{aligned} \tag{2.3.3}$$

where

$$h(u) = \begin{cases} \frac{1}{T}e^{-u/T}, & u \geqslant 0, \\ 0, & u < 0. \end{cases}$$

Hence the output $y(t)$ may be written as a *weighted sum* of past values of the input $x(t)$, that is, the output is a *convolution* of the input with the *weighting function* $h(u)$.

It may be shown in general [3] that the solution to any linear time-invariant differential equation may be written as in (2.3.3) or, by making a change in variable, as

$$y(t) = \int_{-\infty}^{\infty} h(u)x(t - u)\,du. \tag{2.3.4}$$

The weighting function completely characterizes the behavior of the system, just as the differential equation does.

Time-invariant linear systems. Equations (2.3.3) and (2.3.4) are general representations of what are known as *time-invariant linear systems* or filters. These are characterized by the following properties:

(a) *Linear property*: If $x_1(t)$ and $x_2(t)$ are two inputs to the system, and $y_1(t)$, $y_2(t)$ the corresponding outputs, then a linear compound $\mu_1 x_1(t) + \mu_2 x_2(t)$ of the inputs produces the *same* linear compound of the outputs, $\mu_1 y_1(t) + \mu_2 y_2(t)$.

(b) *Time-invariant property*: If the input $x(t)$ is delayed by an amount τ to give $x(t - \tau)$, the output is delayed by the same amount and is $y(t - \tau)$.

It is property (b) which ensures that the weight function $h(u)$ is independent of time. A linear system without the time-invariant property would have a weight function which depended on time t.

Systems which can be described by means of linear differential equations with constant coefficients can be shown to have the time-invariant representation (2.3.3). However, many non-linear systems can be *linearized* so that, for small perturbations in the input, (2.3.3) can be used as an approximate representation of the system.

2.3.2 Step and impulse functions

For any physical system the weight function $h(u)$ must be zero for negative values of u, which means that the system cannot respond to inputs it has not yet received. This is called the condition of physical realizability. For physically realizable systems, equations (2.3.3) and (2.3.4) may be written

$$y(t) = \int_{0}^{\infty} h(u)x(t - u)\,du, \tag{2.3.5}$$

or

$$y(t) = \int_{-\infty}^{t} x(u)h(t - u)\,du. \tag{2.3.6}$$

Impulse response functions. Suppose that the system is given a sharp impulse at time $t = 0$ so that $x(t) = \delta(t)$. Then

$$y(t) = \int_{-\infty}^{\infty} h(u)\,\delta(t - u)\,du, \tag{2.3.7}$$

which equals $h(t)$ using (2.2.5). The weight function $h(t)$ is called the *impulse response* function [4] of the system since it measures the output at time t of a system subjected to an impulse at $t = 0$.

The impulse responses for a number of simple systems are given in the first column of Table 2.6. Figure 2.7 shows the impulse responses for three of these systems. The first example (a) is a system with simple delay for which the output or impulse response is another impulse at time τ later. The second (b) is a system described by a single time-constant and represented by the differential equation (2.3.2) for which the impulse response is the exponential curve of Figure 2.7(b). The third example (c) is a second-order system represented by the differential equation

$$\frac{1}{\omega_n^2}\frac{d^2y}{dt^2} + \frac{2\zeta}{\omega_n}\frac{dy}{dt} + y = x(t), \tag{2.3.8}$$

for which the impulse response is a damped sine wave as in Figure 2.7(c).

Step response functions. It is also possible to characterize a linear system by its response to the unit step function (2.2.7). Suppose, for example, that the input is the flow rate of cold water into a heat exchanger and the output the outlet temperature. Then the step response is the variation of outlet temperature with time when a unit change is made to the input flow rate. From (2.3.5), the response at time t to a unit step at time $t = 0$ is

$$y(t) = \int_{0}^{t} h(u)\,du, \tag{2.3.9}$$

so the step response is the integral of the impulse response.

From Figure 2.7 it may be seen that the step response to a system with a pure delay τ is another step starting at time τ later, as shown in (a). For the exponential impulse response, the step response builds up exponentially to its ultimate value, as in (b). For the second-order system (c), the step response overshoots its ultimate value and then oscillates about this value with decreasing amplitude.

As t tends to infinity, the step response (2.3.9) tends to its limiting or ultimate value

$$g = \int_{0}^{\infty} h(u)\,du, \tag{2.3.10}$$

which is called the *steady state gain* of the system, since it measures the ultimate amplification of a unit step change after the system has been allowed to settle down to its new steady state value.

Stability. A system is said to be *stable* [4] if a bounded input produces a bounded output. Clearly this is desirable, since otherwise output would

TABLE 2.6: Impulse, step, and frequency responses for some simple systems

System	Impulse response $h(t)$	Step response $y(t)$	Frequency response $H(f)$	Gain $G(f)$	Phase $\phi(f)$
1. gain	$g\,\delta(t)$	g	g	g	0
2. delay	$\delta(t-\tau)$	$0,\ t<\tau$ $1,\ t\geq\tau$	$e^{-j2\pi f\tau}$	1	$-2\pi f\tau$
3. integration	1	t	$\frac{1}{j2\pi f}$	$\frac{1}{2\pi f}$	$-\frac{\pi}{2}$
4. single exponential	$\frac{1}{T}e^{-t/T}$	$(1-e^{-t/T})$	$\frac{1}{1+j2\pi fT}$	$\frac{1}{\{1+(2\pi fT)^2\}^{1/2}}$	$-\arctan 2\pi fT$
5. single exponential plus delay	$0,\ t<\tau$ $\frac{1}{T}e^{-(t-\tau)/T},\ t\geq\tau$	$0,\ t<\tau$ $(1-e^{-(t-\tau)/T}),\ t\geq\tau$	$\frac{e^{-j2\pi f\tau}}{(1+j2\pi fT)}$	$\frac{1}{\{1+(2\pi fT)^2\}^{1/2}}$	$-2\pi f\tau-\arctan 2\pi fT$
6. two time constants	$\frac{e^{-t/T_1}-e^{-t/T_2}}{T_1-T_2}$	$\left(1-\frac{T_1e^{-t/T_1}-T_2e^{-t/T_2}}{T_1-T_2}\right)$	$\frac{1}{(1+j2\pi fT_1)(1+j2\pi fT_2)}$	$\frac{1}{\{[1+(2\pi fT_1)^2][1+(2\pi fT_2)^2]\}^{1/2}}$	$-\arctan 2\pi fT_1$ $-\arctan 2\pi fT_2$
7. quadratic lag	$\frac{\omega_n e^{-t\omega_n t}\sin(\omega_n\sqrt{1-\zeta^2}\,t)}{\sqrt{1-\zeta^2}}$	$1-\frac{e^{-t\omega_n t}\sin(\omega_n\sqrt{1-\zeta^2}\,t+\phi)}{\sqrt{1-\zeta^2}}$ $\sin\phi=\sqrt{1-\zeta^2}$	$\frac{1}{1-(f/f_n)^2+j2\zeta(f/f_n)}$ $\omega_n=2\pi f_n$	$\frac{1}{\{[1-(f/f_n)^2]^2+[2(\zeta f/f_n)]^2\}^{1/2}}$	$-\arctan\frac{2\zeta(f/f_n)}{[1-(f/f_n)^2]}$
8. differentiation	$\delta'(t)$	$-\delta(t)$	$j2\pi f$	$2\pi f$	$\frac{\pi}{2}$
9. linear lead	$\delta(t)+T\,\delta'(t)$	$1-T\,\delta(t)$	$1+j2\pi fT$	$\{1+(2\pi fT)^2\}^{1/2}$	$\arctan 2\pi fT$

increase without limit. Suppose that $|x(t)|\leq K_1$ in (2.3.5), where K_1 is a finite constant. Then

$$|y(t)| = \left|\int_{-\infty}^{\infty} h(u)x(t-u)\,du\right| \leq \int_{-\infty}^{\infty} |h(u)|\,|x(t-u)|\,du$$
$$\leq K_1\int_{-\infty}^{\infty}|h(u)|\,du,$$

so a sufficient condition for the system to be stable is that

$$\int_{-\infty}^{\infty}|h(u)|\,du < K_2, \qquad (2.3.11)$$

where K_2 is also a finite constant. An alternative form for the stability condition will be given in the next section.

2.3.3 *Frequency response functions*

For inputs more complicated than an impulse or a step, the calculation of the output by means of the convolution integral (2.3.5) is tedious. This problem is considerably simplified by making use of Fourier analysis. The approach consists of splitting up $s(t)$ into its Fourier components $S(f)$, as in (2.1.24), finding the response of the system to a periodic signal $s_1(t) = e^{j2\pi ft}$, and then adding up the responses according to (2.1.22) to give

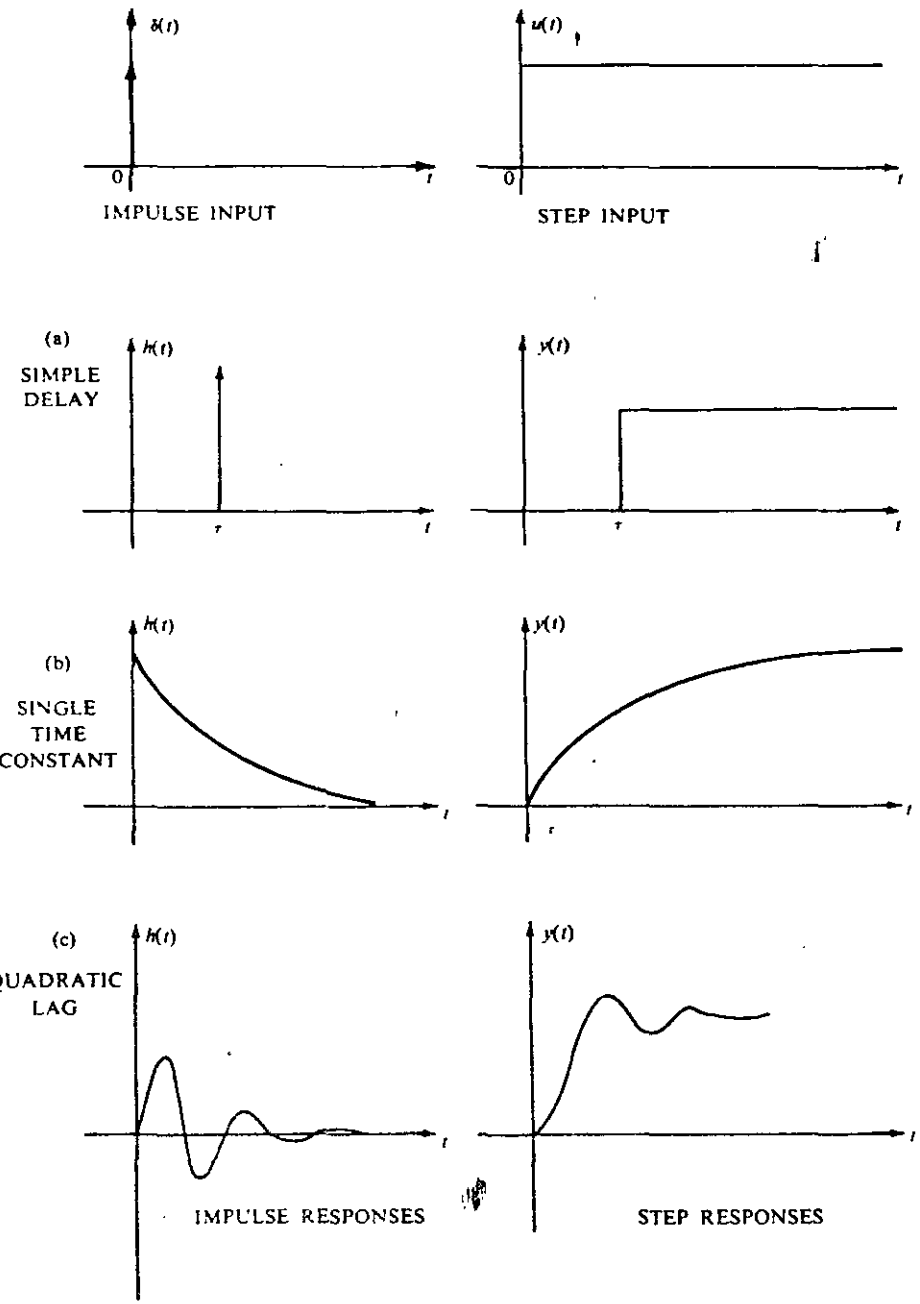


FIG. 2.7: Impulse and step responses for some simple systems

the final output. First it is necessary to know the response of the system to an input $x(t) = \cos 2\pi ft$. Substitution of this input in (2.3.5) yields

$$\begin{aligned} y(t) &= \int_0^\infty h(u) \cos 2\pi f(t-u)\, du \\ &= \int_0^\infty h(u) [\cos 2\pi ft \cos 2\pi fu + \sin 2\pi ft \sin 2\pi fu]\, du \\ &= A(f) \cos 2\pi ft + B(f) \sin 2\pi ft, \end{aligned} \tag{2.3.12}$$

where

$$A(f) = \int_0^\infty h(u) \cos 2\pi fu\, du \tag{2.3.13}$$

and

$$B(f) = \int_0^\infty h(u) \sin 2\pi fu\, du. \tag{2.3.14}$$

Alternatively, (2.3.12) may be rewritten

$$y(t) = G(f) \cos [2\pi ft + \phi(f)], \tag{2.3.15}$$

where

$$G(f) = \sqrt{A^2(f) + B^2(f)}$$

and

$$\phi(f) = \arctan -\frac{B(f)}{A(f)}.$$

Hence the response to a cosine wave of frequency f is a cosine wave at the same frequency but scaled in amplitude by a factor $G(f)$ called the *gain* and shifted in phase by an amount $\phi(f)$ called the *phase angle.*

As before, it is operationally more convenient to consider the response to a complex input

$$e^{j2\pi ft} = \cos 2\pi ft + j \sin 2\pi ft,$$

of frequency f. In this case the output is

$$y(t) = H(f)\, e^{j2\pi ft} = G(f)\, e^{j[2\pi ft + \phi(f)]}, \tag{2.3.16}$$

where

$$H(f) = G(f)\, e^{j\phi(f)} = \int_0^\infty h(u)\, e^{-j2\pi fu}\, du \tag{2.3.17}$$

is called the *frequency response function* of the system. Hence the frequency response function is the Fourier transform of the impulse response function.

Bode plots. The frequency response functions, gains and phases of some simple systems are given in Table 2.6 and the gains and phases are plotted in Figure 2.8. It is customary to plot the logarithm of gain against the logarithm of

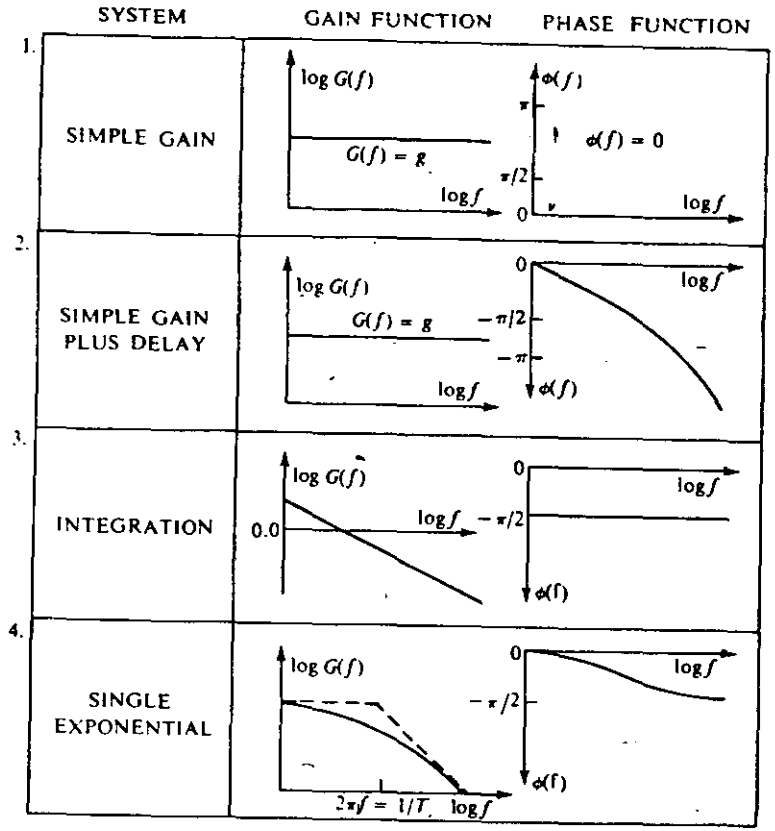


FIG. 2.8: Bode plots for some simple systems

frequency and the phase against the logarithm of frequency. These graphs are usually called *Bode plots* [5]. The diagrams in Figure 2.8 fall naturally into four categories:

(1) Numbers 1 and 2 have a gain which is constant with frequency and are called *all-pass* systems.

(2) Numbers 3, 4, 5 and 6 are such that high frequencies are rejected or attenuated by the system and low frequencies transmitted with different gains. These systems therefore behave as low pass filters and are associated with some form of integration or smoothing of the input.

(3) Number 7 corresponds to an oscillatory system, described by equation (2.3.8). Here the gain plot has a *resonance* or peak at a frequency $f = f_n(1 - 2\zeta^2)^{1/2}$, where f_n is the natural resonant frequency of the system.

(4) Numbers 8 and 9 have gain plots which are such that the lower frequencies are attenuated and the higher frequencies passed. These systems act

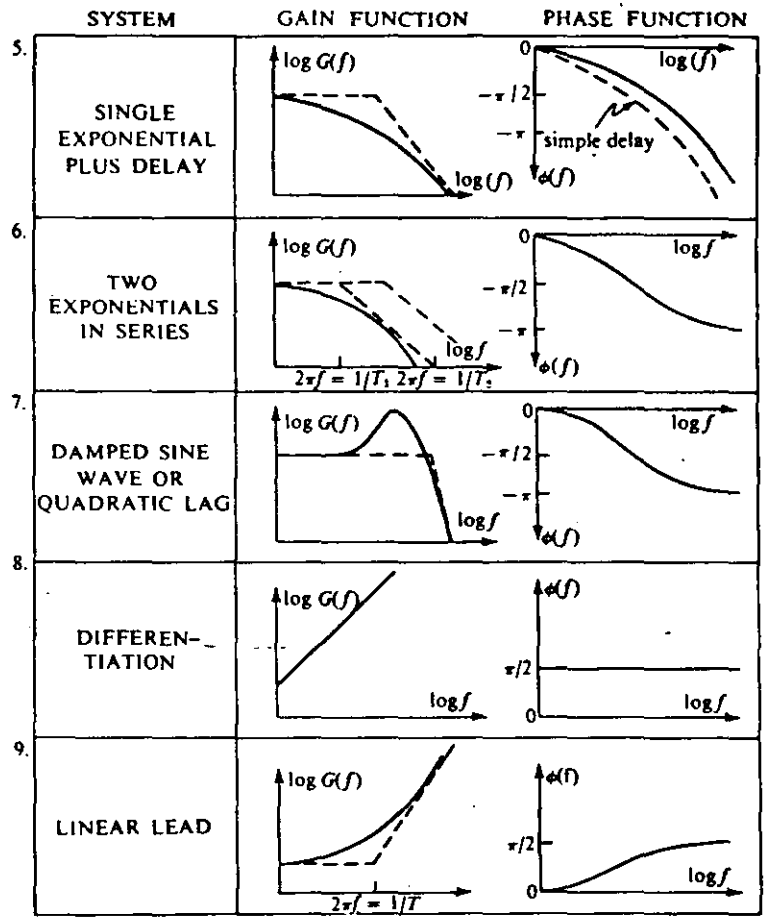


FIG. 2.8 (cont.): Bode plots for some simple systems

as *high pass filters* and involve differentiation of the input. A further difference between categories *(2)* and *(4)* is that in *(2)* integration of the input results in negative phases $\phi(f)$, that is, the output *lags* behind the input. On the other hand, in *(4)* differentiation of the input yields positive phases, so the output *leads* the input as in plot number 9.

Bandwidth. A convenient way of describing the gain function of a linear system is by its *bandwidth* [5]. Various definitions of bandwidth have been suggested, the simplest of which is the half-power bandwidth. For a system which has its maximum gain at f_0, this is defined as the frequency difference $f_2 - f_1$, where f_1 and f_2 are chosen so that

$$G^2(f_1) = G^2(f_2) = \tfrac{1}{2}G^2(f_0).$$

For example, the maximum gain occurs at $f_0 = 0$ for the single exponential system, and the half-power gain occurs at $f = 1/(2\pi T)$. Hence if T is large the bandwidth is very small, as may be seen from Figure 2.8. The impulse response will thus be very wide, and small in amplitude. On the other hand, for T small the bandwidth is large and the impulse response very tall and narrow. In the limiting case as T tends to zero, the bandwidth becomes infinite, as for the simple gain of Figure 2.8, and the impulse response tends to a delta function. Hence wide bandwidths are associated with narrow impulse response functions and, conversely, narrow bandwidths are associated with wide impulse response functions.

Stability. The systems of Table 2.6 can be represented by a differential equation of the general form

$$a_m \frac{d^m y(t)}{dt^m} + \cdots + a_1 \frac{dy(t)}{dt} + a_0 y(t) = b_n \frac{d^n x(t-\tau)}{dt^n} + \cdots + b_1 \frac{dx(t-\tau)}{dt} + b_0 x(t-\tau). \quad (2.3.18)$$

Substituting $x(t) = e^{j2\pi ft}$, $y(t) = H(f)\, e^{j2\pi ft}$ in (2.3.18), it follows that the frequency response function is

$$H(f) = \frac{b_n(j2\pi f)^n + \cdots + b_1(j2\pi f) + b_0}{a_m(j2\pi f)^m + \cdots + a_1(j2\pi f) + a_0} \cdot e^{-j2\pi f\tau}. \quad (2.3.19)$$

Substituting $p = j2\pi f$ in (2.3.19) and equating the denominator to zero gives the *characteristic equation* of the system, namely

$$a_m p^m + \cdots + a_1 p + a_0 = 0. \quad (2.3.20)$$

The condition (2.3.11) that the system be stable may be shown [4] to be equivalent to the condition that the roots $\pi_1, \pi_2, \ldots, \pi_m$ of the characteristic equation (2.3.20) have negative real parts.

2.3.4 *Response to an arbitrary input*

If it is known that the response of a system to an input $x(t) = e^{j2\pi ft}$ is $y(t) = H(f)\, e^{j2\pi ft}$, it is possible to find the response to an arbitrary input. The first step is to Fourier transform the input to give

$$X(f) = \int_{-\infty}^{\infty} x(t)\, e^{-j2\pi ft}\, dt. \quad (2.3.21)$$

The Fourier component of the output at frequency f is

$$Y(f) = \int_{-\infty}^{\infty} e^{-j2\pi ft} \left[\int_{-\infty}^{\infty} h(u)x(t-u)\, du \right] dt = \int_{-\infty}^{\infty} h(v)\, e^{-j2\pi fv}\, dv \int_{-\infty}^{\infty} x(v)\, e^{-j2\pi fv}\, dv, \quad (2.3.22)$$

where $v = t - u$, that is,

$$Y(f) = H(f)X(f). \quad (2.3.23)$$

Equation (2.3.23) states that the component at frequency f in the output is obtained by multiplying the component at the *same* frequency in the input by $H(f)$, the frequency response function at that frequency. Finally, to recover $y(t)$, it is necessary to synthesize or add up the contributions from all frequencies at the same value of t, which gives

$$y(t) = \int_{-\infty}^{\infty} Y(f)\, e^{j2\pi ft}\, df = \int_{-\infty}^{\infty} X(f)H(f)\, e^{j2\pi ft}\, df. \quad (2.3.24)$$

Equations (2.3.22) to (2.3.24) show that *convolution* in the time domain is equivalent to *multiplication* in the frequency domain. Hence, if a relationship between two variables exists in the form of a differential equation (2.3.18), the solution is (2.3.24), where the frequency response function is given by (2.3.19). Hence, the Fourier transform provides an extremely useful operational method for solving linear differential equations.

The solution may be expedited by using tables of transforms. A table of transforms of generalized functions is given in [1] while Fourier transforms of more common functions are given in [6].

Several linear systems in series. Consider k non-interacting linear systems in series, as shown in Figure 2.9. Repeated use of (2.3.23) gives

$$Y(f) = H_k(f)\, H_{k-1}(f) \cdots H_1(f)\, X(f), \quad (2.3.25)$$

which shows that for linear systems in series the overall frequency response function is the product of the frequency response functions of the individual

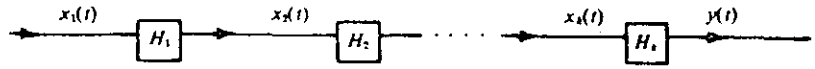


FIG. 2.9: Several linear systems in series

systems. Using (2.3.17) it follows that the overall gain is the *product* of the individual gains

$$G(f) = G_1(f)\, G_2(f) \cdots G_k(f), \quad (2.3.26)$$

and the overall phase shift is the *sum* of the individual phase shifts

$$\phi(f) = \phi_1(f) + \phi_2(f) + \cdots + \phi_k(f). \quad (2.3.27)$$

The output of the system may now be computed by adding up the contributions at all frequencies in the form

$$y(t) = \int_{-\infty}^{\infty} H_1(f)\, H_2(f) \cdots H_k(f)\, X(f)\, e^{j2\pi ft}\, df. \quad (2.3.28)$$

Note that only one integration is required, whereas the time-domain calculation would have required the evaluation of k convolution integrals.

2.3.5 Linear difference equations

In the preceding sections, it was shown that a system described by a linear differential equation may also be described by means of an impulse response function $h(u)$ or a frequency response function $H(f)$, where $h(u)$ and $H(f)$ form a Fourier transform pair. The functions $h(u)$ and $H(f)$ are easily obtained from the differential equation describing the system. In this section it is shown how impulse and frequency response functions may be used to describe a system which is characterized by a linear *difference* equation.

A linear difference equation is an equation of the form

$$y_r = \alpha_1 y_{r-1} + \alpha_2 y_{r-2} + \cdots + \alpha_m y_{r-m} + \beta_0 x_r + \cdots + \beta_n x_{r-n}, \quad (2.3.29)$$

and has a general solution

$$y_r = \sum_{k=0}^{\infty} h_k x_{r-k}. \quad (2.3.30)$$

The quantities $y_r, y_{r-1}, \ldots, y_{r-m}$ and $x_r, x_{r-1}, \ldots, x_{r-n}$ could be values of continuous signals $y(t)$ and $x(t)$ at the instants $t = r\Delta$, $(r-1)\Delta, \ldots, (r-m)\Delta$, $(r-n)\Delta$ respectively, that is,

$$y(t) = \alpha_1 y(t-\Delta) + \alpha_2 y(t-2\Delta) + \cdots + \alpha_m y(t-m\Delta) + \beta_0 x(t) + \cdots + \beta_n x(t-n\Delta). \quad (2.3.31)$$

The Fourier transform of (2.3.31) may be arranged in the form

$$Y(f) = \frac{\beta_0 + \beta_1 e^{-j2\pi f\Delta} + \cdots + \beta_n e^{-j2\pi fn\Delta}}{1 - \alpha_1 e^{-j2\pi f\Delta} - \cdots - \alpha_m e^{-j2\pi fm\Delta}} X(f),$$

so that the frequency response function $H(f)$ of the system is, from (2.3.23),

$$H(f) = \frac{\beta_0 + \beta_1 e^{-j2\pi f\Delta} + \cdots + \beta_n e^{-j2\pi fn\Delta}}{1 - \alpha_1 e^{-j2\pi f\Delta} - \cdots - \alpha_m e^{-j2\pi fm\Delta}}. \quad (2.3.32)$$

The frequency response function $H(f)$ and the discrete impulse response function h_k are related by

$$H(f) = \sum_{k=0}^{\infty} h_k e^{-j2\pi fk\Delta} \quad (2.3.33)$$

and

$$h_k = \Delta \int_{-1/(2\Delta)}^{1/(2\Delta)} H(f) e^{j2\pi fk\Delta}\, df. \quad (2.3.34)$$

$\mathcal{Z}$ transforms. The frequency response function (2.3.32) is best handled by making a substitution of the form $\mathcal{Z} = e^{j2\pi f\Delta}$ yielding

$$H(\mathcal{Z}) = \frac{\beta_0 + \beta_1 \mathcal{Z}^{-1} + \cdots + \beta_n \mathcal{Z}^{-n}}{1 - \alpha_1 \mathcal{Z}^{-1} - \cdots - \alpha_m \mathcal{Z}^{-m}} = \sum_{k=0}^{\infty} h_k \mathcal{Z}^{-k}, \quad (2.3.35)$$

which is the $\mathcal{Z}$ transform [7] of the impulse response function h_k.

From an *operational* point of view, $\mathcal{Z}$ in (2.3.35) may be regarded as a *shift operator* with the property

$$\mathcal{Z}^{-k} x_r = x_{r-k}. \quad (2.3.36)$$

Hence the difference equation (2.3.29) may be written

$$(1 - \alpha_1 \mathcal{Z}^{-1} - \alpha_2 \mathcal{Z}^{-2} - \cdots - \alpha_m \mathcal{Z}^{-m}) y_r = (\beta_0 + \beta_1 \mathcal{Z}^{-1} + \cdots + \beta_n \mathcal{Z}^{-n}) x_r, \quad (2.3.37)$$

that is,

$$y_r = \frac{(\beta_0 + \beta_1 \mathcal{Z}^{-1} + \cdots + \beta_n \mathcal{Z}^{-n})}{(1 - \alpha_1 \mathcal{Z}^{-1} - \cdots - \alpha_m \mathcal{Z}^{-m})} x_r = H(\mathcal{Z}) x_r,$$

where $H(\mathcal{Z})$ is the *transfer function* of the discrete system. Expanding $H(\mathcal{Z})$ in powers of $\mathcal{Z}^{-1}$ gives

$$y_r = \sum_{k=0}^{\infty} h_k \mathcal{Z}^{-k} x_r = \sum_{k=0}^{\infty} h_k x_{r-k},$$

which is the general solution (2.3.30).

Stability. Factoring $\mathcal{Z}^{-m}$, substituting $p = \mathcal{Z}$ and equating the denominator of (2.3.35) to zero gives the characteristic equation of the discrete system

$$p^m - \alpha_1 p^{m-1} - \cdots - \alpha_m = 0. \quad (2.3.38)$$

The stability condition corresponding to (2.3.11) is

$$\sum_{k=0}^{\infty} |h_k| < K_2. \quad (2.3.39)$$

Similarly, the stability condition corresponding to (2.3.20) is that the roots $\pi_1, \ldots, \pi_m$ of the characteristic equation (2.3.38) lie inside the unit circle.

Example. Consider the second-order difference equation

$$y_r = \alpha_1 y_{r-1} + \alpha_2 y_{r-2} + x_r. \quad (2.3.40)$$

This has the $\mathcal{Z}$ transform

$$(1 - \alpha_1 \mathcal{Z}^{-1} - \alpha_2 \mathcal{Z}^{-2}) y_r = x_r,$$

and hence the transfer function

$$H(\mathcal{Z}) = \frac{1}{1 - \alpha_1 \mathcal{Z}^{-1} - \alpha_2 \mathcal{Z}^{-2}}. \quad (2.3.41)$$

The characteristic equation is

$$p^2 - \alpha_1 p - \alpha_2 = 0,$$

which has roots

$$\pi_1 = \frac{\alpha_1 - \sqrt{\alpha_1^2 + 4\alpha_2}}{2}, \qquad \pi_2 = \frac{\alpha_1 + \sqrt{\alpha_1^2 + 4\alpha_2}}{2}. \tag{2.3.42}$$

The impulse response function for this system takes the form

$$h_k = \frac{1}{\pi_1 - \pi_2}\{\pi_1^{k+1} - \pi_2^{k+1}\} \tag{2.3.43}$$

when the roots are real, that is, when $\alpha_1^2 \geqslant -4\alpha_2$. When the roots are complex, that is, $\alpha_1^2 < -4\alpha_2$,

$$h_k = R^k \frac{\sin 2\pi f_0(k+1)}{\sin 2\pi f_0}, \tag{2.3.44}$$

where

$$\pi_1 = R\, e^{i2\pi f_0},$$
$$\pi_2 = R\, e^{-i2\pi f_0}.$$

The system is stable provided $|\pi_1| < 1$, $|\pi_2| < 1$, that is, provided α_1 and α_2 lie in the triangular region

$$\begin{aligned} \alpha_1 + \alpha_2 &< 1, \\ \alpha_1 - \alpha_2 &> -1, \\ -1 < \alpha_2 &< 1. \end{aligned} \tag{2.3.45}$$

2.4 APPLICATIONS TO TIME SERIES ANALYSIS

2.4.1 *Finite-length records*

In practice it is only possible to obtain finite lengths of records. The statistical questions to be discussed later stem from the fact that it is necessary to estimate the accuracy of various functions obtained from finite amounts of data. Even if $s(t)$ is a deterministic function, a bias or truncation error arises if $s(t)$ is only known in a finite interval $-T/2 \leqslant t \leqslant T/2$. To see the effect of this truncation, consider the *data window* defined by

$$w(t) = \begin{cases} 1, & |t| \leqslant \dfrac{T}{2} \\ 0, & |t| > \dfrac{T}{2}. \end{cases} \tag{2.4.1}$$

If $s(t)$ is a deterministic signal in the range $-\infty \leqslant t \leqslant \infty$, the signal actually measured in the finite interval may be written

$$s_T(t) = s(t)w(t). \tag{2.4.2}$$

Thus, the operation of taking a finite length record is equivalent to multiplying the actual signal $s(t)$ by the data window $w(t)$. Using (A2.1.8), it follows that the finite-interval transform $S_T(f)$ is the convolution of the transforms of $s(t)$ and $w(t)$,

$$S_T(f) = \int_{-\infty}^{\infty} S(g)W(f-g)\,dg, \tag{2.4.3}$$

where the *spectral window* $W(f)$ is the Fourier transform of the data window $w(t)$ and in this case is

$$W(f) = T\frac{\sin \pi fT}{\pi fT}. \tag{2.4.4}$$

It is not necessary to restrict the data window to the form (2.4.1). Any reasonable data window $w(t)$ will produce a spectral window $W(f)$ which is concentrated about $f = 0$ but with side lobes or minor peaks which damp out as f gets further away from zero. For small T, $S_T(f)$ may give a very distorted picture of $S(f)$ since the window $W(f-g)$ will be wide and hence values of $S(g)$ far removed from $g = f$ will contribute to $S_T(f)$ in the integral (2.4.3). As T becomes large the distortion will be reduced. Finally, as T tends to infinity, the transform component at frequency f can be fully determined, since data windows will tend to the generalized function 1 as T tends to infinity. Hence, as T tends to infinity, $W(f-g)$ tends to a delta function centered at $g = f$ and so $S_T(f)$ tends to $S(f)$.

The effect of window shape and width on the measured transform is illustrated in Figure 2.10 for a particular input $s(t)$ whose Fourier transform consists of three delta functions centered at f_0, f_1 and f_2. Note that:

(1) Only two main peaks appear in the output transform for windows (a), (c) and (d) because the two input peaks at frequencies f_1 and f_2 are fused into one. This is the result of using a data window which is too narrow.

(2) The output transforms for windows (a) and (b) have several false peaks between the main real peaks. The false peaks are caused by the sharp edges of the data window.

(3) The ability to pick out peaks (resolvability) depends on the data window *width*, as is illustrated by the output transforms for windows (a) and (b), which are of the same shape, but have different widths.

(4) The ability to pick out peaks also depends on the data window *shape*, as is illustrated by the output transforms for data windows (b), (c) and (d), which are of the same width but have different shapes. It will be shown in Chapters 6 and 7 that the window width and shape produce similar effects in spectral analysis.

In Figure 2.10, the frequency spacing $(f_2 - f_1)$ was chosen to be $1/T$. Figure 2.10(a) shows that with a rectangular data window of length T it is not possible

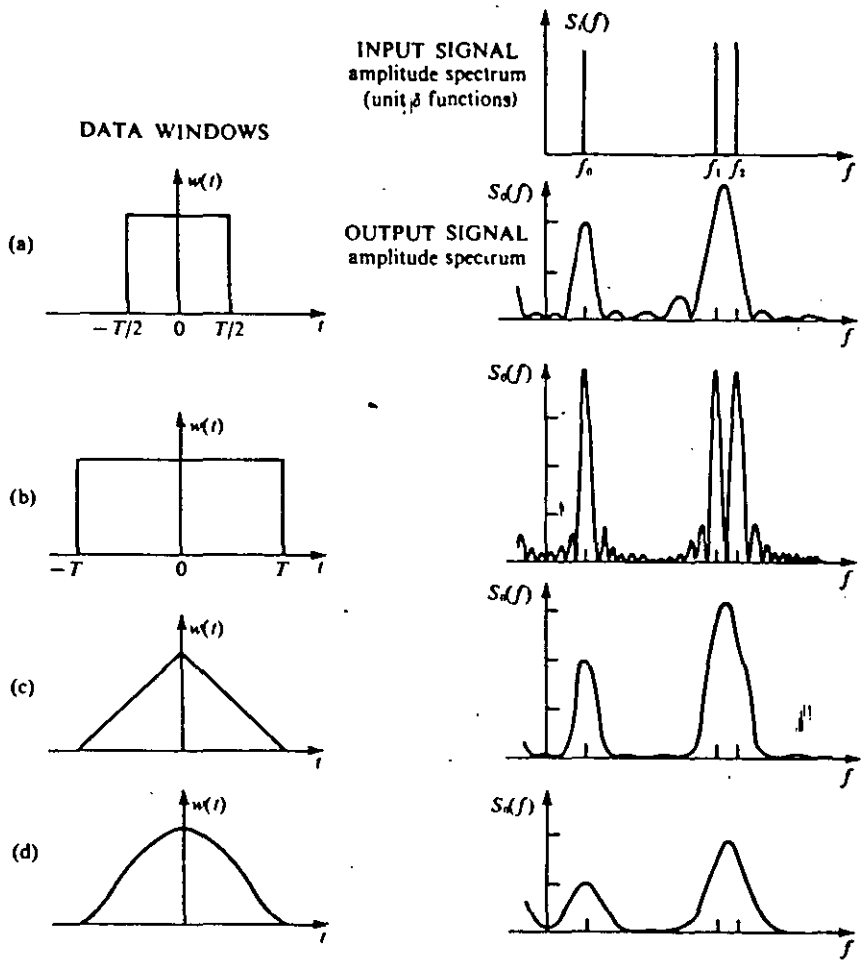


FIG. 2.10: Effect of data window shape and duration on signal spectra

to distinguish the two peaks at f_1 and f_2. However, with a rectangular window of length $2T$, the peaks are easily distinguishable. Hence, to separate two peaks at frequencies f_1 and f_2 it is necessary to use a record length T of order

$$T \geqslant \frac{1}{f_2 - f_1}, \tag{2.4.5}$$

for the rectangular data window. Figures 2.10(c) and (d) show that for non-rectangular windows, the widths must be greater than $2/(f_2 - f_1)$ to be able to distinguish peaks. Further discussion of the record length necessary to distinguish peaks is given in Section 6.4.4.

2.4.2 Time sampling and aliasing

Impulse modulation. For purposes of analysis, most continuous signals $s(t)$ will be read at some fixed interval Δ and converted into *sampled* signals which can then be used for digital calculations. The sampled signal may be regarded as the result of multiplying the original continuous signal by a signal $i(t)$, which consists of a *train of impulse* or *delta functions*,

$$i(t) = \sum_{n=-\infty}^{\infty} \delta(t - n\Delta). \tag{2.4.6}$$

This produces a sampled or impulse-modulated signal $s_i(t)$ where

$$s_i(t) = s(t)i(t). \tag{2.4.7}$$

Hence, using the convolution theorem (A2.1.8),

$$S_i(f) = \int_{-\infty}^{\infty} S(f - g)I(g)\, dg, \tag{2.4.8}$$

where $I(g)$ is the transform of $i(t)$. Using the expression (2.2.16) for $I(g)$, (2.4.8) becomes

$$\begin{aligned} S_i(f) &= \int_{-\infty}^{\infty} S(f - g)\, \frac{1}{\Delta} \sum_{n=-\infty}^{\infty} \delta\left(g - \frac{n}{\Delta}\right) dg \\ &= \frac{1}{\Delta} \sum_{n=-\infty}^{\infty} S\left(f - \frac{n}{\Delta}\right). \end{aligned} \tag{2.4.9}$$

Equation (2.4.9) shows that the sampled or impulse-modulated signal $s_i(t)$ has a transform with period $1/\Delta$, and that if $S(f)$ is zero when $|f| \geqslant 1/(2\Delta)$, then $S_i(f)$ is simply a periodic version of $S(f)$ as in (b) or (c) of Figure 2.11. This means that it is possible to recover $S(f)$ from $S_i(f)$ by multiplying $S_i(f)$ by $H(f)$, where

$$H(f) = \begin{cases} \Delta, & |f| \leqslant \dfrac{1}{2\Delta} \\ 0, & |f| > \dfrac{1}{2\Delta}. \end{cases} \tag{2.4.10}$$

Since multiplication in the frequency domain corresponds to convolution in the time domain, it follows that

$$s(t) = \int_{-\infty}^{\infty} \frac{\sin(\pi u/\Delta)}{(\pi u/\Delta)}\, s_i(t - u)\, du. \tag{2.4.11}$$

The function $\sin(\pi u/\Delta)/(\pi u/\Delta)$ is the ideal filter for recovering a continuous signal $s(t)$ from a sampled signal $s_i(t)$. Alternatively, $\sin(\pi u/\Delta)/(\pi u/\Delta)$ is the

ideal interpolating function for equally spaced ordinates, and (2.4.11) is sometimes referred to as *Whittaker's interpolation formula.*

Aliasing. If the sampling interval is such that $S(f)$ falls off to zero before $|f| = 1/(2\Delta)$ as in (b) or (c) of Figure 2.11, then it is possible to recover $s(t)$ from $s_i(t)$. On the other hand, if $S(f)$ is not zero above $f_N = 1/(2\Delta)$, frequency

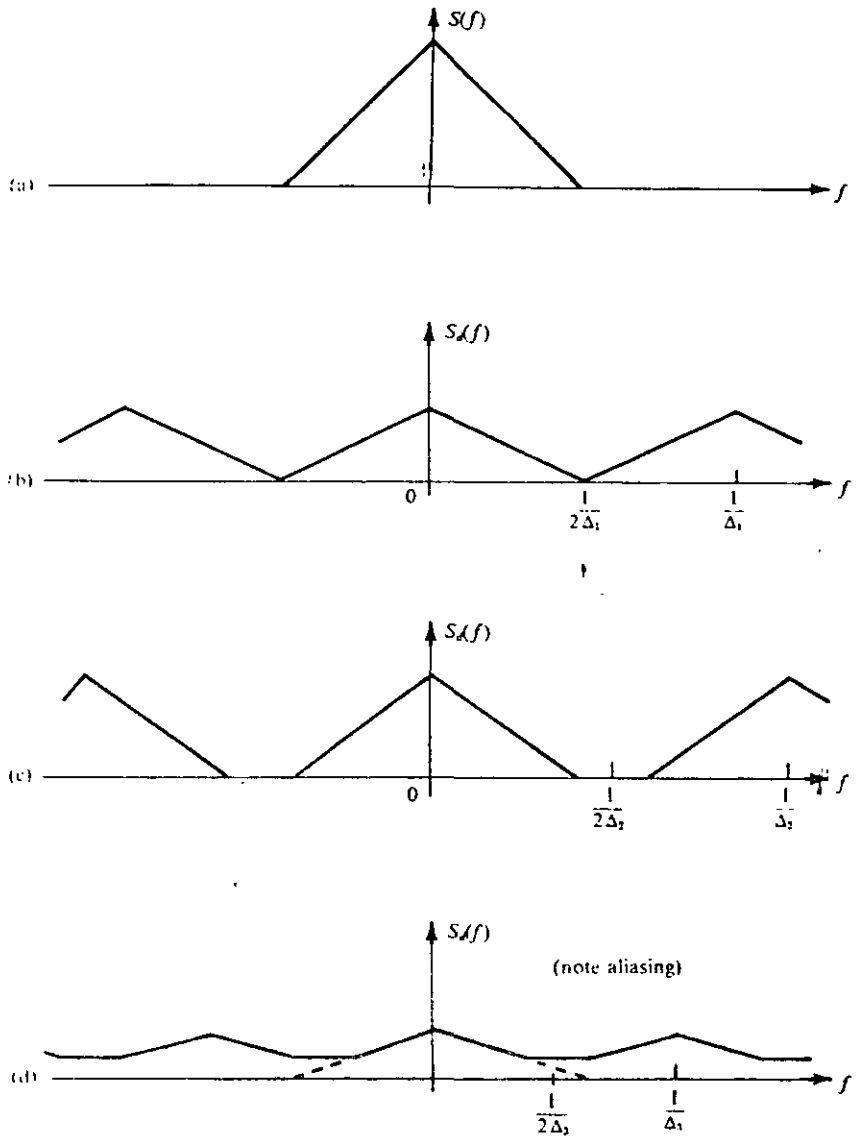


FIG. 2.11. Transforms of a signal and of the sampled signal for various sampling periods

components above $1/(2\Delta)$ in $S(f)$ appear in $S_i(f)$, as in (d) of Figure 2.11, in the range $-1/(2\Delta) \leqslant f \leqslant 1/(2\Delta)$. The frequency $f_N = 1/(2\Delta)$ is called the *Nyquist frequency* and is the highest frequency which can be detected with data sampled at intervals Δ.

For example, if $\Delta = 0.1$ seconds, the Nyquist frequency is 5 cps. The discrete transform $S_i(f)$ at 4 cps will then be made up of contributions from $S(f)$ at 4 cps, $10 + 4 = 14$ cps, $-10 + 4 = -6$ cps, $20 + 4 = 24$ cps, $-20 + 4 = -16$ cps and so on. These other frequencies are usually referred to as *aliases* of the frequency 4 cps, and the effect on the Fourier transform is called aliasing. It follows that in sampling continuous time series, adequate care must be taken to ensure that a high enough sampling frequency $f_N = 1/(2\Delta)$ is chosen so that misleading values of $S_i(f)$ will be avoided.

The phenomenon of aliasing arises in a number of practical contexts, such as in the use of a stroboscope or in movie films. For example, the wheels of a stage coach which is accelerating from zero appear to rotate in the correct direction with increasing speed, then they appear to be rotating in the opposite direction with decreasing speed until they stop, then they begin to rotate with increasing speed in the forward direction and so on.

An example. To illustrate the considerations discussed in this section, suppose that it is desired to calculate the length of record T and sampling interval Δ required to achieve certain objectives. Suppose it is known that the signal under study contains two sinusoidal components at frequencies of 100 cps and 99 cps. Then if it is desired to distinguish these peaks in the Fourier transform of the finite length of record, (2.4.5) shows that $1/T$ must be of the order of $100 - 99 = 1$ cps, that is, T must be of the order of 1 sec. In order to estimate a frequency as high as 100 cps, then $1/(2\Delta)$ must be at least 100 cps and hence $\Delta \leqslant 5$ millisec. Thus at least 200 data values would be needed.

If it were required to separate two frequency components at 999 and 1000 cps, then a record of 1 sec duration would still be necessary, but the sampling interval would now have to be 0.5 millisec, so 2000 data values would be required.

Hence the *length* T of the record determines the extent to which peaks in the Fourier transform may be distinguished. On the other hand, the *sampling interval* Δ determines the maximum frequency which can be distinguished.

REFERENCES

[1] M. J. Lighthill, *An Introduction to Fourier Analysis and Generalized Functions*. Cambridge University Press, Cambridge, 1959.

[2] A. Papoulis, *The Fourier Integral and its Applications*. McGraw-Hill, New York, 1962.

[3] R. Courant, *Differential and Integral Calculus*. Vol. II. Blackie and Son, London, 1952.
[4] H. M. James, N. B. Nichols and R. S. Phillips, *Theory of Servomechanisms*. McGraw-Hill, New York, 1947.
[5] H. W. Bode, *Network Analysis and Feedback Amplifier Design*. Van Nostrand, New York, 1945.
[6] G. A. Campbell and R. M. Foster, *Fourier Integrals for Practical Applications*. Van Nostrand, New York, 1942.
[7] E. I. Jury, *Theory and Application of the Z-Transform Method*. John Wiley, New York, 1964.

APPENDIX A2.1 OPERATIONAL PROPERTIES OF FOURIER TRANSFORMS

During the course of this book, it will be necessary to perform various operations with Fourier transforms. These are now summarized.

Time scaling and shift of origin. If $s(t)$ has the Fourier transform $S(f)$, then the Fourier transform of $s(\alpha t + \beta)$ is

$$\frac{1}{|\alpha|} e^{j(2\pi f\beta/\alpha)} S\left(\frac{f}{\alpha}\right). \tag{A2.1.1}$$

Example: From Table 2.5, the Fourier transform of $e^{-\pi t^2}$ is

$$S(f) = e^{-\pi f^2}.$$

Hence the transform of

$$\exp\left[-\frac{1}{2}\left(\frac{t-\mu}{\sigma}\right)^2\right] = s\left(\frac{t-\mu}{\sqrt{2\pi}\,\sigma}\right)$$

is

$$|\sqrt{2\pi}\,\sigma|\, e^{j2\pi f(-\mu)} S(\sqrt{2\pi}\,\sigma f) = \sqrt{2\pi}\,\sigma\, e^{-j2\pi f\mu}\, e^{-2\pi^2\sigma^2 f^2},$$

where

$$\alpha = 1/\sqrt{2\pi}\,\sigma, \qquad \beta = -\mu/\sqrt{2\pi}\,\sigma, \qquad \beta/\alpha = -\mu.$$

Differentiation. If $s(t)$ has the Fourier transform $S(f)$, then the mth derivative $s^{(m)}(t)$ has the Fourier transform

$$(j2\pi f)^m S(f), \tag{A2.1.2}$$

provided the derivative exists.

Example: As in the preceding example, using the transform pair

$$s(t) = e^{-\pi t^2}, \qquad S(f) = e^{-\pi f^2},$$

from Table 2.5, the Fourier transform of

$$\frac{d}{dt} s(t) = \frac{d}{dt}(e^{-\pi t^2}) = -2\pi t\, e^{-\pi t^2}$$

is

$$j2\pi f S(f) = j2\pi f\, e^{-\pi f^2}.$$

Integration. If $s(t)$ has the Fourier transform $S(f)$ then the Fourier transform of $I^m s(t)$, where

$$Is(t) = \int_{-\infty}^{t} s(u)\, du,$$

is

$$\left(\frac{1}{j2\pi f}\right)^m S(f) + K_1\,\delta(f) + K_2\,\delta'(f) + \cdots + K_m\,\delta^{m-1}(f). \tag{A2.1.3}$$

The constants $K_1, K_2, \ldots, K_m$ in (A2.1.3) may be determined by using the values of the functions $s(t), ds/dt, \ldots, d^m s/dt^m$ evaluated at $t = 0$ since, for example,

$$s(0) = \int_{-\infty}^{\infty} S(f)\, df.$$

Example: From the preceding example, the function

$$s(t) = -2\pi t\, e^{-\pi t^2}$$

has the Fourier transform

$$S(f) = j2\pi f\, e^{-\pi f^2}.$$

Hence the Fourier transform of

$$s_1(t) = \int_{-\infty}^{t} s(u)\, du = e^{-\pi t^2}$$

is

$$S_1(f) = \frac{1}{j2\pi f} S(f) + K_1\delta(f)$$
$$= e^{-\pi f^2} + K_1\delta(f).$$

Integrating both sides over f gives

$$s_1(0) = \int_{-\infty}^{\infty} e^{-\pi f^2}\, df + K_1$$
$$= 1 + K_1.$$

But $s_1(0) = 1$ and hence $K_1 = 0$.

Symmetry. If $S(f)$ is the Fourier transform of $s(t)$, then $s(f)$ is the Fourier transform of $S(-t)$.

Example: The Fourier transform of

$$s(t) = \begin{cases} e^{-t}, & t \geq 0 \\ 0, & t < 0 \end{cases}$$

is $S(f) = 1/(1 + j2\pi f)$. Hence the Fourier transform of $s(t) = 1/(1 - j2\pi t)$ is

$$S(f) = \begin{cases} e^{-f}, & f \geq 0 \\ 0, & f < 0. \end{cases}$$

Similarly, the Fourier transform of $s(t) = 1/(1 + j2\pi t)$ is

$$S(f) = \begin{cases} e^{f}, & f \leq 0 \\ 0, & f > 0. \end{cases}$$

Hence the Fourier transform of

$$\frac{2}{1 + (2\pi t)^2} = \frac{1}{1 - j2\pi t} + \frac{1}{1 + j2\pi t}$$

is

$$e^{-|f|}, \quad -\infty \leqslant f \leqslant \infty.$$

Convolution and Parseval's theorem. This is a more general form of the theorem than the results (2.1.16, 20, 26) derived in Section 2.1. The generalized form states that if $s_1(t)$ and $s_2(t)$ are two complex signals with Fourier transforms $S_1(f)$ and $S_2(f)$ respectively, then

$$\int_{-\infty}^{\infty} s_1(t)s_2^*(t)\,dt = \int_{-\infty}^{\infty} S_1(f)S_2^*(f)\,df, \tag{A2.1.4}$$

where the asterisk denotes a complex conjugate.

Three special cases of (A2.1.4) are sometimes useful:

(a) When $s_2^*(t) = h(u - t)$, (A2.1.4) reduces to

$$\int_{-\infty}^{\infty} s_1(t)h(u - t)\,dt = \int_{-\infty}^{\infty} S_1(f)\,H(f)\,e^{j2\pi fu}\,df. \tag{A2.1.5}$$

(b) When $s_1(t)$ and $s_2(t)$ are real, (A2.1.4) reduces to

$$\int_{-\infty}^{\infty} s_1(t)s_2(t)\,dt = \int_{-\infty}^{\infty} S_1(f)S_2(-f)\,df. \tag{A2.1.6}$$

(c) When $s_1(t) = s_2(t) = s(t)$, (A2.1.4) reduces to

$$\int_{-\infty}^{\infty} |s(t)|^2\,dt = \int_{-\infty}^{\infty} |S(f)|^2\,df. \tag{A2.1.7}$$

The form (A2.1.7) of Parseval's theorem includes the form (2.1.26) which was derived in Section 2.1.

Note that because of the symmetry of the Fourier transform, the roles of the signal and its transform may be reversed. For example,

(a) $$\int_{-\infty}^{\infty} S_1(f)S_2(g - f)\,df = \int_{-\infty}^{\infty} s_1(t)s_2(t)\,e^{j2\pi gt}\,dt, \tag{A2.1.8}$$

(b) $$\int_{-\infty}^{\infty} S_1(f)S_2(f)\,df = \int_{-\infty}^{\infty} s_1(t)s_2(-t)\,dt, \tag{A2.1.9}$$

while the symmetry of the relation given by (c) is already evident.

It should be noted that the above operational properties apply equally well to finite and infinite Fourier series. The three forms of Parseval's theorem derived in Section 2.1 provide examples.

CURSO INGENIERIA MARITIMA.
MODULO: "FUNDAMENTOS DE HIDRAULICA Y MATEMATICAS"
DEL 11 DE MARZO AL 7 DE JUNIO.
MEXICO, D.F.

THE SPECTRUM.

PROF. ING. JOSE RAYNAL.
JUNIO 1985

6

The Spectrum

In Chapter 5 it was shown that a stationary stochastic process is simply described by its autocovariance function. In this chapter it is shown that an equivalent description is provided by its power spectrum, which is the Fourier transform of the autocovariance function. The power spectrum curve shows how the variance of the stochastic process is distributed with frequency.

In Section 6.1 it is shown that classical Fourier methods fail when applied to time series. Thus the estimator of the spectrum obtained from considerations of Fourier analysis, namely the sample spectrum, has the unfortunate property that its variance does not decrease as the length of the time series increases. Hence the methods of Chapter 2 must be modified for time series analysis. These modifications lead to the definition of a spectrum appropriate to stochastic processes, which is given in Section 6.2. This section also deals with the spectra of linear moving average and autoregressive processes.

In Section 6.3 it is shown that an improved estimator of the spectrum can be obtained by smoothing the sample spectrum. The more one smooths the smaller is the variance but the larger is the bias or distortion. Hence a compromise has to be achieved between bias and variance.

Further properties of the smoothed estimators, including the notion of bandwidth, are derived in Section 6.4. It is also shown that confidence intervals are easily obtained at each frequency by using the logarithm of the spectrum estimate.

6.1 THE SAMPLE SPECTRUM

6.1.1 Fourier methods applied to time series

Fourier analysis. It was shown in Chapter 2 that the variance or average power of a signal $x(t)$ in the range $-T/2 \leqslant t \leqslant T/2$ can be decomposed into contributions at harmonics $f_m = m/T$ of the fundamental frequency $f_1 = 1/T$ according to

$$s_T^2 = \frac{1}{T}\int_{-T/2}^{T/2} x^2(t)\,dt = \sum_{m=-\infty}^{\infty} |X_m|^2. \qquad (6.1.1)$$

X_m is called the complex amplitude at the harmonic frequency $f_m = m/T$ and measures the amplitudes of the sine and cosine terms at frequency f_m in $x(t)$. The complex amplitude may be computed from

$$X_m = \frac{1}{T}\int_{-T/2}^{T/2} x(t)\, e^{-j2\pi mt/T}\, dt \tag{6.1.2}$$

by writing

$$e^{-j2\pi mt/T} = \cos\frac{2\pi mt}{T} - j\sin\frac{2\pi mt}{T}.$$

Recall that the Fourier decomposition of $x(t)$ is

$$x(t) = \sum_{m=-\infty}^{\infty} X_m\, e^{+j2\pi mt/T}.$$

Similarly, for a discrete signal observed at times $t = -n\Delta, -(n-1)\Delta, \ldots, (n-1)\Delta$, the average power is decomposed into contributions at a *finite* number of harmonics of the fundamental frequency $f_1 = 1/N\Delta$ ($N = 2n$), and the relations corresponding to (6.1.1) and (6.1.2) are

$$s_T^2 = \frac{1}{N}\sum_{t=-n}^{n-1} x_t^2 = \sum_{m=-n}^{n-1} |X_m|^2, \tag{6.1.3}$$

$$X_m = \frac{1}{N}\sum_{t=-n}^{n-1} x_t\, e^{-j2\pi mt\,\Delta/N\Delta} = \frac{1}{N}\sum_{t=-n}^{n-1} x_t\, e^{-j2\pi mt/N}. \tag{6.1.4}$$

The contribution $|X_m|^2$ to the average power at frequency f_m is called the *intensity* at this frequency, and the plot of $|X_m|^2$ versus m is called the *Fourier line spectrum*, an example of which is given in Figure 2.2.

The power spectrum of deterministic signals. It is when the record length tends to infinity that the main difference between the analysis of deterministic and stochastic signals arises. This distinction is not explained in many engineering texts, which use an argument similar to the following. From (6.1.1), the variance of the infinite record is

$$\sigma^2 = \lim_{T\to\infty}\frac{1}{T}\int_{-T/2}^{T/2} x^2(t)\, dt = \lim_{T\to\infty}\sum_{m=-\infty}^{\infty}(T|X_m|^2)\frac{1}{T} = \int_{-\infty}^{\infty}\Gamma(f)\, df,$$

where

$$\Gamma(f) = \lim_{T\to\infty} T|X_m|^2 \tag{6.1.5}$$

is called a Fourier "*power spectrum.*" Using (6.1.2), the function $T|X_m|^2$ may be written

$$T|X_m|^2 = C_{xx}(f) = \frac{1}{T}\left|\int_{-T/2}^{T/2} x(t)\, e^{-j2\pi ft}\, dt\right|^2. \tag{6.1.6}$$

Note that $C_{xx}(f)$ is defined for a continuous range of frequencies $-\infty \leqslant f \leqslant \infty$ and is called the *sample spectrum.* For the discrete case the sample spectrum is

$$C_{xx}(f) = \frac{\Delta}{N}\left|\sum_{t=-n}^{n-1} x_t\, e^{-j2\pi ft\,\Delta}\right|^2$$

$$= \frac{\Delta}{N}\left\{\left(\sum_{t=-n}^{n-1} x_t\cos 2\pi ft\Delta\right)^2 + \left(\sum_{t=-n}^{n-1} x_t\sin 2\pi ft\Delta\right)^2\right\},$$

$$\frac{1}{2\Delta} \leqslant f < \frac{1}{2\Delta}. \tag{6.1.7}$$

The frequency $1/2\Delta$ in (6.1.7) is the Nyquist frequency discussed in Chapter 2 and is the highest frequency which can be detected from data spaced at intervals Δ seconds apart.

Note that if $x(t)$ has a well-behaved Fourier transform, the limit (6.1.5) for $\Gamma(f)$ is zero. This is because $x(t)$ must tend to zero as t tends to $\pm\infty$ if its Fourier transform is to exist. However, provided $x(t)$ does not dissipate itself in this way, $C_{xx}(f)$ will usually tend to a well-defined limit $\Gamma(f)$. For deterministic signals, the convergence of $C_{xx}(f)$ to $\Gamma(f)$ is smooth in the sense that the function $C'_{xx}(f)$ obtained by increasing the record length from T to T' would be a smoother version of the function $C_{xx}(f)$ based on the record of length T.

It will be shown in the next section that the definition (6.1.5) is not a satisfactory one when $x(t)$ is a realization of a stochastic process. The basic difference between the Fourier analysis of a deterministic and stochastic signal is that the plot of $C'_{xx}(f)$ obtained from a record of length $T' > T$ of a stochastic signal is just as erratic as that obtained from a record of length T, that is, $C_{xx}(f)$ *does not converge in any statistical sense* to a limiting value as T tends to infinity.

6.1.2 The sample spectrum of a white noise process

To illustrate the effect of applying Fourier analysis to a stochastic process, a series of 400 random Normal deviates (Gaussian white noise) was generated. The sample spectrum $C_{zz}(f)$ was computed from the series consisting of the first 50 terms, the first 100, the first 200 and the entire series of 400 terms. Figure 6.1 shows the values of the sample spectrum $C_{zz}(f)$ computed according to (6.1.7) at the frequencies $f = 0.02, 0.04, \ldots, 0.50$ cps for the cases $N = 50$ and $N = 100$, with $\Delta = 1$ sec. Also indicated in the figure is the theoretical spectrum, which is shown in Section 6.2.3 to be a constant in the range $-\frac{1}{2} \leqslant f < \frac{1}{2}$.

Figure 6.1 shows that the fluctuations in $C_{zz}(f)$ are so erratic that it would be difficult to conclude on the basis of this diagram that the true spectrum

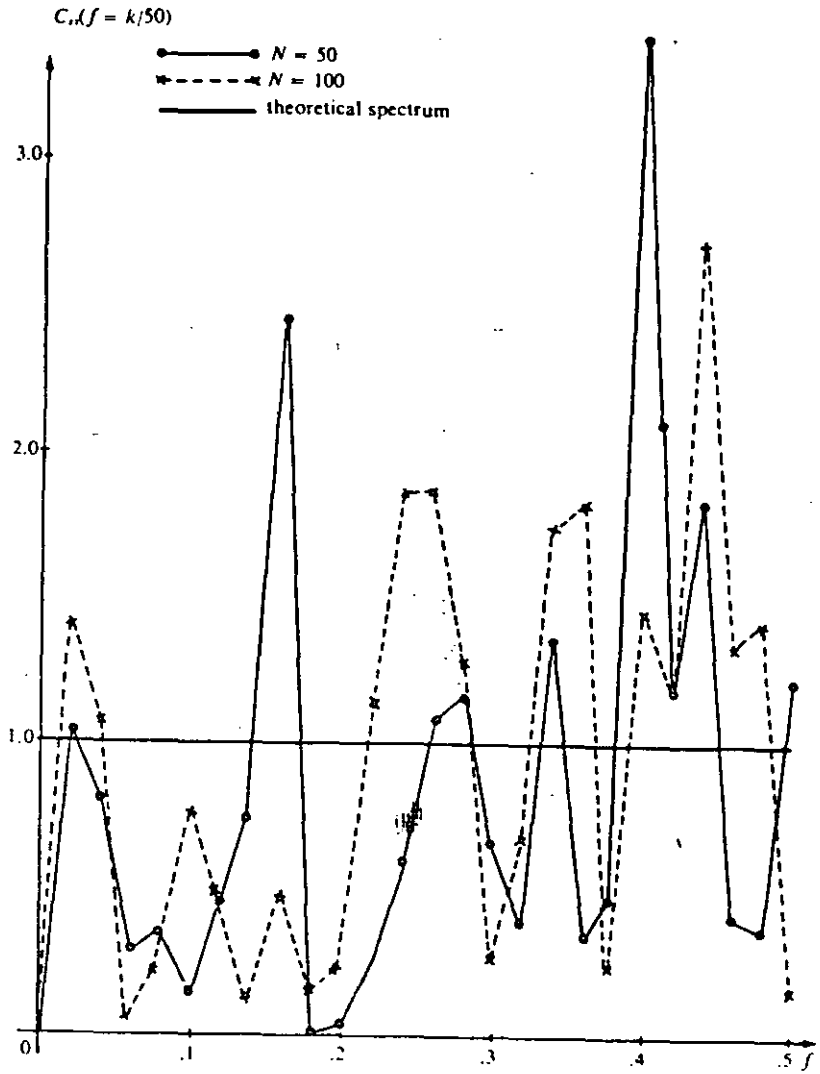


FIG. 6.1: Sample spectra for the first half ($N = 50$) and the whole ($N = 100$) of a realization of discrete Normal white noise

was a constant and hence that the time series was white noise. Note also that the fluctuations in $C_{zz}(f)$ for $N = 100$ are just as large as for $N = 50$, indicating lack of statistical convergence of any kind.

Table 6.1 summarizes the results obtained for the sample spectrum for series with $N = 50$, 100, 200 and 400 terms. Since the theoretical spectrum has the same value at all frequencies, the fluctuations in $C_{zz}(f)$ can be summarized by evaluating the mean, variance and mean square error over frequency. It is seen that the mean values for each series are close to unity, the

value of the theoretical spectrum. Hence the $C_{zz}(f)$ seem to be clustering about some central value. However, Table 6.1 shows that the variances do not decrease as N increases, showing that the estimate of the spectrum obtained from a sample of $N = 100$, 200 or 400 is no better than that for $N = 50$.

TABLE 6.1: Behavior of sample spectra of white noise as the record length is increased

N	50	100	200	400
mean	0.85	1.07	1.00	0.95
variance	0.630	0.777	0.886	0.826
mean square error	0.652	0.782	0.886	0.828

It was seen in Chapter 4 that well-behaved estimators have the property that their variance decreases as N increases. Hence it can be concluded that $C_{zz}(f)$ is not a good estimate of the spectrum, at least not in its present form.

To show that the sample spectra for random processes other than white noise do not converge in any statistical sense, consider the autoregressive process generated according to (5.3.36). The theoretical and sample acf's for a realization of 400 terms were shown in Figure 5.13. The theoretical spectrum and the sample spectrum for the same realization of 400 terms are shown in Figure 6.2. As for the white noise example, the sample spectrum is extremely erratic and bears little resemblance to the theoretical spectrum.

Summary. For deterministic signals, the spectrum is the limit, in the usual mathematical sense, of the sample spectrum $C_{xx}(f)$ as the record length tends to infinity. However, the white noise example above shows that when $C_{xx}(f)$ is used to analyze a time series its behavior is so erratic as to render it useless for estimation purposes. The basic reason why Fourier analysis breaks down when applied to time series is that it is based on the assumption of *fixed* amplitudes, frequencies and phases. Time series, on the other hand, are characterized by *random* changes of frequencies, amplitudes and phases. Therefore it is not surprising that Fourier methods need to be adapted to take account of the random nature of a time series.

6.1.3 *The relation between the sample spectrum and the sample autocovariance function*

Before giving a more precise definition of the spectrum of a stationary stochastic process, a fundamental relationship connecting the sample spectrum and the sample autocovariance function is proved.

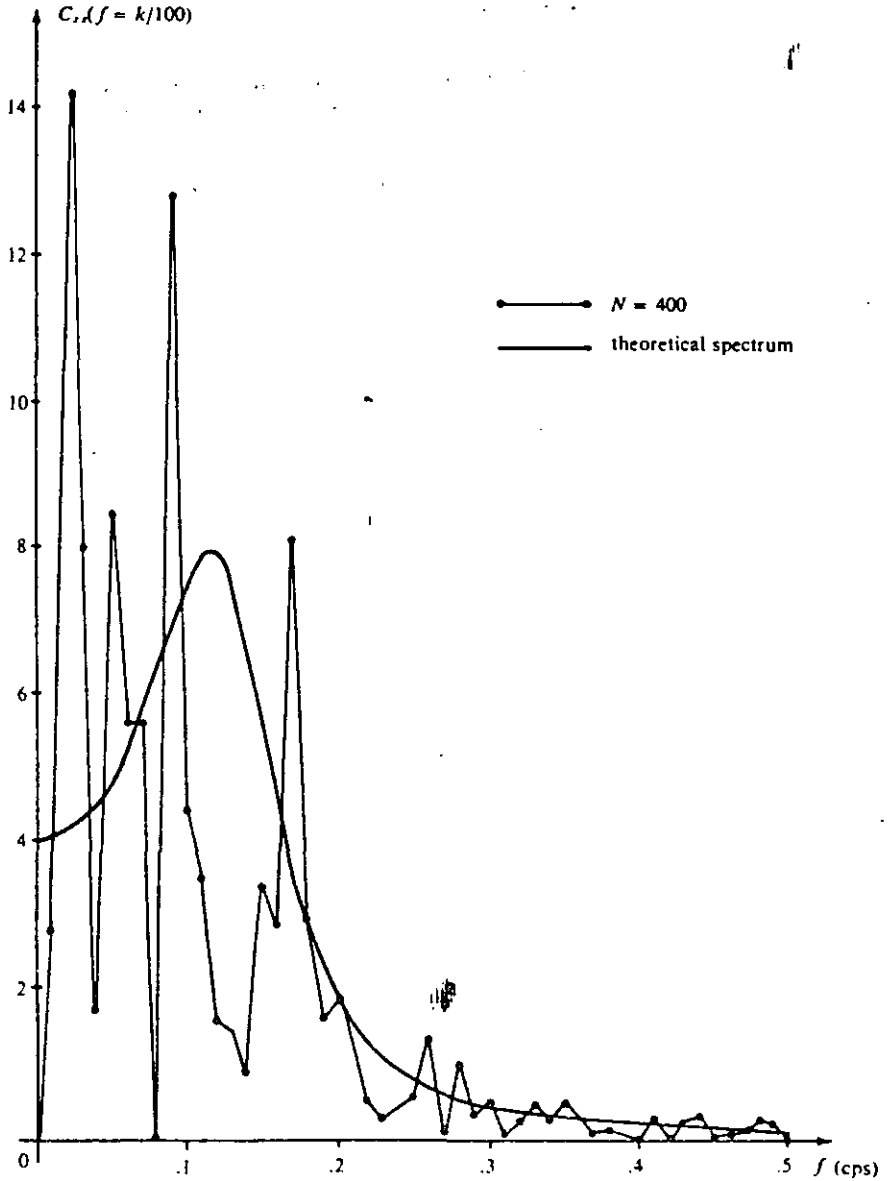


FIG. 6.2: Sample spectrum for a realization of a second-order autoregressive process

Using the definition (6.1.6) of the sample spectrum,

$$C_{xx}(f) = \frac{1}{T}\int_{-T/2}^{T/2} x(t)\, e^{-j2\pi ft}\, dt \int_{-T/2}^{T/2} x(t')\, e^{+j2\pi ft'}\, dt'. \qquad (6.1.8)$$

If the transformation

$$u = t - t', \qquad v = t'$$

is made in the double integral (6.1.8), the region of integration is transformed as shown in Figure 6.3. Then (6.1.8) becomes

$$C_{xx}(f) = \int_0^T \left[\frac{1}{T}\int_{-T/2}^{(T/2)-u} x(v)x(v+u)\, dv\right] e^{-j2\pi fu}\, du + \int_{-T}^{0} \left[\frac{1}{T}\int_{-(T/2)-u}^{T/2} x(v)x(v+u)\, dv\right] e^{-j2\pi fu}\, du.$$

Using the definition (5.3.8) of $c_{xx}(u)$, this becomes

$$C_{xx}(f) = \int_{-T}^{T} c_{xx}(u)\, e^{-j2\pi fu}\, du, \quad -\infty \leqslant f \leqslant \infty. \qquad (6.1.9)$$

Hence *the sample spectrum is the Fourier transform of the sample acvf.* The inverse Fourier transform of (6.1.9) may be written

$$c_{xx}(u) = \int_{-\infty}^{\infty} C_{xx}(f)\, e^{j2\pi fu}\, df, \quad -T \leqslant u \leqslant T, \qquad (6.1.10)$$

which, for $u = 0$, becomes

$$c_{xx}(0) = s_T^2 = \int_{-\infty}^{\infty} C_{xx}(f)\, df. \qquad (6.1.11)$$

Thus *the sample spectrum shows how the variance or average power of $x(t)$ is distributed over frequency.*

In discrete time, the sample spectrum is

$$C_{xx}(f) = \Delta \sum_{k=-(N-1)}^{(N-1)} c_{xx}(k)\, e^{-j2\pi fk\Delta}, \quad -\frac{1}{2\Delta} \leqslant f < \frac{1}{2\Delta}, \qquad (6.1.12)$$

which corresponds to (6.1.9). The inverse transform of (6.1.12) is

$$c_{xx}(u) = \int_{-1/2\Delta}^{1/2\Delta} C_{xx}(f)\, e^{j2\pi fu}\, df, \quad -N\Delta \leqslant u \leqslant N\Delta, \qquad (6.1.13)$$

which corresponds to (6.1.10).

The Fourier transform pairs (6.1.9, 10) and (6.1.12, 13) are mathematical identities which hold whether $x(t)$ is deterministic *or* a realization of a stochastic process. In the next section, an interpretation of the *limiting* value of $C_{xx}(f)$ is given when the function $x(t)$ is a realization of a stationary stochastic process.

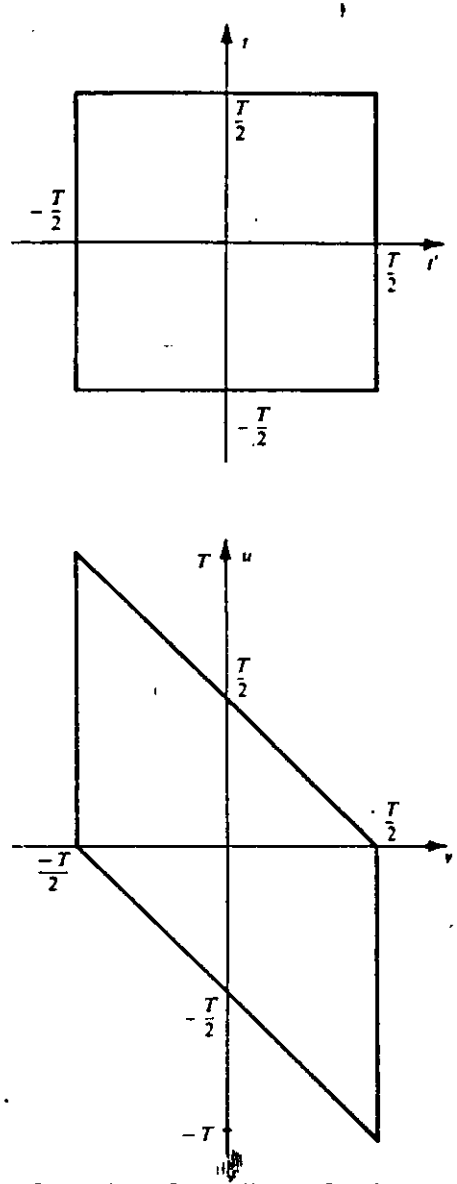


FIG. 6.3: Transformation of coordinates for the sample spectrum

6.2 THE SPECTRUM

6.2.1 *Definition of the spectrum of a stochastic process*

To describe the variability in $C_{xx}(f)$ demonstrated in Section 6.1.2, it is necessary to regard the record $x(t)$, $-T/2 \leqslant t \leqslant T/2$, as being one of many possible time series which might have been observed, that is, as a realization

of a stochastic process. Thus, the variability in the record is characterized by rv's $X(t)$ for $-T/2 \leqslant t \leqslant T/2$, as indicated in Chapter 5. The sample spectrum $C_{xx}(f)$ is then regarded as a realization of the *rv* $C_{XX}(f)$, just as the sample covariance function $c_{xx}(u)$ was regarded as a realization of the rv $c_{XX}(u)$. By deriving the distribution of $C_{XX}(f)$, or its moments, the erratic behavior of $C_{xx}(f)$ demonstrated in Figures 6.1 and 6.2 can be explained.

Using (6.1.9), the first moment of the sample spectrum estimator $C_{XX}(f)$ is

$$E[C_{XX}(f)] = \int_{-T}^{T} E[c_{XX}(u)]\, e^{-j2\pi fu}\, du,$$

which becomes

$$E[C_{XX}(f)] = \int_{-T}^{T} \gamma_{XX}(u) \left(1 - \frac{|u|}{T}\right) e^{-j2\pi fu}\, du, \tag{6.2.1}$$

using (5.3.13). Thus (6.2.1) gives the average distribution (over all possible time series of length T) of power with frequency. As the record length T increases, $E[C_{XX}(f)]$ tends in the limit to

$$\Gamma_{XX}(f) = \lim_{T \to \infty} E[C_{XX}(f)] = \int_{-\infty}^{\infty} \gamma_{XX}(u)\, e^{-j2\pi fu}\, du. \tag{6.2.2}$$

The mathematics associated with this limiting operation is discussed more fully in [1].

The function $\Gamma_{XX}(f)$ is called the *power spectrum*, but from now on it will be referred to more concisely as the *spectrum*.

Equation (6.2.2) shows that the spectrum is the Fourier transform of the acvf of the $X(t)$ process. Using Table 2.3, there is an inverse transform

$$\gamma_{XX}(u) = \int_{-\infty}^{\infty} \Gamma_{XX}(f)\, e^{j2\pi fu}\, df. \tag{6.2.3}$$

Setting $u = 0$ in (6.2.3) gives

$$\gamma_{XX}(0) = \sigma_X^2 = \int_{-\infty}^{\infty} \Gamma_{XX}(f)\, df, \tag{6.2.4}$$

and hence $\Gamma_{XX}(f)$ *shows how the variance of the* $X(t)$ *process is distributed over frequency* in the same way that (6.1.9) shows how the variance of *one particular sample of length* T is distributed over frequency. Specifically, the variance of the $X(t)$ process which is due to frequencies in the range f to $f + df$ is approximately $\Gamma_{XX}(f)\, df$. Note from the definition (6.1.6) that $\Gamma_{XX}(f)$ is non-negative for all f.

For discrete time, the relations corresponding to (6.2.1) to (6.2.3) are

$$E[C_{XX}(f)] = \Delta \sum_{k=-(N-1)}^{N-1} \gamma_{XX}(k) \left(1 - \frac{|k|}{N}\right) e^{-j2\pi fk\Delta}, \quad -\frac{1}{2\Delta} \leqslant f < \frac{1}{2\Delta}, \tag{6.2.5}$$

$$\Gamma_{XX}(f) = \lim_{N\to\infty} E[C_{XX}(f)] = \Delta \sum_{k=-\infty}^{\infty} \gamma_{XX}(k)\, e^{-j2\pi k f \Delta}, \quad -\frac{1}{2\Delta} \leqslant f < \frac{1}{2\Delta}, \tag{6.2.6}$$

and

$$\gamma_{XX}(k) = \int_{-1/(2\Delta)}^{1/(2\Delta)} \Gamma_{XX}(f)\, e^{j2\pi f k \Delta}\, df, \quad k = 0, \pm 1, \pm 2, \ldots. \tag{6.2.7}$$

Some examples. To provide some insight into the information contained in the spectrum, Figures 6.4 and 6.5 show the theoretical spectra of the first-order ar processes whose acfs were shown in Figures 5.7 and 5.8. Analytical expressions for the spectra of ar processes will be derived in Section 6.2.5.

Figure 6.4 shows that when the autoregressive parameter $\alpha_1 = 0.9$, the series is smooth, and this is reflected in an acf which damps out smoothly with lag. It is seen that the corresponding spectrum is large at low frequencies and small at high frequencies. Hence smooth series are characterized by spectra which have most of their power at low frequencies. Note that in Figures 6.4, 6.5 and 6.6, the spectrum is plotted on a *logarithmic scale*, which shows more detail in the spectrum over a wider amplitude range. Other reasons for plotting the logarithm of the spectrum will be given later on.

Figure 6.5 shows that when $\alpha_1 = -0.9$, the series oscillates very quickly, and this is reflected in an acf which changes sign. The corresponding spectrum has large power at high frequencies and small power at low frequencies. Hence quickly oscillating series are characterized by spectra which have most of their power at high frequencies.

Figure 6.6 illustrates a second-order autoregressive process. As discussed in Section 5.2.4, this is a quasi-periodic series with an "average" period of about 8 seconds. The acf reflects the periodic behavior and consists of a damped sine wave with a period of 8 seconds. The corresponding spectrum now has a peak at the frequency $f_0 = 0.125$ cps. Since the $X(t)$ process is not truly periodic, the spectrum is not concentrated at the single frequency $f = 0.125$ cps, but instead is spread over all frequencies in the range $-0.5 \leqslant f < +0.5$ cps. Most of the power, however, is near the frequency $f_0 = 0.125$ cps.

The spectral density function. It is sometimes necessary to compare time series which have different scales of measurement, and in these circumstances it is useful to normalize $\Gamma_{XX}(f)$ by dividing by the variance σ_X^2. The function

$$\frac{\Gamma_{XX}(f)}{\sigma_X^2}$$

is called the *spectral density function*. From (6.2.2) it follows that

$$\frac{\Gamma_{XX}(f)}{\sigma_X^2} = \int_{-\infty}^{\infty} \rho_{XX}(u)\, e^{-j2\pi f u}\, du, \tag{6.2.8}$$

so that the spectral density function is the Fourier transform of the acf.

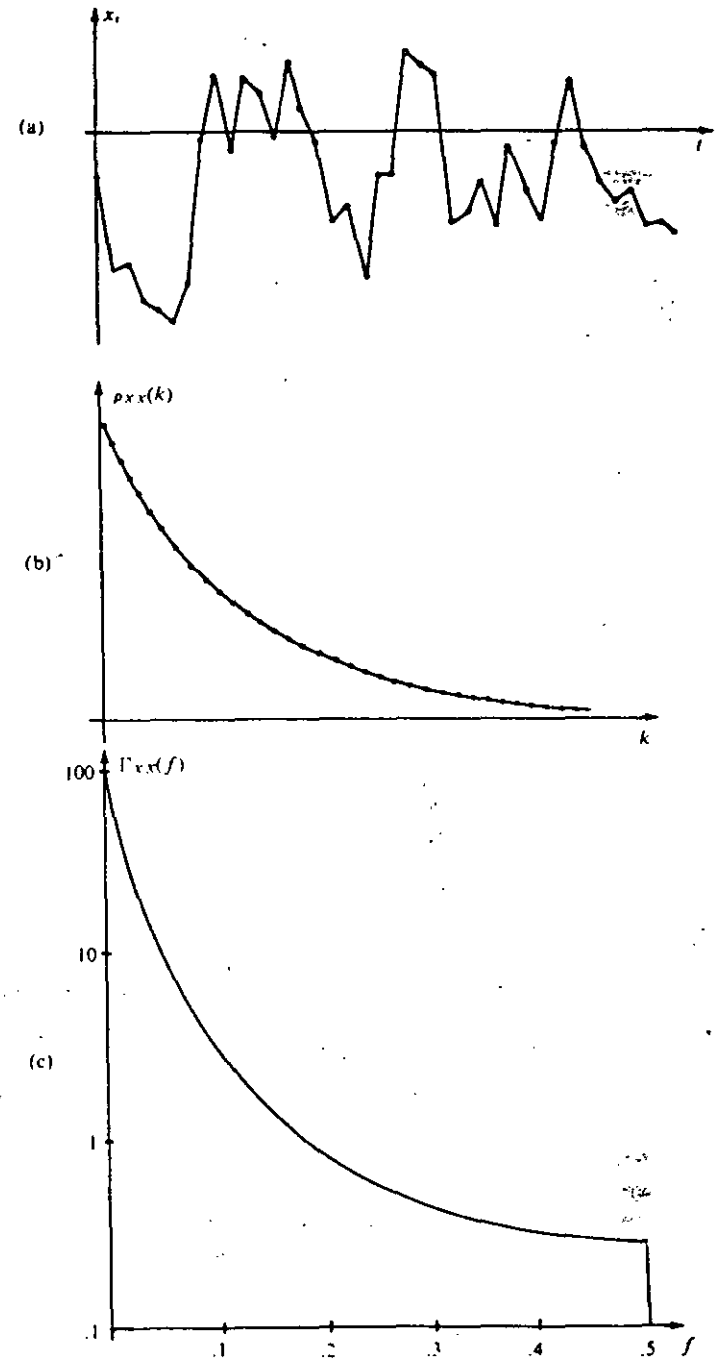


FIG. 6.4: A realization (a), the autocorrelation function (b) and spectrum (c) of a discrete first-order autoregressive process ($\alpha_1 = +0.9$)

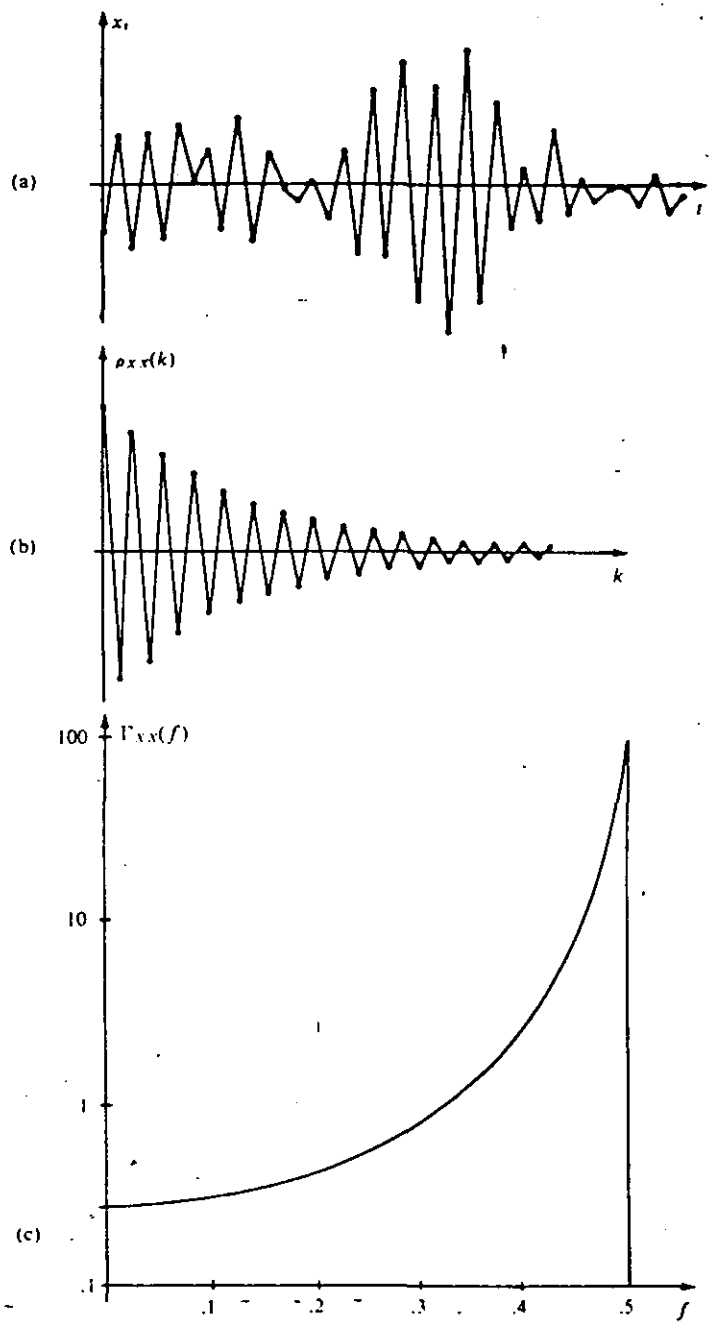


FIG. 6.5: A realization (a), the autocorrelation function (b) and spectrum (c) of a discrete first-order autoregressive process ($\alpha_1 = -0.9$)

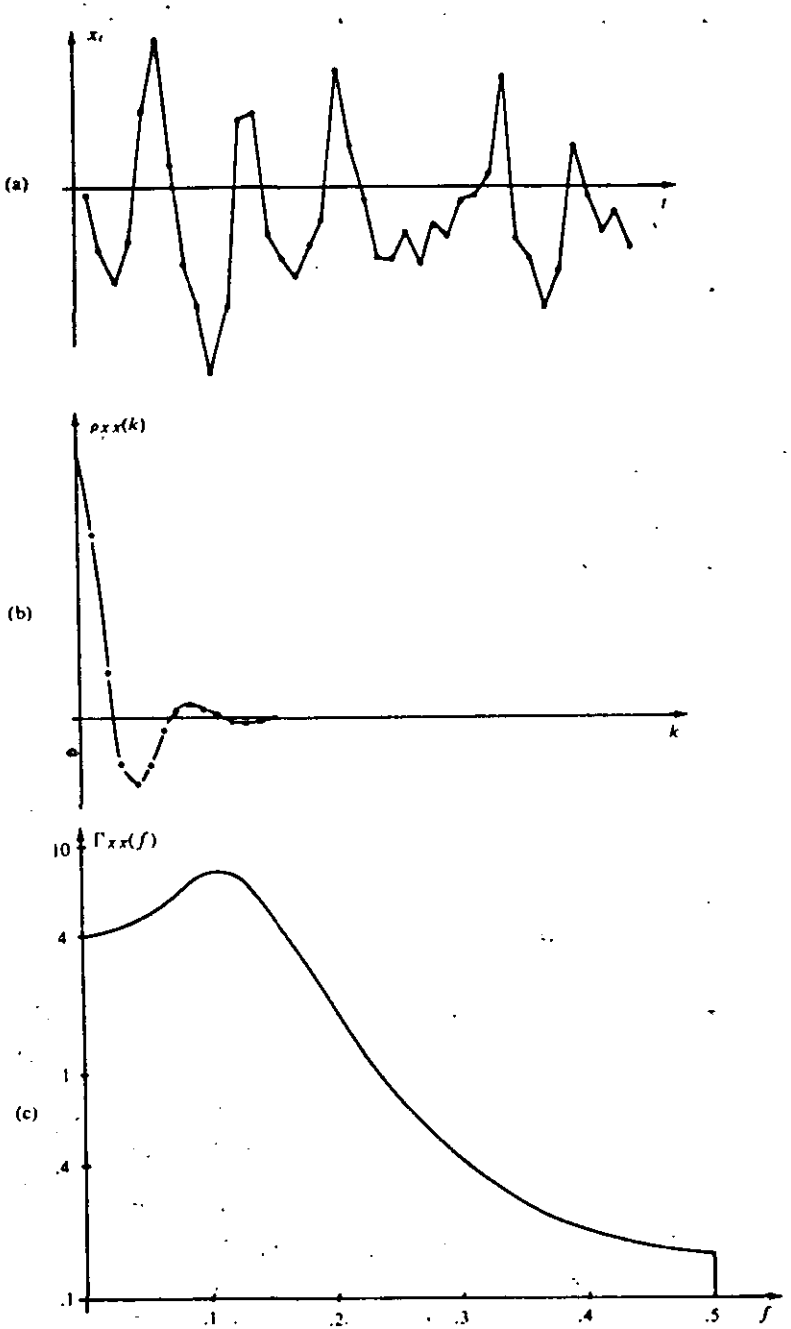


FIG. 6.6: A realization (a), the autocorrelation function (b) and spectrum (c) of a discrete second-order autoregressive process ($\alpha_1 = 1.0$, $\alpha_2 = -0.5$)

Now, the spectral density function, being the limit of a non-negative function, is also non-negative. Since it also integrates to unity, it follows that from a mathematical point of view it satisfies the same properties (3.1.8) as a pdf. It will be shown in Section 6.3 that the similarities between the spectral density function and the pdf also extend to the estimation of the two functions from finite lengths of records.

The method used to define the spectrum in this section is by no means unique. An alternative method based on the latent roots of the covariance matrix of the stochastic process is given in Section 11.1.2.

Comments on definitions of the spectrum used by engineers. Criticisms have already been given in Section 6.1.1 of the definition

$$\Gamma(f) = \lim_{T\to\infty} C_{xx}(f)$$

of the spectrum which is usually given in electrical engineering texts, for example [2] and [3]. The objection to this definition is that if $x(t)$ is a realization of a stationary stochastic process, then the corresponding random variable $C_{xx}(f)$ does not converge in any statistical sense to a limiting value.

A further source of confusion stems from the fundamental identity (6.1.9) proved above. It is falsely argued that since the sample autocovariance function $c_{xx}(u)$ tends in a well-behaved statistical sense to $\gamma_{xx}(u)$ as T tends to infinity, then it is permissible to state that

$$\begin{aligned}\lim_{T\to\infty} C_{xx}(f) &= \int_{-T}^{T} \lim_{T\to\infty} c_{xx}(u)\, e^{-j2\pi fu}\, du \\ &= \int_{-\infty}^{\infty} \gamma_{xx}(u)\, e^{-j2\pi fu}\, du \\ &= \Gamma_{xx}(f).\end{aligned}$$

As shown in Section 5.3.3, it is true that the mean squared error of the sample acvf estimator $c_{xx}(u)$ is of order $1/T$ and hence its distribution tends to be more clustered about $\gamma_{xx}(u)$ as T tends to infinity. Thus $c_{xx}(u)$ is a *consistent estimator* of $\gamma_{xx}(u)$. Another way of stating this fact is that the *ensemble average* $\gamma_{xx}(u)$ can be estimated by the *time average* $c_{xx}(u)$. This is usually referred to as the *ergodic property* and requires that $\gamma_{xx}(u)$ tend to zero at a sufficiently fast rate.

However, the fact that the ergodic property applies to $c_{xx}(u)$ *in no way implies that it holds for its Fourier transform* $C_{xx}(f)$. In fact, it is usually not true that, if there is a consistent estimator of a statistical parameter, its Fourier transform is a consistent estimator of the Fourier transform of that parameter. In other words $C_{xx}(f)$ is an example of a sample function for which the ergodic property does not hold.

An intuitive way of looking at this situation is to consider what happens to $c_{xx}(u)$ for fixed value of the lag u when the record length T increases. In this

case, more and more information in the form of products $x(t)x(t+u)$ are included in $c_{xx}(u)$ and hence the information about $\gamma_{xx}(u)$ contained in $c_{xx}(u)$ increases indefinitely at T tends to infinity. It will be seen later that the information contained in $C_{xx}(f)$ concerning $\Gamma_{xx}(f)$ is spread over a band of frequencies with effective width $\pm 1/T$ about f. As T increases, the total information contained in $C_{xx}(f)$ is distributed over an increasing number of bands of decreasing width. The net result is that as T increases it is possible to estimate the average power in narrower and narrower frequency bands: however, the efficiency of the estimate of the power in the narrowing band does not improve.

6.2.2 The integrated spectrum

Cases where no spectrum exists. In the previous section the power spectrum was defined by

$$\lim_{T\to\infty} E[C_{xx}(f)] = \lim_{T\to\infty} \int_{-T}^{T} \gamma_{xx}(u) \left(1 - \frac{|u|}{T}\right) e^{-j2\pi fu}\, du,$$

provided this limit exists. Clearly, if $\Gamma_{xx}(f)$ is to be finite it is sufficient that

$$\begin{aligned}|\Gamma_{xx}(f)| &= \left| \int_{-\infty}^{\infty} \gamma_{xx}(u)\, e^{-j2\pi fu}\, du \right| \\ &\leqslant \int_{-\infty}^{\infty} |\gamma_{xx}(u)|\, du \leqslant M,\end{aligned} \tag{6.2.9}$$

where M is a finite constant. Hence if the spectrum is to be finite, it is sufficient (but not necessary) for $\gamma_{xx}(u)$ to tend to zero as the lag u tends to infinity at a sufficiently fast rate so that the integral (6.2.9) is finite.

As an example of a stochastic process for which this condition is not met, consider the process

$$X(t) = A\cos 2\pi f_0 t + B\sin 2\pi f_0 t = R\cos(2\pi f_0 t + \phi), \tag{6.2.10}$$

where A and B are independent rv's with mean zero and variance σ^2. For a given realization, $x(t)$ is a cosine wave, $R\cos(2\pi f_0 t + \phi)$ having constant amplitude, frequency and phase. Across the ensemble, the amplitude and phase vary randomly, while the frequency remains fixed. From (6.2.10),

$$E[X(t)] = E[A]\cos 2\pi f_0 t + E[B]\sin 2\pi f_0 t = 0.$$

Hence

$$\begin{aligned}\gamma_{xx}(u) &= E[X(t)X(t+u)] \\ &= E[(A\cos 2\pi f_0 t + B\sin 2\pi f_0 t)(A\cos 2\pi f_0(t+u) + B\sin 2\pi f_0(t+u))] \\ &= \sigma^2[\cos 2\pi f_0 t\cos 2\pi f_0(t+u) + \sin 2\pi f_0 t\sin 2\pi f_0(t+u)] \\ &= \sigma^2\cos 2\pi f_0 u.\end{aligned}$$

$\gamma_{xx}(u)$ does not tend to zero as u tends to infinity, and so the integral (6.2.9) diverges. However, a spectrum can be defined in terms of delta functions, using (2.2.12), so that

$$\Gamma_{xx}(f) = \frac{\sigma^2}{2}[\delta(f - f_0) + \delta(f + f_0)].$$

Hence the spectrum of the stochastic process (6.2.10) may be regarded as two delta functions of area $\sigma^2/2$ centered at the frequencies $f = \pm f_0$.

The integrated spectrum. Even when the spectrum contains delta functions, it is meaningful to talk about the variance of the series accounted for by frequencies less than or equal to some frequency f'. This may be obtained formally by integrating the expression for the spectrum. For example, integrating (6.2.2) from $f = -f'$ to $f = f'$ gives the *integrated spectrum*

$$I_{xx}(f') = \int_{-f'}^{f'} \Gamma_{xx}(f)\, df$$

$$= \int_{-\infty}^{\infty} \gamma_{xx}(u) \frac{\sin 2\pi f' u}{\pi u}\, du, \quad 0 \leqslant f' \leqslant \infty. \tag{6.2.11}$$

This function is analogous to the cumulative distribution function in the same way that the spectral density function is analogous to the probability density function. Thus

$$I(0) = 0,$$
$$I(\infty) = \sigma^2$$

and

$$I(f_1) \leqslant I(f_2)$$

when $f_1 \leqslant f_2$. If the spectral density has a delta function at $f = f_0$, that is, $\gamma_{xx}(u)$ contains a component $k \cos 2\pi f_0 u$, then the integrated power spectrum jumps by an amount k at the frequency f_0.

For the discrete case, the integrated spectrum is

$$I_{xx}(f') = \int_{-f'}^{f'} \Gamma_{xx}(f)\, df$$

$$= \Delta \sum_{k=-\infty}^{\infty} \gamma_{x'x}(k) \frac{\sin 2\pi k f' \Delta}{\pi k \Delta}, \quad 0 \leqslant f' < \frac{1}{2\Delta}. \tag{6.2.12}$$

6.2.3 *The spectrum of white noise*

In Section 5.2.1, the completely random or white noise process $Z(t)$ was defined as one with an acvf $\gamma_{zz}(u) = \sigma_z^2\, \delta(u)$. This has infinite variance and as such cannot represent a valid stochastic process. However, it was

shown that it could be regarded as the limit as τ tends to zero of the Bachelier–Wiener process $Y(t)$ which has the acvf

$$\gamma_{YY}(u) = \begin{cases} 0, & |u| > \tau \\ \dfrac{\sigma_z^2}{\tau}\left(1 - \dfrac{|u|}{\tau}\right), & |u| \leqslant \tau. \end{cases}$$

Hence using the definition (6.2.2), the Bachelier–Wiener process has the spectrum

$$\Gamma_{YY}(f) = \int_{-\tau}^{\tau} \frac{\sigma_z^2}{\tau}\left(1 - \frac{|u|}{\tau}\right) e^{-i2\pi f u}\, du$$

$$= \sigma_z^2 \left(\frac{\sin \pi f \tau}{\pi f \tau}\right)^2, \quad -\infty \leqslant f \leqslant \infty.$$

In the limit as τ tends to zero, $\Gamma_{YY}(f)$ tends to

$$\lim_{\tau \to 0} \Gamma_{YY}(f) = \Gamma_{zz}(f) = \sigma_z^2, \tag{6.2.13}$$

that is, a constant for all f.

The $Z(t)$ process is termed *white noise* by analogy with the optical spectrum of white light, which has all optical frequencies present with approximately the same intensity. True white noise is as physically impossible as an impulse function; indeed, it may be regarded as the statistician's analog of the engineer's impulse function.

Methods of generating white noise. There are no difficulties in defining white noise in discrete time since the acvf of discrete white noise Z_t is

$$\gamma_{zz}(u) = \begin{cases} \sigma_z^2, & u = 0 \\ 0, & u = \pm\Delta, \pm 2\Delta, \pm 3\Delta, \ldots. \end{cases}$$

Using (6.2.6) it follows that

$$\Gamma_{zz}(f) = \sigma_z^2 \Delta, \quad -\frac{1}{2\Delta} \leqslant f < \frac{1}{2\Delta}. \tag{6.2.14}$$

so that all frequencies in the range $-1/2\Delta \leqslant f < 1/2\Delta$ contribute the same amount of power or variance.

Discrete white noise may be generated from non-white continuous noise very simply. For example, suppose a source of continuous non-white noise is available which has an acvf which is zero when $u > u_0$. Clearly, sampling the $X(t)$ process at spacing $\Delta > u_0$ produces a Z_t process with the same acvf as (6.2.14).

The frequency-domain interpretation of this method of generating discrete white noise from continuous non-white noise is that the sampling frequency $1/\Delta$ is so low that many many aliasings of the spectrum $\Gamma_{xx}(f)$ occur (see Section 2.4.2). Hence, the spectrum of the discrete (sampled)

signal which is the sum of the aliased segments tends to a flat-topped function, that is, $\Gamma_{zz}(f)$ tends to a constant in the range $-1/2\Delta \leqslant f < 1/2\Delta$. The flattening of the aliased spectrum is demonstrated in Figure 2.11 for one particular case. Note that in discussing white noise, nothing has been said about the pdf of $Z(t)$, that is, *white noise may have any amplitude pdf whatever.*

True white noise can never exist physically, but very good approximations to it can and do. For example, the fluctuating current in an electron tube provides a very good approximation, having an essentially flat power spectrum from 0 to 100 megacycles per second. This noise is usually referred to as *shot noise* and is due to the random emission of electrons from the cathode of the tube.

Another example of a physical source of noise which is approximately white over a wide frequency range is *thermal noise.* This is the voltage or current in a wire of resistance R due to the thermal motion of electrons. Its power spectrum is approximately constant over a wide range of frequencies and is equal to

$$\Gamma_{xx}(f) = 4RkT,$$

where T is the absolute temperature and k is Boltzmann's constant. A more detailed discussion of shot and thermal noise is given in [2].

6.2.4 *The spectrum of a linear process*

An expression is now derived for the spectrum of the output from a stable linear system when the input is a stationary process. When the input is white noise, the output spectrum is the spectrum of a stationary linear process.

Consider the output process $X(t)$ from a stable linear system with impulse response $h(u)$ and input process $Z(t)$. From (5.2.8), the acvf of the $X(t)$ process is

$$\gamma_{xx}(u) = \int_0^\infty \int_0^\infty h(v)h(v')\gamma_{zz}(u + v - v')\, dv\, dv',$$

and hence from (6.2.2) the spectrum of the output is

$$\begin{aligned}\Gamma_{xx}(f) &= \int_{-\infty}^{\infty} \gamma_{xx}(u)\, e^{-j2\pi fu}\, du \\ &= \int_{-\infty}^{\infty} e^{-j2\pi fu} \int_0^\infty \int_0^\infty h(v)h(v')\gamma_{zz}(u + v - v')\, dv\, dv'\, du \\ &= \int_0^\infty h(v)\, e^{j2\pi fv}\, dv \int_0^\infty h(v')\, e^{-j2\pi fv'}\, dv' \int_{-\infty}^{\infty} \gamma_{zz}(y)\, e^{-j2\pi fy}\, dy,\end{aligned}$$

where $y = u + v - v'$.

Hence

$$\begin{aligned}\Gamma_{xx}(f) &= H(-f)H(f)\Gamma_{zz}(f) \\ &= |H(f)|^2\Gamma_{zz}(f), \quad -\infty \leqslant f \leqslant \infty. \end{aligned} \tag{6.2.15}$$

This fundamental property states that *the spectrum of the output from a linear system is obtained from the spectrum of the input by multiplying by the square of the modulus of the frequency response function.*

If $Z(t)$ is white noise with spectrum $\Gamma_{zz}(f) = \sigma_z^2$, acvf $\gamma_{zz}(u) = \sigma_z^2\, \delta(u)$, then $X(t)$ is a linear process with spectrum

$$\Gamma_{xx}(f) = \sigma_z^2 |H(f)|^2, \quad -\infty \leqslant f \leqslant \infty. \tag{6.2.16}$$

In discrete time, the relation corresponding to (6.2.15) is

$$\Gamma_{xx}(f) = |H(f)|^2\Gamma_{zz}(f), \quad -\frac{1}{2\Delta} \leqslant f < \frac{1}{2\Delta}, \tag{6.2.17}$$

where $H(f) = \sum_{k=0}^{\infty} h_k\, e^{-j2\pi fk\Delta}$.

When the input is the completely random process with variance σ_z^2, the output is a linear process with spectrum

$$\Gamma_{xx}(f) = \Delta\sigma_z^2 |H(f)|^2, \quad -\frac{1}{2\Delta} \leqslant f < \frac{1}{2\Delta}. \tag{6.2.18}$$

From (6.2.15) or (6.2.18) it is seen that, given a white noise source and a suitably versatile analog (or digital) filter it is possible to generate a stochastic process with any given spectrum. In the next section some examples are given of the wide variety of spectra which may be generated by linear filtering of white noise.

6.2.5 *The spectra of autoregressive and moving average processes*

Continuous first-order autoregressive processes. Consider the continuous first-order ar process

$$T\frac{dX(t)}{dt} + (X(t) - \mu) = Z(t),$$

where $Z(t)$ is white noise. This has the impulse response function

$$h(u) = \begin{cases} \frac{1}{T}\, e^{-u/T}, & 0 \leqslant u \leqslant \infty \\ 0, & u < 0 \end{cases}$$

and frequency response function

$$H(f) = \frac{1}{1 + j2\pi fT}.$$

Hence, using (6.2.16), the spectrum of $X(t)$ is

$$\Gamma_{xx}(f) = \frac{\sigma_z^2}{1 + (2\pi fT)^2}, \quad -\infty \leqslant f \leqslant \infty. \tag{6.2.19}$$

The function (6.2.19) has been plotted in Figure 2.3 (a), which shows that most of the power or variance is concentrated at low frequencies.

Discrete first-order autoregressive processes. In discrete time, the first-order ar process is

$$X_t - \mu = \alpha_1(X_{t-1} - \mu) + Z_t,$$

with

$$h_k = \alpha_1^k, \quad k = 0, 1, \ldots, \infty,$$

and

$$H(f) = \frac{1}{1 - \alpha_1 e^{-j2\pi f\Delta}}, \quad -\frac{1}{2\Delta} \leqslant f < \frac{1}{2\Delta}.$$

Hence, using (6.2.18), the spectrum of the X_t process is

$$\Gamma_{XX}(f) = \frac{\Delta\sigma_Z^2}{1 + \alpha_1^2 - 2\alpha_1 \cos 2\pi f\Delta}, \quad -\frac{1}{2\Delta} \leqslant f < \frac{1}{2\Delta}. \tag{6.2.20}$$

The spectrum (6.2.20) has been plotted in Figures 6.4 and 6.5 for $\alpha_1 = +0.9$ and $\alpha_1 = -0.9$ respectively, $\Delta = 1$ and $\sigma_Z^2 = 1$. As discussed in Section 6.2.1, when α_1 is positive the spectrum has most power at low frequencies and when α_1 is negative the spectrum has most power at high frequencies. Note from (6.2.20) that $\Gamma_{XX}(f)$ for $\alpha_1 > 0$ equals $\Gamma_{XX}(1/2\Delta - f)$ for $\alpha_1 < 0$.

Continuous second-order autoregressive processes. Consider the continuous second-order ar process

$$a_2 \frac{d^2X}{dt^2} + a_1 \frac{dX}{dt} + a_0(X(t) - \mu) = Z(t).$$

This has the frequency response function

$$H(f) = \frac{1}{a_2(j2\pi f)^2 + a_1(j2\pi f) + a_0},$$

and hence the spectrum

$$\Gamma_{XX}(f) = \frac{\sigma_Z^2}{(a_0 - a_2 4\pi^2 f^2)^2 + (2\pi f a_1)^2}. \tag{6.2.21}$$

In addition to generating low-frequency spectra (a_1 or a_2 large), (6.2.21) can also generate a spectrum with a peak if the roots of the characteristic equation $a_2p^2 + a_1p + a_0 = 0$ are complex.

Discrete second-oraer autoregressive processes. The discrete second-order ar process (5.2.31), namely,

$$X_t - \mu = \alpha_1(X_{t-1} - \mu) + \alpha_2(X_{t-2} - \mu) + Z_t,$$

has the frequency response function

$$H(f) = \frac{1}{1 - \alpha_1 e^{-j2\pi f\Delta} - \alpha_2 e^{-j4\pi f\Delta}}, \quad -\frac{1}{2\Delta} \leqslant f < \frac{1}{2\Delta},$$

and hence the spectrum

$$\Gamma_{XX}(f) = \frac{\Delta\sigma_Z^2}{1 + \alpha_1^2 + \alpha_2^2 - 2\alpha_1(1 - \alpha_2)\cos 2\pi f\Delta - 2\alpha_2 \cos 4\pi f\Delta}, \quad -\frac{1}{2\Delta} \leqslant f < \frac{1}{2\Delta}. \tag{6.2.22}$$

For certain values of the parameters α_1, α_2, (6.2.22) can give rise to a low-frequency or high-frequency spectrum like the first-order discrete process. In addition to these spectra, it is possible to obtain spectra with a peak or a trough at an intermediate frequency f_0. This occurs when $|\alpha_1(1 - \alpha_2)| < |4\alpha_2|$. The frequency f_0 at which the peak or trough occurs is given by

$$\cos 2\pi f_0\Delta = -\frac{\alpha_1(1 - \alpha_2)}{4\alpha_2}.$$

For example, the time series shown in Figure 6.6 is based on a second-order process with parameters $\alpha_1 = 1$, $\alpha_2 = -0.5$ which produces a spectrum with a peak at $f_0 = 0.125/\Delta$ cps.

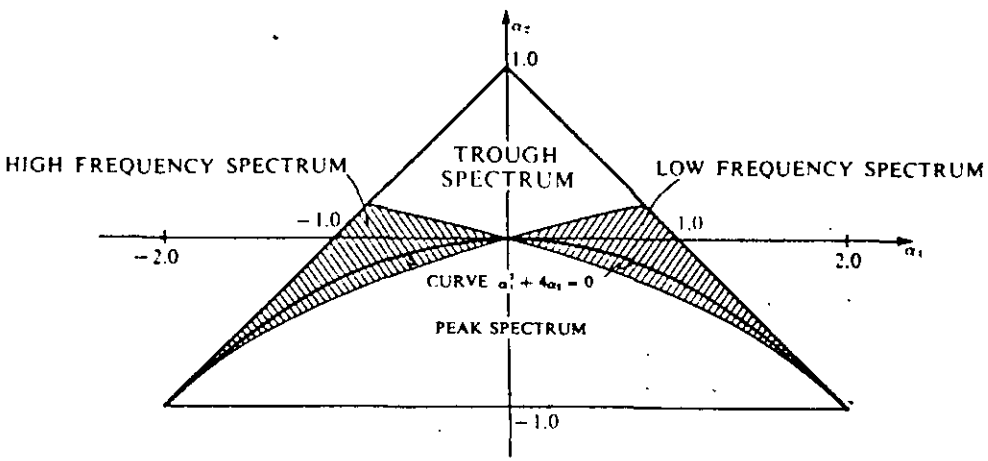


FIG. 6.7: Stability region and classification of spectra for discrete second-order autoregressive processes

The four types of spectra which can be generated with a second-order ar process are summarized in Figure 6.7. An interesting feature which is brought out by this diagram is that the region for which $\alpha_1^2 + 4\alpha_2 < 0$ (and hence for which the acf is a damped sine wave) lies partially in the region for which $|\alpha_1(1 - \alpha_2)| \geqslant |4\alpha_2|$ where the spectrum has no intermediate peak. For the high-frequency spectrum this is not surprising, since a first-order ar process

with $\alpha_1 < 0$ has an oscillatory acf even though the spectrum does not have an intermediate peak. For the low-frequency spectrum, however, the acf can be oscillatory and yet no intermediate peak is evident. It is commonly assumed that a periodicity in the acf will appear as a peak in the spectrum, but this example shows that the amplitude of the damped sine wave must be large enough.

General autoregressive–moving average processes. The general continuous ar–ma process (5.2.51)

$$a_m \frac{d^m X}{dt^m} + \cdots + a_1 \frac{dX}{dt} + a_0(X(t) - \mu) = b_l \frac{d^l Z}{dt^l} + \cdots + b_1 \frac{dZ}{dt} + b_0 Z(t)$$

has the power spectrum

$$\Gamma_{XX}(f) = \sigma_Z^2 \left| \frac{b_0 + b_1 j2\pi f + \cdots + b_l (j2\pi f)^l}{a_0 + a_1 j2\pi f + \cdots + a_m (j2\pi f)^m} \right|^2, \quad -\infty \leqslant f \leqslant \infty. \tag{6.2.23}$$

Similarly, the discrete mixed process (5.2.50), namely

$$X_t - \mu = \alpha_1(X_{t-1} - \mu) + \cdots + \alpha_m(X_{t-m} - \mu) + Z_t + \beta_1 Z_{t-1} + \cdots + \beta_l Z_{t-l},$$

has the spectrum

$$\Gamma_{XX}(f) = \Delta\sigma_Z^2 \left| \frac{1 + \beta_1 e^{-j2\pi f \Delta} + \cdots + \beta_l e^{-j2\pi f \Delta l}}{1 - \alpha_1 e^{-j2\pi f \Delta} - \cdots - \alpha_m e^{-j2\pi f \Delta m}} \right|^2, \quad -\frac{1}{2\Delta} \leqslant f < \frac{1}{2\Delta}. \tag{6.2.24}$$

The expression (6.2.23) for the spectrum of the continuous process shows that if $\Gamma_{XX}(f)$ is to be a proper spectrum in the sense that it integrates out to a finite constant σ_X^2 corresponding to the variance of the $X(t)$ process, then l must satisfy $l \leqslant m - 1$. Note that there is no restriction on l for the discrete case.

The results (6.2.23) and (6.2.24) are obtained by substituting the frequency response functions (2.3.19) and (2.3.32) in (6.2.15) and (6.2.18) respectively. In general, these spectra may have multiple peaks or troughs if the roots of the corresponding characteristic equations are complex.

6.3 SPECTRAL ESTIMATORS

6.3.1 *Distribution properties of sample spectral estimators for white noise*

Introduction. Table 6.1 suggests that the sample spectrum estimator

$$C_{zz}(f) = \frac{\Delta}{N}\left[\left\{\sum_{t=-n}^{n-1} Z_t \cos 2\pi f t\Delta\right\}^2 + \left\{\sum_{t=-n}^{n-1} Z_t \sin 2\pi f t\Delta\right\}^2\right], \quad -\frac{1}{2\Delta} \leqslant f < \frac{1}{2\Delta}, \tag{6.3.1}$$

for a purely random discrete process (discrete white noise) has a variance which is independent of the number of observations N. On the other hand, the average of the sample spectrum over frequency is close to the theoretical value of the spectrum. These results suggest that the sample spectrum is not a consistent estimator in the sense that its distribution does not tend to cluster more closely about the true spectrum as the sample size increases.

To see why this is so, consider the rv's associated with the real and imaginary Fourier components of a discrete process Z_t, $-n \leqslant t \leqslant n-1$. These are

$$A(f) = \sum_{t=-n}^{n-1} Z_t \cos 2\pi f t\Delta,$$

$$B(f) = \sum_{t=-n}^{n-1} Z_t \sin 2\pi f t\Delta, \quad -\frac{1}{2\Delta} \leqslant f < \frac{1}{2\Delta}. \tag{6.3.2}$$

The sample spectrum estimator (6.3.1) is then

$$C_{zz}(f) = \frac{\Delta}{N}[A^2(f) + B^2(f)], \quad -\frac{1}{2\Delta} \leqslant f < \frac{1}{2\Delta}. \tag{6.3.3}$$

By investigating the properties of the rv's $A(f)$ and $B(f)$ it is possible to derive the sampling properties of the sample spectrum. In this section it is shown that if the Z_t process is a purely random Normal process with mean zero and variance σ_Z^2, then at the harmonic frequencies $f_k = k/N\Delta$:

(1) The random variables

$$Y(f_k) = \frac{2C_{zz}(f_k)}{\Delta\sigma_Z^2}, \quad k = \pm 1, \pm 2, \ldots, \pm(n-1), \tag{6.3.4}$$

are distributed as χ_2^2 rv's.

(2) When $f_k = 0$ or $f_k = -1/2\Delta$, the rv's

$$Y(f_k) = \frac{C_{zz}(f_k)}{\Delta\sigma_Z^2} \tag{6.3.5}$$

are distributed as a χ_1^2.

(3) The rv's $Y(f_k)$ are mutually independent for $k = 0, \pm 1, \pm 2, \ldots, \pm(n-1), -n$.

These results will be used to derive a test for white noise in Section 6.3.2. Section 6.3.3 contains a summary of more general results for the distribution properties of the sample spectrum estimators which apply for all frequencies f and for non-Normal and non-white processes. Proofs of these results are given in Appendix A9.1.

The chi-squared property of the sample spectrum estimator. Since $E[Z_t] = 0$, it follows from (6.3.2) that

$$E[A(f)] = 0 = E[B(f)].$$

Hence at the harmonic frequencies $f_k = k/N\Delta$,

$$\text{Var}\,[A(f_k)] = E[A^2(f_k)]$$

$$= \sigma_Z^2 \sum_{t=-n}^{n-1} \cos^2 2\pi f_k t\Delta = \begin{cases} \sigma_Z^2 \dfrac{N}{2}, & k = \pm 1, \pm 2, \ldots, \pm(n-1) \\ \sigma_Z^2 N, & k = 0, -n. \end{cases} \tag{6.3.6}$$

Similarly,

$$\text{Var}\,[B(f_k)] = \begin{cases} \sigma_Z^2 \dfrac{N}{2}, & k = \pm 1, \pm 2, \ldots, \pm(n-1) \\ 0, & k = 0, -n. \end{cases} \tag{6.3.7}$$

Furthermore, when $k \neq l$

$$\text{Cov}\,[A(f_k)A(f_l)] = \sigma_Z^2 \sum_{t=-n}^{n-1} \cos 2\pi f_k t\Delta \cos 2\pi f_l t\Delta$$

$$= 0, \tag{6.3.8}$$

$$\text{Cov}\,[B(f_k), B(f_l)] = 0.$$

In addition, for all k and l

$$\text{Cov}\,[A(f_k), B(f_l)] = 0. \tag{6.3.9}$$

Now since $A(f_k)$ and $B(f_k)$ are linear functions of Normal rv's, $A(f_k)$ and $B(f_k)$ are also distributed Normally. Hence the rv's

$$\frac{A^2(f_k)}{\text{Var}\,[A(f_k)]} = \frac{2A^2(f_k)}{N\sigma_Z^2}, \quad \frac{B^2(f_k)}{\text{Var}\,[B(f_k)]} = \frac{2B^2(f_k)}{N\sigma_Z^2}$$

are each distributed as χ_1^2. Moreover, (6.3.8) and (6.3.9) show that these rv's are uncorrelated and thus independent, since $A(f_k)$ and $B(f_k)$ are Normal rv's. Hence their sum

$$\frac{2}{N\sigma_Z^2}(A^2(f_k) + B^2(f_k)) = \frac{2C_{zz}(f_k)}{\Delta\sigma_Z^2} = Y(f_k)$$

is distributed as χ_2^2.

When $k = 0, -n$, $B(f_k)$ is identically zero. Hence

$$Y(f_k) = \frac{A^2(f_k)}{\text{Var}\,[A(f_k)]} = \frac{C_{zz}(f_k)}{\Delta\sigma_Z^2}, \quad k = 0, -n,$$

is distributed as χ_1^2. Equations (6.3.8) and (6.3.9) imply that the rv's $Y(f_k)$ at distinct frequencies are uncorrelated and hence independent because of the Normal assumption. Thus the results (1), (2) and (3) mentioned above have been established.

One immediate consequence of these results is that the erratic behavior of the sample spectrum in Figure 6.1 can be explained. Thus, for the purely random process, it was shown in Section 6.2.3 that the spectrum is

$$\Gamma_{zz}(f) = \sigma_Z^2\Delta, \quad -\frac{1}{2\Delta} \leq f < \frac{1}{2\Delta}.$$

Using (3.3.6) and the results just derived,

$$E\left[\frac{2C_{zz}(f_k)}{\Delta\sigma_Z^2}\right] = 2,$$

that is,

$$E[C_{zz}(f_k)] = \sigma_Z^2\Delta = \Gamma_{zz}(f_k).$$

Hence, the sample spectrum at the harmonic frequencies provides an *unbiased estimator* of the spectrum for white noise. This explains why the mean values in Table 6.1 are close to the theoretical value.

Similarly, using (3.3.6),

$$\text{Var}\left[\frac{2C_{zz}(f_k)}{\Delta\sigma_Z^2}\right] = 4,$$

that is,

$$\text{Var}\,[C_{zz}(f_k)] = \sigma_Z^4\Delta^2 = \Gamma_{zz}^2(f_k). \tag{6.3.10}$$

Equation (6.3.10) shows that at the harmonic frequencies, at least, the variance of the estimator is a constant *independent of the sample size*. This explains the failure of the sample estimates of variance of $C_{zz}(f_k)$ to decrease with sample size as shown in Table 6.1. It is important to note that even if the Z_t process is not Normal, the rv's $A(f)$ and $B(f)$ will be very nearly Normal by the Central Limit Theorem. Hence the distribution of $C_{zz}(f)$ will be very nearly a χ_2^2 regardless of the distribution of the Z_t process.

Analysis of variance. The significance of the above results may be more readily appreciated by considering the decomposition of the total sum of squares of the rv's Z_t. Thus, using Parseval's theorem (6.1.3),

$$\sum_{t=-n}^{n-1} Z_t^2 = \sum_{t=-n}^{n-1} \frac{C_{zz}(f_k)}{\Delta}.$$

Making use of the fact that $C_{zz}(f_k) = C_{zz}(-f_k)$,

$$\frac{1}{\sigma_Z^2}\sum_{t=-n}^{n-1} Z_t^2 = \frac{1}{\Delta\sigma_Z^2}[C_{zz}(0) + 2\sum_{k=1}^{n-1} C_{zz}(f_k) + C_{zz}(f_n)]. \tag{6.3.11}$$

Since the Z_t/σ_Z are independent Normal rv's with zero means and unit standard deviations, the left-hand side of (6.3.11) will be distributed as χ_N^2. The above results then show that this χ_N^2 rv is decomposed into two χ_1^2 rv's and $(n-1)\,\chi_2^2$ rv's. Thus the total degrees of freedom is decomposed according to

$$N = 2n = 1 + 2(n-1) + 1.$$

When N is odd, the single degree of freedom component corresponding to $k = -n$ does not appear in (6.3.11). The decomposition is a special case of a technique called the *analysis of variance* by statisticians. When $E[Z_t] \neq 0$, the

above analysis holds, but now the decomposition (6.3.11) is more conveniently written

$$\frac{1}{\sigma_z^2}\sum_{t=-n}^{n-1}(Z_t-\bar{Z})^2=\frac{1}{\Delta\sigma_z^2}\left[2\sum_{k=1}^{n-1}C_{zz}(f_k)+C_{zz}(f_n)\right], \qquad (6.3.12)$$

where $\bar{Z}$ is the sample mean of the rv's Z_t.

6.3.2 *A test for white noise*

Need for a test. Situations often occur in practice where it is necessary to test whether an observed time series could be regarded as a realization of a white noise process. One such example has been given in Section 5.3.5, where a test for white noise was applied to the Normal deviates generated by an automatic computer. Another example is to test the adequacy of a fitted model, for example, the ar process (5.2.39). The model can be regarded as being adequate if the residuals from the fitted model constitute a white noise process.

The test for white noise given in Section 5.3.5 is useful when one wants to detect "local correlation," that is, whether neighboring points of the time series are correlated. It is sometimes necessary to detect departures from whiteness caused by periodic effects. For example, after fitting a model to an economic time series containing seasonal variation, the inadequacy of the model might reflect itself in residuals which are periodic. In this situation a frequency domain test based on the sample spectrum is more appropriate. Such a test will now be described and should be regarded as complementary to the test based on the acf described in Section 5.3.5.

The test criterion. Equation (6.2.14) shows that the spectrum of a discrete white noise process is

$$\Gamma_{zz}(f)=\Delta\sigma_z^2, \quad -\frac{1}{2\Delta}\leqslant f<\frac{1}{2\Delta}.$$

Hence the integrated spectrum

$$I_{zz}(f)=\int_{-f}^{f}\Gamma_{zz}(g)\,dg=2\Delta\sigma_z^2 f, \quad 0\leqslant f\leqslant\frac{1}{2\Delta},$$

is a linear function of frequency.

Suppose that the sample spectrum estimate $C_{zz}(f)$ has been computed at the harmonic frequencies $f_k=k/N\Delta$, $k=0,1,\ldots,N/2$. Then $I(f_k)$ may be estimated from the sample integrated spectrum

$$I(f_k)=\frac{1}{N\Delta}\sum_{l=1}^{k}C_{zz}(f_l). \qquad (6.3.13)$$

Note that $C_{zz}(0)$ is zero if the mean is subtracted out. Since $E[C_{zz}(f_k)]=\Gamma_{zz}(f_k)=2\Delta\sigma_z^2$, it follows that

$$E[I(f_k)]=2\Delta\sigma_z^2 f_k=2\Delta\sigma_z^2\frac{k}{N\Delta},$$

and hence $I(f_k)$ is an unbiased estimator of $I_{zz}(f_k)$. It is convenient in practice to normalize $I(f_k)$ by dividing by σ_z^2: in this case, $I(1/2\Delta)=1$. In practice σ_z^2 is not known and has to be replaced by its estimator S_z^2 so that the final form of the estimator is $I(f_k)/S_z^2$. Thus if the estimate

$$\frac{I(f_k)}{s_z^2}=\frac{1}{N\Delta s_z^2}\sum_{l=1}^{k}C_{zz}(f_l)$$

obtained from an observed time series is plotted against $2\Delta f_k$, the points should be scattered about the straight line through the points (0, 0), (1, 1). Since $I(f_k)$ is the sum of rv's with the same distribution, the Kolmogoroff-Smirnov probability limits [4] applicable to cumulative distribution functions can be used to assess when a significant departure from linearity occurs.

Two examples. Table 6.2 gives the values of $C_{zz}(f_k)$ for one of the samples of random Normal deviates used for the calculations of Table 6.1. Here $N=100$, $\Delta=1$ and hence $f_k=0.01, 0.02, \ldots, 0.50$. Figure 6.8 shows the plot of $I(f_k)/s^2$ versus k for this series, and from it can be seen that the deviations from the straight line are not large. To judge these deviations more precisely, a test of significance due to Kolmogoroff and Smirnov [4] may be used when N is large. This consists of constructing a band $\pm\lambda/((N/2)-1)^{1/2}$ about the theoretical line. For significance levels of 0.95 and 0.75, λ is equal to 1.36 and 1.02 respectively. For the present example, $N/2=50$; hence the 95 percent

TABLE 6.2: Sample spectrum estimates at the harmonic frequencies for a sample of white noise

f_k	$C_{zz}(f_k)$	f_k	$C_{zz}(f_k)$	f_k	$C_{zz}(f_k)$
0.01	1.13	0.17	1.91	0.34	1.75
0.02	1.41	0.18	0.15	0.35	0.25
0.03	0.74	0.19	0.85	0.36	1.84
0.04	1.08	0.20	2.49	0.37	3.98
0.05	1.28	0.21	3.89	0.38	0.22
0.06	0.06	0.22	1.13	0.39	1.52
0.07	0.85	0.23	0.53	0.40	1.48
0.08	0.23	0.24	1.86	0.41	0.44
0.09	0.71	0.25	0.47	0.42	1.16
0.10	0.79	0.26	1.87	0.43	1.20
0.11	0.51	0.27	1.35	0.44	2.73
0.12	0.46	0.28	1.29	0.45	1.66
0.13	1.38	0.29	0.06	0.46	1.34
0.14	0.11	0.30	0.24	0.47	0.17
0.15	0.37	0.31	0.56	0.48	1.43
0.16	0.50	0.32	0.68	0.49	1.03
		0.33	0.44		

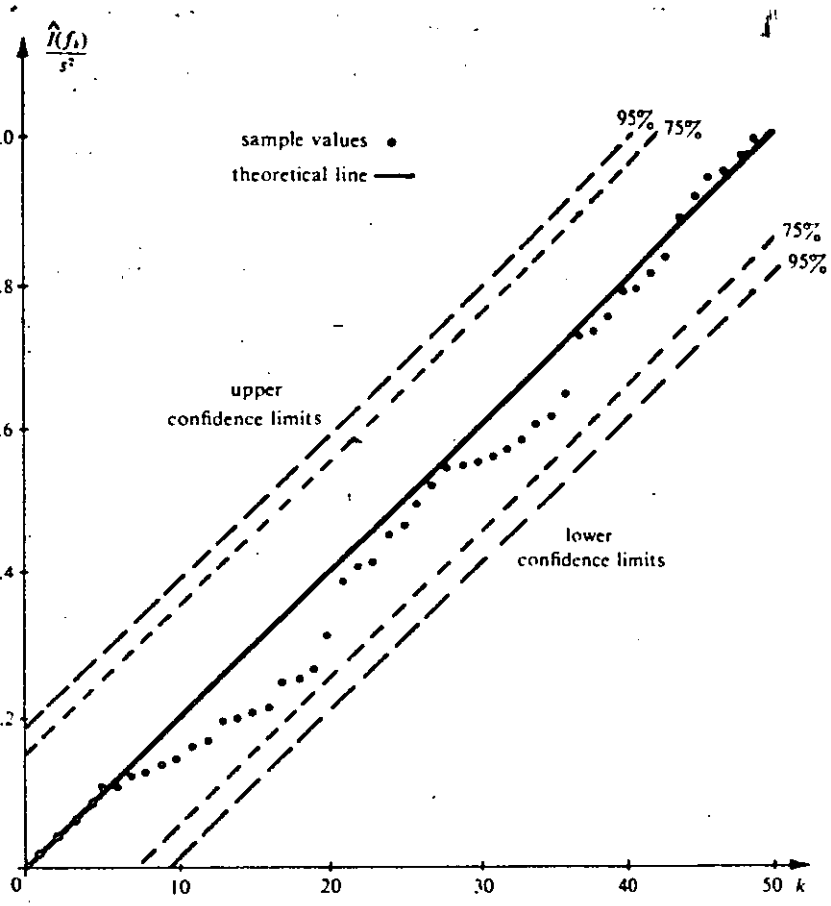


FIG. 6.8: A test for white noise using the sample integrated periodogram

ts are $\pm 1.36/\sqrt{49} = \pm 0.19$ and the 75 percent limits are ± 0.15. These ts are shown as broken lines in Figure 6.8, and it is seen that $I(f_k)/s^2$ falls within them. Hence there is no evidence that the sample does not come n a white noise source. The interpretation of the 75 percent limits, for mple, is that in one out of four plots, on average, the *maximum deviation* n the theoretical line will lie outside the limits if the process is in fact te noise.

Table 6.3 sets out the calculations for this test applied to the ionospheric a of Table 2.1, for which $s_z^2 = 196.4$. The values of $C_{zz}(f_k)$ in Table 6.3 y be obtained by multiplying the contributions to the mean square in le 2.2 by $N = 12$.

Figure 6.9 shows that the sample integrated spectrum deviates widely n a straight line, since $I(f_1)/s^2$ is approximately twice the expected value white noise and $I(f_2)/s^2$ is approximately three times the expected value.

The significance limits quoted above are not applicable here since N is so small. In fact, no test of significance is required in this case, since the values of $I(f_k)$ are so large when $f_k = 0.083$ and 0.166.

TABLE 6.3: White noise test applied to ionospheric data

f_k	$C_{zz}(f_k)$	$\frac{1}{Ns^2}\sum_{i=1}^{k} C_{zz}(f_i)$
0.083	753.6	0.32
0.166	1322.4	0.88
0.250	38.4	0.90
0.333	18.0	0.91
0.417	78.0	0.94
0.500	146.0	1.00

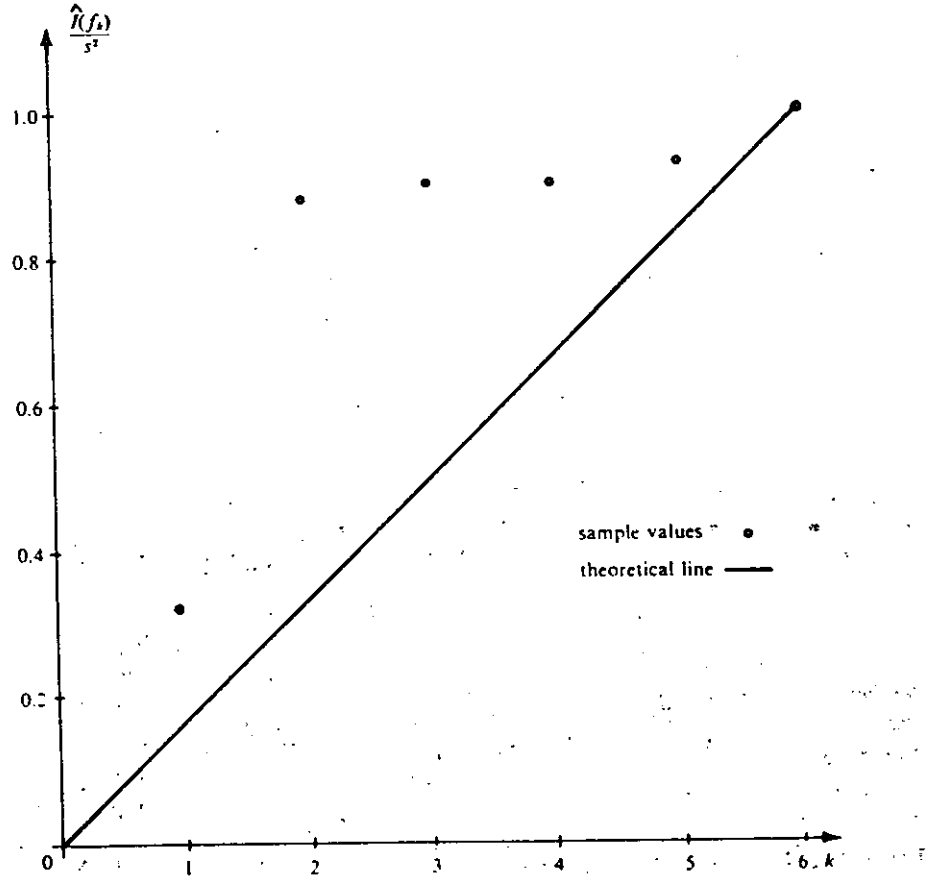


FIG. 6.9: Test of ionospheric data for randomness

6.3.3 General results for sample spectra for white noise

In Section 6.3.1 the mean and covariance of the sample spectrum estimator were derived for the harmonic frequencies $f_k = k/N\Delta$ under the assumption that the Z_t process is Normal. In Appendix A9.1 are derived more general results which apply for all frequencies and for non-Normal processes.

Moments of sample spectrum estimators for white noise. In discrete time these general results are

$$E[C_{zz}(f)] = \Gamma_{zz}(f) = \sigma_z^2\Delta, \quad -\frac{1}{2\Delta} \leqslant f < \frac{1}{2\Delta}, \tag{6.3.14}$$

and

$$\begin{aligned}\text{Cov}\,[C_{zz}(f_1), C_{zz}(f_2)] &= \frac{K_4\Delta^2}{N} + \sigma_z^4\Delta^2\left[\left(\frac{\sin \pi N\Delta(f_1+f_2)}{N\sin \pi\Delta(f_1+f_2)}\right)^2 + \left(\frac{\sin \pi N\Delta(f_1-f_2)}{N\sin \pi\Delta(f_1-f_2)}\right)^2\right], \\ &\qquad -\frac{1}{2\Delta} \leqslant f_1, f_2 < \frac{1}{2\Delta}, \end{aligned} \tag{6.3.15}$$

where K_4 is the fourth cumulant of the distribution of Z_t. It may be verified that (6.3.15) is zero when f_1 and f_2 are multiples of the fundamental frequency $1/N\Delta$ and the Z_t process is Normal, so that $K_4 = 0$. Hence, under these assumptions, the sample spectrum estimators at the harmonic frequencies are independent, as shown in Section 6.3.1.

For white noise in continuous time, the general results are

$$E[C_{zz}(f)] = \Gamma_{zz}(f) = \sigma_z^2, \quad -\infty \leqslant f \leqslant \infty, \tag{6.3.16}$$

and

$$\begin{aligned}\text{Cov}\,[C_{zz}(f_1), C_{zz}(f_2)] &= \frac{K_4}{T} + \sigma_z^4\left[\left(\frac{\sin \pi T(f_1+f_2)}{\pi T(f_1+f_2)}\right)^2 + \left(\frac{\sin \pi T(f_1-f_2)}{\pi T(f_1-f_2)}\right)^2\right], \\ &\qquad -\infty \leqslant f_1, f_2 \leqslant \infty, \end{aligned} \tag{6.3.17}$$

where K_4 is the fourth cumulant of the $Z(t)$ process.

Note that the covariance between spectral estimators is of order $1/T$ for non-Normal processes, that is, when $K_4 \neq 0$. Observe also that for Normal processes $K_4 = 0$ and the covariance between spectral estimators is of order $1/T^2$. In the special case where f_1 and f_2 are multiples of $1/T$, the covariance is zero. Further, the variance of spectral estimators, neglecting terms of order $1/T$ and higher, is

$$\text{Var}\,[C_{zz}(f)] = \sigma_z^4.$$

Hence, the variance is dominated by a constant term which remains finite as T tends to infinity. This shows in general that $C_{zz}(f)$ is not a consistent estimator of $\Gamma_{zz}(f)$.

Chi-squared properties of sample spectrum estimators for white noise. It was shown in Section 6.3.1 that for Z_t Normal white noise, $2C_{zz}(f)/\Delta\sigma_z^2$ was distributed as a χ_2^2 at the harmonic frequencies $f_k = k/N\Delta$. In Appendix A9.1 this result is generalized to show that for Normal white noise processes, $2C_{zz}(f)/\Delta\sigma_z^2$ is exactly distributed as a χ_2^2, while for non-Normal processes, $2C_{zz}(f)/\Delta\sigma_z^2$ is approximately distributed as a χ_2^2 *for all* f when N is large. The results for continuous time are identical except that they refer to $C_{zz}(f)/\sigma_z^2$.

6.3.4 Smoothing of spectral estimators

Bartlett's smoothing procedure. One device which can be used to produce spectrum estimators which have smaller variance than $C_{zz}(f)$ is due to Bartlett [5]. Suppose that instead of computing $C_{zz}(f)$ for a sample of white noise of length $N = 400$, as was done in Section 6.1.2, the series is split up into $k = 8$ series of length $N/k = 50$ and a sample spectrum $C_{zz}^{(i)}(f)$, $i = 1, 2, \ldots, 8$, evaluated for each sub-series. The mean of the eight sample spectra at the frequency f is

$$\bar{C}_{zz}(f) = \frac{1}{8}\sum_{i=1}^{8} C_{zz}^{(i)}(f), \quad -\frac{1}{2\Delta} \leqslant f < \frac{1}{2\Delta}, \tag{6.3.18}$$

and is called a *smoothed spectral estimate* at the frequency f.

Figure 6.10 shows $\bar{C}_{zz}(f)$ and $C_{zz}(f)$ based on all 400 terms, plotted at the frequencies $f = 0, 0.02, \ldots, 0.5$ cps. Note that $\bar{C}_{zz}(f)$ is smoother than $C_{zz}(f)$ and much closer to $\Gamma_{zz}(f)$. Table 6.4 shows the mean, variance and

TABLE 6.4: Moments of unsmoothed and smoothed spectral estimates

	Mean	Variance	Mean square error
$C_{zz}(f)$	0.95	0.826	0.828
$\bar{C}_{zz}(f)$	0.94	0.139	0.143

mean square error of $\bar{C}_{zz}(f)$ and $C_{zz}(f)$ averaged over frequency. From (6.3.10), the variance of each $C_{zz}^{(i)}(f)$ is σ_z^4. Because Z_t is white noise, it follows that the sub-series are independent and hence the variance of $\bar{C}_{zz}(f)$ is $\sigma_z^4/8$. The observed ratio (0.139/0.826 = 1/5.94) of the two variances from Table 6.4 is not inconsistent with the expected ratio 1/8. Hence by averaging or smoothing over sub-series, the variance of the spectral estimator can be reduced to whatever size is required. As an extreme case, sub-series of size 2 could be used, in which case the variance would be reduced to $2\sigma_z^4/N$. To see why this is not sensible, it is necessary to examine the above smoothing procedure more closely and to derive the moments of the smoothed estimator.

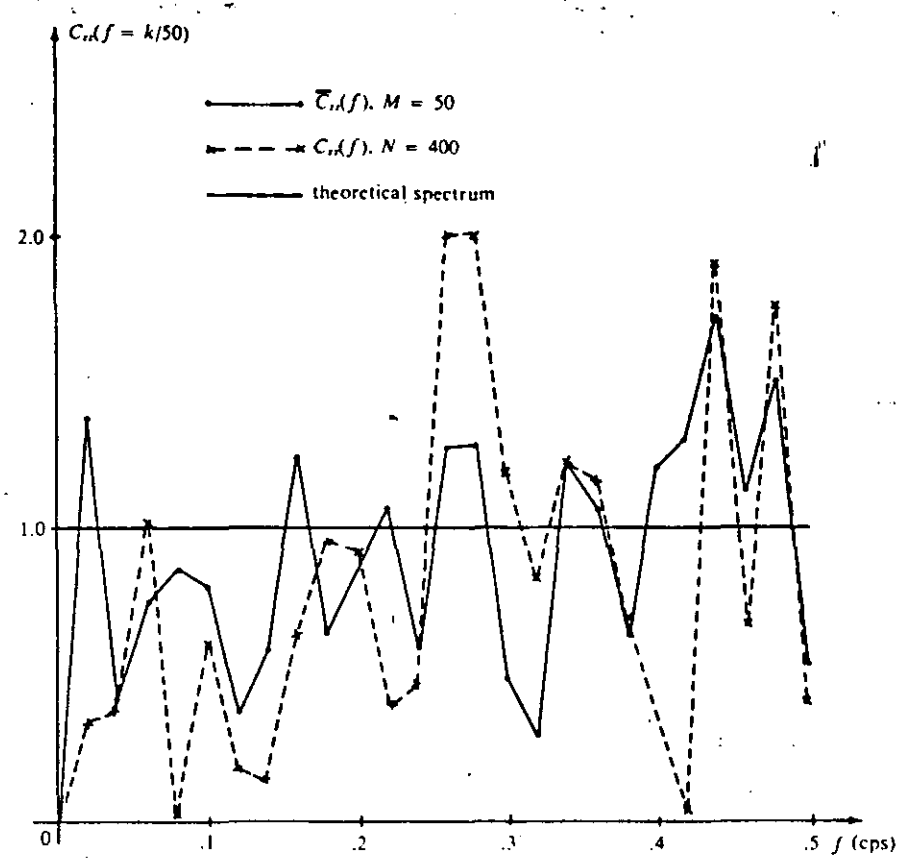


FIG. 6.10: Sample spectrum and smoothed spectral estimates for Normal white noise

Lag and spectral windows. From (6.2.1), the mean of the sample spectrum estimator is

$$E[C_{xx}(f)] = \int_{-T}^{T} \left(1 - \frac{|u|}{T}\right) \gamma_{xx}(u)\, e^{-j2\pi fu}\, du. \tag{6.3.19}$$

This is the Fourier transform of the product of $\gamma_{xx}(u)$ and the function

$$w(u) = \begin{cases} 1 - \dfrac{|u|}{T}, & |u| \leq T \\ 0, & |u| > T. \end{cases} \tag{6.3.20}$$

Hence using the convolution theorem (2.4.3),

$$E[C_{xx}(f)] = \int_{-\infty}^{\infty} T \left\{\frac{\sin \pi Tg}{\pi Tg}\right\}^2 \Gamma_{xx}(f - g)\, dg, \tag{6.3.21}$$

since the Fourier transform of $w(u)$ is

$$W(f) = T\left(\frac{\sin \pi Tf}{\pi Tf}\right)^2. \tag{6.3.22}$$

Equation (6.3.21) shows that the sample spectrum estimator has an expected value which corresponds to looking at the theoretical spectrum $\Gamma_{xx}(f)$ through the *spectral window* $W(f)$. In the terminology of Chapter 2, $E[C_{xx}(f)]$ corresponds to having passed the theoretical spectrum $\Gamma_{xx}(f)$ through a filter with an "impulse response" $W(f)$. The terms *spectral window* for $W(f)$ and *lag window* for $w(u)$ were introduced by Blackman and Tukey [6].

Since $W(f)$ in (6.3.22) behaves like a delta function for large T, it follows from (6.3.21) and (2.2.5) that

$$\lim_{T \to \infty} E[C_{xx}(f)] = \Gamma_{xx}(f),$$

that is, $C_{xx}(f)$ is an asymptotically unbiased estimator of $\Gamma_{xx}(f)$. For finite lengths of record, however, (6.3.21) shows that the sample spectrum is equivalent to looking at the true spectrum through the window (6.3.22). This means that $C_{xx}(f)$ is a biased estimator of $\Gamma_{xx}(f)$ with bias

$$B(f) = E[C_{xx}(f)] - \Gamma_{xx}(f).$$

For white noise, $\Gamma_{xx}(f) = \Delta\sigma_z^2$, and (6.3.21) reduces to

$$E[C_{xx}(f)] = \Delta\sigma_z^2$$

for all T. Hence for white noise, the sample spectrum estimator is an unbiased estimator of the spectrum for all T.

The spectral window $W(f)$ is a slit with width of order $1/T$, so that for large T it is reasonable to assume that $\Gamma_{xx}(f)$ is approximately constant over the slit. Hence (6.3.21) reduces to

$$E[C_{xx}(f)] \approx \Gamma_{xx}(f) \int_{-\infty}^{\infty} T\left(\frac{\sin \pi Tg}{\pi Tg}\right)^2 dg = \Gamma_{xx}(f).$$

Thus, the bias of the sample or unsmoothed spectrum estimate will always be small provided T is reasonably large.

The Bartlett spectral window. Now consider the expected value of the estimator $\bar{C}_{xx}(f)$ used in Bartlett's smoothing procedure. From (6.1.9), for the k sub-series, each of length M,

$$C_{xx}^{(i)}(f) = \int_{-M}^{M} c_{xx}^{(i)}(u)\, e^{-j2\pi fu}\, du.$$

Hence, the smoothed estimator is

$$\bar{C}_{xx}(f) = \frac{1}{k} \sum_{i=1}^{k} C_{xx}^{(i)}(f) = \int_{-M}^{M} \bar{c}_{xx}(u)\, e^{-j2\pi fu}\, du, \tag{6.3.23}$$

where

$$\tilde{c}_{XX}(u) = \frac{1}{k}\sum_{i=1}^{k}\left\{\frac{1}{M}\int_{(i-1)M}^{iM-u} X(t)X(t+u)\,dt\right\}, \quad u \geqslant 0, \qquad (6.3.24)$$

and it is defined in a similar way to (5.3.9) when $u < 0$. The expected value of $\tilde{c}_{XX}(u)$ is then

$$E[\tilde{c}_{XX}(u)] = \gamma_{XX}(u)\left(1 - \frac{|u|}{M}\right)$$

and

$$E[\bar{C}_{XX}(f)] = \int_{-M}^{M}\left(1 - \frac{|u|}{M}\right)\gamma_{XX}(u)\,e^{-j2\pi fu}\,du$$
$$= \int_{-\infty}^{\infty}\Gamma_{XX}(f-g)\,M\left(\frac{\sin \pi gM}{\pi gM}\right)^2 dg. \qquad (6.3.25)$$

Hence, subdividing the record of length T into k section of length $M = T/k$ and forming the smoothed spectral estimator (6.3.23) is equivalent to smoothing the sample spectrum by the spectral window

$$W(f) = M\left(\frac{\sin \pi fM}{\pi fM}\right)^2. \qquad (6.3.26)$$

In the time domain, this is equivalent to multiplying the covariance function by the lag window

$$w(u) = \begin{cases} 1 - \dfrac{|u|}{M}, & |u| \leqslant M \\ 0, & |u| > M. \end{cases} \qquad (6.3.27)$$

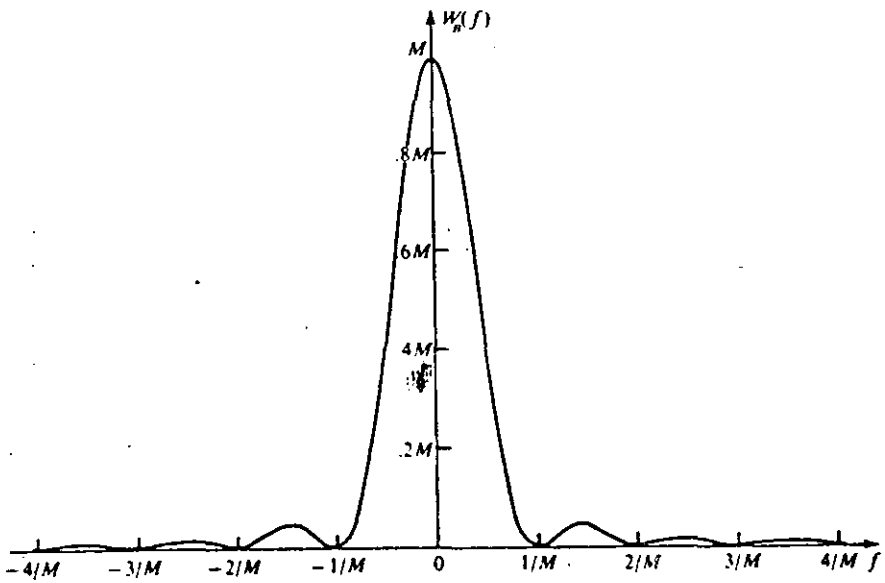


FIG. 6.11: The Bartlett spectral window $W_B(f) = M(\sin \pi fM/\pi fM)^2$

The windows (6.3.26) and (6.3.27) are called the Bartlett spectral window and Bartlett lag window. The Bartlett spectral window is plotted in Figure 6.11 and is seen to be symmetric about the origin and has zeros at $f = \pm 1/M$, $\pm 2/M, \ldots$. Thus the base width of the window, that is, the distance between the first zeros on either side, is $2/M$. Hence by controlling the length M of the sub-series it is possible to regulate the base width of the spectral window. It has already been shown that by making M small the variance can be made small. Since this corresponds to making the base width large, it follows that a small variance can be obtained by using a spectral window with a large base width. However, a large base width implies smoothing over a wide range of frequencies, that is, the "impulse response" $W(f)$ is very wide and so the bias $B(f) = E[\bar{C}_{XX}(f)] - \Gamma_{XX}(f)$ may be large. Thus, as with all statistical estimators, one is forced to compromise between variance and bias. In the next section this compromise is examined for a more general method of smoothing sample spectra.

6.3.5 *Spectral windows and smoothed spectral estimators*

A general class of smoothed spectral estimators. The Bartlett smoothing procedure described above shows that the large variance of the sample spectrum estimator can be reduced by introducing the lag window (6.3.27). More generally, this suggests that a smoothed spectral estimator of the form

$$\bar{C}_{XX}(f) = \int_{-\infty}^{\infty} w(u)c_{XX}(u)\,e^{-j2\pi fu}\,du = \int_{-\infty}^{\infty}\tilde{c}_{XX}(u)\,e^{-j2\pi fu}\,du \qquad (6.3.28)$$

will have a smaller variance than the sample spectrum estimator $C_{XX}(f)$. The lag window $w(u)$ in (6.3.28) is a function satisfying the conditions

$$\begin{aligned} &(1) \quad w(0) = 1, \\ &(2) \quad w(u) = w(-u), \\ &(3) \quad w(u) = 0, \quad |u| > T. \end{aligned} \qquad (6.3.29)$$

In practice, condition *(3)* is replaced by

$$(4) \quad w(u) = 0, \quad |u| \geqslant M, \quad M < T,$$

since covariances need then be computed only up to lag M. Examples of lag windows which are widely used in spectral analysis are given in Table 6.5 and plotted in Figure 6.12. The Fourier transforms of these lag windows, that is, the spectral windows $W(f)$, are shown in Figure 6.13.

Using the convolution property (2.4.3), (6.3.28) may be written

$$\bar{C}_{XX}(f) = \int_{-\infty}^{\infty} W(g)C_{XX}(f-g)\,dg, \qquad (6.3.30)$$

TABLE 6.5: Lag and spectral windows

Description	Lag window	Spectral window
rectangular	$w_R(u) = \begin{cases} 1, & \lvert u\rvert \leqslant M \\ 0, & \lvert u\rvert > M \end{cases}$	$W_R(f) = 2M\left(\dfrac{\sin 2\pi fM}{2\pi fM}\right), \quad -\infty \leqslant f \leqslant \infty$
Bartlett	$w_B(u) = \begin{cases} 1 - \dfrac{\lvert u\rvert}{M}, & \lvert u\rvert \leqslant M \\ 0, & \lvert u\rvert > M \end{cases}$	$W_B(f) = M\left(\dfrac{\sin \pi fM}{\pi fM}\right)^2, \quad -\infty \leqslant f \leqslant \infty$
Tukey	$w_T(u) = \begin{cases} \frac{1}{2}\left(1 + \cos\dfrac{\pi u}{M}\right), & \lvert u\rvert \leqslant M \\ 0, & \lvert u\rvert > M \end{cases}$	$W_T(f) = M\left\{\dfrac{\sin 2\pi fM}{2\pi fM} + \dfrac{1}{2}\dfrac{\sin 2\pi M(f+\frac{1}{2}M)}{2\pi M(f+\frac{1}{2}M)} + \dfrac{1}{2}\dfrac{\sin 2\pi M(f-\frac{1}{2}M)}{2\pi M(f-\frac{1}{2}M)}\right\}$ $= M\left(\dfrac{\sin 2\pi fM}{2\pi fM}\right)\left(\dfrac{1}{1-(2fM)^2}\right), \quad -\infty \leqslant f \leqslant \infty$
Parzen	$w_P(u) = \begin{cases} 1 - 6\left(\dfrac{u}{M}\right)^2 + 6\left(\dfrac{\lvert u\rvert}{M}\right)^3, & \lvert u\rvert \leqslant \dfrac{M}{2} \\ 2\left(1 - \dfrac{\lvert u\rvert}{M}\right)^3, & \dfrac{M}{2} < \lvert u\rvert \leqslant M \\ 0, & \lvert u\rvert > M \end{cases}$	$W_P(f) = \frac{3}{4}M\left(\dfrac{\sin \pi fM/2}{\pi fM/2}\right)^4, \quad -\infty \leqslant f \leqslant \infty$

FIG. 6.12: Some common lag windows

where $C_{xx}(f)$ is the sample spectrum defined by (6.1.6) or (6.1.9) and

$$W(f) = \int_{-\infty}^{\infty} w(u)\, e^{-i2\pi fu}\, du. \tag{6.3.31}$$

The inverse transform

$$w(u) = \int_{-\infty}^{\infty} W(f)\, e^{i2\pi fu}\, df \tag{6.3.32}$$

enables the lag window $w(u)$ to be calculated from the spectral window $W(f)$. Analogous to the properties (6.3.29), the spectral window $W(f)$ satisfies the conditions

$$\begin{aligned} &(1) \quad \int_{-\infty}^{\infty} W(f)\, df = w(0) = 1, \\ &(2) \quad W(f) = W(-f), \\ &(3) \quad W(f) \text{ is a slit with base width of order } 2/M. \end{aligned} \tag{6.3.33}$$

The expected value of a smoothed spectral estimator. Taking expectations in (6.3.30) gives

$$E[\bar{C}_{xx}(f)] = \int_{-\infty}^{\infty} W(g) E[C_{xx}(f-g)]\, dg. \tag{6.3.34}$$

However, for large T, (6.3.21) shows that

$$E[C_{xx}(g)] \approx \Gamma_{xx}(g),$$

and so

$$E[\bar{C}_{xx}(f)] \approx \int_{-\infty}^{\infty} W(g)\Gamma_{xx}(f-g)\, dg = \bar{\Gamma}_{xx}(f). \tag{6.3.35}$$

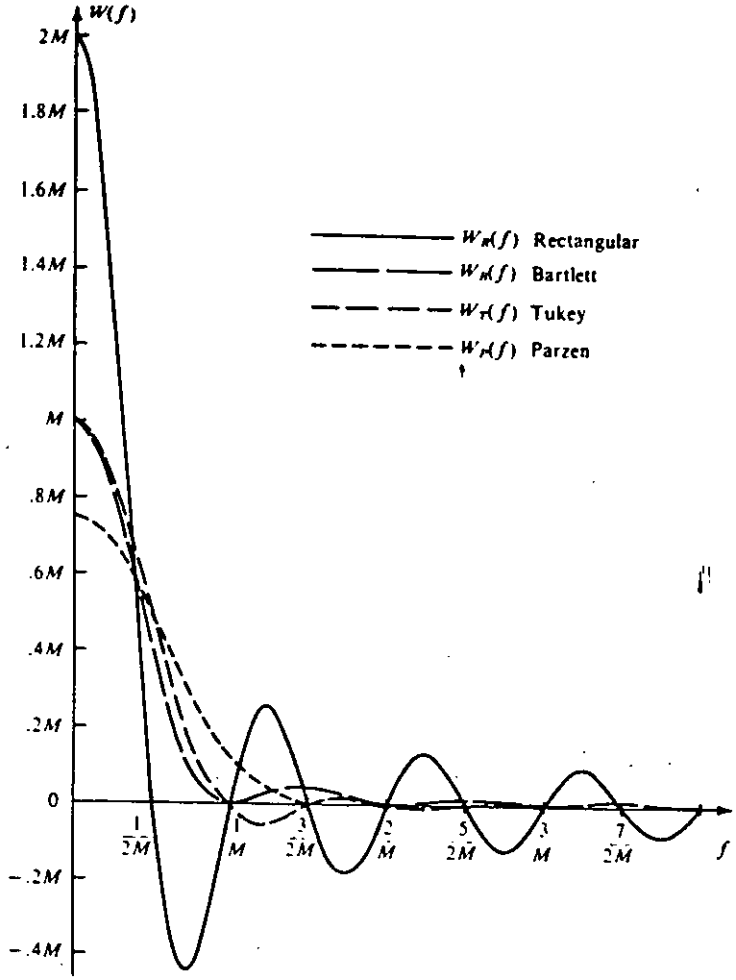


FIG. 6.13: Some common spectral windows

$\bar{\Gamma}_{xx}(f)$ will be called the *mean smoothed spectrum.*

The material of Section 2.4.1 is now relevant. Since the spectral window $W(f)$ satisfies condition *(3)* of (6.3.33), $\bar{\Gamma}_{xx}(f)$ will be a distorted version of $\Gamma_{xx}(f)$. This effect is demonstrated in Figure 2.10, where $\Gamma_{xx}(f)$ corresponds to $|S_i(f)|$, $\bar{\Gamma}_{xx}(f)$ corresponds to $|S_0(f)|$ and the lag windows $w(u)$ correspond to the data windows $w(t)$. From Figure 2.10 it is seen that the narrower the base width of the lag window, the more $\bar{\Gamma}_{xx}(f)$ differs from $\Gamma_{xx}(f)$. Hence, in order to keep the bias

$$B(f) = E[\bar{C}_{xx}] - \Gamma_{xx}(f) = \bar{\Gamma}_{xx}(f) - \Gamma_{xx}(f)$$

small, M must be large. This is contrary to the previous requirement that M

be small in order to keep Var $[\bar{C}_{xx}(f)]$ small. It has been shown in Section 4.2.3 that it is necessary to compromise between the variance and bias of an estimator. Similar considerations apply to estimators of the spectrum. The bias can only be made small by making $W(f)$ narrow or as close to a delta function as possible. On the other hand, a narrow spectral window $W(f)$ results in a large variance. Hence a sensible procedure is to compromise by making the mse

$$\text{Var}\,[\bar{C}_{xx}(f)] + B^2(f)$$

as small as possible [7].

The exact nature of the compromise which has to be made will depend on the degree of smoothness of the spectrum $\Gamma_{xx}(f)$. For example, if $\Gamma_{xx}(f)$ is very smooth, then the variance may be reduced by using a wide window without having serious effect on the bias. In particular, if $\Gamma_{xx}(f)$ is smooth over the range $-1/M \leqslant (f-g) \leqslant 1/M$, then (6.3.35) is approximately

$$\begin{aligned} E[\bar{C}_{xx}(f)] &\approx \Gamma_{xx}(f) \int_{-\infty}^{\infty} W(g)\, dg \\ &= \Gamma_{xx}(f), \end{aligned} \tag{6.3.36}$$

by virtue of (6.3.33) and (6.3.34). Hence, if the spectrum is sufficiently smooth, a virtually unbiased estimator is obtained even though the spectral window has been made wide to reduce the variance.

Approximate expressions for bias. If the spectrum is not smooth relative to the spectral window, it is possible following Parzen [8] to compute approximations to the bias corresponding to a particular spectral window. Using (6.3.28) and (5.3.13), the bias may also be written

$$\begin{aligned} B(f) &= E\left[\int_{-\infty}^{\infty} w(u) c_{xx}(u)\, e^{-j2\pi fu}\, du\right] - \int_{-\infty}^{\infty} \gamma_{xx}(u)\, e^{-j2\pi fu}\, du \\ &\approx \int_{-\infty}^{\infty} (w(u) - 1)\gamma_{xx}(u)\, e^{-j2\pi fu}\, du, \end{aligned} \tag{6.3.37}$$

for T large. Substituting for the lag windows $w(u)$ from Table 6.5, the following expressions are obtained for the bias associated with these windows:

$$\begin{aligned} B_B(f) &\approx \frac{1}{M} \int_{-\infty}^{\infty} -|u|\gamma_{xx}(u)\, e^{-j2\pi fu}\, du, \\ B_T(f) &\approx \frac{\pi^2}{4M^2} \int_{-\infty}^{\infty} -u^2 \gamma_{xx}(u)\, e^{-j2\pi fu}\, du + 0\left(\frac{1}{M^4}\right) \\ &= \frac{0.063}{M^2}\, \Gamma_{xx}^{(2)}(f) + 0\left(\frac{1}{M^4}\right). \\ B_P(f) &\approx \frac{6}{M^2} \int_{-\infty}^{\infty} -u^2 \gamma_{xx}(u)\, e^{-j2\pi fu}\, du + 0\left(\frac{1}{M^3}\right) \\ &= \frac{0.152}{M^2}\, \Gamma_{xx}^{(2)}(f) + 0\left(\frac{1}{M^3}\right). \end{aligned} \tag{6.3.38}$$

In the above expressions, $\Gamma^{(2)}_{XX}(f)$ is the second derivative of the spectrum at frequency f. These formulae show that:

(1) When $\Gamma^{(2)}_{XX}(f)$ is negative, as in the neighborhood of a peak, the bias is negative and hence peaks will tend to be underestimated. Conversely, when $\Gamma^{(2)}_{XX}(f)$ is positive, as in the neighborhood of a trough, the bias is positive, and hence troughs will be overestimated.

(2) The narrower the peak or trough, the larger $\Gamma^{(2)}_{XX}(f)$ and hence the greater the bias.

(3) The bias $B_h(f)$ for the Bartlett window is of order $1/M$ and hence will be larger in general than the bias of the Tukey and Parzen windows, which are of order $1/M^2$.

(4) The bias is reduced as M increases, that is, as the base width of the window is decreased.

(5) For the same truncation point M, the Parzen window has a larger bias than the Tukey window. This is because the Parzen spectral window is wider than the Tukey spectral window, as is seen from Figure 6.13. However, the Parzen estimator has a smaller variance than the Tukey estimator for the same values of M, as will be shown in Section 6.4.1.

The results (6.3.38) are useful in describing the qualitative features of bias but to obtain a quantitative picture it is necessary to plot the mean smoothed spectrum as a function of frequency, as will be shown in Section 7.1.

6.4 FURTHER PROPERTIES OF SMOOTHED SPECTRAL ESTIMATORS

One important property of a spectral estimator has already been derived, namely its bias. Another important property is its variance, for which an approximation was given in Section 6.3.4 for the special case of white noise using a Bartlett window. This result is now extended to any window and any stochastic process. Knowing the variance it is possible to construct a confidence interval for the true spectrum at any frequency. In this section it is shown that the covariance between two estimators at a sufficiently wide frequency spacing is approximately zero. Hence independent confidence intervals may be constructed at this frequency separation.

6.4.1 Covariance between smoothed spectral estimators

The precise derivation of the result for the covariance between smoothed spectral estimators at two frequencies is rather complicated. Hence the derivation given in this section is heuristic, but a more detailed analysis is given in Appendix A9.1.

The approach here will be to use the result (5.2.6), namely, that any stochastic process $X(t)$ with spectrum $\Gamma_{XX}(f)$ can be regarded as having been generated by passing white noise $Z(t)$ through a linear filter. Combining (5.2.6) with the results of Section 6.3.3 for the covariances between the sample spectrum estimators for white noise enables one to derive expressions for the covariances between sample spectrum estimators for any stochastic process. It is then a simple step to derive expressions for the covariance between smoothed spectral estimators.

Covariance of sample spectral estimators. Consider a stochastic process $X(t)$ with spectrum $\Gamma_{XX}(f)$, generated according to

$$X(t) = \int_0^\infty h(u)Z(t-u)\,du, \quad -\infty \leqslant t \leqslant \infty, \tag{6.4.1}$$

where $Z(t)$ is a white noise process. Hence, from (6.2.16), the spectrum can be written

$$\Gamma_{XX}(f) = \sigma_Z^2 |H(f)|^2, \qquad -\infty \leqslant f \leqslant \infty. \tag{6.4.2}$$

For a finite segment of the $X(t)$ process, (6.4.1) may be written

$$X(t) = \int_0^\infty h(u)Z(t-u)\,du \approx X_T(t), \quad -T/2 \leqslant t \leqslant T/2, \tag{6.4.3}$$

where

$$X_T(t) = \int_0^\infty h(u)Z_T(t-u)\,du, \quad -T/2 \leqslant t \leqslant T/2. \tag{6.4.4}$$

In (6.4.4), $Z_T(t)$ is a finite segment of the $Z(t)$ process. Over the interval $-T/2 \leqslant t \leqslant T/2$, the two processes $X(t)$ and $X_T(t)$ will be identical except near the beginning of the interval, provided the impulse response $h(u)$ tends to zero in a short time compared with T. It will now be assumed that the "beginning effect" can be neglected.

The sample spectrum for the $X(t)$ process,

$$C_{XX}(f) = \frac{1}{T}\left|\int_{-T/2}^{T/2} X(t)\,e^{-j2\pi ft}\,dt\right|^2,$$

may then be written approximately

$$\begin{aligned} C_{XX}(f) &\approx \frac{1}{T}\left|\int_{-T/2}^{T/2} X_T(t)\,e^{-j2\pi ft}\,dt\right|^2 \\ &= \frac{1}{T}\left|\int_{-T/2}^{T/2}\int_0^\infty h(u)Z_T(t-u)\,du\,e^{-j2\pi ft}\,dt\right|^2 \\ &= |H(f)|^2 C_{ZZ}(f). \end{aligned} \tag{6.4.5}$$

Thus, the sample spectrum for the $X(t)$ process is approximately equal to the sample spectrum for the white noise process multiplied by the squared

of the modulus of the filter frequency response function. Since $2C_{zz}(f)/\sigma_z^2$ is approximately distributed as a χ_2^2 for all f, (6.4.5) implies that

$$\frac{2C_{xx}(f)}{\sigma_z^2|H(f)|^2} = \frac{2C_{xx}(f)}{\Gamma_{xx}(f)} \tag{6.4.6}$$

is approximately distributed as a χ_2^2. The results of Section 6.3.3 for the properties of spectral estimators of white noise may then be used to yield the following results. Since $E[C_{zz}(f)] = \sigma_z^2$, from (6.3.16),

$$E[C_{xx}(f)] \approx |H(f)|^2\sigma_z^2 = \Gamma_{xx}(f). \tag{6.4.7}$$

Similarly, since

$$\begin{aligned}\operatorname{Cov}[C_{xx}(f_1), C_{xx}(f_2)] &= \operatorname{Cov}[|H(f_1)|^2C_{zz}(f_1), |H(f_2)|^2C_{zz}(f_2)] \\ &= |H(f_1)|^2|H(f_2)|^2 \operatorname{Cov}[C_{zz}(f_1), C_{zz}(f_2)],\end{aligned}$$

it follows from (6.3.17) that

$$\begin{aligned}&\operatorname{Cov}[C_{xx}(f_1), C_{xx}(f_2)] \\ &\quad\approx |H(f_1)|^2\,|H(f_2)|^2\,\sigma_z^4\left\{\left(\frac{\sin \pi T(f_1+f_2)}{\pi T(f_1+f_2)}\right)^2 + \left(\frac{\sin \pi T(f_1-f_2)}{\pi T(f_1-f_2)}\right)^2\right\},\end{aligned} \tag{6.4.8}$$

on neglecting the K_4 term. Since $\Gamma_{xx}(f) = \sigma_z^2|H(f)|^2$, the covariance between the two sample spectrum estimators of the linear process at two distinct frequencies f_1 and f_2 is

$$\begin{aligned}&\operatorname{Cov}[C_{xx}(f_1), C_{xx}(f_2)] \\ &\quad\approx \Gamma_{xx}(f_1)\Gamma_{xx}(f_2)\left\{\left(\frac{\sin \pi T(f_1+f_2)}{\pi T(f_1+f_2)}\right)^2 + \left(\frac{\sin \pi T(f_1-f_2)}{\pi T(f_1-f_2)}\right)^2\right\}.\end{aligned} \tag{6.4.9}$$

The result (6.4.9) shows that *for any Normal stochastic process* $X(t)$,

$$\operatorname{Cov}[C_{xx}(f_1), C_{xx}(f_2)] \approx 0\left(\frac{1}{T^2}\right), \qquad f_1 \neq f_2,$$

$$\operatorname{Var}[C_{xx}(f)] \approx \Gamma_{xx}^2(f),$$

and hence is an extension of the results of Section 6.3.3, which only applied to white noise processes. Note that for T large, the quantity in braces in (6.4.9) behaves like a delta function of area $1/T$. Furthermore, the covariance is identically zero when the frequencies $(f_1 \pm f_2)$ are multiples of $1/T$.

Covariance of smoothed spectral estimators. From (6.3.30), the smoothed spectral estimator $\bar{C}_{xx}(f)$ for the $X(t)$ process is

$$\bar{C}_{xx}(f) = \int_{-\infty}^{\infty} C_{xx}(g)W(f-g)\,dg.$$

and hence the covariance between $\bar{C}_{xx}(f_1)$ and $\bar{C}_{xx}(f_2)$ is

$$\begin{aligned}&\operatorname{Cov}[\bar{C}_{xx}(f_1), \bar{C}_{xx}(f_2)] \\ &\quad= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} W(f_1-g)W(f_2-h)\operatorname{Cov}[C_{xx}(g), C_{xx}(h)]\,dg\,dh.\end{aligned}$$

Substituting for $\operatorname{Cov}[C_{xx}(g), C_{xx}(h)]$ from (6.4.9) and integrating over h gives

$$\begin{aligned}&\operatorname{Cov}[\bar{C}_{xx}(f_1), \bar{C}_{xx}(f_2)] \\ &\quad\approx \frac{1}{T}\int_{-\infty}^{\infty} \Gamma_{xx}^2(g)W(f_1-g)[W(f_2+g) + W(f_2-g)]\,dg,\end{aligned} \tag{6.4.10}$$

provided T is large so that the $(\sin \pi fT/\pi fT)^2$ terms behave as delta functions. Equation (6.4.10) is the final result, but a useful approximation may be derived by making the assumption that $\Gamma_{xx}(f)$ is a smooth function over the width of the spectral window $W(f)$. Under this assumption, (6.4.10) becomes

$$\begin{aligned}&\operatorname{Cov}[\bar{C}_{xx}(f_1), \bar{C}_{xx}(f_2)] \\ &\quad\approx \frac{\Gamma_{xx}^2(f)}{T}\int_{-\infty}^{\infty} W(f_1-g)[W(f_2+g) + W(f_2-g)]\,dg,\end{aligned} \tag{6.4.11}$$

where $f_1 \leqslant f \leqslant f_2$.

Equation (6.4.11) shows that the covariance between smoothed spectral estimators is proportional to the amount of overlap of the spectral windows centered at f_1 and f_2. Hence, if the spectral windows overlap only slightly, the covariance will be very small. Some numerical values for the covariances of spectral windows will be given in Section 7.2.

Variance of smoothed spectral estimators. When $f_1 = f_2 = f$, (6.4.10) reduces to

$$\operatorname{Var}[\bar{C}_{xx}(f)] \approx \frac{\Gamma_{xx}^2(f)}{T}\int_{-\infty}^{\infty} W^2(g)\,dg, \tag{6.4.12}$$

neglecting the term in $\int_{-\infty}^{\infty} W(g)W(g+2f)\,dg$ which is small compared to $\int_{-\infty}^{\infty} W^2(g)\,dg$. Using Parseval's theorem, (6.4.12) may be written in the equivalent form

$$\operatorname{Var}[\bar{C}_{xx}(f)] \approx \frac{\Gamma_{xx}^2(f)}{T}\int_{-\infty}^{\infty} w^2(u)\,du = \Gamma_{xx}^2(f)\,\frac{I}{T}. \tag{6.4.13}$$

For example, for the Bartlett window $w_B(u)$ of Table 6.5,

$$I = \int_{-M}^{M}\left(1 - \frac{|u|}{M}\right)^2 du = \tfrac{2}{3}M,$$

and hence

$$\operatorname{Var}[\bar{C}_{xx}(f)] \approx \frac{\Gamma_{xx}^2(f)}{T}\left(\tfrac{2}{3}M\right).$$

This shows that the variance of the smoothed spectral estimator can be reduced by making the truncation point M of the lag window small. As discussed in Section 6.3.5, however, reducing M increases the bias and causes more distortion of the theoretical spectrum, since then the spectral window is wider. In this event, (6.4.10) shows that adjacent spectral estimators will be more highly correlated since the spectral windows will overlap more. Hence, the exact choice of M is a vital matter, and is discussed in Chapter 7.

Note that since Var $[C_{xx}(f)] \approx \Gamma_{xx}^2(f)$, the ratio

$$\frac{\text{Var}\,[\bar{C}_{xx}(f)]}{\Gamma_{xx}^2(f)} = \frac{1}{T}\int_{-\infty}^{\infty} w^2(u)\,du = I/T \qquad (6.4.14)$$

represents the proportional reduction in variance as a result of using a smoothed estimator as compared to the sample spectrum estimator. The values of the ratio (6.4.14) corresponding to the spectral windows of Table 6.5 are given in column 3 of Table 6.6. It is seen that they are all of the form $c(M/T)$, where c is a constant depending on the window.

TABLE 6.6: Properties of spectral windows

Description	Spectral window	Variance ratio I/T	Degrees of freedom	Standardized bandwidth b_1
rectangular	$2M\,\frac{\sin 2\pi fM}{2\pi fM}$	$2\,\frac{M}{T}$	$\frac{T}{M}$	0.5
Bartlett	$M\left(\frac{\sin \pi fM}{\pi fM}\right)^2$	$0.667\,\frac{M}{T}$	$3\,\frac{T}{M}$	1.5
Tukey	$M\left(\frac{\sin 2\pi fM}{2\pi fM} \times \frac{1}{1-(2fM)^2}\right)$	$0.75\,\frac{M}{T}$	$2.667\,\frac{T}{M}$	1.333
Parzen	$\frac{3}{4}M\left(\frac{\sin(\pi fM/2)}{\pi fM/2}\right)^4$	$0.539\,\frac{M}{T}$	$3.71\,\frac{T}{M}$	1.86

Suppose, for example, that the truncation point M is $0.1T$. Then I/T is $\frac{2}{3}(0.1) = 0.067$ for the Bartlett window. Hence by taking a truncation point which is 10% of the record length, the variance of the smoothed spectral estimator is reduced to 6.7% of the variance of the sample spectrum. The corresponding values for the Tukey and Parzen windows are 7.5% and 5.4% respectively. Hence for a given number of lags M, the Parzen window achieves the smallest variance among these three windows. Inspection of Figure 6.13 shows that this is due to the fact that the Parzen window is wider and flatter than the other two windows. As a result it tends to produce a larger bias and hence comparisons made on the basis of variance alone are misleading, as will be seen later.

6.4.2 *The chi-squared approximation to the distribution of smoothed spectral estimators*

In Section 6.3.5 it was shown that the sample spectrum estimator $C_{xx}(f)$ is such that $2C_{xx}(f)/\Gamma_{xx}(f)$ is approximately distributed as χ_2^2. In this section it is shown that the corresponding result for the smoothed spectral estimator

is that $\nu\bar{C}_{xx}(f)/\Gamma_{xx}(f)$ is approximately distributed as χ_ν^2 where $\nu > 2$. This means that the smoothed spectral estimators will have many more degrees of freedom than the sample spectrum estimator, a consequence of their smaller variance.

The sample spectrum $C_{xx}(f)$ is the Fourier transform of the covariance function estimator $c_{xx}(u)$, where $c_{xx}(u)$ is zero outside the interval $-T \leqslant u \leqslant T$. If $c_{xx}(u)$ is represented over the interval $-T \leqslant u \leqslant T$ by a periodic function $c_{xx}^p(u)$, where $c_{xx}^p(u) = c_{xx}^p(u + 2T)$, then the function $c_{xx}^p(u)$ has a Fourier series representation

$$c_{xx}^p(u) = \sum_{l=-\infty}^{\infty} C_{xx}^p\left(\frac{l}{2T}\right) e^{j(2\pi lu/2T)}.$$

Since the lag window $w(u) = 0$ for $|u| \geqslant M$, the functions $\tilde{c}_{xx}(u) = c_{xx}(u)w(u)$ and $\tilde{c}_{xx}(u) = c_{xx}^p(u)w(u)$ are identical over all u, and so the smoothed spectrum $\bar{C}_{xx}(f)$ has the equivalent representations

$$\bar{C}_{xx}(f) = \int_{-\infty}^{\infty} W(f-g)C_{xx}(g)\,dg$$

and

$$\bar{C}_{xx}(f) = \sum_{l=-\infty}^{\infty} W\left(f - \frac{l}{2T}\right) C_{xx}^p\left(\frac{l}{2T}\right).$$

But

$$\frac{C_{xx}(l/2T)}{2T} = C_{xx}^p\left(\frac{l}{2T}\right),$$

and hence

$$\bar{C}_{xx}(f) = \frac{1}{2T}\sum_{l=-\infty}^{\infty} C_{xx}\left(\frac{l}{2T}\right) W\left(f - \frac{l}{2T}\right).$$

Thus the smoothed spectral estimator is a weighted sum of rv's $C_{xx}(l/2T)$ at the subharmonic frequencies $l/2T$. These rv's are distributed as a χ_2^2 and hence, using the results of Section 3.3.5, the distribution of $\bar{C}_{xx}(f)$ may be approximated by $a\chi_\nu^2$. Using (3.3.14, 15), the constants a and ν may be evaluated from

$$\nu \approx \frac{2\{E[\bar{C}_{xx}(f)]\}^2}{\text{Var}\,[\bar{C}_{xx}(f)]}, \qquad (6.4.15)$$

$$a \approx \frac{E[\bar{C}_{xx}(f)]}{\nu}. \qquad (6.4.16)$$

Under the assumption that the spectrum is smooth with respect to the spectral window,

$$E[\bar{C}_{xx}(f)] \approx \Gamma_{xx}(f),$$

from (6.3.36), and

$$\text{Var}\,[\bar{C}_{xx}(f)] \approx \frac{\Gamma_{xx}^2(f)}{T}\int_{-\infty}^{\infty} w^2(u)\,du,$$

from (6.4.13). Hence on substituting these results in (6.4.15) and (6.4.16),

$$\nu = \frac{2T}{\int_{-\infty}^{\infty} w^2(u)\,du} = \frac{2T}{I} \tag{6.4.17}$$

$$a = \frac{\Gamma_{xx}(f)}{\nu}. \tag{6.4.18}$$

Hence the rv $\nu\bar{C}_{xx}(f)/\Gamma_{xx}(f)$ is distributed as a χ^2_ν with degrees of freedom ν given by (6.4.17). Thus the degrees of freedom of the smoothed spectral estimator depend on the window $w(u)$.

Column 4 of Table 6.6 shows the degrees of freedom associated with the spectral windows of column 2. For example, if the Bartlett window $w_B(u)$ is used with a truncation point M equal to one-tenth of the record length so that $M/T = 0.1$, the degrees of freedom of the estimator is $3/0.1 = 30$. The larger the number of degrees of freedom, the more reliable the estimator in the sense that its variance is smaller. However, there has to be a compromise between degrees of freedom and bias, as discussed above.

Referring to Table 6.6, it is seen that a wide window such as the Parzen window $W_P(f)$ provides a smaller variance and hence a larger number of degrees of freedom than a narrower window such as the Bartlett window $W_B(f)$. This is in accord with the previous remark that the wider the window, the smaller the variance.

6.4.3 Confidence limits for the spectrum

Since $\nu\bar{C}_{xx}(f)/\Gamma_{xx}(f)$ is distributed according to a χ^2_ν where ν is given by (6.4.17), it follows that

$$\Pr\left\{x_\nu\left(\frac{\alpha}{2}\right) < \frac{\nu\bar{C}_{xx}(f)}{\Gamma_{xx}(f)} \leqslant x_\nu\left(1-\frac{\alpha}{2}\right)\right\} = 1-\alpha, \tag{6.4.19}$$

where $\Pr\{\chi^2_\nu \leqslant x_\nu(\alpha/2)\} = \alpha/2$. Hence by a similar argument to that used in Section 3.3.2, the interval between

$$\frac{\nu\bar{C}_{xx}(f)}{x_\nu(1-(\alpha/2))}, \qquad \frac{\nu\bar{C}_{xx}(f)}{x_\nu(\alpha/2)} \tag{6.4.20}$$

is a $100(1-\alpha)\%$ confidence interval for $\Gamma_{xx}(f)$. For a specified value of the ratio T/M the value of ν corresponding to a given spectral window may be obtained from column 4 of Table 6.6, and then the confidence interval obtained from (6.4.20) and Figure 3.10, which gives the multiplying factors $\nu/x_\nu(\alpha/2)$ and $\nu/x_\nu(1-(\alpha/2))$. For example, the smoothed spectral estimate shown in Figure 6.10 was based on the Bartlett window with $M/T = 0.125$. Hence from Table 6.5, $\nu = 3/0.125 = 24$. At the frequency $f = 0.1$ cps, $\bar{C}_{zz}(f) = 0.804$, and so the 95% confidence limits for $\Gamma_{zz}(f)$ are, from Figure 3.10,

$$(0.61)(0.804) = 0.49, \qquad (1.94)(0.804) = 1.56.$$

Similarly, the 95% confidence limits for $\Gamma_{zz}(f)$ based on the sample spectrum $C_{zz}(f)$ at $f = 0.1$ cps are

$$(0.27)(0.622) = 0.169, \qquad (39.5)(0.622) = 24.6.$$

These are considerably wider because of the smaller degrees of freedom associated with the estimate.

Note that (6.4.19) gives a confidence interval for $\Gamma_{xx}(f)$ at a *particular* frequency f only. If a confidence interval is given for q frequencies, at which the estimators are independent, the confidence coefficient will be $(1-\alpha)^q$, which will usually be considerably less than $(1-\alpha)$. Note also that the variance will only summarize the properties of the estimators if the bias is small, as discussed in Section 6.3.5. Hence the above confidence intervals will be of value only when the spectral window is sufficiently narrow so that no appreciable bias exists.

Confidence intervals on a logarithmic scale. It will be shown in Section 7.1.2 that spectral estimates should be plotted on a logarithmic scale so that the variation in the spectrum can be accommodated. The logarithmic scale is also a sensible choice from an engineering point of view, since it is usually *proportional* changes in power which are important. From a statistical point of view there is also a good reason to plot spectra on a logarithmic scale, since then the confidence interval for the spectrum is simply represented by a constant interval about the spectral estimate.

Thus using (6.4.20), the confidence interval for $\log \Gamma_{xx}(f)$ is

$$\log \bar{C}_{xx}(f) + \log\frac{\nu}{x_\nu(1-(\alpha/2))}, \qquad \log \bar{C}_{xx}(f) + \log\frac{\nu}{x_\nu(\alpha/2)}. \tag{6.4.21}$$

Therefore, when plotting the estimated spectrum, the confidence interval for all frequencies can be indicated by a single vertical line.

For example, consider the smoothed spectral estimate $\bar{C}_{xx}(f)$ of Figure 6.10 for which $\nu = 24$. From Figure 3.10 and (6.4.21) the 95% confidence intervals for $\log_{10} \Gamma_{xx}(f)$ are

$$\log_{10} \bar{C}_{xx}(f) + \log_{10}(0.61), \qquad \log_{10} \bar{C}_{xx}(f) + \log_{10}(1.94).$$

With $\bar{C}_{xx}(f)$ plotted on logarithmic paper, the 95% confidence interval would be obtained very simply by plotting the points (0.61, 1.0, 1.94) from Figure 3.10 in a vertical line on the log scale. This procedure will be illustrated in Section 7.2 and elsewhere throughout the book.

6.4.4 Bandwidth of a spectral window

It has been shown in Section 6.4.1 that one useful characteristic of a spectral window is $I = \int_{-\infty}^{\infty} w^2(u)\,du$, since I/T provides a measure of the reduction in variance of the estimator due to smoothing by the spectral window. Hence, to

obtain a small variance it is necessary to choose $w(u)$ so that I is small. For a given window this can be achieved by making M small. A further useful characteristic of a window is its width. It will be shown in later sections that in order to obtain a good estimate of a peak in a spectrum, the "width" of a spectral window must be of the same order as the width of the peak. Since the spectral window is non-zero for most f in the range $-\infty \leqslant f \leqslant \infty$, it is necessary to define more precisely the notion of "width" of a spectral window.

One approach used by statisticians [9] to define the width, or *bandwidth*, of a spectral window is to consider a "bandpass" spectral window

$$W(f) = \frac{1}{h}, \quad -\frac{h}{2} \leqslant f \leqslant \frac{h}{2}.$$

This spectral window is rectangular in the frequency domain and has a unique width h, that is, its bandwidth $b = h$. Using (6.4.13), the variance of a smoothed spectral estimator based on this spectral window is

$$\text{Var}\,[\bar{C}_{xx}(f)] \approx \frac{\Gamma_{xx}^2(f)}{T} \cdot \frac{1}{h} = \frac{\Gamma_{xx}^2(f)}{Tb}.$$

For an estimator based on a spectral window which is not rectangular, it is natural to define the bandwidth of the window as the width of a rectangular window which gives the same variance. That is,

$$\text{Var}\,[\bar{C}_{xx}(f)] \approx \frac{\Gamma_{xx}^2(f)}{T} \cdot \frac{1}{b} = \frac{\Gamma_{xx}^2(f)}{T} \int_{-\infty}^{\infty} w^2(u)\, du. \quad (6.4.22)$$

Hence the bandwidth is

$$b = \frac{1}{I} = \frac{1}{\int_{-\infty}^{\infty} w^2(u)\, du} = \frac{1}{\int_{-\infty}^{\infty} W^2(f)\, df}. \quad (6.4.23)$$

For example, for the rectangular and Bartlett lag windows $w_R(u)$ and $w_B(u)$ of Table 6.5, the bandwidths are $1/2M$ and $3/2M$ respectively.

It is sometimes convenient to define the *standardized bandwidth* b_1 corresponding to $M = 1$, so that

$$b = \frac{b_1}{M} = \frac{1}{\int_{-\infty}^{\infty} w^2(u)\, du}. \quad (6.4.24)$$

For example, the standarized bandwidths for the rectangular and Bartlett lag windows are $\frac{1}{2}$ and $\frac{3}{2}$ respectively. The standardized bandwidths for the windows of Table 6.5 are shown in column 5 of Table 6.6. It is seen that the Parzen window w_P has a standardized bandwidth which is approximately 1.4 times that of the Tukey window w_T.

Engineers will recognize (6.4.23) as being the inverse of the definition of the noise bandwidth of a filter. The precise definition of bandwidth is not very important, for example, some authors [10] use the distance between the half

power points. Our preference for the definition (6.4.23) stems from the fact that it utilizes the full shape of the spectral window and is therefore more likely to distinguish between window shapes than the definition based on half power points. From (6.4.23) it is seen that the variance of the spectral estimator is inversely proportional to the bandwidth of the spectral window. In fact, from (6.4.22) and (6.4.23),

$$\text{Variance} \times \text{Bandwidth} = \text{Constant}. \quad (6.4.25)$$

Hence small variance is associated with large bandwidths and large variance with small bandwidths. Furthermore, (6.4.17) shows that the number of degrees of freedom ν of the smoothed estimator is

$$\nu = \frac{2T}{I} = 2\left(\frac{T}{M}\right) b_1. \quad (6.4.26)$$

Hence a large bandwidth implies that the number of degrees of freedom of the smoothed estimator is large and the variance small. Conversely a small bandwidth implies few degrees of freedom and hence large variance. Since it was shown in Section 6.3.5 that bias is reduced by making M large, it follows that small bias is associated with small bandwidth.

In the next chapter, the concepts introduced here are applied to the practical problem of estimating the spectrum of an observed time series.

REFERENCES

[1] J. L. Doob, *Stochastic Processes*. John Wiley, New York, 1953.

[2] J. H. Laning and R. H. Battin, *Random Processes in Automatic Control*. McGraw-Hill, New York, 1956.

[3] H. M. James, N. B. Nichols and R. S. Phillips, *Theory of Servomechanisms*. McGraw-Hill, New York, 1947.

[4] A. Hald, *Statistical Theory with Engineering Applications*. John Wiley, New York, 1952.

[5] M. S. Bartlett, *An Introduction to Stochastic Processes with Special Reference to Methods and Applications*. Cambridge University Press, Cambridge, 1953.

[6] R. B. Blackman and J. W. Tukey, *The Measurement of Power Spectra from the Point of View of Communications Engineering*. Dover, New York, 1958.

[7] U. Grenander and M. Rosenblatt, *Statistical Analysis of Stationary Time Series*. John Wiley, New York, 1957.

[8] E. Parzen, "Mathematical considerations in the estimation of spectra." *Technometrics* 3, 167 (1961).

[9] G. M. Jenkins, "General considerations in the analysis of spectra." *Technometrics* 3, 133 (1961).

[10] M. B. Priestley, "Basic considerations in the estimation of spectra." *Technometrics* 4, 551 (1962).

CURSO INGENIERIA MARITIMA.
MODULO:"FUNDAMENTOS DE HIDRAULICA Y MATEMATICAS".
DEL 11 DE MARZO AL 7 DE JUNIO.
MEXICO, D.F.

EXAMPLES OF UNIVARIANTE. SPECTRAL ANALYSIS.

PROF. ING. JOSE RAYNAL.
JUNIO 1985.

7

Examples of Univariate Spectral Analysis

In this chapter the theory of Chapter 6 is used to derive practical procedures for estimating spectra from observed time series. To provide the reader with experience in the numerical procedures involved, Section 7.1 illustrates the effect of varying bandwidth and window shape on the spectra of simulated time series. In Section 7.2 a practical method of estimating spectra, called *window closing*, is introduced. This requires the use of a wide bandwidth initially and then progressively smaller bandwidths until all the important detail in the spectrum has been brought out. This procedure can sometimes be badly affected by the instability of the estimate due to the shortness of the time series.

Section 7.3 discusses practical questions which arise in the estimation of spectra and gives a routine procedure which can be followed in practice. The importance of prefiltering the data to remove low frequency trends is emphasized. In Section 7.4 examples of spectral analysis are given in the three application areas of model building, design of experiments and frequency response studies.

7.1 SPECTRAL ANALYSIS USING SIMULATED TIME SERIES

In this section estimates of spectra are computed for simulated time series. This is intended to provide the reader with experience in interpreting spectral estimates. Section 7.1.1 presents the formulae for digital computation of smoothed spectral estimates, together with a sample calculation. Section 7.1.2 then illustrates the effect on spectra of changing the truncation point of the autocorrelation function. This is achieved by comparing $\bar{\Gamma}_{xx}(f)$ with $\Gamma_{xx}(f)$ and $\bar{C}_{xx}(f)$ with $\Gamma_{xx}(f)$ for first- and second-order autoregressive processes. To prepare for a discussion in Section 7.2 of practical methods of smoothing, an investigation is made of the effect of varying the truncation point (Section 7.1.2) and the shape (Section 7.1.3) of spectral windows.

7.1.1 *Discrete estimation formulae*

The statistical theory of spectral estimation has been derived in the previous sections assuming the data $x(t)$ are continuous. There are many situations in which the data are essentially discrete, for example, the batch data of Figure 5.2, and hence digital formulae are required. In addition, because of the accuracy, flexibility and relative availability of digital computers today, it may be assumed that most spectral analyses will be done using digital computers. This implies that the continuous or analog signal will have to be sampled in the manner of Chapter 2 and the sampled values converted to digital form. The process of converting from analog to digital form is termed *quantizing*, and a detailed analysis of its effects on correlation analysis has been given in [1]. It will be assumed that the quantizing is fine enough so that no errors are introduced in the conversion from analog to digital form. In practice this means that the data should be read to the nearest 1/10 to 1/100 of the full range of the signal.

Suppose that the digital data x_t, $t = 1, \ldots, N$, corresponds to values of the signal $x(t)$ at intervals Δ. In this case, the smoothed spectral estimate $\bar{C}_{xx}(f)$ is obtained by replacing the integral (6.3.28) by the corresponding sum

$$\bar{C}_{xx}(f) = \Delta \sum_{k=-(L-1)}^{L-1} w(k) c_{xx}(k)\, e^{-j2\pi f k\Delta}, \quad -\frac{1}{2\Delta} \leqslant f < \frac{1}{2\Delta}. \tag{7.1.1}$$

In (7.1.1) the acvf estimate $c_{xx}(k)$ is

$$c_{xx}(k) = \frac{1}{N} \sum_{t=1}^{N-k} (x_t - \bar{x})(x_{t+k} - \bar{x}), \quad -(N-1) \leqslant k \leqslant N-1, \tag{7.1.2}$$

and $L = M/\Delta$. As before, $w(k)$ is the lag window with truncation point M, but it is now only defined at the discrete time points $u = k\Delta$. Note that it may be necessary to filter the data according to (5.3.27) or (5.3.31) before estimating the acvf.

Since $\bar{C}_{xx}(f)$ is an even function of frequency, it is only necessary to calculate it over the range $0 \leqslant f \leqslant 1/2\Delta$. However, to preserve the Fourier transform relationship between the sample spectrum and the sample acvf, it is necessary to double the power associated with each frequency in the range $0 \leqslant f \leqslant 1/2\Delta$. Hence, the formula generally used is

$$\bar{C}_{xx}(f) = 2\Delta \left\{ c_{xx}(0) + 2 \sum_{k=1}^{L-1} c_{xx}(k) w(k) \cos 2\pi f k \Delta \right\}, \quad 0 \leqslant f \leqslant \frac{1}{2\Delta}. \tag{7.1.3}$$

and the acvf (7.1.2) need only be computed for $k \geqslant 0$.

For purposes of computation it is more convenient to assume that $\Delta = 1$, in which case all sets of data may be processed alike. Thus the computing formula becomes

$$\bar{C}_{xx}(f) = 2\left\{c_{xx}(0) + 2\sum_{k=1}^{L-1} c_{xx}(k)w(k)\cos 2\pi fk\right\}, \quad 0 \leqslant f \leqslant \tfrac{1}{2}. \qquad (7.1.4)$$

If $\Delta \neq 1$, then the estimate (7.1.1) can be recovered from (7.1.4) by multiplying by Δ and plotting the estimate against $f\Delta$ instead of f.

Finally, if autocorrelations are used instead of autocovariances, the smoothed spectral density estimate $\bar{R}_{xx}(f)$ is computed according to

$$\bar{R}_{xx}(f) = 2\left\{1 + 2\sum_{k=1}^{L-1} r_{xx}(k)w(k)\cos 2\pi fk\right\}, \quad 0 \leqslant f \leqslant \tfrac{1}{2},$$

where

$$r_{xx}(k) = \frac{c_{xx}(k)}{c_{xx}(0)}. \qquad (7.1.5)$$

In practice it has been suggested [2] that $\bar{R}_{xx}(f)$ should only be computed at values of f corresponding to $f = 0, 1/L, 2/L, \ldots, \frac{1}{2}$. This frequency spacing is too wide, and it is recommended that $\bar{R}_{xx}(f)$ be evaluated at some fraction of this spacing so that a more detailed plot of $\bar{R}_{xx}(f)$ is obtained. In fact, it is not necessary that the frequency spacing for the values of $\bar{R}_{xx}(f)$ be at all related to the truncation point L. Thus, the f values may be computed at spacing $1/(2F)$, where F is 2 to 3 times L. Hence the final formula for digital computation of the smoothed spectral density function is

$$\bar{R}_{xx}(l) = 2\left\{1 + 2\sum_{k=1}^{L-1} r_{xx}(k)w(k)\cos\frac{\pi lk}{F}\right\}, \quad l = 0, 1, \ldots, F. \qquad (7.1.6)$$

A flow chart for digital computer application is presented in Appendix A7.1.

A sample calculation. The sample autocorrelations for the discrete second-order ar process (5.3.36) are given in Table 5.2. These values may be used to derive an estimate of the spectral density function $2\Gamma_{xx}(f)/\sigma_x^2$ as follows. Using, for illustration, the Bartlett lag window

$$w(k) = \begin{cases}\left(1 - \dfrac{k}{L}\right), & 0 \leqslant k \leqslant L \\ 0, & k > L,\end{cases}$$

the smoothed spectral density function estimate is, from (7.1.6),

$$\bar{R}_{xx}(l) = 2\left\{1 + 2\sum_{k=1}^{L-1}\left(1 - \frac{k}{L}\right)r_{xx}(k)\cos\frac{\pi lk}{F}\right\}, \quad l = 0, 1, \ldots, F.$$

For example, if $L = 3$, the estimate is

$$\bar{R}_{xx}(l) = 2\left\{1 + 2\left(\frac{2}{3}\right)r_{xx}(1)\cos\frac{\pi l}{F} + 2\left(\frac{1}{3}\right)r_{xx}(2)\cos\frac{2\pi l}{F}\right\},$$
$$l = 0, 1, \ldots, F.$$

If the estimates are required at a frequency spacing of 1/16 cps, so that $F = 8$, then using the values of the sample acf from Table 5.2, the calculation may be set out as shown in Table 7.1.

These spectral density estimates are shown in Figure 7.1. It is seen that a well-defined smooth curve can be drawn through the points. Also shown on the graph are the estimates plotted at spacing $\frac{1}{3}$ as suggested in [2]. It is seen that this spacing is too wide for plotting and interpolating purposes.

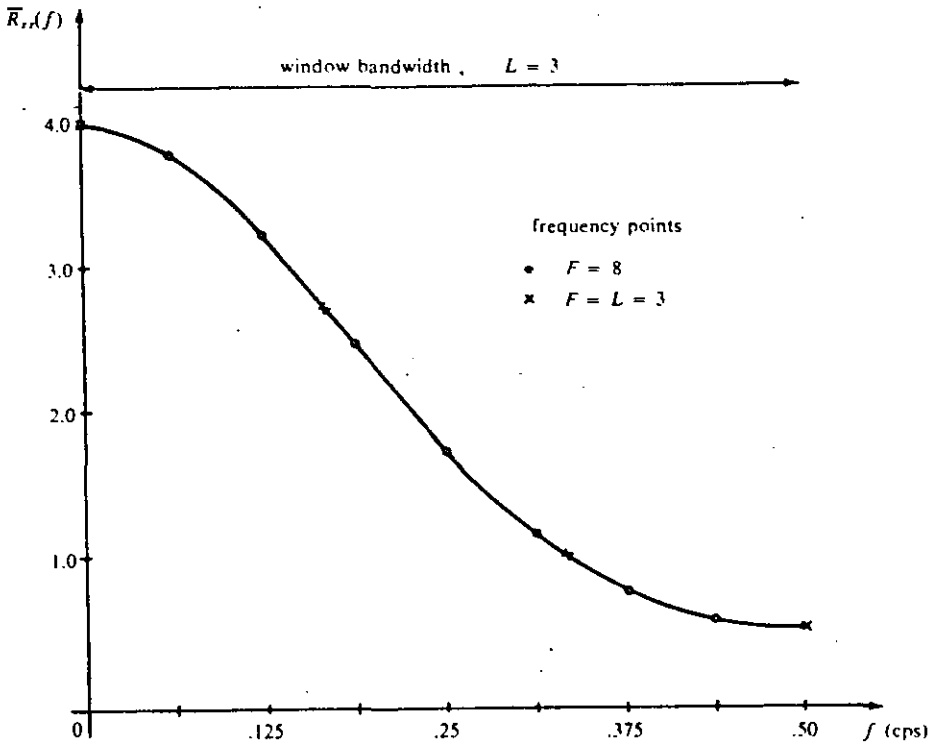


FIG. 7.1: A smoothed spectral density estimate for the second-order ar process ($\alpha_1 = 1.0, \alpha_2 = -0.5$) using the Bartlett window

Also shown on the graph is the bandwidth of the spectral window used. The bandwidth for the Bartlett window, with $L = 3$, is

$$b = b_1/L\Delta = 1.5/3(1) = 0.5 \text{ cps.}$$

This bandwidth is equal to the whole of the frequency range and hence the estimate has been smoothed considerably.

TABLE 7.1: An example of a spectral calculation

		①			②	
l/F	$f = \frac{l}{2F}$	$\cos\frac{\pi l}{F}$	$2\left(\frac{2}{3}\right)(0.645)\cos\frac{\pi l}{F}$	$\cos\frac{2\pi l}{F}$	$2\left(\frac{1}{3}\right)(0.195)\cos\frac{2\pi l}{F}$	$\bar{R}_{xx}(l) = 2(1 + ① + ②)$
0	0	1.0	0.860	1.0	0.130	3.980
$\frac{1}{8}$	$\frac{1}{16}$	0.924	0.794	0.707	0.092	3.773
$\frac{1}{4}$	$\frac{1}{8}$	0.707	0.608	0	0	3.216
$\frac{3}{8}$	$\frac{3}{16}$	0.383	0.329	−0.707	−0.092	2.474
$\frac{1}{2}$	$\frac{1}{4}$	0	0	−1.0	−0.130	1.740
$\frac{5}{8}$	$\frac{5}{16}$	−0.383	−0.329	−0.707	−0.092	1.158
$\frac{3}{4}$	$\frac{3}{8}$	−0.707	−0.608	0	0	0.784
$\frac{7}{8}$	$\frac{7}{16}$	−0.924	−0.794	0.707	0.092	0.596
1	$\frac{1}{2}$	−1.0	−0.860	1.0	0.130	0.540

7.1.2 *Effect of bandwidth on smoothing*

In this section an empirical investigation is made of the effect of varying the bandwidth of the window, or equivalently the truncation point L, on the smoothing of the spectral estimate. The time series used are realizations of first- and second-order ar processes with known spectra $\Gamma_{xx}(f)$. The quantities computed are the mean smoothed spectral density function

$$\frac{\bar{\Gamma}_{xx}(f)}{\sigma_x^2} = 2\left[1 + 2\sum_{k=1}^{L-1} \rho_{xx}(k)w(k)\cos 2\pi fk\right] \tag{7.1.7}$$

and the smoothed spectral density estimate $\bar{R}_{xx}(f)$ (7.1.6). The mean smoothed spectral density $\bar{\Gamma}_{xx}(f)/\sigma_x^2$ is the expected value of the smoothed spectral density estimator, and hence when plotted alongside $\Gamma_{xx}(f)/\sigma_x^2$ shows how the bias varies with frequency. By plotting a series of curves for different values of L, the variation of bias with bandwidth is readily displayed. Similarly, by plotting the smoothed estimate $\bar{R}_{xx}(f)$ alongside $\Gamma_{xx}(f)/\sigma_x^2$ for various values of L, the effect of bandwidth on the variance of the estimator can be displayed visually.

A first-order ar process ($\alpha_1 = -0.4$). The stochastic process used for these computations is the first-order ar process (5.2.26), with $\alpha_1 = -0.4$, $\Delta = 1$, $\mu = 0$, that is,

$$X_t = -0.4X_{t-1} + Z_t.$$

Using (6.2.20), this process has the spectral density function

$$\frac{\Gamma_{xx}(f)}{\sigma_x^2} = \frac{2(0.84)}{1.16 + 0.8\cos 2\pi f}, \quad 0 \leqslant f \leqslant \tfrac{1}{2},$$

shown in Figure 7.2. It is seen that the spectral density function is very smooth with a wide peak at $f = 0.5$ cps, and that it varies only over a small vertical range. Also shown in the figure are the mean smoothed spectral density functions $\bar{\Gamma}_{xx}(f)/\sigma_x^2$ for values of $L = 4$, 8 and 16 based on the Tukey window. These curves show that for $L = 4$, the mean smoothed spectral density is quite biased, particularly near the peak, which is considerably underestimated. Doubling L to 8 improves the correspondence between $\bar{\Gamma}_{xx}(f)$ and $\Gamma_{xx}(f)$, particularly for frequencies less than 0.375 cps, in which cases $\bar{\Gamma}_{xx}(f)$ and $\Gamma_{xx}(f)$ are essentially indistinguishable. Therefore, $L = 8$ would be sufficient to give an estimate with an acceptably small bias over most of the frequency range. However, considerable bias persists near the peak, and hence a larger value of L is necessary to estimate the peak accurately. Doubling L again, to $L = 16$, shows that there is still a little bias near the peak. The reduction in bias achieved by doubling L to 16 is considerably less than that achieved by doubling L to 8. Hence a point of diminishing returns may have been reached.

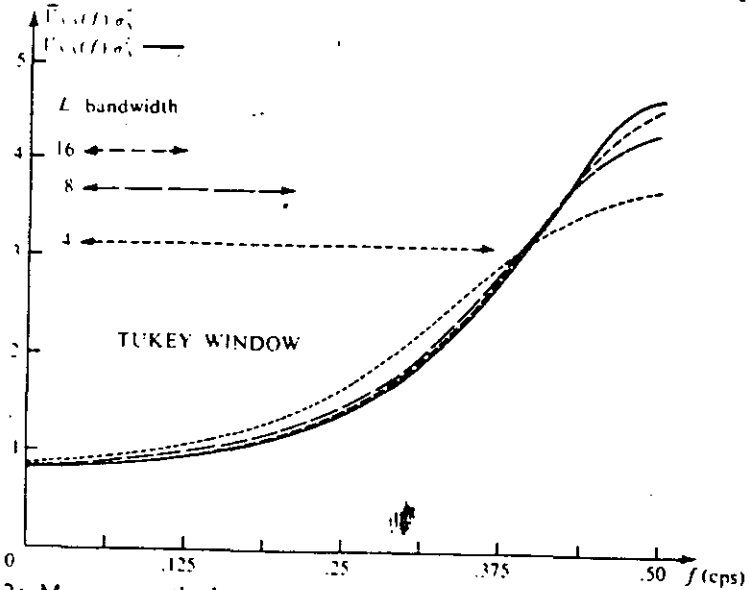


FIG. 7.2: Mean smoothed spectral density functions for a first-order ar process ($\alpha_1 = -0.4$)

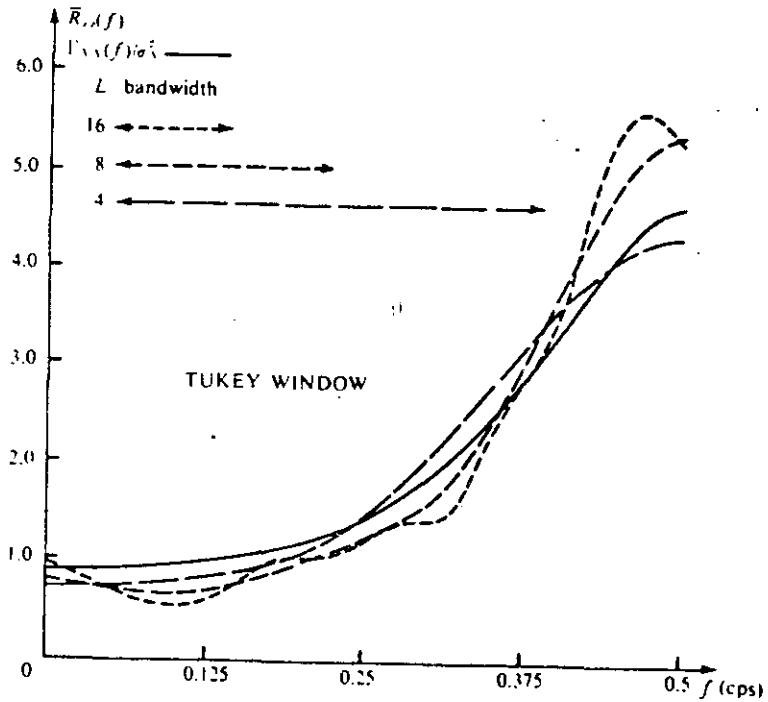


FIG. 7.3: Smoothed spectral density estimates for a first-order ar process ($\alpha_1 = -0.4$; $N = 400$)

This is simply explained by noting that for $k > 8$, $|\rho_{XX}(k)| < (0.4)^8 < 0.001$, and hence increasing L above 8 results in small changes in the sum of the series (7.1.7). Similar conclusions may be drawn from the mean smoothed spectra obtained using the Bartlett and Parzen windows. For the Bartlett window the number of lags which gives a good estimate is approximately $L = 12$ to 16 and for the Parzen window $L = 12$.

Smoothed spectral density estimates using the Tukey window for realizations of $N = 400$ and $N = 100$ terms of this process are shown in Figures 7.3 and 7.4. For $N = 400$, shown in Figure 7.3, changing L from 4 to 8 produces an appreciable change in the spectral estimate. A further increase from $L = 8$ to $L = 16$ produces only minor changes, and hence a reasonable estimate of the spectrum could be obtained with $L = 8$. The number of degrees of freedom per estimate in this case is 133 so that the confidence interval is very narrow.

The situation is different for the estimates, shown in Figure 7.4, obtained from the realization based on $N = 100$. Here, changing L from 4 to 8

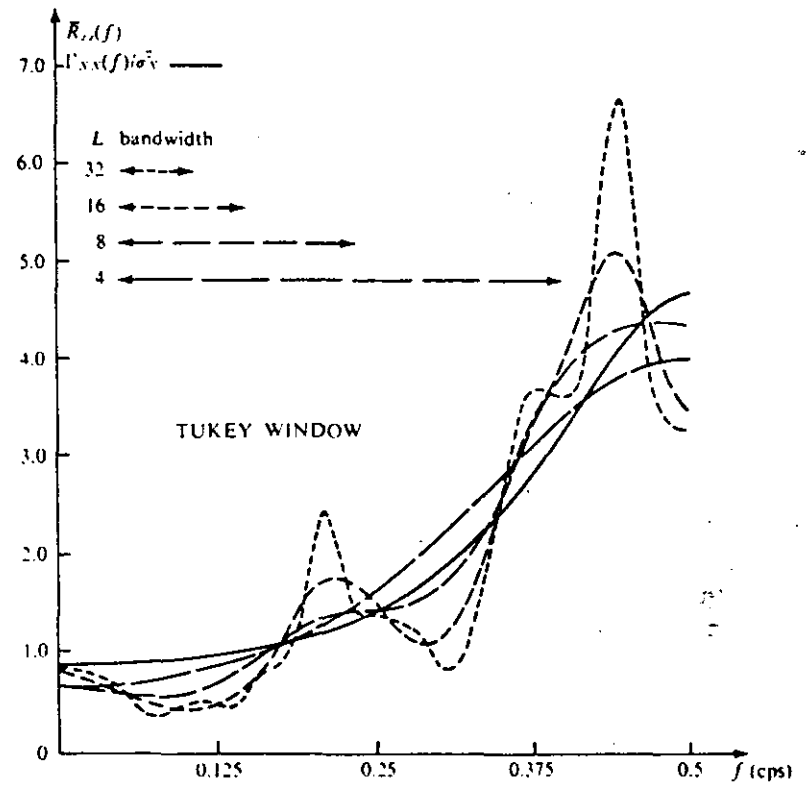


FIG. 7.4: Smoothed spectral density estimates for a first-order ar process ($\alpha_1 = -0.4$; $N = 100$)

produces significant changes in the spectrum, but even greater changes occur if L is increased to 16 or 32. Note that well-defined peaks occur at $f = 0.22$ cps and 0.44 cps which are spurious and are due to the increased variance of the estimator.

Note that the number of degrees of freedom per estimate is 33 for $L = 8$, 17 for $L = 16$ and 8 for $L = 32$. The estimate for $L = 8$ is reasonably close to the theoretical spectrum, but without knowing the correct answer one would be faced with a certain amount of ambiguity in interpreting these estimates. In particular, there would be doubt whether to accept the smooth estimate based on $L = 8$ or to decide for the apparent greater detail, but greater variance, of the estimate based on $L = 16$.

The bandwidths of the windows corresponding to various truncation points have been shown on Figures 7.3 and 7.4. This is a very useful feature since it enables the extent of the detail in the spectrum to be judged in relation to the bandwidth of the window.

A first-order ar process ($\alpha_1 = -0.9$). A similar investigation was conducted using the first-order ar process

$$X_t = -0.9X_{t-1} + Z_t. \tag{7.1.8}$$

This has the theoretical spectral density function

$$\frac{\Gamma_{XX}(f)}{\sigma_X^2} = \frac{2(0.19)}{1.81 + 1.8\cos 2\pi f}, \quad 0 \leq f \leq \tfrac{1}{2},$$

and is shown in Figure 7.5. This spectrum has an extremely narrow peak near $f = 0.5$ cps. Note that for this process the spectrum ranges in value from 0.105 to 38.0 and so the spectral density has been plotted on a logarithmic scale. This has practical merit not only because it gives a better display of the detail in the spectrum but also because, as shown in Section 6.4.3, the confidence limits are the same for all frequencies and hence are easily displayed. Therefore, spectral density estimates should always be plotted on a logarithmic scale. Because the bandwidth is constant with frequency, the frequency scale should be linear.

The mean smoothed spectral densities $\bar{\Gamma}_{XX}(f)/\sigma_X^2$ for the process (7.1.8) are shown in Figure 7.5 for the Bartlett window using truncation points $L = 8$, 16 and 32. As before, considerable bias exists near the peak for all these values of L and the bias decreases with increasing L. No obviously good truncation point is apparent from these plots and hence it is concluded that a larger truncation point, perhaps $L = 48$, is necessary. This is also apparent from consideration of the acf $\rho_{XX}(k)$, since $|\rho_{XX}(k)|$ is 0.034 for $k = 32$ and 0.0012 for $k = 64$. Hence it may be concluded that for this process a very large L will be necessary to produce small bias. Therefore, in order to produce an estimate with small variance, it will be necessary to use a very large number

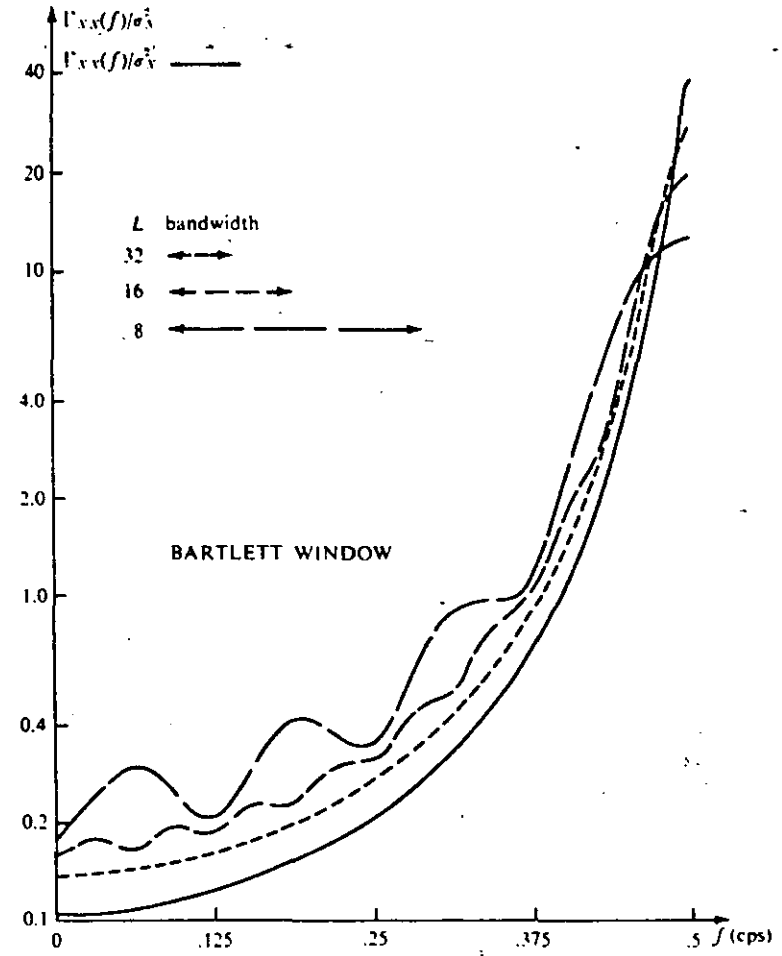


FIG. 7.5: Mean smoothed spectral density functions for a first-order ar process ($\alpha_1 = -0.9$)

of data points. Figure 7.5 shows that for L small there are ripples in the mean smoothed spectrum. These are due to the side lobes of the Bartlett window which will be discussed in Section 7.2.5.

The difficulties of trying to estimate a narrow peak in the spectrum are demonstrated in Figure 7.6. This shows $\bar{R}_{XX}(f)$ for the Bartlett window using $L = 16$, 32 and 48 based on a realization of $N = 100$ terms of the process (7.1.8). For $L = 16$ a reasonably smooth estimate is obtained, but the peak is underestimated by a factor of three. Increasing L to 32 improves the estimate near the peak slightly but more noticeably at lower frequencies. However, oscillations now begin to appear due to the increased variance. For $L = 48$ the spectral estimate becomes very erratic, indicating that more data points

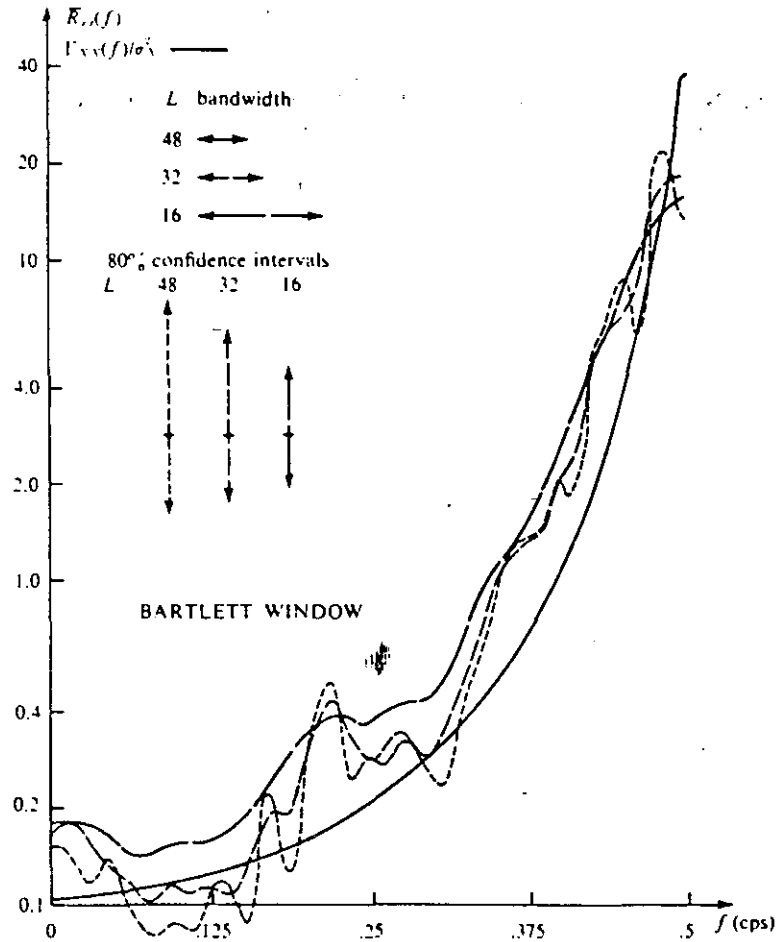


FIG. 7.6: Smoothed spectral density estimates for a first-order ar process ($\alpha_1 = -0.9$; $N = 100$)

are needed in order to get a more reliable estimate in the neighborhood of the peak. Note, however, that it is possible to obtain a satisfactory estimate in the frequency range 0 to 0.45 cps with $L = 32$ or possibly a lower value, say $L = 24$.

Comparison of Figure 7.6 with Figure 7.4, which shows the spectral estimate for $\alpha_1 = -0.4$ and $N = 100$, brings out the important point that a series of length $N = 100$ may provide an acceptable estimate of a smooth spectrum but may be far too short to provide a good estimate of a spectrum which contains a narrow peak.

A second-order ar process. To illustrate the difficulties associated with estimating a spectrum which has a two-sided peak, as opposed to the one-sided peak

in the previous example, consider the second-order ar process (5.3.36), that is,

$$X_t = X_{t-1} - 0.5X_{t-2} + Z_t. \tag{7.1.9}$$

This has the spectral density function (6.2.22), namely

$$\frac{\Gamma_{xx}(f)}{\sigma_x^2} = \frac{2(0.417)}{2.25 - 3\cos 2\pi f + \cos 4\pi f}, \quad 0 \leqslant f \leqslant \tfrac{1}{2},$$

which is plotted in Figure 7.7. Also shown in the figure are the mean smoothed spectral densities $\bar{\Gamma}_{xx}(f)/\sigma_x^2$ based on the Parzen window with $L = 8$, 16 and 32. As before, these curves show that as L increases, $\bar{\Gamma}_{xx}(f)$ more closely resembles $\Gamma_{xx}(f)$. From Table 6.6 it is seen that the bandwidths of the spectral

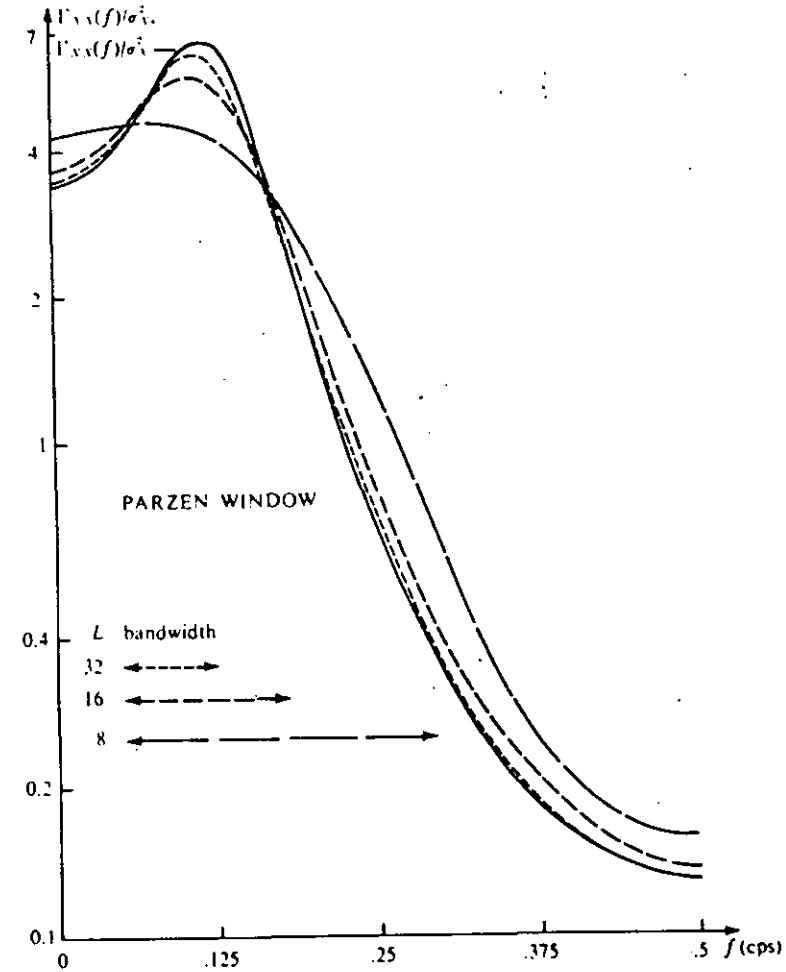


FIG. 7.7: Mean smoothed spectral density functions for a second-order ar process ($\alpha_1 = 1.0$, $\alpha_2 = -0.5$)

windows corresponding to $L = 16$, $L = 32$ are 0.12 cps, 0.06 cps. The width of the peak, defined arbitrarily as the distance between the half power points, is about 0.08 cps. It is seen from Figure 7.7 that the peak is badly estimated when $L = 8$ or 16 but that the bias is acceptably low when $L = 32$. This is due to the fact that for $L = 32$ the bandwidth of the window equals 0.06 cps, which is less than the width 0.08 cps of the peak. This suggests that the number of lags of the acf necessary to give an estimate with acceptably low bias depends on the width of the peaks in the spectrum. This point will be returned to in Section 7.2.4.

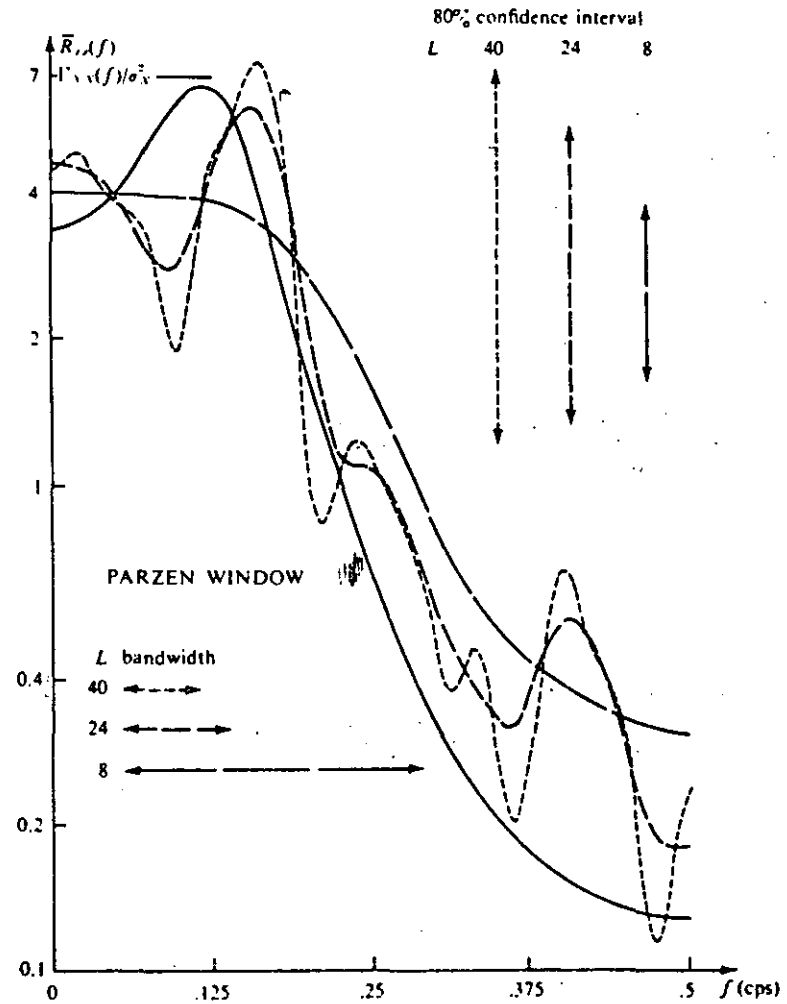


FIG. 7.8: Smoothed spectral density estimates for a second-order ar process ($\alpha_1 = 1.0$, $\alpha_2 = -0.5$; $N = 50$)

Smoothed spectral density estimates $\bar{R}_{xx}(f)$ based on a realization of $N = 50$ terms of the process (7.1.9) using the Parzen window are shown in Figure 7.8. The data used for this example were the first 50 values of the data listed in Appendix A7.4, Table A7.1. The acf for the first 50 values is given in Table A7.2. For $L = 8$ a smooth estimate is obtained but with no indication of a peak. Tripling L to 24 produces a flat peak near 0.125 cps, the position of the true peak.

As L is increased to 40, numerous other small peaks appear due to the increase in variance of the estimate, and hence doubts would remain about the true shape of the spectrum unless the theoretical spectrum were known.

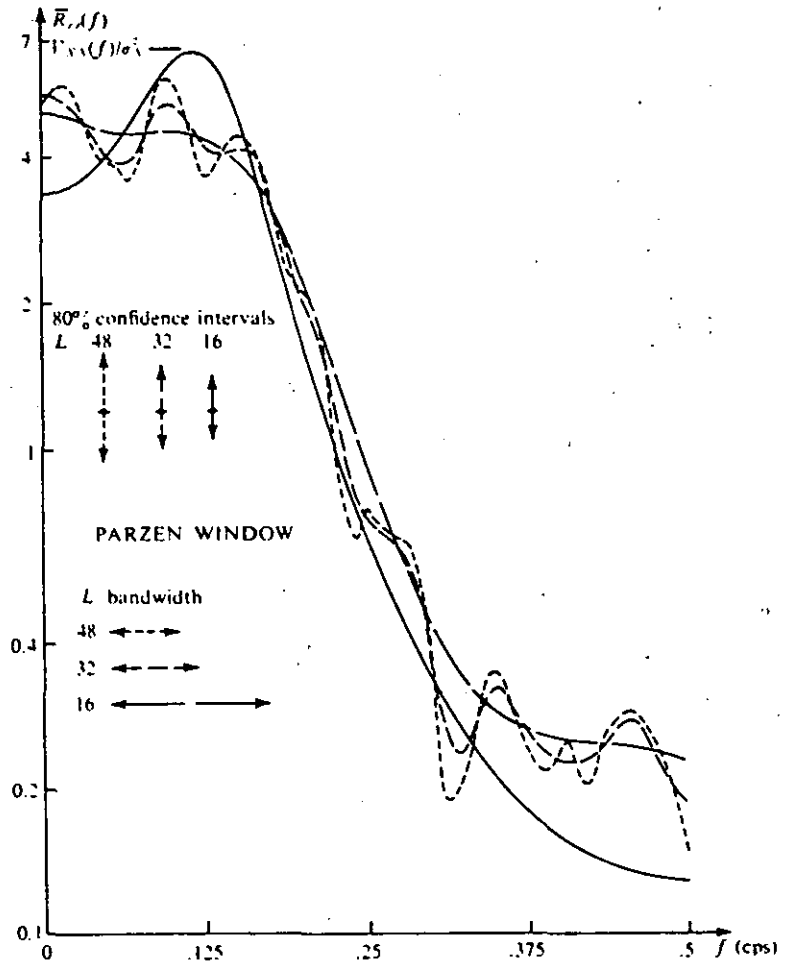


FIG. 7.9: Smoothed spectral density estimates for a second-order ar process ($\alpha_1 = 1.0$, $\alpha_2 = -0.5$; $N = 400$)

Figure 7.9 shows spectral estimates based on $N = 400$ for the same process. The data are given in Table A7.1, and the correlations based on 400 terms in Table A7.3. When $L = 48$ the estimate is closer to the true spectrum than any of the estimates based on $N = 50$, but the improvement is not as great as might be expected. Since the number of degrees of freedom per estimate is 31 when $L = 48$, it would probably be accepted that the peak is real. However, it is noticed that the peak is much narrower than the theoretical peak.

7.1.3 *Effect of window shape on smoothing*

The second aspect of smoothing investigated is the effect of using different spectral windows. Comparisons are made between the Bartlett, Tukey and Parzen windows.

The first-order process (7.1.8) has been used to calculate the mean smoothed spectral densities for the lag windows w_B, w_T and w_P for fixed values of the truncation point. These mean smoothed spectra, denoted by $\bar{\Gamma}_B/\sigma_X^2$, $\bar{\Gamma}_T/\sigma_X^2$ and $\bar{\Gamma}_P/\sigma_X^2$, are shown in Figure 7.10, together with the theoretical spectral density $\Gamma_{XX}(f)/\sigma_X^2$. All smoothed spectra are based on a truncation point $L = 12$.

The large bias and ripples in the estimate based on the Bartlett window for the lower frequencies are evident. However, near the peak the Bartlett window produces small bias. This is to be expected from (6.3.37), which gives the bias for the three windows. Thus the bias of the Bartlett window is related to the first derivative of the spectrum which is small in the neighborhood of a peak but large when the spectrum has large slope. Equation (6.3.37) shows that the important term in the bias of the Tukey and Parzen windows depends on the second derivative of the spectrum, and this is small in a region where the spectrum is linear and relatively large near a peak.

On the whole, the Tukey window has the smallest bias for a given number of lags. If the Parzen and Tukey windows are compared with equal bandwidths, however, the smoothed spectra are almost identical in shape.

The same conclusion can be drawn if comparisons are made between the variances of the estimators corresponding to two windows. From (6.4.25),

$$\text{Bandwidth} \times \text{Variance} = \text{Constant}.$$

Hence two estimators have the same variance if their bandwidths are equal. From Table 6.6 it is seen that the bandwidth of the Parzen window is 1.4 times that of the Tukey window. Hence a Tukey window with truncation point $L = 12$ has the same bandwidth and variance as a Parzen window with $L = 12 \times 1.4 = 16$.

Since both the variance and bias are approximately equal when the bandwidths are equal, it follows that, provided two spectral windows have reasonable shapes, the estimates of the spectrum obtained using equal bandwidths

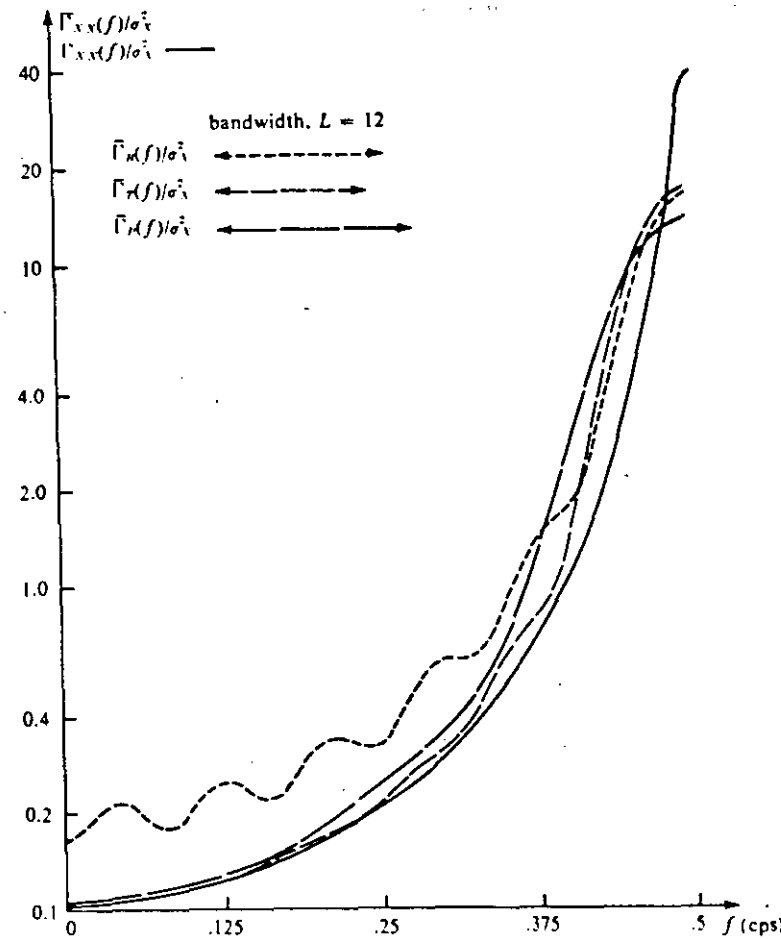


FIG. 7.10: Mean smoothed spectral density functions for a first-order ar process ($\alpha_1 = -0.9$)

should be very similar. Figure 7.11 shows equi-bandwidth comparisons for the Tukey and Parzen windows for the first-order ar process with $\alpha_1 = -0.9$ and $N = 100$. The unbroken line shows the Tukey estimate based on $L = 32$ and the crosses are the Parzen estimate based on $L = 45$. Similarly, the broken line shows the Tukey estimate based on $L = 8$ and the circles are the Parzen estimate based on $L = 12$. The agreement is so close that it can be safely concluded that no significant features in the spectrum would be missed as a result of using one window rather than the other. *Hence the empirical results of this section show that the important question in empirical spectral analysis is the choice of bandwidth and* not *the choice of window.* These questions are discussed more fully in Sections 7.2.4 and 7.2.5.

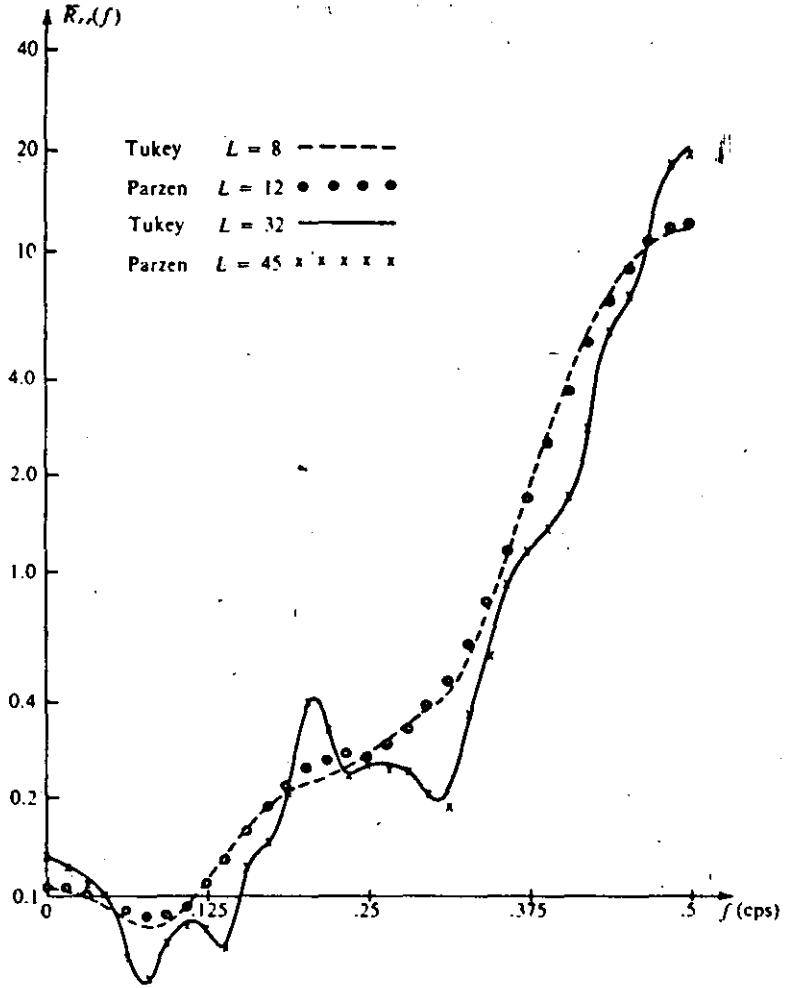


FIG. 7.11: Equi-bandwidth comparison of the Tukey and Parzen windows using a first-order ar process ($\alpha_1 = -0.9$, $N = 100$)

7.2 THE THEORY AND PRACTICE OF SMOOTHING

Section 7.2.1 contains a discussion of the various theoretical attempts to derive optimal methods of smoothing spectra. It is concluded that these theoretical solutions are not satisfactory and that a more empirical approach to spectral estimation is required. The main objectives in estimating spectra are *high stability* and *high fidelity*, and these are summarized in Section 7.2.2. In order to meet these objectives, an empirical approach to spectral estimation is introduced in Section 7.2.3. This process is termed *window closing* and is discussed in detail in Section 7.2.4. The final section deals with *window carpentry*, that is, the choice of spectral window shapes.

7.2.1 Optimal smoothing of spectral estimators

Several attempts have been made to determine smoothed spectral estimators which are optimal in some sense. Ideally, this involves choosing a lag window $w(u)$ in the range $0 \leqslant |u| \leqslant T$ so that some performance criterion is minimized. However, the practical constraint $w(u) = 0, |u| > M$, is usually imposed in order that the autocovariances need only be computed to lag M. Since the bulk of the computing time is spent on evaluating the autocovariances, there are good grounds for keeping M small relative to the record length T.

One criterion for determining an optimal lag window, mentioned in Section 6.3.5, is the mean square error

$$E[\{\bar{C}_{xx}(f) - \Gamma_{xx}(f)\}^2], \tag{7.2.1}$$

suggested in [3]. One criticism which has been made of this approach is that a good estimator for one frequency may not be a good estimator for another frequency and hence a compromise is needed which will be best in some sense for all frequencies.

An overall criterion of optimality proposed in [4] is the integrated mean square error criterion

$$\int_{-\infty}^{\infty} E[\{\bar{C}_{xx}(f) - \Gamma_{xx}(f)\}^2]\, df. \tag{7.2.2}$$

The lag window $w(u)$ which minimizes this criterion is

$$w(u) = \frac{\gamma_{xx}^2(u)}{\gamma_{xx}^2(u) + \text{Var}\,[c_{xx}(u)]}, \quad 0 \leqslant |u| \leqslant T. \tag{7.2.3}$$

Note that this estimator does not require truncation of the acvf. As mentioned above, the main function which truncation serves is to economize in the computation of the autocovariances.

The expression (7.2.3) for the optimum lag window will be recognized by communication and control engineers as being the exact analog of the expression for the frequency response for a minimum mean square error filter. Interpreting (7.2.2) in communication theory terminology, $\Gamma_{xx}(f)$ corresponds to the message and $\bar{C}_{xx}(f)$ corresponds to the message plus noise. Similarly in (7.2.3), $w(u)$ corresponds to the frequency response function of the optimum filter, $\gamma_{xx}^2(u)$ corresponds to the spectrum of the message and Var $[c_{xx}(u)]$ corresponds to the spectrum of the noise.

Another criterion suggested in [5], in an attempt to obtain a compromise estimator for all frequencies, is the expected maximum mean square error

$$E[\max |\bar{C}_{xx}(f) - \Gamma_{xx}(f)|^2]. \tag{7.2.4}$$

A further criterion which could be proposed is to minimize the integral

$$\int_{-\pi}^{\pi} \left\{\frac{E[\{\bar{C}_{xx}(f) - \Gamma_{xx}(f)\}^2]}{\Gamma_{xx}^2(f)}\right\} df. \tag{7.2.5}$$

This differs from (7.2.2) in that the expected mean square error at a particular frequency is weighted in inverse proportion to the value of the theoretical spectrum at that frequency. Thus the integrated proportional mean square error is minimized, with the obvious advantage over (7.2.2) that the estimator is weighted inversely as its variance.

It is shown below that the value of the above criteria in spectral analysis is very limited. The only useful purpose which they serve is to enable spectral windows, such as those suggested by Bartlett, Tukey, Parzen and others, to be ranked according to the various criteria. For example, the rectangular window $w_R(u)$ of Table 6.5 performs badly according to all criteria and hence can be rejected. The other windows of Table 6.5 have similar performances according to these criteria, and hence it may be concluded that these window shapes are generally "good." However, other considerations, for example, the amount of leakage through side lobes, may be used to help to decide on good window shapes. Thus, as shown in Figure 7.10, the Bartlett window is inferior to the Tukey or Parzen windows since it produces large spurious ripples in the mean smoothed spectrum.

Criticisms of the optimality approach to smoothing. Several major criticisms can be made of the optimality approach to smoothing:

(1) The optimality criteria are arbitrary. Therefore each criterion will produce a spectral window which is best is some arbitrary sense.

(2) The optimality criteria represent too rigid a mathematical formulation of the objectives of spectral analysis. For example, a physicist or an engineer may be interested in specific features of the spectrum, such as the width of a peak or the slope of the spectrum over a range of frequencies, and these criteria do not allow for this. Hence it will be shown in Section 7.2.2 that a more useful and flexible formulation of the objectives of spectral analysis is necessary and possible.

(3) Any optimum lag window, for example (7.2.3), will be a function of the unknown spectrum $\Gamma_{xx}(f)$. This criticism is not peculiar to spectral analysis since it is generally true that the best way to design anything must be based on guesses of the true picture. It is therefore of considerable importance to distinguish very clearly between *designing* a spectral analysis in advance of collecting the data and *analyzing* the data once they have been collected. We are therefore prepared to use mean square error or related criteria *in advance* of performing a spectral analysis to decide, for example, how long a record to take. After the data have been collected, it must be recognized that our guesses about $\Gamma_{xx}(f)$ may be wildly wrong. Hence any practical method of spectral analysis should be capable of standing on its own feet and should not depend critically on any major assumptions about $\Gamma_{xx}(f)$. In other words, the data must be allowed to speak for themselves.

(4) Even if there were situations where precise information about $\Gamma_{xx}(f)$ existed, the optimality approach only indicates what is best to do *on average*. Thus, the optimum lag window (7.2.3) is the best window *on average* when judged according to criterion (7.2.2). However, it might be very bad for a particular realization of a stochastic process. For example, it is possible to generate two realizations of equal length from a certain process, one of which requires a large bandwidth to give a good estimate of the spectrum, while the other requires a much narrower bandwidth.

These criticisms are major ones, and hence it must be concluded that a more robust and flexible approach to smoothing is required. In order to be able to suggest a suitable empirical procedure, it is necessary to review the general objectives of spectral analysis and to state them in useful precise terms. This is done in the next section, where the concepts of *fidelity* and *stability* are defined. An empirical procedure for smoothing of spectral estimates is then proposed in Section 7.2.3.

7.2.2 Fidelity and stability

The general objective in any spectral analysis is to estimate the function $\Gamma_{xx}(f)$ as accurately as possible. This involves two requirements:

(1) that the mean smoothed spectrum $\overline{\Gamma}_{xx}(f)$ be as close to $\Gamma_{xx}(f)$ as possible, that is, the bias

$$B(f) = \overline{\Gamma}_{xx}(f) - \Gamma_{xx}(f)$$

should be small. If this is true uniformly for all f, then $\overline{\Gamma}_{xx}(f)$ is said to reproduce $\Gamma_{xx}(f)$ with high *fidelity*.

(2) that the variance of the smoothed spectral estimator

$$\text{Var}\,[\overline{C}_{xx}(f)] \approx \frac{\Gamma_{xx}^2(f)}{T}\left(\frac{M}{b_1}\right)$$

be small. If this is the case, the estimator is said to have high *stability*.

To illustrate how the requirements of high fidelity and high stability conflict, some of the empirical findings of Section 7.1 are now reviewed.

High fidelity. First consider the plot of $\overline{\Gamma}_{xx}(f)$ for the first-order ar process with $\alpha_1 = -0.4$, shown in Figure 7.2. Using the Tukey window, high fidelity can be realized for frequencies less than 0.375 cps with a truncation point $L = 8$, and hence a bandwidth $b = 1.33/8 = 0.167$ cps. The true spectrum has a wide peak centered on $f = 0.5$ cps and in this neighborhood a truncation point of $L = 16$, or $b = 0.083$ cps, is necessary to give comparable fidelity.

The first-order ar process with $\alpha_1 = -0.9$ has a much narrower peak in the neighborhood of 0.5 cps. Figure 7.5 shows that for the Bartlett window, a truncation point of at least $L = 48$, or a bandwidth of 0.03 , is required

to achieve reasonable fidelity. Note, however, that $\Gamma_{xx}(f)$ is plotted on a logarithmic scale so that fidelity is measured by

$$\log \Gamma_{xx}(f) - \log \hat{\Gamma}_{xx}(f),$$

as opposed to

$$\Gamma_{xx}(f) - \hat{\Gamma}_{xx}(f),$$

as in the example above. In our opinion it is more logical to measure fidelity on a logarithmic scale than on a linear scale, since it is proportional changes in power that are relevant. Note that when $L = 32$, the fidelity near the peak is as high as it is in the range 0 – 0.375 cps. Hence a constant bandwidth window is adequate to estimate this spectrum.

The second-order process shown in Figure 7.7 has a more complex spectrum in that the peak is now two-sided and not one-sided as in the previous examples. Figure 7.7 shows $\hat{\Gamma}_{xx}(f)$ for the Parzen window and it is seen that, whereas the plots for $L = 8$ and $L = 16$ tend to underestimate the peak badly, the plot for $L = 32$ reproduces the peak with high fidelity. If the width of the peak is defined by the distance between the half power points, it is seen from Figure 7.7 that this width is approximately 0.08 cps. The bandwidths of the Parzen window with $L = 16$ and 32 are 0.11 cps and 0.06 cps respectively. Hence for $L = 32$ the window bandwidth is less than the width of the peak and high fidelity is obtained.

Figure 7.12 shows an even more complex spectrum which corresponds to a stochastic process consisting of two narrow band white noise sources with

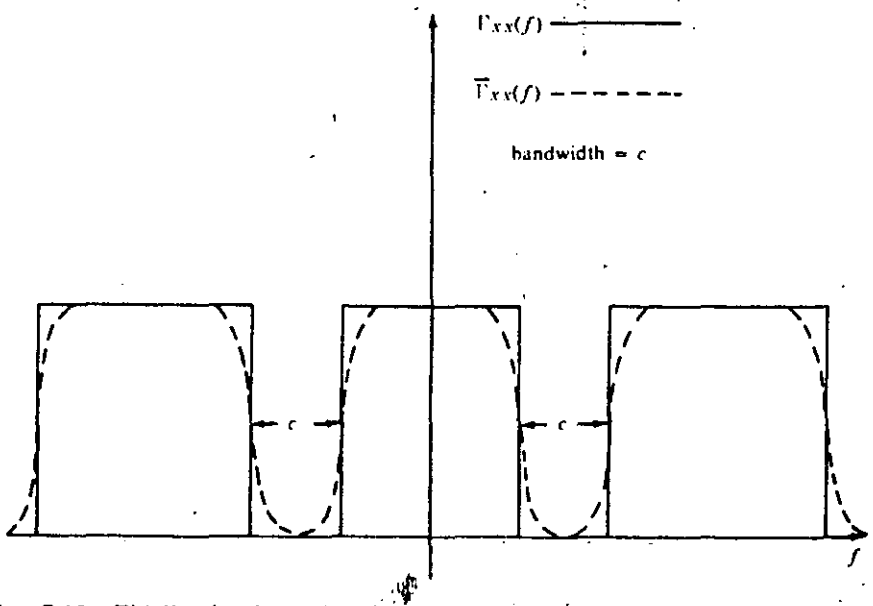


FIG. 7.12: Fidelity in the estimation of two narrow-band white noise sources

power in closely spaced bands. To achieve high fidelity for this spectrum it is necessary to use a spectral window whose bandwidth is of order c, the frequency width of the gap. Hence it can be concluded in general that to achieve high fidelity the *bandwidth of the window must be of the same order as the width of the narrowest important detail in the spectrum.* It follows that to design a spectral analysis in advance of collecting the data, it is useful to be able to guess at the width of the narrowest detail. This will be discussed in Section 7.3.1.

The word *resolution* has been used in [2] to describe a similar concept. This optical analogy presupposes that one is trying to resolve *lines* in the spectrum, that is, the spectrum is of the form

$$\Gamma_{xx}(f) = \frac{A_1}{2}[\delta(f-f_1) + \delta(f+f_1)] + \frac{A_2}{2}[\delta(f-f_2) + \delta(f+f_2)].$$

The delta functions or peaks in the spectrum are then said to be resolved if the bandwidth of the window is less than the frequency separation between the peaks. It is suggested that this is not a very useful concept in spectral analysis because real spectra can never be described in terms of delta functions, that is, the peaks are never of zero width. Further, as shown above, it is the width of the important *details* in the spectrum that are important and not merely the separation between the maxima of the peaks.

High stability. It was shown in Section 7.1 that high fidelity may be possible with a certain bandwidth but that bad estimates of the spectrum may still result if the record length is too small. For example, Figure 7.7 shows that the peak of the second-order process can be estimated with high fidelity when $L = 32$. However, Figure 7.8 shows that the estimate of the spectrum obtained with $N = 50$ gives a very bad picture of the peak. On the other hand, Figure 7.9, based on $N = 400$ terms, shows that a reasonable estimate of the spectrum is possible with $L = 32$. This is predicted by the theory of Section 6.3, since small variance, or high stability, is obtained with a large $T/M = N/L$ ratio. Hence an ideal spectral analysis is one for which M is sufficiently large for high fidelity and T/M is sufficiently large for high stability. This ideal situation is approached in the estimate of the first-order ar spectrum shown in Figure 7.3. However, in many practical problems some form of compromise is necessary between high fidelity and high stability. The practical realization of this compromise is discussed in the next section.

7.2.3 Empirical smoothing of spectral estimates

The discussion of the optimal approach to smoothing in Section 7.2.1, and specifically the criticisms of that approach, indicate the need for an empirical approach to smoothing. In particular, criticisms *(2)*, *(3)* and *(4)* call for an approach which is deliberately left *flexible*, so that many courses of action

suggested by analysis of the data are possible, and a method which allows for the possibility of *learning* from the data enough about the spectrum $\Gamma_{xx}(f)$ to choose adequate smoothing for any frequency ranges of interest.

In the terminology of the preceding section, it must be possible to infer from the data when a reasonable compromise has been achieved between high fidelity and high stability. If it is accepted that economy in the computation of the autocovariances is desirable and that smoothed spectral estimators of the form (6.3.28) are suitable, then the smoothing of a spectral estimator is completely determined by the *shape* or *mathematical form* of the window and the *bandwidth* or, equivalently, the *truncation point* of the window.

Since the effect of window shape on spectral estimates is of secondary importance, as demonstrated in Figure 7.11, an empirical approach must be based on varying the bandwidth. The following is an empirical approach to smoothing which meets the requirements and falls within the framework set out above. First, use a spectral window with an acceptable shape. Then compute smoothed spectral estimates using a wide bandwidth initially and then progressively smaller bandwidths. This empirical approach to spectral analysis was suggested in [6] and illustrated in practice in [7] and [8]. From now on this procedure of using a progressively smaller bandwidth will be referred to as *window closing*. It is discussed more fully in Section 7.2.4. The less important problem of designing spectral windows with acceptable shapes, termed *window carpentry* by J. W. Tukey, is discussed in Section 7.2.5.

7.2.4 Window closing

The technique of window closing involves computing smoothed spectral estimates with a wide bandwidth and then using progressively smaller bandwidths. The first objective of this approach is flexibility, so that any significant features of the spectrum which are of practical interest, or which suggest themselves during the process of analysis, can be explored further.

The method allows one to *learn* about the shape of the spectrum, and thus the initial choice of a wide bandwidth will usually mask a certain amount of detail in the spectrum. By allowing the bandwidth to become smaller, more significant detail can be explored. Finally, as discussed in Section 7.2.2, when the bandwidth of the window is less than the smallest significant detail in the spectrum, there is no point in making the bandwidth smaller. However, practical problems of interpretation arise due to the instability of the estimates and these will be discussed below.

Since certain records yield uninformative spectra in the same way that some uninformative likelihood functions are flat, the method enables the best bandwidth to be chosen to suit the record available.

The important practical question is when to stop the process of narrowing the bandwidth, that is, when should one stop looking for more detail in the

interest of maintaining stability. No rigid rules can be given to answer this question, since the best time to stop will depend on factors such as the degree of detail in the spectrum, the amount of prior knowledge of $\Gamma(f)$ available and the extent to which it is possible to discriminate between real detail and sampling fluctuations due to instability. Nevertheless, it is possible to distinguish between three types of situation which can occur in practice.

(1) It is sometimes possible to narrow the bandwidth sufficiently to reveal most of the significant detail without running into instability. In this case no major changes in the spectrum occur after quite large changes in bandwidth. Such a happy state of affairs is illustrated in Figure 7.3 by the spectral estimates of the first-order ar process. It is seen that only minor changes occur in the shape of the spectrum as a result of a fourfold decrease in bandwidth from $L = 4$ to $L = 16$. It can be concluded that a satisfactory estimate of the spectrum is obtained in the range 0 to 0.375 cps with a truncation point of $L = 8$, but that a higher truncation point, say $L = 12$, is required in the neighborhood of the peak.

(2) Sometimes it is clear that in no sense is the spectrum converging to a stable value. An example of this is given in Figure 7.8, which shows estimates of the second-order ar spectrum based on $N = 50$. The estimate for $L = 8$ is relatively smooth but it is impossible to conclude whether the large change from $L = 8$ to $L = 24$ is due to instability or to the appearance of more detail in the spectrum. Hence it would probably be concluded that the estimate based on $L = 8$ shows the broad features in the spectrum but that a longer series would be required to reveal more detail. Notice, however, that the spectrum based on $L = 8$ contains a great deal of useful information and hence the analysis is by no means worthless.

(3) Usually the situation falls somewhere between *(1)* and *(2)*. For example, consider the spectral estimates of the first-order ar process with $\alpha_1 = -0.4$, $N = 100$, shown in Figure 7.4. Note that well-defined peaks appear at $f = 0.22$ cps and $f = 0.44$ cps when $L = 16$. Without knowledge of the structure of this process it might be tempting to conclude that these peaks were real, since the estimate has approximately 17 degrees of freedom. The peaks become still more well-defined for $L = 32$, so there is doubt as to when to stop the window closing process. Similar comments apply to the estimates shown in Figure 7.6.

These situations are all characterized by a tendency for the estimates to converge initially, but then to diverge due to instability before definite conclusions can be drawn. Since it is not possible to say which of these spectra is closest to the truth, it is suggested that three spectra should be presented in the region where convergence is followed by divergence. However, it is important to remember that as the bandwidth is narrowed, the spectral estimate is a polynomial of high degree in $\cos 2\pi f$ and hence it is easy to

produce spurious peaks. In the limit as the estimate tends to that of the unsmoothed spectrum, it is possible to produce peaks almost anywhere. Hence a certain amount of caution is required in interpreting spectral estimates. Finally, *the spectrum has to make sense physically*, or else the analysis is of little value. In summary, the main objective of window closing is to aid physical insight in the process of estimating and interpreting spectra.

7.2.5 *Window carpentry*

It was shown empirically in Section 7.1 that window closing is much more important than window carpentry. Nevertheless, it is of some importance to pay attention to the design of the window which is to be used. As stated previously, one approach to the design problem is the optimal smoothing approach of Section 7.2.1. However, windows which perform badly on mean square error and related criteria can be shown to have bad shapes on other grounds. In this section a list is given of some of the important properties which spectral windows should possess. An analytic approach to this problem has been given elsewhere [9]; a more descriptive account is given here.

(1) For a given truncation point M, the bias due to the spectral window $W(f)$ will be small if it is concentrated about zero. From Figures 6.12 and 6.13, it is seen that $W_R(f)$ corresponding to the rectangular lag window $w_R(u)$ is more concentrated about the center frequency than any of the other windows. As shown in Table 6.6, the spectral window $W_R(f)$ has the smallest bandwidth. Hence bandwidth gives a *measure of the concentration of the spectral window*.

(2) The spectral window $W_R(f)$ has the smallest bandwidth, but the price which has to be paid is that it also has the largest side lobes, as is seen from Figure 6.13. The effect of side lobes is to permit values of $\Gamma_{XX}(g)$ at frequencies distant from f to make large contributions to the bias at the frequency f. This effect is known as *leakage*. Figure 6.13 shows that windows W_B, W_T and W_P have much smaller side lobes than W_R. However, the Bartlett window W_B has larger side lobes than windows W_T and W_P, and Figure 7.5 shows that these can be troublesome if the spectrum has a narrow peak. If side lobes are to be minimized, then window W_P is preferable to the other windows.

(3) The spectral windows $W_R(f)$, $W_B(f)$ and $W_P(f)$ are of the form

$$W(f) \propto \left\{\frac{\sin(2\pi fM/n)}{2\pi fM/n}\right\}^n, \quad n = 1, 2, 4, \tag{7.2.6}$$

and hence the lag windows $w_R(u)$, $w_B(u)$ and $w_P(u)$ are related by convolution. That is, the lag window $w_B(u) = (1 - |u|/M)$, $0 \leqslant |u| \leqslant M$, may be obtained by convolving the lag window $w_R(u) = 1$, $0 \leqslant |u| \leqslant M/2$, with itself. Similarly, the lag window $w_P(u)$ is obtained by convolving the lag window $w_R(u) = 1$, $0 \leqslant |u| \leqslant M/4$, with itself four times

An equivalent way of saying this is that window w_R is proportional to the pdf of a uniform rv, window w_B is proportional to the pdf of the mean of two uniformly distributed rv's and window w_P is proportional to the pdf of the mean of four uniformly distributed rv's. Clearly, the mean of n uniformly distributed rv's will tend to a Normal or Gaussian distribution as n tends to infinity. Thus the lag window tends to a Normal curve, and hence so does the spectral window (7.2.6), as shown in Chapter 2. In fact, Daniels [10] recommends the use of a Normal window for spectral analysis.

One effect of increasing n is to decrease the height of the side lobes, as is evident from (7.2.6). However, the spectral window also becomes flatter and wider, since the first zero occurs at $f = 2^{n-1}/2M$ and hence requires a large M in order to achieve a specified bandwidth. For example, the Parzen window w_P requires approximately 40% more lags than the Tukey window w_T to achieve a given bandwidth.

(4) The effect of altering the window shape for a given truncation point may be shown by plotting the correlation between smoothed spectral estimators separated in frequency by $f_1 - f_2$. From (6.4.11) this correlation is

$$\rho_{cc}(f_1, f_2) \approx \frac{\int_{-\infty}^{\infty} W(f_1 - g)[W(f_2 + g) + W(f_2 - g)]\,dg}{I}. \tag{7.2.7}$$

Figure 7.13 shows the correlation function (7.2.7) plotted as a function of $f_1 - f_2$ for the windows W_B, W_T and W_P of Table 6.5. It is seen that for a wide window like W_P, the correlation between neighboring estimators is large, whereas the correlation between distant estimators is small. Conversely, for a narrow window like W_B, the correlation between neighboring estimators is relatively small and between distant estimators is relatively large.

Blackman and Tukey [2] have suggested that because of the correlation between neighboring estimators, only the uncorrelated estimates should be plotted. This is a dangerous rule to apply since it is possible, for example, to miss a peak whose frequency lies halfway between the uncorrelated estimates. In our experience it is advisable to plot the estimate at a frequency spacing of at least one-half the frequency spacing between uncorrelated estimators, that is, $F \geqslant 2L$.

The above considerations suggest that windows W_R, W_T and W_P have reasonable shapes but that one should probably reject W_B because of its bad side lobe properties. Windows W_R and W_P always give positive estimates of the spectrum, whereas W_T can sometimes give negative estimates, which is undesirable. Although W_P has smaller side lobes than W_B and W_T, it is a wider window and hence requires more autocovariances to achieve a given bandwidth. This means that if the window closing procedure is applied to W_P, the spectral estimate will take a longer time to settle down to a steady value than if W_T were used.

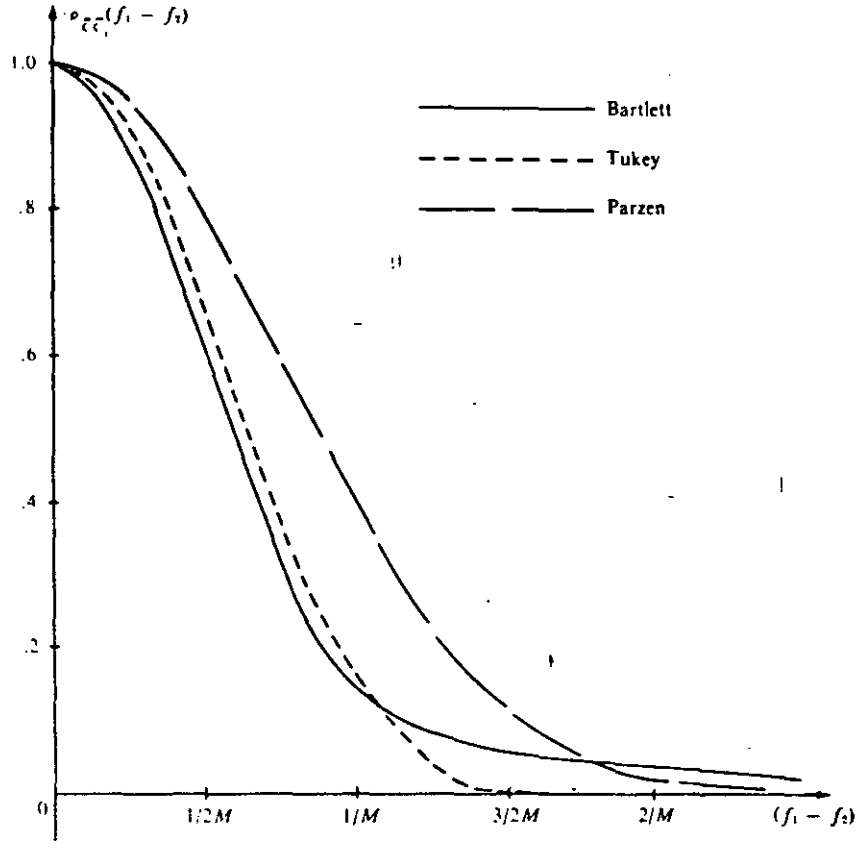


FIG. 7.13: Correlation between smoothed spectral estimators for different windows with the same truncation point

7.3 PRACTICAL ASPECTS OF SPECTRAL ESTIMATION

In this section some of the more practical aspects of spectral estimation are discussed. The first of these, discussed in Section 7.3.1, deals with the planning of a spectral analysis. The next section describes the pilot estimation of a spectrum, and this is illustrated using the batch data of Figure 5.2. Then Section 7.3.3 presents a useful practical procedure to follow in performing a spectral analysis. This procedure is illustrated in Section 7.3.4 by two practical examples which show the usefulness of the window closing technique. In Section 7.3.5 it is shown how digital filtering can be used to improve spectral estimates.

7.3.1 The design of a spectral analysis

It is now shown how a spectral analysis calculation can be designed in advance of collecting the data, and in particular, how the record length can be chosen

to meet certain specifications. There are four basic requirements which have to be met:

(1) The sampling interval Δ must be small enough so that the spectrum can be estimated in the range of interest $0 \leqslant f \leqslant f_0$. Hence Δ must be, at the most, $1/2f_0$.

(2) Care must be taken to avoid *aliasing*. This may be done in one of two ways. The first involves choosing Δ so that $\Gamma_{XX}(f)$ is effectively zero for $f > 1/2\Delta$. This requires initial knowledge of the spectrum which may not be available. Furthermore, $\Gamma_{XX}(f)$ may only be required for frequencies less than some value f_0. If f_0 is much less than the frequency beyond which $\Gamma_{XX}(f)$ is effectively zero, then it may be necessary to read the data at a much finer sampling interval than is really necessary if information is only required for frequencies $f \leqslant f_0$. The second method is to filter the signal *before sampling* so that the power above f_0 is effectively removed. This is most easily done electronically. It should be noted that a certain amount of care at this stage of the data processing may save considerable trouble and expense later on.

(3) Suppose that a guess can be made of the width a of the narrowest important peak in the spectrum, or alternatively that it is required to "detect" detail of width a or more in the spectrum. Then the truncation point M should be chosen so that the bandwidth b is less than a. For example, for the Tukey window this means that $b = 1.33/M \leqslant a$, or $M = L\Delta \geqslant 1.33/a$. The number of lags of the discrete acvf which must be computed is then $L \geqslant 1.33\, a\Delta$. In general the truncation point should be chosen according to

$$M = \frac{b_1}{a}, \tag{7.3.1}$$

where b_1 is the standardized bandwidth. Hence the number of lags required is

$$L = M/\Delta = b_1/a\Delta. \tag{7.3.2}$$

(4) For finite records, the extent to which the width of peaks can be estimated or fine detail detected is also influenced by the variance of the estimator. Hence, to be able to trust the fine structure in the spectrum, it must be possible to tie down the estimate to a given stability. This may be accomplished by specifying the number of degrees of freedom ν desired with each estimate, say 15 to 30, and then determining the length of record T from (6.4.26) and (7.3.1). This gives

$$T = \frac{\nu}{2}\frac{M}{b_1} = \frac{\nu}{2a}, \tag{7.3.3}$$

and hence

$$N = \frac{\nu}{2a\Delta} = \frac{\nu L}{2b_1}, \tag{7.3.4}$$

using (7.3.2). From Figure 3.10 it is then possible to read off the width of the 80% or 95% confidence interval for the given number of degrees of freedom. If the confidence interval so obtained is too large, increasing ν and hence N will reduce it, but at the expense of computing and data-gathering time.

A simple interpretation of (7.3.4) is obtained by considering smoothed spectral estimators separated in frequency by an amount equal to the bandwidth b of the spectral window. The covariance between these estimators will be approximately zero since there is negligible overlap of the spectral window at this spacing. The number of independent smoothed spectral estimators in the frequency band 0 to $1/2\Delta$ is therefore $(1/2\Delta)b = L/2b_1$ since $b = b_1/L\Delta$. However, unsmoothed spectral estimators separated by a frequency spacing $1/T = 1/N\Delta$ are distributed as independent χ^2 with 2 degrees of freedom. Since there are $T/2\Delta$ of these in the interval 0 to $1/2\Delta$, the total number of degrees of freedom associated with each smoothed estimator is $\nu = 2(T/2\Delta)/(L/2b_1) = 2b_1N/L$. Hence $N = \nu L/2b_1$.

The important feature of (7.3.3) is that the record length can be specified independently of the spectral window and that it depends only on ν and a.

An example. Suppose that it is required to estimate the power up to $f_0 = 2$ cps in the spectrum and that it can be safely assumed there is no appreciable power beyond this point. Thus there will be no trouble arising from aliasing. Then according to requirement *(1)*

$$\Delta = \frac{1}{2f_0} = \frac{1}{4} = 0.25 \text{ seconds.}$$

If it can be assumed that the width of the narrowest peak in the spectrum is at least 0.20 cps, and 30 degrees of freedom are deemed adequate, then (7.3.3) shows that the record duration must be at least

$$T = 30/2(0.2) = 75 \text{ seconds.}$$

Hence $N = 300$ data points are required. From Figure 3.10 it is found that 30 degrees of freedom give an 80% confidence interval for $\Gamma_{xx}(f)$ of approximately $(0.7\bar{C}_{xx}(f), 1.3\bar{C}_{xx}(f))$, that is, a 30% proportional error. To decrease this confidence interval to, say, $(0.8\bar{C}_{xx}(f), 1.2\bar{C}_{xx}(f))$ would require $\nu = 80$ and hence $T = 200$ seconds and $N = 800$.

The above calculations are of some value in deciding in advance how long a record to take but it should be emphasized that once the data have been collected, a different approach is required. Thus, if the analysis has been designed to separate peaks of width a, it may be found when the data comes to be analyzed that our guesses about a were wrong. Hence, as described in Section 7.2, it is necessary to tailor the actual analysis of the spectrum to the data available, that is, an attempt must be made to learn about the structure of the spectrum from the data. This forms the basis of the window closing procedure described in Section 7.2.4.

7.3.2 Pilot analysis

It is occasionally useful to obtain a rough estimate of the shape of a spectrum without having to compute first the acvf and then a smoothed estimate of the form (7.1.6). In particular, if it is required to prefilter the data, as will be necessary in certain problems in Chapters 9, 10 and 11, then a rough pilot analysis may be sufficient to make a guess of a good frequency response function for the filter. Since these pilot analyses are easily carried out without using an automatic computer, they also serve as useful exercises to illustrate the information contained in a spectrum.

The form of pilot analysis described below is useful when the number of observations N is 2^p, for some integer p. As will be shown in Section 7.3.5, it is capable of modification for use with any value of N. To explain the procedure, imagine that the first 64 observations of the batch data of Table 5.1 had been obtained from an experiment in which certain process variables had been deliberately varied according to the scheme in Table 7.2. In the experimental

TABLE 7.2: A fictitious experimental design for the batch data

Modification	1								2							
Week	1				2				3				4			
Day	1		2		3		4		5		6		7		8	
Shift	1	2	1	2	1	2	1	2	1	2	1	2	1	2	1	2
Batch 1	47	38	59	56	80	51	44	25	56	45	48	50	43	55	34	68
	64	65	48	40	55	58	57	59	74	54	55	62	52	41	35	38
Batch 2	23	55	71	58	37	50	50	50	50	36	45	44	38	53	54	50
	71	41	35	44	74	60	45	71	58	54	57	64	60	49	45	60

arrangement it is assumed that the yields were influenced by the batch of raw material used, shifts, days, weekly cleaning of the distillation column and two major process modifications. Note that it is only meaningful to make comparisons of the type "between batches within shifts, within days, within weeks, within modifications" since the batches being used are different each time. Hence it is natural to analyze these data by a technique called the nested analysis of variance. Thus the total variance may be decomposed as

$$\frac{1}{N}\sum_{t=1}^{N}(x_t - \bar{x})^2 = \frac{S_R}{N} + \frac{S_B}{N} + \frac{S_S}{N} + \frac{S_D}{N} + \frac{S_W}{N} + \frac{S_M}{N}, \qquad (7.3.5)$$

where the right-hand side of (7.3.5) denotes the contribution to the total variance from differences between replicates, batches, shifts, days, weeks and modifications. Thus

$$S_R = \frac{(x_2 - x_1)^2}{2} + \frac{(x_4 - x_3)^2}{2} + \cdots,$$

$$S_B = \frac{(x_4 + x_3 - x_2 - x_1)^2}{4} + \frac{(x_8 + x_7 - x_6 - x_5)^2}{4} + \cdots,$$

$$S_S = \frac{(x_8 + x_7 + x_6 + x_5 - x_4 - x_3 - x_2 - x_1)^2}{8} + \cdots,$$

$$S_D = \frac{(x_{16} + x_{15} + \cdots + x_9 - x_8 - x_7 - \cdots - x_1)^2}{16} + \cdots,$$

$$S_W = \frac{(x_{32} + x_{31} + \cdots + x_{17} - x_{16} - x_{15} - \cdots - x_1)^2}{32} + \cdots,$$

$$S_M = \frac{(x_{64} + x_{63} + \cdots + x_{33} - x_{32} - x_{31} - \cdots - x_1)^2}{64}.$$

It is seen that S_R is obtained by correlating the data with a square wave of period 2 and then squaring. Hence S_R will tend to be large if the data contains strong periodic components with period 2. Similarly, S_B will be reinforced by components of period 4, and so on. Table 7.3 shows the contributions to the mean square from each of these sources.

TABLE 7.3: Pilot spectral estimate for batch data

Source	Contribution to mean square	Estimate of spectrum	Frequency range
between replications	99.30	397.20	0.25 –0.50 cycles
between batches	15.24	121.92	0.125 –0.25 cycles
between shifts	11.77	188.32	0.0625–0.125 cycles
between days	4.10	131.20	0.0313–0.0625 cycles
between weeks	4.16	266.24	0.0156–0.0313 cycles
between modifications	0.71	90.88	0.0078–0.0156 cycles

The sum of squares S_R contains contributions not only from components with period 2 units but, as will be shown in Section 7.3.5, S_R represents contributions from components with periods between 2 and 4, that is, in the frequency range 0.25 to 0.5 cycles. Hence the average power over the frequency range 0.25 to 0.5 cycles is 99.30/0.25 = 397.2 (yield units)² per cycle. Similarly, S_M represents the total power from frequencies in the range 0.125 to 0.25 cycles, and hence the average power over this band is 15.24/0.125 = 121.92 (yield units)² per cycle.

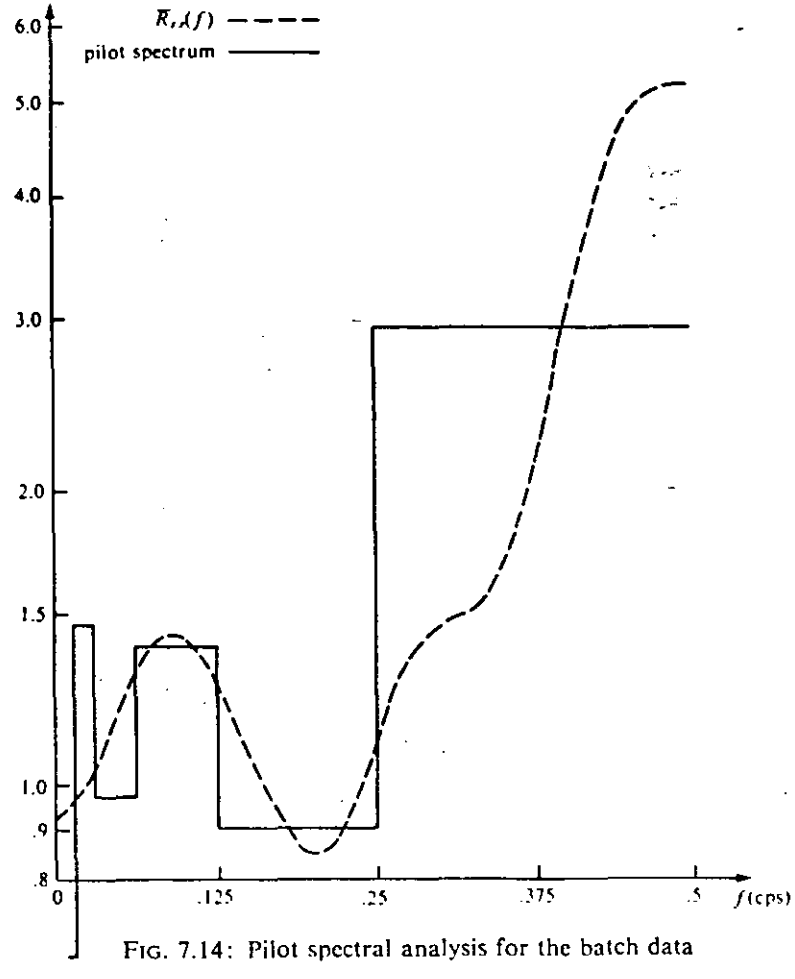


FIG. 7.14: Pilot spectral analysis for the batch data

The estimate of the spectrum obtained from the pilot analysis is plotted in Figure 7.14, together with the spectral estimate obtained from the more refined analysis of Section 7.1.1 using a Tukey window with $M = 12$. It is seen that the pilot spectral estimate agrees quite well with the refined estimate.

7.3.3 A practical procedure for spectral estimation

In this section a useful practical procedure to follow in estimating spectra is described. This consists of the following four stages:

(1) Preliminary stage. The series is examined to see if any obvious trends or periodicities are present. This is helpful in deciding whether to use the original or filtered data as will be described in Section 7.3.5. A particularly useful and

simple form of filter, which will be used extensively from now on, is the first difference filter

$$y_t = x_t - x_{t-1}.$$

A pilot analysis might also be made at this stage. The number of lags L_{max} for which the autocorrelations or autocovariances are to be computed is decided. Initially L_{max} should be chosen to be approximately $N/4$ unless this requires too much computing time. In rare cases where it is found that $N/4$ lags are not sufficient, further autocorrelations can be computed.

(2) First computation stage. The sample acf for the original and differenced data are then computed for $k = 0, 1, \ldots, L_{max}$. Plots of these functions are made to assist in deciding whether to use the original or differenced data, and what range of truncation values to use. The truncation points may be decided by scanning the chosen correlation function to see where it becomes negligible. A set of truncation values L_1, L_2, L_3 to be used in the window closing procedure is chosen to cover a fairly wide range, for example, $L_3/L_1 = 4$.

(3) Second computation stage. The spectral estimates corresponding to these truncation values are computed and plotted on a logarithmic scale, all on the same graph. The frequency spacing $1/2F$ should be chosen so that $F \approx 2L$ or $3L$. Horizontal lines corresponding to the window bandwidths (6.4.24) should be drawn on the graph to indicate the detail in the estimated spectrum relative to these bandwidths. Vertical lines corresponding to the confidence interval (6.4.21) should also be drawn for each bandwidth.

(4) Interpretation of spectral estimates. In general, the composite spectral plot obtained at stage *(3)* will fall into one of three categories, which are described below as ideal, intermediate and poor.

(a) *Ideal spectral analysis.* The variation in the spectral estimates is examined as the truncation point is increased, that is, as the bandwidth is decreased. If only minor changes occur in the estimates when L is changed beyond a certain value L^*, then it can be concluded that the window-closing procedure has revealed most of the detail in the spectrum. If the confidence interval for the spectrum at a single frequency is considered to be small enough, then the estimate based on L^* can be accepted. It is then concluded that the estimate has high fidelity and high stability. Occasionally the largest value of the truncation point L_M will be too small. In this case the spectral estimate may show a tendency to converge to some limiting form, but further estimates based on higher values of $L > L_M$ may have to be computed to confirm this behavior.

(b) *Intermediate spectral analysis.* In practice, situations where high fidelity and high stability can be attained simultaneously are rare. Usually the estimate tends to converge for small values of L but then diverges with

larger values of L. This usually implies that the estimate has become unstable before the fine detail in the spectrum has been revealed. In such situations it is suggested that a series of spectra should be presented covering the intermediate cases where convergence of the spectral estimates gives way to divergence, so that these effects are evident when the spectra are interpreted. As shown in Section 7.1, it is very easy to generate spurious peaks in the spectra by narrowing the bandwidth and hence it is better to be cautious on the side of using small rather than large truncation points.

It sometimes happens that the spectrum converges very quickly in certain frequency ranges where the spectrum is smooth and slowly in other ranges where the spectrum changes quickly. Hence different values of L may be required in different parts of the frequency range.

(c) *Poor spectral analysis.* In some cases the spectral estimates change so markedly as the bandwidth is changed that it is impossible to recommend even a range of spectra. In this badly-behaved situation, the spectral estimate corresponding to a small truncation value may have to be accepted even though it is realized that there may be considerable fine detail masked by the wide spectral window. However, the basic trouble is that N is too small and so the final conclusion should be to collect more data.

It is to be emphasized that the above rules are not rigid and should be regarded only as rough guides. Special problems may dictate other courses of action. For example, one may only be interested in a peak whose frequency is *known* but whose width is only approximately known. Hence the bandwidth could be narrowed at this frequency in order to investigate this particular peak without much concern for problems of stability at other frequencies. The above procedure is now illustrated using two practical examples.

7.3.4 Two practical examples of spectral estimation

Spectral analysis of the batch data

(1) Preliminary analysis. Inspection of the batch data of Figure 5.2 did not reveal any obvious trend in the data. Hence the acvf estimate (7.1.2) was used and computed up to $l_{max} = 18$ lags. Inspection of the pilot spectrum, Figure 7.14, revealed that the spectrum was quite smooth since the range of variation was only about four. Hence differencing was not considered necessary.

(2) First computation stage. The sample acf $r_{xx}(k)$ for these data was plotted in Figure 5.6. From this plot it is seen that the acf is essentially zero for $k > 10$, and hence it was decided to use truncation values of $L = 4$, 8, and 16.

(3) Second computation stage. The spectral estimates for these values of L using the Tukey window were computed and plotted together in Figure 7.15.

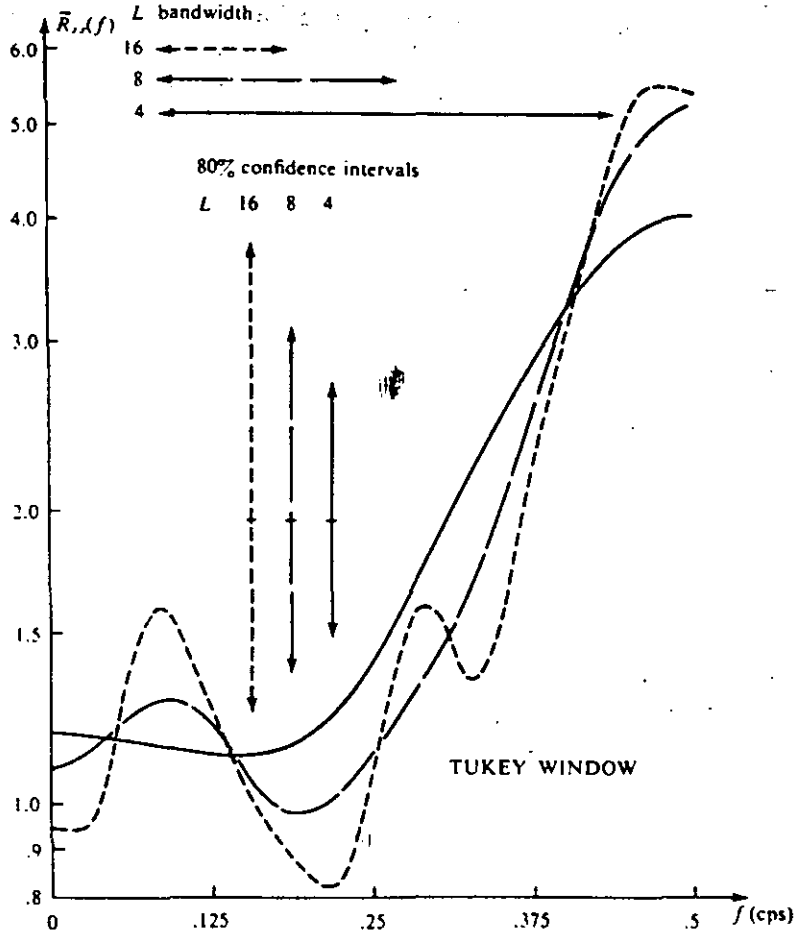


FIG. 7.15: Smoothed spectral density estimates for the batch data

Also shown are the window bandwidths and confidence interval for each value of L.

The similarity in the behavior of this spectrum and the spectrum of the simulated first-order ar process with $\alpha_1 = -0.4$, $N = 100$, shown in Figure 7.4, should be noted. It is clear that the bandwidth corresponding to $L = 4$ is too wide to reveal all the detail in the spectrum, but the changes from $L = 8$ to $L = 16$ indicate that the spectrum is very smooth and that there is no point in closing the window any further. Despite the fact that N is small, it may be concluded that this is a satisfactory spectral analysis and that little would be lost in accepting the value based on $L = 8$. The number of degrees of freedom corresponding to $L = 8$ using the Tukey window is 23, which is acceptable.

Spectral analysis of the radar data. As another example of the approach of Section 7.3.3, Figure 7.16 shows the sample acf of the radar return signal of Figure 5.1. Figure 7.17 shows estimates of the spectral density function using the Bartlett window with $L = 16$, 48 and 60 for a series consisting of $N = 448$ terms. The frequency range is shown as 0 to 0.5 cps, since the actual frequency range is not of great importance. It is seen that the estimate based on $L = 16$ is smooth and does not reveal the peak which might be expected because of the oscillatory acf. For $L = 32$, not shown in the diagram, well-defined peaks appear at approximately $f = 0.07$ cps and $f = 0.25$ cps. Increasing L to 48 reveals these peaks nicely, and it is seen that little change in the spectrum occurs by increasing L to 60. Hence the final spectral window used has an equivalent bandwidth of $1.5/60 = 0.025$ cps, and the estimate at any frequency has $3(448)/60 \approx 22$ degrees of freedom, which is adequate. The confidence interval and bandwidth corresponding to $L = 60$ are shown in the figure as vertical and horizontal lines.

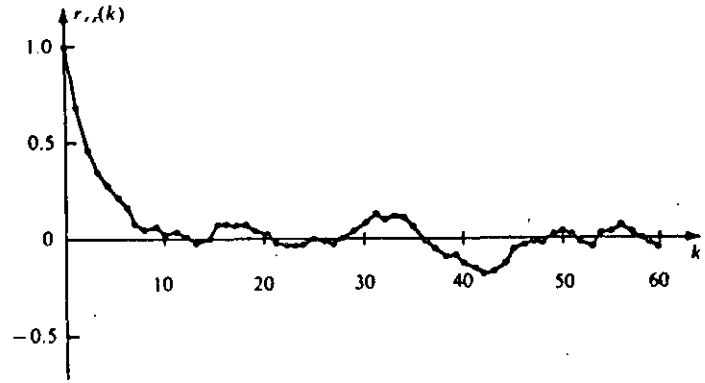


FIG. 7.16: Autocorrelation function estimate for the radar return signal of Figure 5.1

In this particular experiment it was necessary to use 60 lags in order to describe adequately the narrow peak at 0.07 cps. The spectrum for frequencies greater than 0.1 cps, however, is quite adequately defined with only 32 lags. This again illustrates the important practical point that whereas very large values of L may be needed to show up a very narrow peak in a certain part of the spectrum, the remainder of the spectrum may be very successfully analyzed using much smaller values. In this radar example, the peak at $f = 0.07$ cps was of little interest since it was related to the scanning frequency of the radar. In fact, the region of greatest practical interest is that beyond about 0.1 cps, and this could be successfully analyzed with relatively small values of L, such as 32 or 40.

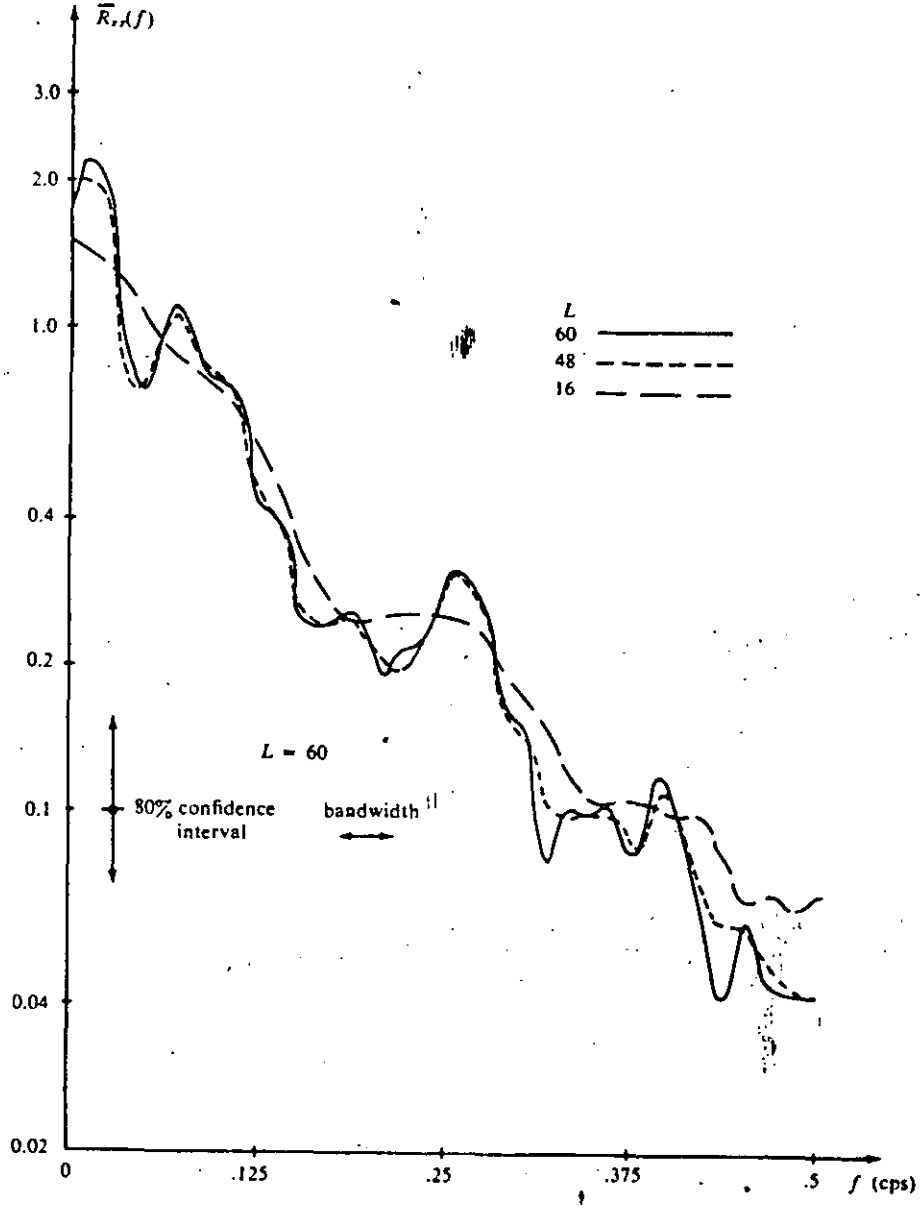


FIG. 7.17: Smoothed spectral density estimates for the radar return signal of Figure 5.1

7.3.5 Digital filtering

It is sometimes apparent from preliminary investigation of data, such as pilot analyses or visual inspection, that the spectrum is badly behaved. By this is meant that most of the power is contained in one or a few very narrow bands. Because of leakage from spectral windows, such peaks can cause errors in the spectral estimates where there is less power. Hence it may be advantageous to filter the data digitally in order to improve the spectral estimates at these frequencies.

Digital filtering is simply the process by which a set of input data x_t is transformed into a set of output data y_t by means of a linear relationship such as

$$y_t = \sum_{l=-\infty}^{\infty} h_l x_{t-l}, \tag{7.3.5}$$

where the h_l are suitably chosen weights. Note that it is not necessary to invoke the condition of "physical realizability" which implies $h_l = 0$, $l < 0$. Hence the filter (7.3.5) can operate on values of x to the left of y_t ("past" values of x) and on values of x to the right of y_t ("future" values of x). As shown in Chapter 2, the transfer function of the digital filter (7.3.5) is

$$H(\mathcal{Z}) = \sum_{l=-\infty}^{+\infty} h_l \mathcal{Z}^{-l}, \tag{7.3.6}$$

using $\mathcal{Z}$ transform notation.

Substituting $\mathcal{Z} = e^{+i2\pi f\Delta}$ in (7.3.6) gives the frequency response function of the filter. A special case of some importance in what follows occurs when $h_l = h_{-l}$. For these *symmetric filters*, the frequency response function is

$$H(f) = h_0 + 2\sum_{l=1}^{\infty} h_l \cos 2\pi f l\Delta, \quad -\frac{1}{2\Delta} \leqslant f < \frac{1}{2\Delta}. \tag{7.3.7}$$

Hence the phase shift between input and output will either be zero or π since (7.3.7) contains no imaginary part. The gain $G(f)$ is obtained by taking the modulus of $H(f)$ in (7.3.6).

It has been noted in Section 7.3.2 that pilot estimation of a power spectrum consists of applying suitable digital filters to the time series and then squaring the output from these filters. An older application of digital filters is the smoothing of time series. For example, economic time series are sometimes smoothed to reduce the effect of short term (high-frequency) fluctuations and hence to enable a study of trends in economic variables to be made.

Examples of digital filters. Some simple digital filters are now defined and their properties discussed. For simplicity, it is assumed in these examples that the sampling interval Δ equals unity.

(1) Smoothing by threes. A time series may be "smoothed by threes" by combining the observations according to

$$y_t = h_{-1}x_{t+1} + h_0x_t + h_1x_{t-1}.$$

If the weights are equal, this reduces to the symmetric form

$$y_t = \tfrac{1}{3}(x_{t+1} + x_t + x_{t-1}),$$

which has the transfer function

$$H(\mathcal{Z}) = \tfrac{1}{3}(\mathcal{Z} + 1 + \mathcal{Z}^{-1}).$$

The frequency response function is

$$H(f) = \tfrac{1}{3}(1 + 2\cos 2\pi f) = \frac{\sin 3\pi f}{3 \sin \pi f}, \quad -\tfrac{1}{2} \leqslant f < \tfrac{1}{2}.$$

Hence the gain and phase functions are

$$G(f) = \left|\frac{\sin 3\pi f}{3 \sin \pi f}\right|, \quad -\tfrac{1}{2} \leqslant f < \tfrac{1}{2}$$

$$\phi(f) = \begin{cases} 0, & |f| \leqslant \tfrac{1}{3} \\ \pi, & \tfrac{1}{3} < |f| < \tfrac{1}{2}. \end{cases}$$

(2) Summing. Consider the sum filter

$$y_t = (x_t + x_{t-1}).$$

This filter has the transfer function

$$H(\mathcal{Z}) = (1 + \mathcal{Z}^{-1}),$$

and hence the frequency response function

$$H(f) = (1 + e^{-i2\pi f}) = 2e^{-i\pi f}\cos \pi f, \quad -\tfrac{1}{2} \leqslant f < \tfrac{1}{2}.$$

The gain and phase functions are

$$G(f) = 2\cos \pi f, \quad -\tfrac{1}{2} \leqslant f < \tfrac{1}{2}$$
$$\phi(f) = \pi f,$$

and hence it acts as a low-pass filter.

(3) Differencing. A difference filter is defined by

$$y_t = (x_t - x_{t-1}) \tag{7.3.8}$$

and has the frequency response function

$$H(f) = 2je^{-i\pi f}\sin \pi f = 2e^{-i\pi(f-1/2)}\sin \pi f, \quad -\tfrac{1}{2} \leqslant f < \tfrac{1}{2}.$$

The gain and phase functions are

$$G(f) = 2|\sin \pi f|, \quad -\tfrac{1}{2} \leqslant f < \tfrac{1}{2}$$
$$\phi(f) = \begin{cases} \pi(f + \tfrac{1}{2}), & -\tfrac{1}{2} \leqslant f < 0 \\ \pi(f - \tfrac{1}{2}), & 0 \leqslant f < \tfrac{1}{2}. \end{cases}$$

Hence the difference filter acts as a high-pass filter. The gain function for the difference filter is shown in Figure 1.4.

(4) Sum and difference filters. Consider now a filter corresponding to m summings and n differencings. From (2.3.26) and (2.3.27), the overall gain and phase functions are

$$G_{m,n}(f) = 2^{m+n}(\cos \pi f)^m |\sin \pi f|^n, \quad -\tfrac{1}{2} \leqslant f < \tfrac{1}{2}; \tag{7.3.9}$$

and

$$\phi(f) = \begin{cases} \pi f(m + n) + n\pi/2, & -\tfrac{1}{2} \leqslant f < 0, \\ \pi f(m + n) - n\pi/2, & 0 \leqslant f < \tfrac{1}{2}. \end{cases}$$

Note that the gain is a maximum at the frequency

$$f_0 = \frac{1}{2\pi}\arccos\left(\frac{m - n}{m + n}\right). \tag{7.3.10}$$

Slutsky's Theorem. Using (7.3.9), if a time series with spectrum $\Gamma_{zz}(f)$ is the input to the above sum and difference filter, the output spectrum is

$$\Gamma_{xx}(f) = 2^{m+n+1}(\cos \pi f)^{2m}(\sin \pi f)^{2n}\Gamma_{zz}(f), \quad 0 \leqslant f \leqslant \tfrac{1}{2}.$$

For white noise, that is, $\Gamma_{zz}(f) = 2$, $0 \leqslant f \leqslant \frac{1}{2}$, it may be verified that as m and n tend to infinity, such that the ratio n/m tends to a constant θ, $\Gamma_{xx}(f)$ tends to a delta function $\delta(f - f_0)$, where

$$\cos 2\pi f_0 = \frac{1 - \theta}{1 + \theta},$$

using (7.3.10). It was shown in Section 6.2.2 that a stochastic process whose spectrum is a delta function is a sine or cosine wave. Hence this result shows that if white noise is summed and differenced sufficiently, the output becomes a sine wave. This result is due to Slutsky [11], who suggested that some apparently periodic or pseudo-periodic behavior in economic time series could be accounted for by the smoothing procedures used on the data.

(5) Pilot analysis filters. The filters used for the pilot analysis in Section 7.3.2 are made up of summing and differencing operations with suitable delays. For example, the $\mathcal{Z}$-transform of the filter associated with the sum of squares S_M in Section 7.3.2 is

$$H(\mathcal{Z}) = \frac{(\mathcal{Z}^{32} - 1)^2}{8(\mathcal{Z} - 1)},$$

and hence the gain function is

$$G(f) = \frac{2(\sin 32\pi f)^2}{|\sin \pi f|}, \quad -\tfrac{1}{2} \leqslant f < \tfrac{1}{2}.$$

Similar expressions may be obtained for the other filters. As noted in Section 7.3.2, the filter corresponding to S_R has a maximum at $f = 0$ [illegible]s and its first

zero at $f = 0.25$ cps. Hence the output from this filter needs to be averaged roughly over the range 0.25 cps to 0.5 cps if an estimate of the average power in this frequency band is required.

(6) Autoregressive–moving average filters. Autoregressive–moving average filters are generalizations of the above filters and are defined by

$$\sum_{i=-m}^{m} \alpha_i y_{t+i} = \sum_{i=-l}^{l} \beta_i x_{t+i}. \tag{7.3.11}$$

The main difference between the filter (7.3.11) and those previously considered is that the output y_t depends on other values of the output as well as on values of the input, that is, these filters use *feedback*. The transfer function for the filter (7.3.11) is

$$H(\mathcal{Z}) = \frac{\beta_{-l}\mathcal{Z}^{l} + \cdots + \beta_0 + \beta_1\mathcal{Z}^{-1} + \cdots + \beta_l\mathcal{Z}^{-l}}{\alpha_{-m}\mathcal{Z}^{m} + \cdots + \alpha_0 + \alpha_1\mathcal{Z}^{-1} + \cdots + \alpha_m\mathcal{Z}^{-m}}.$$

Filters of this form have greater flexibility than those described previously and are also more economical, in the sense that a good approximation to a given filter shape can be achieved with a small number of terms using parameters on both sides of (7.3.11). This fact was illustrated in Figure 5.20, where it was shown that an ar process of two terms gave a better fit to the data than did an ma process of order ten.

For most applications of digital filtering, it is possible to use the sum and difference filters or generalizations of them. However, if special care is required in the design of the filter, then the parameters α_i, β_i in (7.3.11) can be chosen empirically as described in [12]. First the ideal filter shape is specified and then the parameters chosen so as to minimize some performance criterion at a fixed number of frequency points. For example, one could choose the parameters to minimize the mean square error of deviations between the actual and ideal filter shapes at selected frequency points. Alternatively one could use a Tschebysheff criterion, that is, minimize the maximum distance between the actual and ideal filter shape. With the aid of a digital computer, such calculations are easily performed.

Uses of digital filters. Some of the most important uses of digital filters are as follows:

(a) *For pilot estimation of spectra.* This requires a bank of band-pass filters, for example, those given in [13].

(b) *For smoothing data.* This removes high frequency oscillations and requires a low-pass filter.

(c) *For removing trends from data.* This requires a high-pass filter, which may be obtained by using a low-pass filter and subtracting the output of the low-pass filter from the original data. Removal of low-frequency trend is often a necessary preliminary to estimating a spectrum. An example is given below where failure to remove trends produced a serious bias in the spectral estimate.

(d) *For partitioning time series.* In studying relations between time series it is often better to split up the original time series x_t according to

$$x_t = x_t^{(1)} + x_t^{(2)} + \cdots + x_t^{(k)} \tag{7.3.12}$$

by using a bank of band-pass filters. For example, prior information may suggest that the low-frequency components in x_t can be predicted more accurately from a knowledge of the low-frequency components in some other series y_t than from either x_t or y_t directly. Hence each series $x_t^{(i)}$ in (7.3.12) can be used as a separate time series for further analysis. Applications of this approach to the analysis of meteorological time series have been given in [14] and to economic time series in [13].

Binomial filters. A particularly simple set of filters which could be used for this purpose has been given in [15]. These make use of the sum and difference filters introduced earlier. Thus, using $\mathcal{Z}$ transforms,

$$\begin{aligned} x_t &= [\tfrac{1}{2}(1 + \mathcal{Z}^{-1}) + \tfrac{1}{2}(1 - \mathcal{Z}^{-1})]^k x_t \\ &= (\tfrac{1}{2})^k[(1 + \mathcal{Z}^{-1})^k + k(1 + \mathcal{Z}^{-1})^{k-1}(1 - \mathcal{Z}^{-1}) + \cdots \\ &\qquad + \binom{k}{i}(1 + \mathcal{Z}^{-1})^{k-i}(1 - \mathcal{Z}^{-1})^i + \cdots + (1 - \mathcal{Z}^{-1})^k]x_t. \end{aligned}$$

Hence the time series x_t can be filtered into $k + 1$ time series using $(k + 1)$ filters, the ith filter having $\mathcal{Z}$ transform

$$H_i(\mathcal{Z}) = \left(\frac{1}{2}\right)^k \binom{k}{i}(1 + \mathcal{Z}^{-1})^{k-i}(1 - \mathcal{Z}^{-1})^i.$$

Hence the output $x_t^{(i)}$ from this filter is obtained by passing the original series through $k - i$ summation filters and i difference filters and then multiplying by the coefficient $(\frac{1}{2})^k\binom{k}{i}$.

Using (7.3.10) it is seen that the ith filter has a peak frequency at

$$\cos 2\pi f_0 = \frac{k - 2i}{k}, \quad i = 0, 1, 2, \ldots, k.$$

For example, with $k = 4$, the peak frequencies are at 0, 0.167, 0.25, 0.417 and 0.5 cps.

An example of digital filtering. This example relates to the estimation of the spectrum of a radar return signal and is described in greater detail in [16]. For technical reasons the return signal cannot be measured independently of the yawing motion of the aircraft which is being tracked by the radar. Figure 7.18 shows a section of record where the yaw motion is extreme and dominates the high-frequency noise whose spectrum was required.

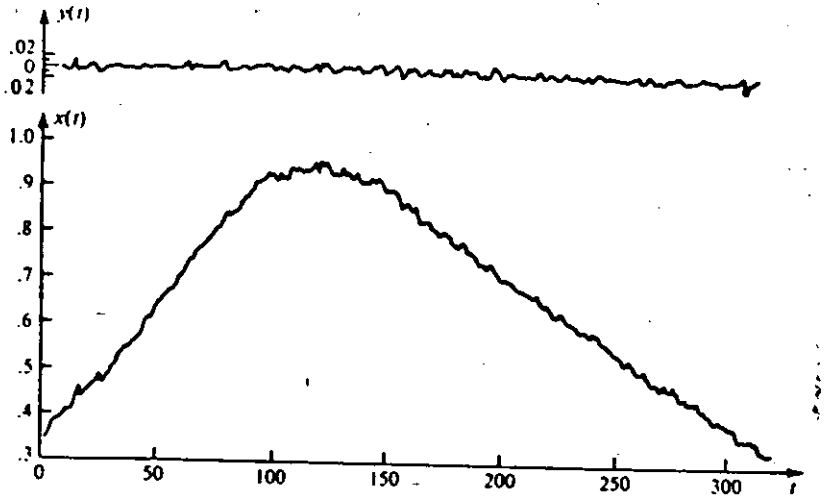


FIG. 7.18: Original and filtered radar return signals

The record was read at 320 points and then filtered using the symmetric filter (7.3.5) with

$$h_0 = 1 - \frac{1}{m+1},$$

$$h_i = h_{-i} = -\frac{1}{m+1}\left(\frac{1}{2} + \frac{1}{2}\cos\frac{i\pi}{m+1}\right), \quad i = 1, 2, \ldots, m,$$

and $m = 9$. These weights are the same as the Tukey window w_T of Table 6.5 but they have been modified to make the filter a high-pass filter and normalized to make the sum of the weights equal to one.

The filtered series y_t is shown above the original series x_t in Figure 7.18 and the effectiveness of the filter in removing the low frequencies can be seen.

The acvf's of the original and filtered series were calculated and spectral estimates obtained for different values of the truncation point L using the Bartlett window. The estimate of the spectrum of the filtered series did not change for values of L greater than 30, but a much higher value of L was required for the original series. To compare the high frequency ends of the two spectra, Figure 7.19 shows the spectra for the same value of L for both series. It is seen that at higher frequencies the original series has a spectrum which is approximately 10 times higher than that of the filtered series. This is because the low-frequency power is so great in the original series that it leaks into the estimates at high frequencies, producing large biases.

Several records of this type exhibiting various degrees of yaw were available, and some contained no yaw motion at all. When the records containing yaw were filtered before estimating the spectrum, excellent agreement was obtained with the spectra estimated when yaw was absent.

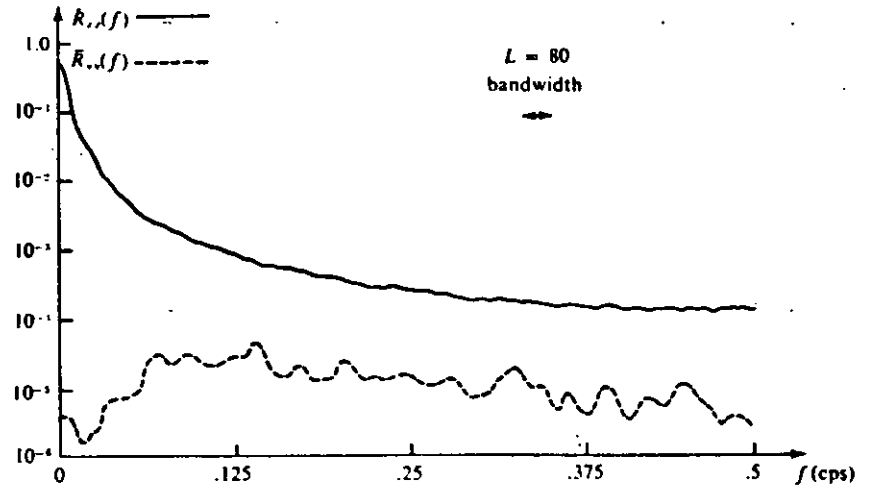


FIG. 7.19: Spectral estimates for the original and filtered radar return signals

7.4 USES AND EXAMPLES OF SPECTRAL ANALYSIS

Since spectral methods were introduced by M. S. Bartlett and J. W. Tukey about fifteen years ago, a wide variety of applications of the technique have been reported in the literature. Most of these applications can be classified according to the following three broad areas: model building, the design of experiments and frequency response studies.

Further applications of spectral analysis will be given later on in the book, but for the present it is convenient to show how knowledge of the spectrum of a single time series is useful in these three areas.

7.4.1 Model building

The shape of a spectrum will sometimes suggest features which need to be explained in any model which may be proposed for the time series. For example, the presence and magnitudes of peaks in the spectrum may throw light on some basic periodicities which require physical interpretation.

In situations where spectra are being studied to obtain a better understanding of the physical mechanism generating the time series, a single spectrum will rarely be very useful. The most important clues concerning the model may be obtained when external conditions change and hence several spectra need to be studied. These external conditions may be beyond one's control as in the first example below, or may be deliberately varied in the form of a planned experiment as in the second example.

Example 1. Figure 7.20 shows the spectrum of the horizontal velocity component of atmospheric turbulence, given in [17]. The up pectrum was

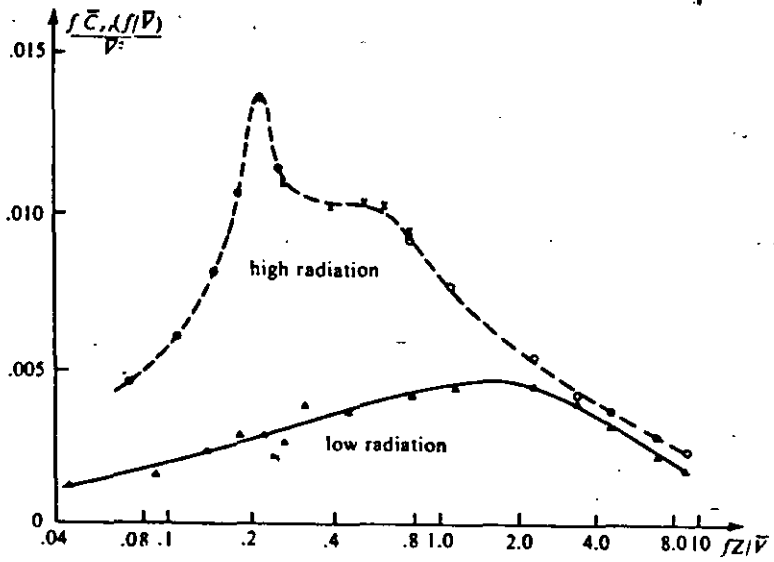


FIG. 7.20: Velocity spectra for horizontal wind components

obtained from measurements made under clear skies (high solar radiation) and the lower spectrum from measurements made under cloudy skies (low solar radiation). Note that the spectrum contains much more power during periods of high radiation and that most of this power appears at low frequencies. In particular, the peak in the spectrum moves toward lower frequencies with increasing radiation and the power at higher frequencies seems to be independent of radiation. These conclusions are borne out in more detailed studies given in [18], where the following physical explanation for this behavior is suggested: at high frequencies the main causes of atmospheric turbulence are *mechanical* or frictional forces, and at low frequencies the main causes are heat *convection* due to solar radiation.

In Figure 7.20 the ordinate plotted is proportional to $f\bar{C}_{xx}(f)$, since the abscissa is $\log f$. As a result, the area under the curve is still the total variance or power. Since the mean wind speed $\bar{V}$ changes the intensity of turbulence, and also its distribution with frequency, in a known manner, the quantities actually plotted in Figure 7.20 are non-dimensional quantities $f\bar{C}_{xx}(fZ/\bar{V})/\bar{V}^2$ and $fZ/\bar{V}$.

Example 2. Figure 7.21 shows three spectra relating to measurements of the vertical velocity component of atmospheric turbulence made at three different heights and described in [18]. The variables actually plotted in Figure 7.21 are the non-dimensional quantities $f\bar{C}_{xx}(f)/\bar{V}^2$ and $fZ/\bar{V}$ where Z is the height from the ground. The figures show that the upper two spectra are very similar in shape and also have maxima at the same frequency. The lower spectrum is not directly comparable to the others. On the basis of these and

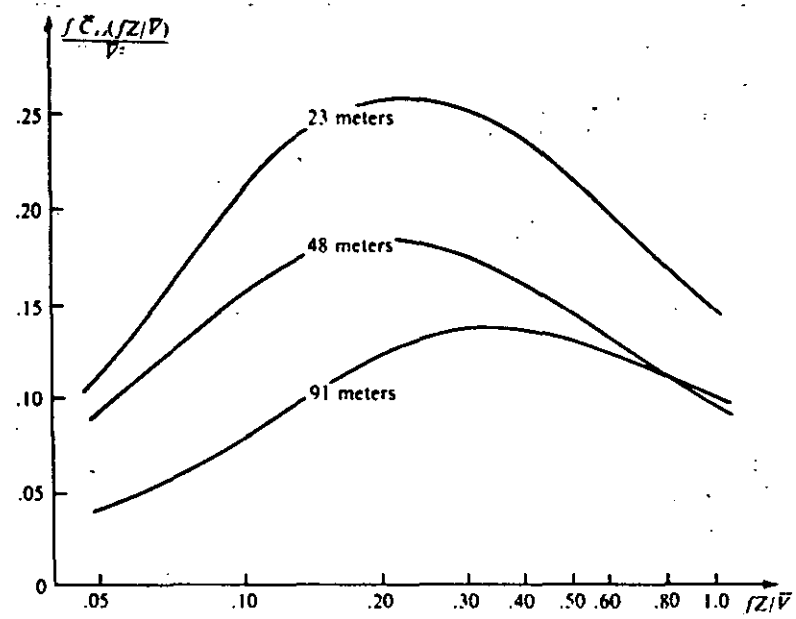


FIG. 7.21: Velocity spectra for vertical wind components

other spectra reported in [18], good agreement was found over the important regions of the spectrum, between the empirical spectra and two suggested theoretical expressions,

$$\Gamma_{xx}(h) = 2h(1 + h)^{-3}$$
$$\Gamma_{xx}(h) = \tfrac{10}{9}h(1 + h)^{-8/3},$$

where $h = \gamma Zf/\bar{V}$ and γ is a constant. Further examples may be found in the literature of explanations for complex physical phenomena which have been suggested or partially verified by means of spectral analysis.

7.4.2 *The design of experiments*

As a simple example of this type of application, suppose that it is required to design an experiment to estimate the slope of a response surface $\eta(v_1, \ldots, v_n)$ with a view to maximizing or minimizing η. For example, when $n = 1$, $\eta(v)$ might be the yield or cost per ton of a chemical product and v the flow rate of feedstock into the reactor. Two situations can be distinguished in practice. In the first, the process is a batch process and the variables v_i are set at the beginning of a run. Alternatively, the process is a continuous process but adjustments are made so infrequently that the dynamics of the process can be neglected. In the second, the process is continuous and the slope is measured on a continuous basis as in a maximum-seeking or hill-climbing

control system [19]. On the basis of the estimated slope the control system can adjust the operating conditions in order to maximize yield or minimize cost.

Suppose in the first situation that adjustments are made to v at unit time intervals according to

$$v_t = a \cos \frac{\pi t}{b}, \quad t = 1, 2, \ldots, N. \tag{7.4.1}$$

Suppose that the amplitude a of the cosine wave is fixed and that it is required to choose its period $2b$ so as to minimize the variance of the estimated slope. Assuming a linear model

$$Y_t = \eta(v_t) + Z_t = \theta_1 v_t + Z_t,$$

where Z_t is a noise or error term, the usual least squares estimate of θ_1 is

$$\hat{\theta}_1 = \frac{\sum_{t=1}^{N} y_t v_t}{\sum_{t=1}^{N} v_t^2} = \frac{2}{Na^2} \sum_{t=1}^{N} y_t v_t.$$

It is shown in appendix A7.2 that the variance of the corresponding estimator $\hat{\Theta}_1$ is approximately

$$\text{Var}\,[\hat{\Theta}_1] \approx \frac{4}{Na^2}\, \Gamma_{zz}\left(\frac{1}{2b}\right). \tag{7.4.2}$$

Hence, for fixed a, the variance is a minimum when the frequency $1/2b$ of the perturbation signal corresponds to the minimum in the spectrum of the noise. In other words, the signal-to-noise ratio $a^2/2\Gamma_{zz}(1/2b)$ is maximized.

For the second situation the perturbation signal is a continuous cosine wave

$$v(t) = a \cos 2\pi f_0 t.$$

It is shown in Appendix A7.2 that the variance of the slope is then minimized when the signal-to-noise ratio

$$\frac{a^2 G^2(f_0)}{2\Gamma_{zz}(f_0)} \tag{7.4.3}$$

is a maximum. In (7.4.3), $G(f_0)$ is the gain of the system at frequency f_0.

Example 3. The batch data of Figure 5.2 was obtained from a process when no deliberate changes were made in the process variables. Thus the smoothed spectrum $\bar{C}_{xx}(f)$ shown in Figure 7.15 gives an estimate of the spectrum $\Gamma_{zz}(f)$ of the noise in the process. This information can be used to design an experiment in which some process variable is deliberately varied according to the cosine wave (7.4.1) or some other periodic signal, say a square wave of period $2b$.

Figure 7.15 shows that the spectrum is approximately flat between $f = 0$ and $f = 0.25$ cps but rises sharply when $f > 0.25$ cps. Since $b = 1$ corresponds to $f = 0.5$ cps and $b \geq 2$ corresponds to the frequency range 0 to 0.25 cps, it is seen that any value of $b \geq 2$ would be acceptable. However, there are strong practical grounds for using as high a frequency as possible, since low frequency drifts or trends are likely to occur and increase the variance at low frequencies. Hence $b = 2$ would be a reasonable choice on the basis of the estimated spectrum.

7.4.3 Frequency response studies

These applications of spectral analysis are centered on the relation (6.2.15) connecting the spectra of the input $Z(t)$ and the output $X(t)$ to a linear system, that is,

$$\Gamma_{xx}(f) = \Gamma_{zz}(f) G^2(f). \tag{7.4.4}$$

It is possible to distinguish between two types of situations which occur in practice. Either the gain function $G(f)$ of the system is fixed and the only quantity which can be altered in (7.4.4) is the input spectrum $\Gamma_{zz}(f)$, or the input spectrum $\Gamma_{zz}(f)$ is fixed but the gain function can be changed.

Example 4. As an example of a system with fixed gain function consider the problem of runway roughness [20]. This presents a problem to aircraft designers which has increased in severity over the last few years since it is responsible for structural failure, reduction of fatigue life of an aircraft, difficulties in reading instruments and passenger discomfort.

The way in which runway roughness affects an aircraft is governed by the frequency response characteristic of the landing gear. For example, the landing gear of a typical commercial aircraft has a gain plot which has a predominant response in the frequency range between 1.5 and 2 cycles per second.

One way of measuring the roughness of a runway is to take readings of the runway elevation at intervals of the order of a foot along the runway. These readings may be used to characterize the runway roughness by computing an estimate of the spectrum of the runway. The spectrum $\Gamma_{zz}(f)$ would then be measured in terms of (elevation units)^2 per cycle per foot, that is, ft^2/cpf. From a knowledge of the gain function of the landing gear and the runway roughness spectrum, it is then possible to investigate how much stress the wings of the aircraft will be subjected to, and so on.

Suppose, for example, that the roughness spectrum is as shown in Figure 7.22(a) and has a definite peak due to a large concentration of roughness at a particular wave length. The squared gain of the landing gear response is shown in Figure 7.22(b) for two different speeds $V = 40$ mph, curve 3, and $V = 20$ mph, curve 2, at constant damping and at two different dampings at the same speed $V = 20$ mph, curve 1 referring to light damping and 2 to

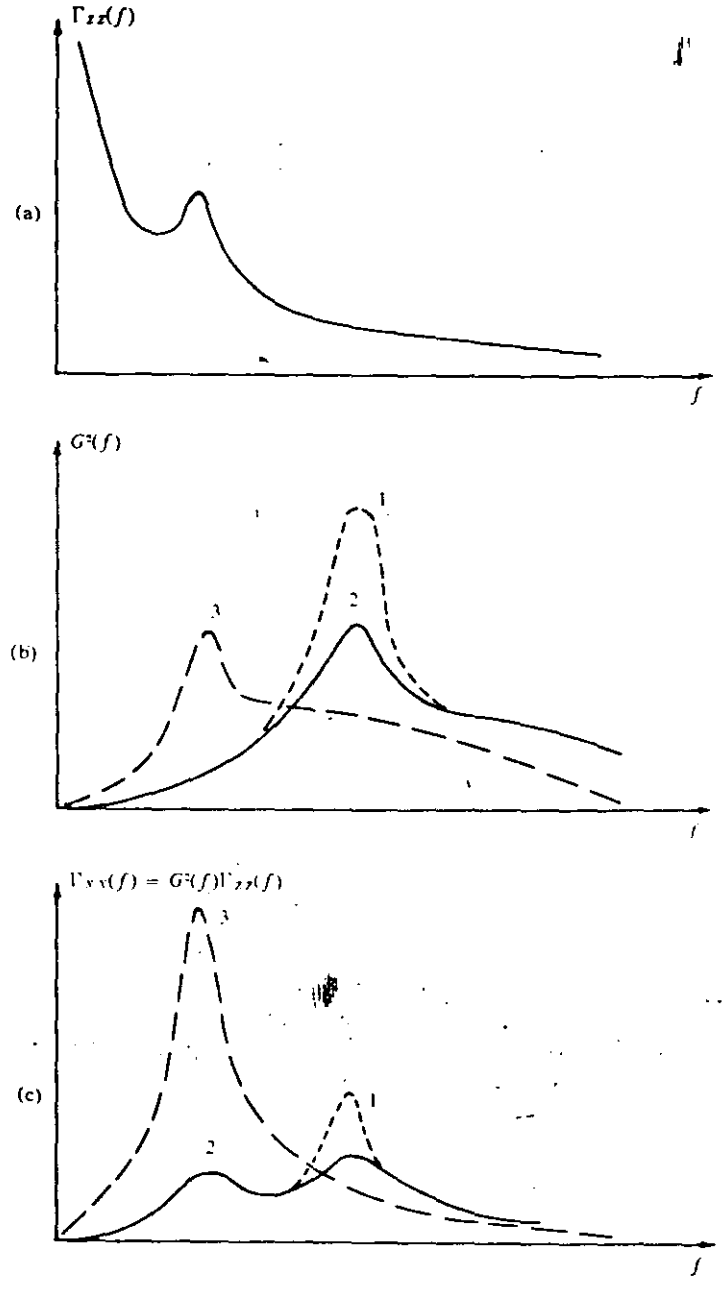


FIG. 7.22: Effects of input peaks, damping and velocity on aircraft response

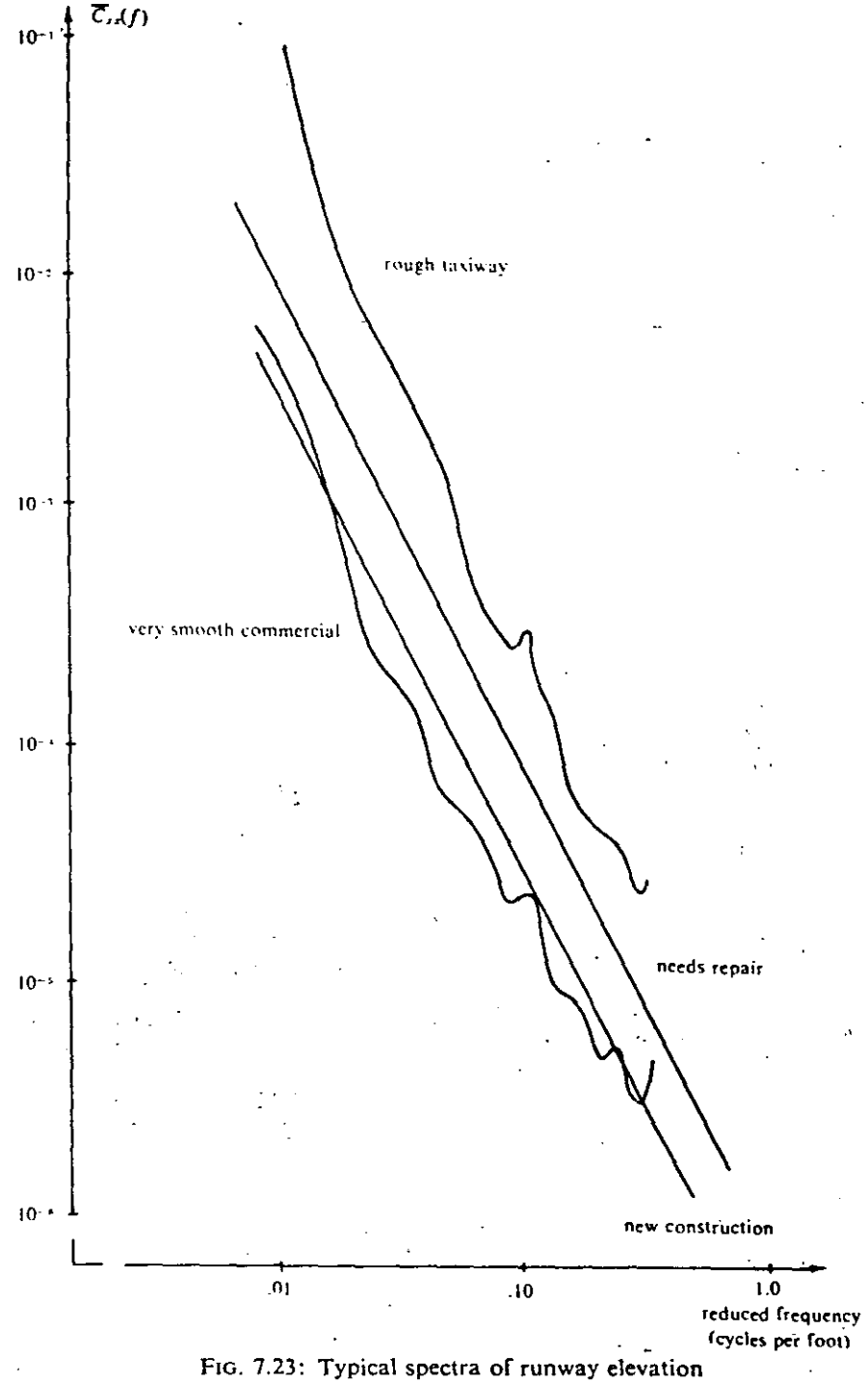


FIG. 7.23: Typical spectra of runway elevation

heavy damping. The predicted response obtained from (7.4.4) is shown in Figure 7.22(c) for these three cases.

Notice that at a given speed the effect of low damping is to increase the power under the spectral curve and hence to produce more damaging effects. Moreover, increasing the speed shifts the maximum in the gain to lower frequencies where the roughness spectrum increases and again the output power is increased. Finally, a shift in the gain plot as a result of an increased speed may interact with the roughness spectrum to produce a very sharp reinforcement in the output spectrum, as in curve 3 of Figure 7.22(c). (An example of this occurs when a car is driven over a washboard road at a speed which produces vehicle resonance.)

Some typical runway elevation spectra are shown in Figure 7.23. On the basis of the above calculations it is possible to lay down norms for the spectra of new runway constructions and for runways needing repair. These norms are indicated by the straight lines in Figure 7.23.

An example of a system with fixed input spectrum is the problem of designing the suspension of motorcycles and cars. Since road surfaces tend to differ widely in different countries, measurements of road surface spectra are beginning to have a considerable influence on the design of frequency response characteristics of cars and motorcycles, especially those designed for export. Another example of this type arises in the design of aircraft to minimize fatigue due to atmospheric turbulence. This is discussed below.

Example 5. Figure 7.24 shows the spectral estimate of the tail shear on an aircraft flown at low altitude in gusty conditions. The spectrum is characterized by a narrow peak at $f = 4.85$ cps, corresponding roughly to the natural

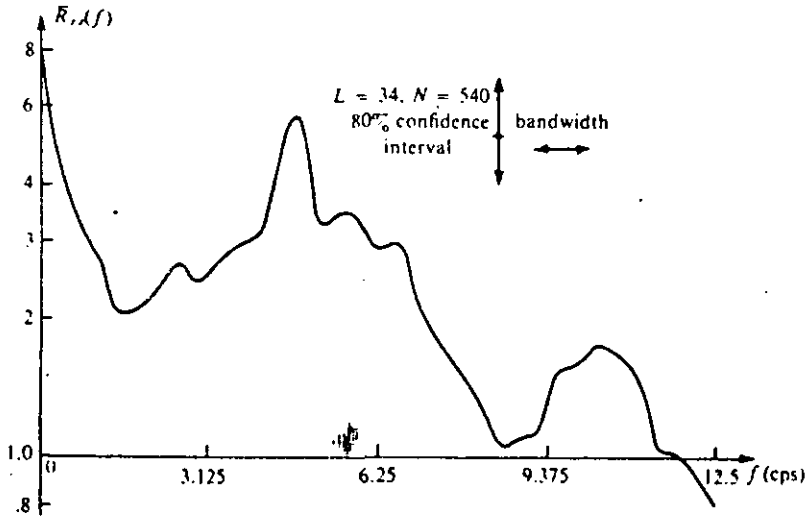


FIG. 7.24: Spectrum of tail shear

frequency of oscillation of the aircraft. There is also a wider peak at $f = 10.3$ cps corresponding to the natural frequency of oscillation of the tail, as would be expected since the measurements were made on the tail. These spectra are useful to design engineers in suggesting how various parts of the aircraft structure need to be modified to minimize the risk of structural damage due to the buffeting of turbulent air.

REFERENCES

[1] D. G. Watts, "A general theory of amplitude quantization with application to correlation determination." *Proc. Inst. Electr. Engrs.*, Part C **109**, 209 (May, 1962).

[2] R. B. Blackman and J. W. Tukey, *The Measurement of Power Spectra from the Point of View of Communications Engineering*. Dover, New York, 1958.

[3] U. Grenander and M. Rosenblatt, *Statistical Analysis of Stationary Time Series*. John Wiley, New York, 1957.

[4] A. A. Lomnicki and S. K. Zaremba, "On estimating the spectral density function of a stochastic process." *Jour. Royal Stat. Soc.* **B 19**, 13 (1957).

[5] E. Parzen, "Mathematical considerations in the estimation of spectra." *Technometrics* **3**, 167 (1961).

[6] G. M. Jenkins, "General considerations in the analysis of spectra." *Technometrics* **3**, 133 (1961).

[7] G. M. Jenkins, "An example of the estimation of a linear open loop transfer function." *Technometrics* **5**, 227 (1963).

[8] G. M. Jenkins, "A survey of spectral analysis." *Applied Statistics* **14**, 2 (1965).

[9] D. G. Watts, "Optimal windows for power spectra estimation." Mathematics Research Center Technical Summary Report 506 (Sept., 1964).

[10] H. E. Daniels, "The estimation of spectral densities." *Jour. Royal Stat. Soc.* **B 24**, 185 (1962).

[11] E. Slutsky, "The summation of random causes as the source of cyclic processes." *Econometria* **5**, 105 (1937).

[12] H. Robertson, "Approximate design of digital filters." *Technometrics* **7**, 387 (1965).

[13] J. M. Craddock, "An analysis of the slower temperature variations at Kew Observatory by means of mutually exclusive band-pass filters." *Jour. Royal Stat. Soc.* **A 120**, 387 (1957).

[14] M. D. Godfrey, "Frequency methods in economic analysis." Ph.D. Thesis. London University (1962).

[15] L. J. Tick, "Some time series techniques useful in life sciences." Proc. I.B.M. Scientific Computing Symposium. *Statistics*, 265 (1963).

[16] A. S. Alavi and G. M. Jenkins, "An example of digital filtering." *Applied Statistics* **14**, 70 (1965).

[17] H. A. Panofsky and I. Van der Hoven, "Spectra and cross-spectra of velocity components in the mesometeorological range." *Quart J. Roy. Meteorol. Soc.* **81**, 603 (1955).

[18] H. A. Panofsky and R. A. McCormick, "Properties of spectra of atmospheric turbulence at 100 meters." *Quart. J. Roy. Meteorol. Soc.* **80**, 546 (1954).

[19] G. E. P. Box and J. R. Chanmugam, "Adaptive optimization of continuous processes." *Industrial and Engineering Chemistry* (*Fundamentals*) **1**, 2 (1962).

[20] J. C. Houbolt, "Runway roughness studies in the aeronautical field." *Jour. Air Transport Div., Amer. Soc. Civ. Engr.* **87**, 11 (1961).

APPENDIX A7.1 FLOW CHART FOR AUTOSPECTRUM SUBROUTINE

The following is a flow chart for a computer subroutine which accepts autocovariance estimates COV(K,J,J) or DCOV(K,J,J), K = 0, MAXM, from program MULTICOR described in Appendix A5.3. Additional inputs are the sampling interval DELTA, the number of frequency points NF at which the smoothed autospectrum estimate is to be computed, and values of lag M ≤ MAXM for which smoothed autospectra are to be computed. In general NF = 2 or 3 times the largest M used. The spectral window used is the Tukey window (Table 6.5). Output consists of printout of the autocovariances (echo check), the smoothed spectra values for each truncation point M and a plot of the logarithm of the smoothed spectrum versus frequency for each truncation point M, overlayed on one graph.

Subroutine AUTOSPEC

1) *Input* parameters N, MAXM, DELTA, NF.

2) *Read* IDENT, COV(K); K = 0, MAXM.

3) *Read* M, *calculate* weights

$$W(K) = 0.5*(1. + COS(\pi K/M)),\ K = 1,\ M-1.$$

4) *Calculate* smoothed autospectral estimate

$$SPEC(I) = 2*DELTA*\left\{COV(0) + 2\sum_{K=1}^{M-1} COV(K)*W(K)*COS\frac{\pi KI}{NF}\right\};\quad I = 0,\ NF.$$

This transform can be performed very rapidly using either the fast Fourier transform [1], or the algorithm listed below in which the transform is obtained as the solution of a difference equation.

5) *Calculate* the logarithm of the spectrum LOGSPEC(I) = LOG10(SPEC(I)); I = 0, NF.

Care must be taken here to ensure that SPEC(I) is positive. If SPEC(I) is negative or zero, set LOGSPEC(I) = −100. for plotting purposes.

6) *Print* the smoothed spectral estimates SPEC(I) I = 0, NF, the bandwidth B = 4/(3M*DELTA) and the degrees of freedom D = (8*N)/(3*M) for the Tukey window and the appropriate values of N, M and DELTA.

7) *Plot* and *overlay* the logarithm of the smoothed spectrum versus frequency for all values of M used.

A search procedure is used to find the maximum value MLOG of LOGSPEC(I), which can then be used in plotting the graph. The procedure used by the authors is to choose the nearest decade value D above MLOG and plot the logspectrum over four decades. For example, suppose that the maximum spectral value was 2, so that MLOG = 0.303 and hence D = 1. Then the logspectrum values would be plotted over the range −3 to 1 corresponding to spectral values from 0.001 to 10. A range of four decades was considered adequate for most purposes, since if more than four decades are required it is probably better to filter the data in order to get a better spectral estimate at lower power levels. A value of LOGSPEC(I) = −100 will automatically be plotted along the lowest decade line.

Algorithm

To find SPEC(I),

Set $C = COS\frac{\pi I}{NF}$, V0 = 0., V1 = 0.

Do 1, K = M − 1, 1

V2 = 2.*C*V1 − V0 + W(K)*COV(K)

V0 = V1

1 V1 = V2.

SPEC(I) = 2.*DELTA*(COV(0) + 2.*(V1*C − V0)).

Example: Consider the example given in Section 7.1.1 for which M = 3, DELTA = 1.0, NF = 8 and COV(0) = 1, W(1)*COV(1) = 0.430, W(2)*COV(2) = 0.065. Then for I = 0, $C = COS\frac{\pi}{8}(0) = 1$, V0 = 0., V1 = 0., and going through the do-loop,

K = 2, V2 = 2(1)(0) − 0 + 0.065 = 0.065

V0 = 0

V1 = 0.065;

K = 1, V2 = 2(1)(0.065) − 0 + 0.430 = 0.560

V0 = 0.065

V1 = 0.560.

Then SPEC(0) = 2(1){1 + 2((0.560)(1) − 0.065)} = 3.980, which agrees with the value obtained in Table 7.1.

For I = 1,

C = COS(π/8) = 0.924, V0 = 0, V1 = 0. Going through the do-loop,

K = 2, V2 = 2(0.924)(0) − 0 + 0.065 = 0.065

V0 = 0

V1 = 0.065;

K = 1, V2 = 2(0.924)(0.065) − 0 + 0.430 = 0.550

V0 = 0.065

V1 = 0.550.

Then SPEC(1) = 2(1)(1 + 2((0.550)(0.924) − 0.065)) = 3.772.

This algorithm, while not as fast as the fast Fourier transform, nevertheless enjoys the advantages of relatively high speed, high accuracy and the need to compute only one cosine function for each frequency point.

APPENDIX REFERENCE

[1] J. W. Cooley and J. W. Tukey, "An algorithm for the machine calculation of complex Fourier series." *Mathematics of Computation*, **19**, 90, 297 (1965).

APPENDIX A7.2 VARIANCE OF SLOPE ESTIMATORS

Discrete Time. The estimate of the slope θ_1 in the model

$$Y_t = \theta_1 x_t + Z_t, \tag{A7.2.1}$$

introduced in Section 7.4.2, is of the form

$$\hat{\theta}_1 = \sum_{t=1}^{N} w_t y_t, \tag{A7.2.2}$$

with $w_t = (2/Na)\cos(\pi t/b)$. Proceeding as in the derivation of (5.2.9), the variance of the corresponding estimator is

$$\text{Var}[\hat{\Theta}_1] = \sum_{t=1}^{N}\sum_{r=1}^{N} w_t w_r \gamma_{YY}(t-r). \tag{A7.2.3}$$

Now if the observations are made at unit time intervals,

$$\gamma_{YY}(k) = \int_{-1/2}^{1/2} \Gamma_{YY}(f)\, e^{i2\pi fk}\, df. \tag{A7.2.4}$$

Substituting (A7.2.4) in (A7.2.3),

$$\text{Var}[\hat{\Theta}_1] = \int_{-1/2}^{1/2} |W(f)|^2 \Gamma_{YY}(f)\, df, \tag{A7.2.5}$$

where

$$W(f) = \sum_{t=1}^{N} w_t\, e^{i2\pi ft}. \tag{A7.2.6}$$

Now, suppose that $w_t = (2/Na)\cos(\pi t/b)$ as in (7.4.1). Then

$$W(f) = \frac{1}{Na}\left\{\frac{1-\exp[j(N+1)2\pi(f-1/2b)]}{1-\exp[j2\pi(f-1/2b)]} + \frac{1-\exp[j(N+1)2\pi(f+1/2b)]}{1-\exp[j2\pi(f+1/2b)]}\right\}.$$

Substituting the modulus squared in (A7.2.5) gives

$$\text{Var}[\hat{\Theta}_1] = \frac{1}{Na^2}\int_{-1/2}^{1/2}\left\{\frac{\sin^2(N+1)\pi(f-1/2b)}{N\sin^2\pi(f-1/2b)} + \frac{\sin^2(N+1)\pi(f+1/2b)}{N\sin^2\pi(f+1/2b)}\right\}\Gamma_{YY}(f)\, df,$$

plus cross-product terms.

Making use of the fact that the quantity in braces tends to

$$\delta(f-1/2b) + \delta(f+1/2b)$$

as N tends to infinity, while the cross-product terms are of order $1/N^2$, gives

$$\lim_{N\to\infty} N\,\text{Var}[\hat{\Theta}_1] = \frac{4}{a^2}\,\Gamma_{YY}\left(\frac{1}{2b}\right). \tag{A7.2.7}$$

For finite N, it follows that

$$\text{Var}[\hat{\Theta}_1] \approx \frac{4}{Na^2}\,\Gamma_{YY}\left(\frac{1}{2b}\right). \tag{A7.2.8}$$

Since the model (A7.2.1) implies that $\Gamma_{YY}(1/2b) = \Gamma_{ZZ}(1/2b)$, (A7.2.8) is equivalent to (7.4.2). Thus the variance is a minimum if the perturbation frequency $1/2b$ is equal to the frequency for which the spectrum of the noise is a minimum.

Continuous Time. In discrete time, the dynamics of the process were ignored. If a continuous sinusoidal perturbation $x(t) = a\cos 2\pi f_0 t$ is applied to an input variable, the model (A7.2.1) needs to be modified to

$$Y(t) = \theta_1 a G(f_0)\cos 2\pi f_0 t + Z(t),$$

where $G(f_0)$ is the gain at frequency f_0. The analysis is then similar to that for the discrete case. The final result is

$$\text{Var}[\hat{\Theta}_1] \approx \frac{4}{Ta^2}\,\frac{\Gamma_{YY}(f_0)}{G^2(f_0)}.$$

Hence the variance is a minimum if the perturbation frequency f_0 corresponds to the maximum of the signal to noise ratio $G^2(f_0)/\Gamma_{YY}(f_0)$.

APPENDIX A7.3 THE FAST FOURIER TRANSFORM

A recent innovation in spectral analysis is the fast Fourier transform (FFT). This is an algorithm for computing discrete Fourier transforms much more quickly than the direct method given in Section 2.1.2, but at the same time retaining accuracy. Thus, using the direct approach, the discrete Fourier transform of a series of N terms would require approximately N^2 operations, whereas the FFT requires only $2N\log_2 N$ operations. Savings in computer time can be very large if one is interested in the Fourier analysis of long series. For example, computation of the Fourier coefficients for a series of $N = 8192$ terms [1] required about 5 seconds on an IBM 8094 computer, as compared with almost 30 minutes using the direct approach.

The relevance of the FFT to spectral analysis is that it is now faster to compute the sample spectrum directly using a FFT, then smooth the ple spectrum

rather than compute the autocorrelation function, smooth with a lag window and finally transform. Despite this computing advantage we do not believe that the case for using the FFT in spectral analysis is as strong as in Fourier analysis for the following reasons:

(1) In the authors' experience, the fast computers which are now available are more than adequate for purposes of spectral analysis. Our present computing facilities are greatly in excess of our ability to make sense of practical data.

(2) We regard the autocorrelation function as an invaluable intermediate stage in spectral analysis. Thus the autocorrelation function of the original and an appropriately differenced series should be plotted to decide

(a) whether differencing is necessary or not

(b) the choice of a suitable range of truncation points

(c) the degree of alignment necessary when cross correlating series.

Description of the Fast Fourier Transform. A complete description of the FFT is given in [1] and its history of discovery and repeated discovery in [2]. These papers are part of a special journal issue [3], which also includes a paper on some additional uses of the FFT [4, 5]. The treatment here follows [1].

Suppose it is required to find the Fourier transform X_m, $m = 0, 1, \ldots, N-1$, of the series x_t, $t = 1, 2, \ldots, N$, where N is even. One approach [5] is to partition the series x_t into two half-series y_t and z_t, where

$$y_t = x_{2t-1};$$
$$z_t = x_{2t}, \qquad t = 1, 2, \ldots, \frac{N}{2}. \tag{A7.3.1}$$

The series y_t, z_t each consist of $N/2$ values and hence have Fourier transforms

$$Y_m^{(N/2)} = \frac{2}{N}\sum_{t=1}^{N/2} y_t\, e^{-i(4\pi tm/N)},$$
$$Z_m^{(N/2)} = \frac{2}{N}\sum_{t=1}^{N/2} z_t\, e^{-i(4\pi tm/N)}, \tag{A7.3.2}$$

where the superscript on the transform denotes the number of terms in the series and the transform. But $X_m^{(N)}$ and $Y_m^{(N/2)}$, $Z_m^{(N/2)}$ are related, since

$$\begin{aligned} X_m^{(N)} &= \frac{1}{N}\sum_{t=1}^{N} x_t\, e^{-i(2\pi tm/N)} \\ &= \frac{1}{N}\sum_{t=1}^{N/2}\left\{y_t\, e^{-i(2\pi m/N)(2t-1)} + z_t\, e^{-i(2\pi m/N)(2t)}\right\} \\ &= e^{i(2\pi m/N)}\frac{1}{N}\sum_{t=1}^{N/2} y_t\, e^{-i(4\pi mt/N)} + \frac{1}{N}\sum_{t=1}^{N/2} z_t\, e^{-i(4\pi mt/N)} \\ &= \frac{e^{i(2\pi m/N)}}{2} Y_m^{(N/2)} + \frac{1}{2} Z_m^{(N/2)}, \qquad 0 \le m \le \frac{N}{2} - 1. \end{aligned} \tag{A7.3.3}$$

Also,

$$Y_{m+(N/2)}^{(N/2)} = Y_m^{(N/2)},$$
$$Z_{m+(N/2)}^{(N/2)} = Z_m^{(N/2)}, \qquad 0 \le m \le \frac{N}{2} - 1,$$

so that

$$\begin{aligned} X_{m+(N/2)}^{(N)} &= \frac{e^{i(2\pi/N)(m+N/2)}}{2} Y_m^{(N/2)} + \frac{1}{2} Z_m^{(N/2)} \\ &= -\frac{e^{i(2\pi m/N)}}{2} Y_m^{(N/2)} + \frac{1}{2} Z_m^{(N/2)}, \qquad 0 \le m \le \frac{N}{2} - 1. \end{aligned} \tag{A7.3.4}$$

Hence, summarizing (A7.3.3) and (A7.3.4),

$$X_m^{(N)} = \frac{e^{i(2\pi m/N)}}{2} Y_m^{(N/2)} + \frac{1}{2} Z_m^{(N/2)}$$
$$X_{m+(N/2)}^{(N)} = -\frac{e^{i(2\pi m/N)}}{2} Y_m^{(N/2)} + \frac{1}{2} Z_m^{(N/2)}, \qquad 0 \le m \le \frac{N}{2} - 1, \tag{A7.3.5}$$

and it is seen that the Fourier transform for the series x_t is easily obtained from the Fourier series of the half-series y_t and z_t. Likewise, if $N/2$ is even, the series y_t and z_t may be partitioned into two series, y_t', z_t' and y_t'', z_t'' respectively, and an appropriate version of (A7.3.5) may be used to construct the transforms $Y_m^{(N/2)}$ and $Z_m^{(N/2)}$ from the transforms of the series of length $N/4$.

For series of length $N = 2^k$, the procedure is followed until partitions of only one term are obtained, for which the Fourier transform equals the term itself. Otherwise the procedure is followed until the reduced series can be simply transformed or until a new factor of N, say $n = 3$, is encountered. The procedure is the same as above then, except that the remaining partitions are partitioned by threes. Details are given in [1]. An example follows.

An example. Consider the ionosphere data of Chapter 2, with $n = 12 = 2^2 3$. The data are as follows:

t	1	2	3	4	5	6	7	8	9	10	11	12
x_t	−6	−20	−28	−8	−1	7	−20	−6	−7	14	19	12

Partitioning into two gives

t	1	2	3	4	5	6
y_t	−6	−28	−1	−20	−7	19
z_t	−20	−8	7	−6	14	12

Partitioning y_t and z_t into two gives

t	1	2	3
y_t'	−6	−1	−7
z_t'	−28	−20	19
y_t''	−20	7	14
z_t''	−8	−6	12

The Fourier transforms of y_t', z_t', y_t'' and z_t'' are easily calculated and each consists of three terms, as shown below.

Transform \ Harmonic m	0	1	2
$Y_m'^{(3)}$	−4.6667	−1.1667 + *j*1.4433	−1.1667 − *j*1.4433
$Z_m'^{(3)}$	−9.6667	14.3333 + *j*2.3093	14.3333 − *j*2.3093
$Y_m''^{(3)}$	0.3333	6.8333 + *j*7.7940	6.8333 − *j*7.7940
$Z_m''^{(3)}$	−0.6667	6.3333 + *j*0.5773	6.3333 − *j*0.5773

The Fourier transforms $Y_m^{(6)}, Z_m^{(6)}, (0 \le m \le 5)$, are then calculated using (A7.3.5). For example,

$$Y_0^{(6)} = \frac{1}{2} Y_0'^{(3)} + \frac{1}{2} Z_0''^{(3)} = -7.1666$$

$$Y_1^{(6)} = \frac{1}{2}\left(\frac{1}{2} + j\frac{\sqrt{3}}{2}\right) Y_1'^{(3)} + \frac{1}{2} Z_1''^{(3)} = 6.2500 + j1.0103$$

$$Y_2^{(6)} = \frac{1}{2}\left(\frac{1}{2} + j\frac{\sqrt{3}}{2}\right)^2 Y_2'^{(3)} + \frac{1}{2} Z_2''^{(3)} = 8.0833 - j1.2990$$

$$Y_3^{(6)} = -\frac{1}{2} Y_0'^{(3)} + \frac{1}{2} Z_0''^{(3)} = -2.5000$$

$$Y_4^{(6)} = -\frac{1}{2}\left(\frac{1}{2} + j\frac{\sqrt{3}}{2}\right) Y_1'^{(3)} + \frac{1}{2} Z_1''^{(3)} = 8.0833 + j1.2990$$

$$Y_5^{(6)} = -\frac{1}{2}\left(\frac{1}{2} + j\frac{\sqrt{3}}{2}\right)^2 Y_2'^{(3)} + \frac{1}{2} Z_2''^{(3)} = 6.2500 - j1.0103.$$

The transform $Z_m^{(6)}$ is obtained similarly, to give the following:

Transform \ Harmonic m	0	1	2	3	4	5
$Y_m^{(6)}$	−7.1667	6.2500 + *j*1.0103	8.0833 − *j*1.2990	−2.5000	8.0883 + *j*1.2990	6.2500 − *j*1.0103
$Z_m^{(6)}$	−0.1666	1.5000 + *j*5.1960	−4.8333 + *j*4.6187	−0.5000	4.8333 − *j*4.6187	1.5000 − *j*5.1960

These values are combined, using (A7.3.5), to give the final transform $X_m^{(12)}$. For example,

$$X_0^{(12)} = \frac{1}{2} Y_0^{(6)} + \frac{1}{2} Z_0^{(6)} = -3.667$$

$$X_1^{(12)} = \frac{1}{2}\left(\frac{\sqrt{3}}{2} + j\frac{1}{2}\right) Y_1^{(6)} + \frac{1}{2} Z_1^{(6)} = 3.204 + j4.598.$$

The complete transform is as follows:

m	0	1	2	3	4	5
$X_m^{(12)}$	−3.667	3.204 + *j*4.598	5.000 + *j*5.485	−0.250 − *j*1.250	−0.167 + *j*0.866	−1.704 − *j*0.598

m	6	7	8	9	10	11
$X_m^{(12)}$	−3.500	−1.704 + *j*0.598	−0.167 − *j*0.866	−0.250 + *j*1.250	5.000 − *j*5.485	3.204 − *j*4.598

Except for the phase shift due to the shift in origin, these are the same as the results in Table 2.2, obtained using the difference equation method of Appendix A7.1 or the straightforward method of Chapter 2.

REFERENCES

[1] Cochran, W. T. et al., "What is the Fast Fourier Transform?" IEEE *Transactions on Audio and Electroacoustics*, Vol. AU-15, No. 2, p. 45, June 1967.

[2] Cooley, J. W., P. A. W. Lewis and P. D. Welch, "Historical Notes on the Fast Fourier Transform," IEEE *Transactions on Audio and Electroacoustics*, Vol. AU-15, No. 2, p. 76, June 1967.

[3] IEEE *Transactions on Audio and Electroacoustics*, Vol. AU-15, No. 2, June 1967.

[4] Cooley, J. W., P. A. W. Lewis and P. D. Welch, "Application of the Fast Fourier Transform to Computation of Fourier Integrals, Fourier Series, and Convolution Integrals," IEEE *Transactions on Audio and Electroacoustics*, Vol. AU-15, No. 2, p. 79, June 1967.

[5] Helms, H. D., "Fast Fourier Transform Method of Computing Difference Equations and Simulating Filters," IEEE *Transactions on Audio and Electroacoustics*, Vol. AU-15, No. 2, p. 85, June 1967.

APPENDIX A7.4: DATA AND AUTOCORRELATIONS FOR A SIMULATED SECOND-ORDER AR PROCESS

TABLE A7.1: 400 values of the second-order ar process $X_t = X_{t-1} - 0.5X_{t-2} + Z_t$

t	x_t									
1– 10	−0.88	−0.12	−0.89	−1.38	−0.07	1.03	2.14	0.35	−1.10	−1.78
11– 20	−2.76	−1.77	0.98	1.00	−0.70	−1.01	−1.30	−0.85	−0.46	1.63
21– 30	0.06	−0.17	−1.01	−1.04	−0.66	−1.12	−0.51	−0.71	−0.20	−0.13
31– 40	0.14	1.59	−0.76	−1.08	−1.77	−1.20	0.45	−0.07	−0.63	−0.35
41– 50	−0.87	−0.62	0.28	1.90	2.14	1.05	0.31	1.07	2.67	2.44
51– 60	1.31	1.10	1.94	0.33	1.82	1.15	0.61	−1.08	−1.62	−0.39
61– 70	0.19	−1.59	−2.25	0.29	1.73	2.30	0.80	−0.40	0.30	−0.50
71– 80	−2.11	−2.43	0.72	3.09	4.96	1.81	−0.46	−0.33	0.04	0.82
81– 90	−1.63	−2.29	−0.77	1.91	1.92	0.85	−0.65	0.35	0.78	1.62
91–100	3.24	1.86	0.76	2.24	0.76	−0.15	0.18	0.60	0.92	−0.70
101–110	−0.03	1.07	0.28	−1.38	−0.63	−1.48	0.19	−1.14	0.31	0.39
111–120	−0.17	0.70	2.14	1.24	0.42	0.61	−0.76	−1.75	−0.37	1.21
121–130	1.40	2.46	1.74	0.78	0.90	1.11	2.20	0.52	−0.22	1.12
131–140	1.02	1.10	1.72	1.80	−0.46	−1.27	0.39	0.93	0.55	−0.45
141–150	−0.87	−0.90	0.64	2.29	2.75	1.43	0.47	1.80	0.46	0.32
151–160	−0.81	−1.81	−2.07	0.96	1.20	0.77	−0.98	−1.46	−1.30	−2.29
161–170	−1.81	−1.61	−1.01	−1.36	−1.78	0.04	1.44	2.58	0.54	0.27
171–180	−0.75	−0.70	0.45	−0.13	−1.03	−1.19	−0.31	1.77	1.89	0.88
181–190	0.58	0.70	−0.32	−1.62	1.08	1.25	0.19	−0.93	−0.61	0.83
191–200	0.46	1.12	0.11	−1.11	−0.85	−1.86	−0.74	−1.04	−0.42	0.16
201–210	0.55	−0.37	−0.62	−1.23	−0.76	−0.79	−1.99	−1.56	−0.36	1.00
211–220	0.02	−0.30	−0.23	−0.63	−1.61	−1.66	−0.80	−1.71	−0.87	−0.74
221–230	1.55	1.39	1.51	2.39	1.68	−0.04	−1.24	−2.24	−1.31	−0.10
231–240	0.46	1.06	1.37	1.67	0.29	−0.31	−2.08	−2.67	−1.50	−1.71
241–250	−0.70	−1.25	−0.25	0.14	1.43	0.47	−1.16	−3.68	−3.41	−1.43
251–260	1.06	2.86	0.72	−1.79	−2.26	−1.87	−1.53	−0.25	1.40	3.37
261–270	0.85	−0.36	0.25	1.57	−0.08	0.78	−0.56	−1.22	0.07	−0.33
271–280	−0.15	1.56	2.23	2.01	0.42	−0.75	−0.47	1.55	3.60	2.07
281–290	1.32	0.06	0.87	0.51	−0.25	0.12	1.54	1.37	1.97	0.81
291–300	−0.67	−2.41	−1.82	−0.45	0.31	0.12	−1.01	−1.12	−1.69	−1.52
301–310	−0.82	−0.81	−0.33	−0.65	−1.86	−0.94	0.50	1.05	1.40	1.52
311–320	0.20	0.64	1.95	1.55	1.74	−0.22	−2.14	−2.33	−1.01	0.42
321–330	2.54	0.86	0.10	−0.04	−1.18	−0.40	−0.53	0.70	−0.14	−0.20
331–340	0.47	1.07	0.85	−0.35	−0.69	−0.63	−2.08	−1.56	−1.00	0.55
341–350	2.08	1.74	−0.34	−1.85	−1.29	1.74	2.58	1.64	1.85	−0.01
351–360	−0.16	−0.29	−0.66	−3.41	−2.33	−2.57	−1.78	−1.31	−2.69	−1.77
361–370	−0.57	1.58	1.78	1.09	−0.54	0.29	−0.26	0.01	1.05	0.94
371–380	−0.91	−2.09	−2.01	−1.12	−0.02	0.98	0.50	2.12	1.68	2.28
381–390	2.59	3.04	1.16	0.50	0.56	0.45	0.35	0.10	2.16	2.60
391–400	1.40	0.62	0.36	−0.09	1.93	1.80	1.13	−1.34	−1.94	−0.89

TABLE A7.2: Autocorrelation estimates for a second-order ar process, $X_t = X_{t-1} - 0.5X_{t-2} + Z_t$, $N = 50$

k	$r_{xx}(k)$	k	$r_{xx}(k)$	k	$r_{xx}(k)$	k	$r_{xx}(k)$
1	0.574	11	−0.068	21	−0.217	31	0.049
2	0.086	12	0.124	22	−0.124	32	−0.051
3	−0.166	13	0.109	23	0.035	33	−0.200
4	−0.130	14	0.000	24	0.165	34	−0.213
5	0.096	15	−0.063	25	0.137	35	−0.097
6	0.225	16	−0.043	26	−0.045	36	0.004
7	0.244	17	0.062	27	−0.136	37	0.007
8	0.032	18	0.103	28	−0.156	38	−0.060
9	−0.180	19	−0.047	29	−0.027	39	−0.096
10	−0.199	20	−0.152	30	0.109	40	−0.086

TABLE A7.3: Autocorrelation estimates for a second-order ar process, $X_t = X_{t-1} - 0.5X_{t-2} + Z_t$, $N = 400$

k	$r_{xx}(k)$	k	$r_{xx}(k)$	k	$r_{xx}(k)$	k	$r_{xx}(k)$
1	0.645	13	0.012	25	−0.075	37	0.019
2	0.196	14	−0.021	26	−0.116	38	−0.020
3	−0.080	15	−0.074	27	−0.078	39	−0.045
4	−0.099	16	−0.119	28	−0.038	40	−0.035
5	−0.009	17	−0.070	29	−0.013	41	0.021
6	0.057	18	0.008	30	−0.036	42	0.083
7	0.066	19	0.064	31	−0.044	43	0.081
8	0.040	20	0.069	32	−0.037	44	0.017
9	0.030	21	0.017	33	−0.015	45	−0.004
10	0.052	22	−0.026	34	0.047	46	0.042
11	0.088	23	−0.032	35	0.080	47	0.081
12	0.051	24	−0.044	36	0.066	48	0.069

8

The Cross Correlation Function and Cross Spectrum

In this chapter the concepts introduced in Chapters 5 and 6 are extended to deal with pairs of time series and stochastic processes. The first generalization, given in Section 8.1, is the cross correlation function of a stationary bivariate stochastic process, which measures the correlation between the two processes at different lags. The second generalization is the bivariate linear process which can be generated by linear operations on two white noise sources. Important special cases of the bivariate linear process are the bivariate autoregressive and moving average processes.

The estimation of the cross correlation function is discussed in Section 8.2, where it is shown that unless a filtering operation is applied to both series to convert them to white noise, spurious cross correlations may arise. The third generalization, the cross spectrum of a stationary bivariate process, is given in Section 8.3. The cross spectrum contains two different types of information about the dependence between the two processes. The first is the coherency spectrum, which effectively measures the correlation between the two processes at each frequency, and the second is the phase spectrum, which measures the phase difference between the two processes at each frequency. In Section 8.4 the information contained in the coherency and phase spectrum is illustrated by simple examples.

8.1 THE CROSS CORRELATION FUNCTION

8.1.1 Introduction

This chapter is concerned with describing *pairs* of time series or *bivariate* time series. The methods used are extensions of those developed in Chapters 5 and 6, and hence all the general considerations relating to time series discussed in Section 5.1 apply. It was briefly mentioned in Section 5.1 under the heading of multivariate time series that the individual series in a multivariate time

series may not arise on equal footing. For example, consider the system shown in Figure 8.1, which has two *inputs* $x_1(t)$, $x_2(t)$, and two *outputs* $x_3(t)$, $x_4(t)$.

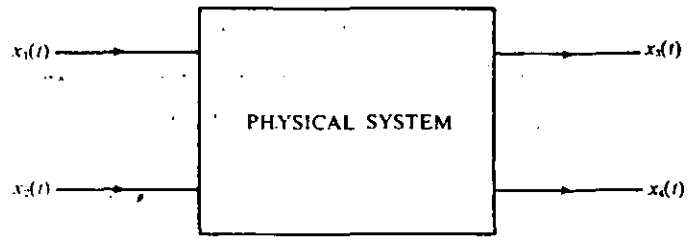


FIG. 8.1: A physical system with two inputs and two outputs

Two situations may be distinguished. The first occurs when the two series arise on a *similar footing*, for example, the two inputs in Figure 8.1. Thus $x_1(t)$, $x_2(t)$ may be two correlated control variables whose interactions are to be studied. An example of a pair of time series which falls into this category is given in Figure 8.2, which shows records of the in-phase and out-of-phase current inputs to a turbo-alternator.

The second situation arises when a pair of series are *causally related*, for example, the output $x_3(t)$ of Figure 8.1 depends on the input $x_1(t)$. In this situation it is usually required to estimate the properties of the system so that the output can be predicted from the input. An example of a pair of time series of this type is given in Figure 8.3, which shows the input gas rate $x_1(t)$ and the concentration $x_2(t)$ of carbon dioxide in the output from a gas furnace. It is seen that the input $x_1(t)$ lags behind the output $x_2(t)$ due to the delay in the transfer of the gases through the reactor.

In the first situation one is interested in describing the interaction or correlation between the two time series so that this interaction can be allowed for in any further studies. For example, if an output $x_3(t)$ is to be controlled by means of two control variables which are cross correlated, then allowance must be made for the cross correlation if the desired effect is to be produced at the output. On the other hand, in the second situation one is interested in relating $x_2(t)$ to $x_1(t)$, for example, by means of a relationship such as

$$x_2(t) = \int_0^\infty h(u)x_1(t-u)\,du,$$

so that $x_2(t)$ can be predicted from $x_1(t)$, as discussed briefly in Section 5.1.5. This chapter and the next will be devoted to series which arise on a similar footing, while series which are causally related will be discussed in Chapter 10.

8.1.2 Cross covariance and cross correlation functions

As in the univariate case discussed in Chapter 5, a useful way of describing pairs of stochastic processes is by their lower-order moments. Thus the

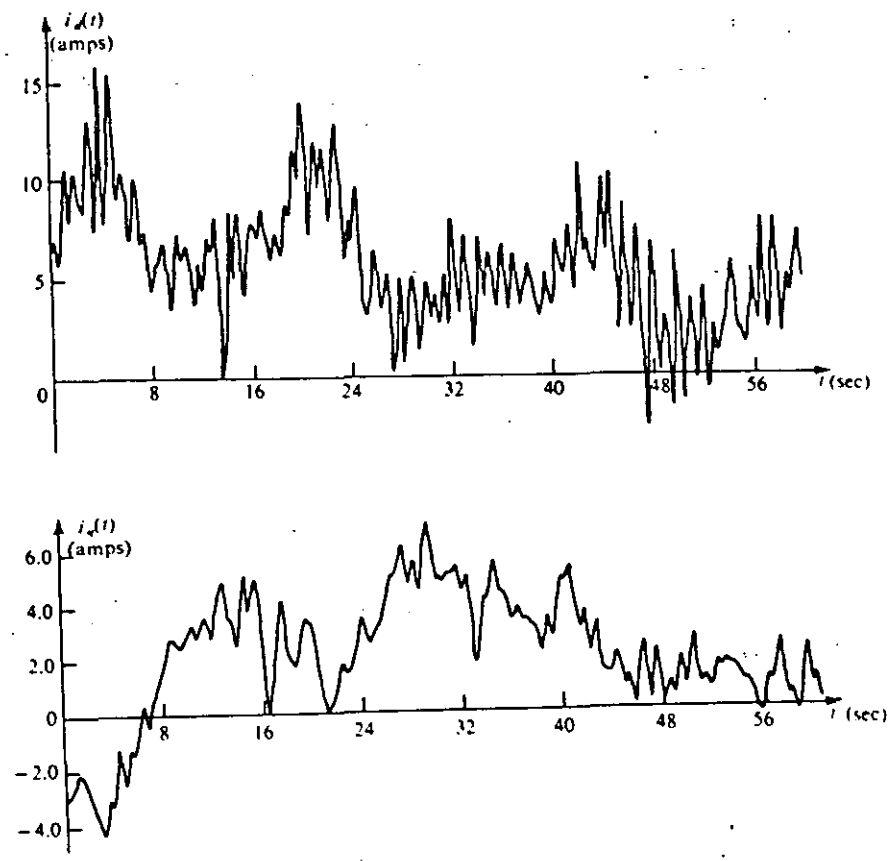


FIG. 8.2: In-phase and out-of-phase current inputs to a turbo-alternator

observed bivariate time series $\{x_1(t), x_2(t)\}$ is regarded as a realization of the bivariate stochastic process $\{X_1(t), X_2(t)\}$. The four rv's $X_1(t)$, $X_2(t)$, $X_1(t+u)$ and $X_2(t+u)$ at times t and $t+u$ will have a joint pdf which can be partially described by its first- and second-order moments. If it is assumed that the processes are *stationary*, then these moments will be functions of time differences u and not of the absolute time t. Thus the first moments are

$$E[X_i(t)] = \mu_i, \quad i = 1, 2,$$

and are independent of time t. The second moments of the joint pdf are the *autocovariance functions*

$$\gamma_{x_1x_1}(u) = E[(X_1(t) - \mu_1)(X_1(t+u) - \mu_1)],$$
$$\gamma_{x_2x_2}(u) = E[(X_2(t) - \mu_2)(X_2(t+u) - \mu_2)]$$

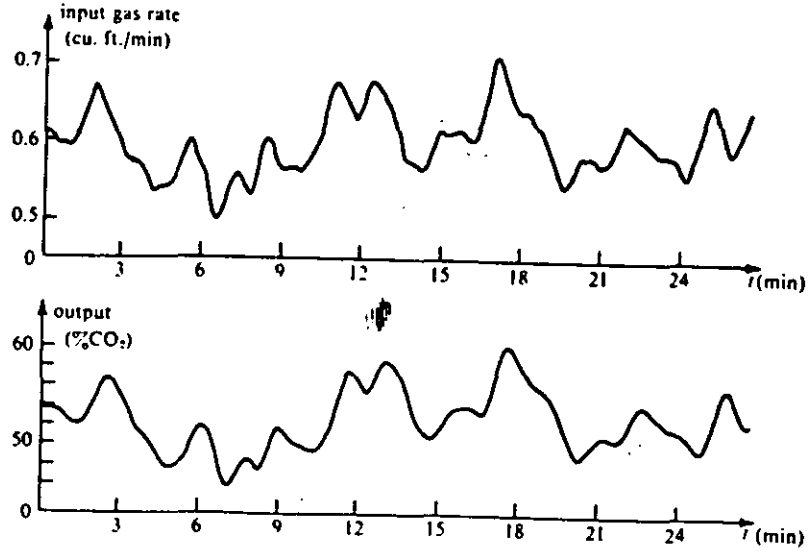


FIG. 8.3: Input and output signals to a chemical reactor

and the *cross covariance functions*

$$\gamma_{X_1X_2}(u) = E[(X_1(t) - \mu_1)(X_2(t+u) - \mu_2)],$$
$$\gamma_{X_2X_1}(u) = E[(X_2(t) - \mu_2)(X_1(t+u) - \mu_1)]. \tag{8.1.1}$$

The function $\gamma_{X_1X_2}(u)$ is called the *cross covariance function* (ccvf) of lag u with $X_2(t)$ leading $X_1(t)$ and $\gamma_{X_2X_1}(u)$ the ccvf of lag u with $X_1(t)$ leading $X_2(t)$. For conciseness when no confusion is likely, the notation $\gamma_{11}(u)$, $\gamma_{22}(u)$, $\gamma_{12}(u)$ and $\gamma_{21}(u)$ will be used to denote the acvf's and ccvf's $\gamma_{X_1X_1}(u)$, $\gamma_{X_2X_2}(u)$, $\gamma_{X_1X_2}(u)$ and $\gamma_{X_2X_1}(u)$ respectively.

Properties of covariance functions. The properties of the acvf's of a real bivariate process are the same as those for the acvf of a univariate process, that is,

$$\left.\begin{aligned}\gamma_{ii}(0) &= \text{Var}\,[X_i(t)] = \sigma_{X_i}^2\\ \gamma_{ii}(u) &= \gamma_{ii}(-u)\end{aligned}\right\} \quad i = 1, 2. \tag{8.1.2}$$

Hence the acvf's are *even* functions of the lag u.

The ccvf $\gamma_{12}(u)$ of two real processes has the property

$$\gamma_{12}(u) = \gamma_{21}(-u), \tag{8.1.3}$$

since

$$\begin{aligned}\gamma_{12}(u) &= E[(X_1(t) - \mu_1)(X_2(t+u) - \mu_2)]\\ &= E[(X_1(t-u) - \mu_1)(X_2(t) - \mu_2)]\\ &= E[(X_2(t) - \mu_2)(X_1(t-u) - \mu_1)]\\ &= \gamma_{21}(-u).\end{aligned}$$

Similarly, $\gamma_{21}(u) = \gamma_{12}(-u)$. Hence the covariance between the two stochastic processes can be described by means of the single ccvf $\gamma_{12}(u)$ where $-\infty \leqslant u \leqslant \infty$. Note that although the acvf is an even function of lag, the ccvf will *not* be an even function in general.

The cross correlation function. In general it may be necessary to study the interactions between two processes with possibly different scales of measurement or different variances. In this situation it is necessary to define the *cross correlation function* (ccf)

$$\rho_{12}(u) = \frac{\gamma_{12}(u)}{\sqrt{\gamma_{11}(0)\gamma_{22}(0)}} = \frac{\gamma_{12}(u)}{\sigma_1\sigma_2}. \tag{8.1.4}$$

The first property of the ccf is that

$$|\rho_{12}(u)| \leqslant 1,$$

which follows from the fact that the rv

$$Y(t) = \lambda_1 X_1(t) + \lambda_2 X_2(t+u)$$

has positive variance. The second property is that

$$\rho_{12}(u) = \rho_{21}(-u),$$

which follows from (8.1.3).

Like the ccvf, the ccf is in general *not* an even function of lag. For example, Figure 8.4 shows the sample ccf of the gas furnace data of Figure 8.3. This ccf has a large peak at $u = 5$ and is clearly not symmetrical about $u = 0$.

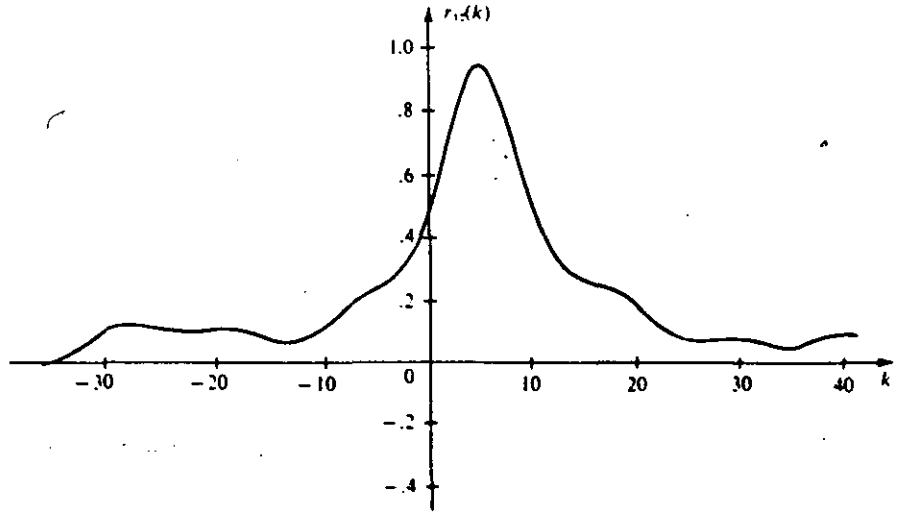


FIG. 8.4: Sample cross correlation function for gas furn[illegible] ata

Note also that most of the cross correlations are positive. This is due to the fact that a positive change in the input gas rate results in a positive change in the output concentration.

The most trivial case of cross correlation between two stochastic processes occurs when the cross correlation function is identically zero for all lags. This implies that the two processes are completely uncorrelated. If, in addition, the processes $\{X_1(t), X_2(t)\}$ are Normal, then they will also be *independent*, as shown in Chapter 3.

Another simple case of cross correlation is when $\rho_{12}(u)$ is non-zero for $u = 0$, and zero for all other lags. This implies that the stochastic processes are correlated at simultaneous times but otherwise are uncorrelated. More general models for describing cross correlation between two stochastic processes will be given in Sections 8.1.3 and 8.1.4.

8.1.3 *The cross correlation function of a linear process*

One of the simplest ways in which two correlated stochastic processes $\{X_1(t), X_2(t)\}$ can occur is when $X_1(t)$ is the input to a linear system and $X_2(t)$ is the corresponding output plus noise, that is,

$$X_2(t) = \int_0^\infty h(u)X_1(t-u)\,du + Z(t). \tag{8.1.5}$$

In discrete time, the corresponding model is

$$X_{2t} = \sum_{r=0}^{\infty} h_r X_{1t-r} + Z_t. \tag{8.1.6}$$

Example 1. As a special case of (8.1.6), consider the simple regression model (4.3.5) which in this notation is written

$$X_{2t} = h_0 X_{1t} + Z_t.$$

If X_{1t} and Z_t have zero means, the ccvf between the input X_{1t} and output X_{2t} is, from (8.1.1),

$$\begin{aligned}\gamma_{12}(k) &= E[X_{1t}(h_0 X_{1t+k} + Z_{t+k})] \\ &= h_0 E[X_{1t}X_{1t+k}] + E[X_{1t}Z_{t+k}].\end{aligned}$$

If it is further assumed that the noise Z_t is uncorrelated with the input X_{1t}, then

$$\gamma_{12}(k) = h_0\gamma_{11}(k),$$

that is, the ccvf is a constant times the acvf of the input. In the special case where $X_1(t)$ is white noise, the only non-zero cross covariance occurs when $k = 0$, that is,

$$\gamma_{12}(0) = h_0\gamma_{11}(0).$$

Note that while the noise Z_t does not seem to enter into these calculations, its effect is to increase the variance of X_{2t}. Thus

$$\begin{aligned}\text{Var}\,[X_{2t}] &= E[X_{2t}^2] \\ &= E[(h_0X_{1t} + Z_t)^2] \\ &= h_0^2\gamma_{11}(0) + \gamma_{zz}(0).\end{aligned}$$

Hence the cross correlation function for the above example is

$$\begin{aligned}\rho_{12}(0) &= \frac{h_0\gamma_{11}(0)}{\sqrt{\gamma_{11}(0)\{h_0^2\gamma_{11}(0) + \gamma_{zz}(0)\}}} \\ &= \frac{h_0}{\sqrt{h_0^2 + (\gamma_{zz}(0)/\gamma_{11}(0))}}, \\ \rho_{12}(k) &= 0, \quad k \neq 0.\end{aligned} \tag{8.1.7}$$

Thus the correlation between the input and output depends on the *signal-to-noise ratio* $\gamma_{11}(0)/\gamma_{zz}(0)$, that is, the ratio of the input and noise variances. If this ratio is large, then $\rho_{12}(0)$ is close to one, but if it is small, the noise dominates and $\rho_{12}(0)$ is correspondingly small.

Example 2. To consider a less trivial example, suppose that

$$X_{2t} = h_0X_{1t} + h_1X_{1t-1} + Z_t,$$

where X_{1t} and Z_t are uncorrelated white noise processes with the same variance σ^2.

Then

$$\begin{aligned}\gamma_{12}(0) &= E[X_{1t}\{h_0X_{1t} + h_1X_{1t-1} + Z_t\}] \\ &= h_0\sigma^2, \\ \gamma_{12}(1) &= E[X_{1t}\{h_0X_{1t+1} + h_1X_{1t} + Z_{t+1}\}] \\ &= h_1\sigma^2, \\ \gamma_{12}(k) &= 0, \quad k \neq 0, 1.\end{aligned}$$

The variances of the two processes are

$$\begin{aligned}\gamma_{22}(0) &= E[(h_0X_{1t} + h_1X_{1t-1} + Z_t)(h_0X_{1t} + h_1X_{1t-1} + Z_t)] \\ &= (h_0^2 + h_1^2 + 1)\sigma^2, \\ \gamma_{11}(0) &= \sigma^2.\end{aligned}$$

Hence the ccf's are

$$\begin{aligned}\rho_{12}(0) &= \frac{h_0}{\sqrt{1 + h_0^2 + h_1^2}}, \\ \rho_{12}(1) &= \frac{h_1}{\sqrt{1 + h_0^2 + h_1^2}}, \\ \rho_{12}(k) &= 0, \quad k \neq 0, 1.\end{aligned}$$

If the weights h_k in (8.1.6) are positive, the two processes $X_1(t)$, $X_2(t)$ will tend to "look alike" and the cross correlation function will be positive. Conversely if the weights are negative, the two processes will tend to be mirror images of each other, that is, positive changes in one process are associated with negative changes in the other.

The cross correlation function of a general linear process. The general expression for the ccvf of the model (8.1.5) may be obtained by multiplying (8.1.5) by $X_1(t-u)$ and taking expectations. If the $X_1(t)$ and $Z(t)$ processes have zero means, the ccvf is

$$\gamma_{12}(u) = E\left[X_1(t-u)\int_0^\infty h(v)X_1(t-v)\,dv + X_1(t-u)Z(t)\right]$$
$$= \int_0^\infty h(v)\gamma_{11}(u-v)\,dv, \quad -\infty \leqslant u \leqslant \infty, \tag{8.1.8}$$

provided $\gamma_{X_1 Z}(u) = 0$ for all u.

Another result required later is the expression for the autocovariance function of the output. This may be obtained from (8.1.5) by squaring and taking expectations. The final result, assuming $E[X_1(t)] = E[Z(t)] = 0$ and $\gamma_{X_1 Z}(u) = 0$, is

$$\gamma_{22}(u) = \int_0^\infty \int_0^\infty h(v)h(v')\gamma_{11}(u+v-v')\,dv\,dv' + \gamma_{ZZ}(u), \quad -\infty \leqslant u \leqslant \infty, \tag{8.1.9}$$

which is a simple extension of (5.2.9).

The cross correlation function is then obtained from

$$\rho_{12}(u) = \frac{\gamma_{12}(u)}{\sqrt{\gamma_{11}(0)\gamma_{22}(0)}}, \tag{8.1.10}$$

with $\gamma_{22}(0)$ obtained from (8.1.9) by setting $u = 0$.

For discrete processes, the results corresponding to (8.1.8), (8.1.9) and (8.1.10) are readily obtained from (8.1.6). Thus

$$\gamma_{12}(k) = \sum_{r=0}^{\infty} h_r\gamma_{11}(k-r), \quad k = 0, \pm 1, \pm 2, \ldots, \tag{8.1.11}$$

$$\gamma_{22}(k) = \sum_{r=0}^{\infty}\sum_{s=0}^{\infty} h_r h_s \gamma_{11}(k+r-s) + \gamma_{ZZ}(k), \quad k = 0, \pm 1, \pm 2, \ldots, \tag{8.1.12}$$

$$\rho_{12}(k) = \frac{\gamma_{12}(k)}{\sqrt{\gamma_{11}(0)\gamma_{22}(0)}}. \tag{8.1.13}$$

8.1.4 Bivariate linear processes

The model (8.1.5) assumed that the fluctuations in $X_1(t)$ *caused* the fluctuations in $X_2(t)$. A more general model for cross correlation between two stochastic processes occurs when it is assumed that the fluctuations in $X_1(t)$ and $X_2(t)$ are caused by two other sources $Z_1(t)$ and $Z_2(t)$ which affect them in different ways. For example, in the simplest possible situation,

$$X_1(t) = h_{11}Z_1(t) + h_{12}Z_2(t),$$
$$X_2(t) = h_{21}Z_1(t) + h_{22}Z_2(t),$$

where $Z_1(t)$, $Z_2(t)$ are uncorrelated white noise processes with variances σ_1^2, σ_2^2. Hence

$$\gamma_{12}(0) = E[\{h_{11}Z_1(t) + h_{12}Z_2(t)\}\{h_{21}Z_1(t) + h_{22}Z_2(t)\}]$$
$$= h_{11}h_{21}\sigma_1^2 + h_{12}h_{22}\sigma_2^2,$$
$$\gamma_{12}(k) = 0, \quad k \neq 0.$$

More generally, suppose that the bivariate stochastic process $\{X_1(t), X_2(t)\}$ is generated according to the lattice diagram of Figure 8.5. Two white noise sources $Z_i(t)$, $i = 1, 2$, are fed into four linear systems with impulse response functions $h_{11}(u)$, $h_{12}(u)$, $h_{21}(u)$ and $h_{22}(u)$ respectively. The outputs from the

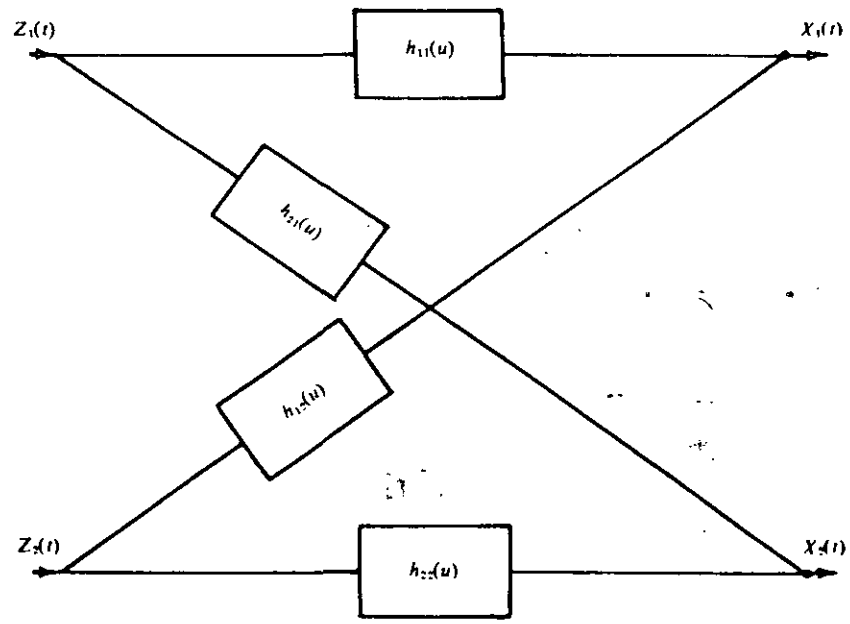


FIG. 8.5: Lattice representation of a bivariate linear ess

first and third systems are summed to give $X_1(t)$, and the outputs from the second and fourth systems are summed to give $X_2(t)$. Thus

$$X_1(t) = \int_0^\infty h_{11}(v)Z_1(t-v)\,dv + \int_0^\infty h_{12}(v)Z_2(t-v)\,dv,$$
$$X_2(t) = \int_0^\infty h_{21}(v)Z_1(t-v)\,dv + \int_0^\infty h_{22}(v)Z_2(t-v)\,dv. \tag{8.1.14}$$

The stochastic process (8.1.14) is called a *bivariate linear process.*

Covariance functions of a bivariate linear process. If the white noise sources are mutually uncorrelated, that is,

$$E[Z_i(t)Z_j(t')] = 0, \quad \text{all } t, t', \quad i = 1, 2, \quad j = 1, 2,$$

then using (5.2.10),

$$\gamma_{11}(u) = \sigma_1^2 \int_0^\infty h_{11}(v)h_{11}(v+u)\,dv + \sigma_2^2 \int_0^\infty h_{12}(v)h_{12}(v+u)\,dv,$$
$$\gamma_{22}(u) = \sigma_1^2 \int_0^\infty h_{21}(v)h_{21}(v+u)\,dv + \sigma_2^2 \int_0^\infty h_{22}(v)h_{22}(v+u)\,dv,$$
$$\gamma_{21}(u) = \sigma_1^2 \int_0^\infty h_{21}(v)h_{11}(v+u)\,dv + \sigma_2^2 \int_0^\infty h_{22}(v)h_{12}(v+u)\,dv, \tag{8.1.15}$$
$$\gamma_{12}(u) = \sigma_1^2 \int_0^\infty h_{11}(v)h_{21}(v+u)\,dv + \sigma_2^2 \int_0^\infty h_{12}(v)h_{22}(v+u)\,dv.$$

In the discrete case the formulae are similar except that integrals are replaced by the corresponding sums. The formulae (8.1.15) show that by adjusting the impulse response functions $h_{ij}(u)$, it is possible to generate a bivariate stochastic process $\{X_1(t), X_2(t)\}$ with any specified cross covariance function and autocovariance functions.

A more general model still is possible by allowing the white noise processes in (8.1.14) to be correlated at simultaneous times, that is,

$$E[Z_1(t), Z_2(t')] = \sigma_{12}\,\delta(t-t').$$

8.1.5 *Bivariate autoregressive and moving average processes*

The simplest type of bivariate linear process occurs when the impulse response functions $h_{ij}(u)$ are zero beyond a certain point. For example, consider the discrete process

$$X_{1t} = Z_{1t} + \beta_{11}Z_{1t-1} + \beta_{12}Z_{2t-1},$$
$$X_{2t} = Z_{2t} + \beta_{21}Z_{1t-1} + \beta_{22}Z_{2t-1}. \tag{8.1.16}$$

Then if the Z_{1t}, Z_{2t} are uncorrelated white noise processes with variances σ_1^2 and σ_2^2, the ccvf of the bivariate process $\{X_{1t}, X_{2t}\}$ is

$$\gamma_{12}(-1) = \beta_{12}\sigma_2^2,$$
$$\gamma_{12}(0) = \beta_{11}\beta_{21}\sigma_1^2 + \beta_{12}\beta_{22}\sigma_2^2,$$
$$\gamma_{12}(1) = \beta_{21}\sigma_1^2,$$
$$\gamma_{12}(k) = 0, \quad k \neq 0, \pm 1.$$

Note that the above ccvf is not one-sided like the one given in example 2 of Section 8.1.3.

Bivariate autoregressive processes. These processes have the property that the impulse response functions $h_{ij}(u)$ in (8.1.14) do not vanish after a finite lag. For example, it is possible to define a continuous first-order process which is a generalization of (5.2.24). Thus, if $Z_1(t)$ and $Z_2(t)$ are white noise processes, correlated at simultaneous times only, a bivariate ar process is defined in continuous time by

$$\frac{dX_1(t)}{dt} + a_{11}X_1(t) + a_{12}X_2(t) = Z_1(t),$$
$$\frac{dX_2(t)}{dt} + a_{21}X_1(t) + a_{22}X_2(t) = Z_2(t), \tag{8.1.17}$$

and in discrete time by

$$X_{1t} = \alpha_{11}X_{1t-1} + \alpha_{12}X_{2t-1} + Z_{1t},$$
$$X_{2t} = \alpha_{21}X_{1t-1} + \alpha_{22}X_{2t-1} + Z_{2t}, \tag{8.1.18}$$

where without loss of generality we assume the processes have zero means.

The computation of the auto- and cross covariances of the continuous process (8.1.17) using (8.1.15) is laborious and is achieved more elegantly using matrix methods to be developed in Chapter 11. For the present it is noted that the auto- and cross covariances of the process (8.1.17) may be written

$$\gamma_{11}(u) = b_{11}\,e^{-a_{11}u} + b_{21}\,e^{-a_{12}u},$$
$$\gamma_{22}(u) = b_{12}\,e^{-a_{21}u} + b_{22}\,e^{-a_{22}u},$$
$$\gamma_{12}(u) = b_{12}\,e^{-a_{11}u} + b_{22}\,e^{-a_{12}u},$$
$$\gamma_{21}(u) = b_{11}\,e^{-a_{21}u} + b_{21}\,e^{-a_{22}u},$$

where the b_{ij} are functions of the a_{ij}. It is interesting to note that the acf of the bivariate first-order process has the same form as the acf (5.2.35) of a univariate *second-order* ar process.

Explicit expressions for the auto- and cross covariances of the disrete process (8.1.18) are derived very simply in Chapter 11 using matrix theory. However, they may also be generated recursively using a scalar recurrence relation for

the covariances, analogous to (5.2.43). Thus, multiplying the first equation in (8.1.18) by X_{2t-k} and taking expectations gives

$$E[X_{2t-k}X_{1t}] = \alpha_{11}E[X_{2t-k}X_{1t-1}] + \alpha_{12}E[X_{2t-k}X_{2t-1}] + E[X_{2t-k}Z_{1t}],$$

or

$$\gamma_{21}(k) = \alpha_{11}\gamma_{21}(k-1) + \alpha_{12}\gamma_{22}(k-1), \quad k \geq 1.$$

Similarly,

$$\begin{aligned}
\gamma_{12}(k) &= \alpha_{21}\gamma_{11}(k-1) + \alpha_{22}\gamma_{12}(k-1), \quad k \geq 1,\\
\gamma_{11}(k) &= \alpha_{11}\gamma_{11}(k-1) + \alpha_{12}\gamma_{12}(k-1), \quad k \geq 1,\\
\gamma_{22}(k) &= \alpha_{21}\gamma_{21}(k-1) + \alpha_{22}\gamma_{22}(k-1), \quad k \geq 1.
\end{aligned} \tag{8.1.19}$$

Hence the values of the covariances at lag k are easily generated from the values at lag $k-1$. To start the process it is necessary to know the values for $k = 0$. These may be obtained by squaring and multiplying the equations (8.1.18) and then taking expectations. Thus

$$\begin{aligned}
\gamma_{11}(0) &= \alpha_{11}^2\gamma_{11}(0) + \alpha_{12}^2\gamma_{22}(0) + 2\alpha_{11}\alpha_{12}\gamma_{12}(0) + \sigma_1^2,\\
\gamma_{22}(0) &= \alpha_{21}^2\gamma_{11}(0) + \alpha_{22}^2\gamma_{22}(0) + 2\alpha_{21}\alpha_{22}\gamma_{12}(0) + \sigma_2^2,\\
\gamma_{12}(0) &= \alpha_{11}\alpha_{21}\gamma_{11}(0) + \alpha_{12}\alpha_{22}\gamma_{22}(0) + (\alpha_{11}\alpha_{22} + \alpha_{12}\alpha_{21})\gamma_{12}(0) + \sigma_{12},
\end{aligned}$$

where $\sigma_1^2 = E[Z_{1t}^2]$, $\sigma_2^2 = E[Z_{2t}^2]$ and $\sigma_{12} = E[Z_{1t}Z_{2t}]$. The values of $\gamma_{11}(0)$, $\gamma_{12}(0)$ and $\gamma_{22}(0)$ may then be obtained by solving the above equations in terms of the known parameters α_{ij}, σ_1^2, σ_2^2 and σ_{12}.

Example 1. A realization of $N = 100$ terms of the process

$$\begin{aligned}
X_{1t} &= 0.6X_{1t-1} - 0.5X_{2t-1} + Z_{1t},\\
X_{2t} &= 0.4X_{1t-1} + 0.5X_{2t-1} + Z_{2t}
\end{aligned} \tag{8.1.20}$$

was generated using two independent sets of random Normal deviates with $E[Z_{it}] = 0$, $E[Z_{1t}Z_{2t}] = 0$, Var $[Z_{it}] = 1$. The values of the two series are given in Appendix A8.1 and plotted in Figure 8.6, where it is seen that the pattern in both series is similar. Thus both series tend to have the same sign, and there is also a tendency for a peak or trough in X_{1t} to be followed after one or two observations by a peak or trough in X_{2t}. To explain this behavior it is necessary to calculate the auto- and cross correlation functions of the bivariate process. Using the above procedure, the recurrence relations for the *covariances* are

$$\left.\begin{aligned}
\gamma_{11}(k) &= 0.6\gamma_{11}(k-1) - 0.5\gamma_{12}(k-1)\\
\gamma_{12}(k) &= 0.4\gamma_{11}(k-1) + 0.5\gamma_{12}(k-1)\\
\gamma_{21}(k) &= 0.6\gamma_{21}(k-1) - 0.5\gamma_{22}(k-1)\\
\gamma_{22}(k) &= 0.4\gamma_{21}(k-1) + 0.5\gamma_{22}(k-1)
\end{aligned}\right\} \quad k \geq 1,$$

with

$$\begin{aligned}
0.64\gamma_{11}(0) - 0.25\gamma_{22}(0) + 0.6\gamma_{12}(0) &= 1,\\
-0.16\gamma_{11}(0) + 0.75\gamma_{22}(0) - 0.4\gamma_{12}(0) &= 1,\\
0.24\gamma_{11}(0) - 0.25\gamma_{22}(0) - 0.9\gamma_{12}(0) &= 0.
\end{aligned}$$

Hence

$$\begin{aligned}
\gamma_{11}(0) &= 1.15/0.52 = 2.21,\\
\gamma_{12}(0) &= 0.04/0.52 = 0.08,\\
\gamma_{22}(0) &= 0.96/0.52 = 1.85.
\end{aligned}$$

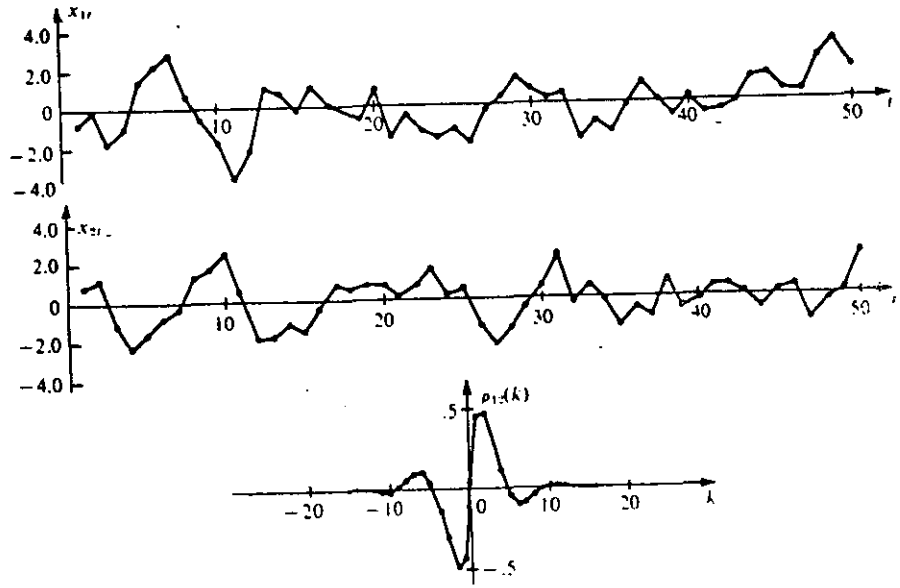


FIG. 8.6: Realization and theoretical cross correlation function of a bivariate autoregressive process

The recurrence relations for the *correlations* are

$$\begin{aligned}
\rho_{11}(k) &= 0.6\rho_{11}(k-1) - 0.5\sqrt{(0.96/1.15)}\rho_{12}(k-1),\\
\rho_{12}(k) &= 0.4\sqrt{(1.15/0.96)}\rho_{11}(k-1) + 0.5\rho_{12}(k-1),\\
\rho_{21}(k) &= 0.6\rho_{21}(k) - 0.5\sqrt{(0.96/1.15)}\rho_{22}(k),\\
\rho_{22}(k) &= 0.4\sqrt{(1.15/0.96)}\rho_{21}(k) + 0.5\rho_{22}(k),
\end{aligned} \tag{8.1.21}$$

with $\rho_{11}(0) = \rho_{22}(0) = 1$ and $\rho_{12}(0) = 0.04/\sqrt{(1.15)(0.96)} = 0.038$.

The correlation function values are shown in Table 8.1 and the ccf is plotted in Figure 8.6. It is seen that whereas $\rho_{12}(0)$ is very small, $\rho_{12}(1)$ and

TABLE 8.1: Theoretical correlations for the bivariate autoregressive process (8.1.20)

k	$\rho_{11}(k)$	$\rho_{22}(k)$	$\rho_{12}(k)$	$\rho_{21}(k)$
0	1.00	1.00	0.04	0.04
1	0.52	0.58	0.46	−0.43
2	0.07	0.14	0.48	−0.50
3	−0.18	−0.14	0.30	−0.33
4	−0.24	−0.22	0.09	−0.11
5	−0.17	−0.17	−0.05	0.04
6	−0.07	−0.08	−0.10	0.10
7	0.01	0.00	−0.09	0.09
8	0.04	0.04	−0.04	0.05
9	0.04	0.04	−0.01	0.01
10	0.03	0.03	0.02	−0.01
11	0.01	0.01	0.02	−0.02
12	0.01	0.00	0.01	−0.02
13	−0.01	−0.01	0.01	−0.01
14	−0.01	−0.01	0.00	0.00
15	−0.01	0.00	—	—
16	0.00	—	—	—
17	—	—	—	—

$\rho_{12}(2)$ are large and positive. This explains the tendency, mentioned earlier, for X_{2t} to lead X_{1t} by one observation. Examination of Figure 8.6 also shows that there is a definite periodicity in the ccf with a period of roughly 10 or a frequency of 0.1 cps. This implies that any pattern in one series will tend to recur or resonate in the other series. The realizations of the two processes shown in Figure 8.6 indicate that there is a tendency for peaks and troughs in X_{1t} to be followed by peaks and troughs at intervals of about 10 in X_{2t}.

Example 2. As a second example of a bivariate linear process consider the process

$$\begin{aligned} X_{1t} &= 0.6X_{1t-1} + Z_{1t}, \\ X_{2t} &= 0.5X_{2t-1} + 2X_{1t-10} + Y_{2t}, \end{aligned} \tag{8.1.22}$$

where

$$Y_{2t} = 0.5Y_{2t-1} + Z_{2t},$$

and Z_{1t} and Z_{2t} are uncorrelated white noise processes with unit variance.

This is an example of the model (8.1.6) given in Section 8.1.3. Thus the process X_{2t} is generated by passing X_{1t} through a linear filter and adding non-white noise. Note that there is an initial delay period of 10 units before X_{1t} begins to affect X_{2t}.

Proceeding as above, the following recurrence relations may be derived for the covariances:

$$\left.\begin{aligned} \gamma_{11}(k) &= 0.6\gamma_{11}(k-1) \\ \gamma_{12}(k) &= 2\gamma_{11}(k-10) + 0.5\gamma_{12}(k-1) \\ \gamma_{21}(k) &= 0.6\gamma_{21}(k-1) \\ \gamma_{22}(k) &= 2\gamma_{21}(k-10) + 0.5\gamma_{22}(k-1) + \gamma_{X_2Y}(k-1) \end{aligned}\right\} \quad k \geq 1, \tag{8.1.23}$$

with

$$\left.\begin{aligned} \gamma_{X_1Y}(k) &= 0, && \text{all } k \\ \gamma_{X_2Y}(k) &= 0.5\gamma_{X_2Y}(k-1) \\ \gamma_{YX_2}(k) &= 0.5\gamma_{YX_2}(k-1) + \gamma_{YY}(k) \\ \gamma_{YY}(k) &= 0.5\gamma_{YY}(k-1) \end{aligned}\right\} \quad k \geq 1,$$

and

$$\begin{aligned} \gamma_{11}(0) &= 1/0.64 = 1.56, \\ \gamma_{YY}(0) &= 1/0.75 = 1.33, \\ -4\gamma_{11}(0) + 0.75\gamma_{22}(0) - 2\gamma_{12}(9) &= \gamma_{YY}(0) + \gamma_{X_2Y}(1), \\ 1.2\gamma_{11}(9) - 0.7\gamma_{12}(0) &= 0, \\ 0.75\gamma_{X_2Y}(0) - 0.5\gamma_{YY}(1) &= 1. \end{aligned} \tag{8.1.24}$$

Solving the equations (8.1.24) for $k = 0$ and substituting the solutions in the recurrence relations (8.1.23) enables the covariances to be calculated. Normalizing these gives the correlations for the process, and these are tabulated in Table 8.2. The ccf has a fairly wide peak centered at lag 10, as would be expected because of the delay of 10 units between the two processes.

Bivariate autoregressive–moving average processes. A more general bivariate process can be obtained by including both ar and ma terms. For example, the discrete process

$$\begin{aligned} X_{1t} &= \alpha_{11}X_{1t-1} + \alpha_{12}X_{2t-1} + Z_{1t} + \beta_{11}Z_{1t-1} + \beta_{12}Z_{2t-1}, \\ X_{2t} &= \alpha_{21}X_{1t-1} + \alpha_{22}X_{2t-1} + Z_{2t} + \beta_{21}Z_{1t-1} + \beta_{22}Z_{2t-1} \end{aligned} \tag{8.1.25}$$

is obtained by combining the models (8.1.16) and (8.1.18).

As stated above, these processes are most easily expressed in matrix form for mathematical conciseness, and a more general treatment of their properties is postponed until Chapter 11.

TABLE 8.2: Theoretical correlations for the bivariate linear process with delay (8.1.22)

k	$\rho_{11}(k)$	$\rho_{22}(k)$	$\rho_{12}(k)$	$\rho_{21}(k)$
0	1.00	1.00	0.00	0.00
1	0.60	0.84	0.01	—
2	0.36	0.62	0.01	—
3	0.22	0.43	0.02	—
4	0.13	0.28	0.04	—
5	0.08	0.18	0.06	—
6	0.05	0.12	0.11	—
7	0.03	0.07	0.18	—
8	0.02	0.04	0.30	—
9	0.01	0.03	0.50	—
10	0.01	0.02	0.83	—
11	0.00	0.01	0.77	—
12	—	0.01	0.59	—
13	—	0.00	0.42	—
14	—	—	0.29	—
15	—	—	0.19	—
16	—	—	0.12	—
17	—	—	0.08	—
18	—	—	0.05	—
19	—	—	0.03	—
20	—	—	0.02	—

8.2 ESTIMATION OF THE CROSS COVARIANCE FUNCTION

8.2.1 The sample cross covariance function

It was shown in Section 5.3.1 that a reasonable estimator of the ccvf at lag u, if the means of the two processes are zero, is

$$c_{X_1X_2}(u) = \begin{cases} \frac{1}{T}\int_{-T/2}^{T/2-u} X_1(t)X_2(t+u)\,dt, & 0 \leqslant u \leqslant T, \\ \frac{1}{T}\int_{-T/2+u}^{T/2} X_1(t)X_2(t+u)\,dt, & -T \leqslant u \leqslant 0. \end{cases} \tag{8.2.1}$$

As with the estimation of autocovariances, the divisor T is preferable to $T-u$ since the estimator has smaller mean square error.

Taking expectations in (8.2.1),

$$E[c_{X_1X_2}(u)] = \left(1 - \frac{|u|}{T}\right)\gamma_{X_1X_2}(u),$$

which shows that $c_{X_1X_2}(u)$ is a biased estimator of $\gamma_{12}(u)$ and only becomes unbiased as T tends to infinity.

If allowance is made for a non-zero mean by using the estimator

$$c_{X_1X_2}(u) = \begin{cases} \frac{1}{T}\int_{-T/2}^{T/2-u} (X_1(t) - \bar{X}_1)(X_2(t+u) - \bar{X}_2)\,dt, & 0 \leqslant u \leqslant T, \\ \frac{1}{T}\int_{-T/2+u}^{T/2} (X_1(t) - \bar{X}_1)(X_2(t+u) - \bar{X}_2)\,dt, & -T \leqslant u \leqslant 0, \end{cases} \tag{8.2.2}$$

where

$$\bar{X}_i = \frac{1}{T}\int_{-T/2}^{T/2} X_i(t)\,dt, \quad i = 1, 2,$$

then a similar calculation to that made in Section 5.3.3 shows that

$$E[c_{X_1X_2}(u)] = \left(1 - \frac{|u|}{T}\right)\gamma_{X_1X_2}(u) + \frac{1}{T}\int_{-T}^{T}\left(1 - \frac{|u|}{T}\right)\gamma_{X_1X_2}(u)\,du.$$

Hence the bias is increased by a term of order $1/T$ by introducing a correction for the mean.

The sample ccvf suffers from the same disadvantages as the sample acvf, namely that neighboring values tend to be highly correlated. It is shown in Appendix A9.1 that the covariance between the estimators $c_{X_1X_2}(u_1)$, $c_{X_1X_2}(u_2)$ at two different lags u_1 and u_2 is given by Bartlett's formula

$$\text{Cov}\,[c_{X_1X_2}(u_1), c_{X_1X_2}(u_2)] = \frac{1}{T^2}\left\{T'\int_{-T'}^{T'}\gamma(r)\left(1 - \frac{|r|}{T'}\right)dr - T''\int_{-T''}^{T''}\gamma(r)\left(1 - \frac{|r|}{T''}\right)dr\right\}, \tag{8.2.3}$$

where

$$T' = T - \frac{|u_1| + |u_2|}{2}, \qquad T'' = \frac{|u_2| - |u_1|}{2},$$

$$\gamma(r) = \gamma_{X_1X_1}\left(r - \frac{u_2 - u_1}{2}\right)\gamma_{X_2X_2}\left(r + \frac{u_2 - u_1}{2}\right) + \gamma_{X_1X_2}\left(r + \frac{u_2 + u_1}{2}\right)\gamma_{X_2X_1}\left(r + \frac{u_2 + u_1}{2}\right) + K(r, u_1, u_2),$$

and $K(r, u_1, u_2)$ is the joint cumulant of the rv's $X_1(t)$, $X_1(t+u_1)$, $X_2(t+r)$ and $X_2(t+r+u_2)$.

For large T, (8.2.3) may be approximated by the following expression, which is analogous to (5.3.22):

$$\text{Cov}\,[c_{X_1X_2}(u_1), c_{X_1X_2}(u_2)] \approx \frac{1}{T}\int_{-\infty}^{\infty}\gamma(r)\,dr. \tag{8.2.4}$$

In the discrete case, the approximation is

$$\text{Cov}\,[c_{X_1X_2}(k), c_{X_1X_2}(l)] \approx \frac{1}{N}\sum_{r=-\infty}^{\infty}\{\gamma_{X_1X_1}(r)\gamma_{X_2X_2}(r+l-k) + \gamma_{X_1X_2}(r+l)\gamma_{X_2X_1}(r-k)\}. \tag{8.2.5}$$

Effect of autocorrelation on the cross correlation between two time series. An interesting case of (8.2.4) occurs when $\gamma_{x_1x_2}(u) = 0$ for all u, that is, the two processes are uncorrelated. Then (8.2.4) becomes

$$\text{Cov}\,[c_{x_1x_2}(u_1), c_{x_1x_2}(u_2)] \approx \frac{1}{T}\int_{-\infty}^{\infty} \gamma_{x_1x_1}\left(r - \frac{u_2 - u_1}{2}\right)\gamma_{x_2x_2}\left(r + \frac{u_2 - u_1}{2}\right)dr, \quad (8.2.6)$$

on neglecting the non-Normality term.

Similarly for two uncorrelated discrete processes, (8.2.5) becomes

$$\text{Cov}\,[c_{x_1x_2}(k), c_{x_1x_2}(l)] \approx \frac{1}{N}\sum_{r=-\infty}^{\infty} \gamma_{x_1x_1}(r)\gamma_{x_2x_2}(r + l - k). \quad (8.2.7)$$

For $X_1(t)$, $X_2(t)$ first-order ar processes with parameters α_1 and β_1 respectively,

$$\gamma_{x_1x_1}(k) = \sigma_1^2\alpha_1^{|k|}, \quad \gamma_{x_2x_2}(k) = \sigma_2^2\beta_1^{|k|},$$

and substitution in (8.2.7) with $l = k$ yields

$$\text{Var}\,[c_{x_1x_2}(k)] \approx \frac{\sigma_1^2\sigma_2^2}{N}\left(\frac{1 + \alpha_1\beta_1}{1 - \alpha_1\beta_1}\right). \quad (8.2.8)$$

For white noise the corresponding result is

$$\text{Var}\,[c_{x_1x_2}(k)] \approx \frac{\sigma_1^2\sigma_2^2}{N}. \quad (8.2.9)$$

Hence if $\alpha_1\beta_1$ is positive, the variance (8.2.8) of the cross covariance estimator is inflated relative to (8.2.9) for two white noise sources, whereas if $\alpha_1\beta_1$ is negative, (8.2.8) is deflated. It is usually the case that $\alpha_1\beta_1$ is positive, that is, the two processes are either both positively correlated or both negatively correlated. Equation (8.2.8) then shows that very large cross covariances, all of them spurious, can be generated *between* two uncorrelated processes as a result of the large autocovariances *within* the two processes.

An example. To illustrate this effect, the sample ccf $r_{x_1x_2}(k)$ is computed for realizations of two independent first-order ar processes with parameters $\alpha_1 = \beta_1 = -0.9$ and $N = 100$. The estimate used is based on the discrete analog of (8.2.2), namely

$$r_{x_1x_2}(k) = \frac{c_{x_1x_2}(k)}{\sqrt{c_{x_1x_1}(0)c_{x_2x_2}(0)}},$$

where

$$c_{x_ix_j}(k) = \frac{1}{N}\sum_{t=1}^{N-k}(x_{it} - \bar{x}_i)(x_{jt+k} - \bar{x}_j), \quad k \geq 0. \quad (8.2.10)$$

$$c_{x_ix_j}(-k) = c_{x_ix_j}(k), \quad k < 0.$$

and

$$\bar{x}_i = \frac{1}{N}\sum_{t=1}^{N} x_{it} \quad i, j = 1, 2.$$

The cross correlation estimate is shown in Figure 8.7 as a broken line. It is seen that values of $r_{x_1x_2}(k)$ as large as ± 0.3 can occur, whereas the theoretical ccf is, of course, zero.

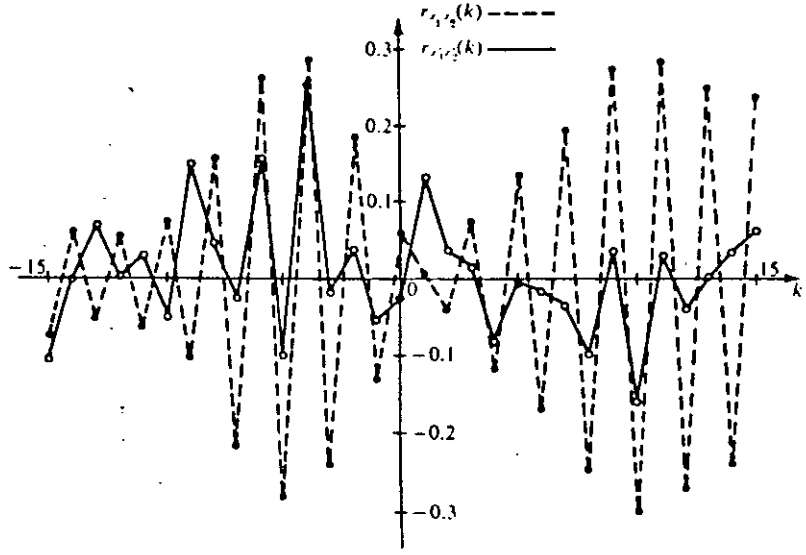


FIG. 8.7: Sample cross correlations of two first-order ar processes before and after filtering

8.2.2 *Improvement of the sample cross correlation function estimator*

To see how the estimation properties of the ccvf estimator can be improved, consider the following model for two cross correlated first-order ar processes:

$$\begin{aligned} X_{1t} &= \alpha_1 X_{1t-1} + Z_{1t}, \\ X_{2t} &= \beta_1 X_{2t-1} + Z_{2t}. \end{aligned} \quad (8.2.11)$$

Suppose that the covariance between Z_{1t} and Z_{2t-k} is zero for k not equal to zero. Then assuming that Z_{1t}, Z_{2t} are bivariate Normal, using (3.1.17), the log likelihood function of the parameters α_1, β_1, γ_{12} is

$$l(\alpha_1, \beta_1, \gamma_{12}(k)) = -(N-k)\log 2\pi - (N-k)\log\sigma_1 - (N-k)\log\sigma_2 - \frac{\sigma_1^2\sigma_2^2}{2(\sigma_1^2\sigma_2^2 - \gamma_{12}^2(k))^2}\left[\sum_{t=1}^{N-k}\left\{\frac{z_{1t}^2}{\sigma_1^2} + \frac{z_{2t-k}^2}{\sigma_2^2} - \frac{2\gamma_{12}(k)}{\sigma_1^2\sigma_2^2}z_{1t}z_{2t-k}\right\}\right].$$

where

$$z_{1t} = x_{1t} - \alpha_1 x_{1t-1},$$
$$z_{2t-k} = x_{2t-k} - \beta_1 x_{2t-k-1}.$$

On differentiating $l(\alpha_1, \beta_1, \gamma_{12}(k))$ with respect to all three parameters, and setting the derivatives equal to zero, the mle of $\gamma_{12}(k)$ is

$$\hat{\gamma}_{12}(k) = \frac{1}{N-k} \sum_{t=1}^{N-k} (x_{1t} - \hat{\alpha}_1 x_{1t-1})(x_{2t-k} - \hat{\beta}_1 x_{2t-k-1}), \quad (8.2.12)$$

where $\hat{\alpha}_1$, $\hat{\beta}_1$ are the mle's of α_1, β_1. The logic behind this estimate is clear:

If a test is required for non-zero correlation between two time series, a filtering operation should be carried out on the two series x_{1t} and x_{2t} to convert them to white noise before computing the cross covariance function.

For the two independent processes in the example of section 8.2.1, the parameters were estimated using the mle (5.4.5). For example,

$$\hat{\alpha}_1 = \frac{\sum_{t=2}^{N} (x_{1t} - \bar{x}_1)(x_{1t-1} - \bar{x}_1)}{\sum_{t=2}^{N} (x_{1t-1} - \bar{x}_1)^2}.$$

Then the two series were filtered according to

$$x'_{1t} = x_{1t} - \hat{\alpha}_1 x_{1t-1},$$
$$x'_{2t} = x_{2t} - \hat{\beta}_1 x_{2t-1},$$

and the ccf of the filtered series (x'_{1t}, x'_{2t}) calculated using (8.2.12). This ccf is plotted as the solid line in Figure 8.7, and it is seen that the values are much smaller than those before filtering. Since the filtered series are random, the standard error of the cross correlation estimate is $\sqrt{1/N} = 0.1$, using (8.2.9). Hence the 95% confidence limits are ± 0.2 about their actual values. It is seen that only one interval fails to enclose zero. Since two intervals would be expected not to enclose zero on average if the two processes were uncorrelated, the observed ccf is not inconsistent with the hypothesis that the two processes are uncorrelated.

8.3 THE CROSS SPECTRUM

This section deals with the frequency-domain description of bivariate time series. It is shown that the sample ccvf discussed in the preceding section has a Fourier transform called the *sample cross spectrum*. This sample cross spectrum is a complex quantity and may be written as the product of a real function called the sample cross *amplitude spectrum* and a complex function

called the sample *phase spectrum*. Similarly, the theoretical cross covariance function has a Fourier transform called the cross spectrum, which can be represented as the product of the cross amplitude spectrum and the phase spectrum. The cross amplitude spectrum shows whether frequency components in one series are associated with large or small amplitudes at the same frequency in the other series. Similarly, the phase spectrum shows whether frequency components in one series lag or lead the components at the same frequency in the other series. The following section contains examples of cross amplitude and phase spectra, based on the cross spectrum of the bivariate linear process (8.1.14). A more useful quantity than the cross amplitude spectrum, namely the coherency spectrum, is then introduced. The coherency spectrum and the phase spectrum are shown to provide a complete description of a bivariate Normal stochastic process.

8.3.1 Fourier analysis applied to bivariate time series

Fourier analysis may be applied to bivariate time series as well as univariate time series. For example, suppose $x_1(t)$, $x_2(t)$ are two cosine waves with the same frequency f_0, different amplitudes A_1 and A_2, and different phases ϕ_1 and ϕ_2 respectively. That is,

$$x_i(t) = A_i \cos(2\pi f_0 t + \phi_i), \quad i = 1, 2. \quad (8.3.1)$$

If records of length T are available, then using (2.2.11) the Fourier transforms of $x_i(t)$, $-T/2 \leqslant t \leqslant T/2$, are

$$X_i(f) = \frac{A_i}{2}\left[e^{i\phi_i}\left\{\frac{\sin \pi(f - f_0)T}{\pi(f - f_0)}\right\} + e^{-i\phi_i}\left\{\frac{\sin \pi(f + f_0)T}{\pi(f + f_0)}\right\}\right], \quad i = 1, 2. \quad (8.3.2)$$

Hence the sample spectra (6.1.6) of the two signals are

$$C_{x_i x_i}(f) = \frac{|X_i(f)|^2}{T}, \quad i = 1, 2,$$

which tend to

$$\tfrac{1}{2}A_i^2[\delta(f - f_0) + \delta(f + f_0)]$$

as T tends to infinity. Thus the variance or average power of a cosine wave is equal to $(1/2)A_i^2$ distributed as delta functions at the frequencies $f = \pm f_0$.

Suppose now that it is required to describe the *covariance* between the two cosine waves. Then a natural function to use is the sample cross power spectrum or, more concisely, the *sample cross spectrum*

$$C_{x_1 x_2}(f) = \frac{X_1^*(f) X_2(f)}{T}, \quad (8.3.3)$$

where the asterisk denotes a complex conjugate. Substituting (8.3.2) in (8.3.3) shows that the sample cross spectrum of two cosine waves is

$$C_{x_1x_2}(f) = \frac{A_1A_2}{4T}\left\{e^{-j\phi_1}\left[\frac{\sin \pi(f-f_0)T}{\pi(f-f_0)}\right] + e^{j\phi_1}\left[\frac{\sin \pi(f+f_0)T}{\pi(f+f_0)}\right]\right\} \times \left\{e^{j\phi_2}\left[\frac{\sin \pi(f-f_0)T}{\pi(f-f_0)}\right] + e^{-j\phi_2}\left[\frac{\sin \pi(f+f_0)T}{\pi(f+f_0)}\right]\right\}, \quad (8.3.4)$$

which tends to

$$\tfrac{1}{4}A_1A_2[e^{-j(\phi_2-\phi_1)}\,\delta(f+f_0) + e^{j(\phi_2-\phi_1)}\,\delta(f-f_0)] \quad (8.3.5)$$

as T tends to infinity.

The definition (8.3.3) is a natural one to adopt since it summarizes all the information about the dependence between the two signals. For the special case of two cosine waves, (8.3.5) shows that this information is summarized by the *phase difference* $\phi_2 - \phi_1$, which shows how one cosine wave leads or lags the other, and the *cross amplitude* A_1A_2, which shows whether a large amplitude at a particular frequency in one wave is associated with a large amplitude at the same frequency in the other.

Sample phase and cross amplitude spectra. More generally, suppose that $x_1(t)$ and $x_2(t)$ are arbitrary real signals with Fourier transforms $X_1(f)$ and $X_2(f)$ respectively. These Fourier transforms give the amplitude and phase distribution of the signals, that is,

$$X_i(f) = A_i(f)\,e^{jF_i(f)}, \quad i = 1, 2, \quad (8.3.6)$$

where $A_i(f)$ is a *positive even function* and $F_i(f)$ is an *odd function*. From (8.3.3), the sample cross spectrum is then

$$C_{x_1x_2}(f) = A_1(f)A_2(f)\,e^{j(F_2(f)-F_1(f))}/T, \quad (8.3.7)$$

which may be written

$$C_{12}(f) = A_{12}(f)\,e^{jF_{12}(f)}. \quad (8.3.8)$$

Hence the covariance between the two series $x_1(t)$ and $x_2(t)$ may be described by the *sample phase spectrum*

$$F_{12}(f) = F_2(f) - F_1(f) \quad (8.3.9)$$

and the *sample cross amplitude spectrum*

$$A_{12}(f) = A_1(f)A_2(f)/T. \quad (8.3.10)$$

The sample phase spectrum $F_{12}(f)$ shows whether the frequency components in one series lead or lag the components at the same frequency in the other series. Similarly, the sample cross amplitude spectrum $A_{12}(f)$ shows whether the amplitude of the component at a particular frequency in one series is associated with a large or small amplitude at the same frequency in the other series. Note that $A_{12}(f)$ is a positive even function of frequency and that $F_{12}(f)$ is an odd function of frequency.

Sample co- and quadrature spectra. Since (8.3.8) is a complex quantity, it may be written as the product of an amplitude function times a phase function as in (8.3.7). An alternative expression for (8.3.8) is as a sum of a real and imaginary part, that is,

$$C_{12}(f) = L_{12}(f) - jQ_{12}(f),$$

where

$$L_{12}(f) = A_{12}(f)\cos F_{12}(f), \qquad Q_{12}(f) = -A_{12}(f)\sin F_{12}(f), \quad (8.3.11)$$

and

$$A_{12}^2(f) = L_{12}^2(f) + Q_{12}^2(f), \qquad F_{12}(f) = \arctan -\frac{Q_{12}(f)}{L_{12}(f)}.$$

Note that $L_{12}(f)$ is an *even* function of frequency and $Q_{12}(f)$ is an *odd* function of frequency because $A_{12}(f)$ is an even and $F_{12}(f)$ an odd function of frequency. To illustrate, consider the bivariate cosine wave example used above.

In the limiting case as T tends to infinity, it may be shown that

$$L_{12}(f) = \frac{A_1A_2}{4}\cos(\phi_2-\phi_1)[\delta(f+f_0) + \delta(f-f_0)] = \left\{\frac{A_1\cos\phi_1A_2\cos\phi_2}{4} + \frac{A_1\sin\phi_1A_2\sin\phi_2}{4}\right\}[\delta(f+f_0) + \delta(f-f_0)].$$

Since the signals $x_i(t)$ may be written

$$\begin{aligned} x_1(t) &= A_1\cos(2\pi f_0t + \phi_1) \\ &= (A_1\cos\phi_1)\cos 2\pi f_0t - (A_1\sin\phi_1)\sin 2\pi f_0t, \\ x_2(t) &= A_2\cos(2\pi f_0t + \phi_2) \\ &= (A_2\cos\phi_2)\cos 2\pi f_0t - (A_2\sin\phi_2)\sin 2\pi f_0t, \end{aligned}$$

it follows that $L_{12}(f)$ measures the covariance between the two cosine components and the covariance between the two sine terms, that is, the covariance between the *in-phase* components. Hence $L_{12}(f)$ is called the sample in-phase or *sample co-spectrum*. Similarly,

$$Q_{12}(f) = \frac{A_1A_2}{4}\sin(\phi_2-\phi_1)[\delta(f+f_0) + \delta(f-f_0)] = \left\{\frac{A_1\cos\phi_1A_2\sin\phi_2}{4} - \frac{A_1\sin\phi_1A_2\cos\phi_2}{4}\right\}[\delta(f+f_0) + \delta(f-f_0)]$$

measures the covariance between the sine and cosine components, that is, the *out-of-phase* or *quadrature* components. Hence $Q_{12}(f)$ is called the *sample quadrature spectrum*.

Similarly for general signals $x_1(t)$, $x_2(t)$, the Fourier representations

$$L_{12}(f) = A_{12}(f)\cos F_{12}(f)$$

and

$$Q_{12}(f) = A_{12}(f)\sin F_{12}(f)$$

measure the covariance between the in-phase and quadrature components at frequency f.

8.3.2 *The relation between the sample cross spectrum and the sample cross covariance function*

It was shown in Chapter 6 that the sample acvf and the sample spectrum are related by the Fourier transform (6.1.9). In this section, (6.1.9) is generalized to show that the sample ccvf and the sample cross spectrum likewise form a Fourier transform pair.

From the definition (8.3.3) of the sample cross spectrum,

$$\begin{aligned} C_{12}(f) &= \frac{1}{T} X_1^*(f)X_2(f) \\ &= \frac{1}{T}\int_{-T/2}^{T/2}\int_{-T/2}^{T/2} x_1(t)x_2(t')\, e^{-j2\pi f(t'-t)}\, dt\, dt'. \end{aligned}$$

Making the transformation $t' - t = u$, $t = v$ and proceeding as in Section 6.1.3,

$$C_{12}(f) = \int_{-T}^{T} c_{12}(u)\, e^{-j2\pi fu}\, du. \tag{8.3.12}$$

Thus the sample cross spectrum is the Fourier transform of the sample ccvf defined by

$$\begin{aligned} c_{12}(u) &= \frac{1}{T}\int_{-T/2}^{T/2-u} x_1(t)x_2(t+u)\, dt, \qquad 0 \leqslant u \leqslant T, \\ &= \frac{1}{T}\int_{-T/2+u}^{T/2} x_1(t)x_2(t+u)\, dt, \qquad -T \leqslant u \leqslant 0, \\ &= 0, \qquad |u| > T. \end{aligned} \tag{8.3.13}$$

The transform (8.3.12) has the inverse

$$c_{12}(u) = \int_{-\infty}^{\infty} C_{12}(f)\, e^{j2\pi fu}\, df. \tag{8.3.14}$$

Substituting (8.3.11) in (8.3.14) yields the important identity

$$\begin{aligned} c_{12}(u) &= \int_{-\infty}^{\infty} (L_{12}(f) - jQ_{12}(f))\, e^{j2\pi fu}\, df \\ &= \int_{-\infty}^{\infty} L_{12}(f)\cos 2\pi fu\, df + \int_{-\infty}^{\infty} Q_{12}(f)\sin 2\pi fu\, df, \end{aligned} \tag{8.3.15}$$

since $L_{12}(f)$ is an even function of f and $Q_{12}(f)$ is an odd function of f. Finally, substituting $u = 0$ in (8.3.15) gives

$$c_{12}(0) = \int_{-\infty}^{\infty} L_{12}(f)\, df, \tag{8.3.16}$$

and hence the sample co-spectrum gives the decomposition of the zero-lag cross covariance with frequency in the same way that the sample spectrum (6.1.11) gives the decomposition of the sample variance with frequency.

Now suppose that (8.3.15) is written

$$c_{12}(u) = l_{12}(u) + q_{12}(u), \tag{8.3.17}$$

where

$$\begin{aligned} L_{12}(f) &= \int_{-T}^{T} l_{12}(u)\cos 2\pi fu\, du, \\ Q_{12}(f) &= \int_{-T}^{T} q_{12}(u)\sin 2\pi fu\, du. \end{aligned} \tag{8.3.18}$$

It may be verified that $l_{12}(u)$ is the even part of $c_{12}(u)$, that is,

$$l_{12}(u) = \tfrac{1}{2}(c_{12}(u) + c_{12}(-u)). \tag{8.3.19}$$

Similarly, $q_{12}(u)$ is the odd part of $c_{12}(u)$, namely,

$$q_{12}(u) = \tfrac{1}{2}(c_{12}(u) - c_{12}(-u)). \tag{8.3.20}$$

8.3.3 *The cross spectrum*

The analysis of the previous section assumed that $x_1(t)$ and $x_2(t)$ were mathematical functions of time t. If $\{x_1(t), x_2(t)\}$ are realizations of a stationary bivariate stochastic process $\{X_1(t), X_2(t)\}$, the same problems arise as for univariate spectra. Thus if the sample co- and quadrature spectra are computed for a realization of a bivariate stochastic process, they do not tend to a limiting value, in any statistical sense, as the record length T tends to infinity. In fact, their behavior is identical to that of the sample spectrum as shown in Figure 6.1. To see why this is so it is necessary to examine the properties of the rv $C_{x_1x_2}(f)$ associated with the sample cross spectrum.

From (8.3.12) the sample cross spectrum estimator is

$$C_{x_1x_2}(f) = \int_{-T}^{T} c_{x_1x_2}(u)\, e^{-j2\pi fu}\, du, \quad -\infty \leqslant f \leqslant \infty.$$

Using (8.2.2), this has mean value

$$E[C_{x_1x_2}(f)] = \int_{-T}^{T}\left(1 - \frac{|u|}{T}\right)\gamma_{x_1x_2}(u)\, e^{-j2\pi fu}\, du, \tag{8.3.21}$$

and as T tends to infinity, this mean value tends to the cross power spectrum or, more concisely, the *cross spectrum*. Thus

$$\lim_{T\to\infty} E[C_{x_1x_2}(f)] = \Gamma_{x_1x_2}(f) = \int_{-\infty}^{\infty} \gamma_{x_1x_2}(u)\, e^{-j2\pi fu}\, du, \quad -\infty \leqslant f \leqslant \infty. \tag{8.3.22}$$

Equation (8.3.22) shows that the cross spectrum is the Fourier transform of the ccvf. Note that in the stochastic definition (8.3.22), the cross spectrum is a continuous function of frequency in the range $-\infty \leqslant f \leqslant \infty$.

It is emphasized once again that the definition

$$\Gamma_{x_1x_2}(f) = \lim_{T\to\infty} C_{x_1x_2}(f)$$

usually given in engineering texts has no meaning for stochastic processes since the variances of the real and imaginary parts of the rv $C_{x_1x_2}(f)$ do not tend to zero as T tends to infinity, as will be shown in Section 9.1.

Co- and quadrature spectra. Writing $\gamma_{x_1x_2}(u)$ as the sum of an even part $\lambda_{12}(u)$ and an odd part $\psi_{12}(u)$ gives

$$\begin{aligned} \lambda_{12}(u) &= \tfrac{1}{2}(\gamma_{12}(u) + \gamma_{12}(-u)), \\ \psi_{12}(u) &= \tfrac{1}{2}(\gamma_{12}(u) - \gamma_{12}(-u)), \end{aligned} \tag{8.3.23}$$

and substitution in (8.3.22) yields

$$\Gamma_{12}(f) = \Lambda_{12}(f) - j\Psi_{12}(f), \tag{8.3.24}$$

where

$$\Lambda_{12}(f) = \int_{-\infty}^{\infty} \lambda_{12}(u) \cos 2\pi fu\, du \tag{8.3.25}$$

and

$$\Psi_{12}(f) = \int_{-\infty}^{\infty} \psi_{12}(u) \sin 2\pi fu\, du. \tag{8.3.26}$$

$\Lambda_{12}(f)$ is called the *co-spectrum* and $\Psi_{12}(f)$ the *quadrature spectrum* of the $\{X_1(t), X_2(t)\}$ process. Equivalent definitions for these quantities may be obtained from (8.3.21) and (8.3.22). Thus

$$\Lambda_{12}(f) = \lim_{T\to\infty} E[L_{12}(f)], \quad \Psi_{12}(f) = \lim_{T\to\infty} E[Q_{12}(f)].$$

Cross amplitude and phase spectra. The cross spectrum may also be written

$$\Gamma_{12}(f) = \alpha_{12}(f)\, e^{j\phi_{12}(f)}, \tag{8.3.27}$$

where $\alpha_{12}(f)$, $\phi_{12}(f)$ are called the *cross amplitude spectrum* and the *phase spectrum* respectively. From (8.3.11), it is seen that

$$\alpha_{12}(f) = \sqrt{\Lambda_{12}^2(f) + \Psi_{12}^2(f)}, \tag{8.3.28}$$

$$\phi_{12}(f) = \arctan -\frac{\Psi_{12}(f)}{\Lambda_{12}(f)}. \tag{8.3.29}$$

TABLE 8.3: Summary of cross spectral formulae

		Theoretical values			Sample values	
Function	Symbol	Definition	Symbol	Definition		
auto-spectrum	$\Gamma_{11}(f)$	$\Gamma_{11}(f) = \int_{-\infty}^{\infty} \gamma_{11}(u)\, e^{-j2\pi fu}\, du$	$C_{11}(f)$	$C_{11}(f) = \int_{-T}^{T} c_{11}(u)\, e^{-j2\pi fu}\, du$		
cross spectrum	$\Gamma_{12}(f)$	$\Gamma_{12}(f) = \int_{-\infty}^{\infty} \gamma_{12}(u)\, e^{-j2\pi fu}\, du$ $= \alpha_{12}(f)\, e^{+j\phi_{12}(f)}$ $= \Lambda_{12}(f) - j\Psi_{12}(f)$	$C_{12}(f)$	$C_{12}(f) = \int_{-T}^{T} c_{12}(u)\, e^{-j2\pi fu}\, du$ $= A_{12}(f)\, e^{+j\Phi_{12}(f)}$ $= L_{12}(f) - jQ_{12}(f)$		
cross amplitude spectrum	$\alpha_{12}(f)$	$\alpha_{12}(f) = \lvert\Gamma_{12}(f)\rvert$ $= \sqrt{\Lambda_{12}^2(f) + \Psi_{12}^2(f)}$	$A_{12}(f)$	$A_{12}(f) = \lvert C_{12}(f)\rvert$ $= \sqrt{L_{12}^2(f) + Q_{12}^2(f)}$		
phase spectrum	$\phi_{12}(f)$	$\phi_{12}(f) = \arctan\left\{-\frac{\Psi_{12}(f)}{\Lambda_{12}(f)}\right\}$	$F_{12}(f)$	$F_{12}(f) = \arctan\left\{-\frac{Q_{12}(f)}{L_{12}(f)}\right\}$		
co-spectrum	$\Lambda_{12}(f)$	$\Lambda_{12}(f) = \int_{-\infty}^{\infty} \lambda_{12}(u)\, e^{-j2\pi fu}\, du$ $= \tfrac{1}{2}\int_{-\infty}^{\infty} \{\gamma_{12}(u) + \gamma_{12}(-u)\} \cos 2\pi fu\, du$	$L_{12}(f)$	$L_{12}(f) = \int_{-T}^{T} l_{12}(u)\, e^{-j2\pi fu}\, du$ $= \tfrac{1}{2}\int_{-T}^{T} \{c_{12}(u) + c_{12}(-u)\} \cos 2\pi fu\, du$		
quadrature spectrum	$\Psi_{12}(f)$	$\Psi_{12}(f) = \int_{-\infty}^{\infty} \psi_{12}(u)\, e^{-j2\pi fu}\, du$ $= \tfrac{1}{2}\int_{-\infty}^{\infty} \{\gamma_{12}(u) - \gamma_{12}(-u)\} \sin 2\pi fu\, du$	$Q_{12}(f)$	$Q_{12}(f) = \int_{-T}^{T} q_{12}(u)\, e^{-j2\pi fu}\, du$ $= \tfrac{1}{2}\int_{-T}^{T} \{c_{12}(u) - c_{12}(-u)\} \sin 2\pi fu\, du$		

Hence the cross amplitude and phase spectra may be calculated by computing $\lambda_{12}(u)$ and $\psi_{12}(u)$ from the cross covariance function according to (8.3.23), by computing the co- and quadrature spectra using (8.3.25) and (8.3.26), and by substituting in (8.3.28) and (8.3.29).

Summary. Table 8.3 summarizes the spectra which have been defined in this chapter.

8.4 CROSS SPECTRA OF LINEAR PROCESSES

8.4.1 Simple examples of cross spectra

Before deriving the cross spectrum of the general bivariate process (8.1.14) it is instructive to consider some simple examples of cross spectra. These examples will be used to demonstrate what information is contained in the cross spectrum. For this purpose the cross spectra of some simple discrete processes are derived. For discrete processes, the cross spectrum is

$$\Gamma_{12}(f) = \Delta \sum_{k=-\infty}^{\infty} \gamma_{12}(k)\, e^{-j2\pi fk}, \quad -\frac{1}{2\Delta} \leqslant f < \frac{1}{2\Delta}. \tag{8.4.1}$$

In the following examples it is assumed that $E[Z_{1t}] = 0$, $E[Z_{2t}] = 0$, $E[Z_{1t}^2] = \sigma_1^2$, $E[Z_{2t}^2] = \sigma_2^2$ and $\Delta = 1$ so that $-\frac{1}{2} \leqslant f < \frac{1}{2}$.

Example 1. Suppose

$$X_{2t} = Z_{2t}, \qquad X_{1t} = Z_{1t},$$

where Z_{1t} and Z_{2t} are mutually uncorrelated white noise sources. Hence

$$\gamma_{12}(k) = E[Z_{1t}Z_{2t+k}] = 0, \quad \text{all } k,$$

and using (8.3.22),

$$\Gamma_{12}(f) = \alpha_{12}(f)\, e^{j\phi} e^{j\phi_{12}(f)} = \sqrt{\Lambda_{12}^2(f) + \Upsilon_{12}^2(f)} = 0.$$

This implies that the cross amplitude spectrum is zero everywhere and hence the co- and quadrature spectra are identically zero. This would seem to imply that the phase spectrum is indeterminate. It may be shown, however, that the phase spectrum is uniformly distributed in the range $(-\pi/2, \pi/2)$ so that the average phase difference between the two processes is zero but the phase difference at any frequency is equally likely to lie anywhere in the range $(-\pi/2, \pi/2)$.

Example 2 (the bivariate equivalent of white noise). Suppose

$$X_{2t} = Z_{2t} + \beta_1 Z_{1t}, \qquad X_{1t} = Z_{1t},$$

so that

$$X_{2t} = \beta_1 X_{1t} + Z_{2t}. \tag{8.4.2}$$

Hence

$$\begin{aligned}\gamma_{12}(0) &= E[Z_{1t}(Z_{2t} + \beta_1 Z_{1t})] \\ &= \beta_1\sigma_1^2, \\ \gamma_{12}(k) &= 0, \quad k \neq 0.\end{aligned}$$

From (8.4.1)

$$\Gamma_{12}(f) = \beta_1\sigma_1^2,$$

which implies that

$$\begin{aligned}\alpha_{12}(f) &= \beta_1\sigma_1^2, & \phi_{12}(f) &= 0, \\ \Lambda_{12}(f) &= \beta_1\sigma_1^2, & \Upsilon_{12}(f) &= 0.\end{aligned}$$

Hence if the two processes are cross correlated only at simultaneous times, the cross amplitude spectrum is a constant, like the spectrum of white noise. Further, the two processes are in phase, since $\phi_{12}(f) = 0$. The cross amplitude and phase spectra for this example are shown in Figure 8.8(a). *Thus (8.4.2) may be regarded as a fundamental model for cross spectra in the same way that white noise is fundamental for univariate spectra.*

Example 3 (the effect of delay). Suppose

$$X_{2t} = Z_{2t} + \beta_1 Z_{1t-d}, \qquad X_{1t} = Z_{1t},$$

so that

$$X_{2t} = \beta_1 X_{1t-d} + Z_{2t},$$

that is, the two series are shifted by a time interval d relative to each other. Hence

$$\gamma_{12}(k) = \begin{cases}\beta_1\sigma_1^2, & k = d \\ 0, & \text{otherwise.}\end{cases}$$

Therefore, from (8.4.1),

$$\Gamma_{12}(f) = \beta_1\sigma_1^2\, e^{-j2\pi fd},$$

$$\begin{aligned}\alpha_{12}(f) &= \beta_1\sigma_1^2, & \phi_{12}(f) &= -2\pi fd, \\ \Lambda_{12}(f) &= \beta_1\sigma_1^2 \cos 2\pi fd, & \psi_{12}(f) &= \beta_1\sigma_1^2 \sin 2\pi fd.\end{aligned}$$

Again the cross amplitude spectrum is a constant, but the phase spectrum is now a linear function of frequency as shown in Figure 8.8(b). This means that a cosinusoidal wave of frequency f cycles per second takes fd cycles to cover the delay time d seconds and hence the phase lag is $2\pi fd$ radians.

Example 4. A more interesting case occurs with the model

$$X_{2t} = \beta_1 Z_{1t} + \beta_2 Z_{1t-1} + Z_{2t}, \qquad X_{1t} = Z_{1t},$$

which may be rewritten

$$X_{2t} = \beta_1 X_{1t} + \beta_2 X_{1t-1} + Z_{2t}.$$

Hence, as shown in the second example of Section 8.1.3,

$$\gamma_{12}(k) = \begin{cases} \beta_1\sigma_1^2, & k = 0, \\ \beta_2\sigma_1^2, & k = 1, \\ 0, & \text{otherwise.} \end{cases}$$

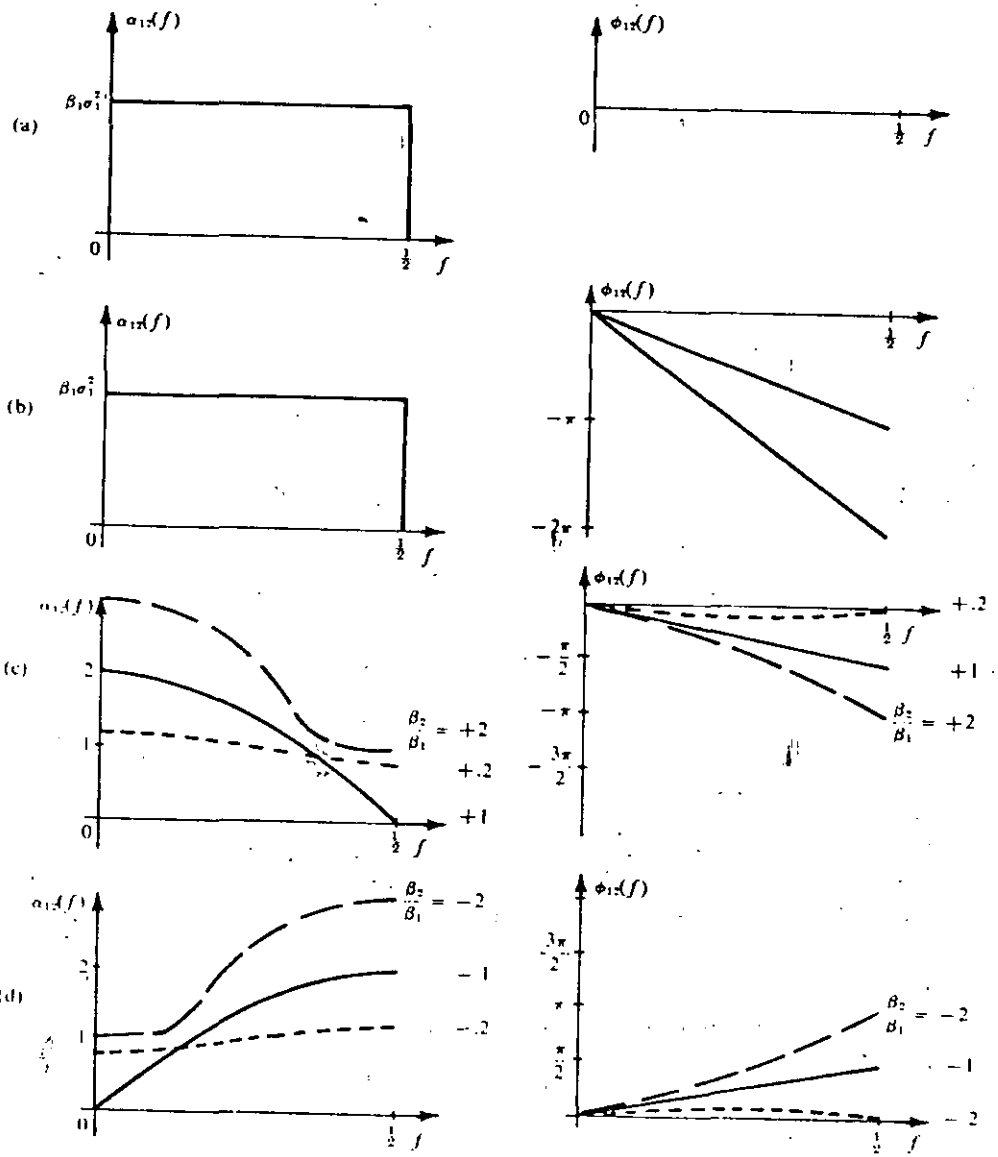


FIG. 8.8: Cross amplitude and phase spectra for some simple bivariate processes

Therefore

$$\Gamma_{12}(f) = \sigma_1^2[\beta_1 + \beta_2 e^{-i2\pi f}],$$
$$\alpha_{12}(f) = \sigma_1^2\sqrt{\beta_1^2 + \beta_2^2 + 2\beta_1\beta_2 \cos 2\pi f},$$
$$\phi_{12}(f) = \arctan\left(\frac{-\beta_2 \sin 2\pi f}{\beta_1 + \beta_2 \cos 2\pi f}\right),$$
$$\Lambda_{12}(f) = \sigma_1^2(\beta_1 + \beta_2 \cos 2\pi f),$$
$$\Psi_{12}(f) = \sigma_1^2\beta_2 \sin 2\pi f.$$

The cross amplitude spectrum shows that the covariance between the two processes is dominated by low frequencies if $\beta_2/\beta_1 > 0$ and by high frequencies if $\beta_2/\beta_1 < 0$. Hence cross correlations all of one sign produce low-frequency cross amplitude spectra and oscillatory cross correlations produce high-frequency cross amplitude spectra. The corresponding phase diagrams are shown in Figure 8.8(c) and (d), and it is seen that when $\beta_2/\beta_1 > 0$, X_{1t} leads X_{2t} and when $\beta_2/\beta_1 < 0$, X_{1t} lags behind X_{2t}.

By using more sophisticated models it is possible to generate a wide variety of cross amplitude and phase spectra. An important point which needs emphasizing at this stage is that examination of the cross amplitude spectra of two empirical time series may indicate that *a different model may be required in different frequency ranges*. For example, a phase diagram made up of straight lines with different slopes would indicate that one series was delayed relative to the other but that the time domain delay between the two series changes from one frequency band to the next.

Hence it is seen that a study of the cross spectra of empirical time series can provide a very flexible tool for suggesting models to describe their behavior. If it is suspected that different models may be operating in different frequency bands, a more efficient analysis results if, as discussed in Section 7.3.5, the two series are filtered into component series using a bank of band-pass filters. Then the cross spectra are calculated using the component series.

8.4.2 *The cross spectrum of a linear system*

It was shown in Section 8.1.3 that two stochastic processes $X_1(t)$, $X_2(t)$ are sometimes related linearly as

$$X_2(t) = \int_0^\infty h(u)X_1(t-u)\,du + Z(t).$$

Thus $X_1(t)$ is the input process to a linear system and $X_2(t)$ is the corresponding output plus an independent noise $Z(t)$. From (8.1.8) the ccvf of the output is

$$\gamma_{12}(u) = \int_0^\infty \gamma_{11}(u-v)h(v)\,dv, \quad -\infty \leqslant u \leqslant \infty. \tag{8.4.3}$$

Transforming (8.4.3) gives

$$\Gamma_{12}(f) = H(f)\Gamma_{11}(f). \tag{8.4.4}$$

Hence the frequency response function may be determined from

$$H(f) = \frac{\Gamma_{12}(f)}{\Gamma_{11}(f)}. \tag{8.4.5}$$

Comparison of equations (8.4.3) and (8.4.4) shows how the analysis of linear systems is simplified by using Fourier techniques. Thus the convolution in (8.4.3) is converted into a multiplication in (8.4.4). Rewriting (8.4.5) in the form

$$H(f) = G(f)\,e^{+j\phi(f)} = \frac{\Lambda_{12}(f) - j\,\Psi_{12}(f)}{\Gamma_{11}(f)}$$

yields expressions for the gain $G(f)$ and phase $\phi(f)$ of the linear system. Thus

$$G(f) = \frac{\sqrt{\Lambda_{12}^2(f) + \Psi_{12}^2(f)}}{\Gamma_{11}(f)} = \frac{\alpha_{12}(f)}{\Gamma_{11}(f)}, \tag{8.4.6}$$

$$\phi(f) = \arctan - \frac{\Psi_{12}(f)}{\Lambda_{12}(f)}. \tag{8.4.7}$$

Remembering that the cross amplitude $\alpha_{12}(f)$ is a measure of the "covariance" between $X_1(t)$ and $X_2(t)$ at frequency f and $\Gamma_{11}(f)$ the "variance" in the input at frequency f, it is seen that the gain $G(f)$ behaves like the regression coefficient (4.3.7) but is now evaluated at each frequency f.

Squared coherency spectrum. This analogy may be pursued further using the relation (8.1.9) for the autocovariance of the output, namely,

$$\gamma_{22}(u) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} h(v)h(v')\gamma_{11}(u + v - v')\,dv\,dv' + \gamma_{zz}(u).$$

On transforming this equation,

$$\Gamma_{22}(f) = G^2(f)\Gamma_{11}(f) + \Gamma_{zz}(f), \tag{8.4.8}$$

which differs from (6.2.15) only in that the spectrum of the noise $Z(t)$ is added to the contribution from the input. Substituting for (8.4.6) in (8.4.8),

$$\Gamma_{22}(f) - \frac{\alpha_{12}^2(f)}{\Gamma_{11}(f)} = \Gamma_{zz}(f)$$

or

$$\Gamma_{zz}(f) = \Gamma_{22}(f)[1 - \kappa_{12}^2(f)], \tag{8.4.9}$$

where

$$\kappa_{12}^2(f) = \frac{\alpha_{12}^2(f)}{\Gamma_{11}(f)\Gamma_{22}(f)} \tag{8.4.10}$$

is called the *squared coherency* between the input and output at frequency f. The plot of $\kappa_{12}^2(f)$ versus f is called the *squared coherency spectrum.*

The resemblance between (8.4.9) and the ordinary correlation coefficient (3.2.19) should be noted. In fact, the coherency plays the role of a correlation coefficient defined at each frequency f. Thus (8.4.9) shows that when the noise spectrum is identical to the output spectrum, the squared coherency is zero. In other words, the squared coherency is zero when the output consists entirely of noise. When $\Gamma_{zz}(f) = 0$, the squared coherency is unity and the output spectrum is simply the input spectrum multiplied by the square of the gain of the system. Eliminating $\Gamma_{22}(f)$ between (8.4.8) and (8.4.10),

$$\kappa_{12}^2(f) = \frac{1}{1 + (\Gamma_{zz}(f)/G^2(f)\Gamma_{11}(f))}. \tag{8.4.11}$$

Equation (8.4.11) shows that the squared coherency is small when the output-signal to noise ratio $G^2(f)\Gamma_{11}(f)/\Gamma_{zz}(f)$ is small and large when the ratio is large.

8.4.3 Cross spectra of bivariate linear processes

It was shown in Section 8.1.4 that a very general model for generating bivariate stochastic processes is obtained by passing two white noise sources $Z_1(t)$ and $Z_2(t)$ through the system shown in Figure 8.5. The auto- and cross covariances of this process are given by (8.1.15). These results will now be used to derive the corresponding auto- and cross spectra. Denote by

$$H_{ij}(f) = \int_{-\infty}^{\infty} h_{ij}(u)\,e^{-j2\pi fu}\,du, \quad i, j = 1, 2,$$

the frequency response functions of the four systems in the lattice diagram of Figure 8.5. The results for the auto- and cross spectra may now be obtained by transforming the equations (8.1.15). Transforming the first two equations and using (6.2.16), the autospectra are

$$\begin{aligned}\Gamma_{11}(f) &= \sigma_1^2|H_{11}(f)|^2 + \sigma_2^2|H_{12}(f)|^2,\\ \Gamma_{22}(f) &= \sigma_1^2|H_{21}(f)|^2 + \sigma_2^2|H_{22}(f)|^2.\end{aligned} \tag{8.4.12}$$

To derive the cross spectrum, note that the last two of the equations (8.1.15) can be combined to give

$$\gamma_{12}(u) = \sigma_1^2\int_{-\infty}^{\infty} h_{11}(v)h_{21}(v + u)\,dv + \sigma_2^2\int_{-\infty}^{\infty} h_{12}(v)h_{22}(v + u)\,dv, \tag{8.4.13}$$

which now holds for $-\infty \leqslant u \leqslant \infty$. On taking Fourier transforms of both sides of (8.4.13), the cross spectrum is

$$\Gamma_{12}(f) = \sigma_1^2 H_{11}^*(f)H_{21}(f) + \sigma_2^2 H_{12}^*(f)H_{22}(f). \tag{8.4.14}$$

Thus the computation of the cross spectrum reduces to computation of the frequency response functions of the associated bivariate linear process (8.1.14).

The frequency response functions $H_{ij}(f)$ may be obtained very simply by transforming equations (8.1.14). On substituting the $H_{ij}(f)$ in (8.4.12) and (8.4.14), explicit expressions for the auto- and cross spectra may be obtained. This procedure is best illustrated by an example.

An example. Consider the continuous bivariate ar process

$$\frac{dX_1(t)}{dt} + a_{11}X_1(t) + a_{12}X_2(t) = Z_1(t),$$
$$\frac{dX_2(t)}{dt} + a_{21}X_1(t) + a_{22}X_2(t) = Z_2(t).$$

Transforming these equations and using the differentiation property (A2.1.2) gives

$$[a_{11} + j2\pi f]X_1(f) + a_{12}X_2(f) = Z_1(f),$$
$$a_{21}X_1(f) + [a_{22} + j2\pi f]X_2(f) = Z_2(f),$$

which may be solved to give

$$X_1(f) = \frac{(a_{22} + j2\pi f)Z_1(f) - a_{12}Z_2(f)}{(a_{11} + j2\pi f)(a_{22} + j2\pi f) - a_{12}a_{21}},$$
$$X_2(f) = \frac{-a_{21}Z_1(f) + (a_{11} + j2\pi f)Z_2(f)}{(a_{11} + j2\pi f)(a_{22} + j2\pi f) - a_{12}a_{21}}.$$

Likewise, transforming (8.1.14) gives

$$X_1(f) = H_{11}(f)Z_1(f) + H_{12}(f)Z_2(f),$$
$$X_2(f) = H_{21}(f)Z_1(f) + H_{22}(f)Z_{22}(f).$$

Hence,

$$H_{11}(f) = \frac{a_{22} + j2\pi f}{D}, \qquad H_{12}(f) = -\frac{a_{12}}{D},$$
$$H_{21}(f) = \frac{-a_{21}}{D}, \qquad H_{22}(f) = \frac{a_{11} + j2\pi f}{D},$$

where $D = a_{11}a_{22} - a_{12}a_{21} - (2\pi f)^2 + j2\pi f(a_{11} + a_{22})$. Finally, using (8.4.12) and (8.4.14), the auto- and cross spectra for the bivariate process are

$$\Gamma_{11}(f) = \frac{[a_{22}^2 + (2\pi f)^2]\sigma_1^2 + a_{12}^2\sigma_2^2}{|D|^2},$$
$$\Gamma_{22}(f) = \frac{a_{21}^2\sigma_1^2 + [a_{11}^2 + (2\pi f)^2]\sigma_2^2}{|D|^2},$$
$$\Gamma_{12}(f) = \frac{-a_{22}a_{21}\sigma_1^2 - a_{11}a_{12}\sigma_2^2 - j2\pi f(a_{12}\sigma_2^2 - a_{21}\sigma_1^2)}{|D|^2}.$$

Cross spectra of discrete bivariate linear processes. Expressions for the auto- and cross spectra of discrete bivariate linear processes may be obtained in a similar way. To illustrate, consider the discrete bivariate process (8.1.20)

$$X_{1t} = 0.6X_{1t-1} - 0.5X_{2t-1} + Z_{1t},$$
$$X_{2t} = 0.4X_{1t-1} + 0.5X_{2t-1} + Z_{2t},$$

where Z_{1t}, Z_{2t} are uncorrelated white noise processes. Taking $\mathcal{Z}$ transforms,

$$X_1(\mathcal{Z}) = 0.6\mathcal{Z}^{-1}X_1(\mathcal{Z}) - 0.5\mathcal{Z}^{-1}X_2(\mathcal{Z}) + Z_1(\mathcal{Z}),$$
$$X_2(\mathcal{Z}) = 0.4\mathcal{Z}^{-1}X_1(\mathcal{Z}) + 0.5\mathcal{Z}^{-1}X_2(\mathcal{Z}) + Z_2(\mathcal{Z}).$$

These equations may be solved to give

$$X_1(\mathcal{Z}) = \frac{(1 - 0.5\mathcal{Z}^{-1})Z_1(\mathcal{Z}) - 0.5\mathcal{Z}^{-1}Z_2(\mathcal{Z})}{1 - 1.1\mathcal{Z}^{-1} + 0.5\mathcal{Z}^{-2}},$$
$$X_2(\mathcal{Z}) = \frac{0.4\mathcal{Z}^{-1}Z_1(\mathcal{Z}) + (1 - 0.6\mathcal{Z}^{-1})Z_2(\mathcal{Z})}{1 - 1.1\mathcal{Z}^{-1} + 0.5\mathcal{Z}^{-2}}.$$

Substituting $\mathcal{Z} = e^{j2\pi f}$ gives the frequency response functions

$$H_{11}(f) = \frac{(1 - 0.5e^{-j2\pi f})}{D}, \qquad H_{12}(f) = \frac{-0.5e^{-j2\pi f}}{D},$$
$$H_{21}(f) = \frac{0.4e^{-j2\pi f}}{D}, \qquad H_{22}(f) = \frac{1 - 0.6e^{-j2\pi f}}{D},$$

where

$$D = 1 - 1.1e^{-j2\pi f} + 0.5e^{-j4\pi f}$$
$$= 1 - 1.1\cos 2\pi f + 0.5\cos 4\pi f + j(1.1\sin 2\pi f - 0.5\sin 4\pi f).$$

Finally, using (8.4.12) and (8.4.14), the auto- and cross spectra for the bivariate process are

$$\Gamma_{11}(f) = \frac{\sigma_1^2(1.25 - \cos 2\pi f) + \sigma_2^2(0.25)}{|D|^2},$$
$$\Gamma_{22}(f) = \frac{\sigma_1^2(0.16) + \sigma_2^2(1.36 - 1.2\cos 2\pi f)}{|D|^2}, \tag{8.4.15}$$
$$\Gamma_{12}(f) = \frac{\sigma_1^2(-0.2 + 0.4\cos 2\pi f) + \sigma_2^2(0.3 - 0.5\cos 2\pi f) - j\sin 2\pi f\,(0.4\sigma_1^2 + 0.5\sigma_2^2)}{|D|^2},$$

where $|D|^2 = 2.46 - 3.3\cos 2\pi f + \cos 4\pi f$.

In the particular case where $\sigma_1^2 = \sigma_2^2 = 1$, (8.4.15) reduces to

$$\Gamma_{11}(f) = \frac{1.5 - \cos 2\pi f}{|D|^2},$$
$$\Gamma_{22}(f) = \frac{1.52 - 1.2\cos 2\pi f}{|D|^2}, \tag{8.4.16}$$
$$\Gamma_{12}(f) = \frac{0.1(1 - \cos 2\pi f) - j(0.9)\sin 2\pi f}{|D|^2}.$$

8.4.4 The squared coherency spectrum

It was shown in Section 8.4.2 that the correlation at frequency f between the input and output of a linear system could be described by the squared coherency $\kappa_{12}^2(f)$. This resembles a correlation coefficient at each frequency and measures the influence of the noise in the system, a large noise spectrum resulting in a low squared coherency and vice versa. It will be shown in Chapter 11 that any bivariate stochastic process has a coherency spectrum. In this section, the basic idea is illustrated by calculating the coherency spectrum of a bivariate linear process.

Consider the bivariate linear process of Figure 8.5, which has auto- and cross spectra

$$\begin{aligned}\Gamma_{11}(f) &= \sigma_1^2|H_{11}(f)|^2 + \sigma_2^2|H_{12}(f)|^2,\\ \Gamma_{22}(f) &= \sigma_1^2|H_{21}(f)|^2 + \sigma_2^2|H_{22}(f)|^2,\\ \Gamma_{12}(f) &= \sigma_1^2 H_{11}^*(f)H_{21}(f) + \sigma_2^2 H_{12}^*(f)H_{22}(f).\end{aligned} \tag{8.4.17}$$

The squared coherency spectrum of the bivariate linear process can then be obtained by substituting (8.4.17) in the definition (8.4.10). That is,

$$\kappa_{12}^2(f) = \frac{\alpha_{12}^2(f)}{\Gamma_{11}(f)\Gamma_{22}(f)} = \frac{|\Gamma_{12}(f)|^2}{\Gamma_{11}(f)\Gamma_{22}(f)}. \tag{8.4.18}$$

Some special cases of (8.4.17) and (8.4.18) are now considered.

Case 1. Suppose that $h_{12}(u) = 0$, $h_{22}(u) = 0$, so that $H_{12}(f) = 0$, $H_{22}(f) = 0$. Then

$$\kappa_{12}^2(f) = \frac{\sigma_1^4|H_{11}^*(f)H_{21}(f)|^2}{\sigma_1^4|H_{11}(f)|^2|H_{21}(f)|^2} = 1.$$

Referring to (8.1.14), if $h_{12}(u) = 0$, $h_{22}(u) = 0$,

$$X_1(t) = \int_0^\infty h_{11}(u)Z_1(t-u)\,du,$$

$$X_2(t) = \int_0^\infty h_{21}(u)Z_1(t-u)\,du.$$

Hence a squared coherency which is everywhere unity means that $X_2(t)$ could be completely recovered from $X_1(t)$. To do this, it would be necessary to convert $X_1(t)$ to white noise $Z_1(t)$ using a filter with frequency response function $1/H_{11}(f)$ and then generate $X_2(t)$ from $Z_1(t)$.

Case 2. Suppose that $h_{21}(u) = 0$, $h_{12}(u) = 0$ in (8.1.14) so that

$$X_1(t) = \int_0^\infty h_{11}(u)Z_1(t-u)\,du,$$

$$X_2(t) = \int_0^\infty h_{22}(u)Z_2(t-u)\,du.$$

From (8.4.17) it is seen that $\Gamma_{12}(f) = 0$ and hence $\kappa_{12}^2(f) = 0$. Since $Z_1(t)$ and $Z_2(t)$ are two different white noise sources, a squared coherency of zero implies that it is impossible to recover or predict $X_2(t)$ from $X_1(t)$.

Case 3 (example 2 of Section 8.4.1). Values of the squared coherency between 0 and 1 correspond to situations where $X_2(t)$ can be partially recovered or predicted from $X_1(t)$. For example, consider the bivariate process (8.4.2) for which

$$\begin{aligned}X_{1t} &= Z_{1t},\\ X_{2t} &= Z_{2t} + \beta_1 Z_{1t}.\end{aligned}$$

Hence $H_{11}(f) = 1$, $H_{22}(f) = 1$, $H_{12}(f) = 0$, $H_{21}(f) = \beta_1$ and

$$\kappa_{12}^2(f) = \frac{\beta_1^2}{1+\beta_1^2},$$

if $\sigma_1^2 = \sigma_2^2$. Thus the squared coherency spectrum is zero if $\beta_1 = 0$ and tends to unity if β_1 tends to infinity. This is to be expected, since when β_1 tends to zero, the noise Z_{2t} dominates, and when β_1 tends to infinity, the signal $\beta_1 Z_{1t}$ dominates.

Case 4 (effect of delay). Consider the bivariate process discussed in the third example of Section 8.4.1, namely,

$$\begin{aligned}X_{2t} &= Z_{2t} + \beta_1 Z_{1t-d},\\ X_{1t} &= Z_{1t}.\end{aligned}$$

When $\sigma_1^2 = \sigma_2^2$, the squared coherency spectrum is

$$\kappa_{12}^2(f) = \frac{\beta_1^2}{1+\beta_1^2},$$

which is identical to that of the preceding case.

However, it was shown in Section 8.4.1 and Figures 8.8(a) and (b) that these two processes have markedly different phase functions. Thus the squared coherency spectrum gives no indication of any phase differences between the two processes, and hence a complete frequency domain description of a bivariate process requires a phase spectrum as well as a coherency spectrum.

Case 5. Consider the discrete bivariate process (8.1.20), which has the auto- and cross spectra (8.4.16) when $\sigma_1^2 = \sigma_2^2$. The corresponding coherency spectrum is

$$\kappa_{12}^2(f) = \frac{0.42 - 0.02\cos 2\pi f - 0.4\cos 4\pi f}{2.88 - 3.32\cos 2\pi f + 0.6\cos 4\pi f}, \quad -\tfrac{1}{2} \leq f < \tfrac{1}{2}, \tag{8.4.19}$$

which is shown in Figure 8.9.

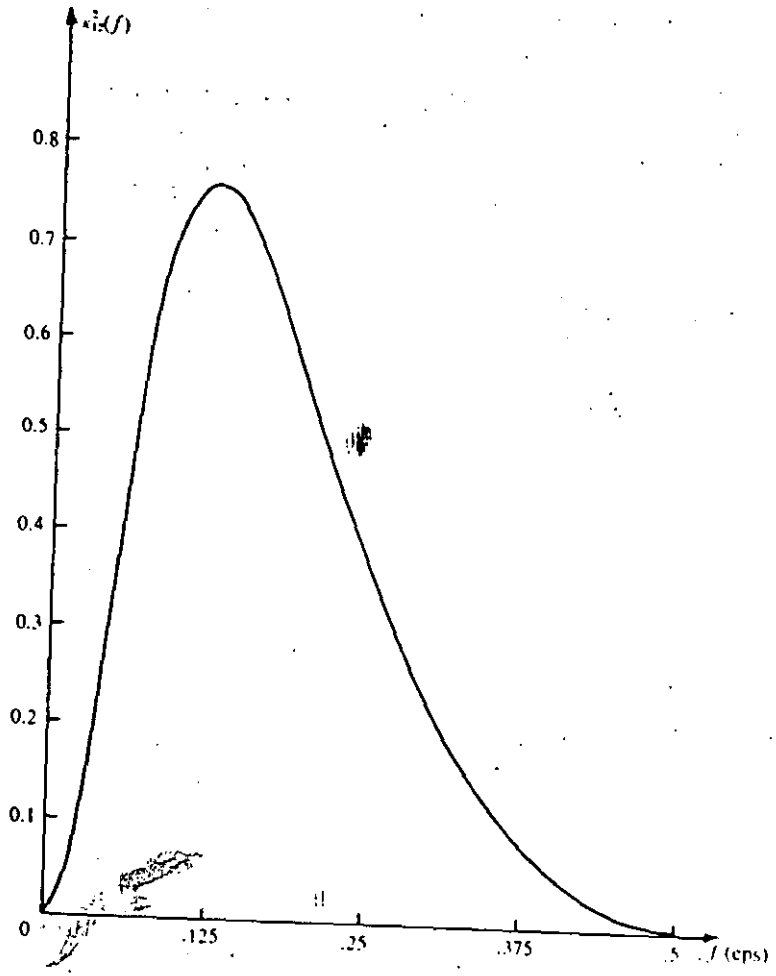


FIG. 8.9: Theoretical coherency spectrum for the bivariate autoregressive process (8.1.20)

The most important features of this coherency spectrum are the large peak at approximately 0.125 cps and the fact that the coherency tends to zero at both low and high frequencies. The peak is to be expected because of the periodicity in the ccf and shows that the bulk of the correlation between the two processes is confined to a band of frequencies in the neighborhood of 0.125 cps. Using (8.4.16), the phase spectrum for this process is

$$\phi_{12}(f) = \arctan \frac{-0.9 \sin 2\pi f}{0.1(1 - \cos 2\pi f)} = \arctan(-9 \cot \pi f), \tag{8.4.20}$$

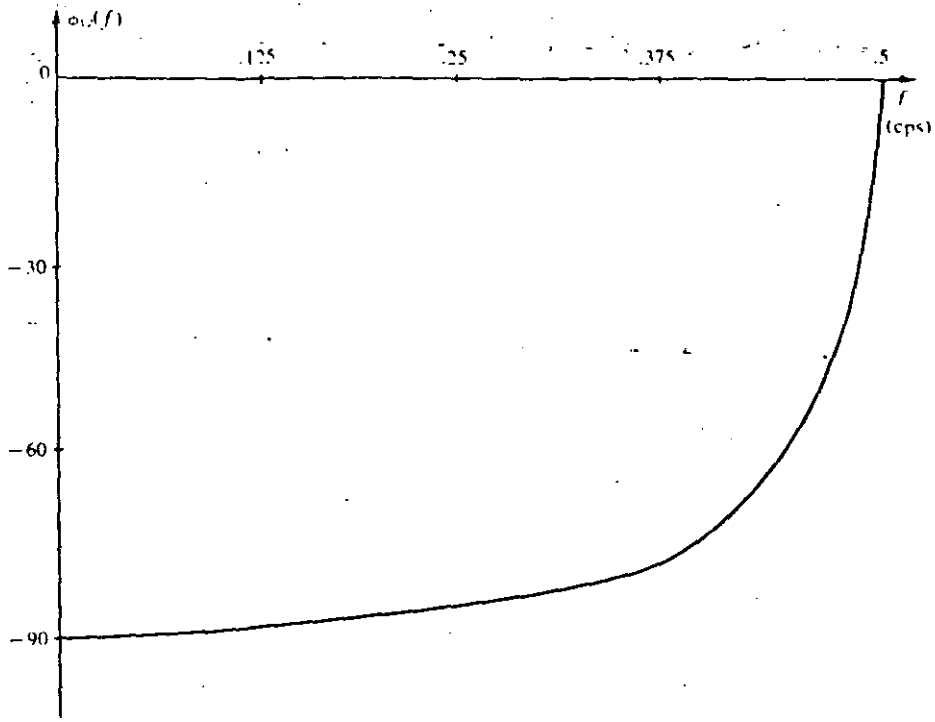


FIG. 8.10: Theoretical phase spectrum for the bivariate autoregressive process (8.1.20)

which is shown in Figure 8.10. This figure shows that the low-frequency components of series 1 lag those of series 2 by essentially 90°, but the phase difference tends to zero at higher frequencies.

Practical use of squared coherency. The coherency spectrum is useful in practice because it provides a non-dimensional measure of the correlation between two time series as a function of frequency. Thus it is to be preferred to the cross amplitude spectrum, which depends on the scale of measurement of $X_1(t)$ and $X_2(t)$. Hence the cross correlation properties of two time series can be described by the squared coherency spectrum $\kappa_{12}^2(f)$ and the phase spectrum $\phi_{12}(f)$. In Section 9.2 it is shown how to estimate these spectra from finite lengths of record.

8.4.5 Linear operations on bivariate time series

In Section 9.4.2 it will be necessary to prefilter the time series before conducting a cross spectral analysis. In this section an investigation is made of the effect of this prefiltering operation on coherency and phase spectra.

It is assumed that different filtering operations are performed on the $X_1(t)$, $X_2(t)$ processes leading to new processes $Y_1(t)$, $Y_2(t)$ according to

$$Y_1(t) = \int_0^\infty h_1(v)X_1(t - v)\,dv,$$
$$Y_2(t) = \int_0^\infty h_2(v)X_2(t - v)\,dv. \tag{8.4.21}$$

Proceeding as in Section 5.2.2, the cross covariance function of the filtered process is

$$\gamma_{Y_1Y_2}(u) = \int_0^\infty \int_0^\infty h_1(v)h_2(v')\gamma_{X_1X_2}(u + v - v')\,dv\,dv'. \tag{8.4.22}$$

On transforming (8.4.22), the cross spectrum is

$$\Gamma_{Y_1Y_2}(f) = H_1^*(f)H_2(f)\Gamma_{X_1X_2}(f). \tag{8.4.23}$$

Since

$$\Gamma_{Y_iY_i}(f) = |H_i(f)|^2\Gamma_{X_iX_i}(f), \quad i = 1, 2,$$

the coherency spectrum of the filtered bivariate series is

$$\kappa^2_{Y_1Y_2}(f) = \frac{|H_1^*(f)|^2|H_2(f)|^2\Gamma_{X_1X_2}(f)}{|H_1^*(f)|^2|H_2(f)|^2\Gamma_{X_1X_1}(f)\Gamma_{X_2X_2}(f)} = \kappa^2_{X_1X_2}(f).$$

Hence the coherency spectrum is unaltered by the filtering.

From (8.3.27) and (2.3.17),

$$\Gamma_{X_1X_2}(f) = \alpha_{X_1X_2}(f)\,e^{j\phi_{X_1X_2}(f)},$$
$$H_i(f) = G_i(f)\,e^{j\phi_i(f)},$$

and

$$\Gamma_{Y_1Y_2}(f) = G_1(f)G_2(f)\alpha_{X_1X_2}(f)\,e^{j(\phi_2(f) - \phi_1(f) + \phi_{12}(f))} = G_1(f)G_2(f)\,e^{j(\phi_2(f) - \phi_1(f))}\Gamma_{X_1X_2}(f).$$

Thus the phase spectrum is changed by $\phi_2(f) - \phi_1(f)$. However, note that if $h_1(u) = h_2(u)$, $\phi_1(f) = \phi_2(f)$ and the phase is unaltered. Hence, if the same filtering operation is applied to both processes, the cross amplitude spectrum is altered by $G^2(f)$ but the coherency and phase spectra are unaltered. The residual spectrum (see Chapter 10) for the filtered data will be affected in the same way as the autospectrum $\Gamma_{Y_2Y_2}(f)$, that is, the residual spectrum for the filtered data equals $G_2^2(f)$ times the residual spectrum for the original data.

APPENDIX A8.1 REALIZATIONS OF TWO BIVARIATE LINEAR PROCESSES

TABLE A8.1: A realization of a bivariate linear process

$x_{1t} = 0.6x_{1t-1} - 0.5x_{2t-1} + z_{1t}$				$x_{2t} = 0.4x_{1t-1} + 0.5x_{2t-1} + z_{2t}$			
t = 1–25	t = 26–50	t = 51–75	t = 76–100	t = 1–25	t = 26–50	t = 51–75	t = 76–100
−0.88	−2.00	0.37	0.15	0.79	−1.40	2.96	2.69
−0.16	−0.22	−0.24	−1.04	1.12	−2.55	1.56	0.57
−1.87	0.38	0.57	0.12	−1.10	−1.66	−0.36	0.29
−1.12	1.31	−0.53	0.08	−2.39	−0.43	−0.59	1.10
1.38	0.71	2.44	0.11	−1.75	0.58	−0.12	0.48
2.13	0.32	1.02	−2.62	−0.82	2.18	3.03	−1.06
2.76	0.48	−0.53	−1.28	−0.36	−0.24	2.11	−2.28
0.56	−1.88	−2.49	1.07	1.27	0.58	0.78	−2.03
−0.69	−0.94	−2.12	3.20	1.75	−0.18	0.89	−0.75
−1.79	−1.54	−1.04	1.92	2.44	−1.55	−1.45	1.00
−3.82	−0.13	−0.12	0.53	0.36	−0.64	−0.36	1.71
−2.38	1.02	−1.88	−1.08	−2.10	−1.09	−0.37	0.58
1.00	0.02	−1.50	0.49	−1.93	0.90	−1.39	1.97
0.70	−0.77	1.54	−0.58	−1.30	−0.66	−4.19	0.99
−0.15	0.11	3.33	0.17	−1.75	−0.35	−0.73	1.94
0.98	−0.60	3.08	1.15	−0.34	0.48	−0.98	2.18
0.11	−0.52	1.71	−0.97	0.74	0.50	0.36	3.14
−0.35	−0.09	0.79	−1.63	0.49	0.05	0.06	0.60
−0.73	1.23	1.55	1.14	0.70	−0.68	−1.94	0.51
0.89	1.46	0.89	−0.67	0.71	0.24	−0.08	1.35
−1.63	0.61	−0.89	−0.88	0.09	0.58	0.17	0.56
−0.44	0.42	−1.18	−0.07	0.59	−1.26	1.00	0.11
−1.37	2.16	0.89	0.24	1.54	−0.25	−0.05	0.00
−1.71	3.18	1.71	0.55	0.14	0.25	0.43	2.34
−1.22	2.10	3.05	−2.16	0.55	2.18	0.15	1.88

TABLE A8.2: A realization of a bivariate linear process with delay

$x_{1t} = 0.6x_{1t-1} + z_{1t}$ $\quad x_{2t} = 0.5x_{2t-1} + 2x_{1t-10} + y_{2t}$, $\quad y_{2t} = 0.5y_{2t-1} + z_{2t}$

t = 1–25	t = 26–50	t = 51–75	t = 76–100	t = 1–25	t = 26–50	t = 51–75	t = 76–100
−2.07	−0.78	0.00	0.20	0.32	−3.65	8.00	4.88
−1.15	0.31	−1.99	−0.42	0.35	−3.38	7.91	3.01
0.69	−0.95	−1.75	1.18	−2.03	−3.04	8.81	6.05
−0.46	−0.90	−0.70	0.82	−4.16	−3.03	2.36	5.67
−1.49	−0.30	0.73	1.50	−3.55	−2.64	5.92	7.23
−0.70	−1.02	1.16	2.92	−1.04	−1.01	2.94	2.32
−1.07	−0.53	0.06	1.18	2.60	2.88	2.05	1.27
−0.69	0.15	−0.02	1.23	0.67	−1.18	−1.52	1.26
−0.68	1.40	1.10	3.16	−0.25	−1.91	−4.87	4.86
1.27	1.22	−0.35	0.79	−0.90	−3.75	−2.96	4.75
−1.05	0.59	−1.67	0.68	−4.69	−3.61	−2.26	3.55
−0.05	0.70	−1.57	1.14	−3.50	−3.08	−4.23	1.50
−0.84	1.70	1.16	1.02	1.62	−4.18	−5.26	3.37
−0.62	2.78	1.84	1.02	0.81	−4.75	−0.96	5.69
−0.49	1.98	3.35	−0.71	−0.95	−2.62	0.37	7.52
−1.29	1.39	0.40	−0.17	−2.24	−2.15	3.58	10.32
−0.49	1.85	0.45	−1.50	−4.50	−1.61	1.63	8.41
−1.06	2.60	1.30	−0.26	−4.55	−1.28	1.05	7.55
−0.38	0.51	0.93	−0.38	−3.85	1.14	3.77	10.38
−0.52	2.77	1.17	0.93	0.78	2.89	1.60	9.14
−0.13	1.16	−1.74	−0.33	−0.02	4.68	−3.67	6.93
1.30	1.07	−1.28	−1.12	−0.72	4.94	−6.24	6.54
−1.51	−0.48	−0.07	−2.95	−1.84	6.41	−1.82	4.13
−0.43	−0.52	1.50	−2.09	−1.78	10.54	2.11	3.49
−1.33	0.37	0.53	−1.11	−2.77	9.06	7.51	0.40

9

Estimation of Cross Spectra

It is shown in Section 9.1 that the sample cross spectrum has the same undesirable property as the sample spectrum, namely that its variance is independent of record length. However, it can be used to construct a frequency-domain test for cross correlation between time series based on the sample integrated cospectrum and the sample phase spectrum. In Section 9.2 expressions are derived for the variances and covariances of the smoothed co- and quadrature spectra and the smoothed coherency and phase spectra. It is shown that these depend on the uncontrollable influence of the theoretical coherency as well as the controllable influence of the smoothing procedure.

Some numerical examples of the estimation of cross spectra are given in Section 9.3, where it is shown that very large biases occur if the cross correlation function does not have a maximum at zero lag. Theoretical analysis of the bias shows that it can be minimized by displacing or aligning the series so that the cross correlation function of the aligned series has a maximum at zero lag. Section 9.4 includes a practical procedure for estimating cross spectra and a practical example.

9.1 PROPERTIES OF THE SAMPLE CROSS SPECTRUM

9.1.1 Moments of the sample cross spectrum for two uncorrelated white noise processes

In this section the means, variances and covariances of the sample co- and quadrature spectral estimators and the sample phase and cross amplitude spectral estimators are derived, under the assumption that the two processes are uncorrelated white noise processes. These results will be useful in two contexts. In Section 9.1.2 they will be used for deriving a test for correlation between two time series, and in Sections 9.1.3 and 9.2.1 they will be used for deriving

the moments of the sample and smoothed cross spectral estimators under fairly general assumptions about the stochastic processes $X_1(t)$ and $X_2(t)$.

As in previous work, the two white noise processes are denoted by $Z_1(t)$, $Z_2(t)$ and are assumed to have zero means. The sample Fourier transforms are denoted by

$$\begin{aligned} Z_i(f) &= \int_{-T/2}^{T/2} Z_i(t)\, e^{-j2\pi ft}\, dt \\ &= A_i(f) - jB_i(f), \end{aligned}$$

where $A_i(f)$, $B_i(f)$ are the cosine and sine transforms of $Z_i(t)$. On dropping the dependence on f in the cosine and sine transforms, the sample auto- and cross spectra are

$$C_{ii}(f) = \frac{|Z_i(f)|^2}{T} = \frac{A_i^2 + B_i^2}{T}, \quad i = 1, 2,$$

$$C_{12}(f) = \frac{|Z_1^*(f)Z_2(f)|}{T} \tag{9.1.1}$$

$$= \frac{1}{T}[(A_1 + jB_1)(A_2 - jB_2)]$$

$$= \frac{1}{T}[(A_1A_2 + B_1B_2) - j(B_2A_1 - B_1A_2)]. \tag{9.1.2}$$

Hence the sample co- and quadrature spectra are

$$L_{12}(f) = \frac{1}{T}(A_1A_2 + B_1B_2), \tag{9.1.3}$$

$$Q_{12}(f) = \frac{1}{T}(B_2A_1 - B_1A_2). \tag{9.1.4}$$

Now it has been shown in Section 6.3.1 that if the $Z_i(t)$ are Normal processes, the A_i and B_i are Normal rv's. It was also shown that if the processes $Z_i(t)$ have zero means,

$$E[A_i] = E[B_i] = 0, \tag{9.1.5}$$

and for the harmonic frequencies $f_m = m/T$,

$$\begin{aligned} \text{Var}\,[A_i] = \text{Var}\,[B_i] &= \frac{T}{2}\sigma_i^2, \\ \text{Cov}\,[A_i, B_i] &= 0, \quad i = 1, 2. \end{aligned} \tag{9.1.6}$$

If, in addition, the two processes $Z_1(t)$ and $Z_2(t)$ are uncorrelated,

$$\begin{aligned} &\text{Cov}\,[A_1, A_2] = 0 = \text{Cov}\,[B_1, B_2], \\ &\text{Cov}\,[A_1, B_2] = 0 = \text{Cov}\,[B_1, A_2]. \end{aligned} \tag{9.1.7}$$

Moments of sample co- and quadrature spectra. Using these results it is possible to derive the moments of the sample co- and quadrature spectra. For example, using (9.1.3) and (9.1.7),

$$E[L_{12}(f)] = \frac{1}{T}\{E[A_1A_2] + E[B_1B_2]\} = 0,$$

and from (9.1.6)

$$\begin{aligned} \text{Var}\,[L_{12}(f)] &= \frac{1}{T^2}E[A_1^2A_2^2 + B_1^2B_2^2 + 2B_1B_2A_1A_2] \\ &= \frac{\sigma_1^2}{2}\frac{\sigma_2^2}{2} + \frac{\sigma_1^2}{2}\frac{\sigma_2^2}{2} + 0 \\ &= \frac{\sigma_1^2\sigma_2^2}{2}. \end{aligned}$$

Similarly,

$$\begin{aligned} E[Q_{12}(f)] &= 0, \\ \text{Var}\,[Q_{12}(f)] &= \frac{\sigma_1^2\sigma_2^2}{2}, \\ \text{Cov}\,[L_{12}(f), Q_{12}(f)] &= 0. \end{aligned}$$

It may also be shown that $L_{12}(f)$ and $Q_{12}(f)$ are uncorrelated with $C_{11}(f)$ and $C_{22}(f)$. Hence the joint distribution of the estimators $C_{11}(f)$, $C_{22}(f)$, $L_{12}(f)$ and $Q_{12}(f)$ may be characterized by the *covariance matrix*

$$\begin{pmatrix} \sigma_1^2 & 0 & 0 & 0 \\ 0 & \sigma_2^2 & 0 & 0 \\ 0 & 0 & \frac{1}{2}\sigma_1^2\sigma_2^2 & 0 \\ 0 & 0 & 0 & \frac{1}{2}\sigma_1^2\sigma_2^2 \end{pmatrix}. \tag{9.1.8}$$

Distribution of the sample cross amplitude spectral estimator. The sample cross amplitude spectrum is defined as $A_{12}(f) = |C_{12}(f)|$, so that, using (9.1.1), the square of the cross amplitude spectral estimator is

$$\begin{aligned} A_{12}^2(f) &= |C_{12}(f)|^2 \\ &= L_{12}^2(f) + Q_{12}^2(f) \\ &= \left(\frac{|Z_1(f)|^2}{T}\right)\left(\frac{|Z_2(f)|^2}{T}\right) \\ &= C_{11}(f)C_{22}(f). \end{aligned}$$

It is convenient now to introduce the rv

$$Y^2(f) = \frac{4A_{12}^2(f)}{\sigma_1^2\sigma_2^2} = \left(\frac{2C_{11}(f)}{\sigma_1^2}\right)\left(\frac{2C_{22}(f)}{\sigma_2^2}\right) = UV.$$

Making use of the chi-squared property of $C_{11}(f)$ discussed in Section 6.3.3 and the independence property of the two processes $Z_1(t)$, $Z_2(t)$, it follows

that the rv $Y^2(f)$ is the product of two independent chi-squared rv's U, V, each with two degrees of freedom. Hence

$$E[Y^2(f)] = E[U]E[V] = 4,$$
$$E[Y^4(f)] = E[U^2]E[V^2] = (8)(8) = 64,$$

using (3.3.6). It follows that

$$\text{Var}\,[Y^2(f)] = 48,$$
$$E[A_{12}^2(f)] = \left(\frac{\sigma_1^2\sigma_2^2}{4}\right) E[Y^2(f)] = \sigma_1^2\sigma_2^2, \tag{9.1.9}$$

and

$$\text{Var}\,[A_{12}^2(f)] = \left(\frac{\sigma_1^4\sigma_2^4}{16}\right) \text{Var}\,[Y^2(f)] = 3\sigma_1^4\sigma_2^4.$$

Note that whereas the variance of the sample spectrum is equal to the square of its mean, the variance of the sample cross amplitude spectrum is three times the square of its mean. The increase in variance is due to the fact that variability is now being introduced by two processes rather than one.

Distribution of the sample phase spectral estimator. Using (9.1.3) and (9.1.4) the sample phase spectral estimator is

$$F_{12}(f) = \arctan\left\{-\frac{Q_{12}(f)}{L_{12}(f)}\right\}$$
$$= \arctan\left\{-\frac{B_2A_1 - B_1A_2}{A_1A_2 + B_1B_2}\right\}.$$

Now consider the random variable $L_{12}(f)$. The rv's A_t, B_t are Normally distributed, so that the range of the rv $A_1A_2 + B_1B_2$ extends from $-\infty$ to $+\infty$ and hence it is reasonable to approximate its distribution by a Normal distribution. Similar considerations apply for the rv $Q_{12}(f)$. Thus $L_{12}(f)$, $Q_{12}(f)$ are approximately Normally and independently distributed and have the same variance. Hence $F_{12}(f)$ is approximately uniformly distributed in the range $-\pi/2$ to $\pi/2$. These results will be used in the next section to derive a test for correlation between two time series.

9.1.2 *A test for correlation between time series*

Situations often arise where it is required to test whether two time series are correlated or uncorrelated. For example, it may be necessary to test whether two control variables are correlated, or whether the residuals are correlated after a suitable model has been fitted to two economic time series. The results of Section 8.2.2 show that, provided both time series have been prefiltered to convert them to white noise, the sample ccf of the filtered series can be used to test whether the two series are correlated. However, the ccf is

useful only in detecting certain types of correlation. For instance, if neighboring points in two time series are cross correlated, it would be expected that the ccf will be large in the neighborhood of the origin and small at values distant from the origin. On the other hand, if there is a tendency for the ccf to contain periodic components, these may not be detected using the ccf. Hence it is necessary to develop a frequency-domain test for correlation between time series which is a generalization of the test for white noise given in Section 6.3.2. This frequency-domain test should be used in conjunction with the test based on the ccf.

Choice of test criteria. The discussion in Section 8.4.4 suggests that the sample coherency spectrum $K_{12}^2(f)$ and the sample phase spectrum $F_{12}(f)$ could be used as the basis for a frequency-domain test for correlation between time series. However, it is noted that

$$K_{12}^2(f) = \frac{|C_{12}(f)|^2}{C_{11}(f)C_{22}(f)} = \frac{\left|\frac{X_1^*(f)X_2(f)}{T}\right|^2}{\frac{|X_1(f)|^2}{T}\frac{|X_2(f)|^2}{T}} = 1.$$

Thus the sample coherency spectrum is identically equal to unity regardless of the nature of the bivariate stochastic process. Hence an alternative approach is required, based on the integrated sample co-spectrum and the sample phase spectrum. These two criteria focus attention on two different aspects of the cross correlation between the two processes.

(*1*) *The integrated sample co-spectrum.* Consider the integrated co-spectrum

$$J_{12}(f) = \int_{-f}^{f} \Lambda_{12}(g)\,dg,$$

which measures the total in-phase covariance between the two processes for all frequencies less than f. Then an estimator of $J_{12}(f)$ is given by the sample integrated co-spectrum

$$\frac{1}{N\Delta}\sum_{t=0}^{k} L_{12}(f_t), \quad f_t = \frac{2t}{N}.$$

However, it is more convenient to use the normalized estimator

$$\hat{J}_{12}(f_k) = \frac{2}{NS_1S_2\Delta}\sum_{t=0}^{k} L_{12}(f_t), \tag{9.1.10}$$

where S_1, S_2 are the estimators of the standard deviations of the two processes. If the two processes are uncorrelated, then $\Lambda_{12}(f)$ is identically zero and hence $J_{12}(f_k)$ is identically zero, but if the two processes are correlated,

$J_{12}(f)$ takes on non-zero values. Using (8.3.16), the normalization in (9.1.10) ensures that $\hat{J}_{12}(f_k) = r_{12}(0)$ when $f_k = 1/2\Delta$ cps.

Figure 9.1 shows the estimate corresponding to (9.1.10) for samples of $N = 100$ from three bivariate Normal processes. These processes were of the form

$$X_{1t} = Z_{1t} + aZ_{2t},$$
$$X_{2t} = aZ_{1t} + Z_{2t},$$

with Z_1, Z_2 random Normal deviates and $a = 0$, 0.1 and 0.3. Thus the cross correlations are zero for all non-zero lags, but $\rho_{12}(0) = 0$, 0.20 and 0.55 respectively for the three cases. Figure 9.1 shows that when $\rho_{12} = 0$ the integrated co-spectrum fluctuates about zero and indicates no correlation between the two processes. When $\rho_{12} = 0.20$, the integrated spectrum rises

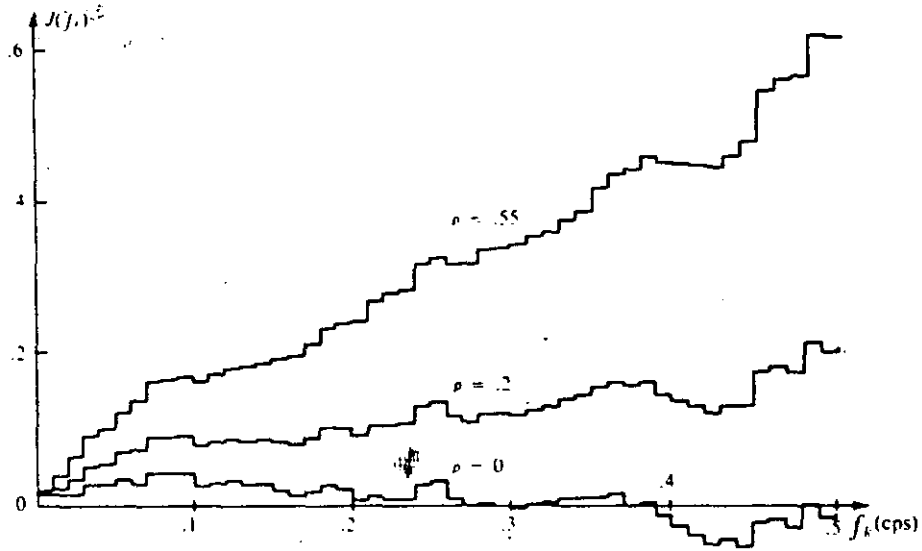


FIG. 9.1: Sample integrated co-spectra for three bivariate correlated series

steadily to the value 0.20 at $f = 0.5$ cps and shows clear evidence of correlation. Since $\hat{J}(1/2\Delta) = \hat{r}_{12}(0)$, the value of 0.20 at $f = 0.5$ cps provides an excellent estimate of $\rho_{12}(0)$, which is known to be 0.20. The curve for $\rho_{12} = 0.55$ shows that there is marked correlation between the two series. Thus the behavior of the three curves in Figure 9.1 confirms that the test is a sensitive indicator of correlation between time series. The value of the sample co- and quadrature spectra when $\rho = 0$ are given in Table 9.1 and could be used by the reader to generate the integrated co-spectrum of Figure 9.1.

(2) *The phase spectrum.* Another indication of the correlation between two time series is provided by the phase spectral estimator $F_{12}(f)$. It was shown in Section 9.1 that if the two processes are uncorrelated, the sample

TABLE 9.1: Sample co- and quadrature spectra for two uncorrelated white noise processes

f_k	$L_{12}(f_k)$	$Q_{12}(f_k)$	$F_{12}(f_k)$	f_k	$L_{12}(f_k)$	$Q_{12}(f_k)$	$F_{12}(f_k)$
0.01	0.08	1.53	1.52	0.26	−1.20	−0.47	0.39
0.02	0.03	−1.44	−1.35	0.27	−0.38	0.08	−0.22
0.03	0.71	−0.67	−0.75	0.28	0.22	1.24	1.40
0.04	−0.10	−0.78	1.44	0.29	−0.21	0.05	−0.22
0.05	−0.39	−1.17	−1.25	0.30	−0.15	−0.45	1.25
0.06	−0.19	0.38	−1.10	0.31	0.26	0.08	0.32
0.07	0.63	0.81	0.91	0.32	0.12	−0.14	−0.89
0.08	0.00	−0.04	1.47	0.33	0.37	−0.52	−0.96
0.09	−0.03	0.28	−1.47	0.34	0.02	−0.48	−1.54
0.10	−0.79	−0.15	0.19	0.35	0.01	1.11	1.56
0.11	0.17	0.43	1.19	0.36	0.16	−0.60	−1.32
0.12	−0.15	−0.35	−1.17	0.37	−0.86	−0.54	0.56
0.13	−0.27	0.18	−0.57	0.38	0.39	0.31	0.67
0.14	0.09	−0.16	−1.09	0.39	−1.05	0.29	−0.27
0.15	−0.40	0.79	−1.10	0.40	−0.63	0.78	−0.89
0.16	−0.42	0.78	−1.08	0.41	−0.50	−0.45	0.73
0.17	0.21	0.31	0.97	0.42	−0.53	0.18	−0.32
0.18	0.49	0.04	0.08	0.43	0.23	0.03	0.11
0.19	−0.15	−0.76	1.38	0.44	−0.41	2.42	−1.40
0.20	−0.79	1.19	−0.98	0.45	1.71	−1.00	−0.53
0.21	0.23	1.49	1.42	0.46	0.04	1.23	1.54
0.22	−0.08	−0.31	1.31	0.47	−0.52	−0.25	0.45
0.23	0.02	0.04	1.06	0.48	1.39	1.36	0.77
0.24	0.90	0.61	0.60	0.49	−0.67	0.08	−0.12
0.25	0.21	0.21	0.78	0.50	0.10	−0.91	0.00

phase spectrum will be approximately uniformly distributed in the range $-\pi/2$, $\pi/2$. Hence the cumulative distribution function of the phase angle will be a straight line in this range.

The numerical values of the phase estimates for the two white noise series with $\rho = 0$ are shown in Table 9.1 alongside the co- and quadrature spectral estimates. The sample cdf of the phase estimates is shown in Figure 9.2, and it is seen that there is good agreement between the actual and theoretical cdf. To guide the eye in deciding whether deviations from linearity are real, 95% confidence limits can be inserted at distances $\pm 1.36/\sqrt{N/2}$ and 75% limits at $\pm 1.02/\sqrt{N/2}$ from the theoretical cdf.

It is seen that the empirical cdf lies well within these limits when $\rho_{12} = 0$. When $\rho_{12} = 0.20$ and 0.55, the cumulative sample phase spectra also lie near the theoretical straight line. Thus when $\rho_{12} = 0$, both the integrated sample co-spectrum and the phase spectrum tests show no evidence of correlation. When $\rho_{12} = 0.20$ and 0.55, the co-spectrum test shows evidence of correlation but the phase spectrum test does not, confirming that the

theoretical phase spectrum is zero. In general, of course, a correlated bivariate process would have a non-zero co-spectrum and a non-zero phase spectrum and hence a realization of such a series would be expected to "fail" both the integrated co-spectrum and phase spectrum tests.

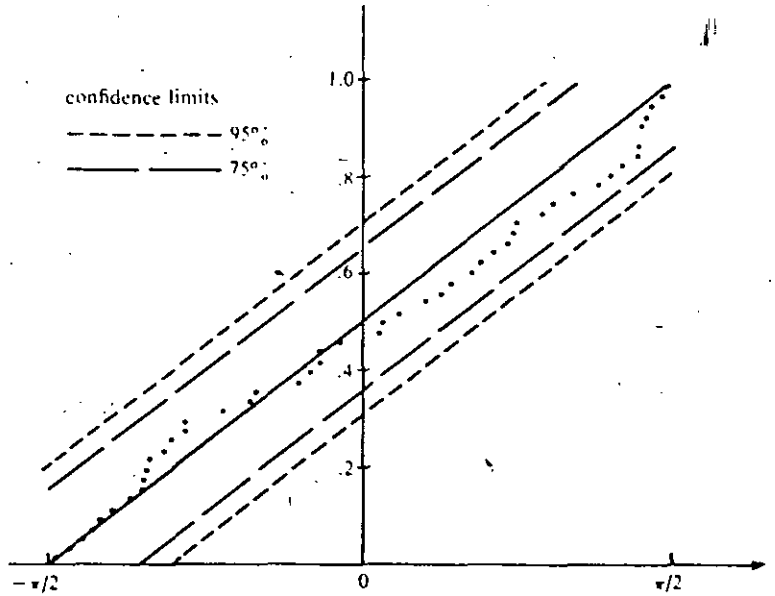


FIG. 9.2: Sample integrated phase spectrum for a bivariate series

9.1.3 General moment properties of sample cross spectral estimators

In this section the results of Section 9.1.1 are generalized to include correlated non-white noise processes. The precise derivation of these results is complicated and is given in Appendix A9.1. The methods used in this section are heuristic and are generalizations of the methods used for univariate spectra in Section 6.4.1.

The computation of the covariance matrix of $C_{11}(f)$, $C_{22}(f)$, $L_{12}(f)$ and $Q_{12}(f)$ for correlated non-Normal processes proceeds in three stages. First, the matrix of covariances between these estimators at two distinct frequencies is evaluated for uncorrelated, non-Normal white noise processes. Then the covariance matrix of $C_{11}(f)$, $C_{22}(f)$, $C_{12}(f)$, $C_{21}(f)$ may be determined. Finally, the values of $C_{11}(f)$, $C_{22}(f)$, $C_{12}(f)$, $C_{21}(f)$ for any bivariate process may be expressed in terms of $C_{11}(f)$, $C_{22}(f)$, $C_{12}(f)$ and $C_{21}(f)$ for uncorrelated white noise processes. Using the covariance matrix derived in the second step, the covariance matrix of the general process may then be calculated.

Generalized covariance matrix of cross spectral estimators for uncorrelated white noises. The results of Section 9.1.1 were derived for the *harmonic frequencies* $f_m = m/T$ for uncorrelated *Normal* white noise processes. In Appendix A9.1 more general results are derived which apply for *all frequencies* and for uncorrelated *non-Normal white* noise processes. These results show that

$$E[L_{12}(f)] = \Lambda_{12}(f) = 0, \qquad E[Q_{12}(f)] = \Psi_{12}(f) = 0. \tag{9.1.11}$$

Further, the matrix

$$\begin{pmatrix} W_0K_4^{(1)} + \sigma_1^4\{W^2(-) + W^2(+)\} & 0 & 0 & 0 \\ 0 & W_0K_4^{(2)} + \sigma_2^4\{W^2(-) + W^2(+)\} & 0 & 0 \\ 0 & 0 & \frac{\sigma_1^2\sigma_2^2}{2}\{W^2(-) + W^2(+)\} & 0 \\ 0 & 0 & 0 & \frac{\sigma_1^2\sigma_2^2}{2}\{W^2(-) + W^2(+)\} \end{pmatrix} \tag{9.1.12}$$

gives the covariances between the estimators $C_{11}(f)$, $C_{22}(f)$, $L_{12}(f)$, $Q_{12}(f)$ at two adjacent frequencies f_1 and f_2. For example, the element in the first row and the first column is Cov $[C_{11}(f_1), C_{11}(f_2)]$, the element in the first row and second column is Cov $[C_{11}(f_1), C_{22}(f_2)]$ and so on. The matrix (9.1.12) is called the *generalized covariance matrix* of the estimators. When $f_1 = f_2 = f$, this reduces to the ordinary covariance matrix between $C_{11}(f)$, $C_{22}(f)$, $L_{12}(f)$ and $Q_{12}(f)$. In (9.1.12), $W_0 = \Delta^2/N$, and

$$W(-) = \frac{\Delta \sin \pi N\Delta(f_1 - f_2)}{N \sin \pi\Delta(f_1 - f_2)}, \quad W(+) = \frac{\Delta \sin \pi N\Delta(f_1 + f_2)}{N \sin \pi\Delta(f_1 + f_2)} \tag{9.1.13}$$

for discrete processes. Similarly, $W_0 = 1/T$ and

$$W(-) = \frac{\sin \pi T(f_1 - f_2)}{\pi T(f_1 - f_2)}, \quad W(+) = \frac{\sin \pi T(f_1 + f_2)}{\pi T(f_1 + f_2)} \tag{9.1.14}$$

in the continuous case. $K_4^{(1)}$ and $K_4^{(2)}$ are the fourth cumulants of the white noise processes $Z_1(t)$ and $Z_2(t)$ respectively. The fourth cumulants vanish if the $Z_i(t)$ are Normal.

The matrix (9.1.12) shows that the spectral estimators are uncorrelated with the other spectral estimators at the same or adjacent frequencies. Furthermore, they are uncorrelated with themselves for sufficiently wide frequency spacing. Note that when $f_1 = f_2$, the variances of the estimators are constants which do not depend on T, the length of the record. Hence, cross spectral estimators show the same bad behavior as autospectral estimators. Note also that the covariances of the co- and quadrature spectra do not involve fourth cumulant terms, so the covariances of the co- and quadrature spectra are always of order $1/T^2$ for $(f_1 - f_2)$ large.

Generalized covariance matrix of $C_{11}(f)$, $C_{22}(f)$, $C_{12}(f)$, $C_{21}(f)$. Since

$$C_{12}(f) = L_{12}(f) - jQ_{12}(f),$$
$$C_{21}(f) = L_{12}(f) + jQ_{12}(f),$$

the covariance matrix of C_{11}, C_{22}, C_{12} and C_{21} is readily derived from the matrix (9.1.12). For example, using elements of (9.1.12),

$$\begin{aligned}\text{Cov}\,[C_{12}(f_1), C_{21}(f_2)] &= E[(L_2(f_1) - jQ_{12}(f_1))(L_{12}(f_2) + jQ_{12}(f_2))] \\ &= \text{Cov}\,[L_{12}(f_1), L_{12}(f_2)] + \text{Cov}\,[Q_{12}(f_1), Q_{12}(f_2)] \\ &= \sigma_1^2\sigma_2^2 W^2(-).\end{aligned}$$

The final result for the generalized covariance matrix of $C_{11}(f)$, $C_{22}(f)$, $C_{12}(f)$, $C_{21}(f)$ is

$$\begin{pmatrix} W_0K_4^{(1)} + \sigma_1^4\{W^2(-)+W^2(+)\} & 0 & 0 & 0 \\ 0 & W_0K_4^{(2)} + \sigma_2^4\{W^2(-)+W^2(+)\} & 0 & 0 \\ 0 & 0 & \sigma_1^2\sigma_2^2W^2(+) & \sigma_1^2\sigma_2^2W^2(-) \\ 0 & 0 & \sigma_1^2\sigma_2^2W^2(-) & \sigma_1^2\sigma_2^2W^2(+) \end{pmatrix} \tag{9.1.15}$$

which will now be used to derive the covariance matrix of the cross spectral estimators for non-white noise processes. Note that unless f_1 and f_2 are very close to one another, all the covariances are approximately zero.

Generalized covariance matrix of cross spectral estimators for general processes. Use is now made of the result given in Section 8.1.4 that any bivariate stochastic process with power spectra $\Gamma_{11}(f)$, $\Gamma_{22}(f)$, $\Gamma_{12}(f)$ can be generated by passing two white noise sources through a lattice network of four linear systems. Thus, on taking Fourier transforms of the equations (8.1.14) and using the same approximations as in (6.4.3),

$$X_1(f) \approx H_{11}(f)Z_{1T}(f) + H_{12}(f)Z_{2T}(f),$$
$$X_2(f) \approx H_{21}(f)Z_{1T}(f) + H_{22}(f)Z_{2T}(f),$$

where $Z_{iT}(f)$ is the Fourier transform of $Z_i(t)$ over $-T/2 \leqslant t \leqslant T/2$. Hence

$$\begin{aligned}C_{11}(f) \approx\ & |H_{11}(f)|^2C_{z_1z_1}(f) + |H_{12}(f)|^2C_{z_2z_2}(f) \\ & + H_{11}^*(f)H_{12}(f)C_{z_1z_2}(f) + H_{11}(f)H_{12}^*(f)C_{z_2z_1}(f),\end{aligned} \tag{9.1.16}$$

$$\begin{aligned}C_{22}(f) \approx\ & |H_{21}(f)|^2C_{z_1z_1}(f) + |H_{22}(f)|^2C_{z_2z_2}(f) \\ & + H_{21}^*(f)H_{22}(f)C_{z_1z_2}(f) + H_{21}(f)H_{22}^*(f)C_{z_2z_1}(f),\end{aligned} \tag{9.1.17}$$

$$\begin{aligned}C_{12}(f) \approx\ & H_{11}^*(f)H_{21}(f)C_{z_1z_1}(f) + H_{12}^*(f)H_{22}(f)C_{z_2z_2}(f) \\ & + H_{11}^*(f)H_{22}(f)C_{z_1z_2}(f) + H_{12}^*(f)H_{21}(f)C_{z_2z_1}(f).\end{aligned} \tag{9.1.18}$$

Hence, from (8.4.17),

$$\begin{aligned}E[C_{11}(f)] &\approx |H_{11}(f)|^2\sigma_1^2 + |H_{12}(f)|^2\sigma_2^2 \\ &= \Gamma_{11}(f),\end{aligned} \tag{9.1.19}$$

$$\begin{aligned}E[C_{22}(f)] &\approx |H_{21}(f)|^2\sigma_1^2 + |H_{22}(f)|^2\sigma_2^2 \\ &= \Gamma_{22}(f),\end{aligned} \tag{9.1.20}$$

$$\begin{aligned}E[C_{12}(f)] &\approx H_{11}^*(f)H_{21}(f)\sigma_1^2 + H_{12}^*(f)H_{22}(f)\sigma_2^2 \\ &= \Gamma_{12}(f).\end{aligned} \tag{9.1.21}$$

The covariance matrix for the $C_{ij}(f)$ is readily obtained from the definitions (9.1.16–18) and the covariance matrix (9.1.15). For example, using (9.1.16) and (9.1.17) and remembering that the covariance matrix is approximately zero unless $f_1 \approx f_2$,

$$\begin{aligned}\text{Cov}\,[C_{11}(f_1), C_{22}(f_2)] \approx\ & |H_{11}(f_1)|^2|H_{21}(f_2)|^2\sigma_1^4\{W^2(-) + W^2(+)\} \\ & + |H_{12}(f_1)|^2|H_{22}(f_2)|^2\sigma_2^4\{W^2(-) + W^2(+)\} \\ & + H_{11}^*(f_1)H_{12}(f_1)H_{21}^*(f_2)H_{22}(f_2)\sigma_1^2\sigma_2^2W^2(+) \\ & + H_{11}^*(f_1)H_{12}(f_1)H_{21}(f_2)H_{22}^*(f_2)\sigma_1^2\sigma_2^2W^2(-) \\ & + H_{11}(f_1)H_{12}^*(f_1)H_{21}^*(f_2)H_{22}(f_2)\sigma_1^2\sigma_2^2W^2(-) \\ & + H_{11}(f_1)H_{12}^*(f_1)H_{21}(f_2)H_{22}^*(f_2)\sigma_1^2\sigma_2^2W^2(+).\end{aligned}$$

The $W^2(+)$ terms are small compared to the $W^2(-)$ terms and may be neglected. Thus, making the approximation $H_{ij}(f_2) \approx H_{ij}(f_1)$ for $f_1 \approx f_2$ and writing $H_{ij}(f_1)$ as H_{ij} gives

$$\begin{aligned}\text{Cov}\,[C_{11}(f_1), C_{11}(f_2)] &\approx \{|H_{11}|^2|H_{21}|^2\sigma_1^4 + |H_{12}|^2|H_{22}|^2\sigma_2^4 \\ &\quad + (H_{11}^*H_{12}H_{21}H_{22}^* + H_{11}H_{12}^*H_{21}^*H_{22})\sigma_1^2\sigma_2^2\}W^2(-) \\ &= |\Gamma_{12}(f_1)|^2W^2(-).\end{aligned}$$

Therefore, on dropping the dependence on f, the generalized covariance matrix of C_{11}, C_{22}, L_{12} and Q_{12} at frequencies f_1 and f_2 for a general process is

$$W^2(-)\begin{pmatrix} \Gamma_{11}^2 & |\Gamma_{12}|^2 & \Gamma_{11}\Lambda_{12} & \Gamma_{11}\Psi_{12} \\ |\Gamma_{12}|^2 & \Gamma_{22}^2 & \Gamma_{22}\Lambda_{12} & \Gamma_{22}\Psi_{12} \\ \Gamma_{11}\Lambda_{12} & \Gamma_{22}\Lambda_{12} & \tfrac{1}{2}\{\Gamma_{11}\Gamma_{22} + \Lambda_{12}^2 - \Psi_{12}^2\} & \Lambda_{12}\Psi_{12} \\ \Gamma_{11}\Psi_{12} & \Gamma_{22}\Psi_{12} & \Lambda_{12}\Psi_{12} & \tfrac{1}{2}\{\Gamma_{11}\Gamma_{22} - \Lambda_{12}^2 + \Psi_{12}^2\} \end{pmatrix} \tag{9.1.22}$$

The matrix (9.1.22) was given in [1] and [2] and is applicable for $\|f_1| - |f_2\|$ very small. For frequency differences greater than $1/T$ the covariances are approximately zero. A more rigorous derivation of these results is given in Appendix A9.1. Note for future reference that for T very large, $W^2(-)$ tends to $(1/T)\,\delta(f_1 - f_2)$ in the continuous case, and to $(1/N\Delta)\,\delta(f_1 - f_2)$ in the discrete case.

9.2 PROPERTIES OF SMOOTHED CROSS SPECTRAL ESTIMATORS

9.2.1 Smoothed cross spectral estimators

It was shown in Section 9.1.3 that the estimators of the sample cross spectrum have the same undesirable properties as the sample spectrum, namely that their variances are dominated by a constant term which does not tend to zero as the record length increases. Hence cross spectral estimators must be smoothed using a spectral window just as it was necessary to smooth auto-spectral estimators.

The smoothed cross spectral estimator is defined by

$$\bar{C}_{12}(f) = \int_{-T}^{T} w(u)c_{12}(u)\, e^{-j2\pi fu}\, du, \tag{9.2.1}$$

where the lag window $w(u)$ has the usual properties (6.3.29). Writing $c_{12}(u)$ in terms of the even and odd parts (8.3.19, 20) gives

$$\begin{aligned}\bar{C}_{12}(f) &= \int_{-T}^{T} w(u)l_{12}(u)\cos 2\pi fu\, du - j\int_{-T}^{T} w(u)q_{12}(u)\sin 2\pi fu\, du \\ &= \bar{L}_{12}(f) - j\bar{Q}_{12}(f),\end{aligned} \tag{9.2.2}$$

where $\bar{L}_{12}(f)$ and $\bar{Q}_{12}(f)$ are the smoothed co- and quadrature spectral estimators.

Expected value of the smoothed cross spectral estimator. The sample cross spectral estimator $C_{12}(f)$ is defined by

$$C_{12}(f) = \int_{-T}^{T} c_{12}(u)\, e^{-j2\pi fu}\, du.$$

From (8.3.21) this has expected value

$$E[C_{12}(f)] = \int_{-T}^{T} \left(1 - \frac{|u|}{T}\right) \gamma_{12}(u)\, e^{-j2\pi fu}\, du,$$

which may be written

$$E[C_{12}(f)] = \int_{-\infty}^{\infty} T\left(\frac{\sin \pi Tg}{\pi Tg}\right)^2 \Gamma_{12}(f-g)\, dg \approx \Gamma_{12}(f). \tag{9.2.3}$$

Equation (9.2.3) follows from the fact that when T is sufficiently large, the spectral window is very narrow. Thus to a good approximation, $C_{12}(f)$ is approximately an unbiased estimator of $\Gamma_{12}(f)$. On taking expectations in

(9.2.1) and using (8.3.21) and (9.2.3), the mean of the smoothed cross spectral estimator is

$$\begin{aligned}E[\bar{C}_{12}(f)] &= \int_{-T}^{T} w(u)\left(1 - \frac{|u|}{T}\right)\gamma_{12}(u)\, e^{-j2\pi fu}\, du \\ &\approx \int_{-\infty}^{\infty} W(g)\Gamma_{12}(f-g)\, dg \\ &= \bar{\Gamma}_{12}(f).\end{aligned} \tag{9.2.4}$$

$\bar{\Gamma}_{12}(f)$ will be called the *mean smoothed cross spectrum*. Since $E[\bar{C}_{12}(f)] = E[\bar{L}_{12}(f)] - jE[\bar{Q}_{12}(f)]$, it follows that the mean smoothed co- and quadrature spectra are

$$\begin{aligned}E[\bar{L}_{12}(f)] &= \int_{-T}^{T} w(u)\left(1 - \frac{|u|}{T}\right)\lambda_{12}(u)\cos 2\pi fu\, du \\ &\approx \int_{-\infty}^{\infty} W(g)\Lambda_{12}(f-g)\, dg \\ &= \bar{\Lambda}_{12}(f)\end{aligned} \tag{9.2.5}$$

and

$$\begin{aligned}E[\bar{Q}_{12}(f)] &= \int_{-T}^{T} w(u)\left(1 - \frac{|u|}{T}\right)\psi_{12}(u)\sin 2\pi fu\, du \\ &\approx \int_{-\infty}^{\infty} W(g)\Psi_{12}(f-g)\, dg \\ &= \bar{\Psi}_{12}(f).\end{aligned} \tag{9.2.6}$$

In form, (9.2.4–6) are similar to (6.3.35) and (6.3.37) for the expected value of the auto-spectral estimator. However, there is an important difference in that the autocovariance function $\gamma_{11}(u)$ in (6.3.37) is symmetric about the origin. Hence, provided $|\gamma_{11}(u)|$ tends to zero fairly quickly, the bias $B_{11}(f) = \bar{\Gamma}_{11}(f) - \Gamma_{11}(f)$ can be expected to tend to zero quickly as the truncation point M of the lag window is increased. The situation is different for the cross spectral estimator since the cross correlation function is not symmetrical about zero. Thus in an extreme case where one process is an exact copy of the other process but delayed by an amount τ, the cross correlation function will be identical to the autocorrelation function but centered at lag τ instead of zero. The effect on (9.2.4) for the mean value of the cross spectral estimator is that there will be appreciable bias if the truncation point M is less than the delay τ. Furthermore if τ is large, the numerical value of the lag window $w(u)$ will be small when $u = \tau$, and hence very large truncation points may be necessary in order to reduce the bias to a reasonable size. This effect will be demonstrated in Section 9.3, where it will be shown that the trouble can be remedied by *aligning* the two series to ensure that the cross correlation function has a maximum in the neighborhood of zero lag.

Covariance matrix of smoothed cross spectral estimators. Using the convolution property (A2.1.8), the smoothed cross spectral estimator (9.2.1) may be written in the alternative form

$$\bar{C}_{12}(f) = \int_{-\infty}^{\infty} C_{12}(f-g)W(g)\,dg. \tag{9.2.7}$$

Hence

$$\begin{aligned}\text{Cov}\,[\bar{C}_{ij}(f_1), \bar{C}_{kl}(f_2)] &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \text{Cov}\,[C_{ij}(f_1-g), C_{kl}(f_2-h)]W(g)W(h)\,dg\,dh. \end{aligned}\tag{9.2.8}$$

Using the covariance matrix (9.1.22), the covariance matrix for the smoothed spectral estimators may be derived. For example, when T is large, (9.1.22) shows that

$$\text{Cov}\,[C_{11}(f_1), C_{22}(f_2)] \approx |\Gamma_{12}(f_1)|^2 \frac{\delta(f_1-f_2)}{T}.$$

Hence

$$\begin{aligned}\text{Cov}\,[\bar{C}_{11}(f_1), \bar{C}_{22}(f_2)] &\approx \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} |\Gamma_{12}(f_1-g)|^2 \frac{\delta(f_1-f_2-g+h)}{T} W(g)W(h)\,dg\,dh \\ &= \int_{-\infty}^{\infty} \frac{|\Gamma_{12}(f_1-g)|^2}{T} W(g)W(f_2+g-f_1)\,dg. \end{aligned}\tag{9.2.9}$$

Making the assumption that $\Gamma_{12}(f)$ is approximately constant over the bandwidth of the spectral window and substituting $h = f_1 - g$ in (9.2.9) gives

$$\text{Cov}\,[\bar{C}_{11}(f_1), \bar{C}_{22}(f_2)] \approx \frac{|\Gamma_{12}(f_1)|^2}{T}\int_{-\infty}^{\infty} W(f_1-h)W(f_2-h)\,dh. \tag{9.2.10}$$

For $f_1 = f_2 = f$, (9.2.10) reduces to

$$\begin{aligned}\text{Cov}\,[\bar{C}_{11}(f), \bar{C}_{22}(f)] &\approx \frac{|\Gamma_{12}(f)|^2}{T}\int_{-\infty}^{\infty} W^2(g)\,dg \\ &= |\Gamma_{12}(f)|^2 \frac{I}{T}, \end{aligned}\tag{9.2.11}$$

using (6.4.13). Similar results are obtained for the other spectral estimators.

Thus the effect of smoothing is to reduce the variances and the covariances of the unsmoothed estimators by the factor I/T. Hence the covariance matrix of the smoothed estimators is obtained by multiplying the covariance matrix (9.1.22) of the sample estimators by I/T instead of by $W^2(-)$. A more precise derivation of these results is given in Appendix A9.1.

The above covariance matrix is not of direct interest but is required as an intermediate stage in the computation of the covariance matrix of the smoothed cross amplitude, squared coherency and phase spectral estimators. This matrix is derived in the next section.

9.2.2 Smoothed cross amplitude, squared coherency and phase estimators

As shown in Section 8.4.4, the correlation between two stochastic processes can be described in the frequency domain by their cross amplitude and phase spectra, or more usefully by their squared coherency and phase spectra.

There are several ways in which smoothed estimators of these spectra may be defined. One simple way is to substitute the smoothed estimators of the co- and quadrature spectra into the expressions for the theoretical spectra. Thus, using (8.3.28), the smoothed cross amplitude estimator is defined by

$$\bar{A}_{12}(f) = \sqrt{\bar{L}_{12}^2(f) + \bar{Q}_{12}^2(f)}. \tag{9.2.12}$$

Similarly, using (8.3.29), the smoothed phase spectral estimator is defined by

$$\bar{F}_{12}(f) = \arctan\left(-\frac{\bar{Q}_{12}(f)}{\bar{L}_{12}(f)}\right). \tag{9.2.13}$$

Finally, using (8.4.18), the smoothed squared coherency estimator is defined by

$$\bar{K}_{12}^2(f) = \frac{\bar{L}_{12}^2(f) + \bar{Q}_{12}^2(f)}{\bar{C}_{11}(f)\bar{C}_{22}(f)}. \tag{9.2.14}$$

Note that even if $\bar{L}_{12}(f)$ and $\bar{Q}_{12}(f)$ were unbiased estimators, the estimators (9.2.12–14) would be biased. However, the bias produced in this way will be small compared with the bias produced by truncation of the cross correlation function and by the fact that it is not symmetrical about zero, and hence it is unlikely to inflate the mean square error. Since all the estimators are non-linear functions of the estimators $\bar{L}_{12}(f)$, $\bar{Q}_{12}(f)$, $\bar{C}_{11}(f)$, $\bar{C}_{22}(f)$, their moments can be derived by expanding in the form of a Taylor series as shown in Section 3.2.5 and in [2]. As an example, the mean and variance of the smoothed cross amplitude estimator (9.2.12) are derived.

For convenience, the dependence on frequency is dropped so that (9.2.12) may be written

$$\bar{A}_{12} = \sqrt{\bar{L}_{12}^2 + \bar{Q}_{12}^2}.$$

Now consider small perturbations $\delta\bar{L}_{12}$ and $\delta\bar{Q}_{12}$ about the expected values $E[\bar{L}_{12}] = \Lambda_{12}$, $E[\bar{Q}_{12}] = \psi_{12}$ so that

$$\begin{aligned}\bar{L}_{12} &= \Lambda_{12} + \delta\bar{L}_{12}, \\ \bar{Q}_{12} &= \Psi_{12} + \delta\bar{Q}_{12}, \\ E[\delta\bar{L}_{12}] &= 0 = E[\delta\bar{Q}_{12}].\end{aligned}$$

Similarly,

$$E[\delta\bar{L}_{12}^2] = \text{Var}\,[\bar{L}_{12}],$$
$$E[\delta\bar{Q}_{12}^2] = \text{Var}\,[\bar{Q}_{12}],$$
$$E[\overline{\delta L}_{12}\,\delta\bar{Q}_{12}] = \text{Cov}\,[\bar{L}_{12}, \bar{Q}_{12}].$$

On expanding (9.2.12) in the form of a Taylor series,

$$\bar{A}_{12} = \sqrt{(\Lambda_{12} + \delta\bar{L}_{12})^2 + (\Psi_{12} + \delta\bar{Q}_{12})^2}$$
$$\approx \alpha_{12}\left(1 + \frac{\Lambda_{12}\,\delta\bar{L}_{12} + \Psi_{12}\,\delta\bar{Q}_{12}}{\alpha_{12}^2}\right).$$

Hence

$$E[\bar{A}_{12}] \approx \alpha_{12}, \tag{9.2.15}$$

$$\text{Var}\,[\bar{A}_{12}] \approx \frac{\Lambda_{12}^2\,\text{Var}\,[\bar{L}_{12}] + \Psi_{12}^2\,\text{Var}\,[\bar{Q}_{12}] + 2\Lambda_{12}\Psi_{12}\,\text{Cov}\,[\bar{L}_{12}, \bar{Q}_{12}]}{\alpha_{12}^2}. \tag{9.2.16}$$

From the covariance matrix (9.1.22), with $W^2(-)$ replaced by I/T,

$$\text{Var}\,[\bar{L}_{12}] \approx \frac{I}{2T}\{\Gamma_{11}\Gamma_{22} + \Lambda_{12}^2 - \Psi_{12}^2\},$$

$$\text{Var}\,[\bar{Q}_{12}] \approx \frac{I}{2T}\{\Gamma_{11}\Gamma_{22} - \Lambda_{12}^2 + \Psi_{12}^2\},$$

$$\text{Cov}\,[\bar{L}_{12}, \bar{Q}_{12}] \approx \frac{I}{T}\Lambda_{12}\Psi_{12}.$$

Substituting these results in (9.2.16), the *variance for the smoothed cross amplitude estimator* is

$$\text{Var}\,[\bar{A}_{12}] \approx \frac{I}{2T}\alpha_{12}^2\left(1 + \frac{1}{\kappa_{12}^2}\right). \tag{9.2.17}$$

Note that when the X_1 and X_2 processes are identical, $\bar{A}_{12} = \bar{C}_{11}$, $\alpha_{12} = \Gamma_{11}$ and $\kappa_{12}^2 = 1$. Hence in this case (9.2.17) yields

$$\text{Var}\,[\bar{C}_{11}] \approx \frac{I}{T}\Gamma_{11}^2,$$

which is identical with the result (6.4.13) derived previously. Similar expressions may be derived for the covariances of the estimators $\bar{A}_{12}$, $\bar{F}_{12}$ and $\bar{K}_{12}^2$. These are:

Variance of smoothed coherency and squared coherency estimator

$$\text{Var}\,[|\bar{K}_{12}|] \approx \frac{I}{2T}(1 - \kappa_{12}^2)^2, \tag{9.2.18}$$

$$\text{Var}\,[\bar{K}_{12}^2] \approx \frac{I}{2T}4\kappa_{12}^2(1 - \kappa_{12}^2)^2. \tag{9.2.19}$$

Variance of smoothed phase estimator

$$\text{Var}\,[\bar{F}_{12}] \approx \frac{I}{2T}\left(\frac{1}{\kappa_{12}^2} - 1\right). \tag{9.2.20}$$

Note that this variance is independent of the theoretical phase function.

Covariance properties

$$\text{Cov}\,[\bar{F}_{12}, \bar{A}_{12}] \approx 0,$$
$$\text{Cov}\,[\bar{F}_{12}, \bar{K}_{12}^2] \approx 0. \tag{9.2.21}$$

The above results will be used in the next section to derive confidence intervals for the phase and coherency spectrum.

9.2.3 *Confidence intervals for squared coherency and phase spectra*

In this section some practical implications of the results of Section 9.2.2 are discussed and used to construct confidence intervals for coherency and phase spectra.

The results (9.2.17) to (9.2.20) show that the variances of these estimators depend on the *smoothing factor* I/T, which can be controlled by window closing, and the *coherency spectrum* $\kappa_{12}^2(f)$ of the two processes $X_1(t)$, $X_2(t)$.

They also show that in all cases, excluding (9.2.17), the variance of the estimator is zero when the coherency is unity and increases as the coherency tends to zero. In fact, the variances of the cross amplitude spectrum and phase spectrum estimators tend to infinity as the coherency tends to zero. This is to be expected, since a low coherency implies a large noise level and hence an inefficient estimate. Thus it is possible to make the important practical observation that the *sampling properties of the phase and cross amplitude estimators may be dominated by the uncontrollable influence of the coherency spectrum rather than by the controllable influence of the smoothing factor I/T.*

Confidence intervals for the coherency spectrum. The covariance property (9.2.21) shows that the phase and coherency estimators are uncorrelated, and hence it is permissible to derive confidence intervals for these spectra separately. Statisticians will recognize that, apart from the effect of smoothing, the variance (9.2.18) of $|\bar{K}_{12}(f)|$ is identical to the variance of an ordinary correlation coefficient. Hence R. A. Fisher's z-transformation [3] can be applied. Thus, using (3.2.28), the estimator

$$\bar{Y}_{12}(f) = \text{arctanh}\,[|\bar{K}_{12}|] = \frac{1}{2}\ln\frac{1 + |\bar{K}_{12}|}{1 - |\bar{K}_{12}|}$$

has a variance

$$\text{Var}\,[\bar{Y}_{12}(f)] \approx \frac{I}{2T}, \tag{9.2.22}$$

which is independent of frequency and which suggests that the estimate $\bar{Y}_{12}(f)$ should be plotted rather than the coherency itself, since the confidence interval can be represented by a constant interval on the Y scale.

To derive the confidence interval, it is reasonable to assume that the rv $\bar{Y}_{12}(f)$ is approximately Normal. Then an approximate $100(1-\alpha)\%$ confidence interval for $\bar{Y}_{12}(f) = \operatorname{arctanh} \kappa_{12}(f)$ is

$$\bar{Y}_{12}(f) \pm \eta\left(1 - \frac{\alpha}{2}\right)\sqrt{\frac{I}{2T}}. \tag{9.2.23}$$

For example, suppose that the observed coherency is $\bar{K}_{12}^2(f) = 0.64$ or $|\bar{K}_{12}(f)| = 0.8$ and that $I/2T = 0.09$. Then $\bar{Y}_{12}(f) = 1.099$ and the 95% confidence interval for $\operatorname{arctanh}[\bar{\kappa}_{12}(f)]$ is $1.099 \pm 1.96\sqrt{0.09} = 0.511, 1.687$.

On transforming back to the original scale, the 95% confidence limits for $\kappa_{12}^2(f)$ are (0.22, 0.87). In practice it is better to plot the coherency on the Y scale and then insert the constant confidence interval (9.2.23) on this diagram. Examples of transformed coherencies and their confidence intervals will be given in Section 9.3.

Confidence intervals for the phase spectrum. It is more difficult to obtain an approximation to the distribution of the phase estimator than that of the coherency estimator. In [1] an accurate approximation to the distribution of this estimator is given, but it is unwieldy. However, it is possible to use the result (9.2.20) to obtain crude confidence intervals for the phase spectrum [2]. More precise joint confidence intervals for the gain and phase will be given in Chapter 10.

It was stated in Section 9.1.1 that if the true coherency is zero, the sample phase estimator is uniformly distributed in the range $(-\pi/2, \pi/2)$. Furthermore, (9.2.20) shows that the effect of smoothing is to reduce the variance of the phase estimator. Hence it is to be expected that smoothing will constrain the estimator to be concentrated in a much narrower band than $(-\pi/2, \pi/2)$. For simplicity, it is desirable to find a transformation such that the distribution of the transformed variable is approximately Normal. The transformation $\tan \bar{F}_{12}$ is suggested because the range of the transformed variable extends from $-\infty$ to $+\infty$. Using (9.2.20) and (3.2.26),

$$\operatorname{Var}[\tan \bar{F}_{12}] \approx \sec^4 \phi \, \frac{I}{2T}\left(\frac{1}{\kappa_{12}^2} - 1\right). \tag{9.2.24}$$

Hence on approximating the distribution of $\tan \bar{F}_{12}$ by a Normal distribution, approximate confidence intervals for $\tan \Phi_{12}$, with confidence coefficient $100(1-\alpha)\%$, may be obtained from

$$\tan \bar{F}_{12}(f) \pm \eta\left(1 - \frac{\alpha}{2}\right)\sqrt{\sec^4 \phi \, \frac{I}{2T}\left(\frac{1}{\kappa_{12}^2} - 1\right)}. \tag{9.2.25}$$

Note that since the true coherency is unknown, it must be replaced by its estimate in (9.2.24). Since (9.2.20) shows that the variance of $\bar{F}_{12}$ is independent of ϕ_{12}, it would be expected that the interval (9.2.25), when converted back to a confidence interval for ϕ_{12}, will be approximately independent of ϕ_{12}.

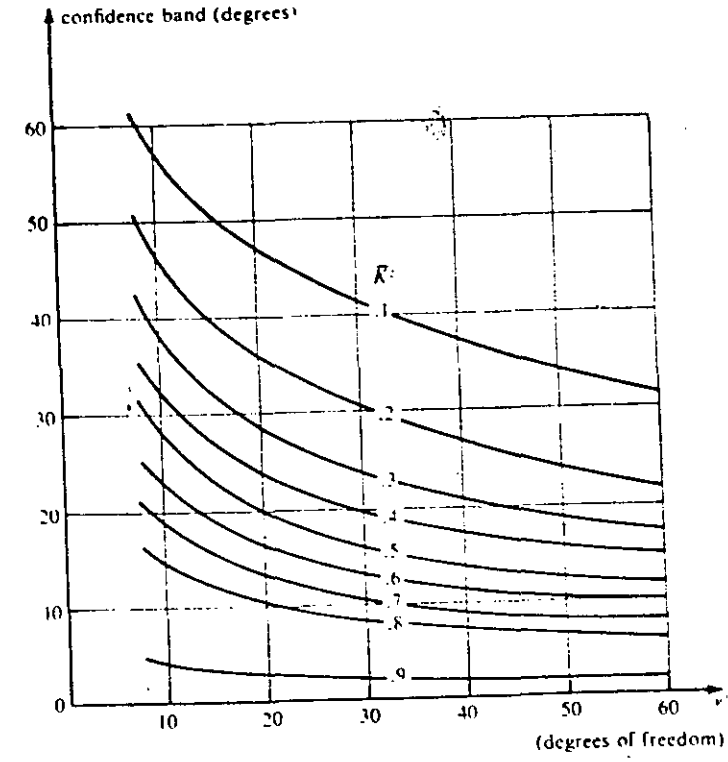


FIG. 9.3: 95% confidence limits for the phase spectrum

Figure 9.3 shows the 95% limits for various degrees of freedom of the spectral estimate. These values are adapted from [4]. For example, if $\nu = 27$ and $\bar{K}_{12}^2 = 0.5$, the confidence interval is $\bar{F} \pm 17°$.

9.3 CROSS SPECTRAL ANALYSIS USING SIMULATED SERIES

In this section the ideas introduced in the previous section are illustrated by comparing the smoothed phase and coherency spectral estimates of simulated series with their known theoretical phase and coherency spectra. The first section presents the formulae for digital computation of the smoothed estimates. Section 9.3.2 is intended to give the reader some experience in interpreting cross spectral estimates. It will then become clear that the procedures described up to the present require modification in order to produce satisfactory estimates. Section 9.3.3 shows that this can be achieved by *aligning* the two series, that is, moving one of the series relative to the other so that their cross correlation function has its maximum at zero lag.

9.3.1 Discrete estimation formulae

The formulae for the estimation of smoothed cross spectra using discrete data are similar to those for autospectra described in Section 7.1.1. As in Section 7.1.1, it is assumed that the data x_{1t}, x_{2t}, $t = 1, \ldots, N$, have been obtained by sampling at an interval of Δ seconds and that the spectral estimates are computed for *positive frequencies* only. For purposes of computation it is convenient to assume $\Delta = 1$ so that $0 \leqslant f \leqslant \frac{1}{2}$ cps. If $\Delta \neq 1$, the correct estimate can be recovered by multiplying the computed estimate by Δ and plotting the computed estimates in the range $0 \leqslant f \leqslant 1/2\Delta$ cps. As in Section 7.1.1, the number of covariance lags used in the spectral calculations is denoted by L. The smoothed spectral estimates are to be computed at the frequencies $0, 1/2F, \ldots, \frac{1}{2}$, where F is of the order of two or three times L. The lag window may be one of the three windows described in Section 7.1.1.

If trends are present, spurious coherencies may occur between the two series. In such cases, it is advisable to difference both series. As shown in Section 8.4.5, the differencing operation will not alter the theoretical coherency and phase spectra. In what follows, it is assumed that the covariances are either of the original or filtered data, as is appropriate in the given situation.

The formulae required and the order in which the computations are performed is as follows:

(1) For the x_{1t} data

(a) *The acvf estimate*

$$c_{11}(k) = \frac{1}{N}\sum_{t=1}^{N-k}(x_{1t} - \bar{x}_1)(x_{1t+k} - \bar{x}_1), \quad 0 \leqslant k \leqslant L-1, \tag{9.3.1}$$

where $\bar{x}_1 = (1/N)\sum_{t=1}^{N} x_{1t}$.

(b) *The smoothed spectral estimate*

$$\bar{C}_{11}(i) = 2\left\{c_{11}(0) + 2\sum_{k=1}^{L-1} c_{11}(k)w(k)\cos\frac{\pi ki}{F}\right\}, \quad 0 \leqslant i \leqslant F. \tag{9.3.2}$$

(2) For the x_{2t} data

(a) *The acvf estimate*

$$c_{22}(k) = \frac{1}{N}\sum_{t=1}^{N-k}(x_{2t} - \bar{x}_2)(x_{2t+k} - \bar{x}_2), \quad 0 \leqslant k \leqslant L-1, \tag{9.3.3}$$

where $\bar{x}_2 = (1/N)\sum_{t=1}^{N} x_{2t}$.

(b) *The smoothed spectral estimate*

$$\bar{C}_{22}(i) = 2\left\{c_{22}(0) + 2\sum_{k=1}^{L-1} c_{22}(k)w(k)\cos\frac{\pi ki}{F}\right\}, \quad 0 \leqslant i \leqslant F. \tag{9.3.4}$$

(3) For the x_{1t} and x_{2t} data

(a) *The ccvf estimate*

$$c_{12}(k) = \frac{1}{N}\sum_{t=1}^{N-k}(x_{1t} - \bar{x}_1)(x_{2t+k} - \bar{x}_2), \quad 0 \leqslant k \leqslant L-1,$$

$$c_{12}(-k) = \frac{1}{N}\sum_{t=1}^{N-k}(x_{1t+k} - \bar{x}_1)(x_{2t} - \bar{x}_2), \quad 0 \leqslant k \leqslant L-1. \tag{9.3.5}$$

(b) *The even and odd ccvf estimates*

$$l_{12}(k) = \tfrac{1}{2}\{c_{12}(k) + c_{12}(-k)\}, \quad 0 \leqslant k \leqslant L-1, \tag{9.3.6}$$

$$q_{12}(k) = \tfrac{1}{2}\{c_{12}(k) - c_{12}(-k)\}, \quad 0 \leqslant k \leqslant L-1. \tag{9.3.7}$$

Note that $q_{12}(0) = 0$.

(c) *The smoothed co- and quadrature spectral estimates*

$$\bar{L}_{12}(i) = 2\left\{l_{12}(0) + 2\sum_{k=1}^{L-1} l_{12}(k)w(k)\cos\frac{\pi ik}{F}\right\}, \quad 0 \leqslant i \leqslant F, \tag{9.3.8}$$

$$\bar{Q}_{12}(i) = 4\sum_{k=1}^{L-1} q_{12}(k)w(k)\sin\frac{\pi ik}{F}, \quad 1 \leqslant i \leqslant F-1, \tag{9.3.9}$$

$$\bar{Q}_{12}(0) = \bar{Q}_{12}(F) = 0.$$

(d) *The smoothed cross amplitude spectral estimate*

$$\bar{A}_{12}(i) = \sqrt{\bar{L}_{12}^2(i) + \bar{Q}_{12}^2(i)}, \quad 0 \leqslant i \leqslant F. \tag{9.3.10}$$

(e) *The smoothed phase spectral estimate*

$$\bar{F}_{12}(i) = \arctan -\frac{\bar{Q}_{12}(i)}{\bar{L}_{12}(i)}, \quad 0 \leqslant i \leqslant F. \tag{9.3.11}$$

(f) *The smoothed squared coherency spectral estimate*

$$\bar{K}_{12}^2(i) = \frac{\bar{A}_{12}^2(i)}{\bar{C}_{11}(i)\bar{C}_{22}(i)}, \quad 0 \leqslant i \leqslant F. \tag{9.3.12}$$

The factor 2 in equations (9.3.2, 4, 8, 9) is to preserve the Fourier transform relationship between the sample spectra and the sample covariances, as in Section 7.1.1. Appendix A9.2 gives a flow chart for bivariate spectral calculations.

Figure 9.4 demonstrates very clearly the effect of window closing on the smoothed coherency estimate, the theoretical value being zero in this example. For $L = 4$ and 8 the coherencies are reasonably smooth and close to zero, but as L is increased, and hence the bandwidth is decreased, very large coherencies begin to appear. It was shown in Section 9.1.3 that this is partly due to the large variance of the estimator when the bandwidth is small. In addition, as shown in Section 9.1.2, as the bandwidth is made smaller the smoothed coherency tends to unity at all frequencies since the sample coherency estimate for unsmoothed data is identically equal to one at all frequencies.

Figure 9.5 shows the smoothed coherency estimate $\bar{K}_{12}^2$ when $L = 16$ for the original series and the series after the filtering operation described in Section 8.2.2. It is seen that filtering of the two series only marginally improves the coherency estimate. This is to be compared with the conclusion drawn in Section 8.2.2 that filtering can lead to greatly improved estimates of the cross correlation function. This behavior of the coherency estimate is explained in Section 9.3.3.

A bivariate autoregressive process. The second process investigated is the bivariate ar process (8.1.20),

$$X_{1t} - 0.6X_{1t-1} + 0.5X_{2t-1} = Z_{1t},$$
$$X_{2t} - 0.4X_{1t-1} - 0.5X_{2t-1} = Z_{2t},$$

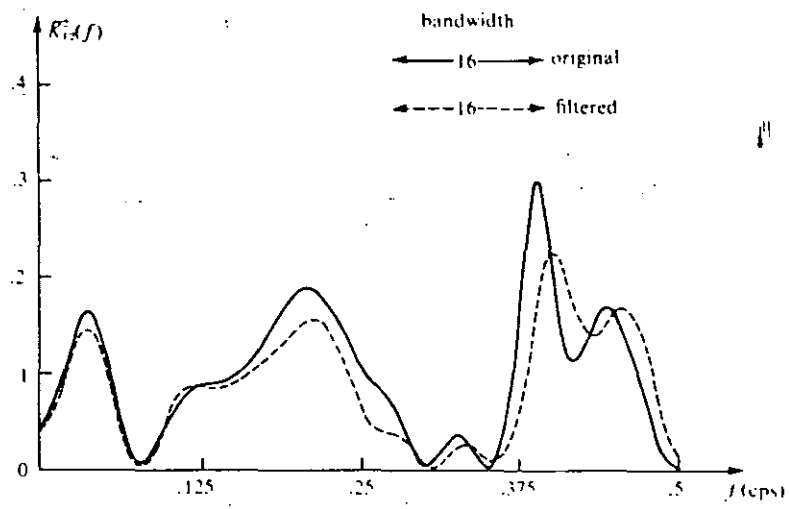


FIG. 9.5: Smoothed coherency estimates for two uncorrelated series: filtered and original series

where Z_{1t}, Z_{2t} are two mutually independent white noise processes.

The theoretical coherency and phase spectra are given by (8.4.19) and (8.4.20) respectively. The theoretical coherency κ_{12}^2 is plotted in Figure 9.6, together with the mean smoothed coherencies $\bar{\kappa}_{12}^2$ for $L = 4$, 8 and 16. It is seen that considerable bias exists for $L = 4$ and 8, the peak of the coherency being displaced by about 0.1 cps for $L = 4$ and 0.05 for $L = 8$. For $L = 16$, good agreement between $\bar{\kappa}_{12}^2$ and κ_{12}^2 is obtained, and for $L = 32$ the theoretical and smoothed spectra are indistinguishable. Hence an estimate of the

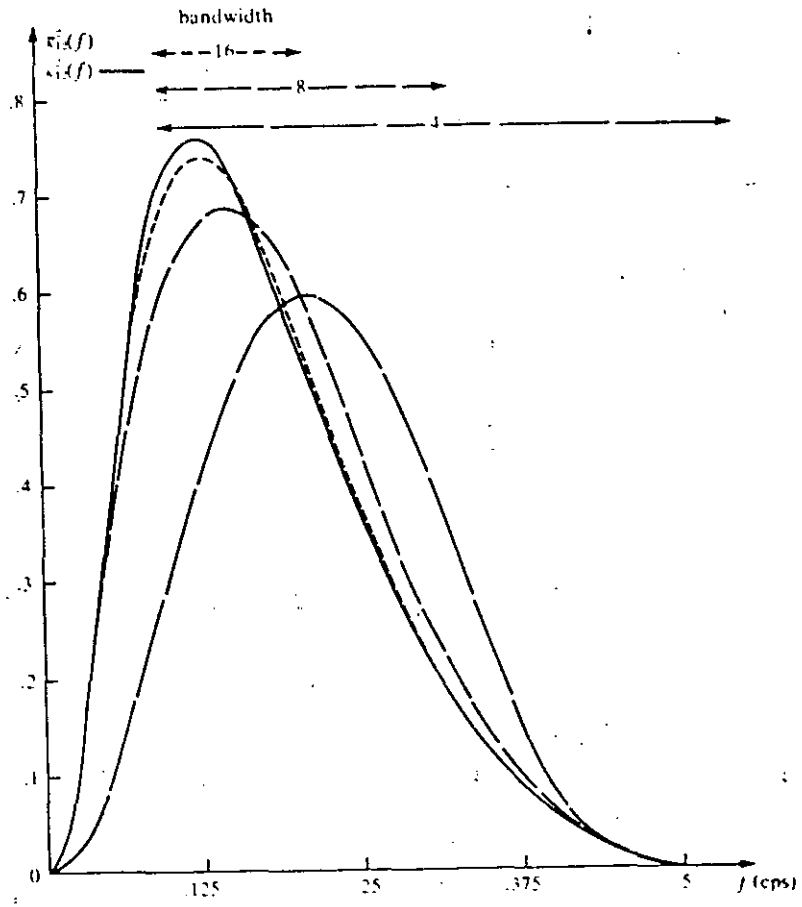


FIG. 9.6: Mean smoothed coherency spectra for the bivariate autoregressive process (8.1.20)

coherency for this process with acceptably small bias should be obtained with $L = 16$.

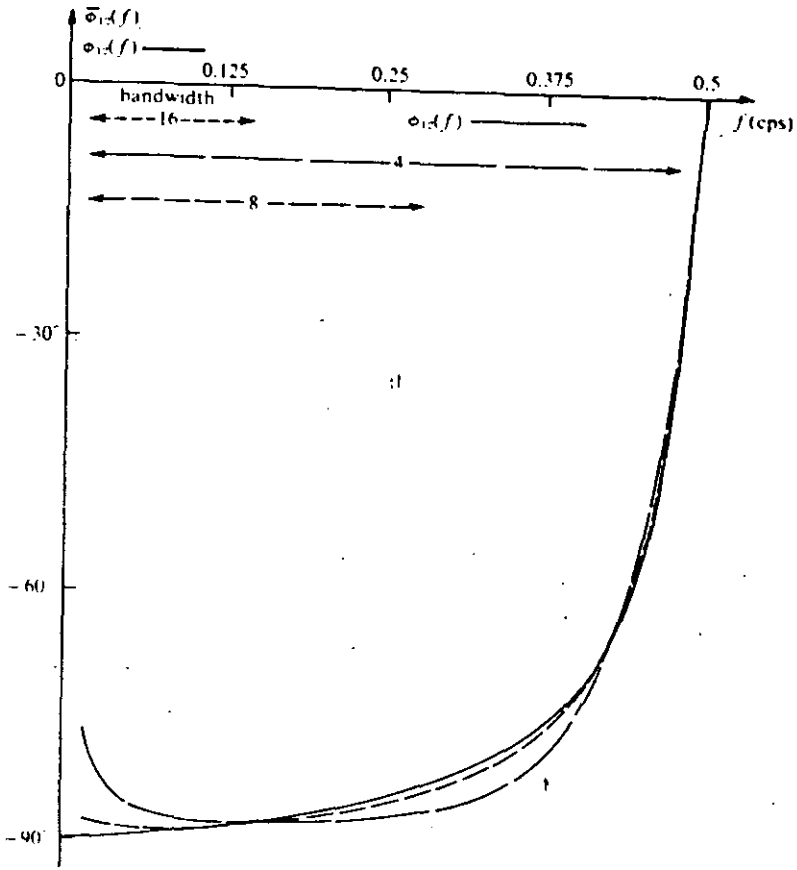


FIG. 9.7: Mean smoothed phase spectra for the bivariate autoregressive process (8.1.20)

Figure 9.7 shows the theoretical and mean smoothed phase spectra for the process (8.1.20) when $L = 4$, 8 and 16. Excellent agreement between ϕ_{12} and ϕ_{12} is obtained for $L = 8$, and for $L = 16$ the mean smoothed phase spectrum is indistinguishable from the theoretical spectrum. Hence the phase spectrum could be estimated with fewer lags than are needed for the coherency spectrum.

Table A9.1 and Figure 9.8 show the sample auto- and cross correlation functions based on a realization of $N = 100$ terms of the process (8.1.20). The original data for these series are given in Table A8.1. Figure 9.9 shows the theoretical coherency and the coherency estimates computed from these correlation functions. The estimate $\bar{K}_{12}^2$ using $L = 4$ is considerably displaced from the theoretical coherency, as is the mean smoothed coherency $\bar{\kappa}_{12}^2$.

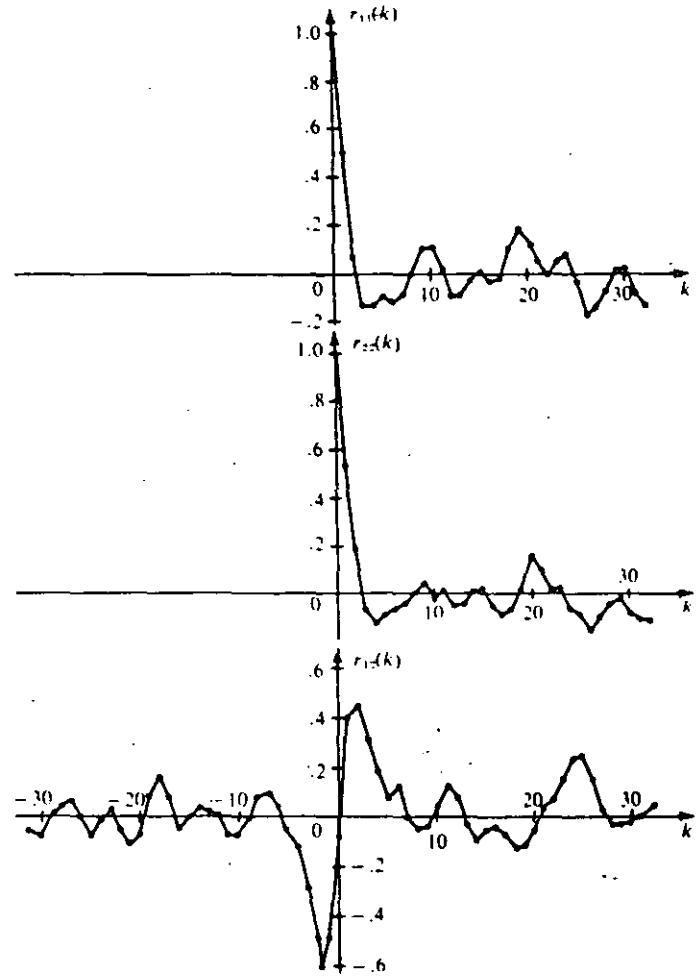


FIG. 9.8: Sample auto- and cross correlations for the bivariate autoregressive process (8.1.20) ($N = 100$)

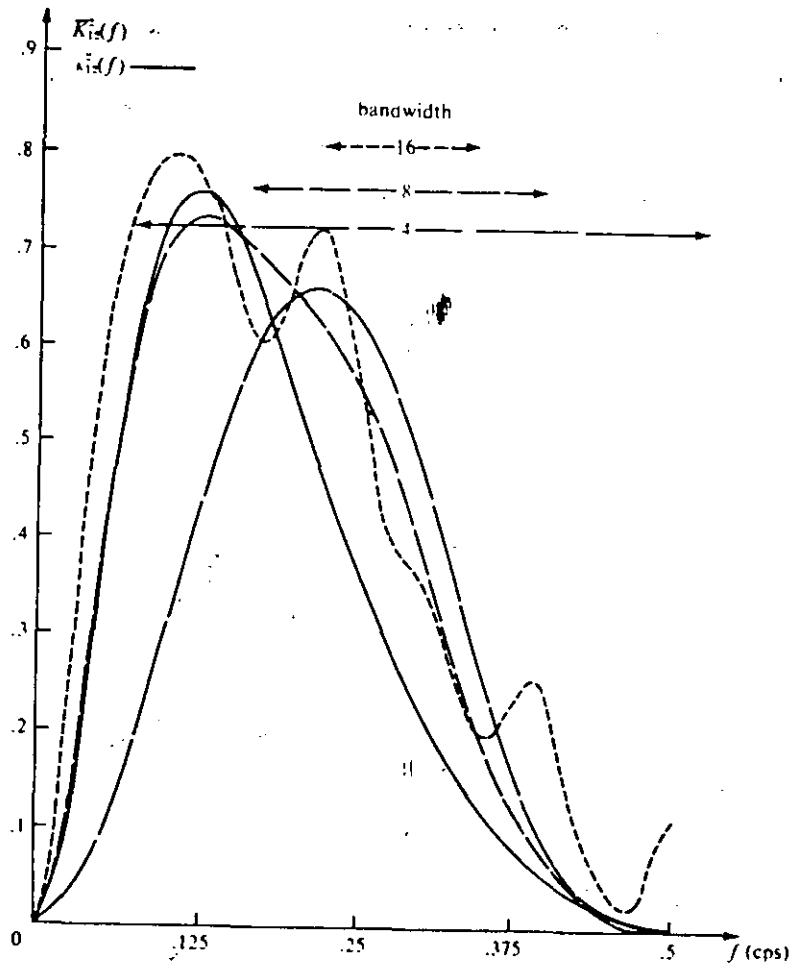


FIG. 9.9: Smoothed coherency estimates for the bivariate autoregressive process (8.1.20) ($N = 100$)

Doubling L to 8 produces a marked change in $\hat{K}_{12}^2$, but increasing L to 16 produces little change. Hence in this case the estimate based on 16 or possibly 12 would be accepted using the window-closing technique of Section 7.2.4. Note that for $L = 16$ there is quite good agreement between $\hat{K}_{12}^2$ and κ_{12}^2. However, increasing L to 32 produces wild oscillations in $\hat{K}_{12}^2$.

The theoretical phase spectrum and the smoothed phase estimates are shown in Figure 9.10 for $L = 4$, 8 and 16. The window-closing procedure suggests that very little change in phase occurs above $L = 8$ and that with $L = 16$, spurious peaks appear in the estimate. Hence the estimate based on

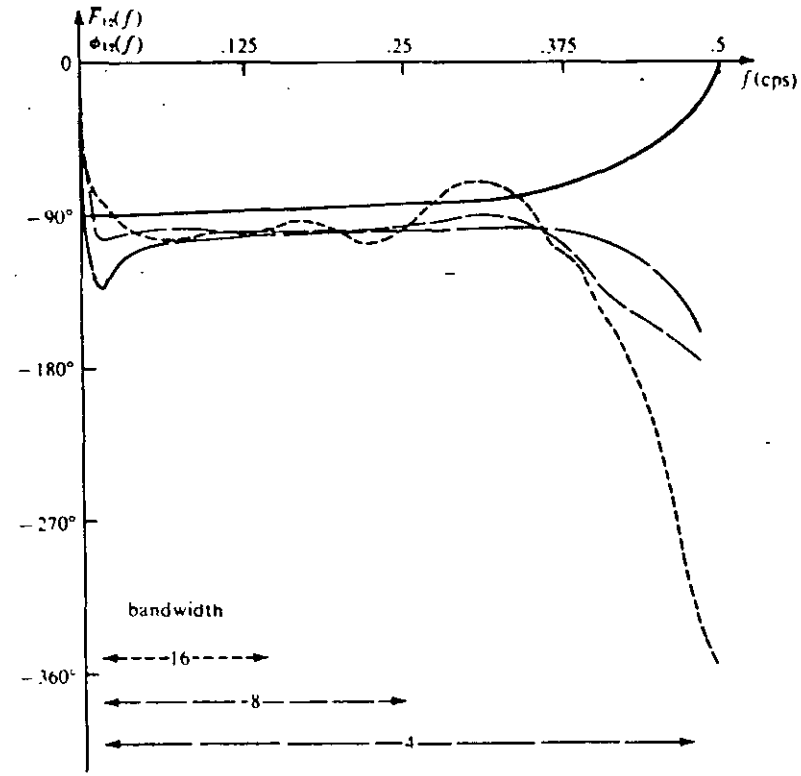


FIG. 9.10: Smoothed phase estimates for the bivariate autoregressive process (8.1.20) ($N = 100$)

$L = 8$ would probably be accepted, and comparison with Figure 9.7 shows that there is good agreement with the theoretical phase spectrum in the range 0 to 0.4 cps. For frequencies above 0.4 cps, the estimates tend to move downward, whereas the theoretical phase shown in Figure 9.7 moves upward. Figure 9.9 shows that $f = 0.4$ cps corresponds to the point where the coherency falls to a small value. Using an average coherency of 0.1 in this region, Figure 9.3 shows that the 95% confidence interval for the phase for $L = 4$ is approximately $\pm 30°$.

Noise through a linear system with delay. The third process investigated is (8.1.22), where X_{2t} is the output from a first-order linear system with a delay of 10 time units

$$X_{2t} = 0.5X_{2t-1} + 2X_{1t-10} + Y_t.$$

The input X_{1t} to the system is a first-order ar process

$$X_{1t} = 0.6X_{1t-1} + Z_{1t}.$$

The noise Y_t is a first-order ar process

$$Y_t = 0.5\,Y_{t-1} + Z_{2t},$$

and Z_{1t}, Z_{2t} are two mutually independent white noise processes. The theoretical correlation functions of this process are given in Section 8.1.4.

The theoretical coherency and phase spectra, derived using the methods of Section 8.4.3, are

$$\kappa_{12}^2(f) = \frac{5 - 4\cos 2\pi f}{6.36 - 5.2\cos 2\pi f},$$

$$\phi_{12}(f) = \arctan\left\{\frac{-\sin 2\pi f}{2 - \cos 2\pi f}\right\} - 20\pi f.$$

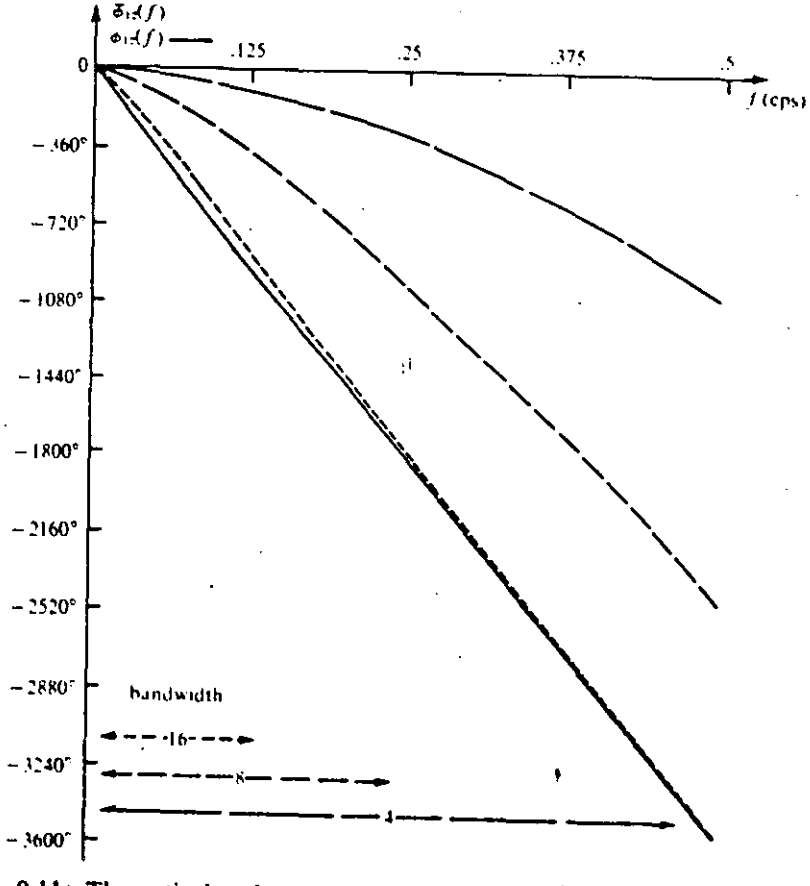


FIG. 9.11: Theoretical and mean smoothed phase spectra for the linear process (8.1.22) (before alignment)

The theoretical and mean smoothed phase spectra for this linear process are shown in Figure 9.11, from which it is seen that good estimates of the phase can be obtained with as few as 12 or 16 lags.

The theoretical coherency spectrum κ_{12}^2 is plotted in Figure 9.12, together with the mean smoothed coherencies $\bar{\kappa}_{12}^2$ for $L = 16$, 24 and 32. It is seen that the mean smoothed coherencies are markedly different from the theoretical coherency even for 32 lags, and that the difference is not attributable to lack of smoothness of the theoretical coherency. The reason is that bias is being introduced because of the large delay between the input and output, as predicted in Section 9.2.1.

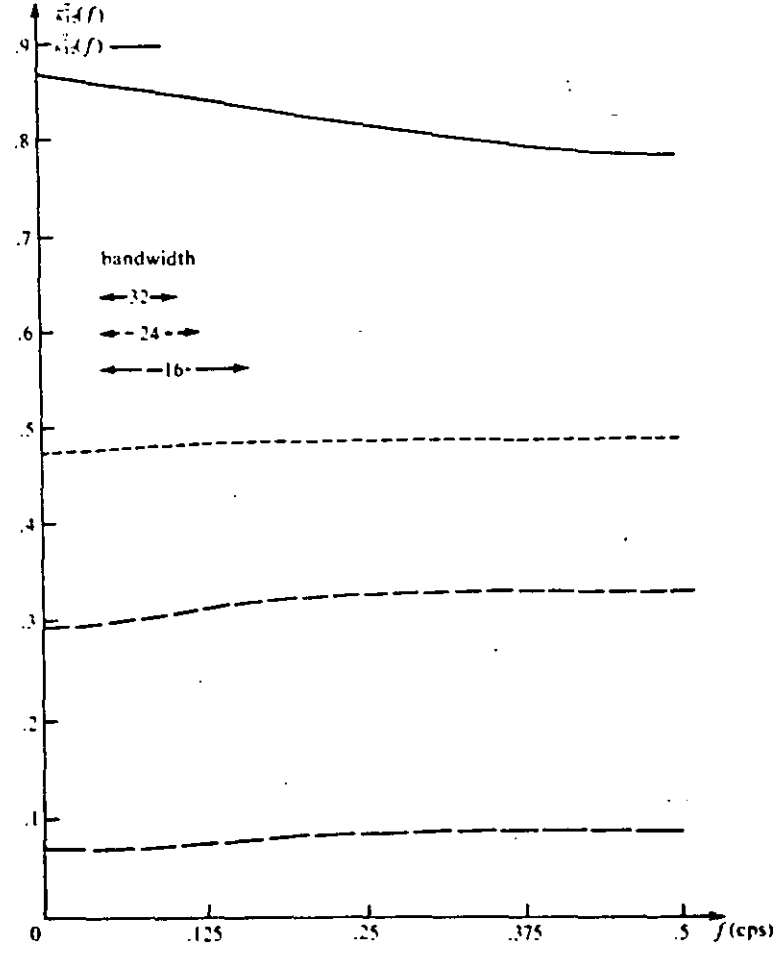


FIG. 9.12: Theoretical and mean smoothed coherency spectra for the linear process (8.1.22) (before alignment)

When $L = 8$, the mean smoothed coherency is approximately zero, as would be expected since the delay between the two series exceeds the number of lags used. Hence in this instance the bias equals the function itself.

It is clear from this theoretical investigation that little would be learned from a spectral analysis of a realization of 100 or so terms of this process unless care is exercised. To demonstrate this conclusion, the auto- and cross correlations of a realization of $N = 100$ terms were computed and are shown in Table A9.2 and Figure 9.13. The original data for this example are given

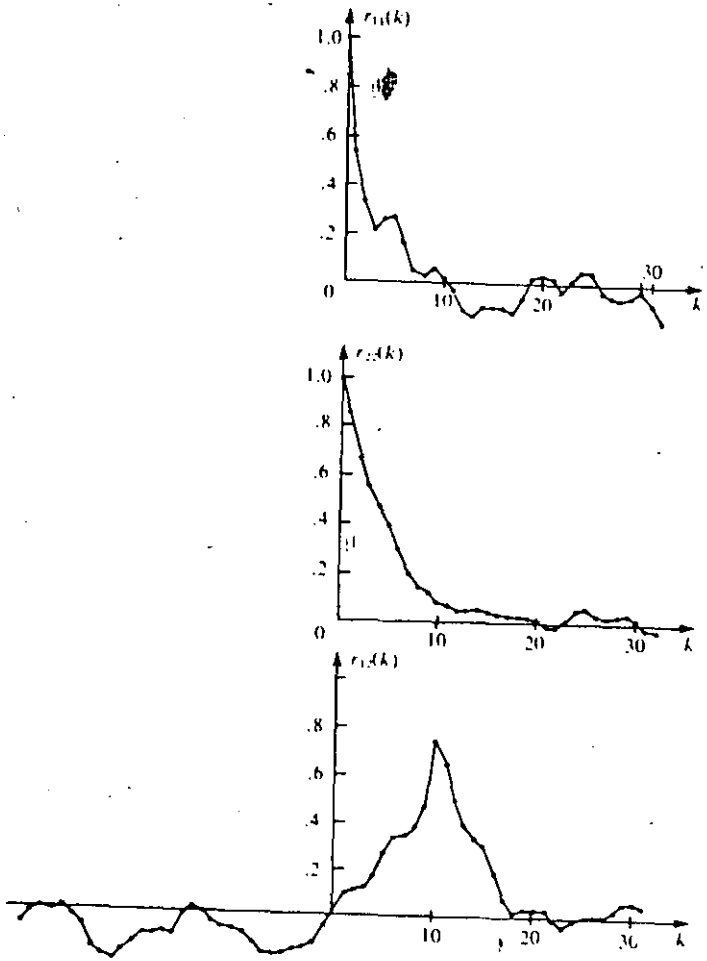


FIG. 9.13: Sample auto- and cross correlations for the linear process (8.1.22) ($N = 100$)

in Table A8.2. Figure 9.14 shows the coherency spectral estimate based on a realization of $N = 100$ terms when $L = 8$, 16 and 32. It is clear that the estimate is not converging to a stable function, and that for $L = 32$ it tends to "blow up." Hence no satisfactory conclusions can be drawn concerning the coherency estimate. The phase estimates for $L = 4, 8, 16$ and 32 are shown in Figure 9.15. It is seen that the estimates are very poor when the number of lags is less than or comparable to the delay period. Once the number of lags is greater than the delay of 10, the estimate improves rapidly and an excellent estimate is obtained with $L = 32$.

The examples of this section illustrate the general point that good estimates of phase may be obtained in situations where the coherency is badly estimated. In the next section it is shown that it is usually possible to obtain considerable

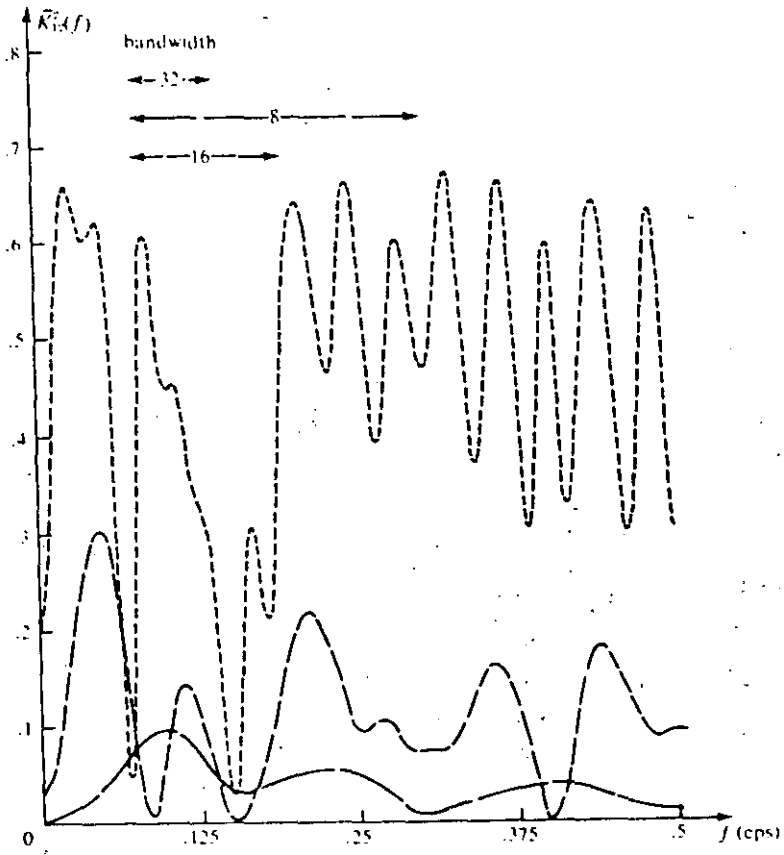


FIG. 9.14: Smoothed coherency estimates for the linear process (8.1.22) ($N = 10$ before alignment)

improvement in the estimates of coherency and phase by aligning the two processes.

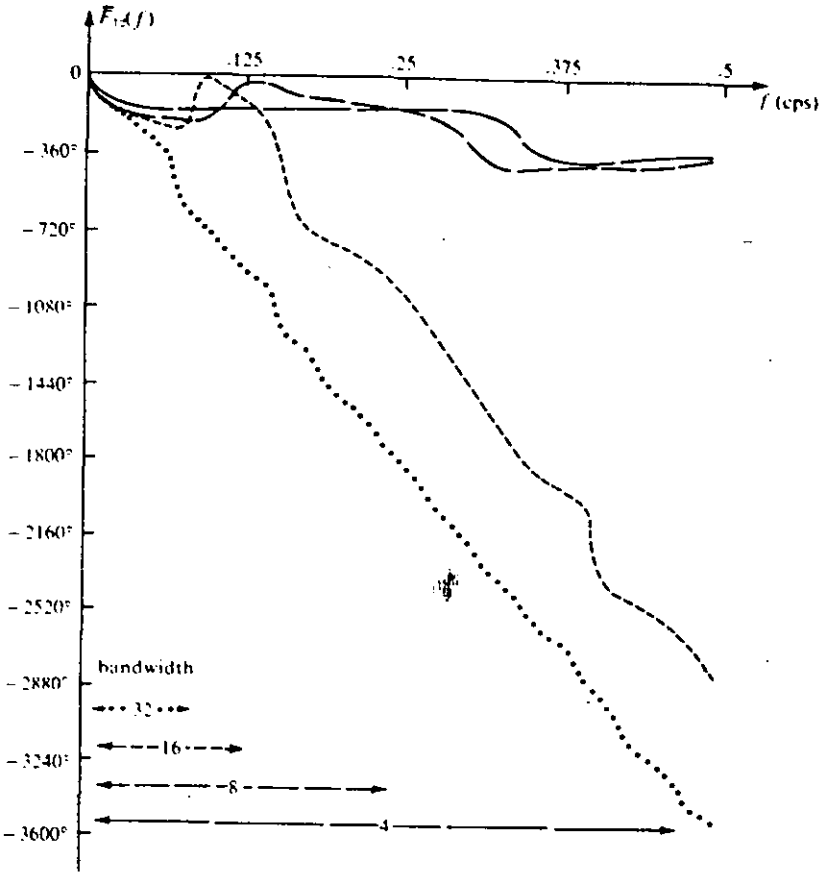


FIG. 9.15: Smoothed phase estimates for the linear process (8.1.22) ($N = 100$, before alignment)

9.3.3 *Improvement of cross spectral estimates*

It was shown in the last section that large bias can occur in the estimation of coherency spectra, especially when there is large delay. In this section the bias in the phase and coherency spectral estimators are evaluated and it is shown that this bias can be considerably reduced by the process of *alignment*. Alignment consists of centering the ccf so that the largest absolute value of the ccf is at zero lag.

Bias in coherency estimators. Approximate expressions for the bias of the smoothed coherency estimators may be obtained by following a procedure

similar to that in Section 6.3.5. For example, the bias in the smoothed coherency estimator is

$$B(f) = E[\bar{K}_{12}^2(f) - \kappa_{12}^2(f)] = E\left[\frac{|\bar{C}_{12}(f)|^2}{\bar{C}_{11}(f)\bar{C}_{22}(f)} - \frac{|\Gamma_{12}(f)|^2}{\Gamma_{11}(f)\Gamma_{22}(f)}\right]. \quad (9.3.17)$$

Using (3.2.23) the bias in the coherency estimator may be approximated by

$$B(f) = \kappa_{12}^2(f)\left[\frac{B_{12}(f)}{\alpha_{12}^2(f)} - \frac{B_{11}(f)}{\Gamma_{11}(f)} - \frac{B_{22}(f)}{\Gamma_{22}(f)}\right], \quad (9.3.18)$$

where $B_{12}(f)$, $B_{11}(f)$, $B_{22}(f)$ are the biases in $|\bar{C}_{12}(f)|^2$, $\bar{C}_{11}(f)$ and $\bar{C}_{22}(f)$ respectively. For the present it is assumed that the autospectra are sufficiently smooth compared with the spectral windows so that the biases in the autospectral estimators may be ignored. First, the bias $B_{12}(f)$ in the squared cross amplitude spectrum is calculated. That is,

$$B_{12}(f) = E[\bar{A}_{12}^2(f) - \alpha_{12}^2(f)].$$

Now,

$$E[|\bar{C}_{12}(f)|^2] = E\left[\int_{-\infty}^{\infty} C_{12}(g)W(f-g)\,dg \int_{-\infty}^{\infty} C_{12}^*(h)W(f-h)\,dh\right] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} W(f-g)W(f-h)E[C_{12}(g)C_{12}^*(h)]\,dg\,dh. \quad (9.3.19)$$

But,

$$E[C_{12}(g)C_{12}^*(h)] = \text{Cov}\,[C_{12}(g), C_{12}^*(h)] + E[C_{12}(g)]E[C_{12}^*(h)],$$

and from (9.1.22) and (9.2.3) this is approximately

$$E[C_{12}(g), C_{12}^*(h)] \approx \frac{1}{T}\Gamma_{11}(g)\Gamma_{22}(g)\,\delta(g-h) + \Gamma_{12}(g)\Gamma_{12}^*(h).$$

Hence

$$E[|\bar{C}_{12}(f)|^2] \approx \int_{-\infty}^{\infty} \frac{W^2(f-g)}{T}\Gamma_{11}(g)\Gamma_{22}(g)\,dg + |\bar{\Gamma}_{12}(f)|^2. \quad (9.3.20)$$

If $\Gamma_{11}(f)$ and $\Gamma_{22}(f)$ are smooth relative to the width of the spectral window, then (9.3.20) reduces to

$$E[|\bar{C}_{12}(f)|^2] \approx \Gamma_{11}(f)\Gamma_{22}(f)\frac{I}{T} + |\bar{\Gamma}_{12}(f)|^2 \quad (9.3.21)$$

and

$$B(f) \approx \frac{I}{T} + \frac{|\bar{\Gamma}_{12}(f)|^2 - |\Gamma_{12}(f)|^2}{\Gamma_{11}(f)\Gamma_{22}(f)}. \quad (9.3.22)$$

using (9.3.18).

Equation (9.3.22) shows that even if the theoretical cross spectrum is zero, the mean smoothed coherency can be large. This explains the large coherency

estimates, obtained in Figure 9.5, for the two independent first-order ar processes discussed in Section 8.2. For example, when $L = 40$,

$$\frac{I}{T} = \frac{(0.75)40}{100} = 0.3,$$

which agrees quite well with the average value estimated from the values of Figure 9.4. Note that as L increases (and I/T increases) the average coherency also increases. As indicated in Section 9.1.2, when $M = L\Delta$ tends to T, the coherency tends to one for all values of f.

Also shown by (9.3.22) is that filtering independent series will not produce improved coherency estimates. This is demonstrated in Figure 9.4 for the original and filtered independent processes, where it is seen that the average coherency for both the original and filtered data is about 0.1. This agrees very well with the value $I/T \doteq (0.75)\ 16/100 = 0.12$ predicted by (9.3.22).

To obtain an explicit expression for the bias (9.3.22), it is necessary to evaluate $|\bar{\Gamma}_{12}(f)|^2 - |\Gamma_{12}(f)|^2$. From (9.2.4),

$$\bar{\Gamma}_{12}(f) = \int_{-\infty}^{\infty} w(u)\gamma_{12}(u)\, e^{-i2\pi fu}\, du,$$

so that

$$|\bar{\Gamma}_{12}(f)|^2 = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} w(u)w(v)\gamma_{12}(u)\gamma_{12}(v)\, e^{-i2\pi f(u-v)}\, du\, dv.$$

Hence

$$|\bar{\Gamma}_{12}(f)|^2 - |\Gamma_{12}(f)|^2 = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \gamma_{12}(u)\gamma_{12}(v)\, e^{-i2\pi f(u-v)}\,(w(u)w(v)-1)\,du\,dv. \tag{9.3.23}$$

Writing

$$w(u)w(v) - 1 = (w(u) - 1)(w(v) - 1) + (w(u) - 1) + (w(v) - 1)$$

gives

$$\begin{aligned}|\bar{\Gamma}_{12}(f)|^2 - |\Gamma_{12}(f)|^2 &= \int_{-\infty}^{\infty} (w(u)-1)\gamma_{12}(u)\, e^{-i2\pi fu}\, du \int_{-\infty}^{\infty} (w(v)-1)\gamma_{12}(v)\, e^{i2\pi fv}\, dv \\ &\quad + \int_{-\infty}^{\infty} (w(u) - 1)\gamma_{12}(u)\, e^{-i2\pi fu}\, du \int_{-\infty}^{\infty} \gamma_{12}(v)\, e^{i2\pi fv}\, dv \\ &\quad + \int_{-\infty}^{\infty} \gamma_{12}(u)\, e^{-i2\pi fu}\, du \int_{-\infty}^{\infty} (w(v) - 1)\gamma_{12}(v)\, e^{i2\pi fv}\, dv.\end{aligned}$$

Using the approximations (6.3.37) for bias gives, for the Tukey window,

$$|\bar{\Gamma}_{12}(f)|^2 - |\Gamma_{12}(f)|^2 \approx \left(\frac{0.063}{M^2}\right)\Gamma_{12}^{(2)}\left(\frac{0.063}{M^2}\right)\Gamma_{12}^{(2)*} + \left(\frac{0.063}{M^2}\right)\{\Gamma_{12}^{(2)}\Gamma_{12}^{*} + \Gamma_{12}\Gamma_{12}^{(2)*}\}, \tag{9.3.24}$$

where $\Gamma_{12}^{(2)}$ is the second derivative of the cross amplitude spectrum at frequency f. Writing $\Gamma_{12}(f) = \alpha_{12}(f)\, e^{i\phi_{12}(f)}$ and taking derivatives with respect to f gives

$$|\bar{\Gamma}_{12}(f)|^2 - |\Gamma_{12}(f)|^2 \approx \frac{0.126}{M^2}\{\alpha_{12}\alpha_{12}^{(2)} - \alpha_{12}^2(\phi_{12}^{(1)})^2\},$$

on neglecting terms of order $1/M^4$. Hence if the biases $B_{11}(f)$, $B_{22}(f)$ in (9.3.18) may be neglected, the bias in the coherency estimator is approximately

$$B(f) \approx \frac{0.75M}{T} + \frac{0.126}{M^2}\left\{\frac{\alpha_{12}\alpha_{12}^{(2)} - \alpha_{12}^2(\phi_{12}^{(1)})^2}{\Gamma_{11}\Gamma_{22}}\right\} \tag{9.3.25}$$

for the Tukey window. For the Parzen window, 0.75 is replaced by 0.54 and 0.126 by 0.304.

The most important feature of (9.3.25) is that the bias is proportional to the square of the first derivative of the phase spectrum. If the constant term and the term in $\alpha_{12}^{(2)}$ are neglected, (9.3.25) reduces to

$$B(f) \approx \frac{0.126}{M^2}(\kappa_{12}^2)(\phi_{12}^{(1)})^2, \tag{9.3.26}$$

so that the bias in coherency is directly proportional to the coherency and the rate of change $\phi_{12}^{(1)}$ of the phase spectrum. Hence if there are large delays between two processes, the coherency estimate can be very poor since $\phi_{12}^{(1)}$ will be large. The bias in cross spectral estimators was first pointed out by Akaike [5].

Alignment. The bias in coherency due to phase shift may be reduced appreciably by *alignment* of the two processes. Thus, suppose the cross correlation function has its largest absolute value, or peak, at lag S. Aligning the processes so that the peak occurs at zero, changes $\Gamma_{12}(f)$ from

$$\alpha_{12}(f)\, e^{i\phi_{12}(f)}$$

to

$$\Gamma'_{12}(f) = \alpha'_{12}(f)\, e^{i\phi'_{12}(f)} = \alpha_{12}(f)\, e^{i(\phi_{12}(f) - 2\pi fS)}.$$

Hence

$$\phi_{12}^{\prime(1)} = \frac{d}{df}(\phi'_{12}(f)) = (\phi_{12}^{(1)} - 2\pi S),$$

and so the bias (9.3.26) may be considerably reduced, as will be demonstrated in the next section.

Use of the phase spectrum to determine the alignment parameter. The choice of the alignment parameter S by locating the peak in the ccf does not always lead to a satisfactory spectral analysis. It may happen that the phase spectrum of the aligned series still has a linear phase component $\phi(f) = 2\pi fd$, showing that a further alignment d is necessary. As a practical procedure, it is

recommended that a first guess of the alignment parameters should be made on the basis of the peak in the ccf. In some cases, the resulting phase spectrum will not contain a linear component and the analysis can be terminated. In those cases where a linear component $\phi(f) = 2\pi fd$ still remains, a modified alignment $S + d$ can be chosen and the analysis repeated.

Bias in phase estimators. Approximate expressions for the bias in phase estimators may be derived using an approach similar to that used above. The final result for the Tukey window is

$$B(f) \approx \frac{0.063}{M^2}\left[\frac{1}{\alpha_{12}^2}\frac{d}{df}(\alpha_{12}^2\phi_{12}^{(1)})\right] = \frac{0.063}{M^2}\left[\phi_{12}^{(2)} + \phi_{12}^{(1)}\frac{d}{df}\ln\alpha_{12}^2\right]. \quad (9.3.27)$$

For the Parzen window, (9.3.27) holds with 0.063 replaced by 0.152.

From (9.3.27) it is seen that the bias is proportional to the second derivative of the phase spectrum and is also proportional to the product of the first derivative of the phase spectrum times the first derivative of the logarithm of the cross amplitude spectrum. For bivariate processes involving large delays the quantity $\phi_{12}^{(1)}$ will probably dominate $\phi_{12}^{(2)}$. Since $\phi_{12}^{(1)}$ is multiplied by $d(\ln\alpha^2)/df$, however, the net bias may be quite small. Such a situation obtains in the linear filter example, shown in Figure 9.11, which gives the theoretical and mean smoothed phase spectra. It is seen that the smoothed and theoretical spectra agree very well, even for 16 lags.

9.3.4 *Discrete estimation formulae for aligned processes*

Suppose that the sample ccvf has its peak at lag S, where S could be positive or negative. Then the aligned estimate is based on the ccvf centered so that its peak is at zero lag. Thus the aligned ccvf estimate is

$$c_{12}''(k) = c_{12}(S + k). \quad (9.3.28)$$

If a lag window of length L is to be used with the aligned estimate, then

$$-L \leqslant k + S \leqslant L,$$

and it is necessary to compute $c_{12}(k)$ for

$$-L - S \leqslant k \leqslant L - S.$$

The usual formulae given in Section 9.3.1 may then be applied using the aligned ccvf estimate. The even and odd ccvf estimates (9.3.6) and (9.3.7) become

$$f_{12}''(k) = \tfrac{1}{2}\{c_{12}(S + k) + c_{12}(S - k)\}, \quad (9.3.29)$$

$$q_{12}''(k) = \tfrac{1}{2}\{c_{12}(S + k) - c_{12}(S - k)\}, \quad (9.3.30)$$

and (9.3.8) to (9.3.13) may then be used without alteration.

9.3.5 *Examples of cross spectral estimation with alignment*

Figure 9.16 shows mean smoothed coherency spectra based on $L = 8$, 16 and 32 for the linear filter example, after alignment with $S = 10$ based on the peak of the ccf in Figure 9.13. Comparison of Figure 9.16 with Figure 9.12, which shows the corresponding smoothed spectra before alignment, indicates that alignment results in a considerable reduction in bias, that is, an improvement in fidelity.

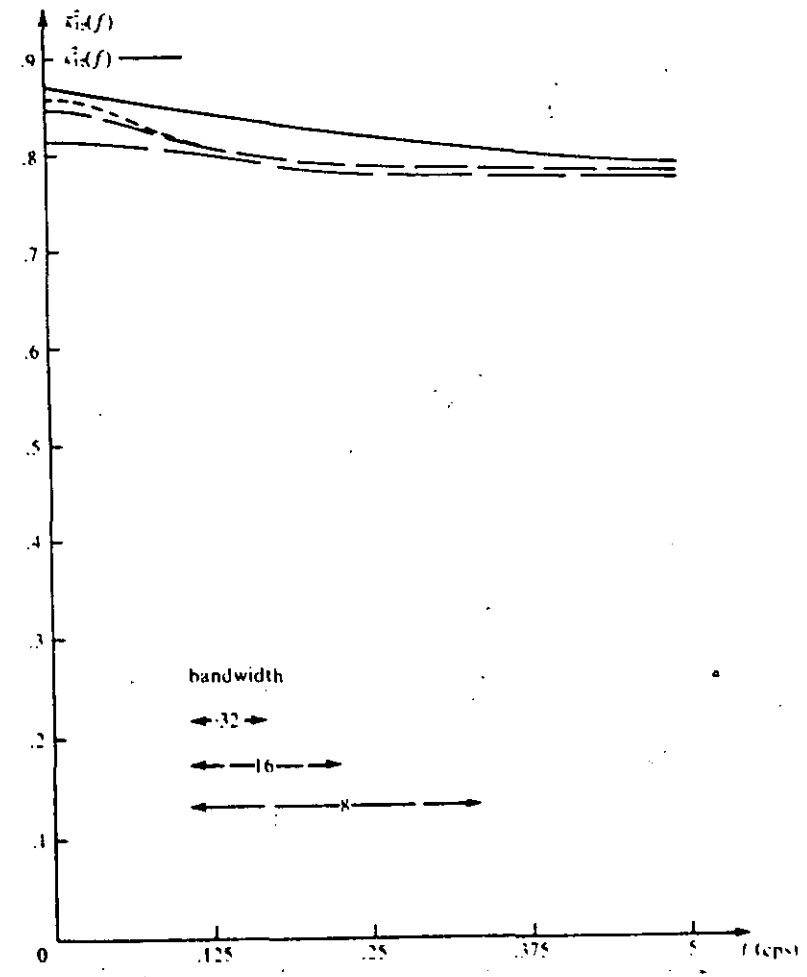


FIG. 9.16: Mean smoothed coherency spectra for the linear process (8.1.22) (after alignment, $S = 10$)

Figure 9.17 shows the mean smoothed phase spectra for $L = 4$ and 32. Comparison with the mean smoothed spectra before alignment, given in Figure 9.11, shows that much more rapid convergence to the true value is obtained with the aligned estimator.

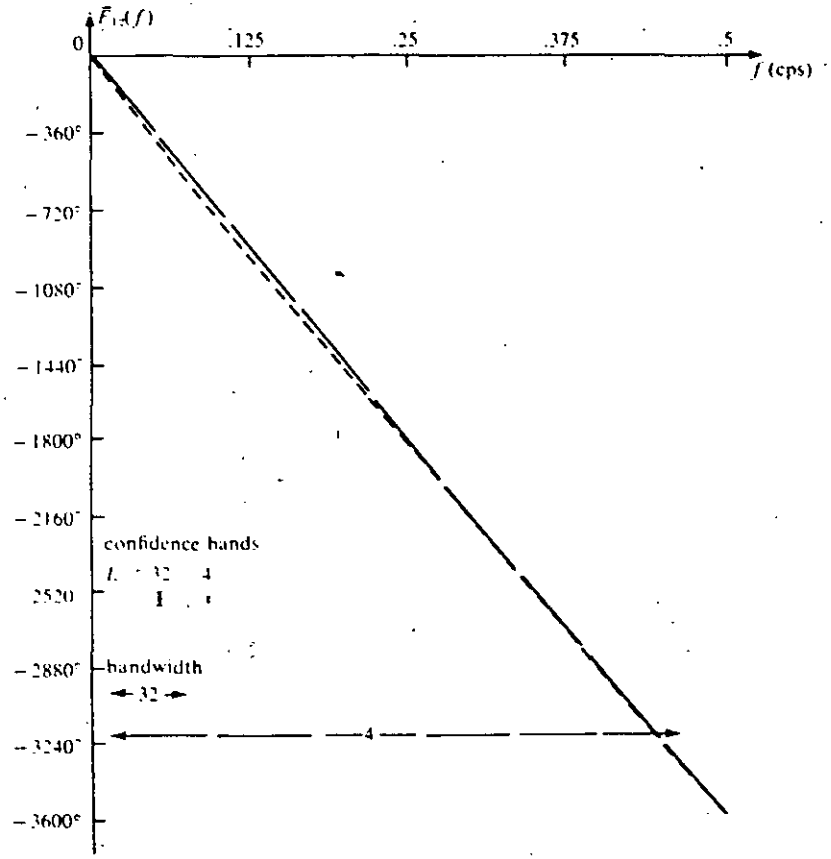


FIG. 9.17: Mean smoothed phase spectra for the linear process (8.1.22) (after alignment, $S = 10$)

Figure 9.18 shows the smoothed coherency estimates after alignment for the linear filter data. Comparison with the unaligned spectra of Figure 9.14 shows the marked improvement in the estimate. Note that oscillations begin to occur in the estimate based on $L = 16$ and hence the estimate based on $L = 8$ or possibly $L = 12$ would be accepted. Note also that the estimate for $L = 8$ is seen to agree very well with the theoretical coherency spectrum. The phase spectra estimates after alignment are not shown, since they agree so closely with the phase spectra shown in Figure 9.17. As noted in Section 9.3.3, the bias in the phase estimator is usually small, and so the difference between

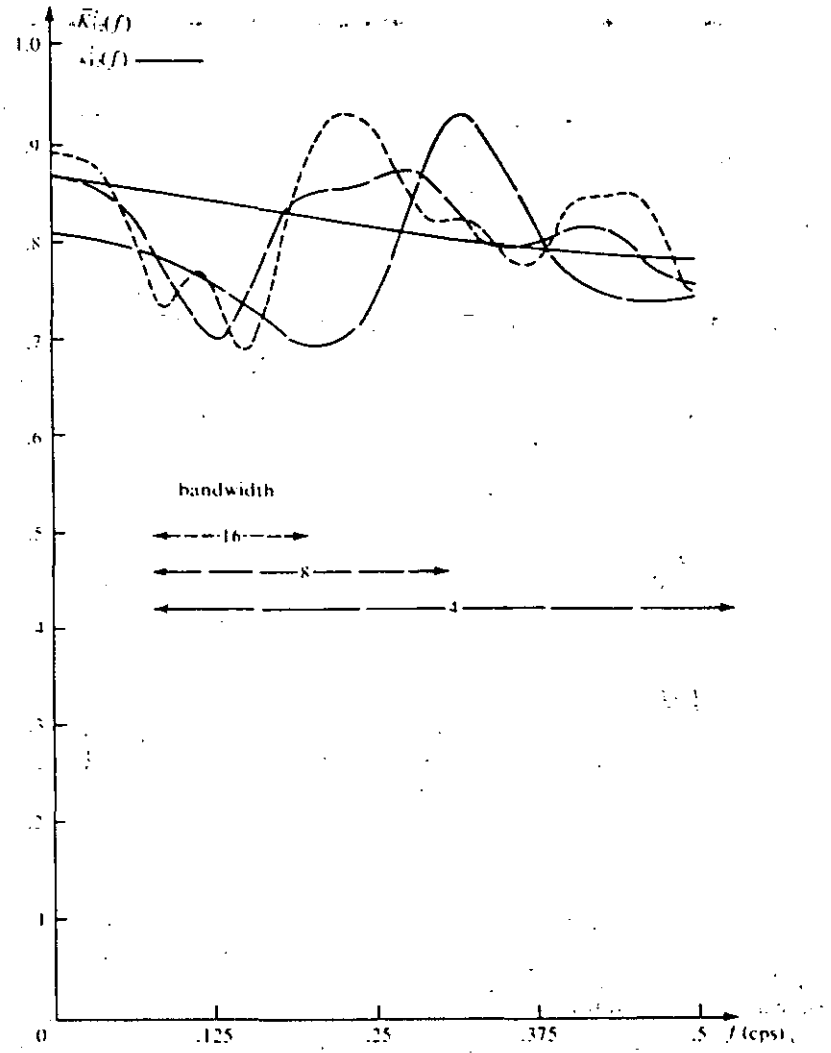


FIG. 9.18: Smoothed coherency estimates for the linear process (8.1.22) ($N = 100$, after alignment, $S = 10$)

the unaligned and aligned phase estimates is not as drastic as the difference between the unaligned and aligned coherency estimates.

Hence it can be concluded from this analysis that the process of alignment has transformed a very bad coherency estimate into a good estimate.

Confidence intervals for coherency and phase. The smoothed coherency estimates of Figure 9.18 have been replotted in Figure 9.19 on the transformed scale $Y_{12} = \operatorname{arctanh} |\bar{K}_{12}|$, and the constant confidence limits computed according

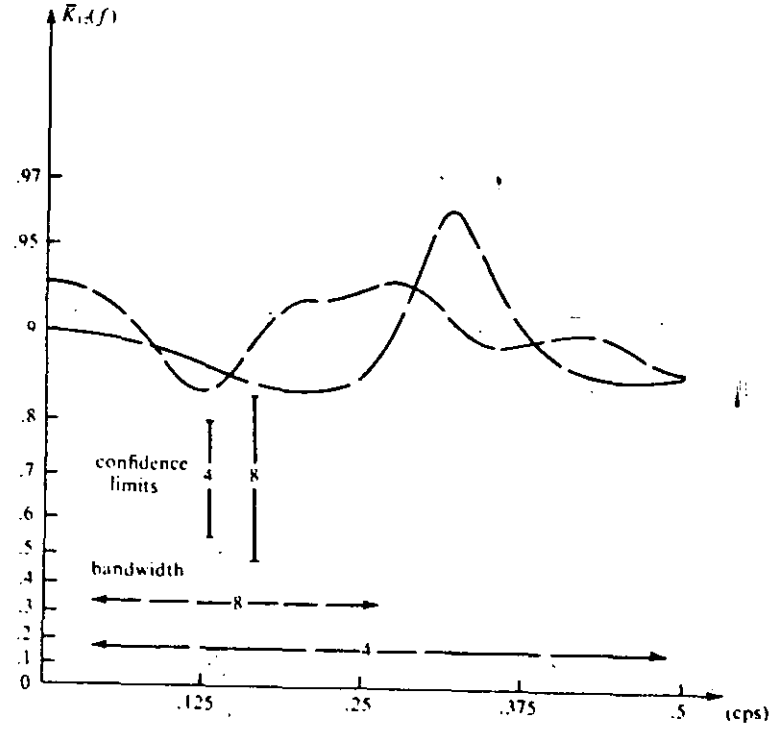


FIG. 9.19: Transformed coherency estimates for the linear process (8.1.22) ($N = 100$, after alignment, $S = 10$)

to (9.2.23) are shown. Since the bias before alignment is so small for the phase estimator, it is justifiable to attach confidence intervals to the unaligned spectra which are shown in Figure 9.17. Taking an average coherency figure of 0.8 over the whole frequency range, the 95% intervals for $L = 4$ and $L = 32$ are 5° and 15° using Figure 9.3.

9.4 PRACTICAL ASPECTS OF CROSS SPECTRAL ESTIMATION

This section contains a discussion of practical problems arising in the estimation of cross spectra. These are analogous to the problems arising in the estimation of autospectra discussed in Section 7.3.

9.4.1 The design of a cross spectral analysis

As in autospectral analysis, it is desirable to consider the choice of record length in advance of collecting the data. In all calculations of this kind, no precise statements can be made, since the best choice of record length depends on knowing the exact spectra. However, if guesses can be made of

how the spectra are likely to behave, the problem of data-gathering can be based on semiquantitative grounds rather than on wild guesses.

As discussed in Section 7.3.1, there are four stages in the design procedure for autospectra:

(1) The sampling interval must be at most $1/2f_0$ where f_0 cps is the maximum frequency of interest.

(2) Electronic filtering may be necessary before sampling the record to ensure that there is no power or cross power at frequencies above f_0 in the continuous traces.

(3) A guess must be made of the width a of the finest detail of interest in the coherency spectrum. As shown in Section 9.3, it is usually more difficult to estimate the coherency spectrum than the phase spectrum and hence the design calculations can be based on the coherency spectrum entirely. High fidelity can be achieved in estimating the coherency spectrum if the bandwidth of the spectral window is less than the width a of the finest detail. Hence the formula (7.3.2) for the number of lags L applies, namely,

$$L = \frac{b_1}{a\Delta},$$

where b_1 is the standardized bandwidth of the window. Note that the above calculation assumes that the two series have been aligned. Since the shift S required to align the series will usually be small compared with the total length of the series, this factor can be ignored at the design stage.

(4) To trust the fine detail in the coherency spectrum, it is necessary that the confidence interval be acceptably small. In Section 9.2.3 it was shown that confidence intervals for the coherency could be based on the result

$$\text{Var}\,[\text{arctanh}\,|\bar{K}_{12}|] \approx \frac{I}{2T} = \frac{1}{\nu} = \frac{L}{2b_1 N}.$$

where b_1 is the standardized bandwidth of the window. The 95% confidence limits for arctanh $|\bar{\kappa}_{12}|$ are given by (9.2.23), that is,

$$\pm 1.96\sqrt{\frac{L}{2b_1 N}} = \pm\frac{1.96}{\sqrt{\nu}}. \tag{9.4.1}$$

Hence, from the desired width of the confidence interval, say $\pm c$, the number of data points required to achieve this width on average is

$$N = \frac{L}{2b_1}\left(\frac{1.96}{c}\right)^2 \tag{9.4.2}$$

or, from (7.3.2),

$$N = \frac{1}{2a\Delta}\left(\frac{1.96}{c}\right)^2. \tag{9.4.3}$$

An example. Suppose that the phase and coherency spectra are needed up to $f_0 = 2$ cps. It is also required to estimate a peak of width $a = 0.20$ cps in the transformed coherency spectrum with a 95% confidence interval of ± 0.2 using the Tukey window.

Then

$$\Delta = \frac{1}{2f_0} = 0.25 \text{ seconds},$$

$$L = \frac{1.333}{(0.2)(0.25)} = 27,$$

$$N = \frac{1}{2(0.2)(0.25)}\left(\frac{1.96}{0.2}\right)^2 = 960.$$

The number of degrees of freedom with each estimate is approximately 95 and the record length should be at least 240 seconds.

9.4.2 *A practical procedure for estimating cross spectra*

It is suggested that the estimation of cross spectra should be conducted in five stages:

(1) Preliminary decision stage

(a) The series are inspected visually to see if there are any obvious trends. If trends are present they can be eliminated using the covariance estimates based on the first difference of each series or one of the digital filters of Section 7.3.5.

(b) It may be necessary to decide whether to analyze the data over a wide frequency range or whether to filter the data into component series relating to distinct frequency bands. If large variations of power can be expected over the total frequency range, then a separate analysis for the low-frequency and high-frequency components would be a minimum requirement. This decision requires some prior knowledge of the spectra. If this is not available, it may be necessary to conduct a pilot analysis. Alternatively, stages *(1)* through *(4)* should be completed and this analysis used as a basis for filtering, followed by reanalysis.

(c) The maximum number of lags $L_{\max}$ chosen for the calculation of the auto- and cross covariances is decided.

(2) First computation stage

(a) The auto- and cross covariances and correlations for the original and differenced data are computed, then plotted.

(b) Even though trends may not be visible in the data, they may be present. They can be detected by a failure of the auto- and cross covariances to damp out. As indicated in Section 7.3.5, trends produce large power at low frequencies which leaks into the spectrum at other frequencies and causes distortion of the spectrum. In cross spectral analysis it also produces spurious coherencies between the two series.

(3) Intermediate decision stage

(a) A decision is made whether to use the original covariances $c_{ij}(k)$ or the detrended covariances $c'_{ij}(k)$.

(b) The cross covariance or correlation function selected in (a) should be examined and the lag S corresponding to its maximum absolute value noted.

(c) A set of three truncation points $L_1 < L_2 < L_3$ should be chosen based on the way the auto- and cross correlation functions damp out.

(4) Second computation stage

(a) The two autospectra and the phase and coherency spectra based on the aligned cross correlations (9.3.28) are computed.

(b) The four spectra should be plotted for the set of three bandwidths. The two autospectra should be plotted on a logarithmic scale, the phase on a linear scale and the coherency on the transformed scale

$$Y = \operatorname{arctanh} |\bar{K}_{12}(f)|.$$

(5) Interpretation stage

(a) The phase spectrum is examined to see if further alignment is needed. If so, the second computing stage is repeated using the new alignment parameter determined from the phase spectrum.

(b) If no further alignment is necessary, the effects of the window-closing procedure should be appraised and the analysis classified as good, intermediate or poor, as described in Section 7.3.3. Final plots for presentation of the spectra should be prepared based on these decisions.

(c) Vertical lines representing the confidence intervals for phase can be added for each bandwidth, using Figure 9.3, and confidence intervals for coherency, using formula (9.2.23).

(d) Horizontal lines corresponding to the bandwidths of the spectral windows should be added so that the detail in the spectrum can be appraised.

9.4.3 *An account of a practical cross spectral estimation*

The procedure of Section 9.4.2 was applied to data, a section of which is shown in Figure 8.2. These data have been analyzed in [6] and a more detailed description will be given in Chapter 11. For the present it is sufficient to state that the variables used for cross spectral analysis are the input in-phase and out-of-phase currents $x_1(t)$, $x_2(t)$ to a turbo-alternator. It is necessary to compute the coherency and phase spectra of these two variables since both are input variables arising on an equal footing. This information will be required in Chapter 11 when it will be used in an input–output analysis to determine the frequency response functions of the turbo-alternator. The data consisted of 4000 pairs of data points read at 8 points per second.

(1) Preliminary decision stage

(a) Inspection of the data did not reveal any obvious trends. However, since the data covered such a long time period, it was expected that the analysis would have to be based on the detrended covariances.

(b) Since the data were read at 8 points per second, the Nyquist frequency is 4 cps. It was known *a priori* that there would be little power above 1 cps, and so the data were filtered to remove power above 1 cps using the filter

$$H(z) = \left\{\frac{1}{7}\sum_{k=-3}^{3} z^k\right\}^4.$$

Since there was negligible power above 1 cps in the filtered record, only two points per second were retained. Thus the final data consisted of 1000 pairs of data points. The first 100 values of the filtered current data are given in Table A11.1.

(c) A value of $L_{max} = 80$ was chosen initially.

(2) First computation stage

(a) The auto- and cross correlations of the data described under (*1-b*) were computed and plotted. Figure 9.20 shows the ccf estimates plotted up to $k = \pm 70$ lags. It is seen that the cross correlations damp out very slowly, as do the autocorrelations (not shown).

(b) The differenced cross correlations are also shown in Figure 9.20. It is seen that the cross correlation function drops to zero very quickly and

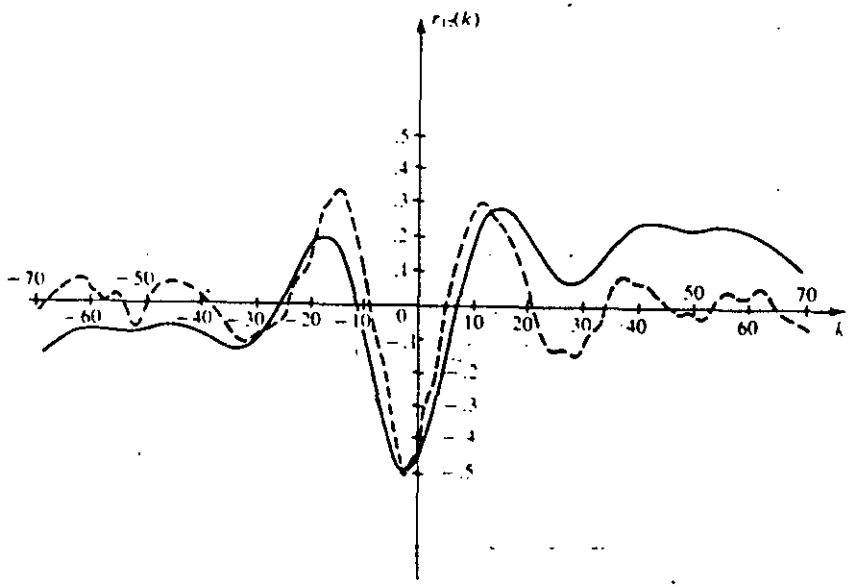


FIG. 9.20: Sample ccf's for original and differenced current data ($N = 1000$)

oscillates about zero with a well-defined period. The main characteristics of the cross correlation function are the delta-function type behavior near the origin and the subsequent periodic behavior. It is seen from Figure 9.20 that the low-frequency trend masks a great deal of detail in the ccf of the original data.

(3) Intermediate decision stage

(a) The above considerations indicate clearly that the differenced cross correlations should be used for spectral analysis.

(b) The ccf is almost symmetric about the origin. The peak occurs at lag -2, and so the shift parameter S was taken to be -2.

(c) Values of $L = 32$, 48 and 64 were chosen initially for the calculation of the spectra.

(4) Second computation stage

(a) The autospectra, the transformed coherency and phase spectra based on $S = -2$ were computed using the Tukey window.

(b) The autospectra are plotted in Figure 9.21 for $L = 64$. The transformed coherency spectra and the phase spectra are shown for $L = 32$ and 64 in Figures 9.22 and 9.23 respectively.

(5) Interpretation stage

(a) The phase spectrum of the aligned series shows no linear trend, and hence further alignment is not considered necessary.

(b) The window-closing procedure suggests that a lag of at least 32 is required for reasonable estimates of all four spectra. For example, Figure 9.22

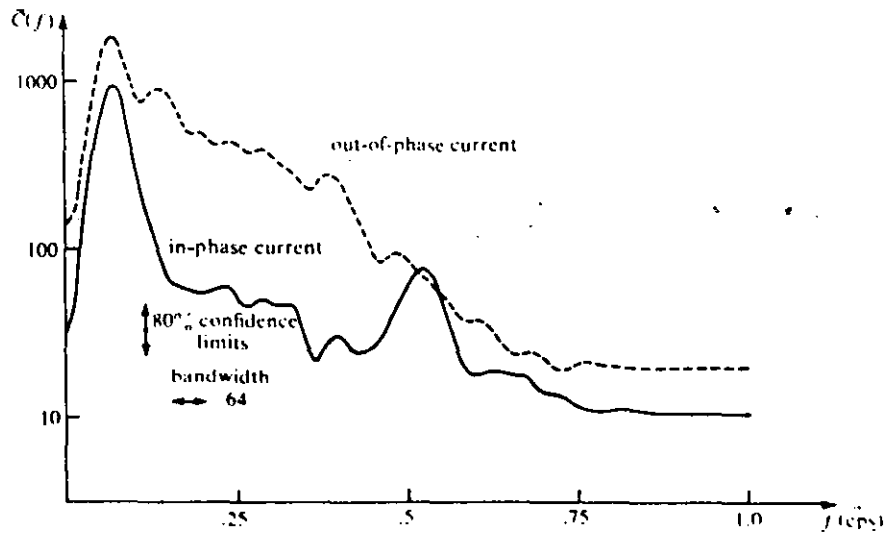


FIG. 9.21: Sample autospectra of differenced current data ($N = 1000$)

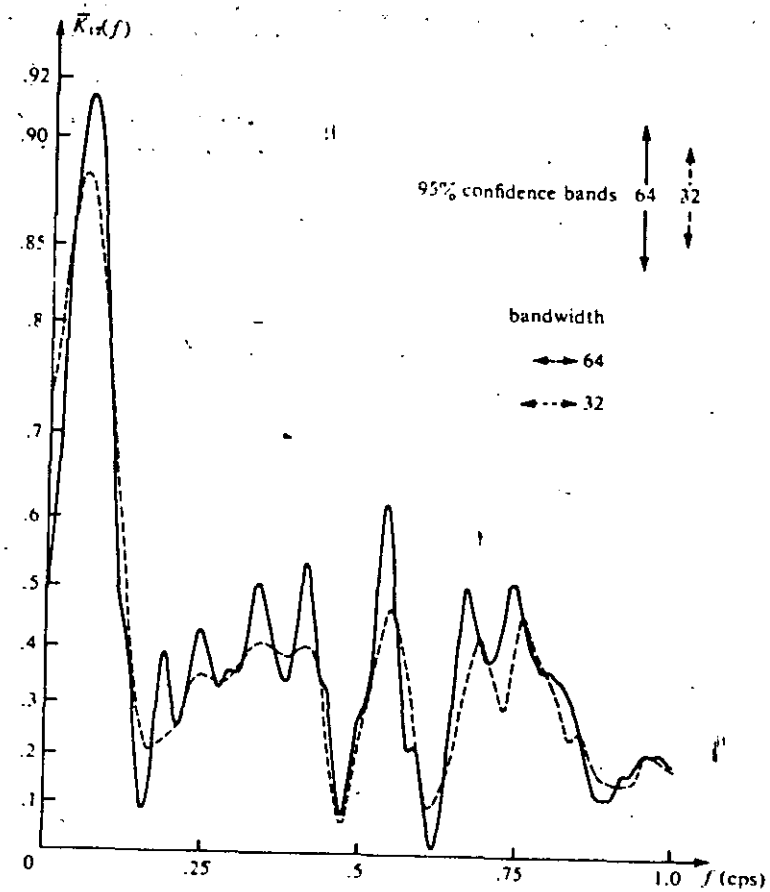


FIG. 9.22: Transformed coherency spectrum for differenced current data ($N = 1000$)

shows that halving the bandwidth by extending the lag from $L = 32$ to $L = 64$ does not change the broad features of the coherency spectrum. With $L = 64$, however, oscillations begin to appear due to instability. Hence a final value of $L = 48$ (not shown on Figure 9.22) is considered adequate. Similar considerations apply to the phase estimates shown in Figure 9.23.

(c) Confidence intervals for phase and transformed coherency were read off from Figure 9.3 and formula (9.2.23) using

$$\nu = \frac{2(1.33)1000}{L},$$

with $L = 32, 64$. The 95% confidence intervals for the transformed coherency can be converted into confidence intervals for $\kappa_{12}^2(f)$, since the coherency scale is marked on Figure 9.22. Since there are 1000 observations in the series, the confidence intervals are quite narrow.

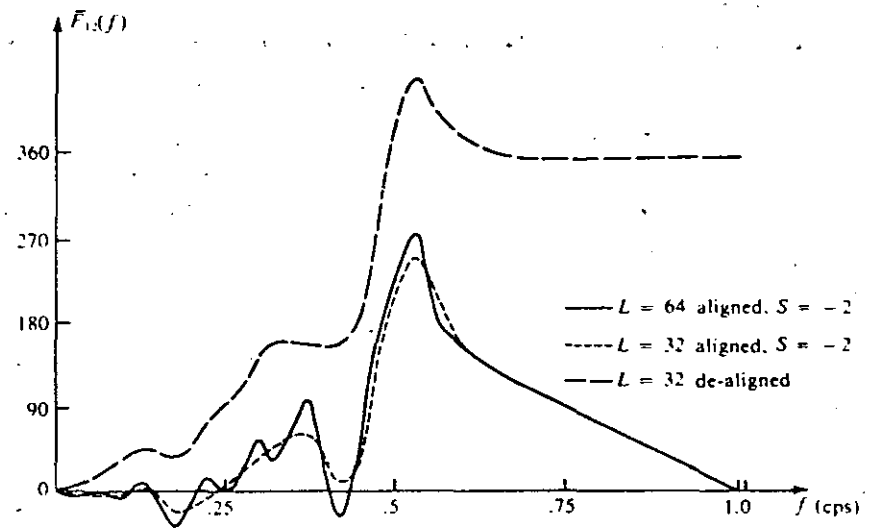


FIG. 9.23: Phase spectrum for differenced current data ($N = 1000$)

(d) Bandwidths are calculated from

$$b = \frac{1.33}{L\Delta}$$

and marked on the figures.

Conclusions. The main features of the cross spectral analysis are the presence of a large peak in the coherency spectrum in the neighborhood of 0.07 cps and a flat region with an average value of approximately $\bar{K}_{12}^2(f) = 0.18$ stretching over the whole frequency range. The large peak near 0.07 cps is associated with the facts that the spectra of both currents have peaks near this frequency and that the band 0 to 0.1 cps contains most of the power. Hence it would be expected that the in-phase and out-of-phase currents are highly correlated in this band. The phase spectra of Figure 9.23 show that the out-of-phase current leads the in-phase current by about 2 seconds. Figure 9.23 also shows the phase spectrum for $L = 32$ aligned and dealigned, that is, corrected by subtracting $2\pi fS\Delta$ from the aligned phase. It is seen that $S = -2$ very effectively removed the linear phase shift between the two currents.

REFERENCES

[1] N. R. Goodman, "On the joint estimation of the spectra, co-spectrum and quadrature spectrum of a two-dimensional stationary Gaussian process." Scientific Paper 10, Engineering Statistics Laboratory, New York University (1957).

[2] G. M. Jenkins, "Cross-spectral analysis and the estimation of linear open loop transfer functions." Chapter 18 of *Time Series*, M. Rosenblatt, ed. John Wiley, New York, 1963.

[3] A. Hald, *Statistical Theory with Engineering Applications.* John Wiley, New York, 1952, p. 609.
[4] C. W. J. Granger, *Spectral Analysis of Economic Time Series.* Princeton University Press, Princeton, 1964.
[5] H. Akaike and Y. Yamamouchi, "On the statistical estimation of frequency response functions." *Ann. Inst. Stat. Math.* **14**, 1 (1962).
[6] K. N. Stanton, "Measurement of turboalternator transfer functions using normal operating data." *Proc. Inst. Electr. Engrs.* **110**, 11 (1963).

APPENDIX A9.1 COVARIANCE OF COVARIANCE FUNCTION ESTIMATORS

In this appendix, a derivation is given of the covariance between the covariance function estimators (8.2.3). The covariance function estimators may be rewritten in the symmetric form

$$c_{ij}(u) = \begin{cases} \dfrac{1}{T}\displaystyle\int_{-(T-|u|)/2}^{(T-|u|)/2} X_i\left(t - \frac{u}{2}\right) X_j\left(t + \frac{u}{2}\right) dt, & -T \leqslant u \leqslant T \\ 0, & |u| > T. \end{cases} \quad \text{(A9.1.1)}$$

It is assumed that the stochastic processes $X_i(t)$, $i = 1, 2, 3, 4$, have the properties

$$E[X_i(t)] = 0, \qquad i = 1, 2, 3, 4, \quad \text{(A9.1.2)}$$

$$\text{Cov}\,[X_i(t), X_j(t+u)] = \gamma_{ij}(u), \quad j = 1, 2, 3, 4, \quad -\infty \leqslant u \leqslant \infty \quad \text{(A9.1.3)}$$

and

$$\text{Cov}\,[X_i(t)X_j(t+u_1), X_k(v)X_l(v+u_2)] = \gamma_{ik}(v-t)\gamma_{jl}(v-t+u_2-u_1) + \gamma_{il}(v-t+u_2)\gamma_{jk}(v-t-u_1) + K(v-t, u_1, u_2), \quad \text{(A9.1.4)}$$

where K is the fourth joint cumulant. The cross spectrum between the processes $X_i(t)$ and $X_j(t)$ is defined by

$$\Gamma_{ij}(f) = \int_{-\infty}^{\infty} \gamma_{ij}(u)\, e^{-j2\pi fu}\, du, \quad \text{(A9.1.5)}$$

$$\gamma_{ij}(u) = \int_{-\infty}^{\infty} \Gamma_{ij}(f)\, e^{j2\pi fu}\, df. \quad \text{(A9.1.6)}$$

Derivation of covariance. From (A9.1.1) and (A9.1.4), the covariance between the estimators $c_{ij}(u_1)$ and $c_{kl}(u_2)$ is

$$\text{Cov}\,[c_{ij}(u_1), c_{kl}(u_2)] = \frac{1}{T^2}\int_{-(T-|u_1|)/2}^{(T-|u_1|)/2}\int_{-(T-|u_2|)/2}^{(T-|u_2|)/2} \left\{\gamma_{ik}\left(v - t - \frac{u_2-u_1}{2}\right)\gamma_{jl}\left(v - t + \frac{u_2-u_1}{2}\right) + \gamma_{il}\left(v - t + \frac{u_2+u_1}{2}\right)\gamma_{jk}\left(v - t - \frac{u_2+u_1}{2}\right) + K(v-t, u_1, u_2)\right\} dv\, dt. \quad \text{(A9.1.7)}$$

Substituting $v - t = r$, $t = s$ transforms the region of integration from a rectangle to a parallelogram, as shown in Figure 5.11. This gives three regions of integration, designated (1), (2) and (3). Hence, for the case $|u_2| > |u_1|$, (A9.1.7) becomes

$$\begin{aligned}\text{Cov}\,[c_{ij}(u_1), c_{kl}(u_2)] = {} & \frac{1}{T^2}\int_{(|u_2|-|u_1|)/2}^{(T-|u_1|+|u_2|)/2} \gamma(r)\, dr \int_{-(T-|u_2|)/2}^{(T-|u_2|)/2} ds \quad \text{[region (1)]} \\ & + \frac{1}{T^2}\int_{-(|u_2|-|u_1|)/2}^{(|u_2|-|u_1|)/2} \gamma(r)\, dr \int_{-((T-|u_2|)/2)-r}^{((T-|u_2|)/2)-r} ds \quad \text{[region (2)]} \\ & + \frac{1}{T^2}\int_{-(T-(|u_1|+|u_2|)/2)}^{-(|u_2|-|u_1|)/2} \gamma(r)\, dr \int_{-(T-|u_1|)/2}^{(T-|u_1|)/2} ds \quad \text{[region (3)]}\end{aligned} \quad \text{(A9.1.8)}$$

where

$$\gamma(r) = \gamma_{ik}\left(r - \frac{u_2-u_1}{2}\right)\gamma_{jl}\left(r + \frac{u_2-u_1}{2}\right) + \gamma_{il}\left(r + \frac{u_2+u_1}{2}\right)\gamma_{jk}\left(r - \frac{u_2+u_1}{2}\right) + K(r, u_1, u_2)$$

Integrating (A9.1.8) with respect to s and combining terms gives

$$\text{Cov}\,[c_{ij}(u_1), c_{kl}(u_2)] = \frac{1}{T^2}\left\{T'\int_{-T'}^{T'}\gamma(r)\left(1 - \frac{|r|}{T'}\right)dr - T''\int_{-T''}^{T''}\gamma(r)\left(1 - \frac{|r|}{T''}\right)dr\right\} \quad \text{(A9.1.9)}$$

where

$$T' = T - \frac{|u_1| + |u_2|}{2}, \qquad T'' = \frac{|u_2| - |u_1|}{2}.$$

When $|u_1| > |u_2|$, the result is (A9.1.9) with $T'' = (|u_1| - |u_2|)/2$.

Simplification of result. Consider now the contribution due to the fourth cumulant term $K(r, u_1, u_2)$. For X_i Normal, $K = 0$, and so the following results are exact for Normal processes. For non-Normal stochastic processes which are linear processes of the form (5.2.6,7), the contribution due to $K(r, u_1, u_2)$ from an integral of the form

$$\int_{-T'}^{T'} K(r, u_1, u_2)\left(1 - \frac{|r|}{T'}\right)dr$$

is of the order $\gamma_{ij}(u_1)\gamma_{kl}(u_2)$, using (5.2.15). Hence the contribution from the fourth cumulant term may be neglected in (A9.1.9).

For T large, the terms of order $1/T^2$ may also be neglected, and (A9.1.9) reduces to

$$\begin{aligned}\text{Cov}\,[c_{ij}(u_1), c_{kl}(u_2)] &\approx \frac{1}{T}\int_{-\infty}^{\infty}\gamma(r)\, dr \\ &= \frac{1}{T}\int_{-\infty}^{\infty}\left\{\gamma_{ik}\left(r - \frac{u_2-u_1}{2}\right)\gamma_{jl}\left(r + \frac{u_2-u_1}{2}\right) + \gamma_{il}\left(r + \frac{u_2+u_1}{2}\right)\gamma_{jk}\left(r - \frac{u_2+u_1}{2}\right)\right\} dr.\end{aligned} \quad \text{(A9.1}$$

Setting $i = j = k = l = 1$ in (A9.1.10) and making the transformation $r = s - (u_2 - u_1)/2$ gives the Bartlett formula (5.3.22). In particular, for $u_2 = u_1$,

$$\text{Var}\,[c_{11}(u)] \approx \frac{1}{T}\int_{-\infty}^{\infty} \gamma_{11}^2(r)\,dr. \tag{A9.1.11}$$

Covariance of spectral estimators. The covariance between spectral estimators $C_{ij}(f_1)$ and $C_{kl}(f_2)$ may be derived from the results of the preceding section as follows:

$$\begin{aligned}\text{Cov}\,[C_{ij}(f_1), C_{kl}(f_2)] &= \text{Cov}\left[\int_{-T}^{T} c_{ij}(u_1)\,e^{-i2\pi f_1 u_1}\,du_1, \int_{-T}^{T} c_{kl}(u_2)\,e^{-i2\pi f_2 u_2}\,du_2\right] \\ &= \int_{-T}^{T}\int_{-T}^{T} \text{Cov}\,[c_{ij}(u_1), c_{kl}(u_2)]\,e^{-i2\pi(f_1 u_1 + f_2 u_2)}\,du_1\,du_2.\end{aligned} \tag{A9.1.12}$$

To simplify (A9.1.12), an alternative form of (A9.1.10) is derived.

Alternative form for Cov $[c_{ij}(u_1), c_{kl}(u_2)]$. From (A9.1.9) and (A9.1.6)

$$\begin{aligned}\text{Cov}\,[c_{ij}(u_1), c_{kl}(u_2)] &= \frac{1}{T^2}\left\{T'\int_{-T'}^{T'} \gamma(r)\left(1 - \frac{|r|}{T'}\right)dr - T''\int_{-T''}^{T''} \gamma(r)\left(1 - \frac{|r|}{T''}\right)dr\right\} \\ &= \frac{1}{T^2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \{\Gamma_{ik}(g_1)\Gamma_{jl}(g_2)\,e^{i\pi u_1(g_1 - g_2)}\,e^{-i\pi u_2(g_1 - g_2)} \\ &\qquad + \Gamma_{il}(g_1)\Gamma_{jk}(g_2)\,e^{i\pi u_1(g_1 - g_2)}\,e^{i\pi u_2(g_1 - g_2)}\} \\ &\quad \times \left[T'\int_{-T'}^{T'} e^{i2\pi r(g_1 + g_2)}\left(1 - \frac{|r|}{T'}\right)dr \right. \\ &\quad \left. - T''\int_{-T''}^{T''} e^{i2\pi r(g_1 + g_2)}\left(1 - \frac{|r|}{T''}\right)dr\right]dg_1\,dg_2.\end{aligned}$$

This reduces to

$$\frac{1}{T^2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \{\Gamma\}\,\frac{\sin^2\pi(g_1 + g_2)T' - \sin^2\pi(g_1 + g_2)T''}{\pi^2(g_1^2 + g_2^2)}\,dg_1\,dg_2,$$

where $\{\Gamma\}$ stands for the quantity in braces above. But

$$\sin^2 aT' - \sin^2 aT'' = \sin a(T - |u_1|)\sin a(T - |u_2|),$$

and hence, substituting $g_1 = f + g$, $g_2 = f - g$ so that $dg_1\,dg_2 = 2\,df\,dg$,

$$\begin{aligned}\text{Cov}\,[c_{ij}(u_1), c_{kl}(u_2)] &= \frac{2}{T^2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \frac{\sin 2\pi f(T - |u_1|)}{2\pi f}\,\frac{\sin 2\pi f(T - |u_2|)}{2\pi f} \\ &\quad \times \{\Gamma_{ik}(f + g)\Gamma_{jl}(f - g)\,e^{i2\pi g(u_1 - u_2)} + \Gamma_{il}(f + g)\Gamma_{jk}(f - g)\,e^{i2\pi g(u_1 + u_2)}\}\,df\,dg,\end{aligned} \tag{A9.1.13}$$

which provides an alternative form for (A9.1.9) and is exact if the X_t are Normal.

For $|u_1|$ and $|u_2|$ small and T large, integrating (A9.2.2) over f gives

$$\begin{aligned}&\text{Cov}\,[c_{ij}(u_1), c_{kl}(u_2)] \\ &\quad = \frac{1}{T}\int_{-\infty}^{\infty} \{\Gamma_{ik}(g)\Gamma_{jl}(-g)\,e^{i2\pi g(u_1 - u_2)} + \Gamma_{il}(g)\Gamma_{jk}(-g)\,e^{i2\pi g(u_1 + u_2)}\}\,dg,\end{aligned} \tag{A9.1.14}$$

since $[(\sin 2\pi fT)/2\pi fT]^2$ tends to $(1/2T)\,\delta(f)$ when T is large.

Substituting (A9.1.13) in (A9.1.12),

$$\begin{aligned}&\text{Cov}\,[C_{ij}(f_1), C_{kl}(f_2)] \\ &\quad = \int_{-T}^{T}\int e^{-i2\pi(f_1 u_1 + f_2 u_2)}\,\frac{2}{T^2}\int_{-\infty}^{\infty}\int \frac{\sin 2\pi f(T - |u_1|)}{2\pi f}\,\frac{\sin 2\pi f(T - |u_2|)}{2\pi f} \\ &\qquad \times \{\Gamma_{ik}(f + g)\Gamma_{jl}(f - g)\,e^{i2\pi g(u_1 - u_2)} + \Gamma_{il}(f + g)\Gamma_{jk}(f - g)\,e^{i2\pi g(u_1 + u_2)}\} \\ &\qquad \times df\,dg\,du_1\,du_2.\end{aligned}$$

Interchanging orders of integration, collecting terms and simplifying

$$\begin{aligned}&\text{Cov}\,[C_{ij}(f_1), C_{kl}(f_2)] \\ &\quad = \frac{1}{T^2}\int_{-\infty}^{\infty} \Gamma_{ik}(x)\,\frac{\sin\pi T(f_1 - x)}{\pi(f_1 - x)}\,\frac{\sin\pi T(f_2 + x)}{\pi(f_2 + x)}\,dx\int_{-\infty}^{\infty} \Gamma_{jl}(-y) \\ &\qquad \times \frac{\sin\pi T(f_1 + y)}{\pi(f_1 + y)}\,\frac{\sin\pi T(f_2 - y)}{\pi(f_2 - y)}\,dy + \frac{1}{T^2}\int_{-\infty}^{\infty} \Gamma_{il}(x)\,\frac{\sin\pi T(f_1 - x)}{\pi(f_1 - x)} \\ &\qquad \times \frac{\sin\pi T(f_2 - x)}{\pi(f_2 - x)}\int_{-\infty}^{\infty} \Gamma_{jk}(-y)\,\frac{\sin\pi T(f_1 + y)}{\pi(f_1 + y)}\,\frac{\sin\pi T(f_2 + y)}{\pi(f_2 + y)}\,dy.\end{aligned} \tag{A9.1.15}$$

Simplification of result. The result (A9.1.15) is exact for Normal rv's. It may be simplified for random processes whose spectra are approximately constant over the range f_1 to f_2, since then the terms $\Gamma_{ik}(f_1)$ may be taken outside the integral. Hence for noise which is approximately white in the frequency range f_1 to f_2

$$\begin{aligned}\text{Cov}\,[C_{ij}(f_1), C_{kl}(f_2)] &\approx \Gamma_{ik}(f_1)\Gamma_{jl}(-f_1)\left\{\frac{\sin\pi T(f_1 + f_2)}{\pi T(f_1 + f_2)}\right\}^2 \\ &\quad + \Gamma_{il}(f_1)\Gamma_{jk}(-f_1)\left\{\frac{\sin\pi T(f_1 - f_2)}{\pi T(f_1 - f_2)}\right\}^2.\end{aligned} \tag{A9.1.16}$$

Equation (A9.1.16) is exact for Normal white noise processes. For non-white and/or non-Normal processes, the result is only approximate.

All the above results apply in the discrete case, except that the terms in $\Gamma(f)$ are multiplied by Δ, and the frequency range is $-1/2\Delta \leq f < 1/2\Delta$. Thus for the discrete case, (A9.1.16) becomes

$$\begin{aligned}\text{Cov}\,[C_{ij}(f_1), C_{kl}(f_2)] &\approx \Delta^2\Gamma_{ik}(f_1)\Gamma_{jl}(-f_1)\left\{\frac{\sin\pi N\Delta(f_1 + f_2)}{N\sin\pi\Delta(f_1 + f_2)}\right\}^2 \\ &\quad + \Delta^2\Gamma_{il}(f_1)\Gamma_{jk}(-f_1)\left\{\frac{\sin\pi N\Delta(f_1 - f_2)}{N\sin\pi\Delta(f_1 - f_2)}\right\}^2.\end{aligned} \tag{A9.1.17}$$

Setting $X_i = X_j = X_k = X_l = Z$ in (A9.1.16) and (A9.1.17) gives the results (6.3.17) and (6.3.15) for white noise, and more generally (6.4.9) for non-white noise.

The results (A9.1.16) and (A9.1.17) show that for T large, the covariance between two spectral estimators is of order $1/T^2$ except when $f_1 = f_2$. Hence to a good degree of approximation, spectral estimators at spacing greater than $1/T$ may be regarded as uncorrelated.

Generalized covariance matrix. The general results (A9.1.16) and (A9.1.17) may be used to derive the generalized covariance matrix (9.1.22). For example, the term Cov $[L_{12}(f), Q_{12}(f)]$ may be derived as follows:

$$L_{12}(f) = \tfrac{1}{2}\{C_{12}(f) + C_{12}(-f)\} = \tfrac{1}{2}\{C_{12}(f) + C_{21}(f)\},$$

$$Q_{12}(f) = \frac{1}{2j}\{C_{12}(f) - C_{12}(-f)\} = \frac{1}{2j}\{C_{12}(f) - C_{21}(f)\}.$$

Hence,

$$\begin{aligned}\text{Cov}\,[L_{12}(f), Q_{12}(f)] &= \text{Cov}\left[\frac{C_{12}(f) + C_{21}(f)}{2}, \frac{C_{12}(f) - C_{21}(f)}{2j}\right] \\ &= \frac{1}{4j}\{\text{Cov}\,[C_{12}(f), C_{12}(f)] - \text{Cov}\,[C_{12}(f), C_{21}(f)] \\ &\quad + \text{Cov}\,[C_{21}(f), C_{12}(f)] - \text{Cov}\,[C_{21}(f), C_{21}(f)]\}.\end{aligned}$$

Substituting for Cov $[C_{ij}(f), C_{kl}(f)]$ from (9.1.15) yields

$$\begin{aligned}\text{Cov}\,[L_{12}(f), Q_{12}(f)] = \frac{1}{4j}\{&\Gamma_{11}(f)\Gamma_{22}(-f)W^2(+) + \Gamma_{12}(f)\Gamma_{21}(-f)W^2(-) \\ &- \Gamma_{12}(f)\Gamma_{21}(-f)W^2(+) - \Gamma_{11}(f)\Gamma_{22}(-f)W^2(-) \\ &+ \Gamma_{21}(f)\Gamma_{12}(-f)W^2(+) + \Gamma_{22}(f)\Gamma_{11}(-f)W^2(-) \\ &- \Gamma_{22}(f)\Gamma_{11}(-f)W^2(+) - \Gamma_{21}(f)\Gamma_{12}(-f)W^2(-)\},\end{aligned}$$

using the notation of (9.1.13) and (9.1.14). This reduces to

$$\begin{aligned}&\text{Cov}\,[L_{12}(f), Q_{12}(f)] \\ &\quad = \frac{1}{4j}\{W^2(+)(-\Gamma_{12}^2(f) + \Gamma_{12}^2(-f)) + W^2(-)(\Gamma_{12}^2(f) - \Gamma_{12}^2(-f))\}.\end{aligned}$$

Since

$$\begin{aligned}&\Gamma_{12}^2(f) - \Gamma_{12}^2(-f) \\ &\quad = \Lambda_{12}^2(f) + 2j\Lambda_{12}(f)\Psi_{12}(f) - \Psi_{12}^2(f) - \{\Lambda_{12}^2(f) - 2j\Lambda_{12}(f)\Psi_{12}(f) - \Psi_{12}^2(f)\} \\ &\quad = 4j\Lambda_{12}(f)\Psi_{12}(f),\end{aligned}$$

on neglecting the terms in $W^2(+)$,

$$\text{Cov}\,[L_{12}(f), Q_{12}(f)] \approx \Lambda_{12}(f)\Psi_{12}(f)W^2(-).$$

Chi-squared properties of autospectral estimators. Consider the stochastic process $X_1(t)$, where $X_1(t)$ is Normal with mean zero and variance σ_1^2. The spectral estimator $C_{11}(f)$ is

$$\begin{aligned}C_{11}(f) &= \frac{1}{T}\left|\int_{-T/2}^{T/2} X_1(t)\, e^{-j2\pi ft}\, dt\right|^2 \\ &= \frac{1}{T}\left[\left\{\int_{-T/2}^{T/2} X_1(t)\cos 2\pi ft\, dt\right\}^2 + \left\{\int_{-T/2}^{T/2} X_1(t)\sin 2\pi ft\, dt\right\}^2\right] \\ &= \frac{1}{T}\{X_c^2(f) + X_s^2(f)\}, \qquad \text{(A9.1.18)}\end{aligned}$$

where $X_c(f)$ and $X_s(f)$ are the Fourier cosine and sine transforms of $X_1(t)$ respectively. Since the Fourier transform is a linear operation and $X_1(f)$ is a Normal process, it follows that $X_c(f)$ and $X_s(f)$ are Normal. But from (A9.1.18),

$C_{11}(f)$ is a sum of two squared Normal rv's and hence $C_{11}(f)$ is distributed as a chi-squared rv with two degrees of freedom. The results of Section 6.3.3 follow immediately.

For non-Normal processes, the Fourier transforms $X_c(f)$ and $X_s(f)$ are weighted sums over time t of the random variables X_t. Hence the above results will be approximately true for non-Normal processes.

Covariance of smoothed spectral estimators. The smoothed spectral estimator $\bar{C}_{ij}(f)$ is related to $C_{ij}(f)$ by

$$\bar{C}_{ij}(f) = \int_{-\infty}^{\infty} C_{ij}(g)W(f-g)\,dg, \qquad \text{(A9.1.19)}$$

where $W(f)$ is the spectral window corresponding to the lag window $w(t)$. Hence

$$\bar{C}_{ij}(f) = \int_{-T}^{T} c_{ij}(u)w(u)\, e^{-j2\pi fu}\, du. \qquad \text{(A9.1.20)}$$

The covariance between smoothed spectral estimators may be derived as follows. Since

$$\bar{c}_{ij}(u) = c_{ij}(u)w(u), \qquad \text{(A9.1.21)}$$

$$\text{Cov}\,[\bar{c}_{ij}(u_1), \bar{c}_{kl}(u_2)] = w(u_1)w(u_2)\,\text{Cov}\,[c_{ij}(u_1), c_{kl}(u_2)]. \qquad \text{(A9.1.22)}$$

Hence, from (A9.1.12) and (A9.1.22),

$$\begin{aligned}&\text{Cov}\,[\bar{C}_{ij}(f_1), \bar{C}_{kl}(f_2)] \\ &\quad = \int_{-T}^{T}\int_{-T}^{T} w(u_1)w(u_2)\,\text{Cov}\,[c_{ij}(u_1), c_{kl}(u_2)]\, e^{-j2\pi(f_1u_1 + f_2u_2)}\, du_1\, du_2. \qquad \text{(A9.1.23)}\end{aligned}$$

Now $w(u)$ is zero for $|u| > M$, where $M \ll T$. Hence the approximation (A9.1.14) may be used for Cov $[c_{ij}(u_1), c_{kl}(u_2)]$ in (A9.1.23).

Thus,

$$\begin{aligned}\text{Cov}\,[\bar{C}_{ij}(f_1), \bar{C}_{kl}(f_2)] \approx &\int_{-T}^{T}\int \frac{w(u_1)w(u_2)}{T}\, e^{-j2\pi(f_1u_1 + f_2u_2)} \\ &\times \int_{-\infty}^{\infty} \{\Gamma_{ik}(g)\Gamma_{jl}(-g)\, e^{j2\pi g(u_1 - u_2)} \\ &+ \Gamma_{il}(g)\Gamma_{jk}(-g)\, e^{j2\pi g(u_1 - u_2)}\}\, dg\, du_1\, du_2. \qquad \text{(A9.1.24)}\end{aligned}$$

Interchanging orders of integration gives

$$\begin{aligned}&\text{Cov}\,[\bar{C}_{ij}(f_1), \bar{C}_{kl}(f_2)] \\ &\quad \approx \frac{1}{T}\int_{-\infty}^{\infty} W(f_1 - g)\{\Gamma_{ik}(g)\Gamma_{jl}(-g)W(f_2 + g) + \Gamma_{il}(g)\Gamma_{jk}(-g)W(f_2 - g)\}\, dg \\ &\qquad \text{(A9.1.25)}\end{aligned}$$

Making the assumption that the Γ's are smooth over the width of the spectral window allows us to remove the Γ's from the integration, and hence (A9.1.25) becomes

$$\begin{aligned}\text{Cov}\,[\bar{C}_{ij}(f_1), \bar{C}_{kl}(f_2)] \approx &\frac{\Gamma_{ik}(f_1)\Gamma_{jl}(-f_1)}{T}\int_{-\infty}^{\infty} W(f_1 - g)W(f_2 + g)\, dg \\ &+ \frac{\Gamma_{il}(f_1)\Gamma_{jk}(-f_1)}{T}\int_{-\infty}^{\infty} W(f_1 - g)W(f_2 - g)\, dg. \\ &\qquad \text{(A9.1.}\end{aligned}$$

Hence, provided the spectral windows are narrow and do not overlap very much, the covariance between smoothed spectral estimators will be very small.

For $f_1 = f_2$, (A9.1.26) becomes

$$\text{Cov}\,[\bar{C}_{ij}(f_1), \bar{C}_{kl}(f_1)] \approx \frac{\Gamma_{ik}(f_1)\Gamma_{jl}(-f_1)}{T}\int_{-\infty}^{\infty} W(f_1 - g)W(f_1 + g)\,dg + \frac{\Gamma_{il}(f_1)\Gamma_{jk}(-f_1)}{T}\int_{-\infty}^{\infty} W^2(f_1 - g)\,dg. \qquad \text{(A9.1.27)}$$

Neglecting the first term on the right-hand side of (A9.1.27), since it represents only a small amount of overlapping,

$$\text{Cov}\,[\bar{C}_{ij}(f_1), \bar{C}_{kl}(f_1)] \approx \frac{\Gamma_{il}(f_1)\Gamma_{jk}(-f_1)}{T}\int_{-\infty}^{\infty} W^2(g)\,dg \doteq \frac{\Gamma_{il}(f_1)\Gamma_{jk}(-f_1)}{T}\int_{-\infty}^{\infty} w^2(u)\,du, \qquad \text{(A.9.1.28)}$$

by Parseval's theorem. The result (A9.1.27) may be used to derive the generalized covariance matrix between smoothed spectral estimators. As stated in Section 9.2.1, this matrix is the same as the matrix (9.1.22) except that the multiplier $W^2(-)$ is replaced by I/T, where

$$I = \int_{-M}^{M} w^2(u)\,du.$$

APPENDIX A9.2 FLOW CHART FOR BIVARIATE SPECTRAL CALCULATIONS

The following is a flow chart for a computer program CROSSPEC which accepts auto- and cross covariance estimates COV(K,I,J) or DCOV(K,I,J), K = 0, MAXM, I = 1,2, J = 1,2 from program MULTICOR described in Appendix A5.3. Additional inputs are DELTA, NF, M ≤ MAXM − |S|, and S, the number of lags necessary to align the two processes so that the largest cross covariance is centered at zero. Output consists of printout of the covariances (echo check), the smoothed autospectra for each truncation point M, and the phase and squared coherency spectra, and plots of the logarithm of the autospectra, the squared coherency, and the phase spectrum with overlays for additional truncation points.

Program CROSSPEC

1) *Input* parameters N, MAXM, DELTA, NF, S.

2) *Read* IDENT(1), COV(K,1,1), K = 0, MAXM
   IDENT(2), COV(K,2,2), K = 0, MAXM
   COV(K,1,2), K = 0, MAXM
   COV(K,2,1), K = 0, MAXM.

3) *Read* M, *set* COV(K) = COV(K,1,1), K = 0, MAXM.

4) *Call* subroutine AUTOSPEC [Appendix A7.1].

5) *Store* SPEC(K,1) = SPEC(K), K = 0, NF.

6) *Set* COV(K) = COV(K,2,2), K = 0, MAXM.

7) *Call* subroutine AUTOSPEC.

8) *Store* SPEC(K,2) = SPEC(K), K = 0, NF.

9) *Call* subroutine EVOD.

Subroutine EVOD

Calculate EV(K) = COV(K + S,1,2) + COV(K − S,2,1), K = 0, M
OD(K) = COV(K + S,1,2) − COV(K − S,2,1), K = 0, M
(NOTE COV(K,1,2) = COV(−K,2,1)).

10) *Call* subroutine CROSPEC.

Subroutine CROSPEC

Calculate smoothed co- and quadrature spectra and the squared cross amplitude,

$$\text{COSPEC(I)} = 2.*\text{DELTA}*\left\{\text{EV(0)} + 2.*\sum_{K=1}^{M-1}\text{EV(K)W(K)}\,\text{COS}\,\frac{\pi\text{KI}}{\text{NF}}\right\}$$

$$\text{QSPEC(I)} = 4.*\text{DELTA}*\sum_{K=1}^{M-1}\text{OD(K)W(K)}\,\text{SIN}\,\frac{\pi\text{KI}}{\text{NF}}$$

$$\text{SQ(I)} = \text{COSPEC(I)}*\text{COSPEC(I)} + \text{QSPEC(I)}*\text{QSPEC(I)}.$$

As in subroutine AUTOSPEC, the Fourier transform can be performed very rapidly using either the fast Fourier transform or the algorithm listed below.

11) *Calculate* PHASE(K) = ARCTAN(−QSPEC(K)/COSPEC(K))
COHSQ(K) = SQ(K)/(SPEC(K,1)*SPEC(K,2)).

12) *Calculate* LOGSPEC(K,1) = LOG10(SPEC(K,1))
LOGSPEC(K,2) = LOG10(SPEC(K,2))
observing the cautionary notes in (5) of AUTOSPEC subroutine.

13) *Print* the smoothed autospectra SPEC(K,1), SPEC(K,2), the phase PHASE(K), the squared coherency COHSQ(K), the bandwidth B and the degrees of freedom D.

14) *Plot* and *overlay*: the LOGSPEC(K,1) for all values of M used
the LOGSPEC(K,2) for all values of M used
the PHASE(K)
the COHSQ(K).

Algorithm

To find COSPEC(I), QSPEC(I),

set $C = COS\ \frac{\pi I}{NF}$, V0 = 0., V1 = 0.

$SN = SIN\ \frac{\pi I}{NF}$, Z0 = 0., Z1 = 0.

Do 1, K = M − 1, 1

V2 = 2.*C*V1 − V0 + W(K)*EV(K)

Z2 = 2.*C*Z1 − Z0 + W(K)*OD(K)

V0 = V1

V1 = V2

Z0 = Z1

1 Z1 = Z2

COSPEC(I) = 2.*DELTA*(EV(0) + 2.*(V1*C − V0))

QSPEC(I) = 4.*DELTA*Z1*SN.

APPENDIX A9.3 SAMPLE CORRELATIONS FOR DATA OF APPENDIX A8.1

TABLE A9.1: Sample correlations for the data of Table A8.1

k								
0–7	1.000	0.505	0.071	−0.139	−0.137	−0.092	−0.124	−0.092
8–15	0.009	0.103	0.110	0.028	−0.097	−0.096	−0.025	0.011
16–23	−0.031	−0.019	0.116	0.195	0.138	0.049	−0.010	0.058
24–31	0.083	−0.031	−0.167	−0.140	−0.069	0.025	0.026	−0.080
	Autocorrelations $r_{11}(k)$							
0–7	−0.075	0.409	0.452	0.320	0.195	0.080	0.123	−0.004
8–15	−0.050	−0.035	0.056	0.133	0.082	−0.033	−0.098	−0.048
16–23	−0.039	−0.071	−0.121	−0.110	−0.059	0.044	0.072	0.156
24–31	0.234	0.245	0.153	0.026	−0.037	−0.027	−0.020	0.004
	Cross correlations $r_{12}(k)$							
0–7	−0.075	−0.498	−0.485	−0.278	−0.114	−0.050	0.042	0.109
8–15	0.087	−0.005	−0.066	−0.067	0.017	0.040	0.043	0.001
16–23	−0.049	0.085	0.175	0.101	−0.077	−0.102	−0.042	0.043
24–31	−0.014	−0.064	0.001	0.081	0.063	0.027	−0.064	−0.049
	Cross correlations $r_{21}(k)$							
0–7	1.000	0.536	0.186	−0.056	−0.121	−0.080	−0.066	−0.048
8–15	0.006	0.054	−0.020	0.005	−0.060	−0.042	0.018	0.025
16–23	−0.051	−0.084	−0.063	0.014	0.166	0.107	0.019	0.032
24–31	−0.063	−0.081	−0.155	−0.045	−0.049	−0.011	−0.080	−0.096
	Autocorrelations $r_{22}(k)$							

TABLE A9.2: Sample correlations for the data of Table A8.2

k							
0–7	1.000	0.540	0.335	0.219	0.262	0.274	0.151
8–15	0.030	0.058	0.017	−0.035	−0.113	−0.139	−0.101
16–23	−0.107	−0.123	−0.066	0.026	0.037	0.024	−0.027
24–31	0.059	0.052	−0.030	−0.043	−0.062	−0.055	−0.028
	Autocorrelations $r_{11}(k)$						
0–7	0.013	0.098	0.114	0.116	0.176	0.264	0.338
8–15	0.369	0.460	0.752	0.661	0.505	0.385	0.328
16–23	0.195	0.081	0.017	0.038	0.031	0.033	−0.012
24–31	−0.016	0.003	0.012	0.008	0.034	0.062	0.074
	Cross correlations $r_{12}(k)$						
0–7	0.013	−0.057	−0.131	−0.147	−0.152	−0.172	−0.171
8–15	−0.120	−0.089	−0.074	−0.071	−0.042	0.003	0.018
16–23	−0.101	−0.095	−0.105	−0.109	−0.143	−0.185	−0.209
24–31	−0.163	−0.079	−0.026	0.003	−0.013	−0.009	−0.007
	Cross correlations $r_{21}(k)$						
0–7	1.000	0.848	0.678	0.546	0.473	0.400	0.299
8–15	0.145	0.124	0.084	0.071	0.043	0.044	0.051
16–23	0.028	0.024	0.026	0.020	0.010	−0.012	−0.018
24–31	0.045	0.058	0.029	0.023	0.030	0.037	0.002
	Autocorrelations $r_{22}(k)$						

Equation (10.3.11) implies that the two degrees of freedom associated with $C_{zz}(f)$ are taken up entirely by the regression term $|X_1(f)|^2|\hat{H}(f) - H(f)|^2$. Thus a consequence of the narrow bandwidth associated with the sample spectrum and cross spectrum is that the information at a given frequency f is absorbed entirely in estimating the gain and phase and gives *no information* about the noise spectrum. It is now shown that this can be rectified by smoothing.

10.3.3 Smoothed least-squares analysis in the frequency domain

In this section it is shown that the least squares analysis of the previous section can be modified to give efficient smoothed estimators of the gain and phase of the linear system and also of the noise spectrum $\Gamma_{zz}(f)$. It is then shown how to use these estimators to derive an approximate test of significance for whether the true coherency is zero, as well as approximate confidence intervals for the gain and phase of the linear system.

Suppose that smoothed estimators of the gain and phase have been computed according to (10.3.7) and (10.3.8). Then (10.3.11) may be replaced by the corresponding smoothed formula

$$\bar{C}_{zz}(f) \approx \bar{C}_{\hat{z}\hat{z}}(f) + \bar{C}_{11}(f)|\bar{H}(f) - H(f)|^2. \tag{10.3.13}$$

In applying a smoothing operation to (10.3.11), it is assumed that the true frequency response function $H(f)$ remains approximately constant over the bandwidth of the spectral window.

As indicated in Section 9.2, this assumption will not be valid unless the two series have been aligned so that the ccf has its peak at zero lag. To derive the approximate distribution properties of the estimators it will now be assumed that the two series have been aligned.

Suppose that the smoothing procedure applied to the auto- and cross spectra produces ν degrees of freedom per estimate. Then

$$\frac{\nu\bar{C}_{zz}(f)}{\Gamma_{zz}(f)}$$

is distributed as a χ^2_ν rv, and (10.3.13) may be written

$$\frac{\nu\bar{C}_{zz}(f)}{\Gamma_{zz}(f)} \approx \frac{\nu\bar{C}_{\hat{z}\hat{z}}(f)}{\Gamma_{zz}(f)} + \frac{\nu\bar{C}_{11}(f)}{\Gamma_{zz}(f)}|\bar{H}(f) - H(f)|^2. \tag{10.3.14}$$

The decomposition (10.3.14) shows that the chi-squared rv with ν degrees of freedom on the left-hand side is decomposed into two component chi-squared rv's. The first component is distributed as $\chi^2_{\nu-2}$ and can be used to estimate the noise spectrum. The second component is distributed as χ^2_2 and can be used to estimate the gain and phase. This result follows from the fact that the right-hand term in (10.3.13) can be shown to have two degrees of freedom, regardless of the amount of smoothing, using the statistical differential technique of Section 3.2.5.

It may be further shown that the two components on the right-hand side of (10.3.14) are statistically independent. Hence, using the additive property for chi-squared rv's given in Section 3.3.5, the other component has $(\nu - 2)$ degrees of freedom.

Finally, using the smoothed version of (10.3.12), the smoothed noise spectral estimator may be calculated from

$$\bar{C}_{\hat{z}\hat{z}}(f) = \bar{C}_{22}(f)(1 - \bar{K}^2_{12}(f)). \tag{10.3.15}$$

The above results are now used to solve the following problem.

A test of significance for non-zero coherency. Suppose that $H(f)$ is identically zero for all f, that is, the input and output process are completely uncorrelated and hence the theoretical coherency is identically zero.

For $H(f) = 0$, the second term on the right-hand side of (10.3.14) is

$$\frac{\nu\bar{C}_{11}(f)\bar{G}^2(f)}{\Gamma_{zz}(f)} = \frac{\nu\bar{C}_{22}(f)\bar{K}^2_{12}(f)}{\Gamma_{zz}(f)}, \tag{10.3.16}$$

using (9.3.12) and (10.3.7).

The decomposition (10.3.14) shows that the rv (10.3.16) is approximately distributed as χ^2_2 and

$$\frac{\nu\bar{C}_{\hat{z}\hat{z}}(f)}{\Gamma_{zz}(f)} = \frac{\nu\bar{C}_{22}(f)(1 - \bar{K}^2_{12}(f))}{\Gamma_{zz}(f)}$$

is approximately distributed as $\chi^2_{\nu-2}$. Hence the rv

$$\frac{(\nu - 2)\bar{K}^2_{12}(f)}{2(1 - \bar{K}^2_{12}(f))} \tag{10.3.17}$$

is approximately distributed as $F_{2,\nu-2}$ if $H(f) = 0$.

As an example, suppose that an estimate of $\bar{K}^2_{12}(f) = 0.3$ has been obtained using a Tukey window with $\Delta = 1$, $L = 20$, $N = 100$. Then

$$\nu = \frac{8}{3}\frac{(100)}{20} = 13$$

and

$$\frac{(\nu - 2)\bar{K}^2_{12}}{2(1 - \bar{K}^2_{12})} = 2.36.$$

The upper 95% point $f_{2,11}(0.95)$ is 4.0, using Figure 3.12. Since the observed value of the criterion (10.3.17) is less than 4.0, it could be concluded that there is no evidence that the true coherency is different from zero.

It is only in rare situations that it will be necessary to apply a test of significance of the above type. For example, it could be used as a rough test for

cross correlation between time series which involves less computation than the test based on the integrated sample co-spectrum described in Section 9.1.2. It happens much more frequently that the object of the investigation is to *estimate* the gain and phase, and to give approximate *confidence intervals* for these functions. This is discussed in the following section.

10.3.4 Confidence intervals for gain and phase functions

The results of the preceding section may be used to derive approximate confidence intervals for gain and phase. From (10.3.14), it is seen that if $H(f)$ is known,

$$\Pr\left\{\frac{\nu-2}{2}\frac{\bar{C}_{11}(f)|\bar{H}(f)-H(f)|^2}{\bar{C}_{\tilde{z}\tilde{z}}(f)} \leqslant f_{2,\nu-2}(1-\alpha)\right\} = 1-\alpha.$$

Hence, if $H(f)$ is unknown, a confidence region for $H(f)$ with confidence coefficient $1-\alpha$ is given by

$$|\bar{H}(f)-H(f)|^2 \leqslant \frac{2}{\nu-2}\frac{\bar{C}_{\tilde{z}\tilde{z}}(f)}{\bar{C}_{11}(f)} f_{2,\nu-2}(1-\alpha). \qquad (10.3.18)$$

In terms of the gain and phase (10.3.18) may be written

$$(G\cos\phi-\bar{G}\cos\bar{F}_{12})^2+(G\sin\phi-\bar{G}\sin\bar{F}_{12})^2 \leqslant \frac{2}{\nu-2}\frac{\bar{C}_{\tilde{z}\tilde{z}}(f)}{\bar{C}_{11}(f)} f_{2,\nu-2}(1-\alpha), \qquad (10.3.19)$$

on dropping the dependence on frequency, f. In terms of the parameters $G\cos\phi$, $G\sin\phi$, (10.3.19) defines a circle of radius k as shown in Figure 10.1(a). When mapped into the (G,ϕ) plane, this becomes the region shown in Figure 10.1(b). As a rough approximation, this region can be enclosed by a rectangle. Thus the confidence interval for the gain is $OT \pm TR$ and the confidence interval for the phase is the angle between the tangents OP and OQ to the circle. This is the region

$$|G-\bar{G}| \leqslant k, \qquad \sin|\phi-\bar{\phi}| \leqslant \frac{k}{\bar{G}}.$$

Noting that

$$\frac{k^2}{\bar{G}^2} = \frac{2}{\nu-2}\frac{1-\bar{K}_{12}^2}{\bar{K}_{12}^2} f_{2,\nu-2}(1-\alpha),$$

the $100(1-\alpha)\%$ confidence intervals for G may be written in the alternative form

$$\bar{G}(f)\left\{1 \pm \sqrt{\frac{2}{\nu-2} f_{2,\nu-2}(1-\alpha)\left(\frac{1-\bar{K}_{12}^2(f)}{\bar{K}_{12}^2(f)}\right)}\right\}, \qquad (10.3.20)$$

and similarly, for ϕ,

$$\bar{F}_{12}(f) \pm \arcsin\sqrt{\frac{2}{\nu-2} f_{2,\nu-2}(1-\alpha)\left(\frac{1-\bar{K}_{12}^2(f)}{\bar{K}_{12}^2(f)}\right)}. \qquad (10.3.21)$$

Note that the confidence intervals are small when the number of degrees of freedom ν is large and when the coherency $\bar{K}_{12}^2(f)$ is large.

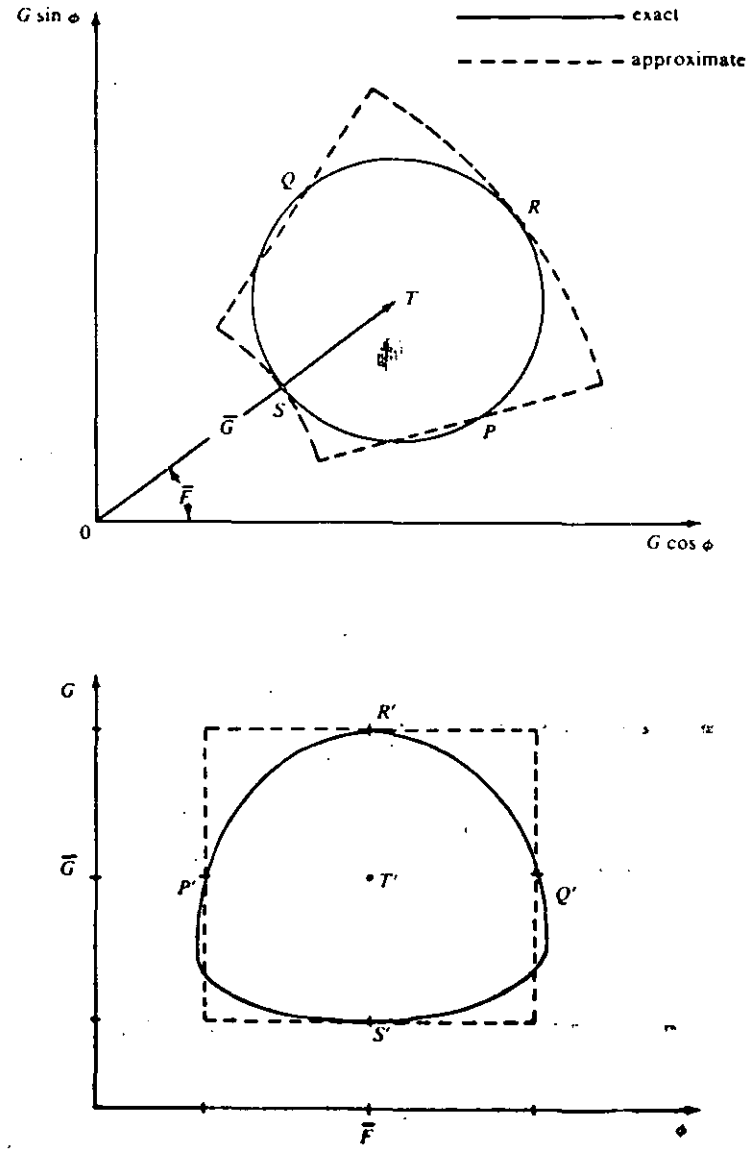


FIG. 10.1: Confidence regions for gain and phase

An example. Suppose that the estimated coherency is $\bar{K}_{12}^2 = 0.8$ and the gain and phase estimates are $\bar{G} = 10$, $\bar{F}_{12} = 70°$, based on $\nu = 17$. Then the 95% confidence intervals (10.3.20) and (10.3.21) are

$$10\left\{1 - \sqrt{\frac{2}{17}(3.68)\left(\frac{0.2}{0.8}\right)}\right\} \leqslant G \leqslant 10\left\{1 + \sqrt{\frac{2}{17}(3.68)\left(\frac{0.2}{0.8}\right)}\right\},$$

$$70 - \arcsin\sqrt{\frac{2}{17}(3.68)\left(\frac{0.2}{0.8}\right)} \leqslant \phi \leqslant 70 + \arcsin\sqrt{\frac{2}{17}(3.68)\left(\frac{0.2}{0.8}\right)},$$

that is,

$$6.7 \leqslant G \leqslant 13.3, \qquad 51° \leqslant \phi \leqslant 89°.$$

Bias in gain estimators. The confidence intervals (10.3.20) and (10.3.21) for the gain and phase were derived under the assumption that the estimators $\bar{G}(f)$ and $\bar{F}(f)$ are unbiased. As shown in Section 9.3.3, the bias in phase and coherency estimators can be minimized by aligning the two series. It is now shown that the bias in the gain estimator can also be reduced by aligning the series. A method of derivation similar to that in Section 9.3.3 is followed.

By definition, the bias in gain is

$$\begin{aligned} B(f) &= E[\bar{G}(f) - G(f)] \\ &= E\left[\frac{\bar{A}_{12}(f)}{\bar{C}_{11}(f)}\right] - \frac{\alpha_{12}(f)}{\Gamma_{11}(f)}. \end{aligned} \tag{10.3.22}$$

Assuming that the bias in the input spectrum is negligible, (10.3.22) becomes

$$B(f) \approx \frac{E[\bar{A}_{12}(f)] - \alpha_{12}(f)}{\Gamma_{11}(f)},$$

so that it is only necessary to determine the bias $E[\bar{A}_{12}(f)] - \alpha_{12}(f)$. Proceeding as in Section 9.3.3, the bias in the gain estimator is

$$B(f) \approx \frac{0.063}{M^2}\frac{1}{\Gamma_{11}}\{\alpha_{12}^{(2)} - \alpha_{12}(\phi_{12}^{(1)})^2 + 2\alpha_{12}^{(1)}\phi_{12}^{(1)}\sin 2\phi_{12}\}, \tag{10.3.23}$$

using the Tukey window.

Note that the term $\alpha_{12}(\phi_{12}^{(1)})^2$ occurs in the expression for the bias in gain, just as it did in the expression (9.3.26) for the bias in coherency. Hence the bias in gain is reduced by alignment for the same reasons as were stated in Section 9.3.

10.4 EXAMPLES OF FREQUENCY RESPONSE ESTIMATION

In this section the methods developed in Section 10.3 are used to estimate frequency response functions of simulated and real physical systems. First the stages in the estimation of frequency response functions are summarized.

10.4.1 A practical procedure for estimating frequency response functions

Discrete estimation formulae. The computations required in the estimation of frequency response functions are almost identical to those described for cross spectra in Section 9.3.1. The only additional calculations required are:

(1) The gain estimate

$$\bar{G}(f) = \frac{\bar{A}_{12}(f)}{\bar{C}_{11}(f)}. \tag{10.4.1}$$

(2) The residual or noise spectrum estimate

$$\bar{C}_{zz}(f) = \bar{C}_{22}(f)(1 - \bar{K}_{12}^2(f)). \tag{10.4.2}$$

(3) Approximate $100(1-\alpha)\%$ *confidence intervals for gain*

$$\bar{G}(f) \pm \bar{G}(f)\sqrt{\frac{2}{\nu - 2}f_{2,\nu-2}(1-\alpha)\left(\frac{1 - \bar{K}_{12}^2(f)}{\bar{K}_{12}^2(f)}\right)}, \tag{10.4.3}$$

where ν is the number of degrees of freedom associated with the smoothing of the output spectrum and $f_{2,\nu-2}(1-\alpha)$ is the upper $100(1-\alpha)\%$ point of the $F_{2,\nu-2}$ distribution.

(4) Approximate $100(1-\alpha)\%$ *confidence intervals for phase*

$$\bar{F}_{12}(f) \pm \arcsin\sqrt{\frac{2}{\nu - 2}f_{2,\nu-2}(1-\alpha)\left(\frac{1 - \bar{K}_{12}^2(f)}{\bar{K}_{12}^2(f)}\right)}. \tag{10.4.4}$$

Stages in the estimation of gain and phase. The five stages suggested in Section 9.4.2 for the estimation of cross spectra apply equally well to the estimation of gain and phase. The only major changes are to the second computation stage which should now read:

(4) Second computation stage

(a) The two autospectra, the noise spectrum and the phase and gain spectra based on the aligned cross correlations (9.3.28) are computed.

(b) The above spectra are plotted for a range of truncation points. The two autospectra and the noise spectrum should be plotted on logarithmic scales. Bode plots should be made for gain and phase, that is, the logarithm of gain should be plotted against the logarithm of frequency, and phase plotted against the logarithm of frequency.

(c) In some cases it may be necessary to plot the transformed coherency. The only other change is that in stage 5.2 the confidence intervals are obtained from (10.4.3) for gain, from (10.4.4) for phase and from (9.2.23) for coherency. A flow chart for frequency response calculations is given in Appendix A10.2.

10.4.2 Analysis of a simulated system

A linear second-order system was simulated in this experiment. The output data x_{2t} was generated according to the model (10.2.1), namely,

$$X_{2t} = X_t + Z_t$$

and

$$X_t = 0.25X_{t-1} - 0.5X_{t-2} + X_{1t}, \tag{10.4.5}$$

where Z_t is white noise and X_{1t} is the input.

The input data x_{1t} used was a realization of the second-order process (7.1.9), that is,

$$X_{1t} = X_{1t-1} - 0.5X_{1t-2} + Z'_t.$$

The input and output data are listed in Appendix A10.1.

The input process has the theoretical spectrum

$$\Gamma_{11}(f) = \frac{2}{2.25 - 3\cos 2\pi f + \cos 4\pi f}, \quad 0 \leqslant f \leqslant 0.5 \text{ cps}, \tag{10.4.6}$$

which is shown in Figure 7.7 as the solid curve. Since the noise spectrum is white,

$$\Gamma_{zz}(f) = 2, \quad 0 \leqslant f \leqslant 0.5. \tag{10.4.7}$$

The theoretical frequency response function of the linear system (10.4.5) is

$$H(f) = \frac{1}{1 - 0.25e^{-i2\pi f} + 0.5e^{-i4\pi f}}, \quad -\tfrac{1}{2} \leqslant f < \tfrac{1}{2}. \tag{10.4.8}$$

Hence the theoretical gain and phase functions are

$$G(f) = \frac{1}{\sqrt{1.3125 - 0.75\cos 2\pi f + \cos 4\pi f}}, \tag{10.4.9}$$

and

$$\phi(f) = \arctan \frac{\sin 2\pi f(4\cos 2\pi f - 1)}{2 + \cos 2\pi f(4\cos 2\pi f - 1)}. \tag{10.4.10}$$

The gain function (10.4.9) is shown as a solid curve in Figure 10.3, and the phase function (10.4.10) appears as a solid curve in Figure 10.4. The theoretical coherency is

$$\kappa_{12}^2 = \frac{1}{1 + (1.3125 - 0.75\cos 2\pi f + \cos 4\pi f)(2.25 - 3\cos 2\pi f + \cos 4\pi f)}.$$

This function is shown as the solid line in Figure 10.5. It is seen that the coherency is fairly high over the range 0 to 0.25 cps, where the input spectra and gain are high. However, the coherency is extremely small over the range 0.3 to 0.5 cps, where both the input spectrum and system gain are small. Hence it is to be expected that good estimates of gain, phase and coherency could be obtained for the frequency range 0 to 0.25 cps and poor estimates for the range 0.3 to 0.5 cps.

Since the data are simulated, many of the decisions required in the five-stage procedure of Sections 9.4.2 and 10.4.1 are not necessary. Nevertheless it is instructive to follow these stages through as if the data were obtained from a real process.

(1) Preliminary decision stage

(a) The data showed no trends, as would be expected since both the input and output were generated from stationary models. Hence no digital filtering to remove trends was needed.

(b) No filtering into distinct frequency bands was needed.

(c) Since the input and output series contained $N = 100$ terms, the auto- and cross correlations were evaluated initially up to 30 lags.

(2) First computation stage

The auto- and cross correlations were computed for the original and differenced data. The correlations of the original data are shown in Figure 10.2, and the covariances are given in Appendix A10.2.

(3) Intermediate decision stage

(a) The original correlations were used for the spectral analysis.

(b) Since the cross correlation function had its maximum value at zero lag, no alignment was required.

(c) Based on the way the correlation functions damp out, truncation points of $L = 8$, 12 and 16 were used for the subsequent analysis.

(4) Second computation stage

(a) The autospectra, gain, phase, coherency and noise spectra based on the Tukey window were calculated from the non-aligned correlations.

(b) The gain, phase, squared coherency and noise spectra are plotted in Figures 10.3, 10.4, 10.5 and 10.6 respectively for $L = 8$, 12 and 16.

(5) Interpretation stage

(a) The window-closing procedure shows that there is some change in all spectra when L is increased from 8 to 12 but little change when L is increased to 16. Hence the final spectra accepted were those based on $L = 16$. Note that the gain, phase and squared coherency estimates agree very well with the theoretical values in the range 0 to 0.35 cps but that large discrepancies appear above this frequency because the coherency drops sharply. The residual spectrum of Figure 10.6 shows excellent agreement with the theoretical value over the whole range 0 to 0.5 cps.

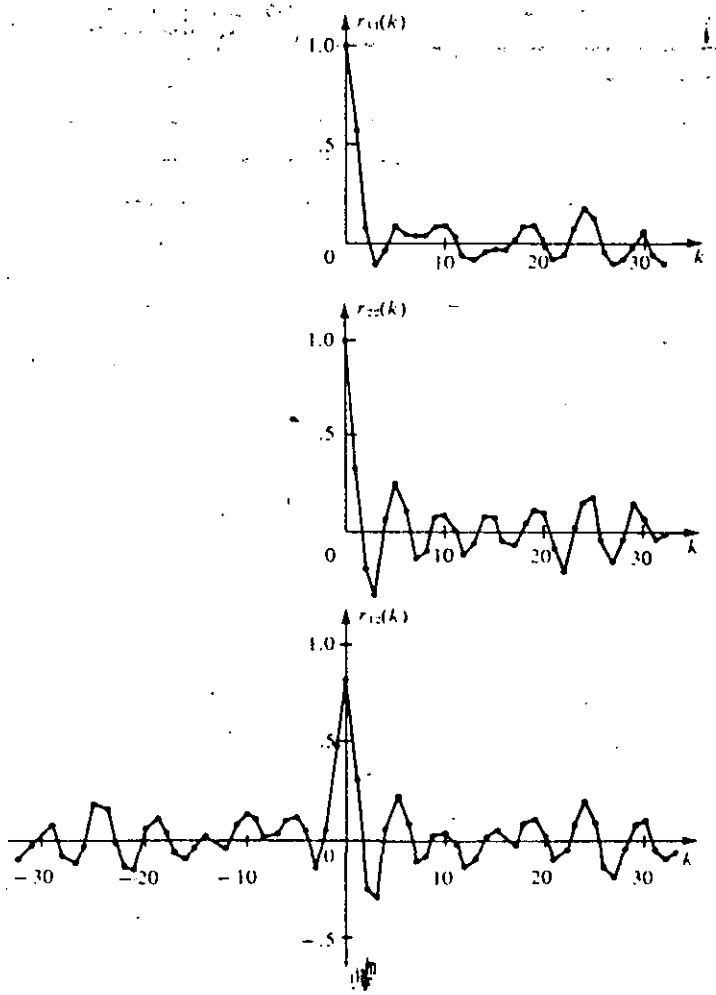


FIG. 10.2: Sample auto- and cross correlations for simulated linear system ($N = 100$)

(b) Confidence intervals for gain and phase based on (10.4.3) and (10.4.4) are shown in Figures 10.3 and 10.4. It is seen that these intervals diverge rapidly above 0.35 cps because of the drop in coherency. Hence the confidence intervals provide valuable guides in the interpretation of the gain and phase plots.

Calculation of the impulse response function. To make direct comparisons with the time-domain estimation procedures used in Section 10.2, the impulse-response function corresponding to the cross spectra based on $L = 16$ was evaluated from

$$h_m = \int_0^{1/2} \frac{\bar{L}_{12}(f)}{\bar{C}_{11}(f)} \cos 2\pi fm \, df + \int_0^{1/2} \frac{\bar{Q}_{12}(f)}{\bar{C}_{11}(f)} \sin 2\pi fm \, df. \quad (10.4.12)$$

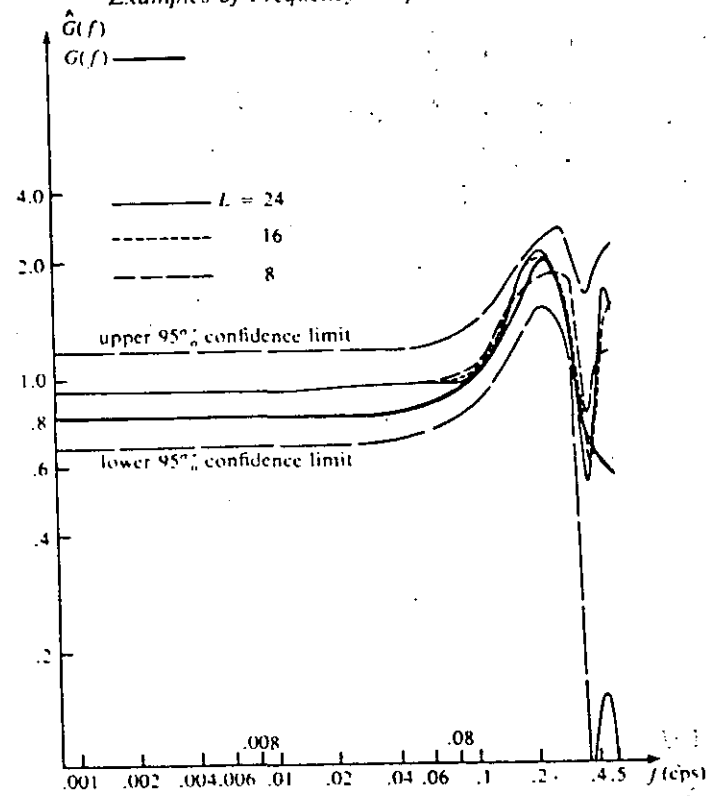


FIG. 10.3: Theoretical and estimated gain functions for a simulated linear system ($N = 100$)

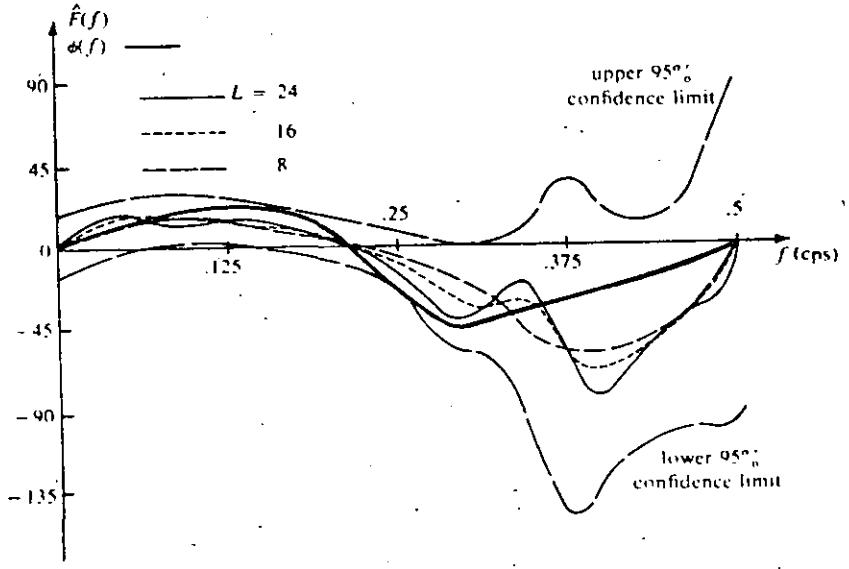


FIG. 10.4: Theoretical and estimated phase functions for a simulated linear system ($N = 100$)

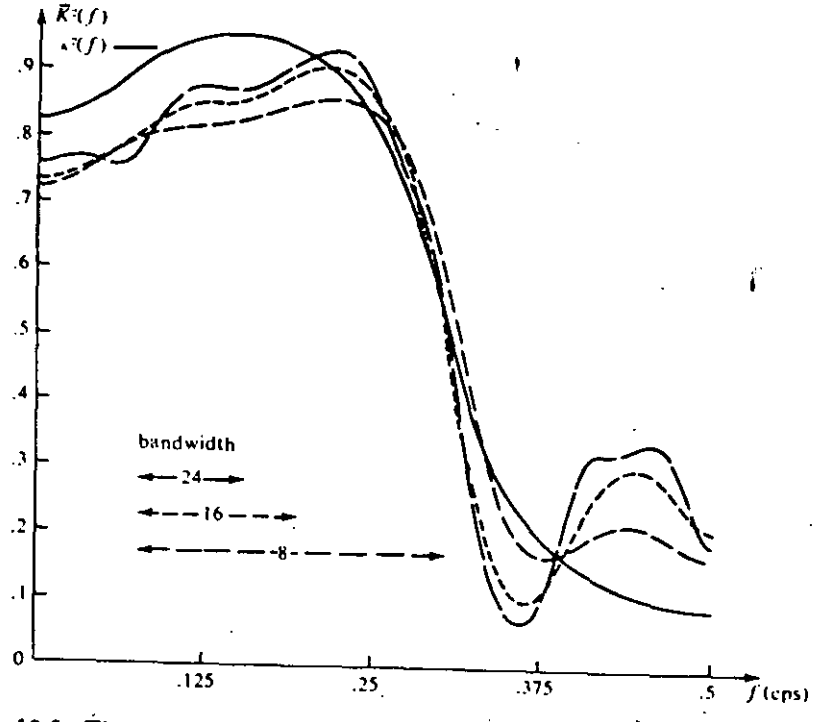


FIG. 10.5: Theoretical and estimated coherency spectra for a simulated linear system ($N = 100$)

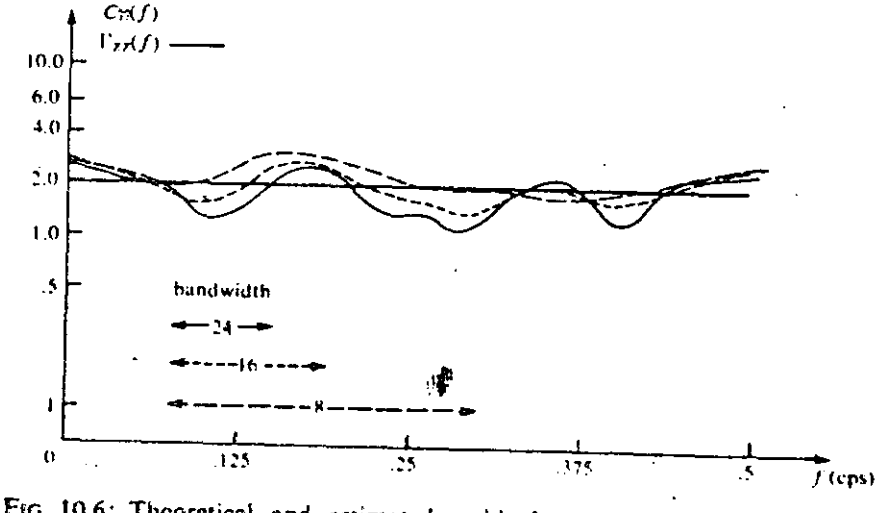


FIG. 10.6: Theoretical and estimated residual spectra for a simulated linear system ($N = 100$)

TABLE 10.4: Comparison of three methods of estimating impulse response functions

	h_m										
m	0	1	2	3	4	5	6	7	8	9	10
theoretical	1.000	0.250	−0.438	−0.234	0.154	0.055	−0.066	−0.044	0.022	0.028	−0.004
parametric est.	1.100	0.276	−0.458	−0.235	0.160	0.153	−0.038	−0.083	−0.003	0.039	0.011
spectral est.	1.108	0.271	−0.468	−0.184	0.139	0.120	−0.87	−0.40	−0.002	0.017	−0.060
direct ($M = 10$)	1.056	0.324	−0.572	0.037	−0.141	0.373	−0.249	0.107	−0.289	0.220	−0.163

This impulse response is compared with the direct and parametric impulse response estimates in Table 10.4. It is seen that it is smoother than the direct estimate and agrees reasonably well with the theoretical impulse response. However, it is not as good as the parametric estimate.

Following a suggestion in [4], values of h_m were also computed for m negative, to determine whether the physical realizability condition was satisfied. The largest value was $h_{-1} = 0.11$, all other values being less in magnitude than 0.1, and so it was concluded that the gain and phase estimates approximately describe a physically realizable system.

10.4.3 Analysis of gas furnace data

Description of the data. The data of Figure 8.3 were obtained from a gas furnace producing carbon dioxide. The output variable is the concentration of carbon dioxide measured as a percentage of the outlet gas from the furnace. The concentration is affected by two input variables, the air rate and the gas rate. In the experiment described here, the input air rate was fixed so that the frequency response function between the input gas rate and the output concentration could be determined.

Continuous measurements were available of both the input and output. Examination of the continuous records revealed that no discernible changes occurred at intervals of less than about 9 to 10 seconds and so the records were read at intervals of 9 seconds, yielding 296 pairs of data points, which are given in Appendix A10.1.

Estimation of gain and phase. A description of the estimation procedure will be given using the basic stages described in Sections 9.4.2 and 10.4.1.

(1) Preliminary decision stage

(a) Inspection of the data (Figure 8.3) did not reveal any obvious trends. However, the auto- and cross correlations of the original and differenced data were computed according to (9.3.13, 14).

(b) The Nyquist frequency corresponding to $\Delta = 9$ seconds is 1/18 cps.

(c) As a preliminary step, it was decided to compute the correlations up to a maximum lag of $L_{max} = 80$.

(2) First computation stage

The auto- and cross correlations were computed and examined. The cross correlations of the original data are shown in Figure 10.7, together with the differenced cross correlations. The approximate covariances for the differenced data are given in Appendix A10.1.

(3) Intermediate decision stage

(a) The presence of trends can be seen in the original ccf, so it was decided to use the differenced correlations.

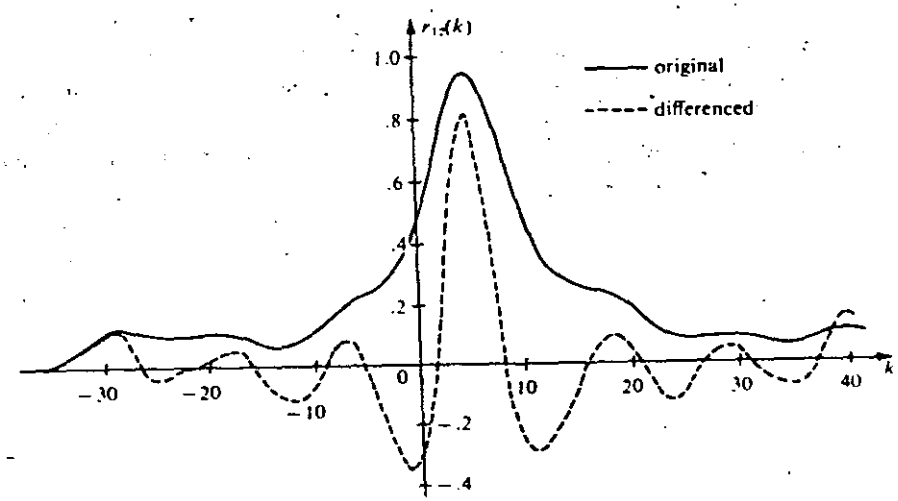


FIG. 10.7: Sample ccf's for original and differenced gas furnace data ($N = 296$)

(b) The ccf of the differenced data has a maximum at $S = 5$.

(c) On the basis of the damping of this ccf, truncation points of $L = 20, 30$ and 40 were chosen for evaluation of the spectra.

(4) Second computation stage

The autospectra, gain, phase and noise spectra were computed for the above truncation points using a shift $S = 5$. The gain function is plotted in Figure 10.8, the phase function in Figure 10.9 and the noise spectrum in Figure 10.10.

(5) Interpretation stage

The window-closing procedure suggests that very little change occurs in these spectra for values of L above 30. The values shown in all figures refer to $L = 30$ and $L = 40$, and it is seen that increasing the bandwidth has little effect on frequencies less than 0.02 cps but produces unstable oscillations above this frequency.

Interpretation of gain estimate. The Bode plot of Figure 10.8 shows that the system is second-order. The estimate of the time constants, obtained by fitting various second-order systems until a good visual fit was obtained, are $T_1 = 6.7$ seconds and $T_2 = 13$ seconds. Note that the 95% confidence intervals diverge rapidly above 0.025 cps because of the falling off in coherency. The gain estimate at zero frequency is 3.1, and hence the dc gain of the system is 0.31.

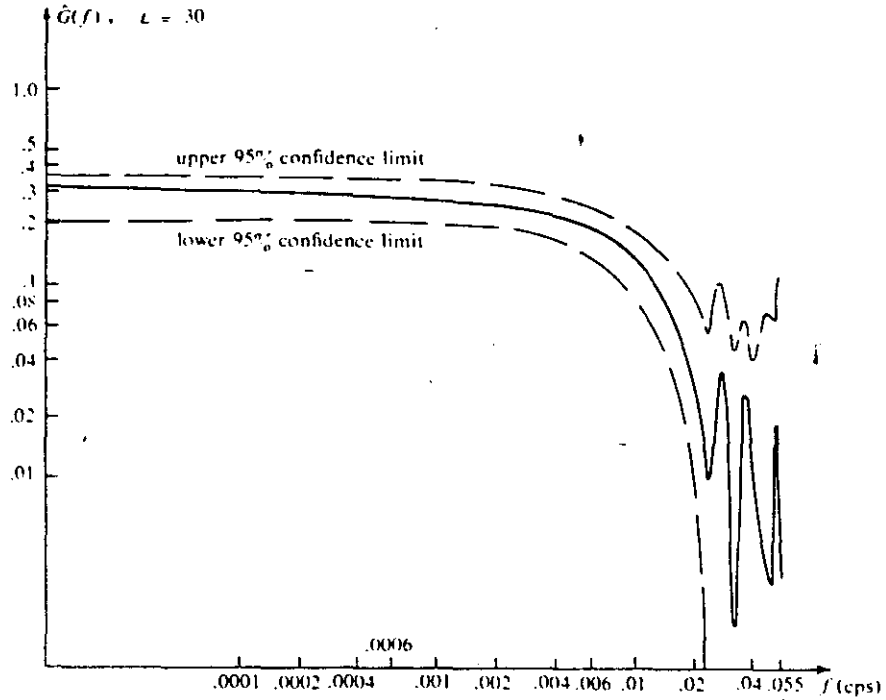


FIG. 10.8: Gain estimate for gas furnace data ($N = 296$)

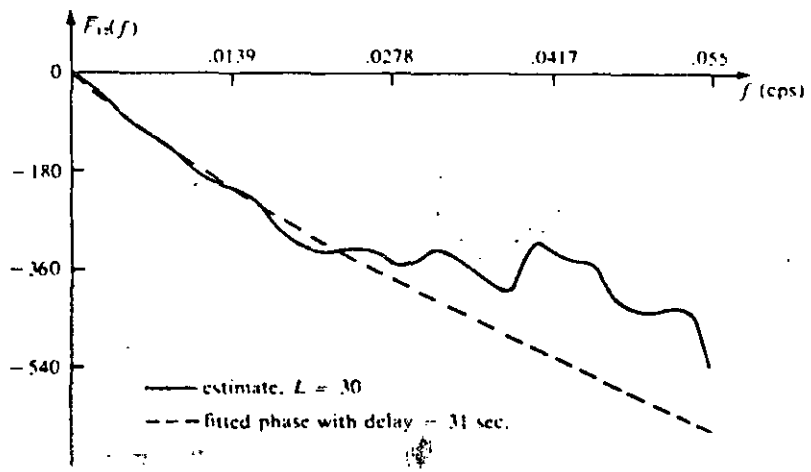


FIG. 10.9: Phase estimate for gas furnace data ($N = 296$)

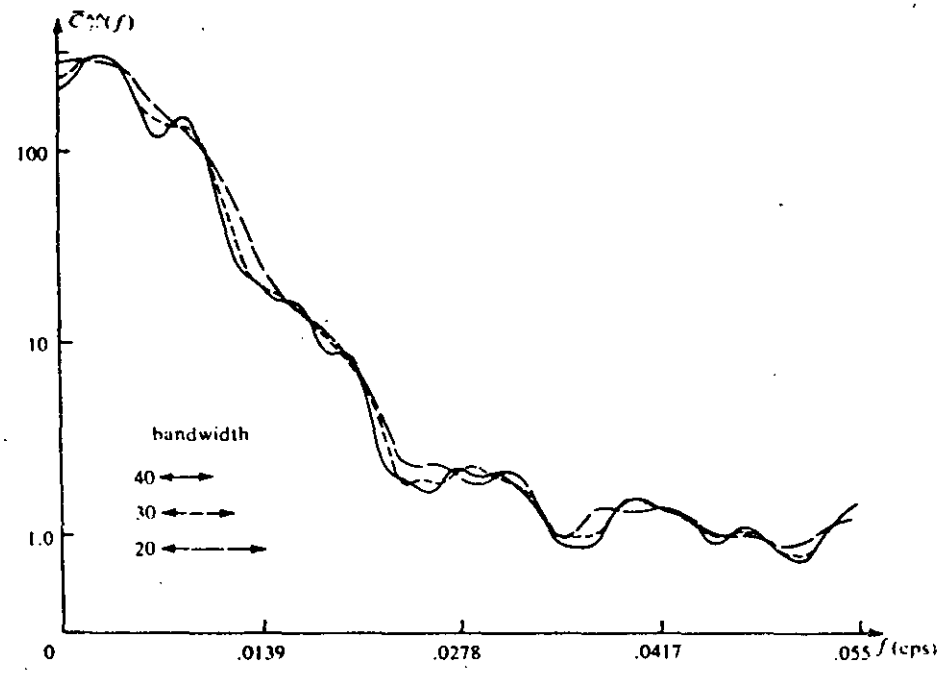


FIG. 10.10: Residual spectrum from gas furnace data ($N = 296$)

Interpretation of phase estimate. A discrete second-order system with $T_1 = 6.7$ seconds and $T_2 = 13$ seconds has the phase function

$$\phi_{12}(f) = \arctan \frac{0.13 \sin 2\pi f - 0.76 \sin 4\pi f}{1 - 0.76 \cos 2\pi f + 0.13 \cos 4\pi f}, \quad 0 \leq f \leq \tfrac{1}{2}. \quad (10.4.13)$$

The phase estimate for $L = 30$ in Figure 10.9 is much larger than this, suggesting that there is dead time or delay. The difference between the phase estimate and the phase (10.4.13) was plotted, and a straight line approximation to this curve was drawn. The approximating line intersected the line $f = 0.055$ at $\phi = -11$, and hence the delay $d = 11/[2\pi(0.055)] = 31$ seconds. The resulting phase function $F(f) = \hat{F}_{12}(f) - 62\pi f$ is shown in Figure 10.9 as a dashed line.

It can be concluded that the furnace behaves like a second-order system with dc gain of 0.31 units, time constants of 6.7 and 13 seconds and a delay of 31 seconds.

Interpretation of noise spectrum. Since the data have been filtered by differencing, it follows that the noise z_t' actually measured is related to the original

noise by $z_t' = z_t - z_{t-1}$. Hence, using (6.2.17), the noise spectrum estimate $\bar{C}_{zz}(f)$ can be recovered from the filtered noise estimate $\bar{C}_{z'z'}(f)$ by using

$$\bar{C}_{zz}(f) = \frac{\bar{C}_{z'z'}(f)}{4 \sin^2 \pi f \Delta}.$$

The spectrum $\bar{C}_{zz}(f)$ is shown in Figure 10.10, which indicates that approximately white noise has been passed through the dynamics of the system.

Further examples of frequency response estimation. An example of the application of cross spectral techniques to the estimation of the frequency response function of a heat exchanger is given in [2]. A wide variety of interesting applications is also given in the collection of papers [3].

As a general conclusion, spectral methods are often very useful in suggesting dynamic models for physical systems, as in the gas furnace example. However, since spectral methods involve estimating a parameter at each frequency, the efficiency of these methods is not high. More positive results may usually be obtained by parameterizing the problem, using models such as (10.1.3).

REFERENCES

[1] G. E. P. Box and G. M. Jenkins, *Statistical Models for Forecasting and Control*. Holden-Day, San Francisco (to be published).

[2] G. M. Jenkins, "An example of the estimation of a linear open-loop transfer function." *Technometrics* 5, 227 (1963).

[3] K. Matusita (ed.), "Studies of the statistical estimation of frequency response functions." Reprinted from *Ann. Inst. Stat. Math.*, Supplement III (1964).

[4] H. Akaike, "Some problems in the application of the cross-spectral method." Seminar on Spectral Analysis of Time Series, Madison, Wisconsin, 1966. Published by John Wiley, New York, 1967, as *Proceedings of an Advanced Seminar on Spectral Analysis of Time Series*, B. Harris, editor.

APPENDIX A10.1 DATA AND COVARIANCES FOR T FREQUENCY RESPONSE ESTIMATIONS

TABLE A10.1: Input and output to a simulated linear system, $N = 100$

1–10	−0.88	−0.12	−0.89	−1.38	−0.08	1.04	2.14	0.36	−1.11	−1.78
11–20	−2.76	−1.78	0.98	1.00	−0.71	−1.01	−1.31	−0.85	−0.47	1.64
21–30	0.06	−0.18	−1.02	−1.05	−0.66	−1.13	−0.52	−0.71	−0.21	−0.14
31–40	0.14	1.59	−0.77	−1.09	−1.77	−1.21	0.46	−0.08	−0.64	−0.36
41–50	−0.88	−0.62	0.29	1.91	2.15	1.05	0.32	1.07	2.67	2.45
51–60	1.31	1.11	1.94	0.34	1.83	1.16	0.62	−1.09	−1.62	−0.40
61–70	0.19	−1.60	−2.26	0.30	1.73	2.31	0.81	−0.40	0.30	−0.51
71–80	−2.12	−2.43	0.73	3.09	4.97	1.81	−0.46	−0.34	0.04	0.82
81–90	−1.64	−2.29	−0.77	1.92	1.93	0.85	−0.65	0.35	0.79	1.62
91–100	3.25	1.87	0.76	2.25	0.77	−0.16	0.19	0.61	0.92	−0.70

Input: $X_{1t} = X_{1t-1} - 0.5X_{1t-2} + Z_{1t}$

1–10	−1.03	−1.09	−0.47	−2.07	−1.52	2.27	3.14	0.26	−1.78	−1.82
11–20	−2.81	0.11	3.22	1.92	0.06	−3.48	−2.37	−0.34	0.43	1.97
21–30	2.01	−2.39	−0.94	−0.67	−1.25	0.05	−1.32	0.52	−1.15	0.40
31–40	0.80	2.09	−0.46	−2.66	−1.95	−0.84	−0.51	0.82	−1.62	−0.16
41–50	0.48	−0.36	−0.55	1.54	2.83	2.68	−0.57	0.54	4.70	1.90
51–60	1.23	−0.40	1.20	−2.02	1.97	−0.93	−0.12	−2.34	−4.42	0.10
61–70	1.06	0.06	−3.35	0.21	2.59	4.48	−0.92	−1.55	0.77	0.34
71–80	−3.27	−3.84	0.53	4.73	5.84	1.24	−3.62	0.59	1.04	3.56
81–90	−0.63	−2.09	−1.38	4.42	3.86	0.03	−2.11	−0.18	4.13	2.67
91–100	2.73	2.02	0.15	1.28	3.49	−0.73	−0.28	−0.84	1.37	−0.34

Output: $X_{2t} = X_t + Z_t$, $X_t = 0.25X_{t-1} - 0.5X_{t-2} + X_{1t}$

TABLE A10.2: Covariance estimates for the data of Table A10.1

k	$c_{11}(k)$	k	$c_{11}(k)$	k	$c_{11}(k)$	k	$c_{11}(k)$
0	1.948	10	0.219	20	0.070	30	0.119
1	1.153	11	0.082	21	−0.136	31	−0.051
2	0.180	12	−0.138	22	−0.136		
3	−0.257	13	−0.174	23	0.142		
4	−0.105	14	−0.095	24	0.391		
5	0.164	15	−0.067	25	0.297		
6	0.123	16	−0.063	26	−0.056		
7	0.049	17	0.000	27	−0.233		
8	0.081	18	0.168	28	−0.173		
9	0.206	19	0.210	29	0.034		

Input autocovariances

k	$c_{12}(k)$	k	$c_{12}(k)$	k	$c_{12}(k)$	k	$c_{12}(k)$
0	2.374	10	0.184	20	0.155	30	0.283
1	1.185	11	−0.007	21	−0.195	31	0.039
2	−0.645	12	−0.349	22	−0.397		
3	−0.972	13	−0.205	23	0.081		
4	−0.064	14	0.107	24	0.526		
5	0.615	15	0.195	25	0.501		
6	0.234	16	0.042	26	−0.064		
7	−0.283	17	−0.019	27	−0.447		
8	−0.244	18	0.126	28	−0.073		
9	0.150	19	0.284	29	0.304		

Input–output cross covariances

TABLE A10.2—*continued*

k	$c_{21}(k)$	k	$c_{21}(k)$	k	$c_{21}(k)$	k	$c_{21}(k)$
0	2.374	10	0.415	20	0.082	30	0.158
1	1.376	11	0.158	21	−0.471	31	−0.109
2	0.095	12	−0.152	22	−0.460		
3	−0.472	13	−0.095	23	0.080		
4	0.101	14	0.040	24	0.608		
5	0.371	15	−0.124	25	0.596		
6	0.331	16	−0.310	26	−0.040		
7	0.044	17	−0.193	27	−0.405		
8	0.030	18	0.155	28	−0.212		
9	0.389	19	0.322	29	0.250		

Output–input cross covariances

k	$c_{22}(k)$	k	$c_{22}(k)$	k	$c_{22}(k)$	k	$c_{22}(k)$
0	4.400	10	0.412	20	0.449	30	0.280
1	1.522	11	0.142	21	−0.564	31	−0.203
2	−0.991	12	−0.605	22	−1.021		
3	−1.521	13	−0.256	23	0.044		
4	0.302	14	0.333	24	0.724		
5	1.220	15	0.349	25	0.903		
6	0.573	16	−0.272	26	−0.338		
7	−0.530	17	−0.233	27	−0.769		
8	−0.499	18	0.073	28	−0.257		
9	0.366	19	0.526	29	0.688		

Output autocovariances

1-9	−1.09	0.00	1.78	3.39	3.73	4.41	4.61	3.48	1.27
10-18	−1.80	−5.88	−10.55	−14.21	−15.20	−13.02	−8.14	−4.75	−1.93
19-27	0.88	4.35	7.71	8.66	8.75	8.91	9.87	12.63	17.75
28-36	19.76	19.34	18.66	18.32	17.67	16.08	12.65	7.90	3.60
37-45	1.15	0.88	3.31	6.45	9.60	14.09	26.70	28.34	28.12
46-54	24.83	19.29	14.85	12.14	12.39	16.08	19.05	20.23	18.15
55-63	5.35	1.22	0.09	1.64	6.71	10.19	11.46	11.55	11.12
64-72	11.21	12.23	12.57	11.57	9.13	6.20	2.55	−2.80	−10.80
73-81	−15.51	−17.99	−18.25	−14.56	−9.44	−5.70	−4.31	−5.77	−9.60
82-90	−16.16	−18.75	−18.91	−17.46	−14.74	−12.01	−9.27	−5.24	0.40
91-99	7.88	9.43	9.30	10.06	11.37	11.98	10.54	5.95	−0.80
100-108	−3.14	−2.88	−1.53	−1.09	−1.87	−2.55	−2.29	−0.07	2.54
109-117	3.30	1.02	−4.23	−11.39	−22.75	−25.94	−27.16	−25.10	−17.90
118-126	−13.46	−10.81	−9.10	−8.76	−8.85	−8.00	−5.44	−4.16	−2.71
127-135	0.00	4.03	8.41	12.85	16.07	17.46	16.83	14.85	9.93
136-144	6.48	5.77	5.77	6.32	7.47	9.00	9.93	9.68	7.90
145-153	3.99	−1.61	−5.53	−6.03	−4.24	−1.94	−0.49	0.60	1.61
154-162	3.01	5.17	5.66	5.60	5.73	5.92	6.71	9.33	13.37
163-171	14.60	13.53	7.72	2.18	−2.37	−7.14	−10.99	−12.69	−11.75
172-180	−6.76	0.33	5.56	6.43	4.84	1.09	−3.10	−6.97	−10.47
181-189	−12.18	−11.83	−8.73	−3.36	0.63	0.84	0.00	0.01	2.09
190-198	5.56	7.82	8.58	9.18	8.62	4.16	−3.36	−9.59	−18.13
199-207	−23.78	−24.99	−24.73	−23.30	−20.53	−17.39	−12.61	−5.69	−1.37
208-216	−0.24	−0.50	−1.35	−2.76	−5.34	−8.71	−12.43	−14.39	−14.22
217-225	−11.75	−8.13	−6.34	−5.82	−6.25	−7.13	−8.48	−10.39	−13.46
226-234	−16.28	−16.19	−11.49	−4.88	−1.60	−0.07	−0.92	−6.20	−10.86
235-243	−15.25	−18.58	−20.29	−20.24	−19.61	−19.52	−17.94	−13.02	−10.30
244-252	−9.18	−7.98	−8.67	−10.47	−11.23	−8.76	−3.95	1.85	6.62
253-261	7.09	6.05	5.01	6.03	9.43	12.23	12.49	8.24	1.02
262-270	0.25	3.82	9.22	10.32	8.66	5.27	0.93	−4.58	−7.48
271-279	−9.47	−10.29	−9.28	−6.45	−4.24	−2.76	−1.58	−0.33	1.02
280-288	2.51	2.80	0.00	−4.93	−7.59	−8.24	−7.40	−5.28	−2.04
289-296	0.34	2.04	2.53	1.95	1.31	0.17	−1.82	−2.62	

[*actual gas rate* (*cu. ft. per min.*) = 0.600 − 0.004 × (*coded value*)]

TABLE A10.4: Output CO_2 concentration (%) from gas furnace

1-9	53.8	53.6	53.5	53.5	53.4	53.1	52.7	52.4	52.2
10-18	52.0	52.0	52.4	53.0	54.0	54.9	56.0	56.8	56.8
19-27	56.4	55.7	55.0	54.3	53.2	52.3	51.6	51.2	50.8
28-36	50.5	50.0	49.2	48.4	47.9	47.6	47.5	47.5	47.6
37-45	48.1	49.0	50.0	51.1	51.8	51.9	51.7	51.2	50.0
46-54	48.3	47.0	45.8	45.6	46.0	46.9	47.8	48.2	48.3
55-63	47.9	47.2	47.2	48.1	49.4	50.6	51.5	51.6	51.2
64-72	50.5	50.1	49.8	49.6	49.4	49.3	49.2	49.3	49.7
73-81	50.3	51.3	52.8	54.4	56.0	56.9	57.5	57.3	56.6
82-90	56.0	55.4	55.4	56.4	57.2	58.0	58.4	58.4	58.1
91-99	57.7	57.0	56.0	54.7	53.2	52.1	51.6	51.0	50.5
100-108	50.4	51.0	51.8	52.4	53.0	53.4	53.6	53.7	53.8
109-117	53.8	53.8	53.3	53.0	52.9	53.4	54.6	56.4	58.0
118-126	59.4	60.2	60.0	59.4	58.4	57.6	56.9	56.4	56.0
127-135	55.7	55.3	55.0	54.4	53.7	52.8	51.6	50.6	49.4
136-144	48.8	48.5	48.7	49.2	49.8	50.4	50.7	50.9	50.7
145-153	50.5	50.4	50.2	50.4	51.2	52.3	53.2	53.9	54.1
154-162	54.0	53.6	53.2	53.0	52.8	52.3	51.9	51.6	51.6
163-171	51.4	51.2	50.7	50.0	49.4	49.3	49.7	50.6	51.8
172-180	53.0	54.0	55.3	55.9	55.9	54.6	53.5	52.4	52.1
181-189	52.3	53.0	53.8	54.6	55.4	55.9	55.9	55.2	54.4
190-198	53.7	53.6	53.6	53.2	52.5	52.0	51.4	51.0	50.9
199-207	52.4	53.5	55.6	58.0	59.5	60.0	60.4	60.5	60.2
208-216	59.7	59.0	57.6	56.4	55.2	54.5	54.1	54.1	54.4
217-225	55.5	56.2	57.0	57.3	57.4	57.0	56.4	55.9	55.5
226-234	55.3	55.2	55.4	56.0	56.5	57.1	57.3	56.8	55.6
235-243	55.0	54.1	54.3	55.3	56.4	57.2	57.8	58.3	58.6
244-252	58.8	58.8	58.6	58.0	57.4	57.0	56.4	56.3	56.4
253-261	56.4	56.0	55.2	54.0	53.0	52.0	51.6	51.6	51.1
262-270	50.4	50.0	50.0	52.0	54.0	55.1	54.5	52.8	51.4
271-279	50.8	51.2	52.0	52.8	53.8	54.5	54.9	54.9	54.8
280-288	54.4	53.7	53.3	52.8	52.6	52.6	53.0	54.3	56.0
					58.3	57.8	57.3	57.0	

TABLE A10.5: Covariance estimates for differenced gas furnace data

k	$c_{11}(k)$	k	$c_{11}(k)$	k	$c_{11}(k)$	k	$c_{11}(k)$
0	10.9	10	0.8	20	−0.6	30	−0.8
1	8.1	11	0.9	21	0.1	31	0.0
2	3.9	12	1.0	22	0.7	32	−0.9
3	−0.2	13	0.9	23	0.9	33	1.7
4	−3.1	14	0.3	24	0.6	34	1.8
5	−3.9	15	−0.4	25	0.3	35	1.4
6	−3.6	16	−0.6	26	−0.1	36	0.6
7	−2.9	17	−1.0	27	−0.5	37	0.0
8	−2.1	18	−1.4	28	−1.1	38	−0.5
9	−1.1	19	−1.3	29	−1.3	39	−1.0

Input autocovariances

k	$c_{12}(k)$	k	$c_{12}(k)$	k	$c_{12}(k)$	k	$c_{12}(k)$
0	0.79	10	0.69	20	−0.12	30	−0.13
1	0.43	11	0.74	21	0.01	31	0.00
2	−0.30	12	0.68	22	0.17	32	0.10
3	−1.23	13	0.55	23	0.30	33	0.15
4	−1.92	14	0.42	24	0.32	34	0.20
5	−2.10	15	0.22	25	0.25	35	0.20
6	−1.70	16	0.00	26	0.10	36	0.17
7	−0.94	17	−0.15	27	−0.06	37	−0.01
8	−0.14	18	−0.23	28	−0.10	38	−0.20
9	0.42	19	−0.23	29	−0.14	39	−0.36

Input–output cross covariances

TABLE A.10.5—*continued*

k	$c_{21}(k)$	k	$c_{21}(k)$	k	$c_{21}(k)$	k	$c_{21}(k)$
0	0.79	10	0.17	20	−0.05	30	−0.27
1	0.89	11	0.26	21	−0.01	31	−0.15
2	0.79	12	0.30	22	0.01	32	0.02
3	0.55	13	0.29	23	0.04	33	0.09
4	0.29	14	0.24	24	0.09	34	0.12
5	0.05	15	0.13	25	0.12	35	0.15
6	−0.15	16	−0.02	26	0.06	36	0.11
7	−0.20	17	−0.12	27	−0.10	37	0.02
8	−0.15	18	−0.13	28	−0.28	38	−0.13
9	0.02	19	−0.08	29	−0.33	39	−0.20

Output–input cross covariances

k	$c_{22}(k)$	k	$c_{22}(k)$	k	$c_{22}(k)$	k	$c_{22}(k)$
0	0.598	10	−0.054	20	−0.060	30	−0.056
1	0.465	11	0.002	21	−0.027	31	−0.054
2	0.294	12	0.034	22	0.004	32	−0.025
3	0.093	13	0.038	23	0.025	33	0.037
4	−0.073	14	0.028	24	0.032	34	0.064
5	−0.168	15	−0.004	25	0.036	35	0.075
6	−0.201	16	−0.030	26	0.033	36	0.061
7	−0.193	17	−0.065	27	0.024	37	0.026
8	−0.156	18	−0.083	28	0.004	38	−0.002
9	−0.116	19	−0.083	29	−0.035	39[a]	−0.020

Output autocovariances

APPENDIX A10.2 FLOW CHART FOR FREQUENCY RESPONSE CALCULATIONS

The following is a flow chart for a computer program FRQRSP which accepts the same inputs as a program CROSSPEC (Appendix A9.2). Printer output from FRQRSP consists of the covariances (echo check), the smoothed autospectra for each truncation point M, and the gain, phase, squared coherency and residual spectra, together with approximate upper and lower 95% confidence intervals, for the gain and phase functions. Plotter output consists of the input, output and residual logspectra, the log gain versus log frequency plot, with upper and lower confidence intervals, plus the phase versus frequency plot, all with overlays for each additional truncation point used.

Program FRQRSP

1) through 11) as in program CROSSPEC, so that at this point in the program the following quantities are available, for K = 0, NF

SPEC(K,1), SPEC(K,2), SQ(K)

Phase(K), COHSQ(K).

12) *Calculate* GAIN(K) = SQRT(SQ(K))/SPEC(K,1)

RESID(K) = SPEC(K,2)(1. − COHSQ(K)).

13) *Calculate* approximate $f_{2,D-2}$ probability point as

D = 8*N/(3*M)

E = D − 2

A = 2.*(2.93 + 11.7/E)/E.

14) *Calculate* upper and lower 95% gain and phase confidence intervals as

G(K) = SQRT(A*RESID(K)*SPEC(K,1))

P(K) = ARCSIN(G(K)/GAIN(K))

GU(K) = GAIN(K) + G(K)

GL(K) = GAIN(K) − G(K)

PU(K) = PHASE(K) + P(K)

PL(K) = PHASE(K) − P(K).

15) *Calculate* logarithms to give

LOGSPEC(K,1) = LOG10(SPEC(K,1))

LOGSPEC(K,2) = LOG10(SPEC(K,2))

LOGRESID(K) = LOG10(RESID(K))

LOGGAIN(K) = LOG10(GAIN(K))

LOGGU(K) = LOG10(GU(K))

LOGGL(K) = LOG10(GL(K)).

16) *Print* autospectra (input, output, residual), squared coherency, phase, gain, plus upper and lower 95% confidence limits for phase and gain.

17) *Plot* and *overlay* logspectra (input, output, residual) versus frequency, phase plus upper and lower 95% confidence limits versus frequency, log gain plus upper and lower 95% confidence limits versus log frequency.

NOTE: Because the phase can change abruptly from +90° to −90°, phase plots can become rather difficult to decipher when overlays are made, and when a logfrequency scale is used. For these reasons, it is suggested that each phase spectrum be plotted on a single graph, and as a function of frequency, not logfrequency.

11

Multivariate Spectral Analysis

In this chapter, the methods of Chapters 8, 9 and 10 are generalized to deal with any number of time series. In particular, it is shown how to describe in the frequency domain q time series arising on an equal footing and how to estimate the multivariate frequency response function of a system having q inputs and r outputs. Up to now matrix theory has been avoided in order to minimize the amount of mathematical technique required to understand the basic ideas in spectral analysis. However, further progress is impossible without the introduction of matrix theory.

In Section 11.1 some of the concepts used in the analysis of univariate and bivariate time series are recast in terms of matrix theory. In particular, the covariance matrix of a time series is defined, and it is shown that the spectrum is intimately connected with the latent roots of this covariance matrix. In Section 11.2 a multivariate linear system is introduced. A linear multivariate process is defined as the output from such a system when the inputs are a set of uncorrelated white noise processes. Important special cases of a multivariate linear process are the bivariate ar and ma processes.

In Section 11.4 the basic ideas in multivariate spectral analysis and frequency response function estimation are developed. To introduce these concepts, the important ideas in multiple regression and multivariate analysis are reviewed in Section 11.3. Finally, in Section 11.5, the more practical aspects of multivariate frequency response estimation are discussed and an example is given based on two inputs and two outputs from a turbo-alternator.

It is conventional to use lower-case letters for vectors and capital letters for matrices. Since this notation has been used to distinguish between time-domain and frequency-domain quantities, it cannot always be used here to distinguish between vectors and matrices. In this book matrices and vectors are denoted by bold-face type. Where possible, upper-case bold letters will refer to matrices, but occasionally they may refer to vectors, the exact meaning being made clear in each case.

11.1 PROPERTIES OF THE COVARIANCE MATRIX

11.1.1 The covariance matrix of a real stochastic process

It was shown in Section 3.1.5 that the second-moment properties of a set of rv's can be summarized by their covariance matrix (3.1.20). A stochastic process can be characterized by an infinite set of rv's, and the second-moment properties of the process can be described by the covariance matrix of the rv's at any subset of times $t_1, t_2, \ldots, t_N$. For a discrete process, these times will be equidistant and so, if the process is stationary,

$$\text{Cov}\,[X(t_i), X(t_j)] = \sigma^2 \rho(i-j), \tag{11.1.1}$$

where $\rho(k)$ is the acf at lag k. The covariance matrix associated with these N times is the array of numbers such that the element in the ith row and the jth column is Cov $[X(t_i), X(t_j)]$. Thus, using (11.1.1), the covariance matrix is

$$\mathbf{V}_N = \sigma^2 \begin{pmatrix} 1 & \rho(1) & \rho(2) & \cdots & \rho(N-1) \\ \rho(1) & 1 & \rho(1) & \cdots & \rho(N-2) \\ \rho(2) & \rho(1) & 1 & \cdots & \rho(N-3) \\ \vdots & \vdots & \vdots & & \vdots \\ \rho(N-1) & \rho(N-2) & \rho(N-3) & \cdots & 1 \end{pmatrix}. \tag{11.1.2}$$

A matrix of the form (11.1.2), which possesses the property that the elements on symmetric diagonals are identical, is called a *Toeplitz* matrix.

Positive semi-definite property. As indicated in Section 3.1.5, the covariance matrix of a set of rv's is positive semi-definite. For a stationary stochastic process, it follows that the matrices $\mathbf{V}_N$ for $N = 2, 3, \ldots, \infty$ are positive semi-definite, that is, all principal minors of the determinant of $\mathbf{V}_N$ are positive or zero. This implies that the autocorrelations of a stationary time series satisfy a wide range of conditions. For example, if $N = 2$, $|\mathbf{V}_2| \geqslant 0$ implies that

$$|\rho(1)| \leqslant 1.$$

Similarly, by considering the rv's $X(t)$, $X(t+s)$,

$$|\rho(s)| \leqslant 1$$

for any s.

A more interesting example occurs when $N = 3$, in which case

$$|\mathbf{V}_3| = \begin{vmatrix} 1 & \rho(1) & \rho(2) \\ \rho(1) & 1 & \rho(1) \\ \rho(2) & \rho(1) & 1 \end{vmatrix} \geqslant 0. \tag{11.1.3}$$

It may be verified that (11.1.3) implies

$$|\rho(2)| \leqslant 1,$$
$$\left|\frac{\rho(2) - \rho^2(1)}{1 - \rho^2(1)}\right| \leqslant 1. \tag{11.1.4}$$

The first of the conditions (11.1.4) does not introduce anything new, but the second defines a constraint which must be satisfied by $\rho(1)$ and $\rho(2)$. To illustrate this constraint, consider the first-order ar process for which $\rho(2) = \rho^2(1)$, $\rho(3) = \rho^3(1)$, and so on. The function

$$\pi(2) = \frac{\rho(2) - \rho^2(1)}{1 - \rho^2(1)} \tag{11.1.5}$$

is zero if the process is first-order ar. Hence, when $\pi(2)$ is non-zero it measures the *excess correlation* over and above that expected if the process were first-order ar. Thus $\pi(2)$ can be used to check whether an empirical time series can be adequately fitted by a first-order process, as described in Section 5.4.3.

More generally, by considering $|\mathbf{V}_k|$, it may be shown that

$$\pi(k-1) = \frac{\|\mathbf{V}_k\|}{|\mathbf{V}_{k-1}|} \tag{11.1.6}$$

lies between -1 and $+1$, where $\|\mathbf{V}_k\|$ is the determinant of the cofactor of the element in the first column and kth row of $\mathbf{V}_k$. The plot of $\pi(k)$ versus k is called the *partial autocorrelation function*. It has the property that if the process is ar of order m,

$$\pi(k) \neq 0, \quad k \leqslant m$$
$$\pi(k) = 0, \quad k > m,$$

and hence it can be used to check whether an empirical time series can be fitted by an ar process of a given order.

11.1.2 Latent roots and the spectrum

In this section it is shown that the latent roots of the covariance matrix $\mathbf{V}_N$ are approximately equal to the power spectrum ordinates at the frequencies i/N. The reader is referred to Appendix A11.1 for some elementary properties of latent roots and vectors which are necessary for this section.

Transformation of correlated random variables to obtain independence. Suppose that the rv's $\mathbf{X}' = (X_1, X_2, \ldots, X_N)$ have covariance matrix $\mathbf{V}$. Now consider linear functions $\mathbf{l}_i'\mathbf{X}$, $\mathbf{l}_j'\mathbf{X}$ of these rv's where $\mathbf{l}_i$, $\mathbf{l}_j$ are left-hand latent vectors of $\mathbf{V}$. Then using (A11.1.7),

$$\text{Cov}\,[\mathbf{l}_i'\mathbf{X}, \mathbf{l}_j'\mathbf{X}] = \mathbf{l}_i'\mathbf{V}\mathbf{l}_j = \begin{cases}\lambda_i, & i = j\\ 0, & i \neq j\end{cases} \tag{11.1.7}$$

Hence the transformations

$$\mathbf{Y}_i = \mathbf{l}_i'\mathbf{X}, \quad i = 1, 2, \ldots, N, \tag{11.1.8}$$

convert the correlated rv's X_i into uncorrelated rv's Y_i. Further, the variance of the rv Y_i is equal to λ_i, the latent root associated with the latent vector $\mathbf{l}_i$. For example, if $N = 2$, the latent roots are obtained from

$$\begin{vmatrix} 1-\lambda & \rho \\ \rho & 1-\lambda \end{vmatrix} = 0, \tag{11.1.9}$$

so that $\lambda_1 = 1 + \rho$, $\lambda_2 = 1 - \rho$. Similarly the latent vectors are

$$\mathbf{l}_1' = \left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right),$$
$$\mathbf{l}_2' = \left(\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}\right). \tag{11.1.10}$$

Hence

$$Y_1 = \frac{1}{\sqrt{2}}(X_1 + X_2),$$
$$Y_2 = \frac{1}{\sqrt{2}}(X_1 - X_2) \tag{11.1.11}$$

are uncorrelated and $\text{Var}\,[Y_1] = 1 + \rho$, $\text{Var}\,[Y_2] = 1 - \rho$, as may be verified directly. The inverse transformation to (11.1.8), namely,

$$\mathbf{X}_i = \mathbf{l}_i\mathbf{Y} = l_{1i}Y_1 + l_{2i}Y_2 + \cdots + l_{Ni}Y_N \tag{11.1.12}$$

shows that the rv X_i may be written as a linear function of the uncorrelated rv's Y_i. Finally, using (3.2.18), the variance of X_i may be decomposed according to

$$\text{Var}\,[X_i] = \sum_{j=1}^{N} l_{ji}^2\lambda_j. \tag{11.1.13}$$

Latent roots of circular stochastic processes. The latent vectors of the covariance matrix (11.1.2) of a stochastic process are, in general, complicated functions of the autocorrelations. However, there is considerable simplification if it is assumed that the process is *periodic* or *circular* with period N. This means that

$$X(t) = X(t + N),$$

and hence the autocorrelation satisfies the further condition

$$\rho(k) = \rho(N - k). \tag{11.1.14}$$

As N tends to infinity, the circular process tends to the actual process. Hence for finite N, the properties of the circular process provide approximations to the properties of the actual process.

The corresponding lagged covariance matrix may be evaluated, using the definition (11.2.1), and is

$$\mathbf{V}_X(k) = \begin{cases} \boldsymbol{\beta}_1\mathbf{V}_Z\boldsymbol{\beta}'_{1+k} + \boldsymbol{\beta}_2\mathbf{V}_Z\boldsymbol{\beta}'_{2+k} + \cdots + \boldsymbol{\beta}_{l-k}\mathbf{V}_Z\boldsymbol{\beta}'_l, & k \leqslant l \\ 0, & k > l. \end{cases} \tag{11.2.25}$$

The process (11.2.24) has the frequency response function

$$\mathbf{H}(f) = \mathbf{I} + \boldsymbol{\beta}_1 e^{-j2\pi f} + \cdots + \boldsymbol{\beta}_l e^{-j2\pi lf},$$

and hence the spectral matrix is

$$\boldsymbol{\Gamma}_X(f) = (\mathbf{I} + \boldsymbol{\beta}_1 e^{j2\pi f} + \cdots + \boldsymbol{\beta}_l e^{j2\pi lf})\mathbf{W}(\mathbf{I} + \boldsymbol{\beta}'_1 e^{-j2\pi f} + \cdots + \boldsymbol{\beta}'_l e^{-j2\pi lf}). \tag{11.2.26}$$

Autoregressive processes. The general discrete ar process may be written

$$\mathbf{X}_t - \boldsymbol{\mu} = \boldsymbol{\alpha}_1(\mathbf{X}_{t-1} - \boldsymbol{\mu}) + \cdots + \boldsymbol{\alpha}_m(\mathbf{X}_{t-m} - \boldsymbol{\mu}) + \mathbf{Z}_t. \tag{11.2.27}$$

and the lagged covariance matrix satisfies the matrix difference equation

$$\mathbf{V}_X(k) = \boldsymbol{\alpha}_1\mathbf{V}_X(k-1) + \cdots + \boldsymbol{\alpha}_m\mathbf{V}_X(k-m), \quad k > 0.$$

In the special case $m = 1$, this has the solution

$$\mathbf{V}_X(k) = \boldsymbol{\alpha}_1^k\mathbf{V}_X(0),$$

so that the lagged covariance matrix is easily obtained by taking powers of the $\boldsymbol{\alpha}_1$ matrix. However, it is necessary to evaluate $\mathbf{V}_X(0)$ by direct methods, as illustrated in the example of Section 8.1.5. Regarding (11.2.27) as a linear system with frequency response matrix

$$\mathbf{H}(f) = [\mathbf{I} - \boldsymbol{\alpha}_1 e^{-j2\pi f} - \cdots - \boldsymbol{\alpha}_m e^{-j2\pi mf}]^{-1},$$

then using (11.2.21), the spectral matrix is

$$\boldsymbol{\Gamma}_X(f) = [\mathbf{I} - \boldsymbol{\alpha}_1 e^{j2\pi f} - \cdots - \boldsymbol{\alpha}_m e^{j2\pi mf}]^{-1} \times \mathbf{W} \times [\mathbf{I} - \boldsymbol{\alpha}'_1 e^{-j2\pi f} - \cdots - \boldsymbol{\alpha}'_m e^{-j2\pi mf}]^{-1}. \tag{11.2.28}$$

Similarly, the lagged covariance matrix of the continuous ar process

$$\mathbf{a}_m\frac{d^m\mathbf{X}(t)}{dt^m} + \cdots + \mathbf{a}_0\mathbf{X}(t) = \mathbf{Z}(t) \tag{11.2.29}$$

satisfies the differential equation

$$\mathbf{a}_m\frac{d^m\mathbf{V}(u)}{du^m} + \cdots + \mathbf{a}_0\mathbf{V}(u) = 0.$$

For example, suppose $m = 1$ and $\mathbf{a}_m = \mathbf{I}$. Then

$$\mathbf{V}(u) = e^{-\mathbf{a}_0 u}\mathbf{V}(0).$$

The spectral matrix corresponding to (11.2.29) may be obtained by regarding it as a linear system with frequency response matrix

$$\mathbf{H}(f) = [\mathbf{a}_m(j2\pi f)^m + \cdots + \mathbf{a}_0\mathbf{I}]^{-1}.$$

Mixed processes. More generally still, it is possible to define a multivariate mixed ar–ma process

$$\mathbf{X}_t - \boldsymbol{\mu} = \boldsymbol{\alpha}_1(\mathbf{X}_{t-1} - \boldsymbol{\mu}) + \cdots + \boldsymbol{\alpha}_m(\mathbf{X}_{t-m} - \boldsymbol{\mu}) + \mathbf{Z}_t + \boldsymbol{\beta}_1\mathbf{Z}_{t-1} + \cdots + \boldsymbol{\beta}_l\mathbf{Z}_{t-l}. \tag{11.2.30}$$

The model (11.2.30) gives rise to a wide class of lagged covariance matrices and hence provides a very powerful model for the fitting of multivariate time series.

11.3 TIME-DOMAIN MULTIVARIATE ANALYSIS

It was shown in Chapters 9 and 10 that cross spectral analysis and frequency response estimation represent extensions to the frequency domain of ordinary correlation and regression analysis. Similarly, multivariate spectral analysis and multivariate frequency response estimation represent extensions of the ideas of multiple correlation analysis and multivariate analysis to the frequency domain. In this section, the basic ideas in multiple correlation and regression analysis and multivariate analysis are reviewed. Complete understanding of the least squares theory of Appendix A4.1 is now assumed.

11.3.1 Multiple regression analysis, single output

The model. Consider a special case of the multivariate dynamic model (11.2.10) in which there is only one output, and assume that the system responds so quickly that it can be effectively described by its steady-state behavior. Then with an appropriate change of notation, (11.2.10) may be written

$$X_{t1q+1} - \mu_{q+1} = h_1(X_{t1} - \bar{X}_1) + h_2(X_{t2} - \bar{X}_2) + \cdots + h_q(X_{tq} - \bar{X}_q) + Z_t, \tag{11.3.1}$$

where Z_t is a noise term. As noted in Section 4.3.4, it is useful to subtract the means from the regression variables in a least squares model to make the parameters orthogonal to the constant term μ_{q+1}.

Normal equations. Assuming that the noise Z_t is white and that observations are made at times $t = 1, 2, \ldots, N$, the normal equations (A4.1.7) are

$$(\mathbf{X}'\mathbf{X})\hat{\mathbf{h}} = \mathbf{X}'\mathbf{x},$$
$$\hat{\mu}_{q+1} = \bar{x}_{q+1}, \tag{11.3.2}$$

where

$$\mathbf{X} = \begin{pmatrix} x_{11} - \bar{x}_1 & x_{12} - \bar{x}_2 & \cdots & x_{1q} - \bar{x}_q \\ x_{21} - \bar{x}_1 & x_{22} - \bar{x}_2 & \cdots & x_{2q} - \bar{x}_q \\ \vdots & \vdots & & \vdots \\ x_{N1} - \bar{x}_1 & x_{N2} - \bar{x}_2 & \cdots & x_{Nq} - \bar{x}_q \end{pmatrix},$$

$$\mathbf{h}' = (h_1, h_2, \ldots, h_q), \tag{11.3.3}$$

$$\mathbf{x}' = (x_{1q+1} - \bar{x}_{q+1}, x_{2q+1} - \bar{x}_{q+1}, \ldots, x_{Nq+1} - \bar{x}_{q+1}).$$

Equations (11.3.2) can be written in terms of covariance estimates

$$c_{ij} = \frac{1}{N}\sum_{t=1}^{N}(x_{ti} - \bar{x}_i)(x_{tj} - \bar{x}_j), \tag{11.3.4}$$

as

$$\begin{aligned} \mathbf{C}_{qq}\hat{\mathbf{h}} &= \mathbf{c}_{q+1}, \\ \hat{\mu}_{q+1} &= \bar{x}_{q+1}, \end{aligned} \tag{11.3.5}$$

where $\mathbf{C}_{qq}$ is the covariance matrix of the inputs, and $\mathbf{c}_{q+1}$ is the vector of cross covariances between the inputs and the output.

An example. For $q = 2$, the model (11.3.1) is

$$X_{t3} - \mu_3 = h_1(X_{t1} - \bar{X}_1) + h_2(X_{t2} - \bar{X}_2) + Z_t, \quad t = 1, 2, \ldots, N,$$

and the estimation equations (11.3.5) are

$$\begin{aligned} \hat{h}_1 c_{11} + \hat{h}_2 c_{12} &= c_{13}, \\ \hat{h}_1 c_{21} + \hat{h}_2 c_{22} &= c_{23}, \\ \hat{\mu}_3 &= \bar{x}_3. \end{aligned}$$

11.3.2 Multiple correlation

The expression (A4.1.11) for the residual sum of squares from the fitted regression is

$$\sum_t z_t^2 = \mathbf{x}'\mathbf{x} - \hat{\mathbf{h}}'(\mathbf{X}'\mathbf{X})\hat{\mathbf{h}}, \tag{11.3.6}$$

or, using (11.3.4) and (11.3.5),

$$\sum_t z_t^2 = Nc_{(q+1)(q+1)} - \hat{\mathbf{h}}'\mathbf{c}_{q+1}. \tag{11.3.7}$$

Equation (11.3.6) shows that the residual sum of squares can be written as the difference between the total or output sum of squares and a positive quantity called the *regression sum of squares.* If the regression sum of squares is expressed as a proportion $r^2_{(q+1)12\ldots q}$ of the total sum of squares, then (11.3.7) may be written

$$\sum_{t=1}^{N} z_t^2 = Nc_{(q+1)(q+1)}(1 - r^2_{(q+1)12\ldots q}), \tag{11.3.8}$$

and $r^2_{(q+1)12\ldots q}$ is the square of the *multiple correlation coefficient* between the output x_{q+1} and the q inputs. Alternatively, the output variance may be written

$$c_{(q+1)(q+1)} = r^2_{(q+1)12\ldots q}c_{(q+1)(q+1)} + (1 - r^2_{(q+1)12\ldots q})c_{(q+1)(q+1)}. \tag{11.3.9}$$

Equation (11.3.9) shows that the output variance can be decomposed into the regression sum of squares, which represents that part of the output which

can be "accounted for" or predicted from the inputs, plus a residual sum of squares which is due to noise and cannot be predicted from the inputs. Hence the square of the multiple correlation coefficient represents that proportion of the output variance which can be accounted for by the inputs.

Equating (11.3.7) with (11.3.8) shows that the multiple correlation coefficient may be estimated from

$$r^2_{(q+1)12\ldots q}c_{(q+1)(q+1)} = \hat{\mathbf{h}}'\mathbf{c}_{(q+1)}. \tag{11.3.10}$$

Substituting for the estimates $\hat{\mathbf{h}}$ as given by (11.3.5) gives an alternative form for the multiple correlation coefficient,

$$r^2_{(q+1)12\ldots q} = 1 - \frac{|\mathbf{C}_{(q+1)(q+1)}|}{c_{(q+1)(q+1)}|\mathbf{C}_{qq}|}, \tag{11.3.11}$$

where $\mathbf{C}_{(q+1)(q+1)}$ is the covariance matrix of all the $q+1$ variables and $\mathbf{C}_{qq}$ the covariance matrix of the inputs only. Equation (11.3.11) may also be written in terms of the corresponding correlation matrices. Thus

$$r^2_{(q+1)12\ldots q} = 1 - \frac{|\mathbf{R}_{(q+1)(q+1)}|}{|\mathbf{R}_{qq}|}. \tag{11.3.12}$$

An example. For $q = 2$, (11.3.10) becomes

$$r^2_{312}c_{33} = (\hat{h}_1, \hat{h}_2)\begin{pmatrix} c_{13} \\ c_{23} \end{pmatrix} = \hat{h}_1 c_{13} + \hat{h}_2 c_{23}. \tag{11.3.13}$$

Using the alternative form (11.3.12),

$$\begin{aligned} r^2_{312} &= 1 - \frac{\begin{vmatrix} 1 & r_{12} & r_{13} \\ r_{21} & 1 & r_{23} \\ r_{31} & r_{32} & 1 \end{vmatrix}}{\begin{vmatrix} 1 & r_{12} \\ r_{21} & 1 \end{vmatrix}} \\ &= \frac{r_{13}^2 + r_{23}^2 - 2r_{12}r_{13}r_{23}}{1 - r_{12}^2}. \end{aligned} \tag{11.3.14}$$

Distribution theory of multiple correlation coefficients. To develop a distribution theory, the sample quantities in the above formulae are replaced by the corresponding rv's. Note that these rv's are assumed to be Normal, as are the residuals Z_t in (11.3.1). Then (11.3.9) corresponds to the decomposition of a χ^2_{N-1} into a χ^2_q and χ^2_{N-q-1}. Thus, under the null hypothesis that all the parameters h_i in the model (11.3.1) are zero, the rv corresponding to

$$\frac{r^2_{(q+1)12\ldots q}}{1 - r^2_{(q+1)12\ldots q}} \cdot \frac{N - 1 - q}{q} \tag{11.3.15}$$

will be distributed as $F_{q,N-1-q}$.

Confidence intervals. More generally, if the parameters in the model are non-zero, their joint confidence region is given by (A4.1.14), which in the present notation becomes

$$(\mathbf{h} - \hat{\mathbf{h}})'\mathbf{C}_{qq}(\mathbf{h} - \hat{\mathbf{h}}) \leqslant \frac{q}{N} f_{q,N-1-q}(1 - \alpha)s^2, \tag{11.3.16}$$

where s^2 is the estimate of the residual variance. For example, if $q = 2$, the confidence region for (h_1, h_2) is

$$(h_1 - \hat{h}_1)^2 c_{11} + (h_2 - \hat{h}_2)^2 c_{22} + 2(h_1 - \hat{h}_1)(h_2 - \hat{h}_2)c_{12} \leqslant \frac{2}{N} f_{2,N-3}(1 - \alpha)s^2. \tag{11.3.17}$$

Formulae for theoretical quantities. The above formulae were derived using sample functions. By suitable interpretation, they apply equally well to the population or theoretical quantities. For example, (11.3.8) becomes

$$\sigma_2^2 = \sigma_{q+1}^2(1 - \rho^2_{(q+1)12\ldots q}), \tag{11.3.18}$$

where $\rho^2_{(q+1)12\ldots q}$ is the theoretical multiple correlation coefficient.

11.3.3 Partial correlation

The multiple correlation coefficient measures the correlation between the output and the best prediction of the output using *all* the inputs. It is also useful to be able to measure the correlation between the output and a single input. This leads to the notion of a partial correlation coefficient.

To illustrate the basic idea, suppose $q = 2$ so that the model (11.3.1) is

$$X_{t3} - \mu_3 = h_1(X_{t1} - \bar{X}_1) + h_2(X_{t2} - \bar{X}_2) + Z_t.$$

Clearly, if both h_1 and h_2 are non-zero, the rv X_3 will be correlated with both X_1 and X_2. However, the correlation coefficients ρ_{31} and ρ_{32} which describe the separate correlations between (X_3, X_1) and (X_3, X_2) are not very meaningful, since X_1 and X_2 may be correlated. As an extreme example, it might happen that X_3 and X_1 are highly correlated with X_2. When allowance is made for ρ_{32} and ρ_{12}, the actual "direct" correlation between X_3 and X_1 may be very small.

Thus it is necessary to remove the influence of the variable X_2 before computing the correlation between X_3 and X_1. This is achieved by conducting a least squares regression of X_3 on X_2 and of X_1 on X_2. The partial correlation coefficient is then defined to be the correlation between the residuals from these two regressions. With the residuals may be associated rv's

$$E_1 = X_1 - \mu_1 - \frac{\gamma_{12}}{\gamma_{22}}(X_2 - \mu_2),$$

$$E_3 = X_3 - \mu_3 - \frac{\gamma_{23}}{\gamma_{22}}(X_2 - \mu_2),$$

where γ_{ik} is the covariance between X_i and X_k. Then

$$\text{Cov}[E_1, E_3] = \gamma_{13} - \frac{\gamma_{12}\gamma_{23}}{\gamma_{22}}, \tag{11.3.19}$$

$$\text{Var}[E_1] = \gamma_{11}(1 - \rho_{12}^2),$$

$$\text{Var}[E_3] = \gamma_{33}(1 - \rho_{23}^2).$$

Hence the correlation between E_3 and E_1 is

$$\rho_{3112} = \frac{\rho_{13} - \rho_{23}\rho_{12}}{\sqrt{(1 - \rho_{12}^2)(1 - \rho_{23}^2)}}, \tag{11.3.20}$$

which is called the *partial correlation* between X_3 and X_1, allowing for X_2. The corresponding sample partial correlation coefficient is obtained by replacing the population correlations ρ_{ij} by their estimates r_{ij}. The partial correlation coefficient ρ_{3211} is obtained by interchanging the suffixes 1 and 2 in (11.3.20).

Note that in the special case where the rv's X_3, X_1 and X_2 refer to three consecutive times of a stationary time series, $\rho_{13} = \rho(2)$, $\rho_{12} = \rho_{23} = \rho(1)$, where $\rho(k)$ is the acf at lag k. In this case (11.3.20) reduces to

$$\frac{\rho(2) - \rho^2(1)}{1 - \rho^2(1)},$$

which is the partial autocorrelation coefficient discussed in Sections 5.4.3 and 11.1.1.

In general, for q input variables the partial correlation coefficient between the output X_{q+1} and any input X_k is defined as the ordinary correlation coefficient between $(X_{q+1} - \hat{X}_{q+1})$ and $(X_k - \hat{X}_k)$, where $\hat{X}_{q+1}$, $\hat{X}_k$ are the least squares predictors of X_{q+1}, X_k obtained from the other variables excluding X_k. These have indices $1, 2, \ldots, k-1, k+1, \ldots, q$ and are denoted by K. It may be shown [1] that the general form for the partial correlation coefficient is

$$\rho_{(q+1)k|K} = \frac{\pi_{(q+1)k}}{\sqrt{\pi_{(q+1)(q+1)}\pi_{kk}}}, \tag{11.3.21}$$

where π_{lm} is the minor of the element ρ_{lm} in the correlation matrix $\mathbf{R}_{(q+1)(q+1)}$ of all $q + 1$ variables.

An example. Equation (11.3.20) may be obtained using (11.3.21) as follows.

$$\mathbf{R}_{33} = \begin{pmatrix} 1 & \rho_{12} & \rho_{13} \\ \rho_{21} & 1 & \rho_{23} \\ \rho_{31} & \rho_{32} & 1 \end{pmatrix},$$

and so

$$\pi_{31} = \rho_{13} - \rho_{12}\rho_{23},$$

$$\pi_{33} = 1 - \rho_{12}^2,$$

$$\pi_{11} = 1 - \rho_{23}^2,$$

which yields (11.3.20).

Analysis of variance. Using (11.3.14) and (11.3.20), with the theoretical correlations replaced by their sample values, it may be shown that

$$(1 - r_{312}^2) = (1 - r_{32}^2)(1 - r_{31|2}^2). \tag{11.3.22}$$

The significance of (11.3.22) may be seen by referring to (11.3.8), which shows that the residual sum of squares after fitting x_1 and x_2 is a proportion $1 - r_{312}^2$ of the total sum of squares. Equation (11.3.22) then shows that the reduction in the sum of squares as a result of fitting x_2 is proportional to $(1 - r_{32}^2)$ and that the *further reduction* as a result of fitting x_1 is $(1 - r_{31|2}^2)$. Note, however, that x_1 could be fitted first so that (11.3.22) has the alternative form

$$(1 - r_{312}^2) = (1 - r_{31}^2)(1 - r_{32|1}^2).$$

Thus the decomposition (11.3.9) may be set out in the form of two analysis of variance tables, as shown in Table 11.1.

TABLE 11.1: Analysis of variance table for a multiple regression analysis

Fitting x_1 first		
Source	Sum of squares	Degrees of freedom
fitting x_1	$r_{31}^2 \sum (x_{t3} - \bar{x}_3)^2$	1
fitting x_2, given x_1	$(r_{312}^2 - r_{31}^2) \sum (x_{t3} - \bar{x}_3)^2$	1
residual	$(1 - r_{312}^2) \sum (x_{t3} - \bar{x}_3)^2$	$N - 3$
total	$\sum (x_{t3} - \bar{x}_3)^2$	$N - 1$

Fitting x_2 first		
Source	Sum of squares	Degrees of freedom
fitting x_2	$r_{32}^2 \sum (x_{t3} - \bar{x}_3)^2$	1
fitting x_1, given x_2	$(r_{312}^2 - r_{32}^2) \sum (x_{t3} - \bar{x}_3)^2$	1
residual	$(1 - r_{312}^2) \sum (x_{t3} - \bar{x}_3)^2$	$N - 3$
total	$\sum (x_{t3} - \bar{x}_3)^2$	$N - 1$

Assuming that x_1 does not contribute to the prediction of x_3, the ratio

$$\frac{(N - 3)r_{31}^2}{1 - r_{312}^2}$$

may be compared with the $F_{1,N-3}$ distribution.

Similarly, under the assumption that x_2 does not contribute to the prediction of x_3 after x_1 has been fitted,

$$(N - 3)\left\{\frac{r_{312}^2 - r_{31}^2}{1 - r_{312}^2}\right\}$$

may be compared with the $F_{1,N-3}$ distribution.

An example. To illustrate the above ideas, the power station data of Figure 11.1 were filtered using the low-pass filter

$$H(\mathcal{Z}) = \{\tfrac{1}{49}(\mathcal{Z}^{24} + \cdots + \mathcal{Z} + 1 + \mathcal{Z}^{-1} + \cdots + \mathcal{Z}^{-24})\}^4.$$

The output variable was the alternator frequency F and the input variables were the in-phase current i_d and the out-of-phase current i_q. Since the filter rejects most of the power above $f = 0.04$ cps, only every twentieth value of the filtered series was retained, giving $N = 41$ values.

The model

$$X_{3t} - \mu_3 = h_1(X_{1t} - \bar{X}_1) + h_2(X_{2t} - \bar{X}_2) + Z_t$$

thus involves the steady-state gains h_1 from i_d to F, and h_2 from i_q to F. The normal equations (11.3.5) are

$$\begin{aligned} 100\hat{\mu}_3 &= 82.558, \\ 5.879\hat{h}_1 + 2.907\hat{h}_2 &= 1.145, \\ 2.907\hat{h}_1 + 43.488\hat{h}_2 &= 2.033, \end{aligned}$$

yielding estimates $\hat{\mu}_3 = 0.8256$, $\hat{h}_1 = -0.2253$ and $\hat{h}_2 = 0.06181$. The multiple correlation coefficient, calculated from (11.3.14), is 0.977 and the partial correlations, calculated from (11.3.20), are $r_{31|2} = -0.98$ and $r_{32|1} = 0.97$. The analysis of variance for this data is given in Table 11.2.

TABLE 11.2: Analysis of variance table for current-frequency data

Source	Sum of squares	Degrees of freedom	Source	Sum of squares	Degrees of freedom
fitting in-phase current	9.1265	1	fitting out-of-phase current	3.8958	1
fitting out-of-phase, given in-phase	6.5825	1	fitting in-phase, given out-of-phase	11.8132	1
residual	0.3655	38	residual	0.3655	38
total	16.0745	40	total	16.0745	40

The F ratios from Table 11.2 are all very large, and hence it may be concluded that both currents contribute strongly to the prediction of frequency. This is apparent from the large values of the partial autocorrelations. Table 11.2 also shows that the in-phase current is the more important single variable, since it accounts for a greater proportion of the total sum of squares. This is because $r_{13} = 0.75$, whereas $r_{23} = 0.49$. However, the large value of the partial correlation $r_{32|1} = 0.97$ shows that the out-of-phase current also makes an important contribution to the prediction of frequency.

11.3.4 Multivariate analysis, multiple outputs

The model. In the preceding sections, it has been assumed that there is one output variable and several input variables. In general there will be several output variables, so that the regression model may be written

$$\begin{aligned}
X_{t(q+1)} - \mu_{q+1} &= h_{(q+1)1}(X_{t1} - \bar{X}_1) + \cdots + h_{(q+1)q}(X_{tq} - \bar{X}_q) + Z_{(q+1)t} \\
X_{t(q+2)} - \mu_{q+2} &= h_{(q+2)1}(X_{t1} - \bar{X}_1) + \cdots + h_{(q+2)q}(X_{tq} - \bar{X}_q) + Z_{(q+2)t} \\
&\vdots \qquad\qquad \vdots \qquad\qquad \vdots \qquad\qquad \vdots \\
X_{t(q+r)} - \mu_{q+r} &= h_{(q+r)1}(X_{t1} - \bar{X}_1) + \cdots + h_{(q+r)q}(X_{tq} - \bar{X}_q) + Z_{(q+r)t}.
\end{aligned} \tag{11.3.23}$$

The branch of statistics which deals with models of the form (11.3.23) is called multivariate analysis, an account of which is given in [1].

Normal equations. It may be shown [1] that the estimates of the parameters which minimize the determinant of the covariance matrix of the estimates are those which separately minimize the residual sum of squares

$$\sum_{t=1}^{N} z_{kt}^2, \quad k = q+1, q+2, \ldots, q+r.$$

This means that, as far as the estimation equations are concerned, multivariate analysis can be reduced to q separate multiple regression analyses. Hence, using (11.3.5), the normal equations are

$$\mathbf{C}_{qq}\hat{\mathbf{h}}_k = \mathbf{c}_{q+k}, \quad k = 1, 2, \ldots, r. \tag{11.3.24}$$

The equations (11.3.24) can be assembled, after transposition, into a single matrix equation

$$\hat{\mathbf{H}}\mathbf{C}_{qq}' = \mathbf{C}_{qr}', \tag{11.3.25}$$

where $\mathbf{C}_{qq}$ is the $q \times q$ matrix of covariances of the inputs and $\mathbf{C}_{qr}$ is the $q \times r$ matrix of cross covariances between inputs and outputs.

An example. Consider the two-input, two-output system for which

$$\begin{aligned}
X_{t3} - \mu_3 &= h_{31}(X_{t1} - \bar{X}_1) + h_{32}(X_{t2} - \bar{X}_2) + Z_{3t}, \\
X_{t4} - \mu_4 &= h_{41}(X_{t1} - \bar{X}_1) + h_{42}(X_{t2} - \bar{X}_2) + Z_{4t}.
\end{aligned}$$

The estimates $\hat{\mu}_3$ and $\hat{\mu}_4$ are

$$\begin{aligned}
\hat{\mu}_3 &= \bar{X}_3, \\
\hat{\mu}_4 &= \bar{X}_4,
\end{aligned}$$

and the normal equations (11.3.24) are

$$\begin{pmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{pmatrix}\begin{pmatrix} \hat{h}_{31} \\ \hat{h}_{32} \end{pmatrix} = \begin{pmatrix} c_{13} \\ c_{23} \end{pmatrix}, \quad \begin{pmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{pmatrix}\begin{pmatrix} \hat{h}_{41} \\ \hat{h}_{42} \end{pmatrix} = \begin{pmatrix} c_{14} \\ c_{24} \end{pmatrix}.$$

The single matrix equation (11.3.25) becomes

$$\begin{pmatrix} \hat{h}_{31} & \hat{h}_{32} \\ \hat{h}_{41} & \hat{h}_{42} \end{pmatrix}\begin{pmatrix} c_{11} & c_{21} \\ c_{12} & c_{22} \end{pmatrix} = \begin{pmatrix} c_{13} & c_{23} \\ c_{14} & c_{24} \end{pmatrix}.$$

Residual covariance matrix. Since the model (11.3.23) represents an interconnected system, the rv's $Z_{(q+k)t}, Z_{(q+l)t}$ are usually correlated at simultaneous times, so they possess a covariance matrix with elements σ_{kl}^2. The estimated covariance matrix $\mathbf{V}_r$ has elements

$$\begin{aligned}
s_{(q+k)(q+l)}^2 &= \frac{1}{N}\sum_{t=1}^{N} z_{(q+k)t} z_{(q+l)t} \\
&= \frac{1}{N}\sum_{t=1}^{N} (X_{t(q+k)} - \hat{h}_{(q+k)1}x_{t1} - \cdots - \hat{h}_{(q+k)q}x_{tq}) \\
&\qquad \times (X_{t(q+l)} - \hat{h}_{(q+l)1}x_{t1} - \cdots - \hat{h}_{(q+l)q}x_{tq}).
\end{aligned} \tag{11.3.2}$$

Using the normal equation (11.3.24), (11.3.26) simplifies to

$$s_{(q+k)(q+l)}^2 = c_{(q+k)(q+l)} - \hat{h}_{(q+k)1}c_{1(q+l)} - \hat{h}_{(q+k)2}c_{2(q+l)} - \cdots - \hat{h}_{(q+k)q}c_{q(q+l)}. \tag{11.3.2}$$

Covariance matrix of estimators. Since the normal equations are obtained by treating each regression in (11.3.23) separately, it follows from (A4.1.9) that the covariance matrices of the estimators associated with the individual equations are

$$\begin{aligned}
\mathbf{V}[\mathbf{h}_{q+k}] &= (\mathbf{X}'\mathbf{X})^{-1}\sigma_{(q+k)(q+k)}^2 \\
&= \mathbf{C}_{qq}^{-1}\sigma_{(q+k)(q+k)}^2.
\end{aligned}$$

Using (11.3.24), the remaining covariances in the covariance matrix of the estimators of all parameters are derived from

$$\begin{aligned}
E[\hat{\mathbf{h}}_{q+k}\hat{\mathbf{h}}_{q+l}'] &= E[(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{x}_{q+k}\mathbf{x}_{q+l}'\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}] \\
&= (\mathbf{X}'\mathbf{X})^{-1}\sigma_{(q+k)(q+l)}.
\end{aligned} \tag{11.3.1}$$

Hence the $qr \times qr$ covariance matrix of the estimators of all parameters may be written

$$\mathbf{V}(\hat{\mathbf{h}}) = \begin{pmatrix} \mathbf{C}_{qq}\sigma^2_{(q+1)(q+1)} & \mathbf{C}_{qq}\sigma^2_{(q+1)(q+2)} & \cdots & \mathbf{C}_{qq}\sigma^2_{(q+1)(q+r)} \\ \mathbf{C}_{qq}\sigma^2_{(q+2)(q+1)} & \mathbf{C}_{qq}\sigma^2_{(q+2)(q+2)} & \cdots & \mathbf{C}_{qq}\sigma^2_{(q+2)(q+r)} \\ \vdots & \vdots & & \vdots \\ \mathbf{C}_{qq}\sigma^2_{(q+r)(q+1)} & \mathbf{C}_{qq}\sigma^2_{(q+r)(q+2)} & \cdots & \mathbf{C}_{qq}\sigma^2_{(q+r)(q+r)} \end{pmatrix}, \tag{11.3.29}$$

where $\hat{\mathbf{h}}' = (\hat{\mathbf{h}}_{q+1}, \hat{\mathbf{h}}_{q+2}, \ldots, \hat{\mathbf{h}}_{q+r})$. This matrix (11.3.29) may be written more concisely as the *direct product matrix*

$$\mathbf{V}(\hat{\mathbf{h}}) = \mathbf{C}_{qq} \odot \mathbf{C}_{zz} \tag{11.3.30}$$

of the input covariance matrix $\mathbf{C}_{qq}$ and the residual covariance matrix $\mathbf{C}_{zz}$. The matrix (11.3.30) can therefore be estimated by substituting the estimates (11.3.27) for the σ^2_{kl} in (11.3.29).

Using (11.3.30), confidence intervals for the extended vector of parameters $\mathbf{h}'$ may be obtained from

$$(\mathbf{h} - \hat{\mathbf{h}})'\mathbf{V}(\hat{\mathbf{h}})(\mathbf{h} - \hat{\mathbf{h}}) \leq \frac{qr}{N} s^2 f_{qr,q(N-1-r)}(1 - \alpha). \tag{11.3.31}$$

An example. For the two-input, two-output system discussed above, the estimated covariance matrix of the residuals is, using (11.3.27),

$$\mathbf{V}_{zz} = \begin{pmatrix} s^2_{33} & s^2_{34} \\ s^2_{43} & s^2_{44} \end{pmatrix} = \begin{pmatrix} c_{33} - \hat{h}_{31}c_{13} - \hat{h}_{32}c_{23} & c_{34} - \hat{h}_{31}c_{14} - \hat{h}_{32}c_{24} \\ c_{43} - \hat{h}_{41}c_{13} - \hat{h}_{42}c_{23} & c_{44} - \hat{h}_{41}c_{14} - \hat{h}_{42}c_{24} \end{pmatrix}.$$

The covariance matrices of the estimates associated with the individual equations are

$$\mathbf{V}[\hat{\mathbf{h}}_3] = \frac{\begin{pmatrix} c_{11} & c_{21} \\ c_{12} & c_{22} \end{pmatrix}}{D} \sigma^2_{33}, \quad \mathbf{V}[\hat{\mathbf{h}}_4] = \frac{\begin{pmatrix} c_{11} & c_{21} \\ c_{12} & c_{22} \end{pmatrix}}{D} \sigma^2_{44},$$

where $D = c_{11}c_{22} - c^2_{12}$.

Hence

$$\text{Var}[\hat{h}_{31}] = \frac{c_{11}}{D}\sigma^2_{33}, \qquad \text{Var}[\hat{h}_{41}] = \frac{c_{11}}{D}\sigma^2_{44},$$

$$\text{Cov}[\hat{h}_{31}, \hat{h}_{32}] = \frac{c_{21}}{D}\sigma^2_{33}, \qquad \text{Cov}[\hat{h}_{41}, \hat{h}_{42}] = \frac{c_{21}}{D}\sigma^2_{44},$$

$$\text{Cov}[\hat{h}_{32}, \hat{h}_{31}] = \frac{c_{12}}{D}\sigma^2_{33}, \qquad \text{Cov}[\hat{h}_{42}, \hat{h}_{41}] = \frac{c_{12}}{D}\sigma^2_{44},$$

$$\text{Var}[\hat{h}_{32}] = \frac{c_{22}}{D}\sigma^2_{33}, \qquad \text{Var}[\hat{h}_{42}] = \frac{c_{22}}{D}\sigma^2_{44}.$$

The covariance matrix (11.3.29) of the estimates of all the parameters is then

$$\mathbf{V}(\hat{\mathbf{h}}) = \begin{pmatrix} \text{Cov}[\hat{h}_{31}, \hat{h}_{31}] & \text{Cov}[\hat{h}_{31}, \hat{h}_{32}] & \text{Cov}[\hat{h}_{31}, \hat{h}_{41}] & \text{Cov}[\hat{h}_{31}, \hat{h}_{42}] \\ \text{Cov}[\hat{h}_{32}, \hat{h}_{31}] & \text{Cov}[\hat{h}_{32}, \hat{h}_{32}] & \text{Cov}[\hat{h}_{32}, \hat{h}_{41}] & \text{Cov}[\hat{h}_{32}, \hat{h}_{42}] \\ \text{Cov}[\hat{h}_{41}, \hat{h}_{31}] & \text{Cov}[\hat{h}_{41}, \hat{h}_{32}] & \text{Cov}[\hat{h}_{41}, \hat{h}_{41}] & \text{Cov}[\hat{h}_{41}, \hat{h}_{42}] \\ \text{Cov}[\hat{h}_{42}, \hat{h}_{31}] & \text{Cov}[\hat{h}_{42}, \hat{h}_{32}] & \text{Cov}[\hat{h}_{42}, \hat{h}_{41}] & \text{Cov}[\hat{h}_{42}, \hat{h}_{42}] \end{pmatrix}.$$

Finally, the estimate of $\mathbf{V}(\hat{\mathbf{h}})$ is given by

$$\mathbf{V}(\hat{\mathbf{h}}) = \mathbf{C}_{qq} \otimes \mathbf{V}_{zz}$$

or

$$\mathbf{V}(\hat{\mathbf{h}}) = \frac{1}{D}\begin{pmatrix} c_{11}s^2_{33} & c_{11}s^2_{34} & c_{12}s^2_{33} & c_{12}s^2_{34} \\ c_{11}s^2_{43} & c_{11}s^2_{44} & c_{12}s^2_{43} & c_{12}s^2_{44} \\ c_{21}s^2_{33} & c_{21}s^2_{34} & c_{22}s^2_{33} & c_{22}s^2_{34} \\ c_{21}s^2_{43} & c_{21}s^2_{44} & c_{22}s^2_{43} & c_{22}s^2_{44} \end{pmatrix}.$$

11.4 FREQUENCY DOMAIN MULTIVARIATE ANALYSIS

In this section the methods of Section 11.3 are generalized so that they can be applied in the frequency domain. There are two main points of difference between the models used in this section and those in Section 11.3. First, the models in Section 11.3 described regressions and correlations between processes at simultaneous times and so they describe only the steady-state behavior of systems. The models considered in this section are the dynamic generalizations of these models. Second, the noise, or residuals, in the models of Section 11.3 were assumed to be white. In this section, the noises can be quite general stationary time series.

11.4.1 Multiple frequency response analysis, single output

In this section it is shown how to estimate the frequency response functions associated with the model

$$X_{(q+1)}(t) - \mu_{(q+1)} = \int_{-\infty}^{\infty} h_{(q+1)1}(u)(X_1(t-u) - \bar{X}_1)\,du + \cdots + \int_{-\infty}^{\infty} h_{(q+1)q}(u)(X_q(t-u) - \bar{X}_q)\,du + Z(t), \tag{11.4.1}$$

which is the dynamic generalization of the steady-state model (11.3.1). To clarify ideas, it is assumed that infinite lengths of record are available for the single output $X_{(q+1)}(t)$ and the q inputs. To simplify the problem still further consider the special case of $q = 2$ inputs.

Proceeding as in Appendix A5.1, it may be shown that the minimum mse estimates of the impulse response functions $h_{31}(u)$ and $h_{32}(u)$ are obtained by solving the simultaneous Wiener–Hopf equations

$$\gamma_{13}(u) = \int_{-\infty}^{\infty} h_{31}(v)\gamma_{11}(u - v)\,dv + \int_{-\infty}^{\infty} h_{32}(v)\gamma_{12}(u - v)\,dv,$$
$$\gamma_{23}(u) = \int_{-\infty}^{\infty} h_{31}(v)\gamma_{21}(u - v)\,dv + \int_{-\infty}^{\infty} h_{32}(v)\gamma_{22}(u - v)\,dv. \qquad (11.4.2)$$

Note that the equations (11.4.2) can also be obtained by multiplying throughout in (11.4.1), first by $(X_1(t - u) - \bar{X}_1)$ and then by $(X_2(t - u) - \bar{X}_2)$, and taking expectations.

Taking Fourier transforms of (11.4.2) gives the frequency domain equations

$$\Gamma_{13}(f) = H_{31}(f)\Gamma_{11}(f) + H_{32}(f)\Gamma_{12}(f)$$
$$\Gamma_{23}(f) = H_{31}(f)\Gamma_{21}(f) + H_{32}(f)\Gamma_{22}(f). \qquad (11.4.3)$$

Solving for $H_{31}(f)$ and $H_{32}(f)$ gives the following expressions for the frequency response functions, in terms of the auto- and cross spectra:

$$H_{31}(f) = \frac{\Gamma_{13}(f)\Gamma_{22}(f) - \Gamma_{23}(f)\Gamma_{12}(f)}{\Gamma_{11}(f)\Gamma_{22}(f) - |\Gamma_{12}(f)|^2},$$
$$H_{32}(f) = \frac{\Gamma_{23}(f)\Gamma_{11}(f) - \Gamma_{13}(f)\Gamma_{21}(f)}{\Gamma_{11}(f)\Gamma_{22}(f) - |\Gamma_{12}(f)|^2}. \qquad (11.4.4)$$

To obtain expressions for the gains and phases, it is necessary to take the modulus and argument of (11.4.4). For example,

$$G_{31} = \sqrt{A_{31}^2 + B_{31}^2}, \qquad \phi_{31} = \arctan -\frac{B_{31}}{A_{31}}, \qquad (11.4.5)$$

where

$$A_{31} = \frac{(\Lambda_{13}\Gamma_{22} + \Psi_{23}\Psi_{12} - \Lambda_{23}\Lambda_{12})}{(\Gamma_{11}\Gamma_{22} - |\Gamma_{12}|^2)},$$
$$B_{31} = \frac{(\Psi_{13}\Gamma_{22} - \Psi_{23}\Lambda_{12} - \Lambda_{23}\Psi_{12})}{(\Gamma_{11}\Gamma_{22} - |\Gamma_{12}|^2)}. \qquad (11.4.6)$$

For q inputs, (11.4.2) may be written

$$\boldsymbol{\gamma}_{(q+1)}(u) = \int_{-\infty}^{\infty} \boldsymbol{\gamma}_{qq}(u - v)\mathbf{h}_{(q+1)}(v)\,dv, \qquad (11.4.7)$$

where $\boldsymbol{\gamma}_{(q+1)}(u)$ is the lagged vector of cross covariances between the output $q + 1$ and all the inputs, $\mathbf{h}'_{(q+1)} = (h_{(q+1)1}, h_{(q+1)2}, \ldots, h_{(q+1)q})$ and $\boldsymbol{\gamma}_{qq}(u)$ is the lagged covariance matrix of the inputs. Transforming (11.4.7) gives

$$\boldsymbol{\Gamma}_{(q+1)}(f) = \boldsymbol{\Gamma}_{qq}(f)\mathbf{H}_{(q+1)}(f), \qquad (11.4.8)$$

where $\boldsymbol{\Gamma}_{(q+1)}(f)$ is the vector of cross spectra between the output $X_{(q+1)}(t)$ and the inputs, $\boldsymbol{\Gamma}_{qq}(f)$ is the spectral matrix of the inputs and $\mathbf{H}'_{(q+1)}(f) = (H_{(q+1)1}(f), H_{(q+1)2}(f), \ldots, H_{(q+1)q}(f))$. The gain and phase functions may be obtained by solving these equations and taking their modulus and argument as described above.

11.4.2 The multiple coherency spectrum

In this section the multiple coherency spectrum is defined. This is a generalization to the frequency domain of the multiple correlation coefficient introduced in Section 11.3.2. First it is necessary to derive an expression for the residual or noise spectrum, which is one of the basic quantities of interest, apart from the estimates of gain and phase.

The residual spectrum. To evaluate the residual spectrum $\Gamma_{zz}(f)$ corresponding to the model (11.4.1), it is first necessary to compute the autocovariance function of the residual process. Proceeding as in Section 11.3.2, the autocovariance function of $Z(t)$ is

$$\gamma_{zz}(u) = \gamma_{(q+1)(q+1)}(u) - \int_0^{\infty} h_{(q+1)1}(v)\gamma_{(q+1)1}(u - v)\,dv - \cdots - \int_0^{\infty} h_{(q+1)q}(v)\gamma_{(q+1)q}(u - v)\,dv.$$

Taking Fourier transforms gives the residual spectrum

$$\Gamma_{zz}(f) = \Gamma_{(q+1)(q+1)}(f) - H_{(q+1)1}(f)\Gamma_{(q+1)1}(f) - \cdots - H_{(q+1)q}(f)\Gamma_{(q+1)q}(f). \qquad (11.4.9)$$

which is the frequency domain analog of (11.3.7).

The squared multiple coherency spectrum. Proceeding as in Section 11.3.2, (11.4.9) may be written

$$\Gamma_{zz}(f) = \Gamma_{(q+1)(q+1)}(f)[1 - \kappa^2_{(q+1)12\ldots q}(f)], \qquad (11.4.10)$$

where

$$\kappa^2_{(q+1)12\ldots q}(f) = H_{(q+1)1}(f)\Gamma_{(q+1)1}(f) + \cdots + H_{(q+1)q}(f)\Gamma_{(q+1)q}(f)$$

is called the *squared multiple coherency spectrum* of the output process and the q input processes. The multiple coherency spectrum measures the proportion of the output spectrum which can be predicted from the inputs. As shown by (11.4.10), the remaining proportion $[1 - \kappa^2_{(q+1)12\ldots q}(f)]$ of the output spectrum is noise.

On substituting in (11.4.9) the expressions (11.4.8) for the frequency response functions, an alternative form for the squared multiple coherency spectrum is obtained which is analogous to (11.3.11), namely,

$$\kappa^2_{(q+1)12\ldots q}(f) = 1 - \frac{|\boldsymbol{\Gamma}_{(q+1)(q+1)}(f)|}{\Gamma_{(q+1)(q+1)}(f)\,|\boldsymbol{\Gamma}_{qq}(f)|}. \qquad (11.4.11)$$

In (11.4.11), $\mathbf{\Gamma}_{(q+1)(q+1)}(f)$ is the spectral matrix of all $q+1$ variables, and $\mathbf{\Gamma}_{qq}(f)$ is the spectral matrix of the q input variables. When $q = 2$, (11.4.11) becomes, on dropping the dependence on f,

$$\kappa^2_{3|12} = 1 - \frac{\begin{vmatrix} \Gamma_{11} & \Gamma_{12} & \Gamma_{13} \\ \Gamma_{21} & \Gamma_{22} & \Gamma_{23} \\ \Gamma_{31} & \Gamma_{32} & \Gamma_{33} \end{vmatrix}}{\Gamma_{33}\begin{vmatrix} \Gamma_{11} & \Gamma_{12} \\ \Gamma_{21} & \Gamma_{22} \end{vmatrix}}, \tag{11.4.12}$$

which corresponds to (11.3.14). When expanded, (11.4.12) becomes

$$\kappa^2_{3|12} = \frac{\Gamma_{22}|\Gamma_{31}|^2 + \Gamma_{11}|\Gamma_{32}|^2 - 2\,\mathrm{Re}\,[\Gamma_{12}\Gamma_{23}\Gamma_{31}]}{\Gamma_{33}(\Gamma_{11}\Gamma_{22} - |\Gamma_{12}|^2)} \tag{11.4.13}$$

where

$$\mathrm{Re}\,[\Gamma_{12}\Gamma_{23}\Gamma_{31}] = \Lambda_{12}\Lambda_{23}\Lambda_{13} + \Lambda_{12}\Psi_{23}\Psi_{13} - \Psi_{12}\Psi_{23}\Lambda_{13} + \Psi_{12}\Psi_{13}\Lambda_{23} \tag{11.4.14}$$

may be expressed in terms of the co- and quadrature spectra of the three processes.

11.4.3 Partial cross, squared coherency and phase spectra

As in multiple regression analysis, it is useful to be able to measure the cross spectrum between the output and one of the input processes after allowance is made for the effect of the other input processes. This leads to the partial cross spectrum, which is the frequency domain analog of the partial correlation coefficient (11.3.21).

To illustrate the basic idea, it is assumed that there are just two input variables. Generalizing the approach of Section 11.3.3, the output $X_3(t)$ is first predicted from past values of the $X_2(t)$ process only, leading to residuals

$$\epsilon_3(t) = (X_3(t) - \mu_3) - \int_0^\infty \hat{g}_{32}(u)(X_2(t-u) - \mu_2)\,du$$

where $\hat{g}_{32}(u)$ is given by the solution of the appropriate Wiener–Hopf integral equation. Similarly, the input $X_1(t)$ is predicted from $X_2(t)$ only, leading to residuals

$$\epsilon_1(t) = (X_1(t) - \mu_1) - \int_0^\infty \hat{g}_{12}(u)(X_2(t-u) - \mu_2)\,du.$$

The partial cross covariance function. The partial cross covariance function between $X_1(t)$ and $X_3(t+u)$, after allowing for $X_2(t)$, is then defined by

$$\begin{aligned}\gamma_{13|2}(u) &= \mathrm{Cov}\,[\epsilon_1(t), \epsilon_3(t+u)] \\ &= \gamma_{13}(u) - \int_0^\infty \hat{g}_{32}(v)\gamma_{12}(u-v)\,dv - \int_0^\infty \hat{g}_{12}(v)\gamma_{23}(u+v)\,dv \\ &\quad + \int_0^\infty \int_0^\infty \hat{g}_{12}(v)\hat{g}_{32}(w)\gamma_{22}(u+v-w)\,dv\,dw. \end{aligned} \tag{11.4.15}$$

The partial cross spectrum. The partial cross spectrum is derived by taking the Fourier transform of (11.4.15), and substituting $\hat{G}_{12} = \Gamma_{12}/\Gamma_{22}$, $\hat{G}_{32} = \Gamma_{23}/\Gamma_{22}$, namely,

$$\Gamma_{13|2}(f) = \Gamma_{13}(f) - \frac{\Gamma_{23}(f)\Gamma_{12}(f)}{\Gamma_{22}(f)}. \tag{11.4.16}$$

The *partial cross spectral density function*

$$\kappa_{13|2}(f) = \frac{\Gamma_{13|2}(f)}{\sqrt{\Gamma_{11}(f)(1-\kappa^2_{12}(f))\Gamma_{33}(f)(1-\kappa^2_{23}(f))}} \tag{11.4.17}$$

is obtained by normalizing the partial cross spectrum.

Partial coherency spectrum. The squared partial coherency spectrum is the squared modulus of $\kappa_{13|2}(f)$. It is most simply calculated using the spectral analog of (11.3.22), namely,

$$1 - \kappa^2_{13|2}(f) = \frac{1 - \kappa^2_{3|12}(f)}{1 - \kappa^2_{23}(f)}. \tag{11.4.18}$$

The partial coherency $\kappa^2_{13|2}(f)$ measures the squared covariance "at frequency f" between the processes $X_3(t)$ and $X_1(t)$ when allowance is made for the influence of $X_2(t)$.

Partial phase spectrum. The partial phase spectrum $\phi_{13|2}(f)$ is the argument of (11.4.16) or (11.4.17) and is given by

$$\phi_{13|2}(f) = \arctan\left\{\frac{\Lambda_{12}\Psi_{23} - \Psi_{12}\Lambda_{23} - \Psi_{13}\Gamma_{22}}{\Lambda_{12}\Lambda_{23} + \Psi_{12}\Psi_{23} - \Lambda_{13}\Gamma_{22}}\right\}. \tag{11.4.19}$$

Similar expressions for the partial squared coherency $\kappa^2_{23|1}(f)$ and the partial phase $\phi_{23|1}(f)$ may be obtained by interchanging the indices 1 and 2 in (11.4.18) and (11.4.19).

The difference between the partial phase $\phi_{13|2}(f)$ and the phases $\phi_{31}(f)$, $\phi_{32}(f)$, obtained from the frequency response model (11.4.1), should be noted. The phase $\phi_{31}(f)$ measures the phase difference between $X_3(t)$ and $X_1(t)$ when a sinusoidal change is made in $X_1(t)$ but there is no change in $X_2(t)$. However, the partial phase $\phi_{13|2}(f)$ measures the "direct" phase difference between $X_1(t)$ and $X_3(t)$ after allowing for the phase differences between $X_2(t)$ and $X_3(t)$ and between $X_2(t)$ and $X_1(t)$. When there is just one input variable, the partial phase angle is equivalent to the ordinary phase angle.

For q inputs, the partial cross spectrum is given by the spectral analog of (11.3.21), namely,

$$\kappa_{(q+1)k|K}(f) = \frac{\pi_{(q+1)k}(f)}{\sqrt{\pi_{(q+1)(q+1)}(f)\pi_{kk}(f)}}, \tag{11.4.20}$$

where π_{lm} is the minor of the element Γ_{lm} in the spectral matrix of all $(q + 1)$ variables.

The partial coherency and phase spectra may then be calculated from (11.4.20).

Summary. As in the analysis of bivariate time series, interest is focused on different types of spectral estimates, depending on whether all the series in a multiple time series arise on an equal footing or whether some are inputs and some outputs to a physical system. If all the series arise on an equal footing, then the main spectrum of interest is the multiple coherency. This is usually supplemented by information about the partial coherency and partial phase spectra between selected pairs of the variables. If the series represent inputs and corresponding outputs to some physical system, the most important part of the analysis is concerned with the estimation of the frequency response functions of the system. The other important estimate is the residual spectrum which describes the noise in the system. In this case, the multiple coherency is only of interest insofar as it affects the confidence intervals for the gain and phase matrix. The estimation of the multiple coherency spectrum is discussed in Section 11.4.5. Confidence intervals for gain and phase functions are derived in Section 11.4.6.

11.4.4 Multivariate frequency response analysis, multiple outputs

The model. In this section the time-domain multivariate analysis discussed in Section 11.3.4 is extended to the frequency domain. The steady-state model (11.3.23) is generalized to the dynamic model

$$\mathbf{x}_{(q+r)}(t) - \boldsymbol{\mu}_{(q+r)} = \int_{-\infty}^{\infty} \mathbf{h}(u)(\mathbf{x}_q(t-u) - \bar{\mathbf{x}}_q)\,du + \mathbf{z}_{(q+r)}(t), \quad (11.4.21)$$

where $\mathbf{x}_{(q+r)}(t)$ is the vector of output variables, $\mathbf{x}_q(t)$ is a vector of input variables and $\mathbf{z}_{(q+r)}(t)$ is a vector of noise variables. For example, when there are $q = 2$ inputs and $r = 2$ outputs (11.4.21) is

$$\begin{aligned} X_3(t) - \mu_3 &= \int_{-\infty}^{\infty} h_{31}(u)(X_1(t-u) - \bar{X}_1)\,du + \\ &\qquad \int_{-\infty}^{\infty} h_{32}(u)\,(X_2(t-u) - \bar{X}_2)\,du + Z_3(t), \\ X_4(t) - \mu_4 &= \int_{-\infty}^{\infty} h_{41}(u)(X_1(t-u) - \bar{X}_1)\,du + \\ &\qquad \int_{-\infty}^{\infty} h_{42}(u)(X_2(t-u) - \bar{X}_2)\,du + Z_4(t). \end{aligned} \quad (11.4.22)$$

As in the preceding sections, it is assumed initially that infinite lengths of record are available for all the inputs and outputs.

Estimation equations. As in Section 11.3.4, the minimum mse estimates of the impulse response functions $h_{ij}(u)$ are obtained by separately minimizing the mean square errors

$$E\left[\int_{-\infty}^{\infty} Z_i^2(t)\,dt\right].$$

This procedure leads to a set of equations of the form (11.4.7), namely,

$$\boldsymbol{\gamma}_{(q+k)}(u) = \int_{-\infty}^{\infty} \boldsymbol{\gamma}_{qq}(u-v)\mathbf{h}_{(q+k)}(v)\,dv, \quad k = 1, 2, \ldots, r. \quad (11.4.23)$$

These may be assembled into a single matrix equation

$$\boldsymbol{\gamma}_{qr}(u) = \int_{-\infty}^{\infty} \boldsymbol{\gamma}_{qq}(u-v)\mathbf{h}'(v)\,dv, \quad (11.4.24)$$

where $\boldsymbol{\gamma}_{qr}(u)$ is the lagged cross covariance matrix between the q inputs and r outputs, $\mathbf{h}(u)$ is the impulse response matrix and $\boldsymbol{\gamma}_{qq}(u)$ is the lagged covariance matrix of the inputs.

On taking Fourier transforms of (11.4.23), the estimation equations for the frequency response functions may be written

$$\boldsymbol{\Gamma}_{(q+k)}(f) = \boldsymbol{\Gamma}_{qq}(f)\mathbf{H}_{(q+k)}(f), \quad k = 1, 2, \ldots, r. \quad (11.4.25)$$

As in (11.4.24), the equations (11.4.25) may be assembled into the single matrix equation

$$\boldsymbol{\Gamma}_{qr}(f) = \boldsymbol{\Gamma}_{qq}(f)\mathbf{H}'(f). \quad (11.4.26)$$

The residual spectral matrix. In addition to estimating the frequency response matrix, it is necessary to characterize the properties of the noise. This is done by calculating the residual or noise spectral matrix $\boldsymbol{\Gamma}_{zz}(f)$ whose elements are the cross spectra $\Gamma_{kl}(f)$ between the processes $Z_k(t)$ and $Z_l(t)$. Proceeding as in Section 11.3.4, the cross covariance function between the processes is

$$\begin{aligned} \gamma_{Z_kZ_l}(u) = \gamma_{(q+k)(q+l)}(u) &- \int_{-\infty}^{\infty} h_{(q+k)1}(v)\gamma_{(q+l)1}(u-v)\,dv - \\ &\cdots - \int_{-\infty}^{\infty} h_{(q+k)q}(v)\gamma_{(q+l)q}(u-v)\,dv. \end{aligned} \quad (11.4.27)$$

On transforming (11.4.27), the cross spectrum is given by

$$\Gamma_{Z_kZ_l}(f) = \Gamma_{(q+k)(q+l)}(f) - H_{(q+k)1}(f)\Gamma_{(q+l)1}(f) - \cdots - H_{(q+k)q}(f)\Gamma_{(q+l)q}(f). \quad (11.4.28)$$

Assembling a matrix whose elements are (11.4.28) gives

$$\boldsymbol{\Gamma}'_{zz}(f) = \boldsymbol{\Gamma}'_{(q+r)(q+r)}(f) - \mathbf{H}(f)\boldsymbol{\Gamma}'_{qr}(f). \quad (11.4.29)$$

Estimation equations. Estimation equations may be obtained by replacing the spectra in (11.4.26) by their smoothed estimators. Thus

Note that (11.4.29) enables the residual spectral matrix to be calculated when the frequency response matrix $\mathbf{H}(f)$ and the theoretical values of the output spectral matrix $\mathbf{\Gamma}_{(q+r)(q+r)}(f)$ and input–output cross spectral matrix $\mathbf{\Gamma}_{qr}(f)$ are known. The corresponding estimation problem is considered in Section 11.4.6.

11.4.5 Estimation of multivariate spectra

It was shown in Sections 11.4.2 and 11.4.3 how to compute multiple and partial coherency spectra in terms of the input and output spectra and cross spectra. In this section the problem of estimating these spectra is considered when finite lengths of record are available. The analysis is a straightforward extension of that given in Section 9.3.1 and hence the details are omitted.

Estimation of multiple coherency. The definition (11.4.11) expresses the square of the multiple coherency in terms of auto- and cross spectra. The sample estimator of the multiple coherency is obtained by replacing the theoretical spectra by their smoothed estimators. For example, when $q = 2$, the smoothed multiple coherency estimator is

$$\bar{K}^2_{312}(f) = \frac{\bar{C}_{22}\,|\bar{C}_{13}|^2 + \bar{C}_{11}\,|\bar{C}_{23}|^2 - 2\mathrm{Re}\,[\bar{C}_{12}\bar{C}_{23}\bar{C}_{31}]}{\bar{C}_{33}(\bar{C}_{11}\bar{C}_{22} - |\bar{C}_{12}|^2)}, \tag{11.4.30}$$

where

$$\mathrm{Re}\,[\bar{C}_{12}\bar{C}_{23}\bar{C}_{31}] = \bar{L}_{12}\bar{L}_{23}\bar{L}_{13} + \bar{L}_{12}\bar{Q}_{23}\bar{Q}_{13} - \bar{Q}_{12}\bar{Q}_{23}\bar{L}_{13} + \bar{Q}_{12}\bar{Q}_{13}\bar{L}_{23}.$$

Since the estimator (11.4.30) is a function of spectra and cross spectra, its variance can be evaluated using the statistical differential technique of Section 3.2.5 and the result (A9.1.28), namely,

$$\mathrm{Cov}\,[\bar{C}_{ij}(f_1), \bar{C}_{kl}(f_1)] \approx \Gamma_{il}(f_1)\Gamma_{jk}(f_1)\frac{I}{T}. \tag{11.4.31}$$

Proceeding as in Section 9.2.2, the final result is

$$\mathrm{Var}\,[\bar{K}^2_{312}] \approx \frac{1}{2T}\cdot 4\kappa^2_{312}(1 - \kappa^2_{312})^2,$$

which is the same as the variance (9.2.19) for the ordinary coherency estimator $\bar{K}^2_{12}$. In Section 11.4.6, an appropriate distribution for $\bar{K}^2_{312}$ will be derived by regarding the problem as one of multiple regression analysis in the frequency domain.

Estimation of partial coherency and phase. The smoothed partial cross spectral estimators are obtained by substituting smoothed estimators for the cross spectra in (11.4.19). The smoothed partial coherency spectra and partial phase spectra can then be obtained by taking the modulus squared and argument of the smoothed cross spectral estimators. For example, when $q = 2$,

These lead to the discrete estimation equations

$$\bar{C}_{1k}(f) = \hat{H}_{k1}(f)\bar{C}_{11}(f) + \hat{H}_{k2}(f)\bar{C}_{12}(f),$$

the two partial coherencies may be estimated from the smoothed multiple coherency $\bar{K}^2_{312}$ using the relations

$$1 - \bar{K}^2_{2311} = \frac{1 - \bar{K}^2_{312}}{1 - \bar{K}^2_{13}},$$

$$1 - \bar{K}^2_{1312} = \frac{1 - \bar{K}^2_{312}}{1 - \bar{K}^2_{23}}. \tag{11.4.32}$$

Similarly, the smoothed partial phase spectra are

$$\bar{F}_{1312} = \arctan\left[\frac{\bar{L}_{12}\bar{Q}_{23} - \bar{Q}_{12}\bar{L}_{23} - \bar{Q}_{13}\bar{C}_{22}}{\bar{L}_{12}\bar{L}_{23} + \bar{Q}_{12}\bar{Q}_{23} - \bar{L}_{13}\bar{C}_{22}}\right], \tag{11.4.33}$$

$$\bar{F}_{2311} = \arctan\left[\frac{\bar{L}_{12}\bar{Q}_{13} + \bar{Q}_{12}\bar{L}_{13} - \bar{Q}_{23}\bar{C}_{11}}{\bar{L}_{12}\bar{L}_{13} - \bar{Q}_{12}\bar{Q}_{13} - \bar{L}_{23}\bar{C}_{11}}\right]. \tag{11.4.34}$$

11.4.6 Estimation of multiple frequency response functions

In this section it is shown how to estimate the frequency response functions corresponding to the model (11.4.1) and to derive confidence intervals for the gains and phases. The results are obtained by a simple extension of those derived in Section 10.3.3.

Estimates based on sample spectra. As in previous work, rv's are associated with the transforms of the data, namely,

$$X_i(f) = \int_0^T X_i(t)\, e^{-j2\pi ft}\, dt.$$

On transforming the model (11.4.1), and making the usual assumptions that the impulse response functions $h_{(q+1)i}(u)$ tend to zero quickly compared with the record length, the output transform is

$$X_{(q+1)}(f) \approx H_{(q+1)1}(f)X_1(f) + \cdots + H_{(q+1)q}(f)X_q(f) + Z_{(q+1)}(f). \tag{11.4.35}$$

Proceeding as in Section 10.3.1, the least squares estimators of the impulse response functions are obtained by replacing the theoretical auto- and cross correlations in (11.4.7) by sample values. Thus

$$\mathbf{c}_{(q+1)}(u) = \int_{-T}^{T} \mathbf{C}_{qq}(u - v)\,\hat{\mathbf{h}}_{(q+1)}(v)\, dv, \quad -T \leqslant u \leqslant T. \tag{11.4.36}$$

On taking Fourier transforms of (11.4.36), the frequency domain estimation equations are

$$\mathbf{C}_{(q+1)}(f) = \mathbf{C}_{qq}(f)\hat{\mathbf{H}}_{(q+1)}(f), \tag{11.4.37}$$

which are the same as the equations (11.4.8) but with theoretical spectra replaced by the corresponding sample spectra.

Residual sample spectrum. Proceeding as in Section 10.3.2, the residual sample spectrum is

$$C_{zz}(f) \approx C_{\hat{z}\hat{z}}(f) + \frac{1}{T}|X_1(f)\{H_{(q+1)1}(f) - \hat{H}_{(q+1)1}(f)\} + \cdots + X_q(f)\{H_{(q+1)q}(f) - \hat{H}_{(q+1)q}(f)\}|^2. \quad (11.4.38)$$

Proceeding as in Section 10.3.4, the decomposition (11.4.43) leads to a joint confidence region

$$\frac{\bar{C}_{11}|H_{31} - \hat{H}_{31}|^2 + \bar{C}_{22}|H_{32} - \hat{H}_{32}|^2 + 2|\bar{C}_{12}||(H_{31} - \hat{H}_{31})(H_{32} - \hat{H}_{32})|}{\bar{C}_{\hat{z}\hat{z}}}$$

11.5.2 A practical procedure for estimating multivariate spectra

The stages in the procedure for multivariate frequency response estimation are very similar to those given for the estimation of cross spectra in Section 9.4.2 and for the single-input, single-output frequency response estimation in Section 10.4.1. Hence only a brief summary of the steps is presented here.

(1) Preliminary decision stage. As in Section 9.4.2, plots of the data are inspected for obvious trends, for possible filtering into different frequency bands and for deciding on the maximum number of lags for the computation of the acvf's and ccvf's.

(2) First computation stage. The auto- and cross correlations of the original and differenced data are computed and the correlations plotted. The correlation plots are inspected for failure to damp out, in which case the need for detrending is indicated, and for delays indicated by the peaks in the ccf.

(3) Intermediate decision stage. A decision is made whether to use the original or differenced covariances. The shift values for alignment are determined and three truncation points for future spectral calculations are selected.

(4) Second computation stage. The spectral calculations are performed and corresponding spectra overlayed for the three truncation values used.

(5) Interpretation stage. The plots are analyzed and interpreted, or additional spectral analyses are performed using the information gleaned from the present analysis.

11.5.3 Analysis of turbo-alternator data

These data have been analyzed previously in [3] and consist of normal operating data collected from a 50 megawatt turbo-alternator operating in parallel with an interconnected system having a capacity of approximately 5000 megawatts. A turbo-alternator may be regarded as a two-input, two-output system as shown in Figure 8.1. The input variables are the in-phase (or active) power and the out-of-phase (or reactive) power which measure the load on the system from the grid. The output variables are the amplitude and frequency of the voltage generated at the stator terminals. Knowledge of the transfer functions relating these variables is important in the design of control systems, particularly for load distribution and frequency.

By a change of variables, the in-phase and out-of-phase power can be replaced by the corresponding currents. Hence the variables measured were the deviations from the rated values of the in-phase and out-of phase currents and the corresponding deviations from the rated amplitude and frequency of the voltage. The currents and voltage were sampled 8 times per second,

yielding 4808 data values, and the frequency was sampled 2 times per second, yielding 1202 data values.

Linearization of the theoretical system equations [2] showed that the turbo-alternator could be represented approximately as a linear bivariate system with the in-phase and out-of-phase current deviations as inputs to the system and the corresponding voltage and frequency deviations as outputs. In the notation of Section 11.4.4, $x_1(t)$ is the in-phase current, $x_2(t)$ the out-of-phase current, $x_3(t)$ the output voltage and $x_4(t)$ the output frequency.

An analysis given in [2] on similar data from the same system showed that there was little power in the current and voltage signals above 2.5 cps and, for the frequency data, above 0.8 cps. Hence it was decided to filter the current and voltage data using a digital filter with transfer function

$$H(\mathcal{Z}) = \{\tfrac{1}{7}(\mathcal{Z}^3 + \mathcal{Z}^2 + \mathcal{Z} + 1 + \mathcal{Z}^{-1} + \mathcal{Z}^{-2} + \mathcal{Z}^{-3})\}^4,$$

which has the frequency response function

$$H(f) = \left\{\frac{\sin 7\pi f\Delta}{7 \sin \pi f\Delta}\right\}^4, \quad -\frac{1}{2\Delta} \leqslant f < \frac{1}{2\Delta}.$$

This frequency response function is low pass, with negligible transfer of power above 0.75 cps, and so the filtered data was sampled at a rate of 2 points per second.

The first 1000 values of the filtered currents and voltage and the frequency variations were then analyzed, following the stages outlined in Section 11.5.2.

(1) Preliminary decision stage. The plots of the data given in Figure 11.1 were inspected for trends and other conspicuous behavior. No trends were apparent, but it was expected that because of the large number of observations some trend removal might be required. A maximum of 125 lags was selected initially.

(2) First computation stage. Auto- and cross covariances for the original and differenced data were computed, and the correlation functions were plotted.

(3) Intermediate decision stage. The acf's $r_{11}(k)$ and $r_{44}(k)$ for the in-phase current and the frequency showed a fairly strong oscillation with a frequency of approximately 12 to 15 cps and very little trend. By contrast, the acf's $r_{22}(k)$ for the out-of-phase current and $r_{33}(k)$ for the voltage showed considerable trend and very little oscillatory behavior.

The ccf $r_{12}(k)$ is plotted in Figure 11.2, $r_{13}(k)$ and $r_{23}(k)$ in Figure 11.3 and $r_{14}(k)$ and $r_{24}(k)$ in Figure 11.4. These again confirm that trends are present and that the correlations of the differenced data should be used. Truncation points of 32, 48 and 64 lags were chosen initially since it was felt that they would be sufficient to reveal the peaks apparent from some of the correlation functions. On the basis of the peaks in the ccf's, the shift parameters were chosen to be $S_{12} = -2$, $S_{13} = 0$, $S_{23} = 2$, $S_{14} = 3$ and $S_{24} = 5$.

(4) Second computation stage. The spectral calculations described in Section 11.5.1 were performed, using the shifted correlations of the differenced data, and the spectra were plotted.

(5) Interpretation stage. Since there are so many spectral plots in this type of analysis, only the important spectra are presented here for discussion. For all spectra, the window-closing procedure showed very little change in going from $L = 32$ to $L = 64$ lags. Since the analysis was based on 1000 data points, the value $L = 32$ was finally accepted, giving 83 degrees of freedom per estimate for the autospectra.

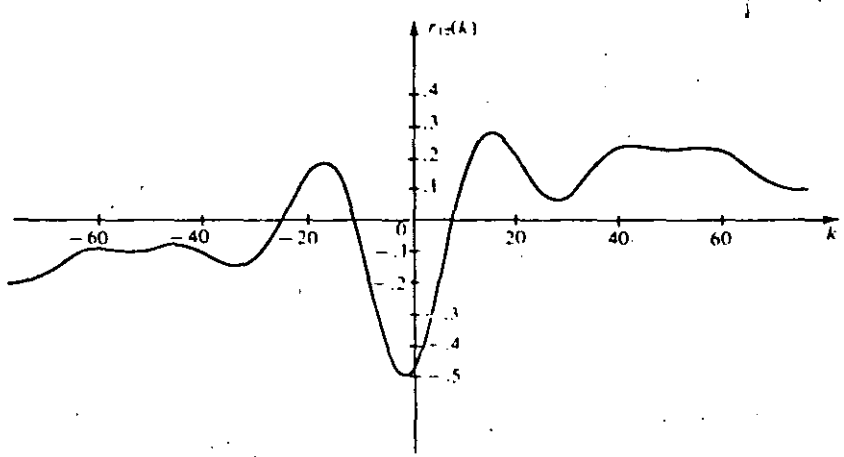


FIG. 11.2: Cross correlation function between input currents

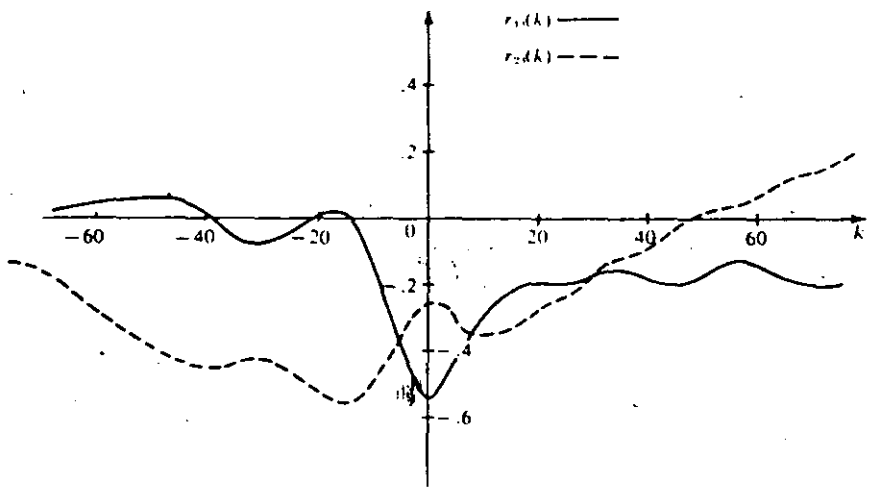


FIG. 11.3: Cross correlation functions between currents and voltage

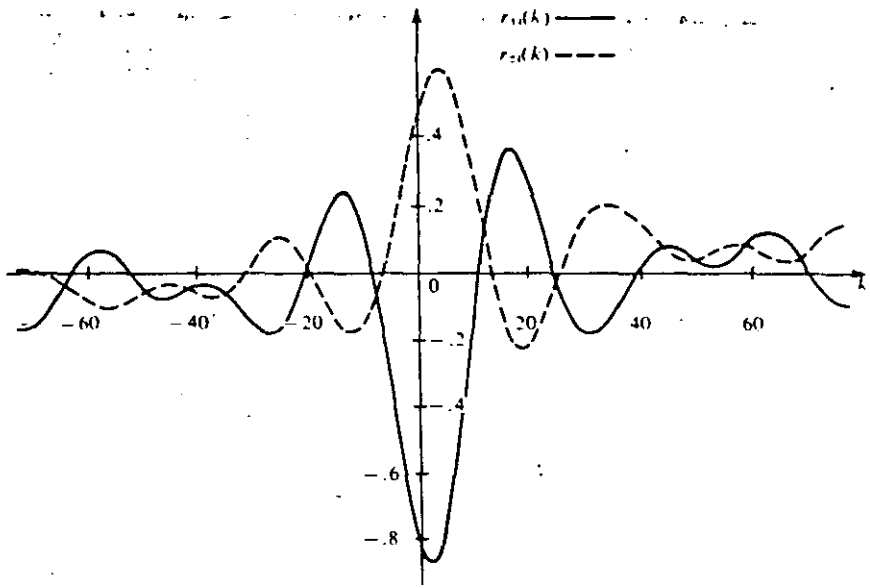


FIG. 11.4: Cross correlation functions between currents and frequency

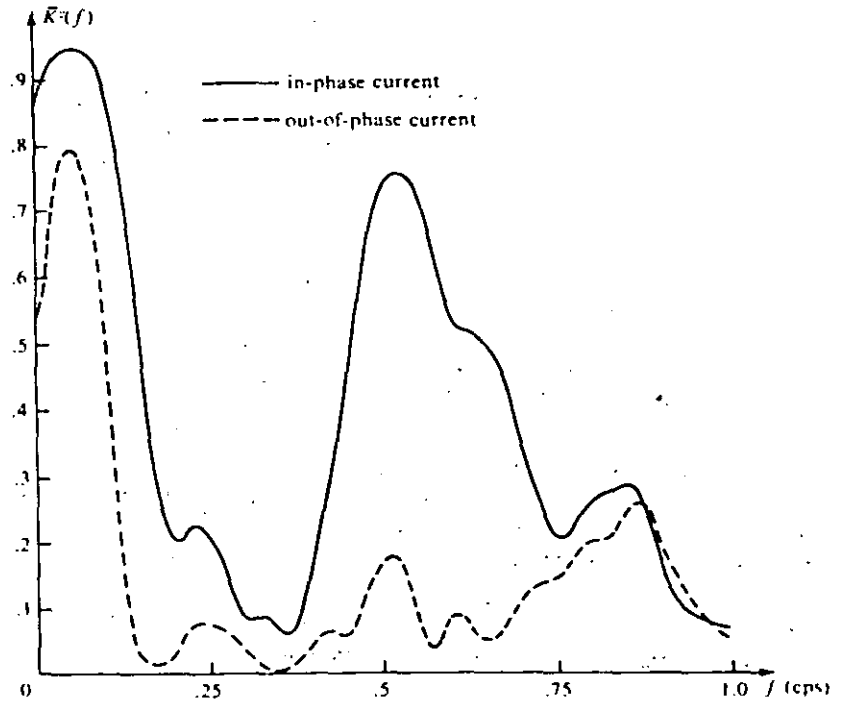


FIG. 11.5: Currents to frequency coherency spectra

Coherency spectra. The squared coherency spectra between the in-phase, out-of-phase currents and frequency are shown in Figure 11.5 and the multiple and partial squared coherency spectra between the two currents and frequency in Figure 11.6.

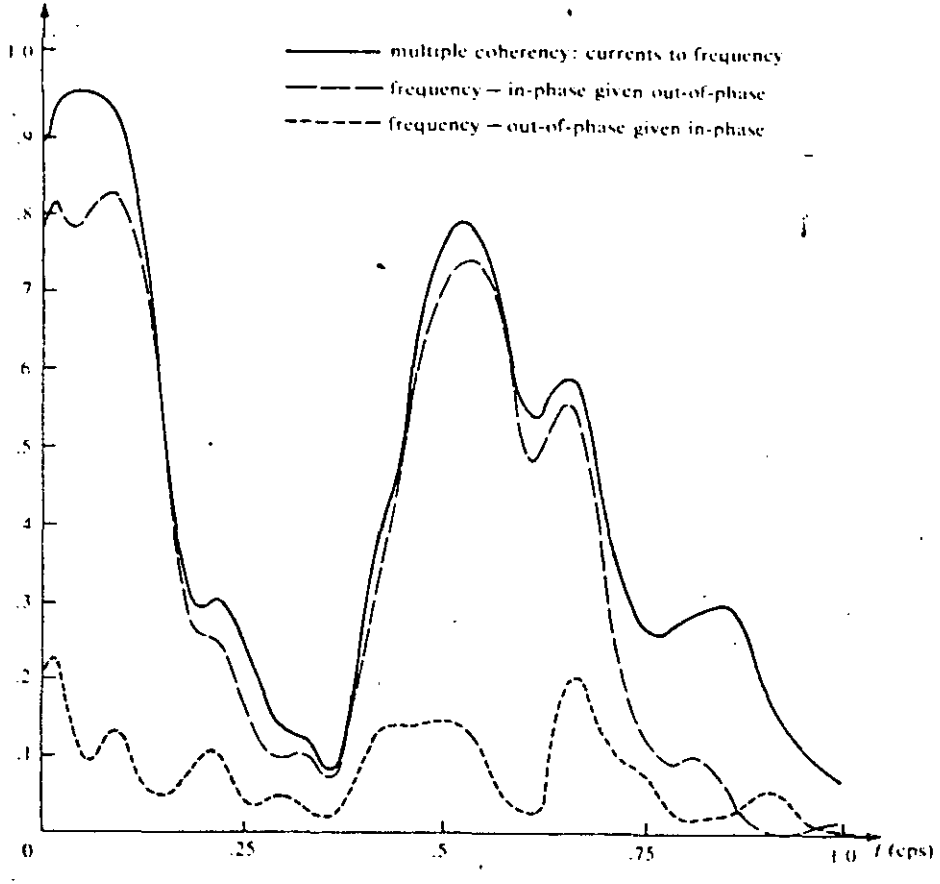


FIG. 11.6: Multiple and partial coherency spectra: currents to frequency

These figures show that the in-phase current and frequency deviations have high coherency over much of the frequency range. Figure 11.5 also suggests that there is a high coherency between the out-of-phase current and frequency. However, the very low partial coherency spectrum between frequency and out-of-phase current given the in-phase current, which is shown in Figure 11.6, indicates that this high coherency is due to the very high coherency between the currents and not to any direct relationship between frequency and out-of-phase current. Hence uninformative gain and phase estimates between the out-of-phase current and frequency variables are to be expected.

The squared coherency, multiple coherency and partial coherency spectra between the two currents and the voltages are shown in Figures 11.7 and 11.8. The in-phase to voltage coherencies are relatively high and uniform to 0.5 cps, where they drop and then begin to rise again at 0.7 cps. The extremely high coherencies above 0.75 cps are probably spurious and are discounted because of the extremely low power levels at these high frequencies. The out-of-phase to voltage coherencies are generally lower, with high coherencies near zero and in the range 0.25 to 0.5 cps. Again, the extremely high coherencies above 0.75 cps are probably spurious.

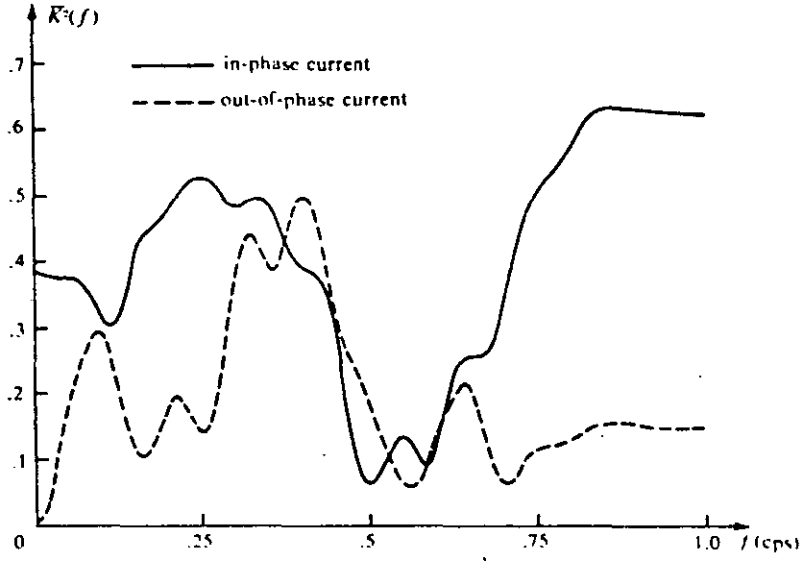


FIG. 11.7: Currents to voltage coherency spectra

Gain functions. The four gain functions are shown in Figures 11.9 to 11.12. The in-phase current to frequency gain function shown in Figure 11.9 has a peak at approximately 0.12 cps with a slope of 2 decades per decade, suggesting a second-order system. The damping factor for the system, obtained using the ratio of peak gain to dc gain, is about 0.2. The out-of-phase to frequency gain function shown in Figure 11.10 is extremely erratic, as expected from the partial squared coherency function. It is concluded that the out-of-phase current and frequency are probably not linearly related.

The gain functions from the two currents to voltage are similar, but the peaks are much less clearly defined. Thus the in-phase current to voltage gain function shown in Figure 11.11 has a low flat peak at about 0.025 cps and a slope of 2 decades per decade, suggesting a slightly oscillatory system with a break point of 0.05 cps and a damping factor of 0.6. Thr t-of-phase to

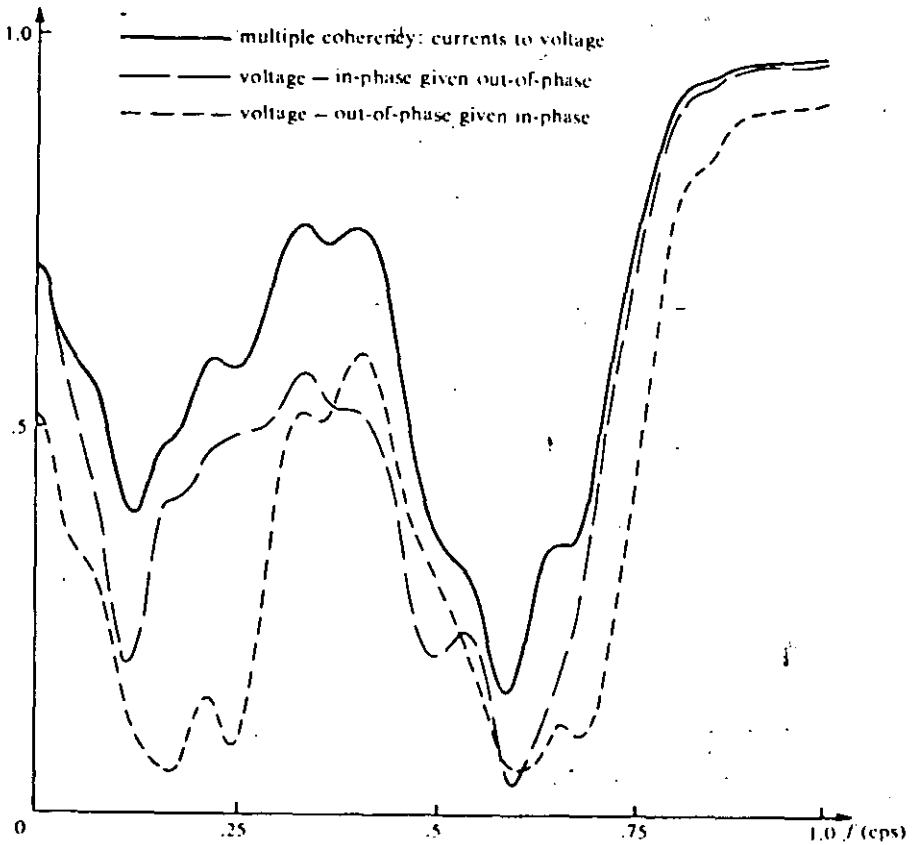


FIG. 11.8: Multiple and partial coherency spectra: currents to voltage

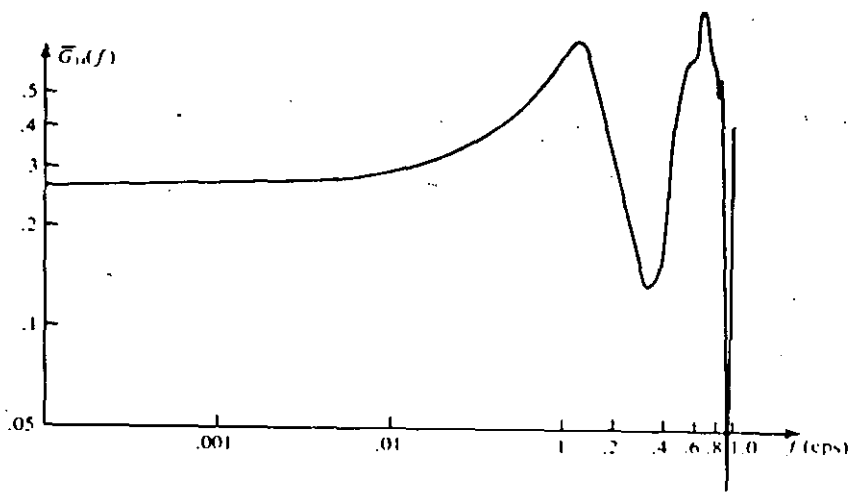


FIG. 11.9: In-phase current to frequency gain function

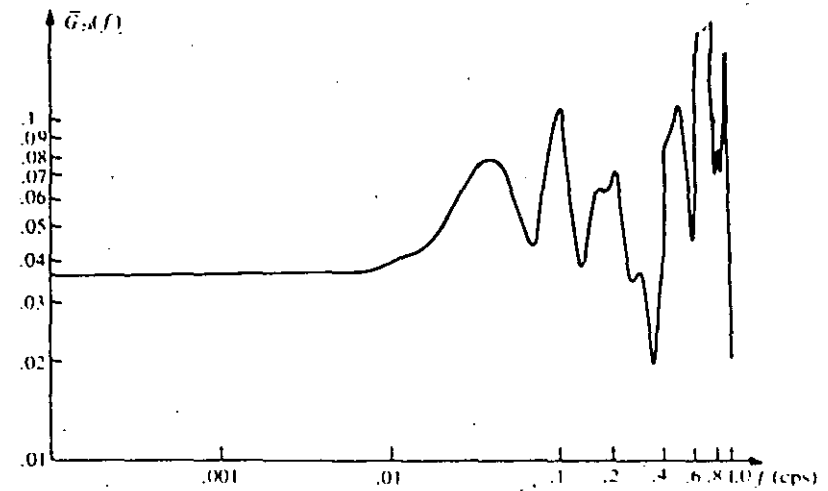


FIG. 11.10: Out-of-phase current to frequency gain function

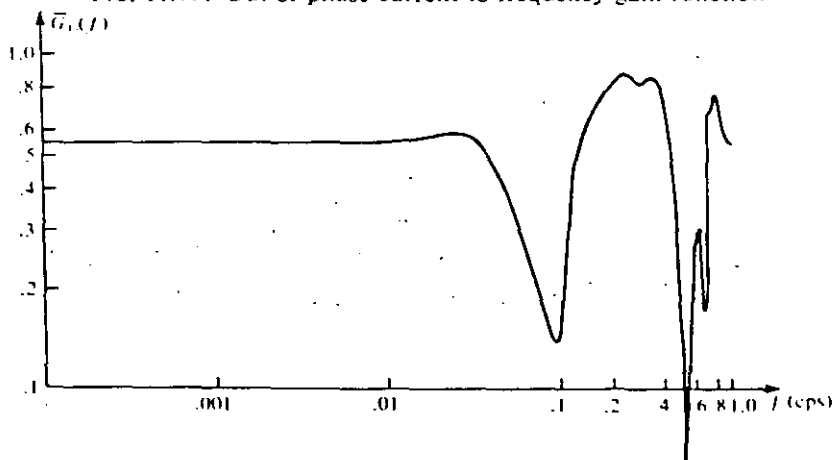


FIG. 11.11: In-phase current to voltage gain function

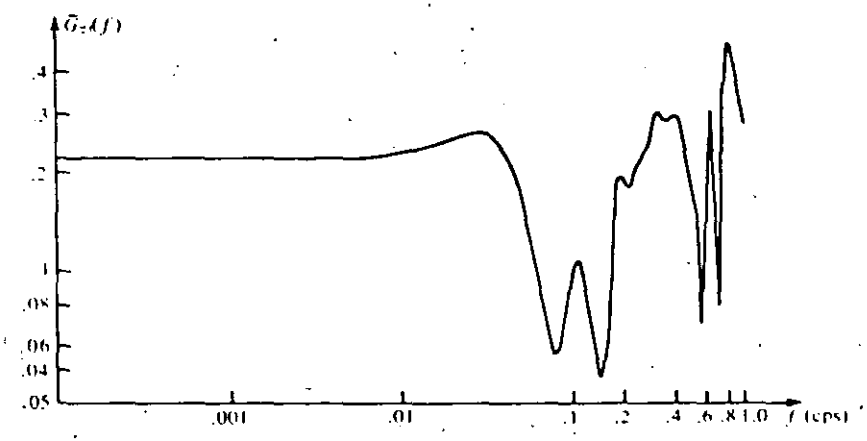


FIG. 11.12: Out-of-phase current to voltage gain function

voltage gain plot suggests a third-order system consisting of a single time constant of 20 seconds and a second-order system with break point at 0.04 cps and damping factor of 0.5. In this analysis, the gain information above 0.1 cps has been discounted because of the low power levels in the input currents above 0.1 cps.

Phase spectra. A difficulty arises in the interpretation of the phase spectra obtained from the shifted correlations in that cross spectra calculated with different shifts enter into the calculation of the phases. To avoid this difficulty, the phase functions shown in Figures 11.13 to 11.16 were based on the unaligned correlation functions. Since the phase spectra are less sensitive to bias due to non-alignment of the time series, this was not considered to be serious.

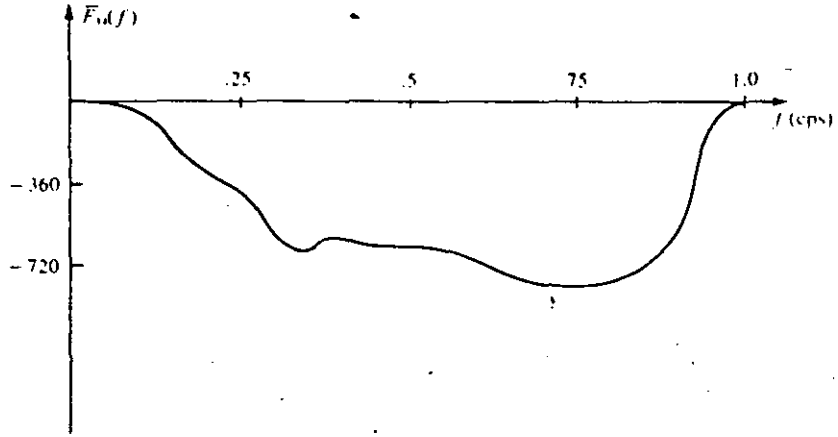


FIG. 11.13: In-phase current to frequency phase function

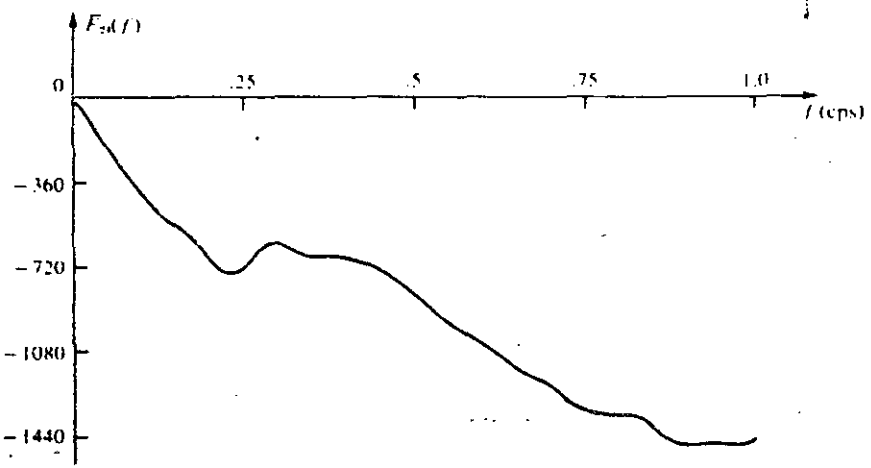


FIG. 11.14 Out-of-phase current to frequency phase function

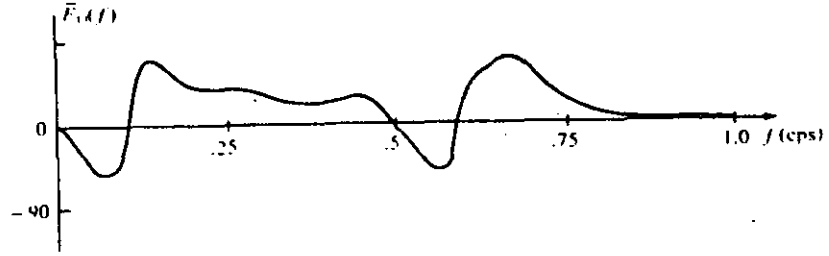


FIG. 11.15: In-phase current to voltage phase function

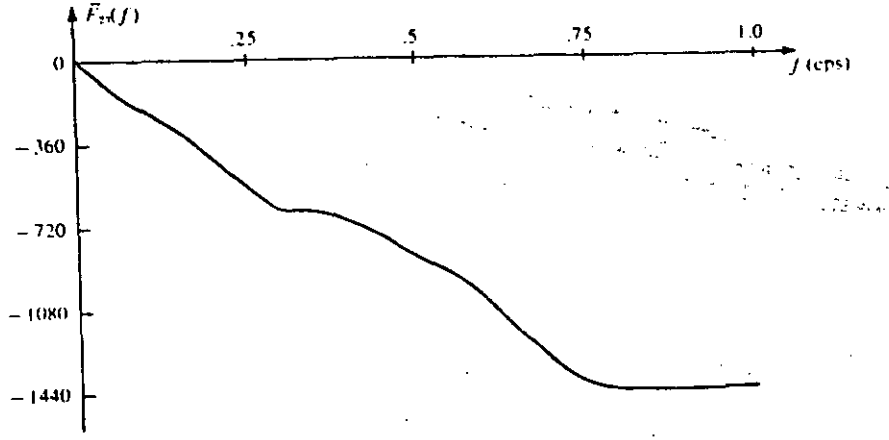


FIG. 11.16: Out-of-phase current to voltage phase function

11.5.4 Summary

The spectral analysis of the power station data may be conveniently summarized in terms of frequency response functions derived from the gain and phase plots.

In-phase current to voltage frequency response function. Using Figures 11.11 and 11.15, the frequency response function $H_{31}(f)$ is estimated to be

$$\hat{H}_{31}(f) = \frac{0.55}{1 + j48f - 1600f^2},$$

that is, a second-order system with damping factor 0.6, resonance frequency 0.025 cps, dc gain 0.55 and delay 0.

In-phase current to frequency frequency response function. From Figures 11.9 and 11.13,

$$\hat{H}_{41}(f) = \frac{0.26\, e^{-j4\pi f}}{1 + j3.3f - 70f^2},$$

that is, a second-order system with damping factor 0.2, resonance frequency 0.12 cps, dc gain 0.26 and delay 2 sampling intervals, equa sec.

Out-of-phase current to voltage frequency response function. From Figures 11.12 and 11.16

$$\hat{H}_{32}(f) = \frac{0.22\, e^{-j6\pi f}}{(1 + j125f)(1 + j25f - 625f^2)}$$

that is, a second-order system with damping factor 0.5, resonance frequency 0.04 cps, dc gain 0.22, delay 1.5 sec and a first-order system with time constant 20 sec.

It should be emphasized that the graphical estimation procedures used for the estimation of the parameters in these models are not very efficient. Therefore, they should be regarded as tentative models which should now be fitted, if necessary, by parametric methods like those illustrated in Chapter 5. However, for many control purposes, knowledge of the gain and phase plots given by Figures 11.9 to 11.16 would be sufficient.

REFERENCES

[1] T. W. Anderson, *Introduction to Multivariate Statistical Analysis.* John Wiley, New York, 1958.

[2] K. N. Stanton, "Estimation of turbo-alternator transfer functions using normal operating data." *Proc. Inst. Electr. Engrs.* **112**, (9), 1713 (1965).

[3] K. N. Stanton, "Measurement of turbo-alternator transfer functions using normal operating data." *Proc. Inst. Electr. Engrs.* **110**, (11), 2001 (1963).

APPENDIX A11.1 LATENT ROOTS AND VECTORS

The *latent roots* λ_i, $i = 1, \ldots, N$, and corresponding *right-hand latent vectors* $\mathbf{r}_i$ of a matrix $\mathbf{A}$ satisfy

$$\mathbf{A}\mathbf{r}_i = \lambda_i \mathbf{r}_i. \tag{A11.1.1}$$

The equations (A11.1.1) written in scalar form imply that

$$\begin{aligned} a_{11}r_{1i} + a_{12}r_{2i} + \cdots + a_{1N}r_{Ni} &= \lambda_i r_{1i} \\ a_{21}r_{1i} + a_{22}r_{2i} + \cdots + a_{2N}r_{Ni} &= \lambda_i r_{2i} \\ \vdots \qquad\quad \vdots \qquad\quad \vdots \qquad\quad \vdots \\ a_{N1}r_{1i} + a_{N2}r_{2i} + \cdots + a_{NN}r_{Ni} &= \lambda_i r_{Ni}. \end{aligned}$$

Geometrically (A11.1.1) means that in the N-dimensional vector space, the vectors $\mathbf{r}_i$, $i = 1, 2, \ldots, N$, are invariant under the linear transformations $\mathbf{y}_i = \mathbf{A}\mathbf{r}_i$. Related to the invariant right-hand vectors $\mathbf{r}_i$ are the *dual* or *adjoint* system of vectors $\mathbf{l}_i$ which satisfy the left-hand equations

$$\mathbf{l}_i'\mathbf{A} = \lambda_i \mathbf{l}_i'. \tag{A11.1.2}$$

It is possible to show that the left-hand vectors are orthogonal, or at right angles, to the right-hand vectors. Thus on postmultiplying (A11.1.2) by $\mathbf{r}_j$,

$$\mathbf{l}_i'\mathbf{A}\mathbf{r}_j = \lambda_i \mathbf{l}_i'\mathbf{r}_j.$$

Using (A11.1.1), this implies

$$(\lambda_i - \lambda_j)\mathbf{l}_i'\mathbf{r}_j = 0.$$

Hence, assuming that the latent roots are distinct,

$$\mathbf{l}_i'\mathbf{r}_j = 0 \quad \text{if} \quad i \neq j,$$
$$\mathbf{l}_i'\mathbf{r}_j \neq 0 \quad \text{if} \quad i = j.$$

By suitable normalization,

$$\mathbf{l}_i'\mathbf{r}_i = 1,$$

that is, the right-hand and left-hand vectors are orthonormal. If the latent vectors are assembled into matrices

$$\mathbf{L} = (\mathbf{l}_1, \mathbf{l}_2, \ldots, \mathbf{l}_N) \quad \text{and} \quad \mathbf{R} = (\mathbf{r}_1, \ldots, \mathbf{r}_N)$$

so that the columns of $\mathbf{L}(\mathbf{R})$ are the left- (right-) hand latent vectors of $\mathbf{A}$, then the above conditions imply

$$\mathbf{L}\mathbf{R} = \mathbf{I}, \tag{A11.1.3}$$

where $\mathbf{I}$ is the identity matrix consisting of ones on the diagonal and zeros elsewhere.

Symmetric matrices. The covariance matrix of a set of real rv's is real and symmetric. Hence the latent roots and vectors have the properties that the latent roots are real, the latent vectors are orthogonal and the matrices $\mathbf{L}$ and $\mathbf{R}$ are orthogonal. The second property implies that the vectors, which are invariant under transformation, are orthogonal. To demonstrate, it is convenient to assemble the equations (A11.1.1) and (A11.1.2) in matrix form

$$\mathbf{A}\mathbf{R} = \mathbf{R}\boldsymbol{\Lambda}, \tag{A11.1.4}$$

where $\boldsymbol{\Lambda}$ is the diagonal matrix of latent roots. Similarly,

$$\mathbf{L}'\mathbf{A} = \boldsymbol{\Lambda}\mathbf{L}'. \tag{A11.1.5}$$

On transposing (A11.1.5), using the fact that $\mathbf{A}' = \mathbf{A}$, $\mathbf{L}' = \mathbf{L}$,

$$\mathbf{A}\mathbf{L}' = \mathbf{L}'\boldsymbol{\Lambda}.$$

Hence on comparing with (A11.1.4)

$$\mathbf{R} = \mathbf{L}'$$

and (A11.1.3) becomes

$$\mathbf{L}\mathbf{L}' = \mathbf{I} = \mathbf{R}'\mathbf{R}, \tag{A11.1.6}$$

that is, the latent vectors are orthogonal. Hence, on premultiplying (A11.1.4) by $\mathbf{R}'$,

$$\mathbf{R}'\mathbf{A}\mathbf{R} = \mathbf{R}'\mathbf{R}\boldsymbol{\Lambda} = \boldsymbol{\Lambda}. \tag{A11.1.7}$$

The geometrical significance of (A11.1.4) when $\mathbf{A}$ is a symmetric matrix is that the quadratic form

$$\mathbf{y}'\mathbf{A}\mathbf{y} = C \tag{A11.1.8}$$

represents a central conic whose axes intersect at the origin. Under a transformation

$$\mathbf{y} = \mathbf{R}\mathbf{x}, \tag{A11.1.9}$$

(A11.1.8) becomes

$$\mathbf{x}'\mathbf{R}'\mathbf{A}\mathbf{R}\mathbf{x} = C,$$

that is,

$$\mathbf{x}'\Lambda\mathbf{x} = C$$

or

$$\lambda_1 x_1^2 + \lambda_2 x_2^2 + \cdots + \lambda_N x_N^2 = C.$$

Hence the conic is transformed to canonical form by the transformation (A11.1.9), and the latent roots give the inverse of the lengths of the principal axes of the conic.

APPENDIX A11.2 FLOW CHART FOR MULTIVARIATE FREQUENCY RESPONSE CALCULATIONS

The following is a flow chart for a computer program MULTSPEC which accepts covariance estimates for two input series $X1(t)$, $X2(t)$ and two output series $X3(t)$, $X4(t)$ and computes the frequency response functions $H31$, $H32$, $H41$, $H42$ relating them. Printer output of important functions is provided, but the most important output is via the plotter. Plotted output includes all autospectra (logspec versus frequency), both residual spectra (logspec versus frequency), all total and partial squared coherencies, and all gain and phase plots, with overlays for additional truncation points.

Program MULTSPEC

1) *Input* N, MAXM, NF, DELTA, M.
2) *Input* COV(K,I,I), K = 0, MAXM, I = 1,2.
3) *Calculate* autospectra, SPEC(K,I), K = 0, NF, I = 1,2 using AUTOSPEC subroutine (Appendix A7.1).
4) *Read* COV(K,3,3), K = 0, MAXM.
5) *Calculate* SPEC(K,3) using AUTOSPEC subroutine.
6) *Read* COV(K,1,2), COV(K,2,1), K = 0, MAXM, shift S12
   COV(K,1,3), COV(K,3,1), K = 0, MAXM, shift S13
   COV(K,2,3), COV(K,3,2), K = 0, MAXM, shift S23.
7) *Use* EVOD subroutine (Appendix A9.2) to calculate even and odd parts.
8) *Call* CROSSPEC subroutine (Appendix A9.2) to calculate
   COSPEC(K,1,2), QSPEC(K,1,2), SQ(K,1,2)
   COSPEC(K,1,3), QSPEC(K,1,3), SQ(K,1,3)
   COSPEC(K,2,3), QSPEC(K,2,3), SQ(K,2,3).
9) *Calculate* DENOM(K) = SPEC(K,1)*SPEC(K,2) − SQ(K,1,2).
10) *Call* subroutine ENDALL.
11) *Store* all quantities in ENDALL which are marked with a marginal asterisk, for future plotting.

12) *Read* COV(K,4,4), K = 0, MAXM.
13) *Calculate* SPEC(K,4) using AUTOSPEC subroutine
14) *Read* COV(K,1,4), COV(K,4,1), K = 0, MAXM, shift S14
    COV(K,2,4), COV(K,4,2), K = 0, MAXM, shift S24.

Repeat steps 7 through 11. Then read another value of M, and use the stored autocovariances and even and odd parts of cross covariances in the subroutines AUTOSPEC, CROSPEC, and ENDALL to calculate the gains, phases, squared coherencies and residual spectra. When the desired values of M have been used, plot and overlay all logspectra and log residual spectra versus frequency, all phases and squared coherencies versus frequency, and all log gains versus log frequency.

Subroutine ENDALL

This subroutine calculates gain, phase, total and partial squared coherencies, and residual spectra. Dropping the index K, the calculations are as follows:

A31 = COSPEC(1,3)*SPEC(2) + QSPEC(2,3)*QSPEC(1,2) − COSPEC(2,3)*
COSPEC(1,2)

SQA31 = A31*A31

B31 = QSPEC(1,3)*SPEC(2) − QSPEC(2,3)*COSPEC(1,2) − COSPEC(2,3)*
QSPEC(1,2)

SQB31 = B31*B31

* PHASE 31 = ARCTAN (− B31/A31)

GAIN 31 = SQRT (SQA31 + SQB31)/DENOM

* LOGGN31 = LOG10(GAIN 31)

A32 = COSPEC(2,3)*SPEC(1) − COSPEC(1,2)*COSPEC(1,3) − QSPEC(1,2)*
QSPEC(1,3)

SQA32 = A32*A32

B32 = QSPEC(2,3)*SPEC(1) − COSPEC(1,2)*QSPEC(1,3) + COSPEC(1,2)*
QSPEC(1,3)

SQB32 = B32*B32

* PHASE 32 = ARCTAN (− B32/A32)

GAIN 32 = SQRT (SQA32 + SQB32)/DENOM

* LOGGN32 = LOG10(GAIN 32)

* COHSQ(1,3) = SQ(1,3)/(SPEC(1)*SPEC(3))

* COHSQ(2,3) = SQ(2,3)/(SPEC(2)*SPEC(3))

```
      R = COSPEC(1,2)*COSPEC(2,3)*COSPEC(1,3)
          + COSPEC(1,2)*QSPEC(2,3)*QSPEC(1,3)
          - QSPEC(1,2)*QSPEC(2,3)*COSPEC(1,3)
          + QSPEC(1,2)*QSPEC(1,3)*COSPEC(2,3)
*     COHSQ
          = (SPEC(2)*SQ(1,3) + SPEC(1)*SQ(2,3) - 2.*R)/(SPEC(3)*DENOM)
      COMP = 1 - COHSQ
      RESID = COMP*SPEC(3)
*     LOGRESID = LOG10(RESID)
*     COH 132 = 1 - (COMP/(1 - COHSQ(2,3)))
*     COH 231 = 1 - (COMP/(1 - COHSQ(1,3))).
```

APPENDIX A11.3 DATA FOR POWER STATION EXAMPLE

TABLE A11.1: 100 values of coded power station data

0.92	0.95	0.99	1.07	1.13	1.16	1.15	1.08	0.97
0.89	0.85	0.76	0.65	0.59	0.59	0.58	0.58	0.60
0.60	0.55	0.54	0.60	0.62	0.54	0.50	0.58	0.65
0.65	0.69	0.74	0.76	0.72	0.69	0.71	0.81	0.96
1.09	1.10	1.05	1.04	1.05	1.03	1.00	0.92	0.81
0.73	0.60	0.48	0.45	0.44	0.37	0.30	0.27	0.30
0.33	0.33	0.34	0.36	0.38	0.41	0.46	0.50	0.50
0.47	0.44	0.47	0.52	0.52	0.50	0.50	0.49	0.48
0.47	0.46	0.43	0.42	0.46	0.52	0.58	0.61	0.66
0.72	0.72	0.68	0.70	0.77	0.75	0.64	0.56	0.50
0.40	0.28	0.24	0.26	0.21	0.16	0.18	0.19	0.18
0.19								

In-phase current deviations

−0.75	−0.92	−1.04	−1.01	−0.83	−0.64	−0.55	−0.46	−0.31
−0.14	−0.02	0.12	0.32	0.54	0.68	0.72	0.77	0.82
0.87	0.91	0.98	1.10	1.17	1.09	1.00	1.06	1.17
1.18	0.96	0.64	0.60	0.80	0.84	0.68	0.63	0.74
0.81	0.67	0.38	0.17	0.17	0.29	0.41	0.52	0.67
0.77	0.80	0.86	1.01	1.22	1.40	1.49	1.49	1.47
1.54	1.64	1.60	1.46	1.41	1.41	1.39	1.26	1.04
0.89	0.99	1.21	1.31	1.25	1.13	1.04	0.98	0.93
0.86	0.79	0.80	0.91	1.10	1.26	1.26	1.10	0.92
0.79	0.67	0.53	0.44	0.43	0.39	0.32	0.29	0.34
0.38	0.37	0.33	0.26	0.25	0.31	0.40	0.47	0.45
0.36								

Out-of-phase current deviations

(continued)

TABLE A.11.1—*continued*

−2.11	−2.08	−2.10	−2.22	−2.39	−2.50	−2.49	−2.43	−2.37
−2.30	−2.25	−2.21	−2.20	−2.22	−2.21	−2.20	−2.20	−2.22
−2.22	−2.21	−2.22	−2.25	−2.26	−2.20	−2.15	−2.18	−2.21
−2.25	−2.27	−2.28	−2.28	−2.26	−2.19	−2.13	−2.15	−2.25
−2.33	−2.32	−2.25	−2.21	−2.23	−2.28	−2.32	−2.33	−2.34
−2.32	−2.26	−2.22	−2.23	−2.27	−2.30	−2.30	−2.28	−2.26
−2.26	−2.25	−2.23	−2.21	−2.23	−2.25	−2.26	−2.22	−2.14
−2.10	−2.15	−2.23	−2.27	−2.26	−2.22	−2.20	−2.19	−2.17
−2.14	−2.10	−2.09	−2.14	−2.24	−2.32	−2.34	−2.30	−2.28
−2.28	−2.26	−2.22	−2.20	−2.20	−2.17	−2.11	−2.07	−2.05
−1.98	−1.90	−1.83	−1.78	−1.77	−1.78	−1.81	−1.83	−1.82
−1.77								

Voltage deviations

4.75	4.75	4.74	4.75	4.75	4.77	4.78	4.76	4.81
4.81	4.84	4.88	4.94	4.96	4.99	5.05	5.11	5.11
5.10	5.10	5.05	5.00	5.01	5.00	5.00	5.01	5.04
5.09	5.09	5.06	5.06	5.04	5.01	4.99	4.97	4.99
5.00	5.01	4.97	4.88	4.81	4.81	4.79	4.79	4.85
4.88	4.90	5.00	5.11	5.15	5.24	5.20	5.20	5.20
5.20	5.21	5.20	5.15	5.15	5.14	5.10	5.11	5.10
5.10	5.09	5.11	5.08	5.10	5.10	5.15	5.15	5.15
5.10	5.15	5.10	5.10	5.10	5.11	5.11	5.10	5.11
5.11	5.08	5.05	5.01	4.90	4.90	4.95	4.94	4.90
4.96	4.97	4.99	5.08	5.16	5.21	5.19	5.19	5.20
5.16								

Frequency deviations

INGENIERIA MARITIMA
DEL 1o. DE JULIO AL 30 DE NOVIEMBRE
1985
MEXICO, D.F

INTRODUCCION A LOS MODELOS HIDRAULICOS

DR. GABRIEL ECHAVEZ ALDAPE
ING. VICTOR BURGUET.

Introducción a los MODELOS HIDRAULICOS de fondo fijo

DEPARTAMENTO DE ESTUDIOS EXPERIMENTALES
COMISION FEDERAL DE ELECTRICIDAD

GABRIEL ECHAVEZ ALDAPE

Mayo 1982

I N D I C E

V. MODELOS CON FRONTERAS RIGIDAS

5.1 Domina la geometría

5.2 Flujo laminar e hidráulicamente liso

5.3 Flujo de transición e hidráulicamente rugoso

VI. MODELOS CON SUPERFICIE LIBRE

6.1 Domina la geometría

6.2 Flujo laminar e hidráulicamente liso

6.3 Flujo de transición e hidráulicamente rugoso

VII. CONSTRUCCION Y OPERACION

7.1 Componentes básicos

7.2 Modelo. Selección y Construcción

7.3 Operación del modelo

VIII. MEDICIONES

8.1 Sistema de medición

8.2 Funcionamiento de un sistema de medición

8.3 Instrumento de orden cero

8.4 Instrumento de primer orden

8.5 Instrumento de segundo orden

8.6 Aplicación a problemas de medición

8.7 Mediciones de tirante, velocidad y gasto

8.7.1 Medición de tirante

8.7.2 Medición de velocidad

8.7.3 Medición de gasto

I. INTRODUCCION

Si bien la experimentación es un arte que solo se puede dominar practicándola, es indispensable entender los principios básicos y conocer una serie de recomendaciones derivadas de la experiencia.

En este manual, orientado a Ingenieros, se presentan los conceptos básicos para hacer experimentación en hidráulica y se dan criterios que ayudan a la solución de problemas en modelos hidráulicos de fondo fijo. Como los resultados dependerán no solo de lo acertado del modelo sino también de la calidad de las mediciones, se complementa el tema anterior con la inclusión de teoría de mediciones y manejo de datos. El primer tema, teoría de modelos, se desarrolla en el aspecto básico en los capítulos 2 y 3, y en la parte aplicada, en los capítulos 4, 5 y 6. El tema de mediciones se presenta en los capítulos 7, 8 y 9. En el último capítulo, 10, se muestran varios problemas que sirvan de ejemplo ilustrativo y en el apéndice se incluyen gráficas y tablas con información necesaria al trabajar con modelos hidráulicos.

II. ANALISIS DIMENSIONAL

1. Objetivo. El Análisis Dimensional es una técnica que da información utilizando las dimensiones de las variables que intervienen en un problema.

En Ingeniería se usa frecuentemente en la verificación de fórmulas, comprobando que las unidades de los dos lados de la ecuación sean iguales y en la transformación de fórmulas de un sistema de unidades a otro, usualmente del inglés al métrico.

Otro uso, que es el que se verá con detalle más adelante, permite reducir el número de variables de un problema, simplificándolo en forma importante.

Antes de iniciar el estudio del Análisis Dimensional es conveniente aclarar algunos de los conceptos que se utilizan.

2. Antecedentes. Cuando una ecuación describe un fenómeno físico, las variables, ligadas entre sí por símbolos matemáticos y constantes de tal manera que se logre una igualdad, representan las características o propiedades del fenómeno.

Cada variable además de tener un valor numérico tiene una dimensión, o sea una combinación de unidades ~~llamadas~~ de referencia, que denota la forma en que se definió la variable.

2.1. Unidades de referencia. Las unidades de referencia usuales son: M, masa; L, longitud y T, tiempo; o en forma equivalente F, fuerza; L y T. Donde F y M están ligadas por la segunda ley de Newton

$$[F] = [M\ L \cdot T^{-2}]$$

y los paréntesis rectangulares indican que se trata de una igualdad entre dimensiones.

Conviene darse cuenta que la selección de estas unidades de referencia es arbitraria y aunque probablemente son las unidades más fáciles de medir, podrían haberse escogido otras variables como unidades, por ejemplo: E, energía; V, velocidad; T, tiempo. El único requisito es que el grupo de unidades de referencia sea completo e independiente (ref 1).

En problemas donde intervenga la temperatura o variables eléctricas, éstas se reducen a combinaciones de F, L y T o se introducen más unidades de referencia, como: θ, temperatura, Q, carga eléctrica, etc. (ver Apéndice).

En la Tabla 1 se presentan las variables más comunes y sus dimensiones. los exponentes a los que hay que elevar MLT o FLT para obtener sus dimensiones

TABLA 1

Variables más comunes y sus dimensiones

	M	L	T	F	L	T
Masa	1	0	0	1	-1	2
Longitud	0	1	0	0	1	0
Tiempo	0	0	1	0	0	1
Area	0	2	0	0	2	0
Volumen	0	3	0	0	3	0
Angulo	0	0	0	0	0	0
Velocidad	0	1	-1	0	1	-1
Gasto	0	3	-1	0	3	-1
Aceleración	0	1	-2	0	1	-2
Densidad	1	-3	0	1	-4	2
Peso específico	1	-2	-2	1	-3	-2
Fuerza	1	1	-2	1	0	0
Presión	1	-1	-2	1	-2	0
Par	1	2	-2	1	1	0
Energía, Trabajo	1	2	-2	1	1	0
Potencia	1	2	-3	1	1	-1
Cantidad de movimiento	1	1	-1	1	0	1
Impulso	1	1	-1	1	0	1
Tensión superficial	1	0	-2	1	-1	0
Viscosidad dinámica	1	-1	-1	1	-2	1
Viscosidad cinemática	0	2	-1	0	2	-1
Módulo de elasticidad	1	-1	-2	1	-2	0
Compresibilidad	-1	1	2	-1	2	0
Frecuencia	0	0	-1	0	0	-1

2.2. Variables dependientes e independientes. En experimentación las variables se clasifican en dos tipos: variables dependientes e independientes. La variable dependiente es la que interesa determinar; en hidráulica es común que ésta sea la velocidad o la presión, o alguna variable derivada de estas dos, como: gasto, empuje, sustentación, etc.

La variable independiente es la que junto con otras interviene en un problema y en función de las cuales interesa determinar la variable dependiente. En hidráulica generalmente las variables que describen la geometría del flujo, como dimensiones de un ducto cerrado o variables que representan propiedades del fluido, como: densidad, viscosidad, etc.

Si en un cierto problema o experimento una variable se conserva constante se le llama parámetro.

Así, en la fórmula para obtener la velocidad, V, del agua en un orificio a una profundidad h.

$$V = \sqrt{2gh}$$

donde g aceleración de la gravedad.

V es la variable dependiente

g parámetro

h variable independiente

En forma general, la ecuación anterior se puede representar como

$$V = f\ (g,\ h)$$

También hay que observar que el aceptar que una ecuación representa un fenómeno físico implica restricciones en esa ecuación; así, además de que se debe satisfacer la igualdad numérica se tiene que:

a) Las dimensiones de los términos deben ser iguales, o sea debe ser dimensionalmente homogénea.

b) La ecuación sólo es válida para cierto rango de las variables.

c) No se pueden hacer todas las operaciones matemáticas con ella, por ejemplo: si V, velocidad; s, desplazamiento; t, tiempo y

$$V = \frac{s}{t}$$

la operación

$$\log V = \log s - \log t$$

es inválida. Para poder efectuar este tipo de operaciones, habría que hacer adimensionales a las variables.

Una consecuencia de la limitación a) es que la ecuación será independiente del sistema de unidades en que se exprese.

2.3. Matriz de los exponentes. Si se ponen en un renglón las variables que intervengan en un problema y en una columna sus dimensiones, se podrá hacer la matriz de los exponentes escribiendo en el cruce respectivo el exponente de la unidad correspondiente; así, si F, fuerza, V, velocidad, ρ, densidad, y ℓ, longitud; la matriz de los exponentes respectiva, ver Tabla 1, es:

	F	V	ρ	ℓ
M	1	0	1	0
L	1	1	-3	1
T	-2	-1	0	0

El rango de esta matriz será el orden de la mayor matriz cuadrada cuyo determinante sea no nulo (ref. 1). Para el caso anterior el rango será

$$r = 3$$

pues el determinante de la matriz de orden 3, formada por ejemplo con las tres primeras columnas, es diferente de cero.

El teorema básico que permite reducir el número de variables de un problema es el teorema de Buckingham-Vaschy o teorema Π que se enunciará a continuación y cuya demostración se puede ver en las refs 3 ó 4.

3. Teorema de Buckingham-Vaschy. Una ecuación dimensionalmente homogénea en donde intervengan m variables y cuyo rango de la matriz de los exponentes sea r, podrá expresarse como una combinación de m-r nuevas variables adimensionales o productos adimensionales, esto es en cada uno de ellos las unidades fundamentales se combinan de tal manera que se simplifican las del numerador con las del denominador.

Aunque en la mayoría de los problemas r es igual al número de unidades de referencia, es conveniente comprobarlo construyendo la matriz de exponentes y determinando r.

Otra forma de presentar el teorema es:

Si se tiene una función que relaciona m variables Xi

$$f\ (X_1, X_2, \ldots\ldots X_m) = 0 \qquad (1)$$

y r es el rango de la matriz de los exponentes de las Xi; existirá otra función

$$f\ (\Pi_1, \Pi_2, \ldots\ldots \Pi_{m-r}) = 0 \qquad (2)$$

donde las Πi serán productos adimensionales.

Esta reducción de variables no solo tiene la ventaja de simplificar grandemente al problema sino que, frecuentemente, da

naturaleza de los productos adimensionales da información útil acerca del comportamiento del fenómeno.

Para hacer esta simplificación, o sea pasar de la ecuación (1) a la (2), se hace lo siguiente:

1) Se forma la matriz de los exponentes y se determina su rango.

(Como ejemplo se proseguirá con una función en donde intervenga fuerza, velocidad, densidad y longitud, o sea m = 4, cuya matriz de los exponentes está en el inciso anterior, y cuyo rango es 3; por lo que habrá una función equivalente con m - r = 1 variable adimensional).

2) Se construye un sistema de ecuaciones utilizando como coeficiente los elementos de la matriz, así:

$$X_1 + X_3 = 0$$

$$X_1 + X_2 - 3X_3 + X_4 = 0$$

$$-2X_1 - X_2 = 0$$

Obsérvese que X_1 representa el exponente de F, X_2 el de V, X_3 el de ρ y X_4 el de ℓ: y que cada unidad de referencia generó una ecuación.

3) Se encuentran m - r soluciones a este sistema. Cada solución mostrará los exponentes a los que se deben elevar las variables originales (V, F, ρ, ℓ) para obtener un producto adimensional.

Como el sistema es homogéneo, sin términos independientes, y habrá además, en general, más incógnitas que ecuaciones, hay que asignar valores a las variables redundantes por lo que es necesario seguir ciertos criterios que faciliten el encontrar productos adimensionales significativos.

4. Recomendaciones para formar productos adimensionales.

4.1 Primera regla. Una regla es hacer que la variable dependiente aparezca en el numerador de un solo producto adimensional, de preferencia con el exponente 1, y no aparezca en ningún otro producto. Esto se logra haciendo el coeficiente respectivo igual a la unidad, para encontrar un producto adimensional, y después cero en todas las otras soluciones.

Una excepción importante a esta recomendación es la velocidad que, debido al papel central que juega en los problemas, aparece frecuentemente en la variable dependiente y en las independientes.

Para el caso que se está usando como ejemplo y aceptando que F es la variable dependiente, deberá hacerse $X_1 = 1$ en una ocasión y en todas las demás igual a cero.

O sea, para $X_1 = 1$, el sistema se reducirá a:

$$X_3 = -1$$

$$X_2 - 3X_3 + X_4 = -1$$

$$X_2 = -2$$

que tiene una solución inmediata

$$X_1 = 1$$

$$X_2 = -2$$

$$X_3 = -1$$

$$X_4 = -2$$

por lo que el producto buscado será:

$$\Pi_1 = \frac{F}{\rho V^2 \ell^2}$$

el cual es adimensional.

4.2 Segunda Regla. Una segunda recomendación es tratar de formar productos adimensionales ya conocidos, o estándar, como por ejemplo: Número de Reynolds, de Froude, etc. En la Tabla 2 se muestran los productos adimensionales comunes en donde intervienen las variables dependientes y en la Tabla 3 lo mismo para variables independientes.

Aplicando esta regla al ejemplo que se ha estado desarrollando, se puede ver que Π_1 es proporcional al número de Euler, pues

$$Eu = \frac{V}{\sqrt{2\Delta p/\rho}}$$

que es igual a $\Pi_1^{-1/2}$ si se exceptúa la constante $2^{-1/2}$, por lo que también se podría haber obtenido directamente sin necesidad de resolver el sistema de ecuaciones.

TABLA 2

Productos adimensionales con variables dependientes

Nombre	Símbolo	Definición
Número de Euler	$\mathbb{E}$	$V/\sqrt{2\Delta p/\rho}$; $\frac{\Delta h}{v^2/2g}$
Número de Strouhal	$\mathbb{S}$	$V/n\mathcal{D}$
Número de Thoma	$\mathbb{T}$	$(h-hv)/(V^2/2g)$
Coeficiente de presión	C_p	$1/\mathbb{E}^2$
Coeficiente de arrastre	C_A	$F_A/\frac{1}{2}\rho V^2 A$
Coeficiente de sustentación	C_s	$F_s/\frac{1}{2}\rho V^2 A$
Coeficiente de cortante	C_τ	$\tau/\frac{1}{2}\rho V^2$

A	área
F_A	fuerza de arrastre
F_S	fuerza de sustentación
V	velocidad
$\mathcal{D}$	diámetro, longitud
h, Δh	carga de presión
hv	carga de vaporización
n	frecuencia
Δp	incremento de presión
ρ	densidad
τ	esfuerzo cortante

TABLA 3

Productos adimensionales con variables independientes

Nombre	Símbolo	Definición
Número de Reynolds	$\mathbb{R}$	$V D \rho / \mu$
Número de Froude	$\mathbb{F}$	$V / \sqrt{gh}$
Número de Mach	$\mathbb{M}$	$V / \sqrt{E/\rho}$
Número de Weber	$\mathbb{W}$	$V / \sqrt{\sigma / \rho \ell}$
Número de von Kármán	$\mathbb{K}$	$\frac{k\sqrt{g R S}}{\nu} = \frac{v_* k}{\nu}$

E	módulo de elasticidad		k	rugosidad (longitud)	
R	radio hidráulico	→	$\mathcal{L}$	longitud	
S	pendiente		μ	viscosidad dinámica	
V	velocidad		ν	viscosidad cinemática	
D	diámetro o longitud característica		ρ	densidad	
g	gravedad		σ	tensión superficial	
h	tirante				

4.3 Tercera regla. La tercera recomendación es que las variables que describen la geometría del problema, y cuya dimensión es una longitud, se pueden hacer adimensionales simplemente seleccionando una longitud como característica y dividiendo cada variable entre esa longitud. Por ejemplo: en un problema donde intervengan ℓ, x, y, dos productos adimensionales serán

$$\frac{x}{\ell} \quad , \quad \frac{y}{\ell}$$

En forma similar es fácil construir números del tipo

$$\frac{\rho_1}{\rho_2} \quad , \quad \frac{\gamma_1}{\gamma_2} \quad , \quad \frac{T_1}{T_2}$$

cuando aparecen variables con las mismas dimensiones.

Finalmente, en los casos en que después de aplicar los tres criterios anteriores aún quedan productos por determinar, será necesario asignar valores a los otros coeficientes para obtener las soluciones faltantes, observando de preferencia las indicaciones siguientes:

Las variables que representan las propiedades del flujo, exceptuando a la densidad, como: viscosidad, módulo de elasticidad, etc. deben aparecer solas en un sólo producto adimensional.

Así, un producto en donde aparezcan viscosidad y módulo de elasticidad, o viscosidad y tensión superficial debe evitarse. Igualmente sólo debe haber un producto adimensional en donde esté la viscosidad, tensión superficial, etc.

Hay que buscar productos no solo correctos sino también que satisfagan ciertos criterios de sencillez y naturalidad.

Aunque

$$\frac{F^{1/3}}{\rho^{1/3} \; V^{2/3} \; \ell^{2/3}}$$

también es solución del sistema de ecuaciones anterior, no se recomienda pues es evidente que se puede simplificar. En casos en donde intervengan más variables, o más bien donde m − r sea

mayor que 1, aunque los productos de los números también sean solución, son poco naturales y por lo tanto deben preferirse otros.

No deben aparecer demasiadas variables en un producto adimensional. Generalmente intervienen 3 ó 4, aunque hay algunos con más de 4.

En los productos adimensionales deben estar todas las variables que aparecen en el problema. Si falta alguna obviamente la solución está equivocada.

5. Ejemplos. A continuación se resolverán dos problemas que servirán como aplicación de los conceptos anteriores.

Ejemplo 1. Pérdida de energía en una tubería. En un tramo ℓ de tubería de diámetro D, fig. 1, con una rugosidad equivalente k por la que escurre un fluido de viscosidad dinámica μ y densidad ρ, a una velocidad V; la pérdida de carga Δh, depende de:

$$\Delta h = f\ (V,\ D,\ \ell,\ k,\ \mu,\ \rho,\ g) \qquad (3)$$

donde g es la aceleración de la gravedad.

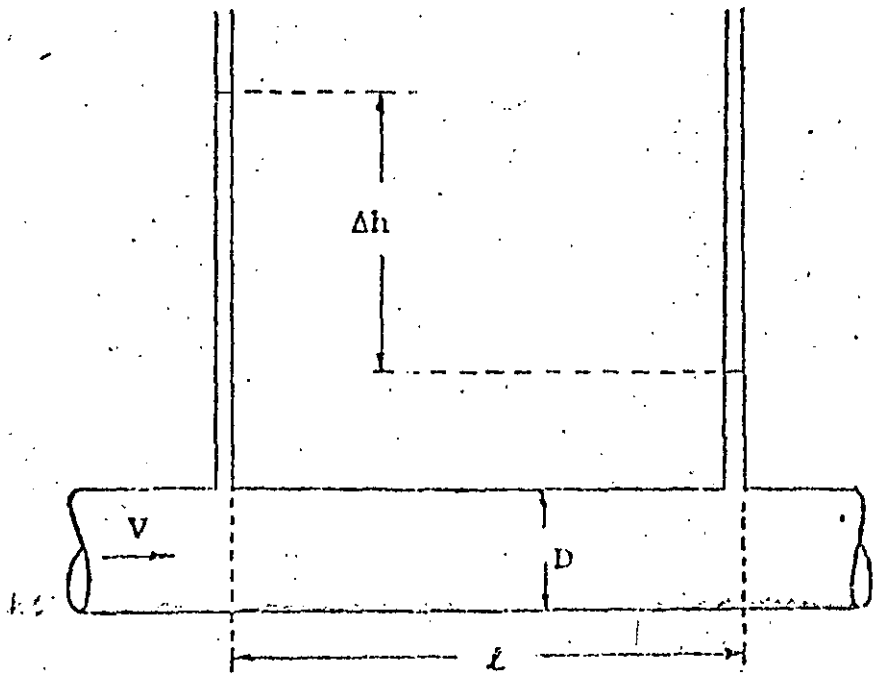


Fig 1. Pérdida de carga en una tubería

$$c' = \frac{c}{c_0}, \quad t' = \frac{t}{t_0}, \quad x' = \frac{x}{x_0} \quad y \quad D' = \frac{D}{D_0}$$

Ahora se sustituyen estas nuevas variables en la ecuación (1), lo que se hará término por término con fines explicativos. Como c' y x' son función únicamente de t y x, respectivamente, se tendrá que:

$$\frac{\partial c}{\partial t} = \frac{\partial c}{\partial t'} \frac{\partial t'}{\partial t} \quad \text{o sea} \quad \frac{\partial c}{\partial t'} = c_0 \frac{\partial c'}{\partial t'} \quad y \quad \frac{\partial t'}{\partial t} = \frac{1}{t_0}$$

$$\frac{\partial c}{\partial t'} = \frac{\partial c}{\partial t} \frac{\partial t}{\partial t'} \qquad \frac{\partial t}{\partial t'} = t_0$$

queda por lo que

$$\frac{\partial c}{\partial t} = \frac{c_0}{t_0} \frac{\partial c'}{\partial t'}$$

En forma similar, el segundo término

$$\rightarrow \quad D \frac{\partial^2 c}{\partial x^2} = D_0 D' \frac{\partial}{\partial x} \left(\frac{\partial c}{\partial x'} \frac{\partial x'}{\partial x} \right) \qquad \frac{dc}{dx'} = c_0 \frac{dc'}{dx'}$$

$$\frac{\partial x'}{\partial x} = \frac{1}{x_0}, \; \partial x'$$

$$\frac{\partial c}{\partial t} = D \frac{\partial^2 c}{\partial x^2}$$

$$= D_0 D' \frac{\partial}{\partial x} \left(\frac{c_0}{x_0} \frac{\partial c'}{\partial x'} \right) \qquad \frac{\partial x'}{\partial x} = \frac{1}{x_0} \frac{\partial}{\partial x}$$

$$= \frac{D_0 c_0}{x_0} D' \frac{\partial}{\partial x} \left(\frac{\partial c'}{\partial x'} \right) \qquad \frac{\partial}{\partial x} = \frac{1}{x_0} \frac{\partial}{\partial x'}$$

$$= \frac{D_0 c_0}{x_0^2} D' \frac{\partial^2 c'}{\partial x'^2}$$

por lo tanto que la ecuación 1, es igual a:

$$\frac{c_0}{t_0} \frac{\partial c'}{\partial t'} = \frac{D_0 c_0}{x_0^2} D' \frac{\partial^2 c'}{\partial x'^2}$$

donde los términos con coma son adimensionales. Dividiendo los dos términos entre c_0/t_0, queda

$$\frac{\partial c'}{\partial t'} = \frac{D_o t_o}{X_o^2} \; D' \frac{\partial^2 c'}{\partial x'^2} \quad \ldots (2)$$

Si se acepta que la misma ecuación se aplica tanto al prototipo como al modelo, entonces deberá mantenerse igual el parámetro $\Pi = \frac{D_o t_o}{X_o^2}$ en los dos casos

Entonces se tendrá semejanza si se mantienen este número igual en prototipo y modelo, o sea

$$\frac{c}{c_o} = f\left(\frac{D_o t_o}{X_o^2}, \frac{x}{x'}, \frac{t}{t'}, \text{geometría}\right)$$

Con fines prácticos, obsérvese que para pasar de la ecuación (1) a la adimensional (2) se hicieron operaciones que equivalen en el resultado a simples sustituciones algebraicas, sacando del operador diferencial a las variables de referencia por ser constantes. En el caso de derivadas sucesivas, la variable de referencia del denominador aparece elevada al cuadrado si es de segundo orden, al cubo si es de tercero, etc. Por ejemplo

$$\frac{\partial^2 c}{\partial x^2} = \frac{c_o}{X_o^2} \frac{\partial^2 c'}{\partial x'^2}$$

por lo que por complicada que sea la ecuación, el procedimiento es sencillo.

Para practicar se encontrarán los criterios de semejanza a partir de la ecuación de Navier-Stokes para flujos incompresibles.

La ecuación en dirección x , es:

$$\frac{\partial u}{\partial t} + u\frac{\partial u}{\partial x} + v\frac{\partial u}{\partial y} + w\frac{\partial u}{\partial z} = -\frac{1}{\rho}\frac{\partial p}{\partial x} - g\frac{\partial h}{\partial x} + \nu\nabla^2 u \quad \ldots (3)$$

donde

u, v, w	componentes de la velocidad en dirección x, y, z.
t	tiempo
p	presión
h	distancia en la dirección que actua la gravedad, g
ρ	densidad
ν	viscosidad cinemática

y el Laplaciano se define como

$$\nabla^2 u = \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + \frac{\partial^2 u}{\partial z^2}$$

Si se utilizan como variables de referencia

to	tiempo
Po	presión
Vo,	velocidad
Lo,	longitud

y se consideran ρ, ν y g constantes, se podrán definir las siguientes variables adimensionales:

$$x' = \frac{x}{L_0} \qquad y' = \frac{y}{L_0} \qquad z' = \frac{z}{L_0} \qquad h' = \frac{h}{L_0}$$

$$u' = \frac{u}{V_0} \qquad v' = \frac{v}{V_0} \qquad w' = \frac{w}{V_0}$$

$$t' = \frac{t}{t_0} \qquad p' = \frac{p}{p_0}$$

Sustituyendo estas variables en cada término de la ecuación (3), queda:

$$\frac{\partial u}{\partial t} = \frac{V_0}{t_0}\frac{\partial u'}{\partial t'} \Rightarrow \frac{\partial(u'V_0)}{\partial(t_0 t')} = \frac{V_0}{t_0}\frac{\partial u'}{\partial t'}$$

$$u\frac{\partial u}{\partial x} = \frac{V_0^2}{L_0} u'\frac{\partial u'}{\partial x'} \Rightarrow \frac{u'V_0\,\partial(u'V_0)}{\partial(x'x_0)} = \frac{u'V_0^2}{x_0}\frac{\partial u'}{\partial x'}$$

$$v\frac{\partial u}{\partial y} = \frac{V_0^2}{L_0} v'\frac{\partial u'}{\partial y'} \Rightarrow \frac{v'V_0\,\partial(u'V_0)}{\partial(y'y_0)} = \frac{v'V_0^2}{y_0}\frac{\partial u'}{\partial y'}$$

$$w\frac{\partial u}{\partial z} = \frac{V_0^2}{L_0} w'\frac{\partial u'}{\partial z'} \Rightarrow \frac{w'V_0\,\partial(u'V_0)}{\partial(z'z_0)} = \frac{w'V_0^2}{z_0}\frac{\partial u'}{\partial z'}$$

$$-\frac{1}{\rho}\frac{\partial p}{\partial x} = -\frac{p_0}{L_0}\frac{1}{\rho}\frac{\partial p'}{\partial x'} \Rightarrow -\frac{1}{\rho}\frac{\partial(p'p_0)}{\partial(x'x_0)} = -\frac{p_0}{\rho x_0}\frac{\partial p'}{\partial x'}$$

$$-g\frac{\partial h}{\partial x} = -\frac{L_0}{L_0} g\frac{\partial h'}{\partial x'} \Rightarrow -g\frac{\partial(h'h_0)}{\partial(x'x_0)} = -\frac{g h_0}{x_0}\frac{\partial h'}{\partial x'}$$

$$+\nu\nabla^2 u = \frac{V_0}{L_0^2}\nu\nabla^2 u' \Rightarrow -\nu' V_0\nabla^2(V_0 u') = \frac{V_0}{L^0}\nabla^2 u'$$

(en este último término, aparece Lo^2 pues se está derivando dos veces).

Por lo que la ecuación (3), después de dividir todos los términos entre V_o^2/L_o queda:

$$\frac{L_o}{V_o t_o}\frac{\partial u'}{\partial t'} + u'\frac{\partial u'}{\partial x'} + v'\frac{\partial u'}{\partial y'} + w'\frac{\partial u'}{\partial z'} = -\frac{P_o}{\rho V_o^2}\frac{\partial p'}{\partial x'} - \frac{L_o g}{V_o^2}\frac{\partial h'}{\partial x'} + \frac{\nu}{V_o L_o}\nabla^2 u'$$

Esta ecuación tendrá la misma solución en dos sistemas semejantes geométricamente si los coeficientes adimensionales se mantienen constantes en los dos sistemas.

O sea, llamando α

$$H = \frac{L_o}{V_o t_o} \quad \text{(criterio de homocronía)}$$

y como los otros coeficientes son proporcionales a los números de Euler, Froude y Reynolds, o sea

$$E \sim \frac{P_o}{\rho V_o^2}, \quad F \sim \frac{L_o g}{V_o^2}, \quad R \sim \frac{\nu}{V_o L_o}$$

y se recuerda que el número de Euler es, generalmente, la variable dependiente, se podrá decir que se deberán mantener H, F y R constantes en prototipo y modelo para tener el mismo E en puntos correspondientes, esto es:

$$E = f(H, F, R, \text{geometría})$$

Esta forma de representar la relación entre los coeficientes adimensionales, es conveniente y se continuará usando en los próximos capítulos.

Como casos especiales de esta ecuación se puede poner:

Para flujos permanentes, (que no varían en el tiempo):

$$E = f\,(F, R, \text{geometría})$$

Para flujo permanente con fronteras rígidas, o sea sin superficie libre.

$$E = f\,(R, \text{geometría})$$

Para flujo permanente turbulento con superficie libre.

$$E = f\,(F, \text{geometría})$$

a) donde la geometría es más importante

b) flujo laminar

c) escurrimiento hidráulicamente liso

En el primer caso la geometría determina las características del escurrimiento, por ejemplo: el modelo de una transición en un túnel, de placas orificios para introducir pérdidas, de empujes en obstáculos, etc. En estos casos la geometría juega el papel central y como el flujo es turbulento, o sea las fuerzas de viscosidad son despreciables al lado de las de inercia, el criterio para tener semejanza dinámica es la semejanza geométrica y mantener en el modelo un escurrimiento turbulento. Nótese que no es necesario conservar el número de Reynolds sino simplemente que sea mayor que un valor crítico para garantizar que el escurrimiento en el modelo también sea turbulento.

En este caso la ecuación (1) se simplificará a

$$E = \text{cte}$$

y, por lo tanto, la relación entre velocidad y presión, por ejemplo:

$$\frac{V}{\sqrt{2\Delta p/\rho}} \quad \text{ó} \quad \frac{\Delta h}{V^2/2g} \,,$$

será constante para todos los puntos correspondientes.

b) Flujo laminar.

En este caso se tendrá semejanza dinámica manteniendo el mismo número de Reynolds. La rugosidad, siempre que sea de un orden de magnitud inferior al diámetro ~~o ancho~~ del conducto no influye.

c) Escurrimiento hidráulicamente liso.

Con este escurrimiento las protuberancias de la rugosidad están cubiertas por una subcapa laminar que no permite que éstas influyan. Para saber en qué casos el ~~escurrimiento~~ se puede considerar hidráulicamente liso se utilizan parámetros, que se verán con detalle en el siguiente capítulo. Igual que en el escurrimiento laminar se tendrá semejanza dinámica manteniendo el mismo número de Reynolds.

6.2 Interviene la rugosidad.

La rugosidad es un factor importante en dos casos:

a) escurrimiento de transición

b) flujo hidráulicamente rugoso

En los escurrimientos que están en la zona de transición importa el efecto viscoso y la rugosidad relativa; para ser modelados hay que considerar al número de Reynolds y mantener la semejanza geométrica de la rugosidad. Como esto es difícil habrá que recurrir a técnicas experimentales especiales, que se describirán en el siguiente capítulo.

Para flujos en regimen hidráulicamente rugosos o sea completamente turbulento, la influencia de la viscosidad es mínima y se tendrá semejanza dinámica con la semejanza geométrica, incluyendo a la rugosidad, y manteniendo el número de Reynolds arriba de un valor mínimo para garantizar que tanto modelo como prototipo están en el mismo régimen.

Como modelar la rugosidad es impráctico se puede recurrir a una aproximación que utiliza la rugosidad equivalente y la cual se encuentra tabulada para diferentes tipos de superficie, por ejemplo vidrio, acero, asbesto, etc. Ver Tabla .

7. Modelos con superficie libre

En este tipo de modelos la ecuación a satisfacer es

$$E = f(F, \text{rugosidad, geom.})$$

y también ahora hay casos donde la rugosidad es importante y otros en donde juega un papel secundario.

Antes de ver estos casos con más detalle, es conveniente definir los parámetros que se utilizan para representar el efecto de la rugosidad en escurrimientos a superficie libre.

7.1 Coeficiente de rugosidad

Las fórmulas más comunes en la práctica para encontrar pérdidas, son la Chezy y la de Manning dadas por:

7.1.1 Fórmula de Chezy

$$Q = CA\sqrt{RS}$$

donde

A, área de la sección hidráulica, $[L^2]$

C, coeficiente de Chezy, $[\sqrt{g}] = [L^{1/2}/T]$

Q, gasto, $[L^3/T]$

R, radio hidráulico, $[L]$

S, pendiente, $[0]$.

7.1.2 Fórmula de Manning

$$Q = A \frac{1}{n} R^{2/3} S^{1/2}$$

donde

1 número con dimensiones, $[\sqrt{g}] = [L^{1/2}/T]$

g gravedad, $[L/T^2]$

n coeficiente de Manning, ver Tabla 5, $[L^{1/6}]$

que se pueden relacionar con la fórmula de Darcy-Weisback ampliamente usada en conductos cerrados

$$h_f = f \frac{L}{D} \frac{V^2}{2g}$$

donde

D diámetro del conducto, que para secciones no muy diferentes de la circular se puede sustituir por cuatro veces el radio hidráulico, $[L]$

f coeficiente de Fanning, que se valúa con el diagrama de Moody, fig 7, $[0]$

h_f pérdidas de carga entre dos puntos que distan L, $[L]$,

V velocidad, $[L/T]$

Esta ecuación se puede poner en forma similar a la de Chezy

$$Q = A\sqrt{\frac{8g}{f}}\sqrt{RS}$$

7.1.3 Relación entre los coeficientes de pérdidas

Los coeficientes anteriores satisfacen las siguientes relaciones entre sí,

$$C = \frac{R^{1/6}}{n} = \sqrt{\frac{8g}{f}}$$

así como el cortante medio en la pared del conducto, τ_o, $[M/LT^2]$,

$$\frac{\tau_o}{\rho V^2} = \frac{g}{C^2} = \frac{gn^2}{R^{1/3}} = \frac{f}{8}$$

donde

ρ densidad del fluído, $[M/L^3]$

Los coeficientes n y C se encuentran en función de la rugosidad del conducto, ver Tabla 5, y f de la rugosidad relativa y del número de Reynolds, fig 7. Esto representa una aproximación burda pues en la realidad f, y por lo tanto n ó C, son función de (ref 12):

a) número de Froude
b) geometría de la sección
c) rugosidad relativa
d) uniformidad, en perfil y planta, del canal
e) número de Reynolds
f) variación del escurrimiento en el tiempo.

7.2 No interviene la rugosidad

Cuando la geometría general del prototipo es determinante en el escurrimiento, como por ejemplo: vertedores, obras de toma, compuertas, transiciones, etc. O también cuando el tamaño del modelo impide que se desarrollen las influencias de la rugosidad; no es necesario escalar la rugosidad con cuidado y es suficiente aproximarse a ella, utilizando el material o el acabado que tenga una n adecuada. Recuérdese que de la Tabla 4 cuando se sigue la ley de Froude,

$$n_e = Le^{1/6}$$

y como Le generalmente es mayor que la unidad, también ne lo será, esto es:

$$n_p > n_m$$

por lo que el modelo usualmente es más liso que el prototipo. Cuando esta condición no puede satisfacerse por tenerse una rugosidad en el prototipo y una escala tal que impliquen una rugosidad de modelo menor que la que se puede dar con un material muy liso, por ejemplo si

$$n_m \leq 0.008$$

es suficiente construir el modelo con un acabado lo más liso posible y no tomar en cuenta a la ne. Aunque es recomendable reducir a la Le (construir al modelo un poco mayor) para obviar este problema.

Otros dos casos donde la rugosidad no es importante es cuando el escurrimiento es laminar o es hidráulicamente liso; como en los flujos con superficie libre.

Problemas donde

7.3 Interviene la rugosidad

Esta es importante en los siguientes casos:

a) escurrimiento de transición

b) flujo hidráulicamente rugoso

Para modelar estos escurrimientos hay que calibrar el modelo modificando la rugosidad, por medio de tanteos, hasta que se tenga una semejanza cinemática aproximada.

Solo en el caso, poco común, en que la rugosidad del modelo se pueda reproducir manteniendo semejanza geométrica con la del prototipo, o sea

$$k_e = L_e$$

no es necesario calibrar y se tendrá semejanza dinámica conservando el mismo número de Froude.

7.4 Modelos distorsionados

Debido a que hay problemas en que las distancias horizontales son de un orden de magnitud mayor que las distancias verticales, por ejemplo: ríos, playas, estuarios; se usan modelos distorsionados donde la escala de longitudes horizontales es mayor que la de verticales, o sea

$$X_e > Y_e$$

y el cociente

$$\frac{X_e}{Y_e} = r$$

es la distorsión que se trata de mantener pequeña, por ejemplo menor que 8.

El usar modelos distorsionados dificulta la posibilidad de tener semejanza dinámica y además implica que las relaciones de escala se modifiquen, pues la Y_e es un nuevo parámetro, así:

$$A_e = X_e Y_e$$

y

$$Q_e = X_e Y_e^{3/2} \quad \text{etc.}$$

en la Tabla 7 se muestran las nuevas relaciones que se deben satisfacer.

7.2.1 Ventajas y desventajas de modelos distorsionados

Algunas de las ventajas de usar modelos-distorsionados son:

1) Los tirantes, alturas de ola y pendientes se exageran facilitando su medición.

2) Ayuda a reducir el tamaño del modelo

3) Aumenta el número de Reynolds del modelo

4) En modelos de fondo móvil se facilita el movimiento del material.

y desventajas:

1) Se distorsiona la semejanza dinámica

2) El escurrimiento puede ser diferente que en el prototipo, por ejemplo en modelos de oleaje.

3) Aumenta la posibilidad de error y produce una impresión desfavorable al que lo observa.

4) En problemas de fondo móvil puede ser imposible reproducir las laderas y pendientes si el material es suelto.

TABLA 4

Características del flujo y escalas de semejanza

Característica	Ley de Froude	Ley de Reynolds
Longitud	L_e	L_e
Area	L_e^2	L_e^2
Volumen	L_e^3	L_e^3
Tiempo	$[L_e\rho_e/\gamma_e]^{1/2}$	$L_e^2\rho_e/\mu_e$
Velocidad	$[L_e\gamma_e/\rho_e]^{1/2}$	$\mu_e/L_e\rho_e$
Aceleración	γ_e/ρ_e	$\mu_e^2/\rho_e^2L_e^3$
Gasto	$L_e^{5/2}\,(\gamma_e/\rho_e)^{1/2}$	$L_e\mu_e/\rho_e$
Masa	$L_e^3\rho_e$	$L_e^3\rho_e$
Fuerza	$L_e^3\gamma_e$	μ_e^2/ρ_e
Presión	$L_e\gamma_e$	$\mu_e^2/\rho_eL_e^2$
Impulso y cantidad de movimiento	$L_e^{7/2}\,(\rho_e\gamma_e)^{1/2}$	$L_e^2\mu_e$
Energía y trabajo	$L_e^4\gamma_e$	$L_e\mu_e^2/\rho_e$
Potencia	$L_e^{7/2}\gamma_e^{3/2}/\rho_e^{1/2}$	$\mu_e^3/L_e\rho_e^2$

Notas:

1. Si se usa el mismo fluido en prototipo y modelo:

$$\rho_e = \gamma_e = \mu_e = 1$$

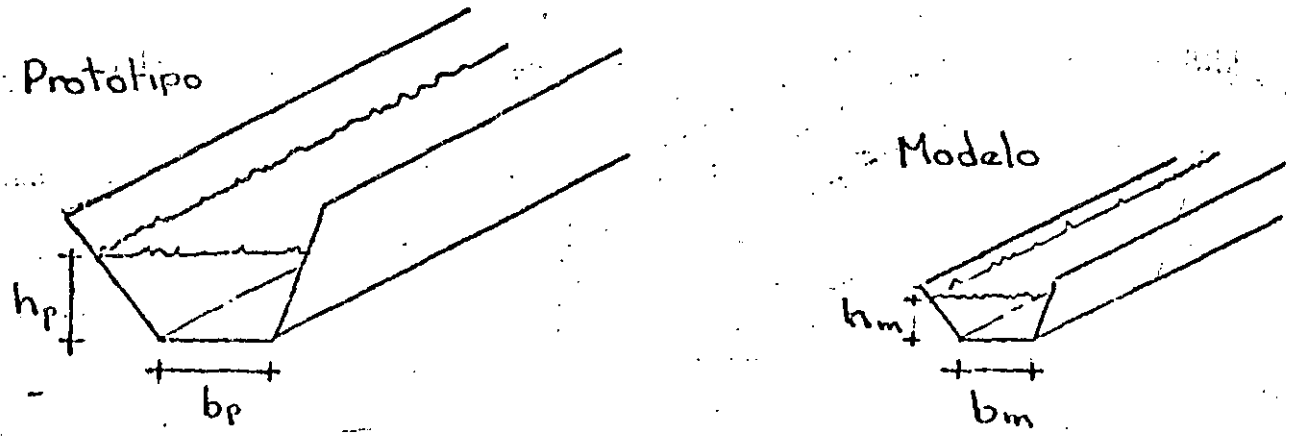
2. Cuando $g_e = 1$, $\gamma_e/\rho_e = 1$

TABLA 5

Valores del coeficiente de Manning (12)

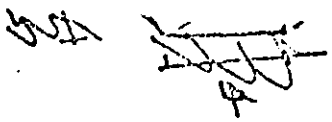
Para Tablas más extensas ver referencias 13 y 14

Tipo del conducto y acabado	
Muy liso. Vidrio, lucita, bronce. Paredes alineadas	0.009 y 0.010
Madera lisa, metal, concreto liso. Paredes alineadas	0.011 y 0.012
Madera lisa, metal, concreto liso. Paredes razonablemente alineadas	0.013
Madera en buen estado, superficies de concreto con algo de curvatura y con algunos depósitos, acabado con plana	0.014
Madera con depósitos, concreto liso, acabado con plana. Metal con proyecciones leves o liso con curvaturas excesivas	0.015
Canales de metal con proyecciones grandes. Madera o concreto con depósitos de algas o fuertes	0.016

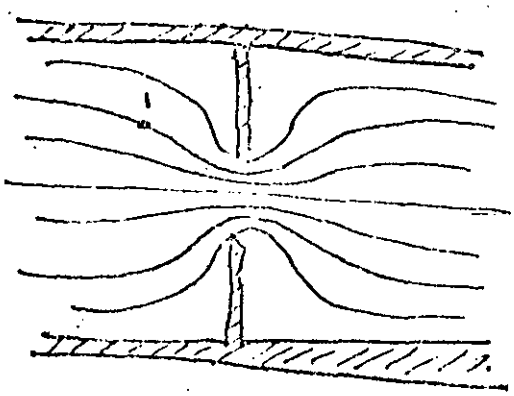


$$\frac{b_p}{b_m} = \frac{h_p}{h_m} = \cdots = L_e \text{ (escala de longitudes)}$$

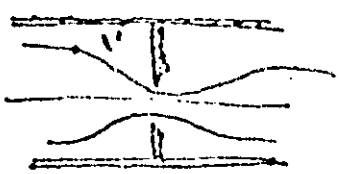
Fig 3. Semejanza geométrica

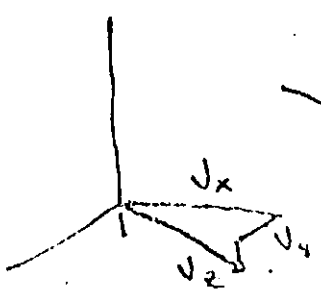


Prototipo (p)



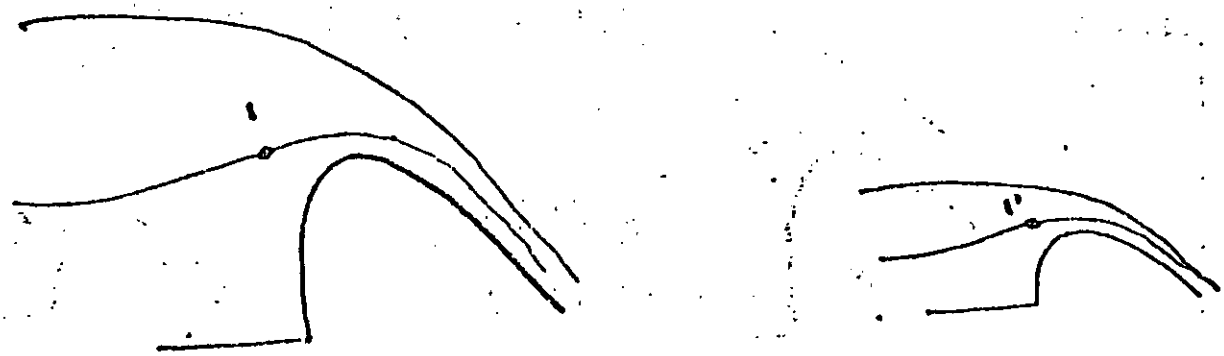
Modelo (m)

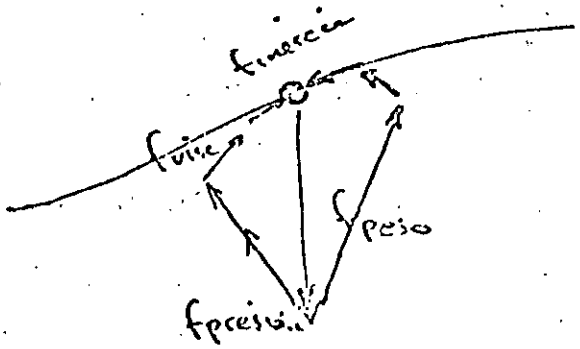




$$\frac{V_{xp}}{V_{xm}} = \frac{V_{yp}}{V_{ym}} = \frac{V_{zp}}{V_{zm}} = \cdots = V_e$$

(escala de velocidad





$$\frac{f_{iP}}{f_{iPn}} = \frac{f_{wP}}{f_{wPn}} = \frac{f_{PP}}{f_{PPn}} = \frac{f_{DP}}{f_{DPn}} = \dots = f_e$$

escala de fuerzas

Fig 5 Semejanza dinámica

Guía general para el Diseño de Sist. Integrales de Saneamiento del Agua Potable.

Separatas correlativas en Hidrología Subterránea

V MODELOS CON FRONTERAS RIGIDAS

Como se vió en el capítulo anterior una primera división en modelos puede ser:

Con fronteras rígidas

Con superficie libre

que obviamente no presenta problema en su identificación.

En este capítulo se expondrá con mayor detalle y dando criterios cuantitativos, las condiciones generales, criterios de semejanza y recomendaciones para las tres categorías de modelos con fronteras rígidas:

1. Donde domina la geometría
2. Flujo laminar e hidráulicamente liso
3. Flujo de transición e hidráulicamente rugoso

FRONTERAS RIGIDAS

1. Domina la geometría

1.1. Condiciones generales

Modelos con fronteras rígidas donde domina el efecto de la geometría. La rugosidad juega un papel secundario. Flujo turbulento. No hay superficie libre. El fluido puede ser líquido o gas. Ver sección 2.3.

1.2. Ejemplos

Válvulas, codos, transiciones (ampliaciones o contracciones) en tuberías y túneles llenos, cambios de dirección, placas orificio, empuje en: rejillas, obstáculos, edificios y cuerpos en general.

1.3. Criterios de semejanza

En estos problemas es usual que el flujo, tanto en prototipo como en modelo, sea francamente turbulento, por lo que se tendrá semejanza dinámica si se tiene:

i) Semejanza geométrica

ii) Escurrimiento turbulento en modelo y prototipo, lo que se logra satisfaciendo que

$$IR_{min} > 4000$$

En general, es recomendable tener no solo escurrimiento turbulento sino que además tengan turbulencia parecida, por lo que conviene que los números de Reynolds no sean muy diferentes, por ejemplo que:

$$iii) \quad IR_{modelo} > \frac{1}{10} IR_{prototipo}$$

aunque esta restricción no es de la importancia de las dos anteriores.

Si se satisfacen estas recomendaciones, la ecuación dinámica queda

$$E = f(\text{geometría}) = cte$$

por lo que se tendrá el mismo número de Euler, en puntos correspondientes de modelo y prototipo, para todas las condiciones de escurrimiento.

Lo anterior significa que basta con un ensayo bien hecho, que permita encontrar los números de Euler en los diversos puntos del problema, para tener la distribución de presiones debidas a cualquier velocidad del fluido.

1.4. Recomendaciones

Es conveniente seguir las siguientes recomendaciones:

1.4.1 Escalas

$$Le = \frac{L_p}{L_m}$$

$5 < Le < 30$ modelos de válvulas, codos, rejillas, etc.

$100 < Le < 200$ modelos de edificios, ampliaciones, etc.

Además se debe satisfacer:

1.4.2 Diámetro, o diámetro equivalente, mayor de 10 cm.

1.4.3 Rugosidad relativa

$$K/D < 0.004$$

rugosidades mayores se deben considerar como fronteras irregulares y por lo tanto escalarse geometricamente.

1.4.4 En el caso que en el modelo o prototipo el fluído sea aire, las

velocidades no deben exceder 100 m/s para evitar efectos de compresibilidad.

1.4.5 El flujo de llegada debe tener una distribución de velocidad lo más parecido posible al prototipo. Si éste tiene velocidad uniforme, el modelo debe reproducir una longitud de 20 a 50 diámetros aguas arriba del punto de interés para que se uniformize el escurrimiento.

1.5. Construcción La zona de aguas abajo del modelo debe ser a conducto lleno, por lo que a veces es necesario controlar el flujo por medio de agujas, válvulas o tirantes de agua a la salida.

Se construye el modelo con material resistente y liso, reproduciendo a escala lo que se vaya a estudiar. Se recomienda usar un material transparente, por ejemplo lucita, o proveer de ventanas de observación. Es conveniente medir la velocidad media y la presión antes de la geometría que se vaya a estudiar para usarlas como referencia, además de colocar piezómetros en puntos críticos o de interés, considerando que, en general, es mejor que sobren a que falten.

1.6. Operación

Una vez establecido el flujo y cebados los aparatos de medición, con la velocidad de referencia se comprueba que se satisfacen los criterios enunciados en la sección 3. En caso de que no se puedan satisfacer o de que se tengan dudas acerca de si influye la viscosidad (o sea que intervenga el número de Reynolds), se hace lo siguiente: se mide la presión en un punto de la geometría a estudiar, p_1, y se encuentra el número de Euler correspondiente utilizando la presión y velocidad de referencia o sea

$$E_1 = \frac{V_o}{\sqrt{2\Delta P/\rho}}$$

donde

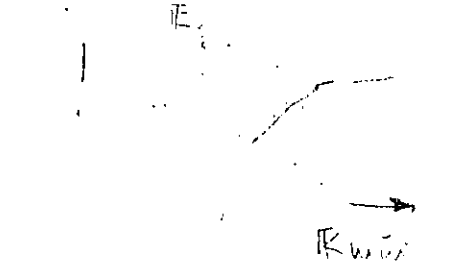
$$\Delta p = |p_1 - p_0|$$

Este valor se gráfica contra el número de Reynolds, ~~ver fig.~~ y se repite el proceso hasta tener una gráfica como la mostrada en la fig. . En el punto en que las E dejan de estar en una horizontal, sin tomar en cuenta las dispersiones naturales del experimento, se encuentra el R_{min} (que sea el límite inferior) con el que se debe operar al modelo.

1.7. Observaciones

Aunque como ya se dijo, si se tiene semejanza geométrica y el flujo es turbulento se tendrá semejanza dinámica no solo entre prototipo y modelo sino tambien para cualquier velocidad, o sea el número de Euler permanecerá constante en puntos correspondientes para diferentes velocidades y gastos, no hay que olvidar que la semejanza dinámica se tendrá mientras las líneas de corrientes del fluído, y no solo las fronteras, sean semejantes geometricamente. Por lo tanto en problemas donde hay separación del escurrimiento (de los puntos de esta) depende de la velocidad ~~y la presión~~ del flujo, por ejemplo en ampliaciones suaves o redondeadas, se debe tener semejanza tambien en la frontera libre (ver fig.) lo que se logra manteniendo números de Reynolds parecidos. Si las fronteras son angulosas la separación ocurrirá en los vértices de la geometría y no dependerá tanto del número de Reynolds.

Afortunadamente para un rango amplio de números de Reynolds los



puntos de separación no varian demasiado, aunque hay casos, como el del flujo alrededor de una esfera o cilindro en donde para números de Reynolds de aproximadamente 3×10^5 el flujo cambia bruscamente de configuración.

Se valua

$$\frac{k_s}{d} = 0.00025 \quad 2.5 \times 10^{-4}$$

del diagrama de Moody, Apéndice, se puede estimar

$$f_s \doteq 0.014$$

despejando de la inecuación (1) y con $\nu = 0.01\ cm^2/s$

$$V < \frac{14\,\nu}{k_s\sqrt{f}} = 237\ cm/s$$

$$\therefore \quad V < 2.37\ m/s$$

O sea que para velocidades hasta de 2.37 m/s la rugosidad no interviene pues el flujo es laminar o hidraulicamente liso.

$$\frac{V k_s \sqrt{f}}{\nu} < 14$$

$$V < \frac{14\,\nu}{k_s\sqrt{f}} = \frac{14 \times 0.01}{5\times10^{-2} \times (0.014)^{1/2}} = 236.64\ cm/s$$

$$V < 2.37\ m/s \quad (OK)$$

3. Flujo de transición o hidraulicamente rugoso

3.1. Condiciones generales

Modelos con fronteras rigidas y sin superficie libre. Intervienen la rugosidad y el efecto viscoso. Velocidades no muy altas y altas.

3.2. Ejemplos

Bombas, turbinas, tuberías y conductos llenos, como: túneles a presión.

3.3. Criterios de semejanza

Para escurrimientos de transición se tendrá semejanza dinámica si se tiene semejanza geométrica, incluyendo a la rugosidad, y se mantiene el mismo número de Reynolds en prototipo y modelo. O sea

$$E = f(R, \text{geom.}, \text{rugosidad})$$

En este tipo de escurrimiento el espesor de la subcapa laminar es del mismo orden que las rugosidades de la frontera, y el criterio para identificarlo es

$$14 < \frac{V k_s \sqrt{f}}{\nu} < 198$$

donde las variables fueron definidas en la sección anterior.

Cuando el escurrimiento es hidráulicamente rugoso, o sea el producto adimensional anterior es mayor que 198, ya no existe subcapa laminar pues la mayor rugosidad no deja que se forme.

En este caso el criterio de semejanza se reduce a mantener la semejanza geométrica, incluyendo a la rugosidad, y tener un número de Reynolds lo más grande posible.

$$E = f\ (\text{geom.}, \text{rugosidad})$$

Como escalar a la rugosidad requiere reproducir las pequeñas irregularidades de la superficie del prototipo de tal manera que satisfagan la misma escala de longitudes L_e, que relaciona las otras dimensiones de prototipo y modelo, y esto es impráctico; se recurre a una aproximación utilizando la rugosidad tabulada, en tal forma que:

$$k_e = L_e$$

o se hace una calibración que requiere medir alguna característica en prototipo y modelo (por ejemplo: pérdidas, velocidades, números de Euler) y si no son semejantes se modifica la rugosidad hasta lograr la semejanza. Esto se describirá con más detalle en la Sección que es donde se usa con más frecuencia.

3.1. Recomendaciones

Si el mantener $k_e = L_e$ implica usar un material en el modelo excesivamente liso se recomienda hacer el modelo lo más grande posible y con la velocidad del gasto mayores que se pueda, de tal forma de tener un Reynolds lo más parecido posible al del prototipo.

3.2. Escalas

$5 < L_e < 10$	Bombas utilizando agua Bombas utilizando aire
$L_e \sim 50$	Turbinas donde se reproduce la tubería a presión o el tubo de aspiración.
$50 < L_e < 100$	Túneles llenos
$100 < L_e < 500$	Conductos a presión.

3.5. Modelos de bombas y turbinas.

Como un caso ilustrativo de este tipo de modelos se verán los criterios para modelar bombas y tubbinas.

En el funcionamiento de una bomba o turbina interesa encontrar la eficiencia η, en función de:

D	diámetro del rodete
H	carga
n	velocidad angular
Q	gasto
g	gravedad
k	rugosidad
~~η~~	~~eficiencia~~
~~ρ~~	~~densidad~~
ν	viscosidad dinámica cinemática

O sea

$$\eta = f(D, H, n, Q, g, k, \nu)$$

Si se construye el modelo lo más grande y liso que se pueda y se opera con el Q mayor posible, la rugosidad se podrá escalar o jugará un papel secundario y el escurrimiento será turbulento, por lo que la viscosidad no será importante. Entonces la ecuación anterior se simplificará a:

$$\eta = f(D, H, n, Q, g)$$

que agrupado en productos adimensionales quedará:

$$\eta = f\left(\frac{D}{H}, \frac{n^2 D}{g}, \frac{Q}{nD^3}\right)$$

y que como la experiencia indica, se puede aproximar con simplificar a:

$$\eta = f\left(\frac{n^2 D^2}{gH}, \frac{Q}{nD^3}\right)$$

Entonces los criterios de semejanza serán:

$$\frac{n_e^2 D_e^2}{g_e H_e} = 1$$

$$\frac{Q_e}{n_e D_e^3} = 1$$

que, haciendo operaciones, y sustituyendo $g_e = 1$ se pueden poner como:

$$\frac{n_e D_e}{\sqrt{H_e}} = 1 \qquad \ldots (1)$$

$$\frac{Q_e}{D_e^2 \sqrt{H_e}} = 1 \qquad \ldots (2) \quad \Rightarrow Q_e = D_e^2 \sqrt{H_e}$$

Otra variable, derivada de las anteriores, de interés, es la potencia dada por:

$$P = \gamma Q H \quad \text{o sea} \quad P_e = \gamma_e Q_e H_e$$

que con igual fluido, $\gamma_e = 1$ y eliminando a la Q_e con la ecuación (2), queda:

$$P_e = \gamma_e H_e = D_e^2 \sqrt{H_e} \cdot H_e$$

$$\frac{P_e}{D_e^2 H_e^{3/2}} = 1 \qquad \frac{P_e}{D_e^2 H^{3/2}} \ldots (3)$$

Las tres ecuaciones anteriores permiten escalar las propiedades entre prototipo y modelo. Sin embargo, en la práctica es común definir "valores unitarios" que corresponden a una turbina o bomba geométricamente semejante con un rodete de 1 m y operando

bajo una carga neta de 1 m.

[Obsérvese que no hay una diferencia fundamental entre usar las ecuaciones (1), (2), y (3), que son adimensionales y los valores unitarios, que no lo son. Se continuará con los valores unitarios por ser la forma usual como encuentra uno resultados].

Entonces, de las ecuaciones anteriores

$$n_1 = \frac{n\,D}{\sqrt{H}} \qquad \frac{n^{1/2}}{n^{1/2}} = L^{1/2}\,T^{-1}$$

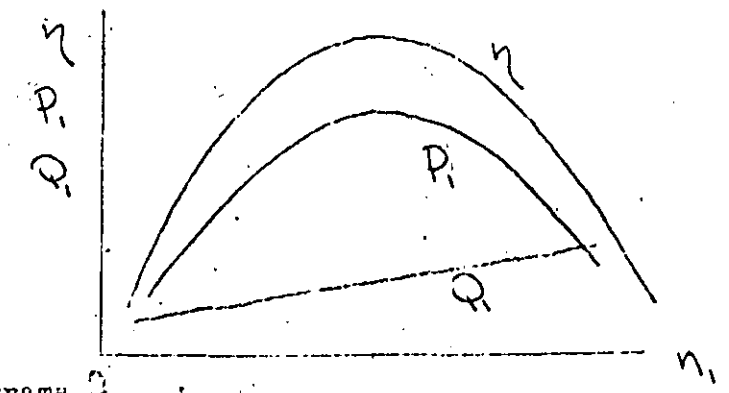
$$Q_1 = \frac{Q}{D^2\sqrt{H}} \Rightarrow \frac{L^3\,T^{-1}}{L^2\,L^{1/2}} = L^{1/2}\,T^{-1}$$

$$P_1 = \frac{P}{D^2 H^{3/2}} \Rightarrow \frac{M L^2 T^{-3}}{L^2 L^{3/2}} = M L^{-3/2} T^{-3}$$

donde el subíndice, 1, indica unitario y las otras variables se miden en el modelo.

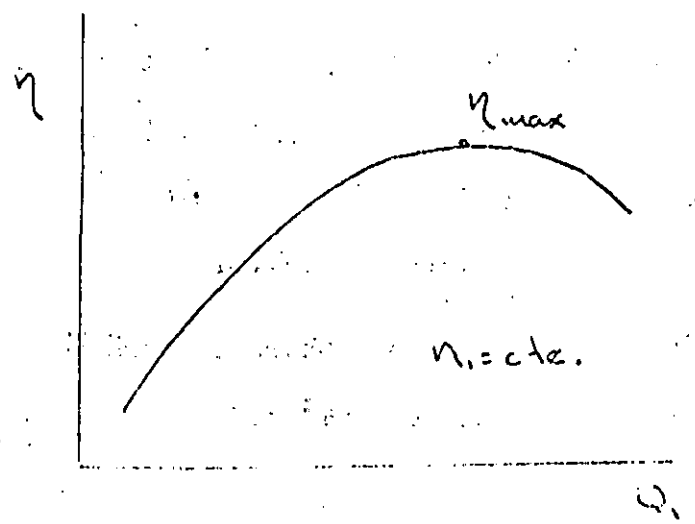
De las pruebas es común obtener las siguientes gráficas.

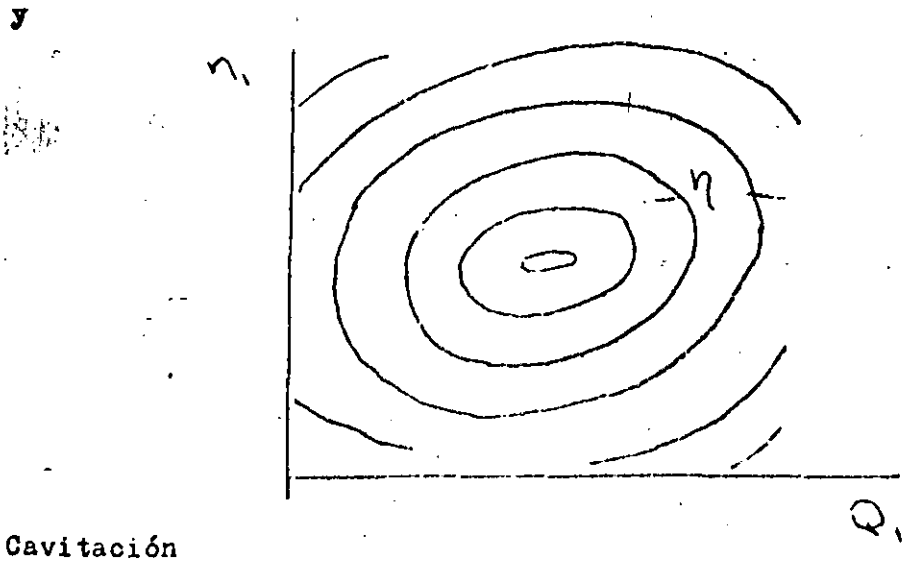
(1)



y de este diagrama

(2)





3.5.1 Cavitación

Si se desea encontrar en modelo las condiciones en que se tendrá cavitación se utiliza el criterio de Thoma dado por:

$$\sigma = \frac{H_a - H_v - H_s}{H}$$

donde

H_a - presión atmosférica

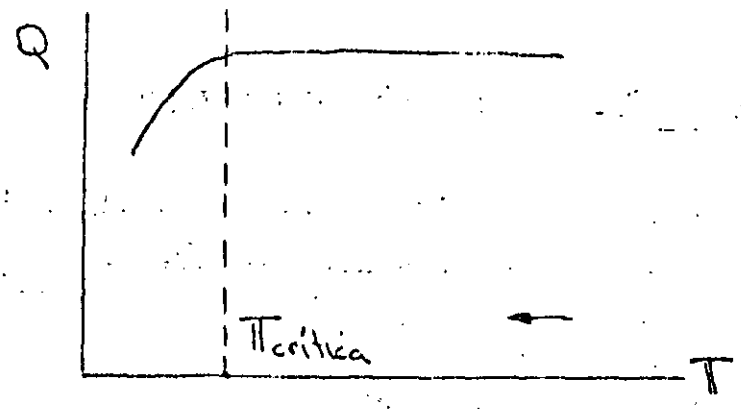
H_v - presión de vapor de agua

H_s - carga de succión

H - carga de la bomba

Para bombas la prueba consiste en mantener la velocidad y carga constantes y en aumentar H_s, (o sea disminuyendo σ), por medio de una válvula en el tubo de succión, hasta que se produce cavitación, la cual se reconoce por el ruido que hace y por la reducción en el gasto. Este valor de σ se denomina el coeficiente de cavitación crítico.

En la figura se muestra una gráfica del gasto y la potencia contra el índice de cavitación.



Para turbinas lo que se mantiene constante es la velocidad y la abertura de los varios móviles, mientras se determina la eficiencia y el gasto para diferentes valores de Π.

VI MODELOS CON SUPERFICIE LIBRE

A continuación se presentan los modelos con superficie libre que como se dijo anteriormente se pueden agrupar en:

1. Domina la geometría
2. Flujo laminar e hidráulicamente liso
3. Flujo de transición e hidráulicamente rugoso

Modelos a Superficie Libre
- No interviene la rugosidad
  - a) Domina la geometría
  - b) Flujo laminar
  - c) Hidráulicamente liso
- Interviene la rugosidad
  - a) Flujo de transición
  - d) Hidráulicamente rugoso

FLUJO A SUPERFICIE LIBRE

1. Domina la geometría

1.1. Condiciones generales

Modelos con superficie libre. No pinta la rugosidad.

Flujo turbulento. El fluido es líquido. (OJO)

1.2. Ejemplos

La mayoría de los modelos que interesan al Ingeniero Hidráulico, como: vertedores, obras de toma, cápidas, pilas, cambios de sección en canales, salto hidrúlico, tanques de amortiguamiento, cubetas deflectoras, saltos de ski. Túneles, alcantarillas y conductos en general con superficie libre.

1.3. Criterios de semejanza

Se debe escalar la geometría y mantener el mismo número de Froude, o sea

$$E = f(F, \text{geometría})$$

Aunque la rugosidad juega un papel secundario es conveniente escalarla lo más posible, en tal forma que:

$$k_e = L_e$$

o más común en escurrimientos a superficie libre,

$$n_e = L_e^{1/6} \quad \text{---- (1)}$$

donde n es el coeficiente de Manning

1.4. Recomendaciones

Como la rugosidad mínima que se puede dar en modelo es $n=0.010$ o a lo mucho $n=0.008$ no conviene hacer el modelo demasiado chico (<---- y por lo tanto L_e y n_e grandes), sin embargo

La ecuación (1) se debe considerar como algo que sería deseable satisfacer cuando se no como un criterio para fijar escalas; o sea, simplemente se recomienda escalar la rugosidad lo más posible.

1.4.1 Escalas recomendadas

$10 < L_e < 100$	Vertedores
$10 < L_e < 60$	Tanques amortiguadores
$3 < L_e < 20$	Estructuras en canales y caídas
$50 < L_e < 100$	Obras de toma y excedencias

1.4.2 Reproducir

lo más posible el flujo de llegada que se espera. Para esto puede ser necesario construir en el modelo una zona amplia aguas arriba, así como tranquilizar y encauzar al agua por medio de enhuacalados, mallas, muros, vanos, etc.

1.4.3 Los tirantes en el modelo deben ser mayores de 3 cm, si el régimen es subcrítico o lento, esto es si

$$F < 1$$

y mayores de 1 cm si es supercrítico o rápido, para pues para tirantes menores el efecto de tensión superficial es importante.

Obsérvese que, a veces, no es posible satisfacer esta condición en todo el modelo por lo que se estará haciendo un modelo

parcial, en el que se espera semejanza dinámica sólo en ciertas zonas del modelo, por ejemplo: las que tengan tirantes mayores, y no en todo el modelo.

Otra zona donde no se estén escalando los efectos es en el perímetro del agua que está en contacto con las paredes del modelo ~~en el perímetro del agua~~ cuando hay régimen lento, pues el menisco obviamente no se presentará aumentado en el prototipo.

1.4.4 Para construcción y operación de este tipo de modelo véase el capítulo [illegible] VII.

FLUJO A SUPERFICIE LIBRE

2. Flujo laminar o hidráulicamente liso

2.1. Condiciones generales

Modelos con superfice libre. Flujo laminar o hidráulicamente liso. Velocidades bajas, dimensiones reducidas o fluídos viscosos. No intervienen la rugosidad de las fronteras.

2.2. Ejemplos

En la Ingeniería Hidráulica son poco comunes. Escurrimiento de lluvia sobre pisos o carreteras. Flujo del vidrio en los tanques de fundición., Mesas para laminar celulosa, en la fabricación del papel.

2.3. Criterios de semejanza.

Se deben mantener los mismos números de Froude y de Reynolds, además de la semejanza geométrica. Como la escala de gravedades no es fácil de modificar, esto es $\gamma_c = 1$ lo anterior obliga a tener diferente líquido en prototipo y modelo o a utilizar una escala de longitudes igual a uno, modelo de mismo tamaño que el prototipo. Esto último no es difícil para problemas de pequeñas dimensiones e inclusive puede ser recomendable hacer el modelo mayor que el prototipo.

9/8/82.83

3. Flujo de transición o hidraulicamente rugoso

3.1. Condiciones generales

Modelos con superficie libre. Interviene la rugosidad. Velocidades medias y altas. Fondo fijo.

3.2. Ejemplos

Canales y cauces de ríos donde la rugosidad es importante. Llanuras de inundación. Playas.

3.3. Criterios de semejanza

Si se tiene que

$$\frac{\sqrt{g R S_f}\, k_s}{\nu} < 70$$

y

$$\frac{V R}{\nu} > 500$$

donde

g	gravedad
R	radio hidráulico o tirante si la sección es muy ancha
S_f	gradiente hidráulico
ν	viscosidad ~~cinemática~~ cinemática
V	velocidad.

el escurrimiento será de transición y, para tener semejanza dinámica habrá que mantener la semejanza geométrica y el mismo número de Froude, Reynolds y escala de rugosidades.

$$E = f(\text{geom.}, F, R, Rg.)$$

es necesario aumentar la velocidad del modelo tantas veces como se redujo su tamaño. Hacer esto no se puede hacer siempre y sólo se dá en el modelo el mayor número de Reynolds posible.

Si se aprovecha que ahora no hay fronteras libres, se puede utilizar aire para simular flujos de agua, y viceversa. Como el cociente entre viscosidades cinemáticas a temperatura ambiente entre el aire y agua es:

$$\frac{\nu_{aire}}{\nu_{agua}} \doteq 10$$

hay que dar velocidades diez veces mayores con aire que con agua.

A pesar de esto, es ventajoso utilizar aire pues como su densidad es alrededor de 800 veces menor que el agua necesita mucho menor energía para moverlo, además no requiere almacenarse, no moja ni mancha, no cuesta y no se acaba.

7.2.2.2 Construcción del modelo

Si el fluido es agua es necesario hacer el modelo con un material resistente. Se puede usar cemento o tubos comerciales aunque hay que poner ventanas de observación, ~~si el modelo es de cemento~~ también se puede hacer todo de cristal o lucita. Hacer el modelo de lucita con las ventajas y desventajas mencionadas en la sección anterior.

Para ensayos con aire el modelo es más barato pues se puede hacer de triplay y cartón aunque la medición es más delicada. En la fig. 17 se muestra un esquema de un túnel

de viento para ensayos con aire.

7.3 Operación del modelo

Aunque la operación del modelo depende de las condiciones específicas de cada tipo de problema, a continuación se darán una serie de recomendaciones generales.

Antes de iniciar los ensayos es conveniente hacer los siguientes ajustes en el modelo para mejorar la correspondencia con el prototipo y obtener mejores resultados. 1. Verificar que el agua llegue al modelo en la forma lo más parecida, o que se crea más parecida, al prototipo.

2. Ver que los aparatos de medición estén colocados de acuerdo a las especificaciones y, además, que no alteren al escurrimiento.

3. Comprobar el funcionamiento de las obras de control usadas para dar los gastos o tirantes deseados. En algunos casos el flujo tarda en estabilizarse demasiado tiempo lo que hace que las mediciones sean tardadas e inexactas. Para evitar esto hay que poner controles eficientes, como compuertas con flujo superior y de suficiente cresta, fig. 12, y evitar controles por medio de válvulas u orificios.

4. Observar que el escurrimiento sea razonablemente parecido al del prototipo o como se espera que éste sea. Comprobar que se pueden dar las condiciones de escurrimiento que interesan, como gasto máximo, tirantes mínimos, etc.

5. Inspeccionar el modelo para encontrar grietas y fugas e impermeabilizarlo si esto es necesario.

En max la etapa de ensayos es importante seguir un plan que permita obtener el máximo de información. Las mediciones deben ser muy cuidadosas, repitiéndolas varias veces de preferencia por dos personas. Es conveniente llevar una bitácora, por ejemplo una libreta de tránsito, donde se registren no solo las mediciones sino cualquier observación o comentario que se ocurra durante la prueba. No es recomendable usar hojas sueltas que fácilmente se desordenan y aún extravían. Una práctica que mejora mucho la confiabilidad de los resultados es hacer las gráficas de las variables medidas; de preferencia en forma adimensional, como: No. de Euler, Froude, Coeficiente de descarga, etc., conforme se vayan haciendo las mediciones. De esta manera es posible darse cuenta de cualquier error o anomalía mientras se está haciendo el ensayo y hacer las rectificaciones necesarias.

Un aspecto esencial es no olvidar nunca el objetivo final de los ensayos que es dar información acerca de un fenómeno real en prototipo y, además, usar esta información en beneficio de la obra, por lo que debe haber un estricto contacto entre el experimentador y los ingenieros de diseño, o en su caso, los encargados de la obra.

Una vez terminados los ensayos, se paran las bombas, se drena el modelo, se retira o protege el equipo de medición. Es prudente que aun terminado el estudio se deje el modelo un cierto tiempo sin destruirse, pues en varios casos se han rehabilitado modelos que se consideraban ya obsoletos y se han usado para resolver algún problema o modificación que surgió durante la construcción, y aun operación de la obra, con lo que se ahorran

con cierto
a pagar su costo y el del espacio ocupado. ~~con creces. Lo~~
~~mismo ocurre cuando se hacen modificaciones a la obra.~~

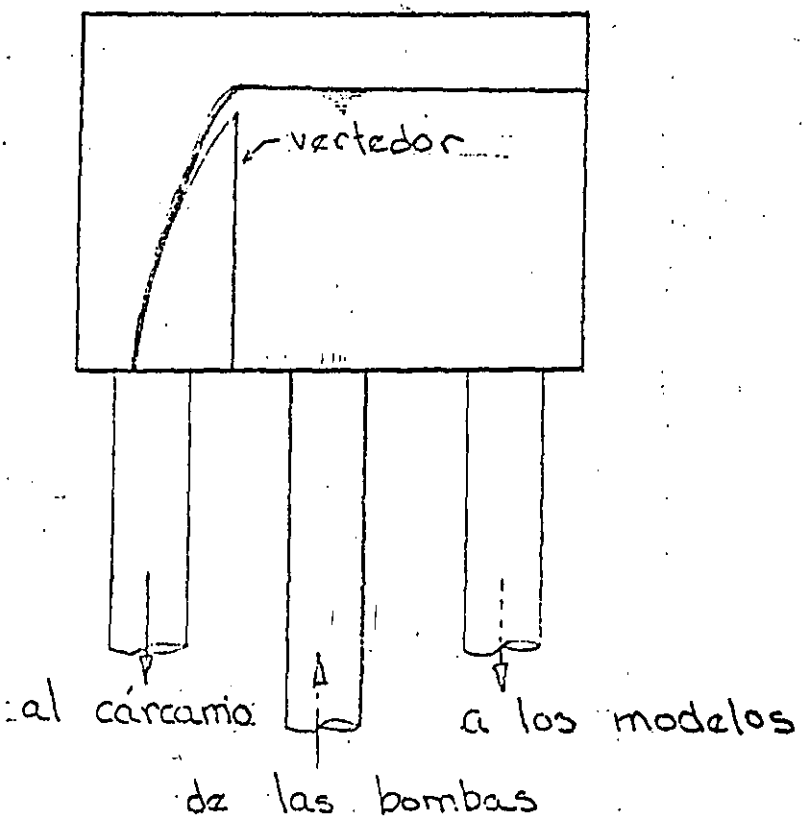


Fig 2 Esquema del tanque de carga constante

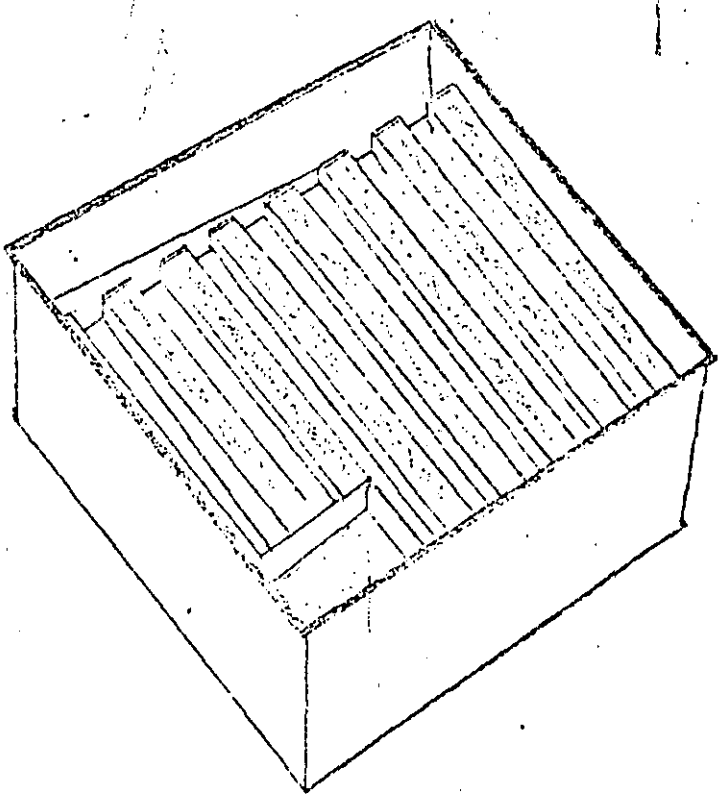


Fig. 3 Tanque de carga constante. Obsérvese la gran longitud del vertedor

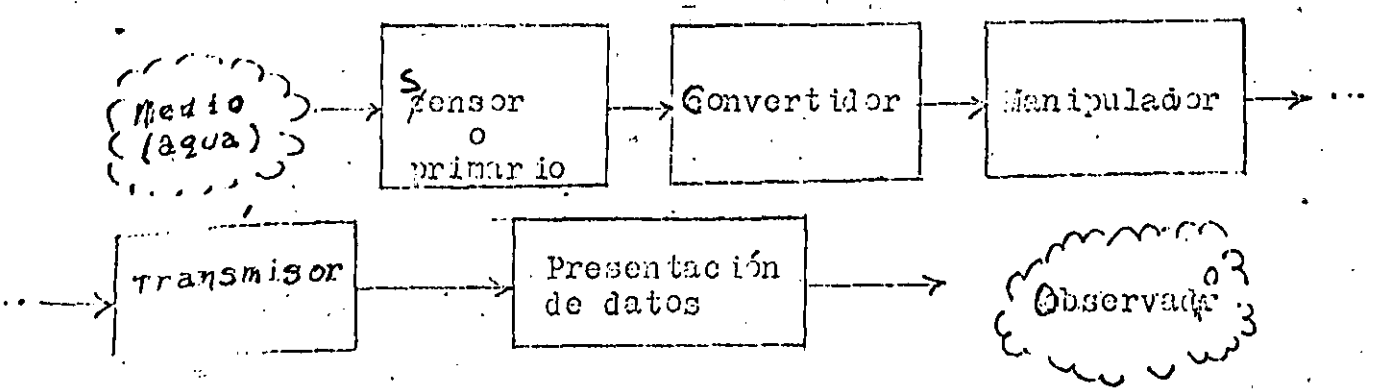
VIII MEDICIONES

Una vez construido el modelo se hacen mediciones en él de acuerdo a un cierto plan. Los resultados dependerán, en gran parte, de la calidad y confiabilidad de estas mediciones así como de su procesamiento.

En este capítulo se presentan los conceptos básicos de lo que es un sistema de medición y se describen las formas de medir las variables más comunes que se presentan en modelos hidráulicos que son: tirantes, velocidades y gastos.

8.1 Sistema de medición.

Es conveniente considerar a un instrumento o a un sistema de medición compuesto, en el caso más general, de los siguientes elementos:



El sensor o primario produce una señal relacionada con la cantidad medida. Usualmente está en contacto con el medio, por ejemplo el agua, y es común que tome energía del mismo medio. Si la energía que extrae es grande puede modificar la

cantidad por medir. Un buen instrumento se diseña en forma que pertube lo menos posible la cantidad medida.

El convertidor transforma la señal del sensor en otra señal más adecuada para su transmisión y procesamiento. Así, un desplazamiento o una presión se pueden convertir en una señal eléctrica.

Como esta señal puede ser muy débil sería necesario amplificarla o, tal vez, convenga eliminar los ruidos que se presentan en ciertas frecuencias por medio de filtros. Este tipo de operaciones se hacen en el elemento manipulador.

El transmisor lleva la señal desde el punto de medición hasta donde está el observador y, finalmente, el elemento de presentación de datos, que puede ser una pantalla, una aguja graficadora, etc. permite que el observador vea a la señal medida. Un ejemplo de lo anterior puede ser la celda de presión, donde el elemento sensor sería el diafragma en contacto con el fluído; y al mismo tiempo, convierte a la señal, presión, en un desplazamiento. Este desplazamiento se vuelve a convertir en una señal eléctrica que es amplificada, por el elemento manipulador, y transportada, por el transmisor, a través de cables hasta donde puede ser observada o registrada por medio del elemento de presentación de datos.

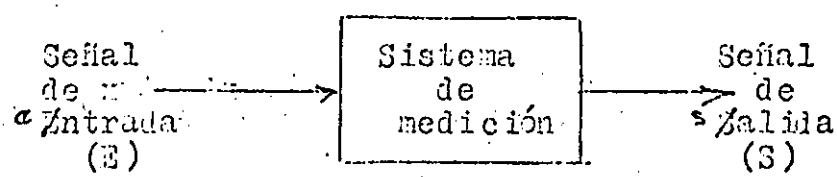
Cada vez que se cambia de medio, por ejemplo aire-agua o má-

quina-hombre se dice que hay una interfase. Como cada elemento y cada interfase puede introducir errores conviene mantener el sistema de medición, o instrumento, lo más sencillo posible.

Cada elemento puede introducir errores por lo que conviene mantener al sistema de medición, o instrumento lo más sencillo posible.

8.2 Funcionamiento de un sistema de medición.

En forma esquemática un sistema de medición se puede representar como:



donde las señales pueden ser, estáticas o dinámicas.

La función que liga la señal de entrada, E, con la salida, S, se llama función de transferencia.

Así, por ejemplo, en un instrumento lineal donde la señal de entrada sea estática, la función de transferencia sería simplemente

$$S = KE$$

donde K es una constante de proporcionalidad.

Obsérvese que las unidades de S y de E pueden ser diferentes por lo que K debería tener las unidades respectivas.

En el caso más general es común aceptar que la función que liga a la señal de entrada con la de salida es una ecuación diferencial lineal ordinaria con coeficiente constantes.

Aunque esta ecuación puede ser de orden n , aquí se considerarán sólo los sistemas de cero, primero y segundo orden, que frecuentemente ocurren en la práctica.

Interesan dos casos:

1. Se tiene un sistema de medición y se desea encontrar sus características dinámicas o sea su función de transferencia.

2. Se conoce la señal de salida, por ejemplo: una lista de mediciones o un registro continuo; y se quiere conocer la señal de entrada que es la variable de interés, como: velocidad o presión.

Para el primer caso es necesario calibrar al aparato utilizando una señal de entrada conocida, como: una señal senoidal, en escalón, etc., y si se acepta que el sistema es de cierto orden, por observación de la señal de salida se pueden encontrar las características dinámicas del sistema.

Una vez conocidas éstas se pasa al segundo caso en donde al hacer una medición se obtienen una señal de salida y lo que interesan es conocer la de entrada, que puede ser muy diferente.

3.3 Instrumento de orden cero.

Ecuación $S = k\,E$ (1)

donde k = sensibilidad estática

$$[k] = [S] / [E]$$

Ejemplos: Potenciómetro de desplazamiento, pantógrafo. Ver fig \

Observaciones: La lectura de salida es proporcional a la señal de entrada sin ninguna distorsión, atraso o defasamiento. Aunque esta condición es ideal hay sistemas que se aproximan lo suficiente como para poder usar este modelo.

Calibración: Si no se conoce k se puede obtener fácilmente usando una señal estática de entrada conocida y comparándola con la de salida.

Determinación de τ:

Una vez conocida la k se puede encontrar una señal de entrada, ya sea estática o dinámica, dividiendo el registro de salida entre la k.

3.4- Instrumento de primer orden (2)

Ecuación

$$\frac{S}{E} = \frac{k}{\tau D + 1} \rightarrow \text{si } \tau = 0 \rightarrow \frac{S}{E} = k \rightarrow \text{orden cero}$$

donde

k sensibilidad estática, $[S/E]$

τ constante de tiempo, s

Ejemplos: termómetro de bulbo.

Observaciones:

Ahora la lectura de salida tiene un defasamiento con respecto a la de entrada. Este defasamiento depende de la constante de tiempo, si éste es cero el instrumento responderá instantaneamente o sea, como un instrumento de orden cero. Si la constante es grande, con respecto a la variación de la señal de entrada, tendrá mucha inercia y siempre irá atrasada lo que hace que la señal de salida pueda ser muy diferente a la de entrada.

Calibración: La sensibilidad estática se puede encontrar con una calibración estática, pues al no variar en el tiempo la ecuación (2) se reduce a la (1). La constante de tiempo se la señal puede encontrar como sigue. Se aplica una función escalón definida por:

$$E \begin{cases} = 0 & \text{para} \quad t \leq 0 \\ = E_1 & \text{para} \quad t > 0 \end{cases}$$

que físicamente se puede obtener variando la señal de cero hasta un valor conocido en forma súbita, por ejemplo: si la señal de entrada es una presión se puede introducir el sensor en un globo y reventar éste en cierto momento.

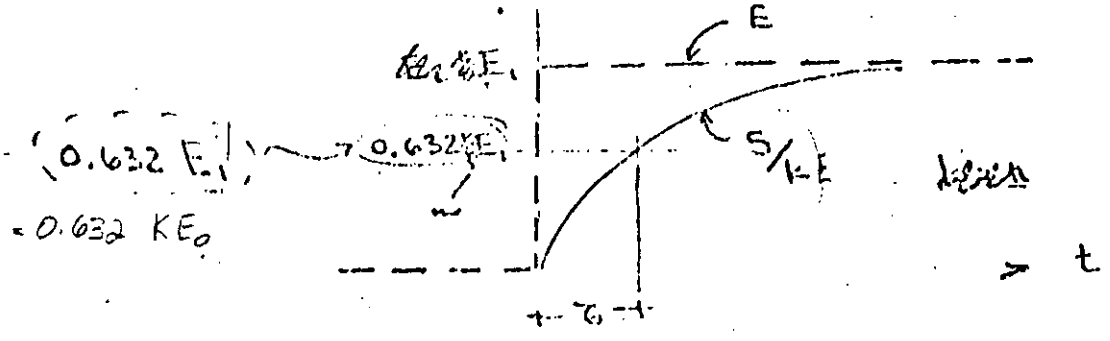


Fig 2. Estimación de la constante de tiempo. Sistema de primer orden

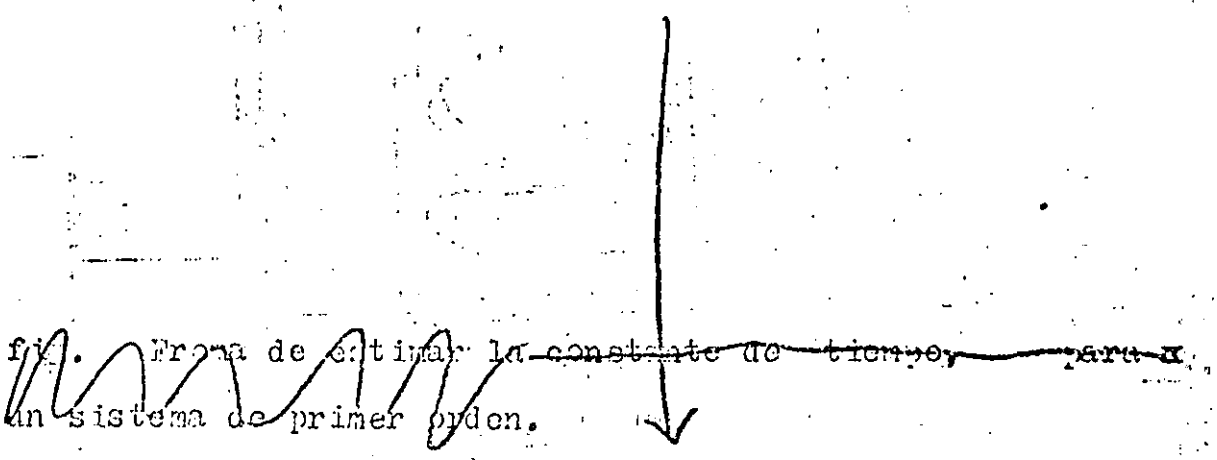


fig. Forma de estimar la constante de tiempo, para un sistema de primer orden.

La señal de salida seguirá la forma de la función escalón en forma suave hasta que llegue, después de cierto tiempo (la constante de tiempo, τ), al valor $S_k = E_1$, será cuando la señal S_k alcance un valor de 63.2 porciento el valor final, E_1.

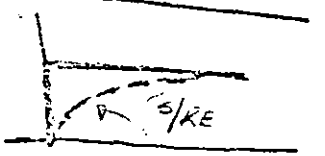
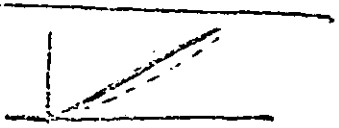
Para calibraciones más completas, o con señales senoidales, que son difíciles de simular físicamente, se sugiere ver la referencia 2.

Determinación de E.

Conocidas k y τ se puede encontrar la señal de entrada por medio de la ecuación (2). En la TABLA I se dan las soluciones para señales tipo y al final de la sección se comentarán las señales aleatorias.

TABLA I

Señales de salida para un instrumento de primer orden

Señal	Entrada (E)	Salida (S)	E(—), S(---)
Escalón	$E=0$ para $t \le t_0$ $E=E_0$ para $t \ge t_0$	$S = K E_0 [1 - e^{-t/\tau}]$	
Rampa	$E=0$ para $t \le t_0$ $E=\dot{E}t$ " $t \ge 0$	$S = K E_0 (\tau e^{-t/\tau} + t - \tau)$	
Senoide	$E = A_e \operatorname{sen} \omega t$	$\frac{A_s}{A_e} = \frac{K}{\sqrt{\omega^2\tau^2+1}}$, $\phi = \tan^{-1}(-\omega\tau)$	
Impulso	Si $E=0$ para $0 \le t \le T$ y $E = A/T$ para $0 \le t \le T$ Impulso y $T \to 0$	$S = \frac{KA}{\tau}[1 - e^{-t/\tau}]$	

Tesis

donde E_0, $\dot{E}$, A_e y τ constantes

A_e y ω amplitud y frecuencia de la señal de entrada senoidal

$\frac{A_s}{A_e}$ y ϕ amplitud y ángulo de fase de la señal de salida senoidal

K y τ_0 sensibilidad estática y constante de tiempo del

3.5 Instrumento de segundo orden.

Ecuación:

$$\frac{S}{E} = \frac{k}{\frac{D^2}{\omega_o^2} + \frac{2\zeta D}{\omega_o} + 1}$$

donde

ω_o frecuencia natural, rad/s

ζ amortiguamiento, adimensional

y las demás variables ya fueron definidas.

Ejemplos: dinamómetro de resorte, celdas de presión, ver fig.3.

Calibración

La k se encuentra con una calibración estática. La y la frecuencia natural y El amortiguamiento se pueden encontrar utilizando una función escalón y aplicando las fórmulas:

$$\zeta = \left[\frac{1}{\left(\frac{\pi}{\ln \frac{a}{kE_i}} \right)^2 + 1} \right]^{1/2}$$

$T = \frac{2\pi}{\omega}$; $\omega = 2\pi f$

$\omega = \frac{2\pi}{T}$

$$\omega_o = \frac{2\pi}{T\sqrt{1-\zeta^2}}$$

donde a, ζ y T se muestran en la fig 4.

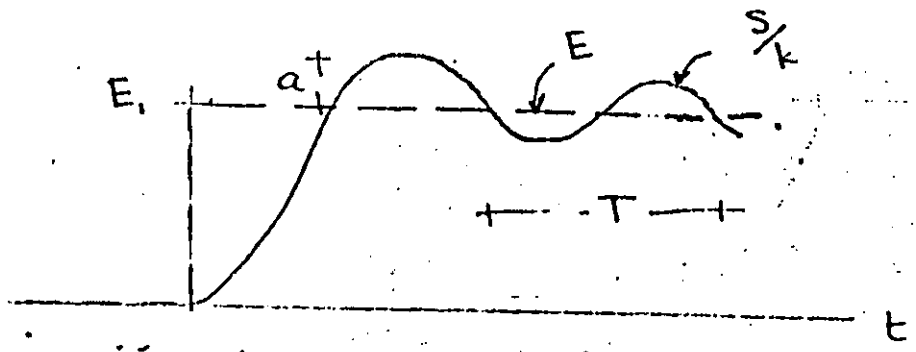


Fig 4. Estimación de a y T. Sistema de segundo orden

También se pueden encontrar los parámetros del sistema de medición usando señales de entrada senoidales, lo que es caro y tardado, o señales impulso lo que se puede ver en la referencia 2.

Determinación de E.

Una vez determinados los parámetros del sistema se relaciona la señal de salida, S, con la de entrada, E, por medio de la ecuación (3). En las TABLAS II a V se dan las soluciones para las señales tipo.

TABLA DE [illegible]

Señales de salida para un instrumento de segundo orden

	Entrada (E)	Salida (S)	E(—) S(---)
Escalón	E=0 para t≤0 E=E₀ para t>0	S=kE₀	
Rampa			
Senoide			
Impulso			

Copiar del Desplegable

8.7.3 Medición de gastos.

Para medir gastos generalmente se usan vertedores de diversas formas. Cuando el gasto es muy pequeño se pueden usar rotámetros o simplemente se mide el volumen en un cierto tiempo.

Se recomienda utilizar vertedores de cresta delgada. Rectangulares para gastos entre ? y m³/s y triangulares para gastos menores de. ?

Los vertedores rectangulares deben tener suficiente aireación para que el chorro despegue libremente de la pared. Es importante colocarlos verticalmente y con la cresta horizontal.

Hay varios tipos de vertedores ver fig. 12, por ejemplo: sin contracciones laterales donde la cresta es del ancho del canal de llegada y cuyo gasto se determina con la fórmula

Con contracciones laterales

donde

Una variante de este tipo de vertedor es el Cipolleti, (fig. 13) donde los lados del vertedor se abren para reducir el efecto de la contracción. En este caso el gasto está dado por:

Se pueden utilizar vertedores triángulares con diferente ángulo central, usualmente se utiliza un ángulo de 90°. En la tabla I se da el gasto correspondiente para diferentes tirantes. Para ángulos diferentes a 90° se utiliza la fórumula de

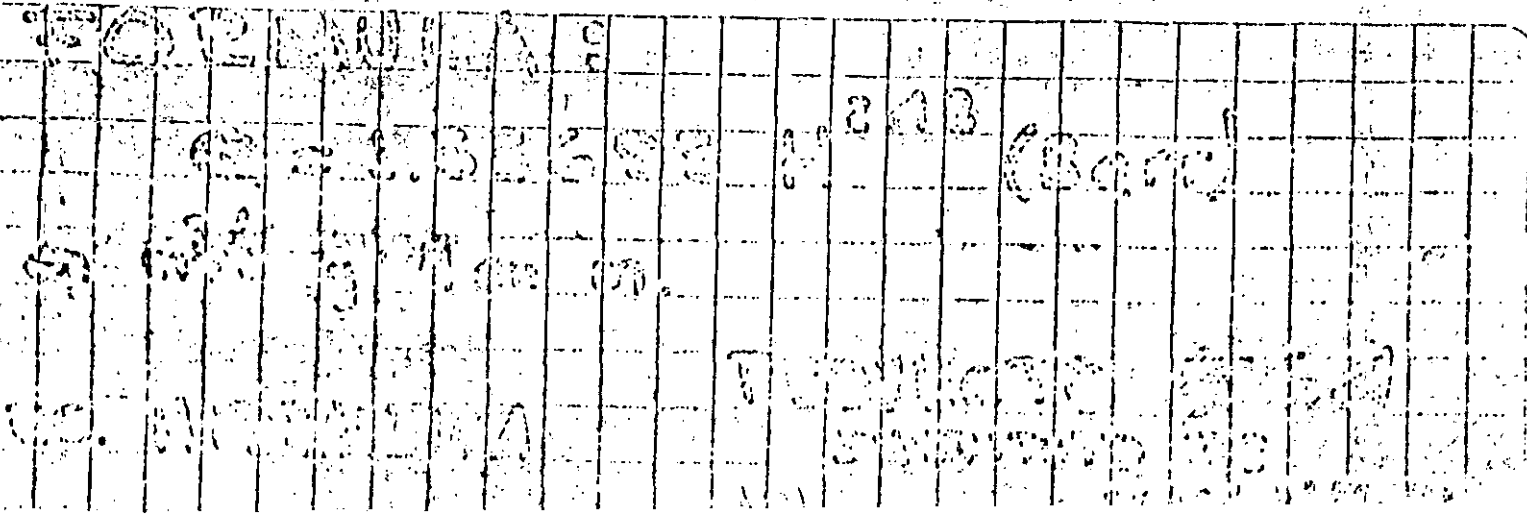
donde

θ angulo central

El tirante se debe medir a partir de la cresta, en vertedores rectangulares, o del vértice del triángulo, y en un punto alejado cuando menos 3 veces el tirante para salirse de la zona donde el agua se abate al adquirir carga de velocidad .

Si la instalación es adecuada se puede esperar una aproximación en la determinación del gasto de alrededor de 3 por ciento (ref.)

0.0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0
0.082	0.094	0.106	0.118	0.130	0.142	0.154	0.166	0.178	0
0.224	0.245	0.266	0.286	0.307	0.328	0.354	0.379	0.405	0
0.456	0.487	0.518	0.549	0.580	0.611	0.648	0.684	0.721	
0.794	0.836	0.878	0.921	0.963	1.006	1.054	1.103	1.152	
1.25	1.30	1.36	1.41	1.47	1.52	1.58	1.64	1.71	1
1.83	1.90	1.97	2.03	2.10	2.17	2.25	2.32	2.40	2
2.55	2.63	2.71	2.80	2.88	2.96	3.05	3.14	3.23	3
3.41	3.51	3.61	3.70	3.80	3.90	4.01	4.11	4.22	4
4.43	4.54	4.66	4.77	4.89	5.00	5.12	5.24	5.37	5
5.61	5.74	5.87	6.00	6.13	6.26	6.40	6.54	6.68	6
6.96	7.11	7.26	7.40	7.55	7.70	7.86	8.02	8.17	8
8.49	8.66	8.82	8.99	9.15	9.32	9.50	9.67	9.85	10
10.20	10.39	10.57	10.76	10.94	11.13	11.32	11.52	11.71	11
12.10	12.31	12.51	12.72	12.92	13.13	13.34	13.56	13.77	13
14.20	14.43	14.65	14.88	15.10	15.33	15.57	15.80	16.04	16
16.51	16.76	17.00	17.25	17.49	17.74	18.00	18.25	18.51	18
19.02	19.26	19.56	19.82	20.09	20.36	20.64	20.92	21.19	21
21.75	22.04	22.33	22.61	22.90	23.19	23.49	23.79	24.10	24
24.70	25.01	25.32	25.64	25.95	26.26	26.58	26.90	27.23	27
27.87	28.21	28.54	28.88	29.21	29.55	29.90	30.25	30.59	30
31.29	31.65	32.01	32.36	32.72	33.08	33.45	33.82	34.20	34
34.99	35.32	35.70	36.09	36.47	36.85	37.24	37.64	38.03	38
38.82	39.23	39.64	40.04	40.45	40.86	41.28	41.70	42.11	42
42.95	43.38	43.82	44.25	44.69	45.12	45.56	46.01	46.45	46
47.34	47.80	48.26	48.71	49.17	49.63	50.10	50.57	51.05	51
51.99	52.48	52.96	53.45	53.93	54.42	54.91	55.41	55.90	56
56.89	57.40	57.91	58.42	58.93	59.44	59.97	60.49	61.02	61
62.07	62.61	63.14	63.68	64.21	64.75	65.30	65.85	66.41	66
67.51	68.08	68.64	69.21	69.77	70.34	70.92	71.50	72.07	7
73.23	73.82	74.41	75.01	75.60	76.19	76.80	77.41	78.01	78
79.23	79.85	80.47	81.09	81.71	82.33	82.97	83.60	84.24	84
85.51	86.16	86.81	87.46	88.11	88.76	89.42	90.09	90.75	91
92.08	92.76	93.44	94.12	94.80	95.48	96.17	96.87	97.56	98
98.95	99.66	100.4	101.1	101.8	102.49	103.2	103.9	104.7	105
106.1	106.8	107.6	108.3	109.1	109.8	110.6	111.3	112.1	112
113.6	114.4	115.1	115.9	116.6	117.4	118.2	119.0	119.7	120
121.3	122.1	122.9	123.7	124.5	125.3	126.1	126.9	127.8	128
129.4	130.2	131.1	131.9	132.8	133.6	134.4	135.3	136.1	137
137.8	138.7	139.5	140.4	141.2	142.1	143.0	143.9	144.7	145
146.5	147.4	148.3	149.2	150.1	151.0	151.9	152.8	153.7	154
155.5	156.4	157.3	158.3	159.2	160.1	161.1	162.0	163.0	163
164.9	165.9	166.8	167.8	168.7	169.7	170.7	171.7	172.6	173
174.6	175.6	176.6	177.5	178.5	179.5	180.5	181.5	182.6	183
184.0									



IX PROCEDIMIENTO EXPERIMENTAL

Al hacer un experimento, o serie de ensayos con el mismo propósito, es importante seguir los siguientes pasos:

1. Planear el experimento
2. Controlar las variables
3. Tomar datos adecuados
4. Detectar los problemas
5. Interpretar los resultados
6. Reporte

En la planeación, además de considerar la factibilidad y el costo, se decide un programa que con mínimo de esfuerzo dé un máximo de información.
Para esto, es esencial tener control sobre las variables independientes, o sea poderles asignar valores prefijados con cierta aproximación y mantener esos valores durante el ensayo.

Para lograrlo, así como para cuantificar la o las variables dependientes, se necesita tomar datos suficientemente exactos haciendo mediciones que sean confiables y precisas.

Aun midiendo con todo cuidddo, y sobre todo en experimentos no tradicionales, es común que se presentan problemas y errores los que hay que detectar para corregirlos o tomarlos en cuenta. Finalmente, los datos obtenidos se interpretan para poder llegar a resultados provechosos.

viene de 7

Para ilustrar la propagación de errores se verán dos ejemplos:

Ejemplo 1

¿Con qué error se estará estimando la n de Manning?, si se usan los valores:

$$R = 2 \pm 0.05\ m$$

$$S = 0.001 \pm 0.00003$$

$$V = 0.40 \pm 0.05\ m/s$$

y la fórmula es

$$n = \frac{R^{2/3} S^{1/2}}{V}$$

$$n = 0.125$$

$$\frac{\partial n}{\partial V} = -\frac{R^{2/3} S^{1/2}}{V} = -0.3137$$

$$\frac{\partial n}{\partial R} = \frac{2}{3}\frac{S^{1/2}}{V R^{1/3}} = 0.04183$$

$$\frac{\partial n}{\partial S} = \frac{1}{2}\frac{R^{2/3}}{V S^{1/2}} = 62.7476$$

por lo que aplicando la ecuación (4) queda:

$$e_n^2 = (-0.3137)^2 (0.05)^2 + (0.04183)^2 (0.05)^2 + (62.7476)^2 (0.00003)^2$$

$$e_n = \pm 0.016$$

así que n será:

$$n = 0.125 \pm 0.016$$

Nota: La e puede ser cualquiera de los índices, como: ϕ ó s ———, sólo debe ser el mismo en todas las variables.

Ejemplo 2.

Encontar la precisión con que se está midiendo el gasto, Q, en un vertedor de cresta delgada con contracciones laterales si la fórmula es:

$$Q = 1.8\, L\, H^{3/2}$$

donde

H carga del vertedor, en m

L longitud efectiva de la cresta dada por $L = L' - 0.2H$, y

L' Longitud real, en m

y las mediciones son:

$L' = 2.50 \pm 0.0005$ m

$H = 0.50 \pm 0.001$ m

Es más sencillo mantener separadas las dos ecuaciones, así que:

$$\frac{\partial L}{\partial L'} = 1 \qquad e_L = \left(\frac{\partial L}{\partial L'}\right)^2 (e_{L'})^2 + \left(\frac{\partial L}{\partial H}\right)^2 (e_H)^2$$

$$\frac{\partial L}{\partial H} = -0.2$$

$$e_L^2 = 1^2 (0.0005)^2 + (-0.2)^2 (0.001)^2$$

$$e_L = \pm 0.000539$$

$$\frac{\partial Q}{\partial L} = 1.8\, H^{3/2}$$

$$\frac{\partial Q}{\partial H} = \frac{3}{2} \times 1.8 L H^{1/2}$$

por lo que $\left(\frac{\partial Q}{\partial L}\right)^2 e_L^2$

$$e_Q^2 = 3.24(0.50)^3(0.000539)^2 + 7.29(2.40)^2(0.50)(0.009)^2$$

$$e_Q = \pm 0.0046$$

Así que el gasto estará dado con la siguientes precisión.

$$Q = 1.53 \pm 0.0046 \ m^3/s$$

Criterio de Chauvenet.

Para rechazar puntos disparados, muchas veces consecuencias de errores espurios, se puede utilizar el criterio de Chauvenet que consiste en lo siguiente: Se rechazan los puntos con probabilidad de ocurrir menor que $1/2n$, donde es el número de puntos.

Para aplicar el criterio anterior hay que tomar en cuenta lo siguiente;

1. Se debe aplicar el criterio una sola vez.
2. No se debe aplicar en puntos extremos de la medición pues a lo mejor lo que pasa es que la función cambia de naturaleza, o sea sólo se recomienda para puntos medios.

Se verán un ejemplo para ilustrar su aplicación.

Ejemplo 1:

Se tienen 8 mediciones de una misma velocidad.

2.3 m/s	2.4
2.2	2.2
2.3	2.6
2.0	2.2

¿Se debe rechazar algún punto?

1. Se encuentra la media, la desviación estandar y 1/2 n.

$$\bar{x} = 2.275 \quad , \quad s = 0.175 \quad , \quad 1/2n = 0.0625$$

2. De la tabla de probabilidades se encuentra la distancia desde la media que corresponde a una probabilidad normal de

$$1 - 0.0625 = 0.9375$$

que es 1.869

como $s = 0.175$, la distancia será $0.175 \times 1.869 = 0.327$

por lo que se rechazarán los puntos que caigan fuera del rango

$$2.275 + 0.327 = 2.602$$

$$2.275 - 0.327 = 1.948$$

Como todos los valores están comprendidos en ese intervalo, no se rechaza ninguno.

Si los puntos experimentales siguen una cierta ley, primero se les ajusta alguna recta o curva con mínimos cuadrados, se encuentra la desviación estandar con respecto a ella y finalmente la franja fuera de la cual se rechazarán los puntos. Una forma gráfica de hacer ésto es trazar una faja que comprenda la mitad de los puntos, lo que corresponde a $2/3$

Después, con la ecuación

$$s = 1.482 \, \delta$$

Se encuentra la desviación estándar y de ahí se procede como en el ejemplo. (a hoja 7)

13

ene de la 7

mejorarla al doble hay que hacer 16 mediciones, una vez que éstas pasan de cierto número es inestable o muy tardado tratar de mejorar la calidad de los resultados.

En modelos hidráulicos se puede considerar como un buen resultado aquel que tiene un error menor o igual a de 5% y aceptable si es menor de 10% así que, en general, no es necesario obtener un gran número de puntos que pueden, incluso, hacer confuso el resultado.

Para determinar el espaciamiento de los puntos experimentales hay que tomar en cuenta dos criterios:

1. La precisión que se requiere en las diferentes regiones del experimento.
2. La forma de la función experimental.

El primero se refiere a que en ciertas regiones, por ejemplo: para valores bajos de las variables, la dispersión de los datos puede ser, proporcionalmente, mucho mayor que para valores altos; por lo que será necesario hacer más mediciones en esa zona. Para cuantificar lo anterior se pueden usar las técnicas expuestas en el análisis de errores.

El segundo criterio considera la forma de la función que se desee determinar. Si ésta es complicada, con varios quiebres, será necesario tener más puntos. Si es sencilla, por ejemplo: una recta o una curva suave, con relativamente pocos puntos se puede determinar.

Como muchas veces el experimentador tiene una idea de como va a ser la curva aún antes de iniciar los ensayos, puede utilizar un espaciamiento tal que los puntos queden, más o menos, a igual distancia unos de otros sobre la curva, verf Fig.4- .

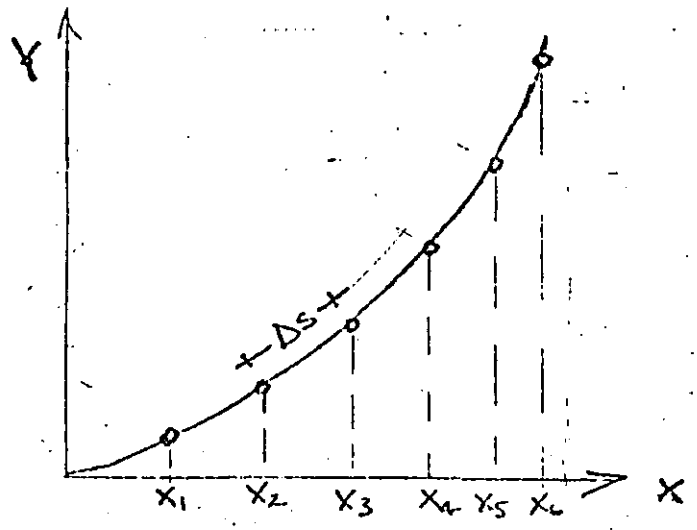


Fig. 4- Igual espaciamiento a lo largo de la curva.

Para hacer ésto se puede trazar la curva en forma aproximada entre los valores máximos y mínimos que se piense medir y espaciar los puntos a ojo. Si se desconoce la curva se pueden medir unos cuantos valores en todo el rango y con ellos trazarla en forma preliminar para poder hacer el espaciamiento.

Orden o secuencia de la medición

La secuencia más recomendables es la aleatoria pues aunque tal vez con ella se tenga mayor dispersión, está será alrededor del valor real y se evitarán las tendencias instrumentales, naturales y humanas que puede haber.

Como una forma usual de presentar resultados es por medio de informes, a continuación se dará una guía adaptada de un boletín de la Sección de Servicios Editoriales del Instituto de Ingeniería de la UNAM.

PORTADA

Título

Debe enunciar lo esencial del tema estudiado; se procurará que sea breve.

Nombres de los autores

El orden lo fijarán los autores, de acuerdo con su participación en el trabajo.

INDICE

Aparecerán los títulos de los capítulos y subcapítulos:

RESUMEN
INTRODUCCION
1. INTRODUCCION
1.1 Significado y propósito del estudio

RESUMEN

Se presenta a continuación del índice y sintetiza el informe. El título enuncia el tema; el resumen lo desarrolla en forma condensada. Conviene traducirlo al ingles (ABSTRACT).

NOMENCLATURA

Consiste en una lista de los símbolos utilizados en el

texto. Conviene apuntarlos alfabéticamente en el siguiente orden:

1. Caracteres latinos:
   - Mayúsculas
   - Mayúsculas con subíndice letra
   - Mayúsculas con subíndice número
   - Minúsculas con subíndice letra
   - Minúsculas con subíndice número
2. Caracteres griego:
   - Igual que los anteriores
3. Números

Las unidades se expresan al final, por ejemplo:

Q_t gasto, m^3/s

t tirante, m

2. ~~Símbolos griegos:~~

~~Igual que las unidades~~

3. ~~Números~~

~~Las unidades se expresan al final, por ejemplo:~~

Q gasto, en m^3/s

t tirante, etc.

4a.- X

INTRODUCCION

En esta parte se hablará de las necesidades que dieron origen al trabajo; de las finalidades y propósitos del mismo; del programa seguido y de los métodos empleados.

ANTECEDENTES

Aquí se expondrán los trabajos previos relacionados con el tema y las consideraciones que se hicieron para realizar el estudio.

TEXTO

Esta formado por varios capítulos que dependen de cada trabajo en particular, como:

3. EQUIPO EMPLEADO
4. PRUEBAS REALIZADAS
5. RESULTADOS
6. CONCLUSIONES
7. RECOMENDACIONES

RECONOCIMIENTO; REFERENCIAS, TABLAS Y FIGURAS

Al final se señala qué personas o instituciones han colaborado para realizar el trabajo; después aparecen las referencias y, por último, tablas y figuras.

Formato

Las distintas secciones del informe van numeradas. Si hay subdivisiones éstas se distinguirán con cifras decimales añadidas al número del capítulo:

1. INTRODUCCION

2. ANTECEDENTES

2.1 Descripción del fenómeno

2.1.1 Gráficas carga-gasto

Para profundizar más en el tema y, de paso, dar un ejemplo de cómo se deben poner las referencias, se recomienda: Zubizarreta A., "La aventura del trabajo intelectual", Fondo Educativo Interamericano, S. A., México (1969)

XI Ejemplos

A continuación se presentarán cuatro ejemplos de problemas resueltos con diferentes tipos de modelos, que ilustran algunas de las ideas expuestas en los capítulos anteriores y muestran su aplicabilidad.

CAPITULO 3. PROYECTO HIDRAULICO EL CARACOL

Antecedentes: El P.H. El Caracol se encuentra en el Estado de Guerrero sobre el río Balsas a 60 kms aguas abajo del punto Mezcala. El area de su cuenca es 40,137 km² con un caudal medio anual de 240 m³/s. La cortina será de materiales graduados con una altura máxima de 134 m y una longitud de corona de 345 m. La obra de toma está en la margen derecha y tendrá 3 turbinas tipo Francis de eje vertical conectadas a generadores con capacidad nominal de 211,000 KVA. La obra de excedencia, también en la margen derecha, consiste de dos canales trapeciales de 52.30 m de ancho en su base y 500 m de longitud terminado en una cubeta deflectora.

Se hicieron estudios en modelo para probar la localización general de las diferentes componentes de la obra, una vez definido ésto, se procedieron a estudiar los detalles del diseño.

Tipo de modelo:

Es un modelo a superficie libre donde domina la influencia de la geometría y del Número de Froude, y la rugosidad es poco importante.

Modelo: Se construyó un modelo no distorsionado escala 1:50 en el laboratorio de Hidráulica de la Comisión Federal de Electricidad en Chicoasén, Chiapas. Se reprodujo en forma fija la cortina, parte del vaso, la obra de exce-

dencia y un tramo del cauce, ver fig. 1. Aguas abajo de la obra de excedencia y un tramo de cauce se ~~[illegible]~~ reprodujo con fondo móvil para estimar la socavación debida al agua y la formación de barras con el material removido por el agua.

Los límites del modelo se construyeron con paredes de piedra y tabique y la topografía con arcilla del lugar recubierta con una capa de concreto de 6 cm. de espesor.

El gasto máximo disponible es de $1m^3/s$, suficiente para reproducir la avenida máxima probable del prototipo de $17750\ m^3/s$. Se puede simular el nivel en el río con una compuerta al final del modelo, se colocaron un limnímetro de gancho para medir el nivel en el vaso y otro móvil, que se puede deslizar a lo largo de la obra de excedencia, de punta y un tubo de Pitot para medir las velocidades.

Ensayos realizados;

Se probaron dos alternatativas de localización de la obra de excedencia que fueron desechadas adoptándose la opción de un vertedor dividido por un muro central.

Después se estudiaron detalles del diseño. Se probó el canal de llamada de la obra de excedencia y se decidió poner un muro central que guíe el escurrimiento, sobre todo cuando la operación de las compuertas es asimétrico. Se vió que el agua para gastos hasta de $6000\ m^3/s$ por

c nal, no rebase los muros laterales ni el central.

Se estudió la estructura de control, se cambió la geometría de las pilas para reducir las perturbaciones de aguas
~~Rampantitta~~ abajo, se modificó el perfil de la plantilla de la rápida para uniformizar al flujo y se probó el funcionamiento hidráulico de la obra en general para diferentes aberturas de las compuertas. Se ensayaron diferentes elevaciones y ángulos de lanzamiento del deflector y se escogió la que disipa mejor la energía del chorro sin que éste choque con la margen contraria del río ni que existan corrientes fuertes pegadas al deflector.

Mediciones en el modelo: se midió el gasto de despegue con diferentes aberturas de la compuerta, se encontró la curva de gastos-elevaciones tanto con el vertedor libre como con las compuertas parcialmente abiertas y se midieron los tirantes y las velocidades en diversas secciones del vertedor, Fig. 3.

En la obra de toma se hicieron varias modificaciones para reducir los vórtices que se forman sobre la toma (ref.).

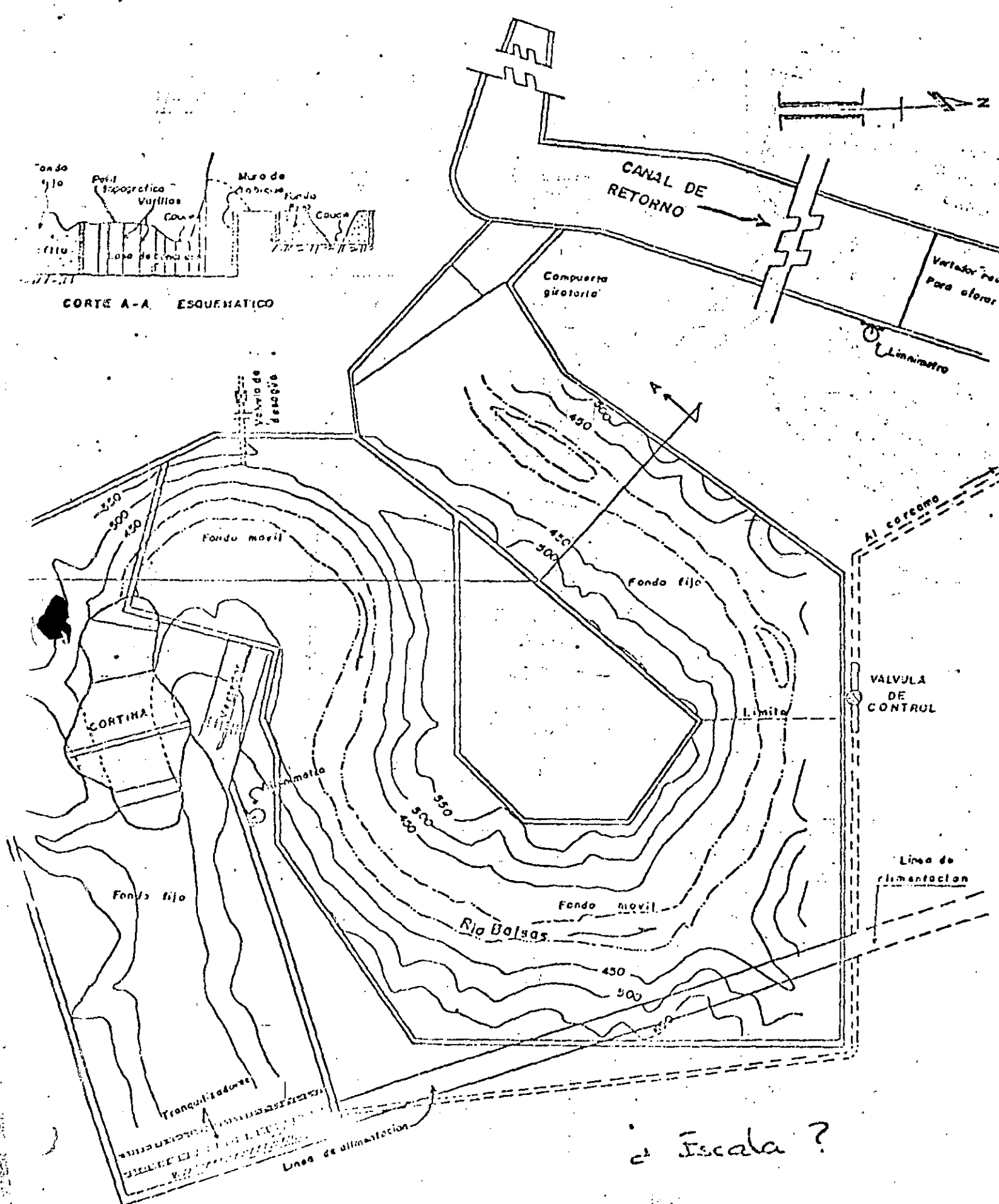


FIG. V-IV-2 PLANTA DEL MODELO

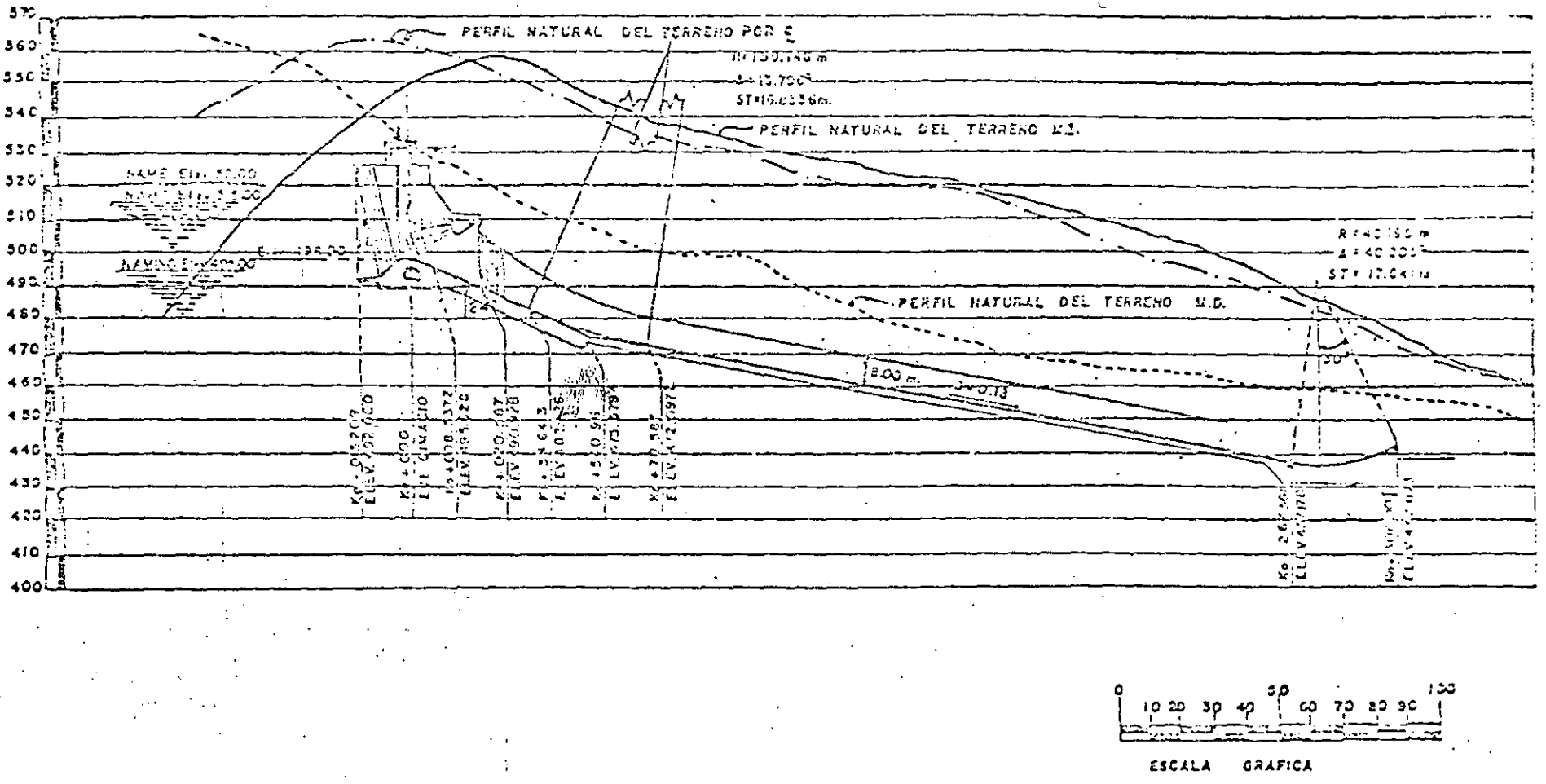


FIG. II. 2 PERFIL LONGITUDINAL DEL VERTEDOR

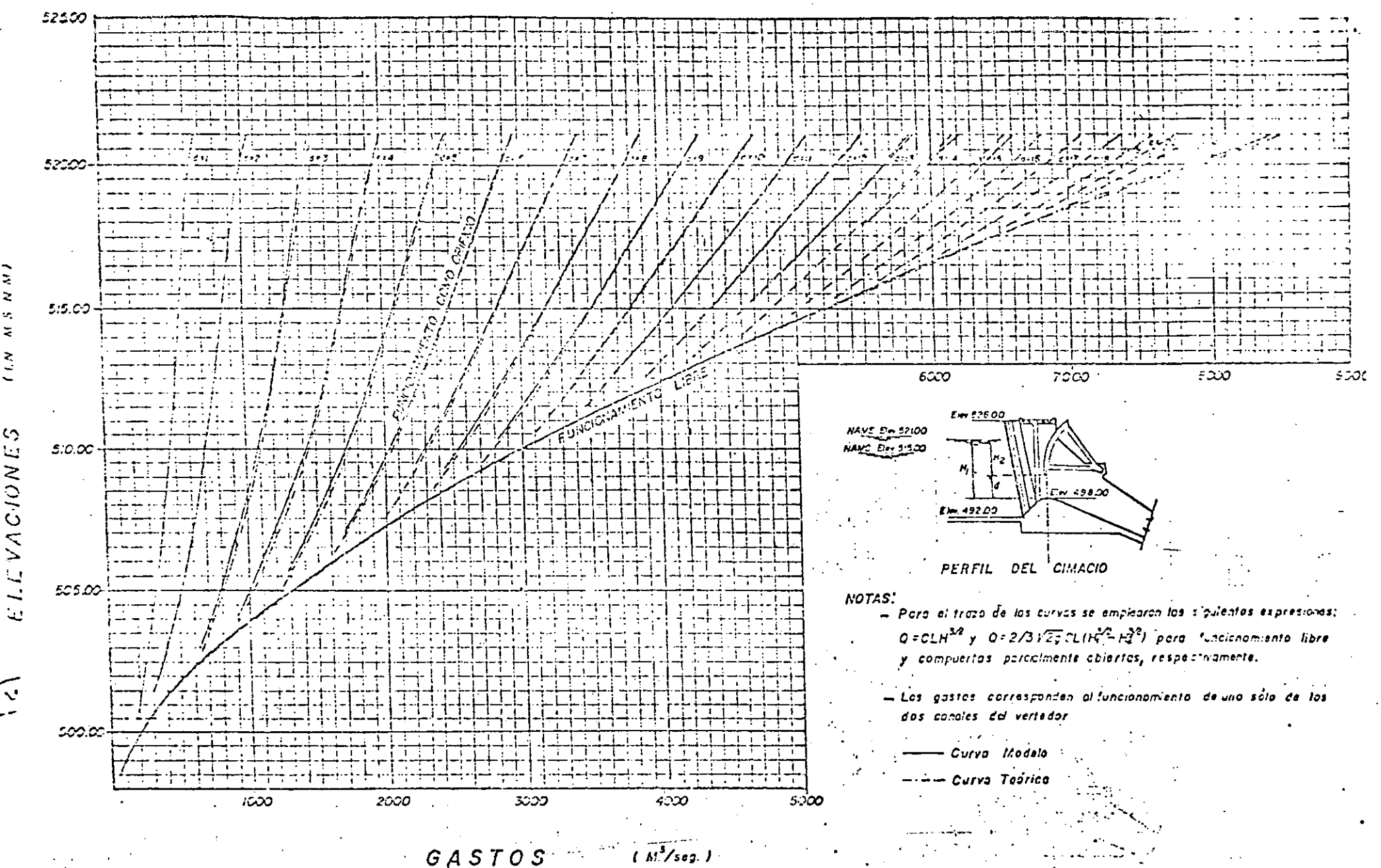


FIG. 3 CURVA DE GASTOS TEORICA-EXPERIMENTAL

12.5 Densidad del agua a la presión atmosférica

Temperatura °C	Densidad kg/m3
0	999.97
10	999.33
20	998.42
30	995.95
60	983.31
100	958.48

12.6 Densidad de algunos líquidos a 20°C de temperatura

Líquido	Densidad, kg/m3
Mercurio	13553
Glicerina	1257.57
Agua de mar	1027.20
Gasolina	720.90
Aceite lubricante	897.12

12.7 Densidad del aire a presión atmosférica

Temperatura, °C	Densidad, kg/m3
-10	1.34
0	1.29
10	1.24
20	1.20
40	1.12
60	1.06
80	0.99
100	0.94

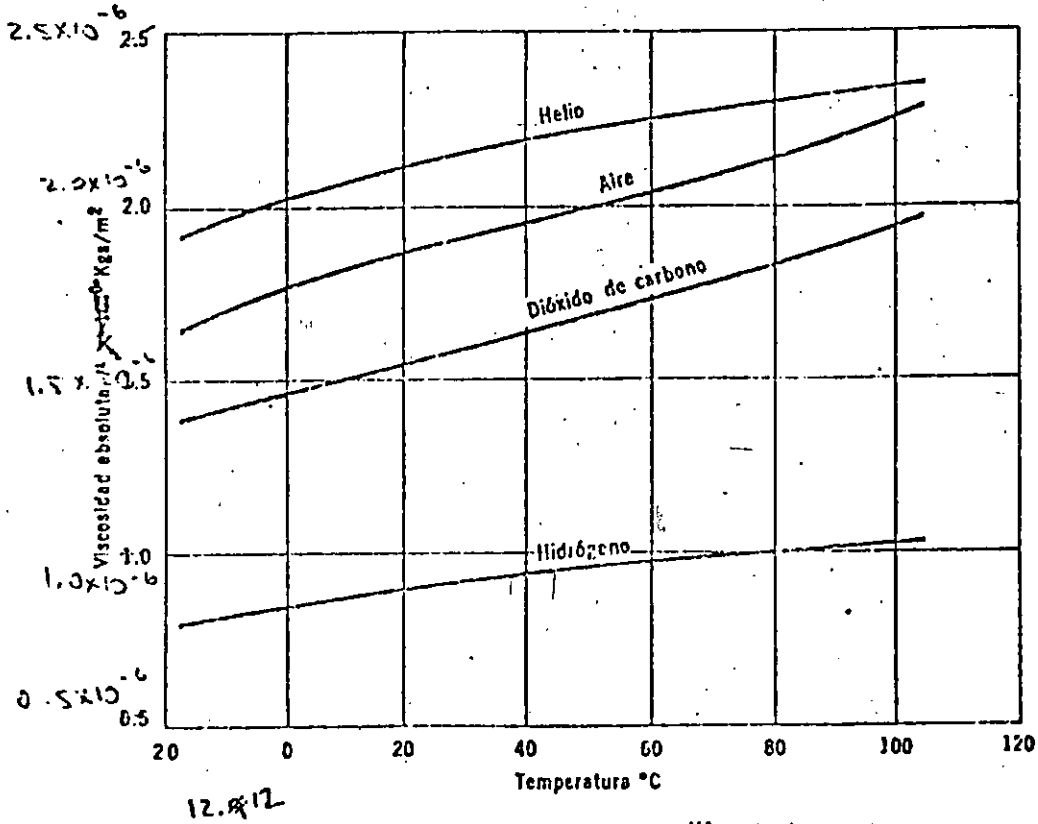
12.8 Densidad de algunos gases a 20°C de Temperatura y presión atmosférica

Gas	Densidad, kg/m3
Anhidrido carbónico	1.8359
Oxígeno	1.3297
Nitrógeno	1.1631
Metano	0.6664

12.9 Módulo de elasticidad del agua y presión de vaporización

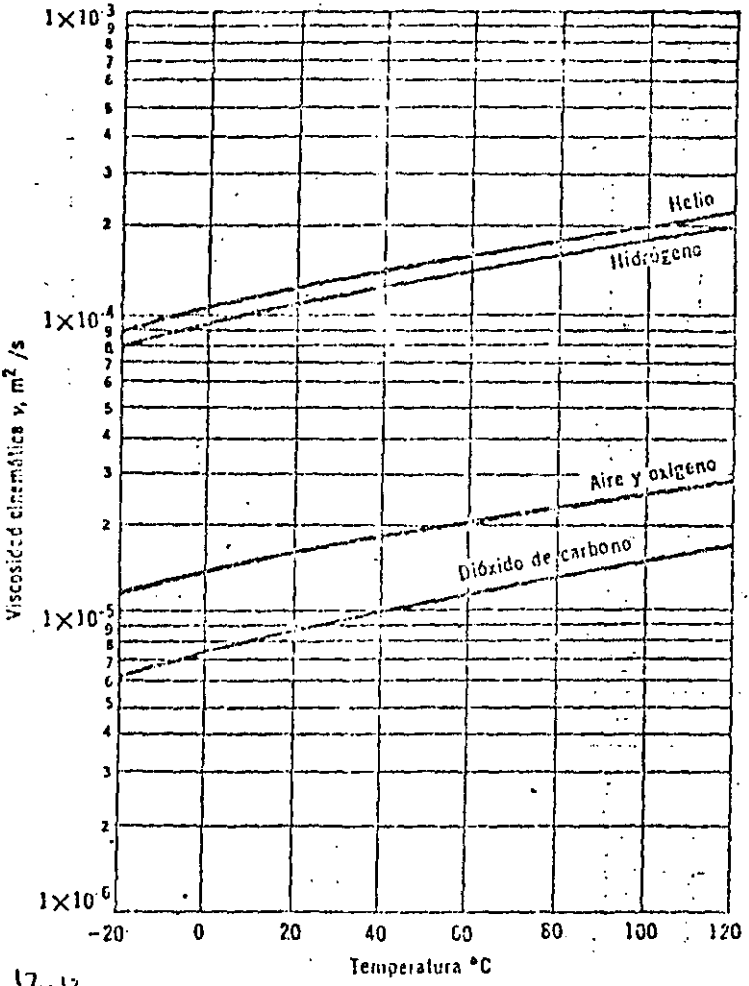
Temperatura	Módulo de elasticidad kg/cm^2	Presión de vaporización m
0	20500	-10.271
20.	22500	-10.095
50	23300	- 9.145
100	21000	0

12.10 Módulo de elasticidad de la lucita



12.11

Fig. 12 Viscosidad absoluta de algunos gases a diferentes temperaturas.



12.13

Fig. 13 Viscosidad cinemática de algunos gases a diferentes temperaturas.

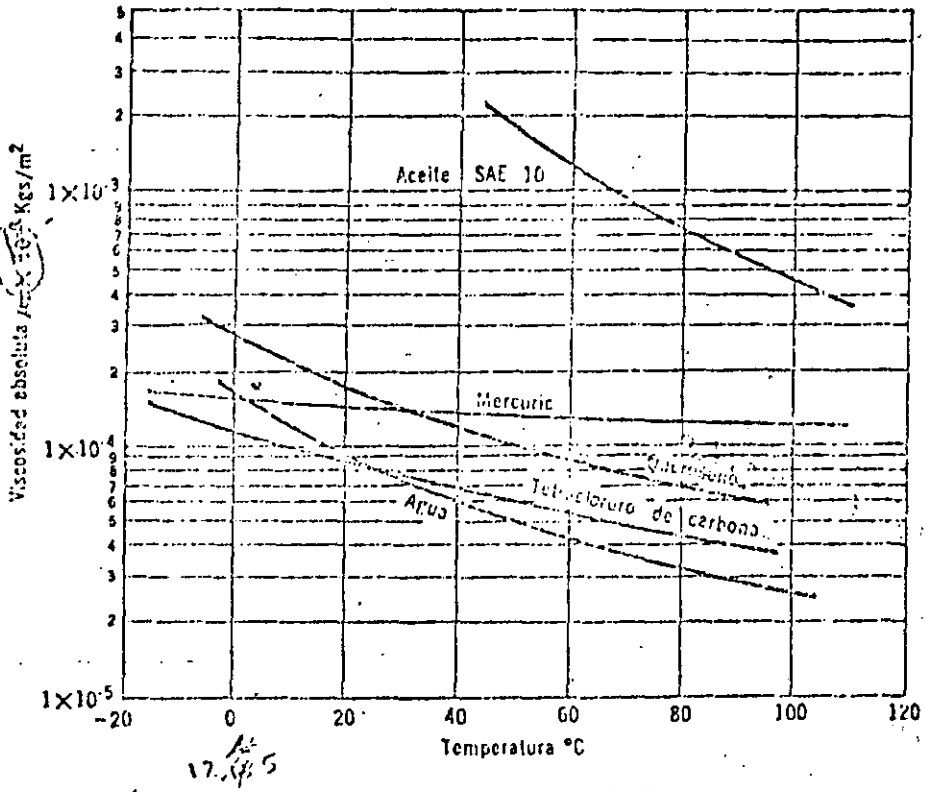


Fig. 1-4 Viscosidad absoluta de algunos líquidos.

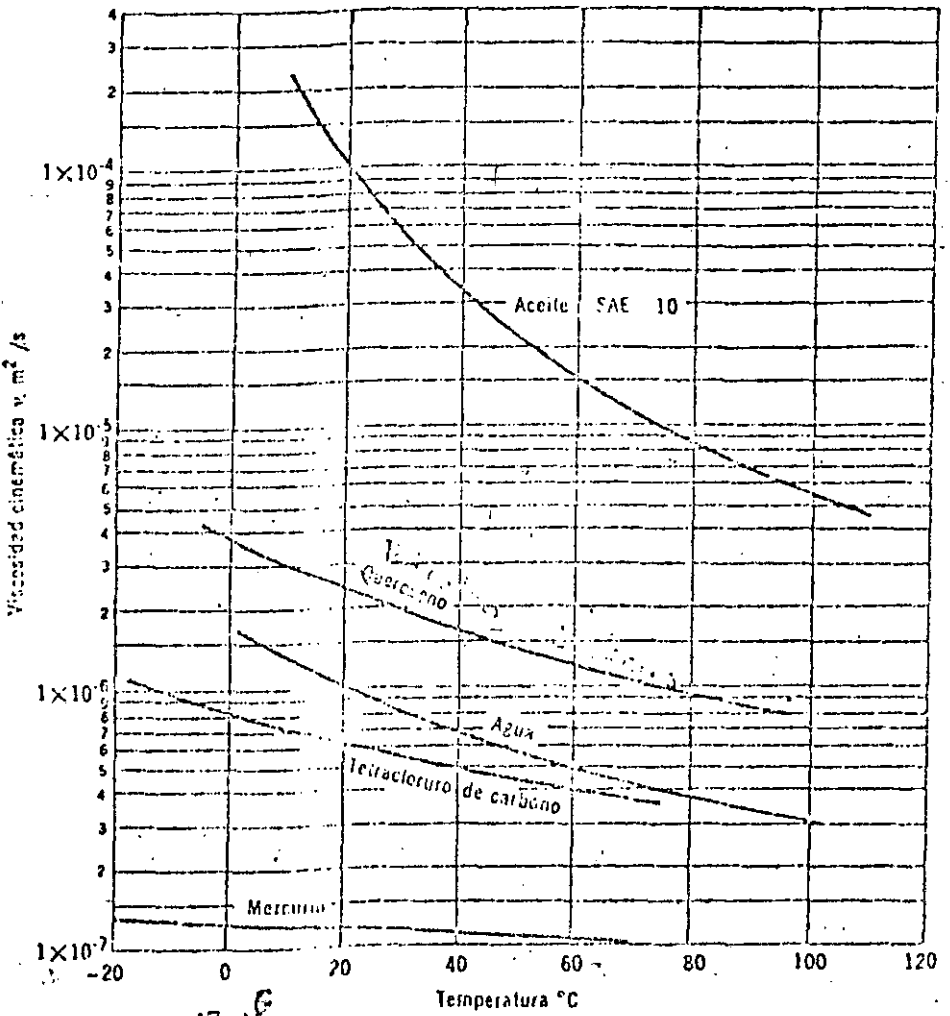


Fig. 1.5 Viscosidad cinemática de algunos líquidos.

INGENIERIA MARITIMA
DEL 1o. DE JULIO AL 30 DE NOVIEMBRE
1985
MEXICO, D.F

THE SIZE OF MAN

ING. VICTOR BURGUET

AMERICAN SCIENTIST, 56, 4, pp. 400–413, 1968

THE SIZE OF MAN

By F. W. WENT

IN THE following article, I want to show how important his physical size is for man and how many of his attainments, such as the development of a technology, were possible only because of his specific size. In the course of these considerations, it will become clear that many physiological and mechanical processes have grave limitations in relation to size. We cannot indiscriminately magnify or reduce processes from the microscopic to the macroscopic level or the reverse without physiological distortion.

To emphasize the significance of size, frequent comparisons will be made between man and ant. In spite of the enormous difference in size (linearly, three orders of magnitude; on the basis of mass, 8 orders of magnitude), both have developed highly complex social systems, and ants (like bees, termites, and other social insects) have also effective communications systems. Whereas the social insects developed millions of years ago to their high degree of specialization, they never developed a technology with which they could have dominated the earth. On the other hand once he attained a certain degree of social differentiation, man explosively developed to his present state in the course of only 10,000 years. Whereas I cannot possibly attempt to discuss all contributing factors to man's emergence, I want to point out that his physical size was a dominating factor.

The relative importance of size was completely unknown until man started to probe into the microscopic world with lenses and microscopes.

When, in the second half of the seventeenth century, man looked for the first time through a microscope, he suddenly became aware of an entirely new world—until then almost unsuspected, the micro-world. The genius of these first microscopists made them realize what an important step they were taking, and the showmanship of some of them, notably Leeuwenhoek, brought the significance of this new world also to the attention of nonscientists. The common man, and royalty, flocked to the door of Leeuwenhoek's simple house in Delft to catch a glimpse of this new world, so full of unsuspected wonders. Suddenly, highly important principles, such as that of preformation and *generatio spontanea*, could be approached visually and experimentally. Curiously enough the basis not only for new understanding but also for a schism was laid.

One of the first men who speculated on the remarkable possibilities which magnification or diminution of physical dimensions provides was Jonathan Swift, who, in *Gulliver's Travels*, drew some conclusions as to what dwarfs and giants would really look like, and what sociological

consequences size would have. Some time ago Florence Moog (*Scientific American*, November 1948) showed that Swift was a "bad biologist," or Gulliver a "poor liar." She showed that a linear reduction in size would carry with it a reduction in the number of brain cells, and hence a reduction in intellectual capacity in Lilliputians, whereas the enormous Brobdingnagians were physically impossible; they could have had physical reality only if their necks and legs had been short and thick. These 90-ton monsters could never have walked on dry land, nor could their tremendous weight have been carried on proportionately-sized feet.

Even though Swift, in his phantasy, committed a number of physical errors, because he was not sufficiently aware of the fact that some physical properties of a body are proportional to the linear dimensions (height), whereas others vary with the third power of linear size (such as weight and cell number), yet he surpassed his medieval predecessors in many respects and drew a number of excellent conclusions, bringing both giants and dwarfs close to physical reality.

Since I want to make an argument, that between the macroscopic world (visible by the naked eye) and the microworld (of microscopic dimensions) a definite discontinuity exists, we first will consider whether the opposite view of a gradual transition has any merit. For instance, the microscopic embryo already has a number of features which the young or mature individual shows; sand grains under a microscope are like rocks, and a copepod seems nothing more than a miniature shrimp. And yet, it is curious that commonly there is not just a quantitative, but also a qualitative difference where we cross the border between macro- and microworld. For instance, all animals in the macro-world have hearts or more primitive means of circulation of body fluids. On the other hand, the cell, the unit of life activity, is almost without exception a microscopic structure. Similarly, the legs of animals, and walking, are achievements of the macro-world, whereas chemical cell reactions can be understood only on the molecular level.

The specific size of each organism is partly determined by its life habits. An elephant who eats branches (which are the food storage organs of many trees) must have enormous teeth to grind them, whereas harvester ants, living on tiny seeds gleaned from the soil, could not possibly be large and still collect enough food. Similarly parasites must be smaller than the organisms they grow on.

In spite of the enormous differences in size between species, cells of all organisms are equal in size and have an average volume of $10^3\mu^3$ in animals and $10^5\mu^3$ in plants. Therefore, size differences in plants and animals are almost exclusively due to differences in cell number. This suggests that fundamentally a cell functions well only within the size range which normally occurs. This is even more clearly indicated when we compare cell sizes within a single organism. In plants the most active

cells, which are growing and dividing and have the greatest metabolic rate, are those in the growing point of stem and root tips, and these are much smaller than the mature cells and more nearly equal in size. They range from 1 to 3 × $10^3\mu^3$ for the meristematic cells of the fastest growing plants, and are in the $10^4\mu^3$ range for the slower growing ones.

This uniform cell size suggests that the most fundamental cell activities are dependent upon size. This conclusion is strengthened by the consideration that the most active cells are smallest and are most nearly alike in size, both in plants and animals. Actually the fastest growing cells are those of bacteria, with a generation time as short as 20 minutes. For cells in the growing point of a plant, the time elapsed between successive divisions is 24 hours or more, or about 100 times as long, whereas

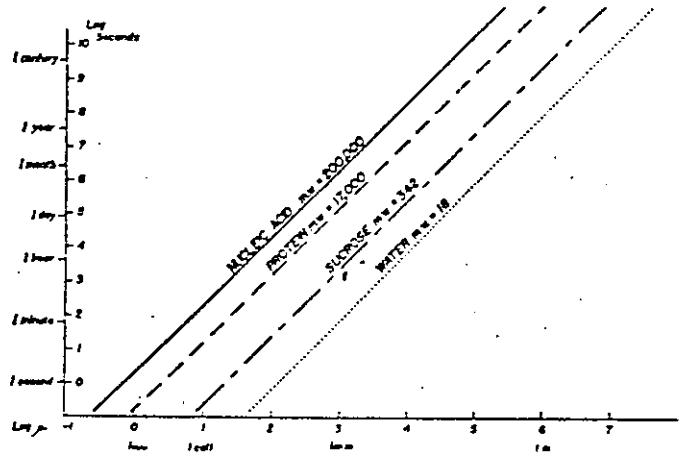


FIG. 1. Relationship between diffusion time (ordinate) and distance where half saturation is reached (abscissa) for a number of substances with different molecular weights. Over short distances diffusion is very fast, whereas over long distances diffusion is fantastically slow. Data based on calculations by Dr. K. Sevier.

the linear dimensions are in the proportion of 1:10. Therefore, we might conclude that the division rate and the activity of the cell are controlled by a process which decreases in rate with the square of the diameter of the cell. This is typical for a diffusion process as shown in Figure 1. Mass reactions would presumably not be much influenced by the degree to which the reaction mixture is divided into smaller cells. From Figure 1 we can also conclude that if a diffusion process in the cell is rate-limiting, this must be a process involving very large molecules, for small molecules, like sugar, become equally divided within a cell by diffusion in terms of seconds. If, on the other hand, food distribution in the body of a man depended on diffusion, it would take a lifetime to get sugar fed into the stomach to diffuse into the feet and hands, and it would be

possible to take only a few steps in a lifetime. This shows that, depending on size, different processes are differentially controlled. At the size of a cell, mixing of even larger molecules is almost instant, and protoplasmic streaming would, for instance, not speed up the equalization of sugar distribution in a cell of about 0.1 mm in diameter. On the other hand, streaming and mass flow become almost imperative to get food distributed in an organism, when it is more than a millimeter in size.

In this connection it is interesting to note that, almost without exception, active animals of more than 1 mm size have a heart or other organ for circulation of substances through the body: above 1 mm size hydraulic flow has to substitute for food distribution by diffusion. And a major reason why insects are all small when compared with the higher animals, is that they depend for oxygen supply to their tissues mainly upon diffusion through tracheae. Some of the larger insects do have ventilation in their tracheae and do not get oxygen only by diffusion. For any larger sized animal, lungs or gills with active ventilation are essential. The exo-skeleton of insects would make lung ventilation very difficult. Arthropods with gills past which they pump water have become considerably larger (crabs, lobsters).

TABLE 1

ACTIVITIES AND ACCOMPLISHMENTS OF ANIMALS COMPARABLE WITH THOSE OF MAN

Tools:	California Sea Otter carries a stone as anvil to crack shellfish
	Birds swallow gravel to grind food in stomach
	Lobsters use sand as otoliths in antennae
	Ant lion pelts victims with sand to let them slide into his sand trap
	The fish, Toxotes jaculator, aims water-drops at insects above the water which then drop into water and are eaten
	Monkeys use sticks to reach fruit
Structures:	Nests (of birds, bees, wasps, ants, termites, Weaver birds)
	Webs and trapdoors of spiders
	Hiding structure of Octopus made with stones
	Container made from sand grains by the protist Difflugia
Cultivation:	Fungi (leaf cutter ant)
	Flowering plants (epiphytic ant gardens)
	Aphids (ants)
Communications:	Sound (mammals and birds)
	Smell (ants)
	Touch (ants)
	Vision: Dance (bees)
	Radar (bats)
Societies:	Ants, bees, termites
	Monkeys
	Beavers
	Weaver birds
	Ground hogs
Food Storage:	Squirrels and many other rodents
	Ants, bees, termites
Air Conditioning:	Ventilation of Atta nests
	Burrows of desert animals

The emergence of a man as a thinking creature with a full-fledged technology, from among the tens of millions of animals which have lived and still are living on earth, has been attributed to a variety of factors, anatomical and mechanical. A favorite among these is his brain volume; others are his erect posture and the development of hands. Then again his use of fire is considered the main factor which differentiates man from beast. I believe that as good a case (or perhaps even a better one) can be made by considering man's physical size as the critical factor which made it possible for him to develop a technology and to use fire.

In most respects there are no differences in principle between man and animals. We will not consider any actions for which the animal body shows special adaptation, or any instinct actions. The accompanying Table 1 shows that many animals have learned the use of tools, make elaborate structures (usually nests), cultivate plants and animals, have means of communicating with each other, develop societies (with hierarchies of authority and division of labor), store food, and make use of air-conditioning.

In most of these respects, however, there are quantitative differences between man and the animals. Man's tools are infinitely more refined, his structures more sophisticated, his communication system incomparably more complex, his food storage and air-conditioning of a completely different order of magnitude. But there are also qualitative differences between man and all other animals. Among these I want to stress:

(1) the use of fire
(2) the capture and use of animal and mechanical energy (horsepower, wind and water power, the steam and internal combustion engines, atomic energy)
(3) the use of the wheel
(4) the production and use, first of stone implements, and later of metals
(5) the development of script
(6) the storage of information, available for education and development.

When we analyze these six functions, we find that each one of them could not have been developed or used by ants or other small animals, no matter how far advanced their society might have become. This is a result of the principles of critical mass and critical size.

Just as an example of what is meant by the latter statement, let us consider for a moment the dimensional limitations of fire. A flame cannot be smaller than several millimeters length (and even then is relatively unstable), and requires a volatile combustible material, such as gas or alcohol. Nonvolatile combustible material such as wood or coal has a much larger critical mass for combustion. The reason for the lower limit of flame size is that the ignition point of gases and vapors lies rather

high, usually many hundreds of degrees centigrade. When the temperature of the flame decreases below the ignition point, the flame extinguishes. Since the flame has to be maintained with the oxygen of the air, and this air is cold, the flame must be large enough to heat the inrushing air to the ignition point, and provide enough heat to volatilize the fuel to be burned. When there is wind, the critical flame size increases because of the strong cooling by the inrushing air. Then candles are impractical, unless shielded from the wind, and torches are required with flames of sufficient size so that the air brought in by the wind can be heated to the ignition point of the burning material.

The critical size of a wood or coal fire is considerably greater than for a flame, because not only the air, but also the basic mass producing the combustible materials has to be heated to the point that either volatile materials are liberated or that the ignition point of the solid material is reached. At that temperature so much heat is radiated that such a fire can only be maintained if the burning piece of wood or coal is surrounded by other burning pieces. Interestingly enough, a wood or coal fire above the critical size produces just the right amount of heat to warm man in a cave, or a room, or a camping site. But ants or small rodents would have to keep too far away to make a fire economical, or rather, they would be unable to bring up enough wood to keep the fire going. Therefore in an ant society fire is not an economical possibility, and they have developed without its benefits, by operating only while outside temperatures are within the physiological range. Man on the other hand has been able to move into very cold areas by using fire.

Let us consider another significant difference between the world of the ant and the world man lives in. This is also a consequence of dimension. Mass and weight decrease with the third power of linear size. The weight of a body in the macro-world of the earth crust has practically nothing to do with its position in relation to other objects since the gravitational pull of the earth completely overshadows all other forces such as the gravitational effects of the surrounding objects.

But as the dimensions decrease another force becomes important, and this is molecular attraction or cohesion, also manifested in absorption phenomena. In Figure 2 we see how gravitational attraction increases with the third power of the linear dimension of an object. Molecular attraction on the other hand increases with the second power of the linear dimension, and decreases with the square root of the distance between the surfaces. Since with a decrease in size there generally is also a decrease in effective distance between bodies in physical contact, the force with which touching bodies are attracted one to another increases approximately with the linear size whereas mass is increasing with the third power of the diameter. This means that a 1 mm^3 cube with a reasonably flat surface (such as a wooden cube cut with a razor blade)

does not drop, but can be supported by adhesion to another flat surface. In this case the gravitational force is barely balanced by adhesion. On the other hand, for a 1 cm³ block, gravity is fifty times greater than adhesion, whereas for a cube of a 0.1 mm rib size adhesion is 50 times stronger than gravity. At still smaller sizes, gravity loses all significance in relation to molecular attraction, regardless of the shape of the body, and particles of dust will stick anywhere to any surface. Basically,

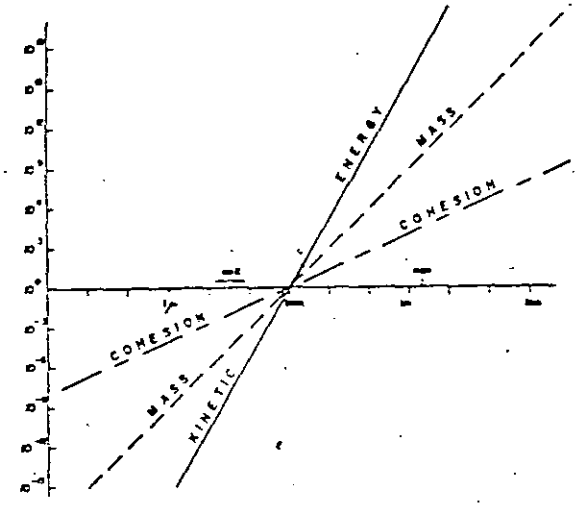


FIG. 2. Dependence of (1) cohesion between surfaces, (2) mass (weight), and (3) kinetic energy (relative values indicated on ordinate) upon linear dimension (abscissa). Both cohesion and kinetic energy values are only approximate, based on average roughness of surface and on average velocities of masses of various sizes.

electrostatic attraction is similar to molecular attraction, since it is proportional to the surface of the charged bodies and inversely proportional to the square of their distance. The absolute value of electrostatic attraction is a function of the charge and is potentially more than a thousandfold greater than molecular attraction. But when bodies are in contact with each other, they lose their charge, and only molecular attraction remains.

The foregoing considerations make it clear why man cannot walk up against a vertical surface since at his size gravitational attraction exceeds molecular adhesion more than a millionfold. A fly, on the other hand, with its flat sole surfaces (not suction cups) is just in the size range, where Figure 2 shows that gravity and adhesion are of the same order of magnitude. Therefore it can walk against surfaces in any position, even upside down against the ceiling. Consequently we find a complete change in the physical world of organisms, between the size where

gravitational forces exceed molecular forces, and the size where molecular forces predominate. I propose to name these two worlds: The macro-world, or Newtonian world where gravitational forces hold sway; and the micro-world or molecular world, where not classical mechanics but thermodynamic and statistical laws are dominant. A good criterion as to whether a process belongs in the classical, mechanical, or thermodynamical category is the degree to which it is dependent upon temperature. In general, we can say that gravitational attraction, or the velocity of an object is largely independent of temperature. Chemical reactions, or osmotic water movement, on the other hand, depend on the activity or specific temperature of the participating molecules.

Considerations about kinetic energy show us another fundamental difference between the macro-world of man and the micro-world of insects and smaller creatures. The numerical values of kinetic energy actually give us a good clue as to the optimal size of man. A 2 m tall man, when tripping, will have a kinetic energy upon hitting the ground 20–100 times greater than a small child who learns to walk. This explains why it is safe for a child to learn to walk; whereas adults occasionally break a bone when tripping, children never do. If a man were twice as tall as he is now, his kinetic energy in falling would be so great (32 times more than at normal size) that it would not be safe for him to walk upright. Consequently man is the tallest creature which could reasonably walk upright on two legs. The larger mammals can become taller, because they are more stable on their four legs. Yet, they break bones more easily when they fall. These are all cases indicating the dangers involved in living in the macro-world, and the limitations it imposes on the macro-animals. The micro-animals, such as insects and micro-organisms live in a totally different world, unattended by the forces (and dangers) of gravity and kinetic energy. However, they depend upon the equally exacting forces of thermodynamics. Perhaps we should call these Gibbsian forces in contradistinction to Newtonian forces because of the emphasis Gibbs placed on the forces and laws of thermodynamics.

From these considerations it seems clear why man could not be much larger than he is, nor could an animal of the size of an ant have conquered the world with a technology. But the question remains open, why could not an animal of rat or rabbit size have come to dominate the world? At that size they could have developed fire, as the amount of wood collected by beavers attests. And from that, every other technological development might conceivably have sprung as it did in the case of man. However, there are other primitive activities for which man-size is required, and without which man could not have reached and crossed the brink of technology.

The earliest tools and weapons of man all are limited in their effectiveness by size. A club of a size to kill a larger animal could not be swung

Although most of the differences between the macro- and micro-worlds are matters of common knowledge, the implications of these differences are often forgotten and thus an amazing number of incorrect interpretations are suggested by magnification of the micro-world into the dimensions of the macro-world. Let me enumerate some of these differences.

In the macro-world we can weigh objects by measuring the gravitational pull on them. It is easy to compare weights in the pound or kilogram class, but when we come below the milligram, weight loses its significance, for molecular attraction so much exceeds gravitational pull that weight is an abstraction rather than a reality. Then we prefer to speak in molecular equivalents.

The same thing holds for length. In the macro-world a meter has a very concrete meaning; it is a definite and fixed distance to be measured. But as we reduce size, the wave length of the used radiation starts to limit the accuracy of our measurements. At a still smaller scale, thermal motion of molecules and molecular aggregates make distance a statistical average; it is no longer a reality.

Our time scale is equally precise in both worlds but the dimensions are entirely different. In the macro-world we are dealing with seconds to millions of light years; in the micro-world we frequently are confronted with milli- or microseconds or less.

When we are considering volume in the micro-world, gas bubbles cannot exist any more. A gas bubble of 1μ diameter in water can exist only when the gas is under pressure of three atmospheres; a bubble of 0.1μ is physically almost impossible because of the extreme surface tension of the water causing a pressure of 30 atm inside the bubble. This results in the well-known property of cohesion of water, preventing it from pulling away from a thoroughly soaked hydrophilic surface such as the wall of a tracheid or vessel in wood.

There has been much discussion during the last years whether water cohesion could actually account for the enormous negative tensions necessary to move water into the highest branches of the tallest trees. In physical models it never was possible to reach through cohesion negative tensions of several hundred atmospheres, as are claimed to exist in, for instance, drying annulus cells of fern sporangia. However, here the data of Figure 2 give the solution. In a microscopic system, the kinetic energy of the mass of water enclosed in a cell remains very small even after a sharp shock of the cell. In a larger system, such as a glass tube used in a physical model, slight vibrations can build up kinetic energies of the water in the tube a million times as great as that in individual plant cells, such as tracheids. Since this kinetic energy can temporarily exceed the cohesive force of the water, under tension, a water column breaks much sooner in a macroscopic than in a microscopic system.

This provides another striking difference between the macro- and micro-world. If a tree had a plumbing system like that of a house, consisting of fairly large and wide tubes, the cohesion of water would be insufficient to raise water to more than 10 m (30 ft) height, and without a pressure system, pushing water beyond the 10 m limit, trees could not reach any greater height. As a matter of fact, without a hydraulic water transport system, no land plants could be more than a few millimeters high nor could they be very different in form from a liverwort or a fern prothallium since osmotic and other thermodynamic systems become practically ineffective if diffusion or osmosis has to move water molecules beyond the microscopic range as seen in Figure 1. Therefore the first successful landplants, raising themselves even slightly above the surface of the wet ground, had to augment their osmotic or, more generally speaking, thermodynamic water moving system with a hydraulic system; or, in other words, they had to graduate from the micro- to the macro-world.

All these examples show the chasm which exists between macro- and micro-world, a chasm which as biologists we have to cross continuously in both directions without having been alerted to the pitfalls of the crossing. For instance, when we observe a cell under the microscope, the magic of wave-mechanics transposes the basic unit of life from the micro-world into a body with the apparent size of a pea or a football, and in our mind it becomes part of the macro-world. Its practical weightlessness, its diminutive size disappear and we think of it carrying out its functions on a macro scale. Let us consider how a sugar crystal introduced in the football-size cell would dissolve and mix through diffusion with the cell contents. It would require weeks before concentration equality would be reached. But a proportionately small sugar particle in the micro-sized cell would dissolve and reach diffusion equilibrium in a minute. This same cell when flaccid can take up the necessary water to become fully turgid in a matter of minutes. But a plasmolyzed football would require days to become turgid again.

Less extreme cases would involve the performance of an ant which can easily lift ten times its own weight which, transposed in human dimensions, would make it simple for milady shopping to carry her compact car off the street into the parking lot. Or a leafcutter ant cutting off a leaf segment high in a tree hundreds of times its own height will fall down with it without batting an eyelid (they do not have any), in contrast to a steeplejack a thousand feet up in a tower. We also often compare a wheat plant 1½ m high, having a stem diameter of 7 mm, with a 2 m wide tower 250 times as high (400 m, about the height of the Empire State Building). The latter example shows that we do not need to move from micro- to macro-world to make impermissible comparisons by proportioning an object from one into another dimension.

3. SOLUCION DE ECUACIONES EN UNA VARIABLE

Ecuaciones no lineales

Bisección

Falsa posición

Newton-Raphson

Secante

Otros (Aitken Δ^2, aceleración de métodos iterativos, Monte Carlo, etc)

Aplicaciones

4. SOLUCION DE SISTEMAS DE ECUACIONES

Sistemas lineales

Eliminación de Gauss

Métodos iterativos

Sistemas no lineales

Aplicaciones

5. ECUACIONES DIFERENCIALES ORDINARIAS

Métodos de un paso

Series

Métodos predictor-corrector

Solución por series

Diferencias finitas

Sistemas de ecuaciones diferenciales ordinarias

Aplicaciones

INGENIERIA MARITIMA. MODULO:
METODOS NUMERICOS EN HIDRAULICA
DEL 1o. DE JULIO AL 6 DE SEPTIEMBRE 1985
MEXICO, D.F.

INTRODUCCION A LOS METODOS NUMERICOS EN HIDRAULICA

M. EN I. OSCAR FUENTES MARILES
SEPTIEMBRE 1985

UNIVERSIDAD NACIONAL AUTONOMA DE MEXICO

INTRODUCCION A LOS METODOS NUMERICOS EN HIDRAULICA

M en I Oscar Fuentes Mariles *
*M en I Polioptro Martínez Austria**

*Profesores de la División de Estudios de Posgrado de la Facultad de Ingeniería

1. INTRODUCCION

Estas notas tratan de ser una introducción a los métodos numéricos que se aplican en la hidráulica, se ha intentado evitar demostraciones, teoremas y otros requerimientos matemáticos con objeto de no hacerlas complicadas, buscando propiciar en el lector el conocimiento y empleo de ciertas técnicas numéricas utilizadas para resolver problemas de hidráulica y mostrar la aplicación de tales técnicas.

Algunas veces en el estudio de presas los ingenieros enfrentan problemas como el tránsito de una avenida. Sea el almacenamiento mostrado en la fig 1.1; para calcular el tránsito de un hidrograma se requiere resolver la ec

$$\frac{dV}{dt} = I - O \qquad 1.1$$

donde V es el volumen almacenado, I y O son gastos de entrada y salida del almacenamiento respectivamente y t es tiempo

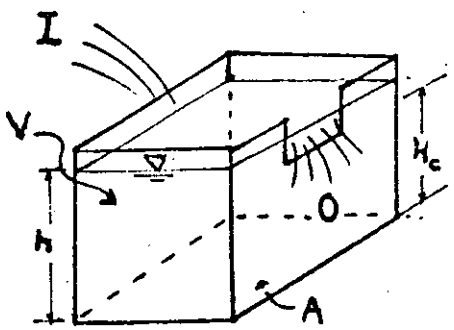


fig 1.1.

De la fig 1.1

$$V = A\,h \qquad (1.2)$$

mientras que el gasto descargado según la ecuación de vertedor rectangular

$$O = K(h-h_c)^{3/2} \qquad (1.3)$$

al tomar en cuenta 1.2 y 1.3 en 1.1

$$\frac{dh}{dt} = \frac{I}{A} - \frac{K}{A}(h-h_c)^{3/2} \qquad (1.4)$$

la cual es una ecuación diferencial ordinaria que no es lineal porque la incognita h aparece elevada a una potencia diferente de la unidad.

Para resolver esta ecuación los métodos analíticos resultan ser complicados de aplicar, sin embargo, una solución aproximada de esta ecuación puede ser encontrada por medio de un método numérico, la cual generalmente es tan útil como la exacta.

Se entiende por método numérico aquel procedimiento por el cual problemas matemáticos se resuelven mediante operaciones aritméticas. Dado el gran número que se requiere de estas

operaciones es conveniente emplear computadoras. La combinación del método y computadora se ha constituido en una poderoso herramienta para resolver problemas matemáticos.

Se entenderá que las soluciones de los problemas matemáticos son dos clases, la analítica o cerrada que corresponden a la solución exacta y la aproximada o numérica o abierta que dará una solución muy parecida a la exacta.

Por ejemplo, la integral $\int senxdx$ tiene por solución exacta a $-cosx+C$ mientras que $\int_a^b \frac{senx}{x}\,dx$ tiene por solución aproximada a $\frac{b-a}{2}(\frac{sena}{a} + \frac{senb}{b})$.

Como se recordará esta última integral no se puede encontrar mediante las técnicas analíticas elementales del cálculo integral, sin embargo, numéricamente es sencillo obtener una solución aproximada.

La mayor parte de los fenómenos que estudia la ingeniería civil quedan representados por modelos matemáticos; esto es, por ecuaciones: lineales, no lineales, diferenciales ordinarias, diferenciales parciales.

En algunas ocasiones, las ecuaciones resultantes son difíciles de resolver; en otras definitivamente no se conoce solución exacta. Afortunadamente, en el propio campo de las matemáticas se han desarrollado métodos numéricos de solución, los

La energía específica se define como:

$$E = y + \frac{V^2}{2g} \qquad (2)$$

y de continuidad:

$$V = \frac{Q}{A} \qquad (3)$$

que sustituída en (2) resulta:

$$E = y + \frac{Q^2}{2gA^2} \qquad (4)$$

y como $A^2 = by$; y definiendo el gasto unitario como:

$$q = \frac{Q}{b}$$

donde b es el ancho del canal, la energía específica finalmente puede expresarse como:

$$E = y + \frac{q^2}{2gy^2} \qquad (5)$$

Con esta expresión, conocido y_1 se calcula E_1.

Escribiendo la ec. 5 en lugar de E_2 en la expresión 1, se obtiene:

$$E_1 = y_2 + \frac{q^2}{2gy_2^2} + \Delta Z \qquad (6)$$

ordenando:

$$y_2 + \frac{q^2}{2gy_2^2} + \Delta Z - E_1 = 0$$

y multiplicando por y_2^2 se obtiene finalmente:

$$y_2^3 + (\Delta Z - E_1)y_2^2 + \frac{q^2}{2g} = 0 \qquad (7)$$

La solución de esta ecuación proporciona el tirante y_2, buscado, puesto que se conocen ΔZ y Q; y E_1 se puede calcular dado y_1.

En casos como este, el método de solución más evidente, pero también el más ineficiente, consiste en asignar valores a la variable mediante "tanteos", hasta que se cumpla la ecuación.

Por otra parte, el problema de resolver una ecuación en una sola variable, cuando ésta es implícita, ha sido atacado por los matemáticos desde hace más de un siglo, existen métodos eficientes para computadora; y aún ahora se siguen proponiendo otros nuevos.

El problema puede plantearse sencillamente como; dada la función

$$f(X) = 0 \qquad (8)$$

encontrar los valores X_1, X_2,..., X_n que la satisfacen. Estos valores se denominan raíces de la ecuación. En la figura 2 se muestra una interpretación geométrica.

El número de raíces depende de la propia ecuación; pudiendo ser infinito, por ejemplo la ecuación $f(X) = SenX = 0$.

A continuación se procederá a presentar algunos de los métodos más usuales para la solución de la ecuación (8).

2.2 METODO DE BISECCION

Considérese el problema de encontrar una raíz de una ecuación como se muestra en la figura 3.a .

Supóngase que se eligen dos valores de X; a y b, a ambos lados de la raíz, y se toma el promedio de ellos (P_1).

El uso de la instrucción DEF FNF, permite además una mayor flexibilidad en el uso del programa, dado que bastará cambiar la función en esta linea del listado para que se pueda resolver con problema diferente.

En la figura 5 se presenta el listado de este programa para el método de bisección, con comentarios suficientes para su empleo.

Para ejemplificar la aplicación de este método, considérese nuevamente el problema de flujo sobre un escalón, planteado en la introducción de este capítulo.

La ecuación para el cálculo del tirante en la sección 2, es la 7:

$$y_2^3 + (\Delta Z - E_1) y_2^2 + \frac{q^2}{2g} = 0 \qquad (7)$$

En el ejemplo, se tenían los siguientes datos:

$$Q = 0.5\ m^3/s \qquad b = 1.5\ m$$

$$y_1 = 0.15\ m \qquad \Delta Z = 0.10\ m$$

El gasto unitario será:

$$q = \frac{Q}{b} = \frac{0.5}{1.5} = 0.333\,(m^3/s)/m$$

y por la ec 5, la energía específica en la sección 1 valdrá:

$$E_1 = y_1 + \frac{q^2}{2g y_1^2} = 0.45 + \frac{0.333^2}{19.62(0.15)^2} = 0.402$$

sustituyendo este valor en la ec. 7, los otros datos y haciendo operaciones, se obtiene finalmente:

$$y_2^3 - 0.352 y_2^2 + 0.00566 = 0 \qquad (8)$$

una de cuyas soluciones es el tirante buscado.

Para ilustrar el funcionamiento del método de bisección, se procederá a operar el algoritmo inicialmente sin auxilio del programa.

El tirante en la sección 1 es supercrítico, por lo que el tirante sobre el escalón lo será también. El tirante crítico es:

$$y_c = \sqrt[3]{\frac{q^2}{g}} = 0.224 \ m$$

Entonces el tirante buscado estará en el intervalo:

$$0.15 < y_1 < 0.224$$

por lo que se puede escoger $a = 0.16$; $b = 0.224$

En la tabla 1 se resúmen los cálculos efectuados con el algoritmo del método de bisección. Si se considera aceptable una tolerancia de 0.0001, con el tercer cálculo se obtendría el resultado: $y_2 \cong 0.184$

Utilizando el programa, el procedimiento sería como sigue:

1. Antes de ejecutar el programa se teclea en la linea 140 la función $\acute{o}(X) = a$ la que se busca raíz en el intervalo a, b.

Iteración	a	b	p	$\acute{o}(p)$
1	0.16	0.224	0.192	-0.0002382
2	0.16	0.192	0.176	0.0002082
3	0.176	0.192	0.184	-0.0000278

$\acute{o}(p) = p^3 - .352p^2 + .00566$

$0.0000278 < 0.0001 \ \therefore \ y \cong 0.184$

Tabla 1 Ejemplo del método de bisección

2. El programa da al usuario información general:

METODO DE BISECCION

ECUACION A RESOLVER

EN LINEA 140

3. El programa pide datos:

INTRODUZCA INTERVALO DE BUSQUEDA

VALOR MINIMO DE BUSQUEDA?

Teclear: 0.16

VALOR MAXIMO DE BUSQUEDA?

Teclear: 0.224

TOLERANCIA?

Teclear: 0.0001

4. El programa informa inicio de proceso y, segundos después, anuncia resultados:

*** SE ENCONTRO SOLUCION***

LA RAIZ ES:

$X = 0.184$

****

```
10 REM PROGRAMA PARA SOLUCION D
   E ECUACION ALGEBRAICA.
20 REM **METODO DE BISECCION**
30 REM
40 CLEAR @ DISP "**************
   **" @ DISP "METODO DE BISECC
   ION" @ DISP
50 DISP "ECUACION A RESOLVER:"
   @ DISP " EN LINEA 140" @ DIS
   P
60 DISP "INTRODUZCA EL INTERVAL
   O DE" @ DISP " BUSQUEDA" @ D
   ISP @ DISP
70 REM SE PIDEN DATOS.
80 REM
90 DISP "VALOR MINIMO DE BUSQUE
   DA":@ INPUT A
100 DISP "VALOR MAXIMO DE BUSQUE
   DA":@ INPUT B
110 CLEAR @ DISP "TOLERANCIA";@
   INPUT T
120 REM
130 REM EN LA SIGUIENTE LINEA SE
    DEFINE LA FUNCION
140 DEF FNF(X) = X^3-.352*X^2+.0
   0566
150 REM
160 REM SE INICIA PROCESO DE CAL
   CULO, SE INFORMA A USUARIO
170 CLEAR @ BEEP @ DISP "SE INIC
   IA PROCESO"
180 P=(A+B)/2
190 IF ABS(FNF(P))<T THEN GOTO 2
   90
200 IF FNF(A)<0 THEN GOTO 240
210 IF FNF(P)<0 THEN GOTO 230
220 A=P @ GOTO 180
230 B=P @ GOTO 180
240 IF FNF(P)<0 THEN GOTO 220
250 GOTO 230
260 REM
270 REM SE ENCONTRO RAIZ, Y SE I
   NFORMA A USUARIO
280 REM
290 CLEAR @ BEEP @ DISP "*** SE
   ENCONTRO SOLUCION ***" @ DIS
   P
300 DISP "LA RAIZ ES:" @ DISP "X
   =" P
310 DISP "**********************"
320 END
```

Figura 5.- Listado de programa método de bisección

Para concluir estas notas sobre el método de bisección conviene mencionar que, si se cumple la restricción de que $f(a)$ y $f(b)$ sean de signos contrarios, el método es siempre convergente, es decir que se aproxima gradualmente a la solución.

Un método no converge cuando se aleja, o no se aproxima, a la solución conforme aumenta el número de iteraciones.

La convergencia es una característica importante del método de bisección, sin embargo es comparativamente lento; es decir que se acerca a la solución con un mayor número de iteraciones que otros métodos.

2.3 METODOS DE FALSA POSICION

El método de falsa posición, o de Regula Falsi, opera de manera similar al de bisección, solamente que en lugar de utilizarse un promedio aritmético para evaluar aproximaciones a la raíz, se emplea una relación de triángulos semejantes.

Considérese que se desea calcular una raíz de una función como se indica en la figura 6, donde se han elegido dos valores de la variable, a y b, tales que se cumple que $f(a)f(b) < 0$.

El punto en que la recta que une a $f(a)$ y $f(b)$ cruza el eje X, se denota por la letra C.

En la figura se pueden definir dos triangulos; el mayor con vértices en $f(a)$, d y $f(b)$; y el menor con vértices

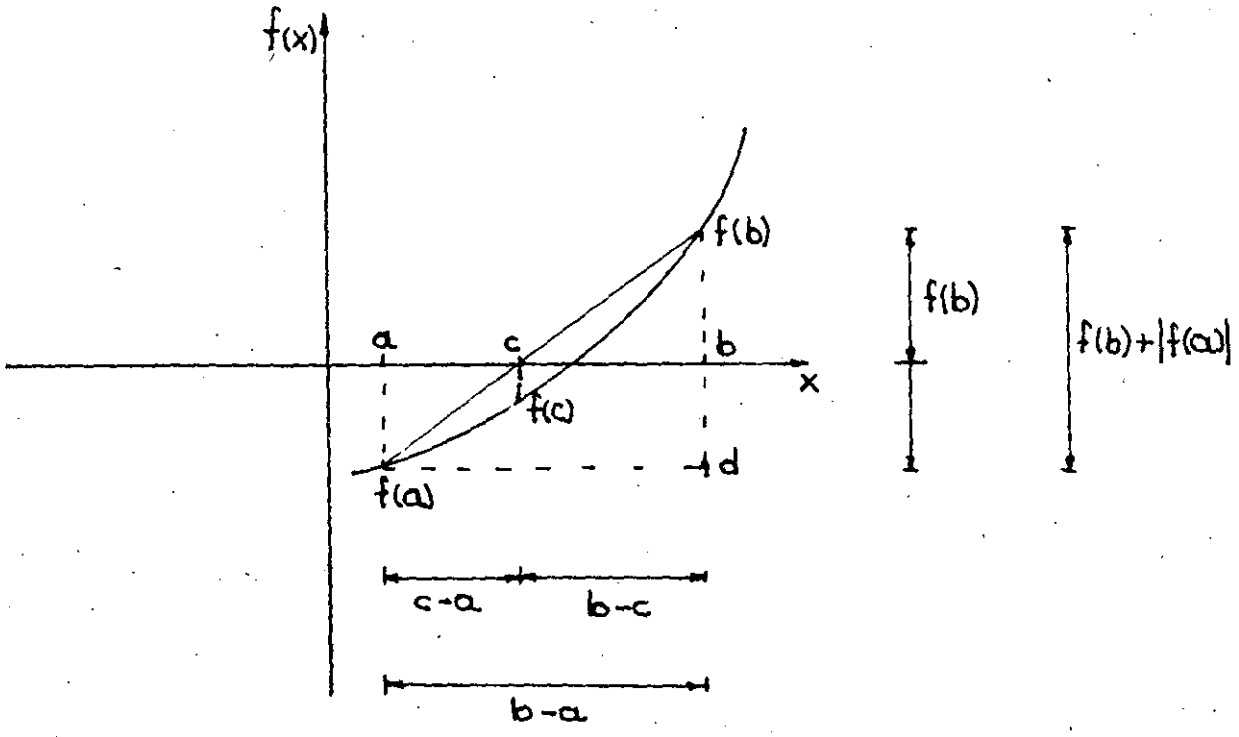


Figura 6.- Geometría de una función para el método de falsa posición

en c, b y $\delta(b)$. Estos triángulos son semejantes, de manera que se puede plantear la relación:

$$\frac{\delta(b)}{b-c} = \frac{\delta(b) - \delta(a)}{b-a} \qquad (9)$$

y despejando el valor de c, se obtiene:

$$c = b - \frac{\delta(b)\,(b-a)}{\delta(b) - \delta(a)} \qquad (10)$$

Calculado C, se aproxima a la raíz por un procedimiento iterativo similar al del método de bisección; es decir cerrando el intervalo de búsqueda paulatinamente; cuidando de conservar la condición de que $\delta(a)\delta(b) < 0$.

Si $\acute{o}(c)$ es negativa, entonces a se iguala a c; en caso contrario es b la que adopta el valor de c.

En el caso en que la curva fuera decreciente, es decir que $\acute{o}(a)$ fuera positiva, entonces la ecuación del método es la siguiente:

$$c = a + \frac{\acute{o}(a)\ (b - a)}{\acute{o}(a) - \acute{o}(b)} \qquad (10')$$

El algoritmo del método sería el siguiente:

ALGORITMO METODO DE FALSA POSICION

1. Localize dos valores de X; que se denominarán a y b, localice tales que $\acute{o}(a)\ \acute{o}(b) < 0$.

   Difina una tolerancia T

2. Calcule $\acute{o}(a)$ y $\acute{o}(b)$

3. Calcule c:

   Si $\acute{o}(a) < 0$ con la ecuación 10
   Si $\acute{o}(a) > 0$ con la ecuación 10'

4. Calcular $\acute{o}(c)$

5. Si $|\acute{o}(c)| < T$ vaya al paso 9
   Si $|\acute{o}(c)| > T$ continúe

6. Si $f(c) < 0$ haga alguna de las siguientes operaciones:

Haga $a = c$ si $f(a) < 0$
Haga $b = c$ si $f(a) > 0$
y vaya a paso 8

7. Si $f(c) > 0$ haga alguna de las siguientes operaciones:

Haga $b = c$ si $f(a) < 0$
Haga $a = c$ si $f(a) > 0$
continúe

8. Regrese a paso 3

9. Algoritmo terminado: la raíz es C.

Para ilustrar la aplicación de este método, se utilizará una vez más el ejemplo del flujo sobre un escalón.

La ecuación a resolver era la número 8:

$$y_2^3 - 0.352y_2^2 + 0.00566 = 0 \qquad (8)$$

y los valores iniciales de cálculo serían:

$$a = 0.16 \quad y \quad b = 0.224$$

En la tabla 2 se resúmen los cálculos efectuados.

Iteración	a	b	δ(a)	δ(b)	c	δ(c)
1	0.16	0.224	0.0007448	-0.0007625	0.1916	- .000229
2	0.16	0.1916	0.0007448	- .000229	0.1842	-0.0000325

Tabla 2 Aplicación del método de falsa posición

Si se fija una tolerancia de *0.0001*, en el segundo cálculo se habrá encontrado la raíz:

$$y_2 = 0.1842$$

En este cálculo se ha utilizado la ecuación 10'.

El algoritmo de falsa posición es también de sencilla programación. En la figura 7 se presenta un diagrama de flujo, y en la 8 el listado correspondiente. Como en el caso del método de bisección, al utilizar la definición de función por el usuario aumenta la generalidad del programa.

Una corrida típica sería como sigue:

```
PROGRAMA METODO DE FALSA POSICION
**********************
FUNCION EN LINEA 130

INTERVALO DE BUSQUEDA
VALOR MINIMO DE BUSQUEDA?
0.16
VALOR MAXIMO DE BUSQUEDA?
0.224
TOLERANCIA?
0.0001
```

```
SE INICIA EL PROCESO

**********************
SE ENCONTRO RAIZ
          X= 0.184188
**********************
```

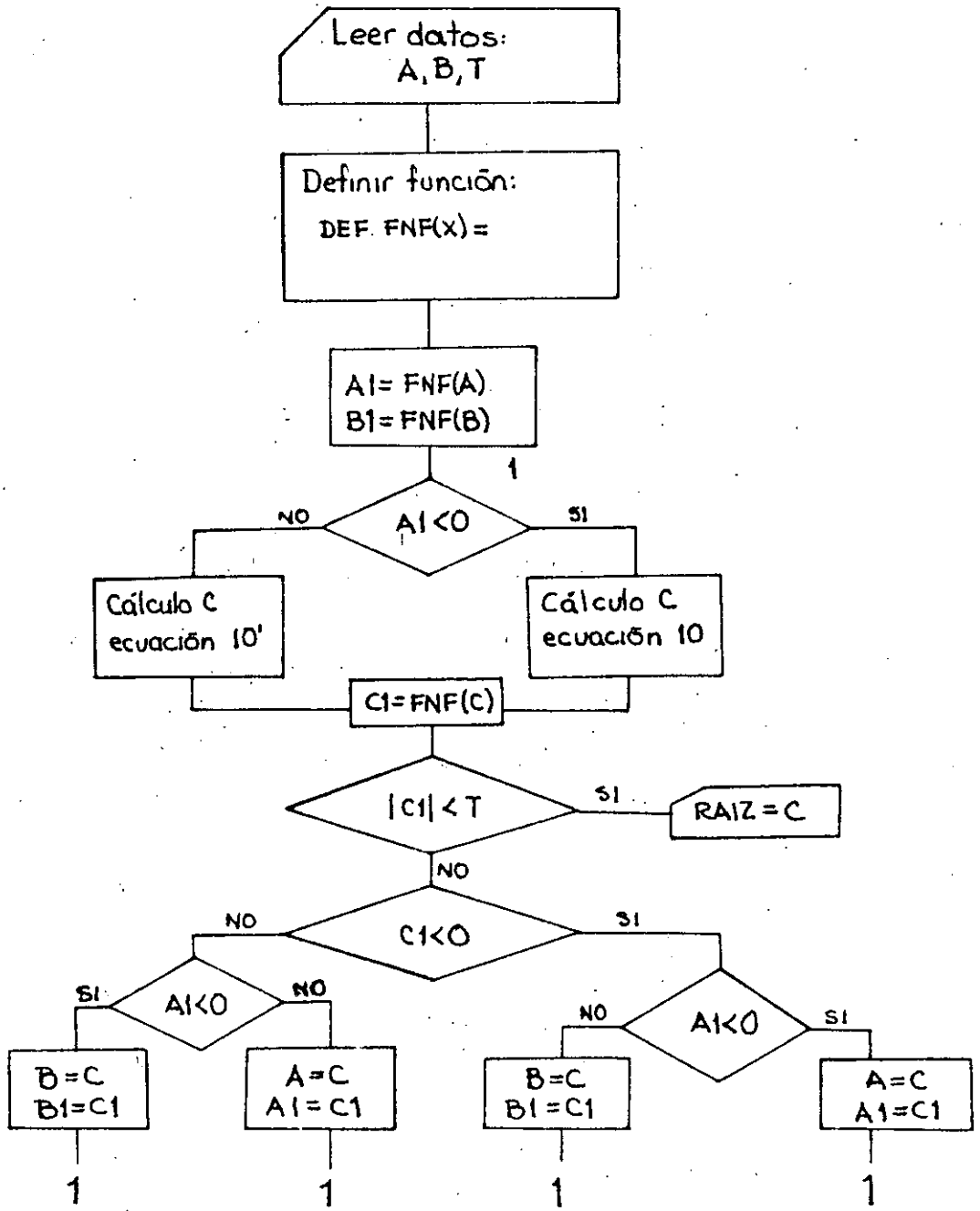


Figura 7 Diagrama de flujo método de falsa posición

```
 10 CLEAR @ DISP " PROGRAMA DE F
    ALSA POSICION "
 20 DISP " ************ " @ BEEP
 30 DISP " FUNCION EN LINEA 130
    "
 40 REM SE PIDEN DATOS
 50 REM
 60 DISP @ DISP "INTERVALO DE BU
    SQUEDA ":@ DISP
 70 DISP " VALOR MINIMO DE BUSQU
    EDA ":@ INPUT A
 80 DISP " VALOR MAXIMO DE BUSQU
    EDA ".@ INPUT B
 90 DISP " TOLERANCIA ":@ INPUT
    T
100 CLEAR
110 DISP " SE INICIA PROCESO " @
     BEEP
120 REM EN LINIA SIGUIENTE SE DE
    FINE FUNCION.
130 DEF FNF(X) = X^3-.352*X^2+.0
    95*6
140 A1=FNF(A) @ B1=FNF(B)
150 IF A1<0 THEN C=-(B1*(B-A)/(B
    1-A1)) ELSE C=A+A1*(B-A)/(A1
    -B1)
160 C1=FNF(C)
170 IF ABS(C1)<T THEN GOTO 260
180 IF C1<0 THEN GOTO 220
190 IF A1<0 THEN GOTO 210
200 A=C @ A1=C1 @ GOTO 150
210 B=C @ B1=C1 @ GOTO 150
220 IF A1<0 THEN GOTO 200
230 GOTO 210
240 REM
250 REM SE ESCRIBEN RESULTADOS
260 CLEAR @ BEEP @ DISP " ******
    ******** "
270 DISP " SE ENCONTRO RAIZ" @ D
    ISP "X=";C
280 DISP "**********"
290 END
```

Figura 8 Listado del programa para método de falsa posición

El método de falsa posición, al igual que el de bisección, es convergente siempre que se cumpla la condición de que $f(a)\ f(b) < 0$, y su velocidad de convergencia es del mismo órden.

2.4 METODO DE NEWTON-RAPHSON

Un método muy conocido de solución de ecuaciones en una variable, es el de Newton-Raphson, que utiliza el concepto de derivada.

Considérese una función como se muestra en la figura 9.a; en la que se ha propuesto un valor X_i para la raíz. En términos generales esta primera elección no será la solución, es decir que en X_i la función tendrá un valor $f(X_i)$. En este punto, la derivada, que define la pendiente de la función, valdrá $f'(X_i)$ y esta recta cruzará el eje X en el punto X_{i+1}.

Los vértices X_{i+1}, X_i y $f(X_i)$ forman un triángulo, y la tangente del ángulo α indicado en la figura vale:

$$tan\ \alpha = \frac{f(X_i)}{X_i - X_{i+1}} \qquad (11)$$

y como $tan\ \alpha = f'(X_i)$; se puede escribir:

$$\frac{f(X_i)}{X_i - X_{i+1}} = f'(X_i) \qquad (12)$$

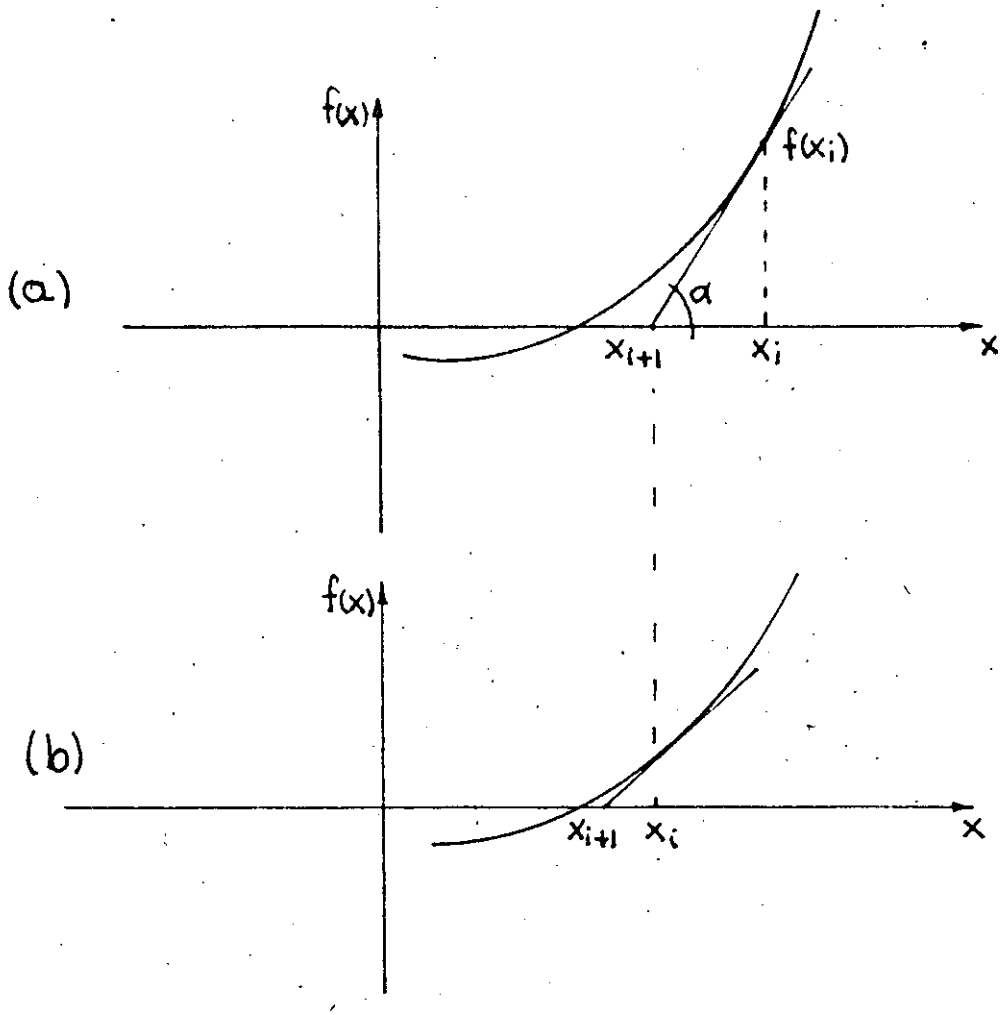


Figura 9 Esquematización de una iteración del método de Newton-Raphson

Despejando a X_{i+1} se obtiene:

$$X_{i+} = X_i - \frac{f(X_i)}{f'(X_i)} \qquad (13)$$

Si en la siguiente iteración se sustituye el valor de X_i, por el de X_{i+1} de la anterior, el valor de la función se aproximará más a la raíz, y con un número suficiente de cálculos, se obtendrá la solución, dada una tolerancia.

El algoritmo del método de Newton-Raphson, puesto en forma de diagrama de bloques, se muestra en la figura 10.

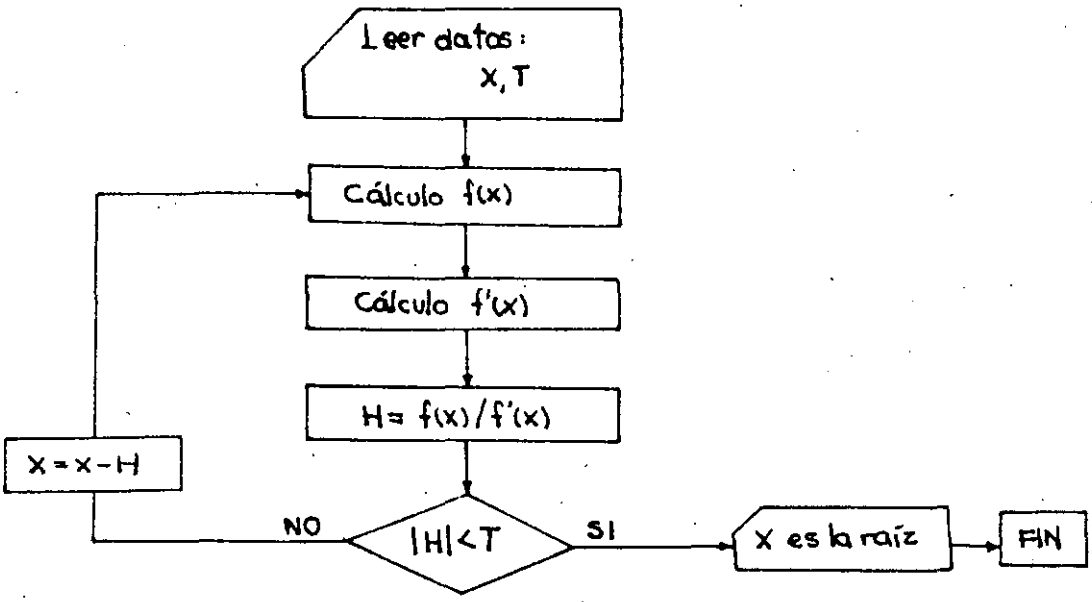


Figura 10.- Diagrama de flujo para el método de Newton-Raphson

Debe observarse que, contrariamente a los métodos anteriores, solo se requiere un valor inicial para comenzar el cálculo.

Para ilustrar la aplicación de este método, considérese el siguiente problema típico en hidráulica.

PROBLEMA:

Calcular el tirante normal de un canal trapecial, para un gasto $Q = 0.75\ m^3/s$, que tiene ancho de plantilla $b = 1.70\ m$ talud $K = 2$ y pendiente $S_o = 0.001$, y coeficiente de rugosidad $n = 0.10$.

Solución:

De la ecuación de Manning:

$$V = \frac{1}{n} R^{2/3} S_o^{1/2} \qquad (14)$$

donde:

R = radio hidráulico

V = velocidad

de la ecuación de continuidad:

$$V = \frac{Q}{A} \qquad (15.a)$$

y

$$R = \frac{A}{P} \qquad (15.b)$$

donde:

A = área hidráulica = $(b + Ky)y$

P = perímetro mojado = $b + 2y\sqrt{1 + K^2}$

Sustituyendo las ecuaciones 15 en las 14 y ordenando:

$$\frac{Qn}{S_o^{1/2}} = \frac{A^{5/3}}{P^{2/3}}$$

es decir que la función a resolver es:

$$\delta(y) = \frac{A^{5/3}}{P^{2/3}} - \frac{Qn}{S_o^{1/2}} = 0 \qquad (16)$$

en la cual A y P son funciones de "Y". Derivando para la aplicación del método de Newton-Raphson:

$$\delta'(Y) = A^{5/3} - \left[\frac{2}{3} P^{-5/3} \frac{dP}{dy}\right] + P^{-2/3}\left[\frac{5}{3} A^{\frac{2}{3}} \frac{dA}{dy}\right]$$

y ordenando:

$$\delta'(y) = \frac{A^{2/3}}{P^{2/3}}\left[\frac{5}{3}\frac{dA}{dy} - \frac{2}{3}\frac{A}{P}\frac{dP}{dy}\right] \qquad (17)$$

y de las expresiones para área hidráulica y perímetro mojado:

$$\frac{dP}{dy} = 2\sqrt{1 + K^2} \qquad (18.a)$$

$$\frac{dA}{dy} = b + 2Ky \qquad (18.b)$$

Entonces, la ecuación recursiva del método se puede escribir:

$$y_{i+1} = y_i - \frac{\delta(y)}{\delta'(y)}$$

donde $\delta(Y)$ se calcula con la ecuación 17 y $\delta'(Y)$ con las ecuaciones 17 y 18.

En la figura 11 se presenta un programa, en lenguaje Basic, para el cálculo del tirante normal con el método de Newton-Raphson. Debe notarse que, por comodidad, en este programa se han definido cuatro funciones; una para el área hidráulica, otra para el perímetro mojado, una más para calcular la función del problema (ec. 16), y otra para su derivada (ec. 17).

Ejecutando este programa para los datos del problema se obtendría:

```
PROGRAMA PARA CALCULO DE TIRANTE NORMAL
      METODO DE NEWTON-RAPHSON
      ***************
      GASTO (m^3/S)?
      0.75
      ANCHO DE PLANTILLA (m)?
      1.70
      TALUD?
      2
```

```
10 CLEAR @ DISP "PROGRAMA CALCU
   LO DE TIRANTE NORMAL"
20 DISP @ DISP "METODO DE NEWTO
   N RAPHSON"
30 DISP "***************"
40 REM -ENTRADA DE DATOS-
50 DISP "GASTO(M^3/S)";@ INPUT
   Q
60 DISP "ANCHO DE PLANTILLA(M)"
   ;@ INPUT B
70 DISP "TALUD";@ INPUT K
80 DISP "COEFICIENTE DE MANNING
   ";@ INPUT N
90 DISP "PROPONGA UN TIRANTE(M)
   ";@ INPUT Y0
100 DISP " PENDIENTE DEL CANAL";
   @ INPUT S
110 DISP "TOLERANCIA";@ INPUT T
120 REM SE DEFINEN FUNCIONES A=(
   X)=AREA;P(X)=PERIMETRO M.
130 REM F(Y)=FUNCION A RESOLVER
   D(Y)=DERIVADA
140 DEF FNA(X) = (B+K*X)*X
150 DEF FNP(X) = B+2*X*SQR(1+K^2
   )
160 DEF FNF(Y) = (FNA(Y)^5/FNP(Y
   )^2)^(1/3)-Q*N/SQR(S)
170 DEF FND(Y) = (FNA(Y)/FNP(Y))
   ^(2/3)*(5/3*(B+2*K*Y)+2/3*(F
   NA(Y) FNP(Y))*2*SQR(1+K^2))
180 REM SE INICIA ALGORITMO DE N
   EWTON-RAPHSON
190 H=FNF(Y0)/FND(Y0)
200 IF ABS(H)<T THEN GOTO 240
210 Y0=Y0-H
220 GOTO 190
230 REM SE ENCONTRO LA RAIZ; SE
   INFORMA AL USUARIO
240 CLEAR @ BEEP @ DISP "*******
   *****"
250 DISP "TIRANTE NORMAL (M)" @
   DISP "Y=";Y0 @ DISP "*******
   *****"
260 END
```

Figura 11.- Programa para el cálculo de tirante normal de un canal trapecial, método de Newton-Raphson

```
COEFICIENTE DE MANNING?
0.10
PROPONGA UN TIRANTE (m)?
0.45
PENDIENTE DEL CANAL?
0.001
TOLERANCIA?
0.000001

******************
TIRANTE NORMAL (m)
Y = 0.9513742
******************
```

El método de Newton-Raphson es mucho más rápido que los métodos de bisección y falsa posición; sin embargo no siempre es convergente.

El método de Newton-Raphson es convergente si se cumple que:

$$\frac{|\delta(X)\ \delta''(X)|}{(\delta'(X))^2} < 1 \qquad (19)$$

Recomendaciones de tipo cualitativo serían que, si se elige un valor X_o como punto inicial, se debe cumplir que:

a) X_o debe ser suficientemente cercano a la raíz

b) $\delta''(X_o)$ no debe ser muy grande

c) $\delta'(X_o)$ no muy próximo a cero

3.5 METODO DE LA SECANTE

El método de la secante es una variante del método de Newton -Raphson, útil para casos en los que la derivada de la función, cuya raíz se busca, es complicada.

En la figura 12 se presenta la gráfica de una función $f(X)$, a la que se desea calcular la raíz, Se han elegido dos valores X_1 y X_2; a los que corresponden valores de la función $f(X_1)$ y $f(X_2)$.

Los valores de X_1 y X_2 no tienen la restricción de que $f(X_1)$ $f(X_2) < 0$.

El arco que une $f(X_1)$ y $f(X_2)$ corta el eje X en un valor X_3 y, como se observa en la figura, se pueden formar dos triángulos semejantes con vértices en X_3, X_2, $f(X_2)$ y $f(X_1)$, c, $f(X_2)$. Entonces se puede plantear la relación:

$$\frac{f(X_2) - f(X_1)}{X_2 - X_1} = \frac{f(X_2)}{X_2 - X_3}$$

y despejando X_3:

$$X_3 = X_2 - \frac{f(X_2)}{f(X_2) - f(X_1)} (X_2 - X_1) \qquad (20)$$

Si el valor de X_3 no se aproxima suficientemente a la raíz, se sustituye en X_2, y se repite el cálculo. Con este procedimiento X_3 se aproximará gradualmente a la solución.

En la figura 13 se presenta un diagrama de bloques del método de la secante. Para ejemplificar su aplicación, considérese un problema de interés en hidráulica: el cálculo del tirante crítico en un canal trapecial.

El tirante crítico se presenta cuando el número de Froude es igual a la unidad, condición que puede escribirse como:

$$\frac{Q^2}{g\frac{A^3}{T}} = 1 \qquad (21)$$

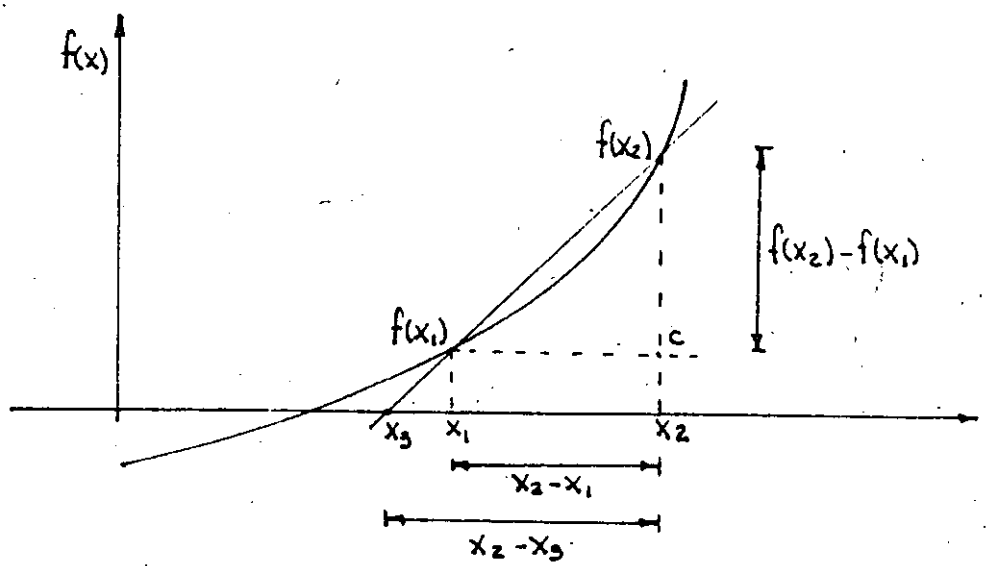


Figura 12.- Esquematización para el método de la secante

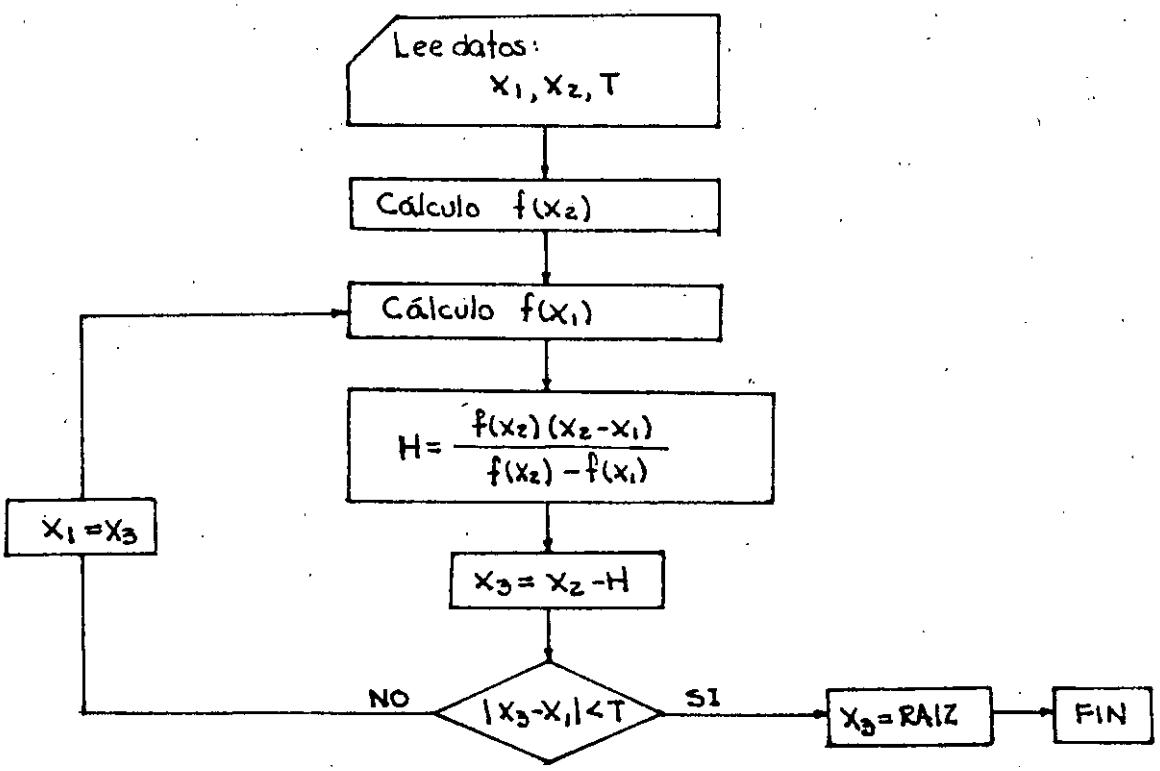


Figura 13.- Diagrama de bloques método de la secante

donde T es el ancho de la superficie libre, que para un canal trapecial está determinada por la ecuación:

$$T = b + 2Ky \qquad (22)$$

La ecuación 21 se puede escribir:

$$\frac{A^3}{T} - \frac{Q^2}{g} = 0$$

y sustituyendo las expresiones de A y T:

$$\delta(y) = \frac{[(b + Ky)y]^3}{b + 2Ky} - \frac{Q^2}{g} = 0 \qquad (23)$$

Esta ecuación es la que debe resolverse para encontrar el tirante crítico.

En la figura 14 se presenta un programa para la solución de este problema con el método de la secante.

Utilizando los mismos datos que en el ejemplo de cálculo del tirante normal, la ejecución de este programa sería como sigue:

```
PROGRAMA CALCULO DE TIRANTE CRITICO
         METODO DE LA SECANTE
         ******************

         GASTO (m^3/S)?
         0.75
         ANCHO DE PLANTILLA (m)?
         1.70
```

```
10 CLEAR @ DISP "PR
   LO DE TIRANTE CR.
20 DISP "METODO DE LA
   @ DISP "***********
30 DISP @ DISP "GASTO(M
    INPUT Q
40 DISP "ANCHO DE PLANTILL
   A";@ INPUT B
50 DISP "TALUD" @ INPUT K
60 DISP "PROGRAMA DOS TIRANTE
   Y1,Y2(M)";@ INPUT Y1,Y2
70 DISP " TOLERANCIA",@ INPUT T
80 DEF FNF(X)= ((B+K*X)*X)^3/(
   B+2*K*X)-Q^2/9.81
90 F2=FNF(Y2)
100 F1=FNF(Y1)
110 H=F2*(Y2-Y1)/(F2-F1)
120 Y3=Y2-H
130 IF ABS(Y3-Y1)<T THEN GOTO 15
    0
140 Y1=Y3 @ GOTO 100
150 CLEAR @ BEEP @ DISP "*******
    *****"
160 DISP "TIRANTE CRITICO" @ DIS
    P "Y=";Y3
170 DISP "***********" @ BEEP
180 END
```

Figura 14.- Programa para cálculo de tirante crítico, método de la secante

```
TALUD?
2
PROPONGA DOS TIRANTES Y1, Y2(m)?
0.20, 0.40
TOLERANCIA?
0.0001
```

```
***************
TIRANTE CRITICO
y = 0.2445
***************
```

Las características del método de la secante son similares a las del método de Newton-Raphson; aunque tiene una velocidad de convergencia ligeramente menor.

Existen otros métodos de solución de ecuaciones algebráicas que pueden encontrarse en textos generales de métodos numéricos .

DIVISION DE EDUCACION CONTINUA
FACULTAD DE INGENIERIA U.N.A.M.

INGENIERIA MARITIMA. MODULO:
METODOS NUMERICOS EN HIDRAULICA
DEL 1o. DE JULIO AL 6 DE SEPTIEMBRE 1985
MEXICO, D.F.

SOLUCION NUMERICA DE SISTEMAS DE ECUACIONES LINEALES

M.EN I. OSCAR FUENTES MARILES
M.EN I. POLIOPTRO MARTINEZ A.

3. SOLUCION NUMERICA DE SISTEMAS DE ECUACIONES LINEALES

3.1 INTRODUCCION

La solución numérica de sistemas de ecuaciones lineales es uno de los problemas fundamentales del análisis numérico, no solamente porque existen muchos problemas que, por si mismos conducen a un sistema de este tipo; sino también porque otros métodos numéricos, como los esquemas implícitos de diferencias finitas, quedan finalmente expresados como un conjunto de ecuaciones lineales.

Los sistemas de ecuaciones lineales tienen solución analítica conocida, como se recordará más adelante; y sin embargo también existen varios métodos numéricos de solución, y se producen continuamente otros. La razón de esta situación es que, dada la magnitud de los sistemas a resolver, y su variedad; debe elegirse entre estos el más eficiente para un problema dado; con el fin de minimizar el tiempo de computo.

En este capítulo se presentarán algunos de los métodos numéricos más conocidos de solución de sistemas de ecuaciones lineales; y se ejemplificará su aplicación a un problema importante en hidráulica: la solución de una red de distribución de agua potable.

3.2 SOLUCION ANALITICA DE SISTEMAS LINEALES

Un sistema lineal es un sistema de ecuaciones de la forma:

Ecuación 1: $a_{11} X_1 + a_{12} X_2 + \text{-------} + a_{1n} X_n = b_1$

Ecuación 2: $a_{21} X_2 + a_{22} X_2 + \text{-------} + a_{2n} X_n = b_2$ (1)

$\text{------------------------------------}$

Ecuación m: $a_{m1} X_1 + a_{m2} X_2 + \text{-------} + a_{mn} X_n = b_m$

c. $a \cdot (b\,A) = (a \cdot b)A$

d. $1 \cdot A = A$

3. Producto de dos matrices

Sean A (a_{ij}) una matríz (m, p) y B (b_{ij}) una matríz (p, n). Su producto C : $C = A \cdot B$ se define como:

$$C_{ij} = a_{i1}\, b_{1j} + a_{i2}\, b_{2j} + \text{---} + a_{ip}\, b_{pj} = \sum_{k=1}^{p} a_{ik} \cdot b_{kj}$$

Propiedades:

a. El producto es asociativo:

$$A \cdot (B \cdot C) = (A \cdot B) \cdot C$$

b. El producto no es conmutativo

$$A \cdot B \neq B \cdot A$$

c. Existe una matríz cuadrada I (la matríz identidad) tal que:

$$I \cdot A = A$$

d. Existen matrices no nulas cuyo producto es la matríz nula

e. El producto es distributivo respecto a la suma:

$$A \cdot (B + C) = AB + AC$$

$$(A + C) \cdot B = A \cdot B + C \cdot B$$

f. El producto por un escalar (a):

$$a(A \cdot B) = A(a \cdot B)$$

La multiplicación de matrices puede ser un cálculo prolongado, dado que cada elemento de la matríz resultado es una sumatoria, que requiere realizar p multiplicaciones y p adiciones. Conviene contar con una subrutina que haga esta p operación; como se muestra en la figura 1.

Para estudiar la solución analítica de un sistema de ecuaciones lineales se requiere recordar previamente algunos conceptos.

Dada una matríz A (m, n); con elementos a_{ij}, su transpuesta A^T (n, m) se obtiene realizando la operación:

$$a^T_{ij} = a_{ji}$$

Es decir intercambiando renglones por columnas de la matríz A.

Una matríz simétrica es una que cumple que $A = A^T$.

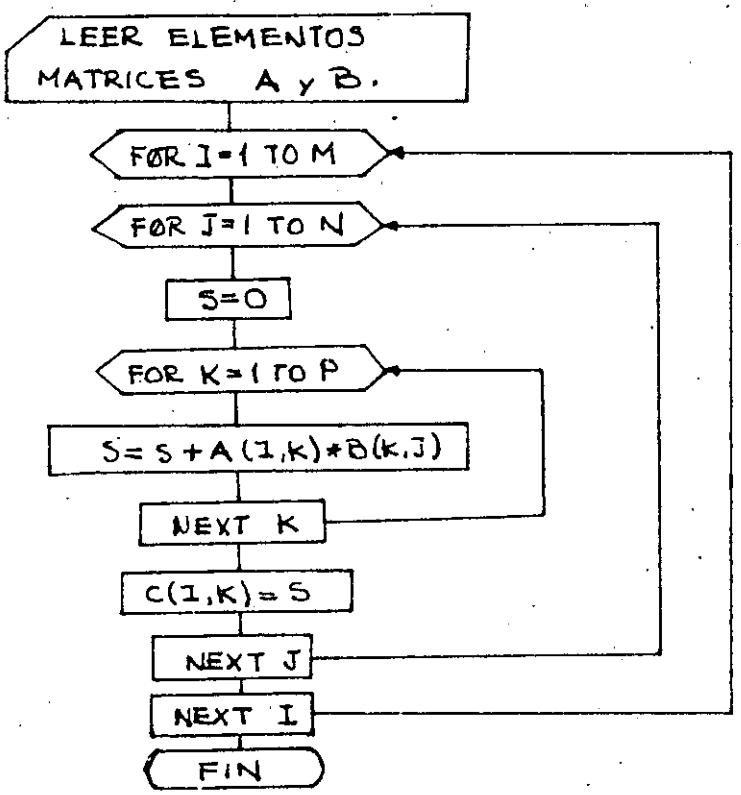


Figura 1. Diagrama de flujo para una subrutina de multiplicación de matrices

Considérese una matríz cuadrada (n, n), A. Denótese por M_{ij} la submatríz cuadrada que se obtiene quitando a A el i-ésimo renglón y la j-ésima columna.

El determinante de esta submatríz ($|M_{ij}|$) se denomina "el menor" del elemento a_{ij} de A; y se define "el cofactor" de a_{ij}, denotado A_{ij}, como el menor con signo dado por la operación:

$$A_{ij} = (-1)^{i+j} |M_{ij}| \qquad (3)$$

Los cofactores integran una nueva matríz, la matríz de cofactores, que se denotará por A^c.

Se define la matríz adjunta como la transpuesta de la de cofactores: $(A^c)^T$.

Si A es una matríz cuadrada su inversa se denota por A^{-1}; con la que se cumple:

$$A \cdot A^{-1} = I$$

y se puede demostrar que:

$$A^{-1} = \frac{1}{|A|} \cdot (A^c)^T \qquad (4)$$

Entonces, si se tiene un sistema lineal $A \cdot \bar{X} = \bar{b}$; este sistema tiene solución (si A es invertible) para cualquier b, dada por:

$$\bar{X} = A^{-1} \cdot \bar{b} \qquad (5)$$

o bien, de acuerdo a lo anterior:

$$\bar{X} = \frac{1}{A} \cdot (A^c)^T \cdot \bar{b} \qquad (5.a)$$

Si existe esta solución cabría preguntar sobre la necesidad de disponer de métodos numéricos como solución alternativa. Este requerimiento se origina en la gran cantidad de operaciones que se deben efectuar para cumplir con (5.a).

Para resolver el sistema se necesitan:

a. Determinar $|A|$ requiere $n \cdot n!$ multiplicaciones y adiciones

b. A^C tiene n^2 terminos, y para calcular cada cofactor se requieren $(n-1) \cdot (n-1)!$ multiplicaciones y adiciones

c. Efectuar la operación $(A^C)^T$ requiere n^2 multiplicaciones

En total se requiere de $n^3(n-1)!$ operaciones; que resulta imposible aún para sistemas pequeños.

3.3 METODO DE GAUSS

Los métodos numéricos de solución de sistemas de ecuaciones lineales se clasifican en dos tipos: directos e indirectos. Los primeros son aquellos en los que la solución se obtiene de manera inmediata, mientras que en los segundos se requieren varias iteraciones.

Los métodos directos aprovechan para la solución las propiedades de un sistema con matríz de coeficientes diagonal o triangular.

Supóngase, en el caso más sencillo, que se tiene un sistema de ecuaciones $A\bar{X} = \bar{b}$; y la matríz A es una matríz diagonal $(a_{ij} = 0 \text{ si } i \neq j)$:

$-m_{31} R_1 + R_3$: $\quad | \; 0 \quad 7 \quad 7.333 \; \vdots \; 41.6667 \; |$

Entonces el nuevo R_3 es:

R_3: $\quad | \; 0 \quad 7 \quad 7.333 \; \vdots \; 41.6667 \; |$

La nueva matríz aumentada es:

$$\begin{bmatrix} 3 & 1.5 & 4 & \vdots & 8 \\ 0 & 0.125 & 0.72 & \vdots & 0.27 \\ 0 & 7 & 7.333 & \vdots & 41.6667 \end{bmatrix}$$

Paso 6. Se hace $i = i+1 = 1+1 = 2$

como $i < n$ $(2 < 3)$ se regresa al paso 3

Paso 3. $a_{ii} = a_{22} = 0.125 \neq 0$. Se continúa otra vez al paso 5

Paso 5. Para $j = i+1 = 3$

$$m_{ji} = m_{32} = \frac{a_{32}}{a_{22}} = \frac{7}{0.125} = 56$$

La operación $(R_j - m_{ji} R_i) \rightarrow R_j$

es: $(R_3 - m_{31} R_2) \rightarrow R_3$

Numericamente:

$m_{32} R_2$: $\quad 56 \; | \; 0 \quad 0.125 \quad 0.72 \; \vdots \; 0.27 \; |$

$-m_{32} R_2$: $\quad | \; 0 \quad -7 \quad -40.32 \; \vdots \; -15.12 \; |$

R_3: | 0 7 7.333 ⋮ 41.667 |

| 0 0 -32.98 ⋮ 26.5467 |

Paso 6. Se hace $i=i+1 = 2+1=3$ y como $i=n$; se continúa al paso 7

La matríz aumentada final es:

$$\left[\begin{array}{ccc:c} 3 & 1.5 & 4 & 8 \\ 0 & 0.125 & 0.72 & 0.27 \\ 0 & 0 & -32.98 & 26.5467 \end{array}\right]$$

Paso 7. La aplicación del procedimiento de vuelta atrás permite calcular la solución:

$$X_3 = \frac{26.5467}{32.98} = -0.805$$

$$X_2 = \frac{(0.27 - 0.72(-.805))}{0.125} = 6.797$$

$$X_1 = \frac{(8 - 4(-.805) - 1.5(6.797)}{3} = 0.3415$$

Paso 8. La solución del sistema es:

$$X_3 = -.805 \qquad X_2 = 6.797 \qquad X_1 = 0.3415$$

La programación del algoritmo del método de Gauss es relativamente sencilla. En la figura 3 se presenta el listado de un programa en BASIC, y se incluye en las páginas siguientes un breve manual de usuario.

Utilizando este programa los resultados del ejemplo anterior serían:

$$X_1 = 0.3419 \qquad X_2 = 6.795 \qquad X_3 = -.8047$$

Como puede verse, existe una ligera diferencia respecto al resultado antes obtenido; y se debe a que la computadora opera con un mayor número de cifras decimales. El error cometido se denomina "de redondeo", y está presente en todos los métodos numéricos.

3.4 METODO DE JACOBI

El procedimiento de Gauss, con ser más eficiente en cuanto a tiempo de cálculo que los procedimientos analíticos, puede aún dar orígen a pérdidas de eficiencia en determinados casos; por ejemplo cuando la matríz de coeficientes tiene muchos términos nulos.

En situaciones como la anterior se puede utilizar algún método indirecto; que son métodos de naturaleza iterativa, es decir probando a partir de un vector inicial propuesto $\bar{X}_o$.

```
  5 CLEAR
 10 DISP "ESTE PROGRAMA RESUELVE
    UN  SISTEMA DE N ECUACIONES
    CON EL ALGO--RITMO DE GAUSS
    "
 15 DISP "********************"
 20 DISP "LOS COEFICIENTES DE LA
    MATRIZ AUMENTADA SE INTRODU
    CEN EN FORMA  DE RENGLON"
 25 DISP "********************"
 30 DISP "TECLEE EL RANGO DE LA
    MATRIZ DE COEFICIENTES"
 40 DISP "N=?"
 50 INPUT N
 60 DIM A(20,21),X(20),C(1,21),M
    (21)
 70 FOR I=1 TO N
 80 DISP "TECLEE LOS COEFICIENTE
    S DEL RENGLON";I
 90 FOR J=1 TO N+1
100 DISP "A(";I;",";J;")=" @ INP
    UT A(I,J)
110 NEXT J
120 NEXT I
130 I=1
140 IF A(I,I)#0 THEN GOTO 260
150 P=1
160 IF A(I,P)#0 THEN GOTO 190
170 IF P>=N THEN GOTO 410
180 P=P+1 @ GOTO 160
190 FOR J=1 TO N+1
200 C(I,J)=A(I,J)
210 NEXT J
220 FOR J=1 TO N+1
230 A(I,J)=A(P,J)
240 A(P,J)=C(I,J)
250 NEXT J
260 FOR J=I+1 TO N
270 M(J)=A(J,I)/A(I,I)
280 NEXT J
285 FOR J=I+1 TO N
290 FOR P=1 TO N+1
292 A(J,P)=A(J,P)-M(J)*A(I,P)
296 NEXT P
298 NEXT J
300 I=I+1
310 IF I<N THEN GOTO 140
320 IF A(I,I)=0 THEN GOTO 410
330 X(N)=A(N,N+1)/A(N,N)
335 FOR I=N-1 TO 1 STEP -1
340 S=0
350 FOR J=I+1 TO N
360 S=S+A(I,J)*X(J)
370 NEXT J
380 X(I)=(A(I,N+1)-S)/A(I,I)
390 NEXT I
400 GOTO 420
410 CLEAR @ DISP "EL SISTEMA NO
    TIENE" @ DISP " SOLUCION UNI
    CA" @ GOTO 460
```

Figura 3. Programa método de Gauss

```
420 CLEAR @ DISP "*** LA SOLUCIO
    N DEL SISTEMA ES "
430 FOR I=1 TO N
440 DISP "X(";I;")=";X(I)
450 NEXT I
460 DISP "**********************
    ***" @ END
```

Figura 3. Continuación --------

La idea fundamental de la mayoría de los métodos iterativos consiste en transformar el sistema original:

$$A\bar{X} = \bar{b}$$

en otro con la forma

$$\bar{X}_1 = T\bar{X}_2 + \bar{C} \qquad (11)$$

que de manera recursiva quedaría

$$\bar{X}^K = T\bar{X}^{K-1} + \bar{C} \qquad (12)$$

siendo K el número de iteración.

Considérese por ejemplo el sistema:

$$\begin{aligned} 10X_1 - 7X_2 + X_3 &= 73 \\ X_1 + 8X_2 - 3X_3 &= 23 \\ X_1 + 3X_2 - 9X_3 &= 57 \end{aligned} \qquad (13)$$

Entonces se tiene:

$$A = \begin{bmatrix} 10 & -7 & 3 \\ 1 & 8 & -3 \\ 1 & 3 & -9 \end{bmatrix} \qquad \bar{b} = \begin{Bmatrix} 73 \\ 23 \\ 57 \end{Bmatrix}$$

Si se despeja en cada ecuación i, la i-ésima variable se obtiene:

$$\begin{aligned} X_1 &= 7.3 + 0.7X_2 - 0.1X_3 \\ X_2 &= 2.875 - .125X_1 + .375X_3 \qquad (14) \\ X_3 &= -6.333 + 0.111X_1 + 0.333X_2 \end{aligned}$$

Es decir que en términos de (12), el sistema (14) tiene T y $\bar{C}$:

$$T = \begin{bmatrix} 0 & 0.7 & -0.1 \\ -.125 & 0 & .375 \\ .111 & .333 & 0 \end{bmatrix} \qquad \bar{C} = \begin{Bmatrix} 7.3 \\ 2.875 \\ -6.333 \end{Bmatrix}$$

y se pueden escribir las ecuaciones recursivas:

$$\begin{aligned} X_1^K &= 7.3 + 0.7X_2^{K-1} - 0.1X_3^{K-1} \\ K_2^K &= 2.875 - .125X_1^{K-1} + .375X_3^{K-1} \qquad (14.a) \\ K_3^K &= -6.33 + .111X_1^K + .333X_2^K \end{aligned}$$

El método podría mejorarse si para calcular X_i^K se utilizaran los X_{i-1}^K ya calculados anteriormente.

La ecuación iterativa para este cálculo sería:

$$X_i^K = \frac{-\sum_{j=i}^{i-1} (a_{ij} X_j^K) - \sum_{j=i+1}^{n} (a_{ij} X_j^{K-1}) + b_i}{a_{ii}} \qquad (16)$$

Este método es conocido como el de Gauss-Seidel; y su algoritmo, muy semejante al de Jacobi; quedaría como sigue:

ALGORITMO METODO DE GAUSS-SEIDEL

Para resolver el sistema $A\bar{X} = \bar{b}$

Paso 1. Proponga una solución inicial $\bar{X}^\circ = (X_1^\circ, X_2^\circ, ---, X_n^\circ)$

Paso 2. Sea $K = 1$

Paso 3. Para todo $i = 1, 2, 3, ---, n$ calcule:

$$X_i^K = \frac{-\sum_{j=i}^{i-1} (a_{ij} X_j^K) - \sum_{j=i+1}^{n} (a_{ij} X_j^{K-1}) + b_i}{a_{ii}}$$

Paso 4. Si $\bar{X}^K$ es suficientemente exacto continúe a paso 5; en caso contrario haga $K = K + 1$ y regrese a paso 3

Paso 5. Procedimiento concluido

Para ejemplificar la aplicación de este método; considérese el sistema antes resuelto con el método de Jacobi; y que ya en la forma $\bar{X}^K = T\bar{X}^{K-1} + \bar{C}$, está dado por:

$$X_1^K = 7.3 + 0.7\,X_2^{K-1} - 0.1\,X_3^{K-1}$$

$$X_2^K = 2.875 - 0.125\,X_1^{K-1} + 0.375\,X_3^{K-1} \qquad (14.a)$$

$$X_3^K = -6.333 + 0.111\,X_1^K + 0.333\,X_2^K$$

El cálculo se organiza en la siguiente tabla:

Iteración K	X_1^K	X_2^K	X_3^K	
0	1	1	1	← Propuesto
1	7.9	2.2625	-4.7027	
2	9.3540	-0.0578	-5.3140	
3	7.7809	-0.0914	-5.4998	
4	7.9599	-0.1824	-5.5102	
5	7.7233	-0.1567	-5.5279	
6	7.7431	-0.1659	-5.5288	
7	7.7368	-0.1654	-5.5293	

Se encuentra la siguiente solución aproximada:

$$X_1 = 7.7368 \qquad X_2 = -.1654 \qquad X_3 = -5.5293$$

y como puede verse se reduce el número de iteraciones.

Es conveniente antes de seguir adelante establecer criterios para determinar cuando un método será convergente; para lo cual se presentarán algunas definiciones y teoremas, a continuación.

Definición:

Una secuencia $\{\bar{X}^K\}_{K=1}^{\infty}$ de vectores se dice que converge a $\bar{X}$, con respecto a la norma $\|\cdot\|$, si dado cualquier $\varepsilon > 0$, existe

un entero $N(\varepsilon)$ tal que:

$$\| \bar{X}^K - \bar{X} \| < \varepsilon \quad \text{para todo} \quad K \geq N(\varepsilon)$$

Pude demostrarse que cualquier norma cumple esta definición.

La norma infinita de una matríz A está definida como:

$$\| A \|_\infty = \max_{1 \leq i \leq n} \sum_{j=1}^{n} |a_{ij}| \qquad (17)$$

Es decir que la norma infinita de una matríz es la máxima suma de los valores absolutos de los elementos de cada renglón. Por ejemplo la norma infinita de la matríz del ejemplo anterior:

$$A = \begin{bmatrix} 10 & -7 & 1 \\ 1 & 8 & -3 \\ 1 & 3 & -9 \end{bmatrix}$$

Se calcula como:

Suma en renglón 1: $10 + 7 + 1 = 18$
Suma en renglón 2: $1 + 8 + 3 = 12$
Suma en renglón 3: $1 + 3 + 9 = 13$

Entonces $\| A \|_\infty = 18$

Se puede demostrar el siguiente teorema:

La secuencia $\{\bar{X}^K\}_{K=0}^{\infty}$, definida por $\bar{X}^K = T\bar{X}^{K-1} + \bar{C}$; para todo $\bar{C} \neq 0$ y $K \geq 1$; converge al vector $\bar{X}$ para cualquier vector inicial $\bar{X}^\circ$; si $\| T \| < 1$, para cualquier norma.

En el ejemplo anterior:

$$T = \begin{bmatrix} 0 & 0.7 & -0.1 \\ -.125 & 0 & 0.375 \\ .111 & .333 & 0 \end{bmatrix}$$

es decir que su norma infinita es:

$$\|T\|_\infty = 0.8 < 1$$

y por lo tanto el método de Jacobi o Gauss-Seidel convergen para el sistema de ecuaciones.

La revisión de convergencia puede fácilmente introducirse en un programa del algoritmo de Gauss-Seidel.

En la figura 4 se presenta un programa para solución de sistemas de ecuaciones con el método de Gauss-Seidel.

3.6 METODO DE SOBRERRELAJACION (SOR)

Tanto en el procedimiento de Jacobi como en el de Gauss-Seidel existe un vector residual; diferencia entre el vector solución real y el vector de aproximación.

Denótese:

$$e_i^K = (e_1{}_i^K,\ e_2{}_i^K, \ldots, e_n{}_i^K) \qquad (18)$$

el vector residual del método de Gauss-Seidel; correspondiente al vector de la K-ésima aproximación:

$$(X_i^K,\ X_2{}^K, \ldots,\ X_{i-1}^K,\ X_i{}^{K-1},\ X_{i-11}^K, \ldots X_n{}^{K-1}) \qquad (19)$$

```
  2   CLEAR @ DISP "**********
      **********"
  5 DISP @ DISP "PROGRAMA PARA S
    OLUCION DE SISTEMA DE ECUAC
    IONES" @ DISP "**************
    ******"
  7 DISP "METODO DE GAUSS-SEIDEL
    " @ DISP "********************
    **"
 10 DIM X(50),A(50,50),B(50),T(5
    0),X1(50),H(50),H1(50)
 20 DISP @ DISP "EL RANGO DEL SI
    STEMA ES:"
 30 INPUT N
 40 DISP @ DISP "ELEMENTOS DE LA
    MATRIZ " @ DISP "DE COEFICI
    ENTES"
 50 FOR I=1 TO N
 60 FOR J=1 TO N
 70 INPUT A(I,J)
 80 NEXT J
 90 NEXT I
100 DISP "ELEMENTOS DEL VECTOR D
    E" @ DISP "TERMINOS INDEPEND
    IENTES"
110 FOR I=1 TO N
120 INPUT B(I)
130 NEXT I
140 DISP "VALORES INICIALES DE X
    (I)"
150 FOR I=1 TO N
160 INPUT X(I)
170 NEXT I
180 DISP "TOLERANCIA"
190 INPUT E
195 CLEAR @ DISP "**************
    *" @ DISP "PROCEDO A VERIFIC
    AR CONVERGENCIA" @ DISP "***
    *****ESPERE"
196 WAIT 3000
200 FOR I=1 TO N
210 S=0
220 FOR J=1 TO N
230 IF J=I THEN GOTO 250
240 S=S+ABS(A(I,J)/A(I,I))
250 NEXT J
260 T(I)=S
270 NEXT I
290 L1=T(1)
300 FOR I=2 TO N
310 F=L1-T(I)
320 IF F>0 THEN GOTO 340
330 L1=T(I)
340 NEXT I
350 IF L1>1 THEN GOTO 740
355 CLEAR @ DISP "**************
    "
360 CLEAR @ DISP "**************"
     @ DISP "EL SISTEMA CONVERGE
    "
```

Figura 4. Programa para solución de sistemas de ecuaciones lineales con el método de Gauss-Seidel

```
370 DISP @ DISP "  NOMBRE ES
    "
375 DISP @ DISP "  INICIA GAUS
    -SEIDEL" @ DISP "**********
    **********ESPERE"
380 WAIT 3000
390 K=1
400 FOR I=1 TO N
410 S=0
420 IF J=I THEN GOTO 450
430 FOR J=1 TO I-1
440 S=S-A(I,J)*X1(J)
450 NEXT J
460 S1=0
470 IF I=N THEN GOTO 500
480 FOR J=I+1 TO N
490 S1=S1-A(I,J)*X(J)
500 NEXT J
510 X1(I)=(S+S1+B(I))/A(I,I)
520 NEXT I
530 C=ABS(X1(1)-X(1))
540 C1=ABS(X1(1))
550 FOR I=2 TO N
560 H(I)=ABS(X1(I)-X(I))
570 H1(I)=ABS(X1(I))
580 NEXT I
590 FOR J=2 TO N
600 G=H(J)-C
610 IF G<0 THEN GOTO 630
620 C=H(J)
630 NEXT J
640 FOR K=2 TO N
650 G1=H1(K)-C1
660 IF G1<0 THEN GOTO 680
670 C1=H1(K)
680 NEXT K
690 D=C/C1
700 IF D<E THEN GOTO 745
710 FOR I=1 TO N
720 X(I)=X1(I)
730 NEXT I
740 K=K+1 @ GOTO 390
745 CLEAR @ DISP "**********" @
    DISP "EL SISTEMA HA CONV
    E" @ GOTO 790
750 CLEAR @ DISP "*******
    ********  @ DISP "LA SOL
       ES "
760 FOR I = 1
770 PRINT "X(
780 NEXT I
790 PRINT "*********
    END
```

Figura 4. .....Continuación

4. ECUACIONES DIFERENCIALES ORDINARIAS

Una ecuación diferencial ordinaria es una ecuación que contiene derivadas ordinarias. Una ecuación tal puede escribirse en términos de diferenciales pero generalmente no es conveniente a menos que la ecuación contenga solamente la primera derivada.

Se dice que se ha resuelto o integrado una ecuación diferencial que contiene a x,y y derivadas de y respecto a x, cuando se ha encontrado una función de x y y que no contiene derivadas, que sustituida en la ecuación diferencial, la reduce a una identidad.

El orden de una ecuación diferencial es igual al de la derivada de más alto orden que hay en ella.

Una ecuación diferencial ordinaria es lineal si contiene x y si y y las derivadas de y respecto a x aparecen a la primer potencia. La forma general de una ecuación diferencial ordinaria lineal de orden n es

$$b_0(x)\frac{d^n y}{dx^n} + b_1(x)\frac{d^{n-1} y}{dx^{n-1}} + \ldots + b_{n-1}(x)\frac{dy}{dx} + b_n(x)y = R(x) \qquad (4.1)$$

Por ejemplo, es lineal

$$x^2 \frac{d^2y}{dx^2} + x \frac{dy}{dx} + (x^2-4)\, y = 5x^4$$

y no es lineal

$$x^2\left(\frac{dy}{dx}\right)^2 + s \frac{dy}{dx} + (x^2-4)y = 5x^4$$

Para encontrar la solución de las ecuaciones diferenciales ordinarias existen muchos procedimientos analíticos, como son el de separación de variables, factor integrante, variación de parámetros entre otros.

Desafortunadamente, muchas, a caso la mayoría, de las ecuaciones diferenciales que se presentan en la práctica no pueden ser integradas por métodos analíticos o cuando lo son, el obtener su integral es muy complicado. Sin embargo, ecuaciones del tipo

$$\frac{dy}{dx} = f(x,y) \quad , \; y = y_a \text{ en } x = x_a \tag{4.2}$$

pueden ser integradas numéricamente. Para conocer la solución en el intervalo $x_a \leq x \leq x_b$, se puede dividir tal intervalo en N intervalos de ancho Δx (fig 4.1), y al considerar que $x_n = \Delta x\, n + x_a$, $y_n = y(x_n)$ se plantea la solución de 4.2 como

$$\int_{y_a}^{y_1} dy = \int_{x_a}^{x_1} f(x,y)dx$$

$$y_1 = y_a + \int_{x_a}^{x_1} f(x,y)dx \quad \text{y del mismo modo}$$

$$y_2 = y_1 + \int_{x_1}^{x_2} f(x,y)dx$$

$$y_3 = y_2 + \int_{x_2}^{x_3} f(x,y)dx \tag{4.3}$$

$$\vdots$$

$$y_{n+1} = y_n + \int_{x_n}^{x_{n+1}} f(x,y)dx$$

$$\vdots$$

$$y_N = Y_{N-1} + \int_{x_{n-1}}^{x_n} f(x,y)dx$$

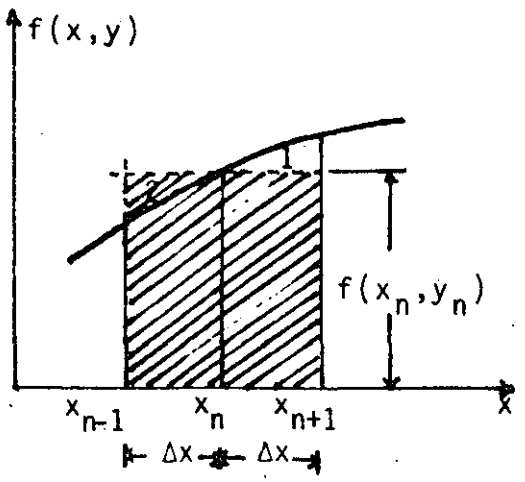


Fig 4.4

Por lo anterior.

$$\int_{x_{n-1}}^{x_{n+1}} f(x,y)dx \doteq f(x_n,y_n)\ 2\Delta x$$

como la integral abarca dos intervalos Δx, ahora se tiene

$$\boxed{y_{n+1} = y_{n-1} + 2f(x_n,y_n)\Delta x} \qquad (4.6)$$

Se observa que esta fórmula trata de compensar el área bajo la curva entre x_{n-1} y y_{n+1} no cubierta con el rectángulo (área identificada con 1) con el área del rectángulo donde $f(x,y)$ es menor a $f(x_n, y_n)$ (área señalada con 2) por lo que es una adecuada representación de la integral.

Este método tiene el inconveniente que no se puede valuar y_1, pero si y_2, y_3,..., etc. Para estimar y_1 se recomienda utilizar la ec. 4.5. Una vez conocida y_1 y con la condición inicial y_0, ya no se tienen limitaciones para utilizar la ec. 4.6.

4.4 Método basado en la serie de Taylor

El método de Euler o bien lo comentado en el subcapítulo 4.2 hace pensar que una estimación de $f(x_{n+1},y_{n+1})$ permite representar la integral de una mejor manera. En efecto, considerese que la fig 4.3 ahora se presenta como en la fig 4.5.

Sea el desarrollo en serie de Taylor

$$y(x+\Delta x) = y(x) + \Delta x y'(x) + \frac{\Delta x^2}{2!} y''(x) + \frac{\Delta x^3}{3!} y''(x) + \ldots$$

si $x = x_n$, $\Delta x = x_{n+1} - x_n$, entonces

---

* Hornbeck R.W., "Numerical Methods", Ed, Quantum, QPI series, USA, 1975.

$$y(x_{n+1}) = y(x_n) + \Delta x\, y'(x_n) + \frac{\Delta x^2}{2!} y''(x_n) + \frac{\Delta x^3}{3!} y'''(x_n) + \ldots \tag{4.7}$$

De acuerdo con la ec. 4.2 se tiene

$$y' = f(x,y)$$

$$y'' = \frac{df}{dx} = f' = \frac{\partial f}{\partial x}\frac{dx}{dx} + \frac{\partial f}{\partial y}\frac{dy}{dx} = \frac{\partial f}{\partial x} + \frac{\partial f}{\partial y}\, y'$$

$$y''' = \frac{df'}{dx} = f''$$

$$\vdots$$

$$y^{(p)} = \frac{df^{(P-2)}}{dx} = f^{(P-1)}$$

Sustituyendo los resultados anteriores en 4.7

$$y(x_{n+1}) = y(x_n) + \Delta x f(x_n,y_n) + \frac{\Delta x^2}{2!} f'(x_n,y_n) + \frac{\Delta x^3}{3!} f''(x_n,y_n) + \ldots \tag{4.8}$$

Si se considera que

$$T^{(P)}(x_n,y_n) = f(x_n,y_n) + \frac{\Delta x}{2!} f'(x_n,y_n) + \frac{\Delta x^2}{3!} f''(x_n,y_n) + \ldots +$$

$$+ \frac{\Delta x^{P-1}}{P!} f^{(P-1)}(x_n,y_n) \tag{4.9}$$

La ec (4.8) se puede escribir como

$$\boxed{y(x_{n+1}) = y(x_n) + \Delta x\, T^{(P)}(x_n,y_n)} \tag{4.10}$$

La ec. anterior corresponde al llamado método de Taylor de orden p. Nótese que el método de Euler corresponde al caso especial en que P=1. En la figura 4.5 se muestra que $T^{(P)}(x_n,y_n)$ corresponde a una estimación de la ordenada del área sombreada

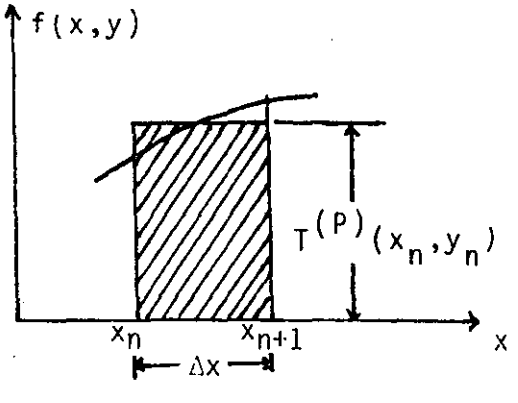


Fig 4.5

El método de Taylor tiene en ocasiones el inconveniente del cálculo de las derivadas f', f", f"' ,... las cuales pueden ser difíciles de evaluar o complicadas de calcular.

4.5 Métodos de Runge-Kutta

La dificultad para conocer el valor de las derivadas $f'(x_n,y_n)$, $f''(x_n,y_n)$, $f'''(x_n,y_n)$,... del método de Taylor, ha sido salvada por Runge (1895) y Kutta (1901) a través de un procedimiento basado en valuar varias veces la función f(x, y) obteniendo una precisión equivalente al método de Taylor. Aparte de ello, los métodos de Runge-Kutta tienen la ventaja de usar una fórmula de suma pesada, similar a la utilizada en integración numérica, con lo cual se logra una adecuada aproximación de área bajo la curva f(x,y) entre x_n y x_{n+1}.

A continuación se hará la derivación del procedimiento conocido como Runge-Kutta de tercer orden.

El problema consiste en planear como ecuación del método a

$$y_{n+1} = y_n + \Delta x\,(ak_1 + bk_2 + ck_3) \qquad (4.11)$$

donde

$$k_1 = f(x_n,\ y_n) \qquad (4.12)$$

$$k_2 = f(x_n + m\Delta x,\ y_n + m\Delta x k_1) \qquad (4.13)$$

$$k_3 = f[x_n + p\Delta x,\ y_n + \Delta x(qk_1 + (p-q)k_2)] \qquad (4.14)$$

e interesa conocer los valores de a,b,c,m,p y q. Los cuales son únicos e independientes de la ecuación diferencial por resolver.

Sea el desarrollo de la serie de Taylor 4.7 que incluye hasta términos de tercer orden

$$y_{n+1} = y_n + \Delta x\, y'_n + \frac{\Delta x^2}{2} y''_n + \frac{\Delta x^3}{6} y'''_n \qquad (4.15)$$

Según lo planteado en el subcapítulo 5.4 se tiene

$$y' = f$$

$$y'' = \frac{df}{dx} = \frac{\partial f}{\partial x} + \frac{\partial f}{\partial y} = f_x + f_y f$$

$$y''' = \frac{df'}{dx} = \frac{d}{dx}(f_x + f_y f) = f_{xx} + 2f_{xy} f + f_{yy} f^2 + f_y (f_x + f_y f)$$

si

$$A = f_x + f_y f \qquad (4.16)$$

y

$$B = f_{xx} + 2f_{xy}\, f + f_{yy} f^2 \qquad (4.17)$$

entonces

$$y'' = A$$

$$y''' = B + f_y A$$

y por lo tanto la ec. 4.15 también se escribe como

$$y_{n+1} = y_n + \Delta x f + \frac{\Delta x^2}{2} A + \frac{\Delta x^3}{6} B + \frac{\Delta x^3}{6} f_y A \qquad (4.18)$$

En esta última ecuación se entiende que f, f_y, A y B se calculan para $x = x_n$ y $y = y_n$

Por otra parte, el desarrollo en serie de Taylor de una función de las variables x y y es

$$f(x+h,y+k) = f(x,y) + h\, f_x(x,y) + k\, f_y(x,y) + \frac{h^2}{2} f_{xx}(x,y) + hk\, f_{xy}$$

$$(x,y) + \frac{k^2}{2} f_{yy}(x,y) \qquad (4.19)$$

Así, al desarrollar 4.13 (siendo $h = m\Delta x$ y $k = m\Delta xf$) se tiene

$$y_{n+1} = y_n + \frac{\Delta x}{6}(k_1 + 2k_2 + k_3) \tag{4.24a}$$

donde

$$k_1 = f(x_n, y_n) \tag{4.24b}$$

$$k_2 = f(x_n + \Delta x/2,\ y_n + \frac{\Delta x}{2} k_1) \tag{4.25c}$$

$$k_3 = f(x_n + \Delta x, y_n + 2\Delta x k_2 - \Delta x k_1) \tag{4.24d}$$

Las ecs. 4.24 corresponden al método de Runge-Kutta de orden tres.

De manera semejante se pueden deducir las ecuaciones del método de Runge-Kutta de orden cuatro:

$$y_{n+1} = y_n + \frac{\Delta x}{6}(k_1 + 2k_2 + 2k_3 + k_4) \tag{4.25a}$$

donde

$$k_1 = f(x_n,\ y_n) \tag{4.25b}$$

$$k_2 = f(x_n + \frac{\Delta x}{2},\ y_n + \frac{\Delta x k_1}{2}) \tag{4.25c}$$

$$k_3 = f(x_n + \frac{\Delta x}{2},\ y_n + \frac{\Delta x k_2}{2}) \tag{4.25d}$$

$$k_4 = f(x_n + \Delta x,\ y_n + \Delta x k_3) \tag{4.25e}$$

y Runge-Kutta de quinto orden:

$$y_{n+1} = y_n + \Delta x\left(\frac{23}{192}k_1 + \frac{125}{192}k_2 - \frac{81}{192}k_5 + \frac{125}{192}k_6\right) \tag{4.26a}$$

donde

$$k_1 = f(x_n, y_n) \tag{4.26b}$$

$$k_2 = f(x_n + \frac{1}{3}\Delta x,\ y_n + \frac{1}{3}\Delta x k_1) \tag{4.26c}$$

$$k_3 = f(x_n + \frac{2}{5}\Delta x,\ y_n + \frac{4}{25}\Delta x k_1 + \frac{16}{25}\Delta x k_2) \tag{4.26d}$$

$$k_4 = f(x_n + \Delta x,\ y_n + \tfrac{1}{2}\Delta x k_1 - 3\Delta x k_2 + \frac{15}{4}\Delta x k_3) \tag{4.26e}$$

$$k_5 = f(x_n + \frac{2}{3}\Delta x,\ y_n + \frac{2}{27}\Delta x k_1 + \frac{10}{9}\Delta x k_2 - \frac{50}{81}\Delta x k_3 + \frac{8}{81}\Delta x k_4) \tag{4.26f}$$

$$k_6 = f(x_n + \frac{4}{5}\Delta x,\ y_n + \frac{2}{25}\Delta x k_1 + \frac{12}{25}\Delta x k_2 + \frac{2}{15}\Delta x k_3 + \frac{-8}{75}\Delta x k_4) \tag{4.26g}$$

4.6 Fórmulas de Adams

El desarrollo de Taylor planteado en 4.8 también permite un enfoque distinto para resolver numéricamente la ecuación diferencial ordinaria 4.2. Se ocurre ahora, basarse en la idea de integración numérica. En general, se propondrán dos clases distintas de ecuaciones, unas donde es explícito el cálculo de y_{n+1} (se llaman cerradas), y las otras en las que se requiere de un método iterativo (se denominan abiertas). Ambos casos corresponden a las fórmulas de Adams.

4.6.1 Fórmulas abiertas de Adams

Considere la expresión 4.8

$$y_{n+1} = y_n + \Delta x f_n + \frac{\Delta x^2}{2!} f_n' + \frac{\Delta x^3}{3!} f_n'' + \ldots \qquad (4.8)$$

o bien

$$y_{n+1} = y_n + \Delta x \left(f_n + \frac{\Delta x}{2!} f_n' + \frac{\Delta x^2}{3!} f_n'' + \ldots\right) \qquad (4.27)$$

Si la serie incluye hasta términos de derivadas de primer orden:

$$y_{n+1} = y_n + \Delta x \left(f_n + \frac{\Delta x}{2} f_n'\right) \qquad (4.28)$$

Ahora, si f_n' se aproxima por medio de una diferencia hacia atrás (cap. 6)

$$fn' = \frac{df}{dx} = \frac{fn - f_{n+1}}{\Delta x} + \frac{\Delta x}{2} f_n'' + \sigma(\Delta x)^2 \qquad (4.29)$$

Sustituyendo 4.29 (sin sus términos de segundo orden) en 4.28

$$\boxed{y_{n+1} = y_n + \Delta x \left[\frac{3}{2} f_n - \frac{1}{2} f_{n-1}\right]} \qquad (4.30)$$

la cual es la fórmula abierta de Adams de segundo orden

Cuando en la expresión 4.27 se considera hasta la derivada de segundo orden:

$$y_{n+1} = y_n + \Delta x\left(f_n + \frac{\Delta x}{2} f'_n + \frac{\Delta x^2}{6} f''_n\right) \qquad (4.31)$$

considerando la ec. 4.29

$$f''_n = \frac{df'}{dx} = \frac{f'_n - f'_{n-1}}{\Delta x} \qquad (4.32)$$

Según 4.29

$$f'_{n-1} \doteq \frac{f_{n-1} - f_{n-2}}{\Delta x} \qquad (4.33)$$

Sustituyendo 4.29 y 4.33 en 4.32

$$f''_n = \frac{f_n - 2f_{n-1} + f_{n-2}}{\Delta x^2} \qquad (4.34)$$

Si ahora, en la ec. 4.29 se considera la derivada de segundo orden y esta se sustituye por 4.34:

$$f'_n = \frac{f_n - f_{n-1}}{\Delta x} + \frac{\Delta x}{2}\left(\frac{f_n - 2f_{n-1} + f_{n-2}}{\Delta x^2}\right) + \sigma(\Delta x)^2$$

Después de simplificar:

$$f'_n = \frac{3}{2} f_n - \frac{3}{2} f_{n-1} + \frac{f_{n-2}}{2} + \sigma(\Delta x)^2 \qquad (4.35)$$

Si se desprecia el término $\sigma(\Delta x)^2$ de 4.35 y se sustituyen 4.35 y 4.34 en 4.31 se obtiene

$$y_{n+1} = y_n + \Delta t\left[f_n + \left(\frac{3}{4} f_n - \frac{3}{4} f_{n-1} + \frac{f_{n-2}}{4}\right) + \left(\frac{f_n}{6} - \frac{2f_{n-1}}{6} + \frac{f_{n-2}}{6}\right)\right]$$

o bien

$$y_{n+1} = y_n + \Delta t \left[\frac{23}{12} f_n - \frac{16}{12} f_{n-1} + \frac{5}{12} f_{n-2}\right] \qquad (4.36)$$

la cual es la fórmula abierta de Adams de tercer orden.

De manera similar se podría obtener la fórmula abierta de Adams de cuarto orden:

$$y_{n+1} = y_n + \Delta t \left[\frac{55}{24} f_n - \frac{59}{24} f_{n-1} + \frac{37}{24} f_{n-2} - \frac{9}{24} f_{n-3}\right] \qquad (4.37)$$

Estas expresiones abiertas también se conocen con el nombre de Fórmulas de Adams-Bashforth. Como tienen la desventaja de no iniciarse por sí mismas, es decir para emplear en un principio a 4.30 se conoce f_0 de las condiciones iniciales $y_0 = y(x_0)$ pero no a f_{-1}, se recomienda utilizarlo a partir de n=1, como en este caso no se sabe el valor de f_1 se sugiere aplicar la ec. 4.5 para conocer y_1 y luego con este valor y $x_1 = x_0 + \Delta x$ valuar f_1, definidas f_0 y f_1 ya no habrá dificultad en utilizar la ec. 4.30

Se recomienda aplicar 4.36 a partir de n=2 y valuar previamente f_1 y f_2 calculando y_1 y y_2 por medio de las ecs. 4.24. Al igual para usar 4.37 es conveniente empezar con n=3 y calcular y_1, y_2 y y_3 por medio de las ecs. 4.25, con ellas se obtienen f_1, f_2 y f_3.

4.6.2 Fórmulas cerradas de Adams

Sea el desarrollo de la serie de Taylor

$$y(x-\Delta x) = y(x) - \Delta x y'(x) + \frac{\Delta x^2}{2!} y''(x) - \frac{\Delta x^3}{3!} y'''(x) + \ldots$$

si $x = x_{n+1}$, $\Delta x = x_{n+1} - x_n$, entonces

$$y(x_n) = y(x_{n+1}) - \Delta x y'(x_{n+1}) + \frac{\Delta x^2}{2!} y''(x_{n+1}) - \frac{\Delta x^3}{3!} y'''(x_{n+1}) + \ldots$$

Como en el subcapítulo 4.4, $y'_{n+1} = f_{n+1}$, $y''_{n+1} = f''_{n+1}$, y''_{n+1} etc. y resolviendo para y_{n+1}, se encuentra:

$$Y_{n+1} = y_n + x[f_{n+1} - \frac{\Delta x}{2} f'_{n+1} + \frac{\Delta x^2}{3!} f''_{n+1} - \frac{\Delta x^3}{4!} f'''_{n+1} + \ldots] \tag{4.38}$$

Como antes, si se considera hasta la primera derivada

$$y_{n+1} = y_n + \Delta x[f_{n+1} - \frac{\Delta x}{2} f'_{n+1}] \tag{4.39}$$

Despreciando los términos de segundo orden, de 4.29 se tiene:

$$f'_{n+1} = \frac{f_{n+1} - f_n}{\Delta x} \tag{4.40}$$

Al sustituir 4.40 en 4.39 y simplificando

$$\boxed{y_{n+1} = y_n + \Delta x[\tfrac{1}{2} f_{n+1} + \tfrac{1}{2} f_n]} \tag{4.41}$$

esta ecuación corresponde a la fórmula cerrada de Adams de segundo orden.

Siguiendo un razonamiento semejante a los del inciso 4.6.1 se obtienen las fórmulas cerradas de Adams de tercer (4.42) y cuarto orden 4.43

$$\boxed{y_{n+1} = y_n + \Delta x[\frac{5}{12} f_{n+1} + \frac{8}{12} f_n - \frac{1}{2} f_{n-1}]} \tag{4.42}$$

$$\boxed{y_{n+1} = y_n + \Delta x[\frac{9}{24} f_{n+1} + \frac{19}{24} f_n - \frac{5}{24} f_{n-1} + \frac{1}{24} f_{n-2}]}$$

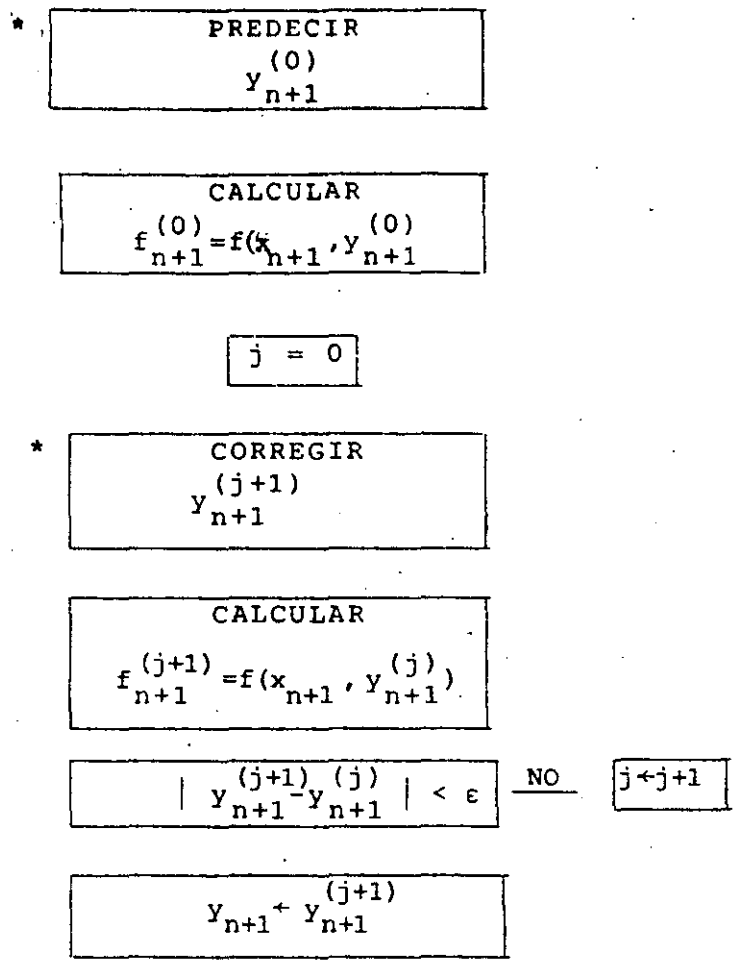
(4.43)

Se observa que en las ecs. cerradas aparte de la dificultad de no iniciarse por sí mismas, que no se conoce f_{n+1}, por lo que se propone resolverlas por iteraciones, y una vez que ya no hay problema con el principio (ver inciso 4.6.1), se resuelven por aproximaciones sucesivas proponiendo un valor de y_{n+1}, con el cual se valúa f_{n+1} y al sustituir en la fórmula cerrada en cuestión se obtiene y_{n+1}, si este es aproximadamente igual al supuesto se ha encontrado y_{n+1} y se incrementa el valor de n para continuar con el siguiente Δx ; en caso contrario se necesita escoger otro valor de y_{n+1} y se repite el proceso.

4.7 Métodos Predictor-Corrector

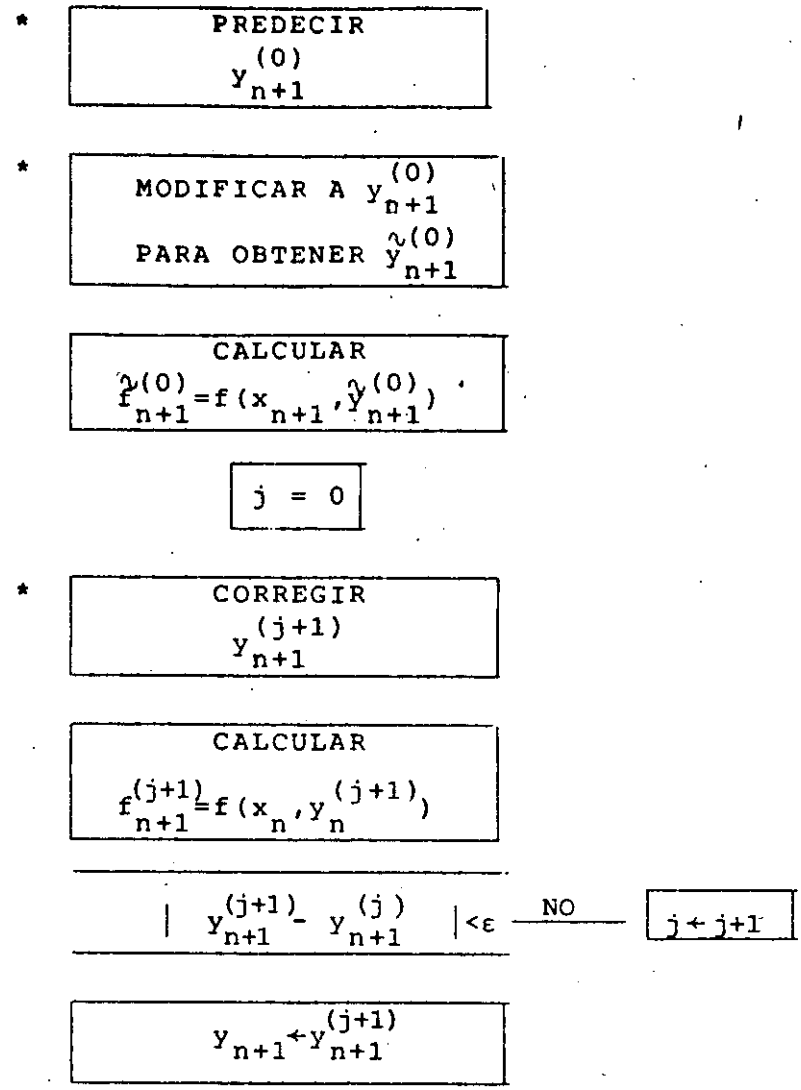
Una ventaja de las fórmulas cerradas de Adams es su precisión, sin embargo, en ocasiones el proceso iterarivo se hace largo y se contraresta esta ventaja. Cuando el valor propuesto a y_{n+1} no es muy diferente del correcto, el número de iteraciones se reduce en forma importante y nuevamente hace útil la fórmula cerrada.

De esto se desprende la idea de escoger un valor inicial de y_{n+1} adecuado. Para ello se sugiere utilizar una ecuación diferente a la fórmula cerrada. Esta ecuación servirá para "predecir" el valor de y_{n+1} con el cual se comienzan las iteraciones. Luego con él se emplea la fórmula cerrada; como en cada iteración se mejora el valor de y_{n+1} se dice que se esta "corrigiendo" este. Por esto a esta clase de procedimientos se les conoce con el nombre de "métodos predictor-corrector". En la fig. 4.7 se muestra la forma de utilizar estos métodos.



* Se usan las fórmulas predictor-corrector
Fig. 4.7

También se ha observado que incluyendo una ecuación que modifique la estimación del predictor, el número de iteraciones se reduce, inclusive en muchos de los casos sólo se requiere una iteración; para esta variante, el diagrama del método predictor-corrector queda como se muestra en la fig. 4.8.



Continuar con el siguiente intervalo Δx.

* Se usan las fórmulas predictor, modificador corrector.

Fig. 4.8

Entre los métodos predictor-corrector se anotan los siguientes:

Método de Adams

Predictor:

$$y_{n+1}^{(0)} = y_n + \Delta x \left[\frac{55}{24} f_n - \frac{59}{24} f_{n-1} + \frac{37}{24} f_{n-2} - \frac{9}{24} f_{n-3}\right] \quad (4.44)$$

Corrector:

$$y_{n+1}^{(j+1)} = y_n + \Delta x\left[\frac{9}{24} f_{n+1}^{(j)} + \frac{19}{24} f_n - \frac{5}{24} f_{n-1} + \frac{1}{24} f_{n-2}\right] \quad (4.45)$$

(Estas ecuaciones ya fueron discutidas en el subcapítulo 4.6, corresponden a las 4.37 y 4.43).

Método de Milne

Predictor:

$$y_{n+1}^{(0)} = y_n + \Delta x\left[\frac{8}{3} f_n - \frac{4}{3} f_{n-1} + \frac{8}{3} f_{n-2}\right] \quad (4.46)$$

Corrector:

$$y_{n+1}^{(j+1)} = y_n + \Delta x\left[\frac{1}{3} f_{n+1}^{(j)} + \frac{4}{3} f_n + \frac{1}{3} f_{n-1}\right] \quad (4.47)$$

Método de Hamming

Predictor:

$$y_{n+1}^{(0)} = y_{n-3} + \Delta x\left[\frac{8}{3} f_n - \frac{4}{3} f_{n-1} + \frac{8}{3} f_{n-2}\right] \quad (4.48)$$

Modificador:

$$\tilde{y}_{n+1}^{(0)} = y_{n+1}^{(0)} + \frac{112}{121}\left(y_n - y_n^{(0)}\right) \quad (4.49)$$

Corrector:

$$y_{n+1}^{(j+1)} = \frac{1}{8}\left(9y_n - y_{n-2}\right) + \Delta x\left(\frac{3}{8}\tilde{f}_{n+1}^{(j)} + \frac{6}{8} f_n - \frac{3}{8} f_{n-1}\right) \quad (4.50)$$

Se observa que las fórmulas predictoras no se inician por sí mismas (ver comentarios a las fórmulas 4.36 y 4.37). Para disponer de los valores iniciales necesarios para su aplicación se recomienda utilizar los métodos de Runge-Kutta de cuarto orden.

Una de las ventajas de los métodos predictor-corrector estriba en el hecho de que casi siempre se requiere una iteración y que, por tanto, se requieren menos cálculos que en los métodos de Runge-Kutta (nótese que para el método de Adams se requiere calcular $y_{n+1}^{(0)}$ y con este a $f_{n+1}^{(0)}$ mientras que para las ecs. 4.25 se necesita valuar k_2, k_3 y k_4 implicando mas operaciones aritméticas)

También los métodos predictor-corrector tienen un aspecto a su favor en lo referente al cálculo del error que se comete con ellas, pues la forma de determinar este error es muy simple.

4.8 Métodos de parámetros indeterminados

Dentro de esta categoría de procedimientos se agrupa el método basado en el cálculo de variaciones (Ritz) y el de Galerkin.

4.8.1 Método de Ritz

Cuando un alambre doblado en forma de una circunferencia, primero se introduce en una solución jabonosa y luego se extrae, se observa que se forma una delgada película de jabón, formando una superficie. Este experimento inspira el siguiente problema: Dada una curva cerrada encontrar la superficie limitada por la misma de modo tal que su área sea mínima.

En Cálculo Diferencial se trata como encontrar un punto donde la función es máxima o mínima. Ahora no se desea definir un punto, sino una función que cumpla con ciertas condiciones y que haga máxima o mínima una propiedad; esto último se estudia por medio del cálculo de variaciones.

Algunos de los problemas del Cálculo de Variaciones consiste en encontrar la función (curva) que une dos puntos dados y que maximiza o minimiza una integral.

Ejemplo

Encontrar el arco $y(x)$ que pasa a través de los puntos $(0,1)$ y $(1,2)$, que minimiza

$$J = \int_0^1 \frac{\sqrt{1+y'^2}}{y}\, dx$$

Generalizando lo anterior, se desea incontrar una función $y(x)$ tal que $y_1 = y(x_1)$ y $y_2 = y(x_2)$ de manera que para una función dada $F(x,y,y')$, la integral

$$J = \int_{x_1}^{x_2} F(x,y,y')dx \qquad (4.51)$$

sea máxima o mínima. La integral que toma un valor numérico para ciertas funciones $y(x)$ se llama funcional.

Para encontrar la función $y(x)$ se propone la familia de funciones

$$Y(x) = y(x) + \varepsilon\eta(x) \qquad (4.52)$$

donde

$$\eta(x_1) = \eta(x_2) = 0 \qquad (4.53)$$

entonces

$$Y(x) = y(x) + \varepsilon\eta(x) \qquad (4.54)$$

$$Y'(x) = y'(x) + \varepsilon\eta'(x) \qquad (4.55)$$

se observa que para $\varepsilon=0$ se tiene la función que hace mínima a 4.51 si se reemplaza y y y' en 4.51 respectivamente por Y y Y', se forma la integral:

$$J(\varepsilon) = \int_{x_1}^{x_2} f(x,Y,Y')dx \qquad (4.56)$$

Para encontrar un extremo (máximo o mínimo) de $J(\varepsilon)$, se deriva respecto a ε y se iguala a cero, tal como sucede en el Cálculo Diferencial, así:

$$\frac{dJ(\varepsilon)}{d\varepsilon} = \int_{x_1}^{x_2} \left(\frac{\partial F}{\partial Y}\frac{\partial Y}{\partial \varepsilon} + \frac{\partial F}{\partial Y'}\frac{\partial Y'}{\partial \varepsilon}\right)dx = 0 \qquad (4.57)$$

y el mínimo es precisamente cuando:

$$\frac{dJ(0)}{d\varepsilon} = 0 \qquad (4.58)$$

según 4.54 y 4.55

$$\frac{\partial Y}{\partial \varepsilon} = \eta(x) \qquad (4.59)$$

$$\frac{\partial Y'}{\partial \varepsilon} = \eta'(x) \qquad (4.60)$$

considerando 4.59, 4.60, y 4.54 y 4.55 para $\varepsilon=0$, se tiene:

$$\frac{\partial J(0)}{\partial \varepsilon} = \int_{x_1}^{x_2} \left(\frac{\partial F}{\partial y}\eta + \frac{\partial F}{\partial y'}\eta'\right)dx \qquad (4.61)$$

$$= \int_{x_1}^{x_2} \frac{\partial F}{\partial y}\eta\, dx + \int_{x_1}^{x_2} \frac{\partial F}{\partial y'}\eta'\, dx$$

integrando por partes la segunda integral:

$$\int_{x_1}^{x_2} \frac{\partial F}{\partial y'}\eta'dx = \frac{\partial F}{\partial y'}\eta\Big]_{x_1}^{x_2} - \int_{x_1}^{x_2} \eta \frac{d}{dx}\left[\frac{\partial F}{\partial y'}\right]dx$$

dientes p(x) y q(x) tales que cumplen con:

$$\int_{x_1}^{x_2} p(x)q(x)\,dx = 0$$

son ortogonales en el intervalo $x_1 \leq x \leq x_2$

El método de Galekin consiste en los siguientes pasos:

1. Sean la ecuación diferencial

$$L(y) - f(x) = 0 \qquad (4.66)$$

y

$$y(x_1) = y_1 \quad y \quad y(x_2) = y_2 \qquad (4.67)$$

2. Escoger un conjunto de funciones linealmente independientes $u_0(x), u_1(x), u_2(x), \ldots, u_n(x)$ donde $u_0(x)$ satisface las condiciones 4.67 y $u_1(x), u_2(x), \ldots, u_n(x)$ se anulan en (x_1, y_1) y (x_2, y_2).

3. Se forma la solución

$$y = u_0(x) + a_1 u_1(x) + a_2 u_2(x) + \ldots + a_n u_n(x) \qquad (4.68)$$

4. Sustituyendo 4.68 en 4.66

$$L[u_0(x) + a_1 u_1(x) + a_2 u_2(x) + \ldots + a_n u_n(x)] - f(x) = R(x)$$

5. Las constantes $a_1, a_2, \ldots, a_n$ se encuentran al considerar que R(x) es ortogonal con las funciones $u_1(x), u_2(x), \ldots, u_n(x)$.

Esto es

$$\int_{x_1}^{x_2} R(x)u_1(x)dx = 0$$

$$\int_{x_1}^{x_2} R(x)u_2(x)dx = 0 \qquad (4.69)$$

$$\vdots$$

$$\int_{x_1}^{x_2} R(x)u_n(x)dx = 0$$

6. Al resolver el sistema de ecuaciones 4.69 se encuentran $a_1, a_2, \ldots, a_n$.

7. Sustituir los valores de $a_1, a_2, \ldots, a_n$ en 4.68, con lo cual queda definida la solución aproximada.

4.9 Ecuaciones diferenciales de orden mayor a uno.

Una ecuación diferencial de orden mayor a uno o un sistema de ecuaciones diferenciales que involucran algunas derivadas de orden alto, pueden deducirse a un conjunto de ecuaciones de primer orden haciendo un cambio de variable simple. La ecuación de orden n:

$$y^{(n)} = f(x,y,y',y'',\ldots y^{(n-1)}) \tag{4.70}$$

se transforma haciendo

$$\begin{aligned} y &= g_0 \\ y' &= g_1 \\ y'' &= g_1' = g_2 \\ y''' &= g_1'', = g_2' = g_3 \\ y^{(n)} &= \ldots = g_{n-1}' = f(x,y,y',y'',\ldots,y^{(n-1)}) \end{aligned} \tag{4.71}$$

Las ecuaciones anteriores se pueden tomar con cualquiera de los métodos descritos. El cálculo se hace en paralelo, se realiza el cálculo para el primer Δx para todas las ecuaciones antes de pasar al siguiente, y así sucesivamente.

En el caso especial del método de Runge-Kutta se especifica para el caso de dos ecuaciones ordinarias de primer orden, el conjunto de ecuaciones siguientes. Ellas corres-

a un método de orden cuatro.

Sean $\frac{dy}{dx} = f(x,y,u)$ (4.72)

$\frac{du}{dx} = h(x,y,u)$ (4.73)

donde y yu son conocidas para $x = x_n$. En particular, las fórmulas de Runge-Kutta se generalizan como sigue:

$$y_{n+1} = y_n + \Delta x(\tfrac{1}{6}k_1 + \tfrac{1}{3}k_2 + \tfrac{1}{3}k_3 + \tfrac{1}{6}k_4) \quad (4.74)$$

$$u_{n+1} = u_n + \Delta x(\tfrac{1}{6}m_1 + \tfrac{1}{3}m_2 + \tfrac{1}{3}m_3 + \tfrac{1}{6}m_4) \quad (4.75)$$

con

$$k_1 = f(x_n, y_n, u_n) \quad (4.76a)$$

$$k_2 = f(x_n + \tfrac{\Delta x}{2},\ y_n + \tfrac{1}{2}\Delta x k_1,\ u_n + \tfrac{1}{2}\Delta x m_1) \quad (4.76b)$$

$$k_3 = f(x_n + \tfrac{\Delta x}{2},\ y_n + \tfrac{1}{2}\Delta x k_2,\ u_n + \tfrac{1}{2}\Delta x m_2) \quad (4.76c)$$

$$k_4 = f(x_n + \Delta x,\ y_n + \Delta x k_3,\ u_n + \Delta x m_3) \quad (4.76d)$$

y

$$m_1 = h(x_n, y_n, u_n) \quad (4.77a)$$

$$m_2 = h(x_n + \tfrac{\Delta x}{2},\ y_n + \tfrac{1}{2}\Delta x k_1,\ u_n + \tfrac{1}{2}\Delta x m_1) \quad (4.77b)$$

$$m_3 = h(x_n + \tfrac{\Delta x}{2},\ y_n + \tfrac{1}{2}\Delta x k_2,\ u_n + \tfrac{1}{2}\Delta x m_2) \quad (4.77c)$$

$$m_4 = h(x_n + \Delta x,\ y_n + \Delta x k_3\ u_n + \Delta x\ m_3) \quad (4.77d)$$

Una consideración de esta forma indica, como cualquiera de las otras fórmulas de los métodos pueden ser usadas.

4.10 Errores en los métodos numéricos para resolver ecuaciones diferenciales.

En la solución de las ecuaciones diferenciales, se entenderá que la diferencia entre la solución exacta de la ecuación diferencial (por ejemplo aquella obtenida por métodos análiticos con todas las cifras decimales) menos la solución obtenida con un método numérico con un número limitado de cifras corresponde al error total E(x).

$$E(x) = S(x) - p(x) \tag{4.78}$$

donde

S(x) es la solución exacta

p(x) es la solución mediante el método numérico con un número limitado de cifras.

Introduciendo Q(x), la solución mediente el método numérico con todas las cifras necesarias, en la ec. 4.78:

$$E(x) = S(x) - Q(x) + Q(x) - p(x)$$

Llamando error de truncado o discretización a D(x) = S(x) - Q(x) y error de redondeo a R(x) = Q(x) - p(x) se tiene

$$E(x) = D(x) + R(x) \tag{4.79}$$

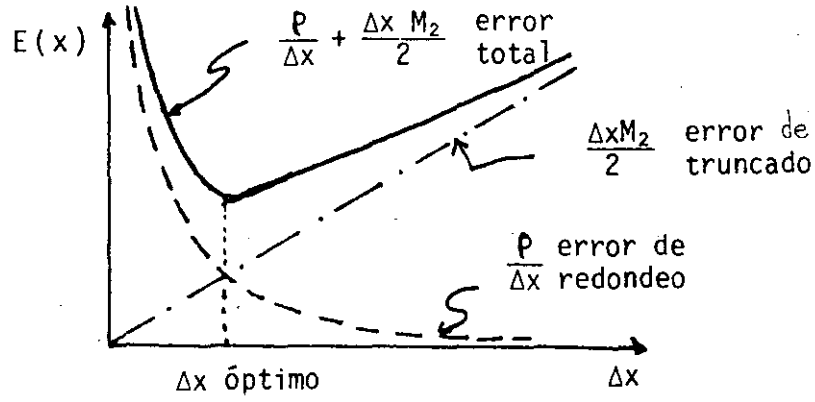


Fig. 4.8

En la elección del tamaño del intervalo de integración Δx aparecen involucrados estos errores, ya que por un lado, al asignar un valor "grande" a Δx se comete un error de truncado grande (a continuación se discute porque) y si por otra parte se escoge muy pequeño puede haber errores por despreciar cifras decimales (error de redondeo), (fig. 4.8).

La elección de valor óptimo de Δx no es sencillo, lo más usado en la práctica es escoger un Δx relativamente pequeño y aplicar el procedimiento numérico, luego se toma otro Δx menor y se utiliza el procedimiento otra vez, si los resultados no difieren mucho, se acepta uno de los cálculos como bueno; en caso contrario, se escogen otros dos valores de Δx, si no se llega a un resultado adecuado suspende el cálculo y quiza convenga probar otro método numérico diferente.

En atención al error de truncado, la ec. 4.8 dice:

$$y_{n+1} = y_n + \Delta x y'(x_n) + \frac{\Delta x^2}{2!} y''(x_n) + \frac{\Delta x^3}{3!} y'''(x_n) + \ldots$$

Al compararla con la ec. del método de Euler, se observa que esta no toma en cuenta los términos de segundo orden en adelante, es decir la ec. 4.8 se ha truncado, y ello implica un error de este tipo.

Según la serie de Taylor, el error de truncado es tal que:

$$D(x) \leq \left| \frac{d^2 y(\xi)}{dx^2} \right|_{máx} \frac{\Delta x^2}{2} \qquad (4.80)$$

$$x_n \leq \xi \leq x_{n+1}$$

Suponiendo que:

$$\left| \frac{d^2 y(\xi)}{dx^2} \right|_{máx} = M$$

Así, en el primer intervalo de integración el error de truncado es:

$$d_1 = M^{(1)} \frac{\Delta x^2}{2}$$

En el segundo intervalo, vuelve aparecer un error de truncado sea

$$d_2 = M^{(2)} \frac{\Delta x}{2}$$

suponiendo que $M^{(1)} = M^{(2)} = M$ para cualquier iteración se tiene que:

$$d_1 = M \frac{\Delta x^2}{2}$$

$$d_1 + d_2 = M \frac{\Delta x^2}{2} + M \frac{\Delta x^2}{2} = M \Delta x^2$$

también

$$d_1 + d_2 + d_3 = \frac{3}{2} M \Delta x^2$$

y así sucesivamente, hasta que en la iteración N (N > 3) el error acumulado:

$$d_1 + d_2 + d_3 + \ldots + d_N = N \frac{M}{2} \Delta x^2 \qquad (4.81)$$

Por otra parte:

$$x_1 = x_0 + \Delta x$$

$$x_2 = x_1 + \Delta x = x_0 + 2\Delta x$$

$$x_N = x_{N-1} + \Delta x = x_0 + N\Delta x$$

de esta última ecuación

$$N = \frac{x_N - x_0}{\Delta x} \qquad (4.82)$$

Sustituyendo 4.82 en 4.81

$$d_1 + d_2 + d_3 + \ldots + d_N = (x_N - x_0)\ \tfrac{1}{2}\ M\Delta x$$

si

$$d_T = d_1 + d_2 + \ldots + d_N$$

y

$$M_2 = (x_N - x_0)\ M$$

entonces el error acumulado de redondeo es

$$d_T = \frac{\Delta x}{2}\ M_2 \qquad (4.83)$$

Y se afirma que en el método de Euler el error acumulado de truncado es proporcional al tamaño del intervalo de integración Δx.

En general, se ha notado que este error tratandose de ecuaciones diferenciales ordinarias es más importante.

Se puede demostrar o encontrar en libros sobre el tema los errores acumulados de truncado para cada uno de los métodos aquí descritos. Según la potencia a la que aparece Δx se dice el "orden del método", que tendrá un menor error de truncado mientras mayor sea el orden (Δx^2, si Δx es menor que 1 es mejor que Δx). En la tabla 4.1 se reporta el orden de algunos de los métodos.

TABLA 4.1

Método	Orden
Euler	1
Euler Modificado	2
Heun	2
Nystrom	2
Serie de Taylor	2,3,4,... según el número de términos

Runge-Kutta	2,3,4,5	
Adams	2,3,4,5	según se especifique
Predictor-Corrector	2,3,4,5	

4.11 Ejemplos

Ejemplos 4.11.1

Calcular el tránsito de una avenida a través del almacenamiento mostrado en la fig. 4.9. Se sabe que la avenida es constante e igual $I = 10 m^3/s$ y que el gasto que sale del almacenamiento esta dado por la ecuación:

$$Q = C_0 a \sqrt{2gh} = 5 \quad h \quad (m^3/s) \qquad (4.84)$$

El área de la base del almacenamiento es 100 m^2. El nivel en el almacenamiento al tiempo t=0 s tiene una carga h = 16 m.

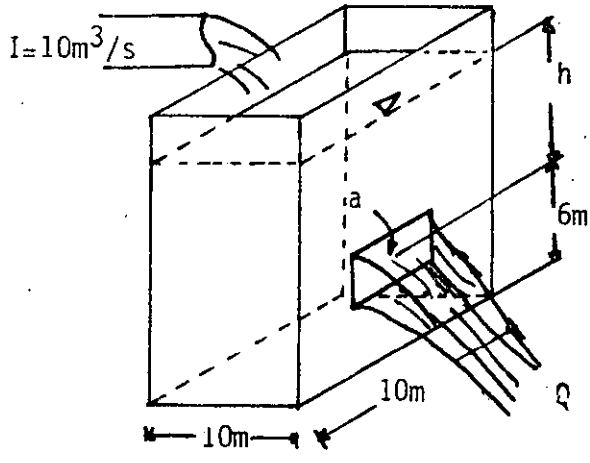


Fig. 4.9

Solución: Se trata de resolver la ecuación de continuidad:

$$\frac{dV}{dt} = I - Q$$

como:

$$V = Ah+6$$

así

$$dV = A\ dh$$

$$\frac{dh}{dt} = \frac{I-Q}{A}$$

sustituyendo valores y la ecuación del gasto de descarga.

k) Comparación de los resultados del método de Euler Mejorado y la solución exacta. Valores de h_n.

tiempo	Δt = 1	Δt = 2	Δt = 3	Δt = 4	EXACTA
1	15.90031				15.90031
2	15.80125	15.80125			15.80125
3	15.70281		15.70283		15.70280
4	15.60498	15.60499		15.60503	15.60498
5	15.50778				15.50777
6	15.41118	15.41120	15.41122		15.41118
7	15.31521				15.31520
8	15.21984	15.21986		15.21994	15.21983
9	15.12508		15.12514		15.12507
10	15.03095	15.03095			15.03092
11	14.93738				14.93738
12	14.84444	14.84447	14.84452	14.84459	14.84443
13	14.75210				14.75209
14	14.66036	14.66039			14.66035
15	14.56921		14.56931		14.56920
16	14.47866	14.47870		14.47885	14.47865
17	14.38871				14.38869
18	14.29934	14.29938	14.29945		14.29933
19	14.21056				14.21055
20	14.12237	14.12242		14.12260	14.12236
21			14.03489		

j) Comparación de los resultados del Método de Euler con distintos intervalos de tiempo y la solución exacta.

	h_n				
tiempo	$\Delta t = 1$	$\Delta t = 2$	$\Delta t = 3$	$\Delta t = 4$	EXACTA
1	15.9				15.90031
2	15.80063	15.8			15.80125
3	15.70188		15.7		15.70280
4	15.60375	15.60251		15.6	15.60498
5	15.50624				15.50777
6	15.40935	15.40751			15.41118
7	15.31308				15.31520
8	15.21742	15.21498		15.21006	15.21983
9	15.12237		15.11690		15.12507
10	15.02793	15.02492			15.03092
11	14.93410				14.93738
12	14.84088	14.83730	14.83369	14.83006	14.84443
13	14.74826				14.75209
14	14.65624	14.65211			14.66035
15	14.56483		14.55598		14.56920
16	14.47401	14.46933		14.45987	14.47865
17	14.38378				14.38869
18	14.29415	14.28894	14.28369		14.29933
19	14.20511				14.21055
20	14.11667	14.11093		14.09934	14.12236
21			14.01679		

i) obtención de la solución exacta de la ecuación diferencial:

$$\frac{dh}{dt} = 0.1 - 0.05\ h$$

Por separación de variables:

$$\frac{dh}{0.1-0.05\sqrt{h}} = dt$$

integrando

$$\int \frac{dh}{0.1-0.05\sqrt{h}} = t + c$$

si

$$x = \sqrt{h}$$

$$x^2 = h$$

$$2xdx = dh$$

así:

$$\int \frac{dh}{0.1-.0.5\sqrt{h}} = \int \frac{2xdx}{0.1-0.05x} = -2 \int \frac{xdx}{0.05x-0.1}$$

como

$$\int \frac{x\ dx}{ax+b} = \frac{x}{a} - \frac{b}{a^2} \ln (ax+b)$$

entonces

$$-2 \int \frac{xdx}{0.05x-0.1} = -2 \left[\frac{x}{0.05} - \frac{0.1}{0.0025} \ln(0.05x-0.1)\right]$$

$$= -40x + 80 \ln(0.05x-0.1)$$

por lo tanto

$$\int \frac{dh}{0.1-0.05\sqrt{h}} = -40\sqrt{h} - 80 \ln (0.05\sqrt{h} - 0.1) = t + c$$

como: $h = 16$ para $t = 0$

$$c = -0 - 40\sqrt{16} - 80 \ln (0.05\sqrt{16} - 0.1) = 24.206807$$

la solución es:

$$\boxed{t = -24.206807 - 40\sqrt{h} - 80 \ln (0.05\sqrt{h} - 0.1)}$$

Ejemplo 4.11.2

Encontrar una solución aproximada de la ecuación $y'' + y = x$ la cual pasa por $(0,0)$ y $(1,0)$ sabiendo que tiene asociado el funcional $J = \int_0^1 (y'^2 - y^2 + 2xy)dx$.

Solución: (Método de Ritz, subcapítulo 4.8.1)

Sea la solución del tipo $y = u_0(x) + a_1 u_1(x)$, donde $u_0(x) = 0$ y $u_1(x) = x - x^2$, ya que $u_0(x)$ satisface las condiciones $y(0) = 0$ y $y(1) = 0$, y la función $u_1(x)$ se anula en $(0,0)$ y $(1,0)$, es decir $u_1(0) = 0 - 0^2 = 0$ y $u_1(1) = 1 - 1^2 = 0$.

$$y = a_1(x - x^2) \qquad \ldots(\alpha)$$

Y el problema consiste en encontrar el valor de a_1.

con base en la ec. α

$$\left.\begin{aligned} y' &= a_1(1-2x) \\ y'' &= a_1^2(1-2x)^2 = a_1^2(1-4x+4x^2) \\ y^2 &= a_1^2(x^2 - 2x^3 + x^4) \\ 2xy &= 2a_1(x^2 - x^3) \end{aligned}\right\} \ldots(\beta)$$

sustituyendo las ecs.β en el funcional

$$J = \int_0^1 [a_1^2(1-4x+4x^2) - a_1^2(x^2-2x^3+x^4) + 2a_1(x^2-x^3)]dx$$

Integrando

$$J = a_1^2 \left[x-2x^2+\frac{4}{3}x^3\right]_0^1 - a_1^2 \left[\frac{1}{3}x^3-\frac{1}{2}x^4+\frac{1}{5}x^5\right]_0^1 + 2a_1\left[\frac{1}{3}x^3-\frac{1}{4}x^4\right]_0^1$$

$$J = a_1^2 \left[1-2+\frac{4}{3}\right] - a_1^2 \left[\frac{1}{3} - \frac{1}{2}+\frac{1}{5}\right] + 2a_1\left[\frac{1}{3} - \frac{1}{4}\right]$$

$$J = a_1^2 \left(\frac{1}{3}\right) - a_1^2 \left(\frac{1}{3}-\frac{3}{10}\right) + a_1\left(\frac{1}{6}\right) = \frac{3}{10}\, a_1^2 + \frac{1}{6}a_1 \qquad \ldots(\delta)$$

Para hacer mínimo el funcional

$$\frac{\partial J}{\partial a_1} = 0$$

así con base en la ec. δ

$$\frac{\partial J}{\partial a} = 2a_1\frac{3}{10} + \frac{1}{6} = 0$$

$$\frac{3}{5}\, a_1 = -\frac{1}{6}$$

$$a_1 = -\frac{5}{18}$$

entonces, una solución aproximada de la ec. diferencial es:

$$\boxed{y = -\frac{5}{18}\,(x-x^2)}$$

Ejemplo 4.11.3

Encontrar una solución aproximada de la ecuación $y'' + x = 0$ de manera que $y(1) = 0$ y $y(2) = -1$.

Solución (Método de Galerkin, subcapítulo 4.8.2)

Como no se dispone del funcional se procederá mediante el método de Galerkin.

Se supone una solución del tipo $y = u_0(x) + a_1u_1(x)$

siendo $u_0(x) = 1-x$ y $u_1(x) = x^2-3x+2$ pues $u_0(1) = 0$ y $u_0(2) = -1$; y tambien $u_1(0) = 0$ y $u_1(2) = 0$, de modo que

$$y = (1-x) + a_1(x^2-3x+2) \qquad \ldots(\alpha)$$

se trata de encontrar el valor de a_1

con base en la ec. α

$$y' = -1 + a_1(2x-3)$$

$$y'' = 2a_1(2x-3) \qquad \ldots(\beta)$$

Al sustituir la ec. β en la ec. diferencial

$$2a_1 + x \neq 0$$

$$2a_1 + x = R(x)$$

Por lo que

$$\int_1^2 R(x)u_1(x)dx = \int_1^2 (2a_1+x)(x^2-3x+2)dx$$

$$= \int_1^2 2a_1(x^2-3x+2)+(x^3-3x^2+2)dx$$

$$2a_1\left[\frac{x^3}{3} - \frac{3x^2}{2} + 2x\right]_1^2 + \left[\frac{x^4}{4} - x^3 + x^2\right]_1^2$$

$$= \left[2a_1(-\frac{1}{6})+(-\frac{37}{12}) = 0\right.$$

$$= -\frac{a_1}{3} - \frac{37}{12} = 0$$

$$a_1 = -\frac{37}{4}$$

por lo que una solución aproximada es

$$\boxed{y = 1-x-\frac{37}{4}(x^2-3x+2)}$$

a) Solución mediante el método de Euler

$h_{n+1} = h_n + \Delta t\, f(t_n,h_n)$; $f(t_n,h_n) = 0.1 - 0.05\sqrt{h_n}$

METODO DE EULER (DT= 1)

N	TN	H(N)	F(TN,HN)	H(N+1)
0	1	16	-.1	15.9
1	2	15.9	-.0993740204	15.800626
2	3	15.800626	-.0987500062	15.701876
3	4	15.701876	-.0981279636	15.603748
4	5	15.603748	-.0975078987	15.5062401
5	6	15.5062401	-.0968898177	15.4093503
6	7	15.4093503	-.0962737266	15.3130766
7	8	15.3130766	-.0956596316	15.2174169
8	9	15.2174169	-.0950475387	15.1223694
9	10	15.1223694	-.094437454	15.027932
10	11	15.027932	-.0938293834	14.9341026
11	12	14.9341026	-.093223333	14.8408792
12	13	14.8408792	-.0926193087	14.7482599
13	14	14.7482599	-.0920173165	14.6562426
14	15	14.6562426	-.0914173622	14.5648252
15	16	14.5648252	-.0908194516	14.4740058
16	17	14.4740058	-.0902235908	14.3837822
17	18	14.3837822	-.0896297854	14.2941524
18	19	14.2941524	-.0890380413	14.2051144
19	20	14.2051144	-.0884483642	14.116666
20	21	14.116666	-.0878607597	14.0288053

METODO DE EULER (DT= 2)

N	TN	H(N)	F(TN,HN)	H(N+1)
0	2	16	-.1	15.8
1	4	15.8	-.0987460692	15.6025079
2	6	15.6025079	-.0975000498	15.4075078
3	8	15.4075078	-.0962619918	15.2149838
4	10	15.2149838	-.0950319448	15.0249199
5	12	15.0249199	-.0938099578	14.8373
6	14	14.8373	-.0925960798	14.6521078
7	16	14.6521078	-.091390359	14.4693271
8	18	14.4693271	-.0901928436	14.2889414
9	20	14.2889414	-.0890035808	14.1109343
10	22	14.1109343	-.0878226175	13.935289
11	24	13.935289	-.0866500001	13.761989
12	26	13.761989	-.0854857745	13.5910175
13	28	13.5910175	-.0843299858	13.4223575
14	30	13.4223575	-.0831826786	13.2559921
15	32	13.2559921	-.0820438967	13.0919043
16	34	13.0919043	-.0809136834	12.930077
17	36	12.930077	-.0797920811	12.7704928
18	38	12.7704928	-.0786791315	12.6131346
19	40	12.6131346	-.0775748754	12.4579848
20	42	12.4579848	-.0764793529	12.3050261

Listado del programa utilizado

```
5  PR# 1: PRINT "";
10  REM  METODO DE EULER (EC 5.4)
15  DIM H(21)
20  READ DT
30  DATA  2
40  REM  SE DEFINE LA FUNCION F(T,H)
50  DEF  FN F(H) = 0.1 - 0.05 *  SQR (H)
60  REM  DATOS INICIALES
65 TN = 0:H(0) = 16
70  PRINT "METODO DE EULER  (DT= "DT" )
71  PRINT " "
72  PRINT "N                TN                    H(N)          F(TN,HN)
        H(N+1)"
74  PRINT " "
80  REM  EMPEIEZA EL METODO
90  FOR N = 0 TO 20
100 F =  FN F(H(N))
110 H(N + 1) = H(N) + DT * F
120 TN = TN + DT
130  PRINT N,TN,H(N),F,H(N + 1)
140  NEXT N
150  END
```

$$h_{n+1}=h_n+[f(t_n,h_n)+f(t_{n+1},\tilde{h}_{n+1})]\frac{\Delta t}{2}$$

METODO DE EULER MODIFICADO (DT= 1)

N	TN	H(N)	F(TN,HN)	F(TN+1,HT)
0	1	16	-.1	-.0993740204
1	2	15.900313	-.0993759828	-.0987519624
2	3	15.801249	-.0987539847	-.0981319896
3	4	15.7028061	-.0981338219	-.0975137482
4	5	15.6049823	-.0975157104	-.0968976043
5	6	15.5077757	-.0968995663	-.0962834436
6	7	15.4111842	-.0962854056	-.0956712726
7	8	15.3152058	-.0956732342	-.0950610968
8	9	15.2198387	-.0950630582	-.0944529224
9	10	15.1250807	-.0944548834	-.0938467552
10	11	15.0309299	-.0938487159	-.093242601
11	12	14.9373842	-.0932445614	-.0926404656
12	13	14.8444417	-.0926424259	-.0920403555
13	14	14.7521003	-.092042315	-.0914422759
14	15	14.660358	-.0914442347	-.0908462324
15	16	14.5692128	-.0908481988	-.0902522311
16	17	14.4786626	-.0902541089	-.0896602777
17	18	14.3887053	-.0896622348	-.0890703777
18	19	14.299339	-.0890723342	-.0884825369
19	20	14.2105616	-.0884844927	-.0878967609
20	21	14.122371	-.0878987159	-.0873130551

METODO DE EULER MODIFICADO (DT= 2)

N	TN	H(N)	F(TN,HN)	F(TN+1,HT)
0	2	16	-.1	-.0987480692
1	4	15.8012539	-.0987539835	-.0975078991
2	6	15.6049921	-.0975157722	-.096277612
3	8	15.4111987	-.0962854982	-.095055299
4	10	15.2198579	-.0950631815	-.0938409837
5	12	15.0309537	-.0938488699	-.0926347321
6	14	14.8444701	-.0926426105	-.0914365751
7	16	14.6603909	-.0914444499	-.090246564
8	18	14.4786999	-.0902544345	-.0890647447
9	20	14.2993808	-.0890726101	-.0878911622
10	22	14.122417	-.087899922	-.0867258615
11	24	13.9477921	-.0867337148	-.0855688869
12	26	13.7754895	-.0855767328	-.0844202812
13	28	13.6054925	-.0844281194	-.0832800879
14	30	13.4377843	-.0832879175	-.0821483492
15	32	13.272348	-.0821561694	-.0810251065
16	34	13.1091667	-.0810329165	-.0799104007
17	36	12.9482234	-.0799181996	-.0788042716
18	38	12.7895009	-.0788120588	-.0777067587
19	40	12.6329821	-.0777145333	-.0766179002
20	42	12.4786497	-.0766256614	-.0755377337

```
5  PR# 1: PRINT "";
10  REM   METODO DE EULER MODIFICADO (EC 5.5)
15  DIM H(21)
20  READ DT
30  DATA  2
35 D1 = DT * 0.5
40  REM  SE DEFINE LA FUNCION F(T,H)
50  DEF  FN F(H) = 0.1 - 0.05 *  SQR (H)
60  REM  DATOS INICIALES
65 TN = 0:H(0) = 16
70  PRINT "METODO DE EULER MODIFICADO (DT= "DT" )"
71  PRINT " "
72  PRINT "N                TN                  H(N)          F(TN,HN)
        F(TN+1,HT)"
74  PRINT " "
80  REM  EMPIEZA EL METODO
90  FOR N = 0 TO 20
100 F1 =  FN F(H(N))
105 HT = H(N) + DT * F1
108 F2 =  FN F(HT)
110 H(N + 1) = H(N) + D1 * (F1 + F2)
120 TN = TN + DT
130  PRINT N,TN,H(N),F1,F2
140  NEXT N
150  END
```

$$h_{n+1} = h_{n-1} + 2\Delta t\, f(t_n, h_n)$$

METODO DE NYSTROM (DT= 1)

N	TN	H(N-1)	F(TN,HN)	H(N)
1	1	16	-.0987539554	15.8012539
2	2	15.8012539	-.0987617414	15.8024921
3	3	15.8024921	-.0975077872	15.6037304
4	4	15.6037304	-.0975314945	15.6074765
5	5	15.6074765	-.0962693776	15.4086674
6	6	15.4086674	-.096309308	15.4149378
7	7	15.4149378	-.0950387706	15.2160488
8	8	15.2160488	-.0950952346	15.2248602
9	9	15.2248602	-.0938160103	15.0258583
10	10	15.0258583	-.0938893254	15.0372282
11	11	15.0372282	-.0926011483	14.8380797
12	12	14.8380797	-.0926916314	14.8520259
13	13	14.8520259	-.0913942033	14.6526964
14	14	14.6526964	-.0915022031	14.6698375
15	15	14.6698375	-.080195242	14.469692
16	16	14.469692	-.0903210909	14.488847
17	17	14.488847	-.0890042979	14.2898498
18	18	14.2898498	-.0891483441	14.3108284
19	19	14.3108284	-.0879214123	14.1107531
20	20	14.1107531	-.0879840127	14.1351956

METODO DE NYSTROM (DT= 2)

N	TN	H(N-1)	F(TN,HN)	H(N)
1	2	16	-.0987539554	15.8012539
2	4	15.8012539	-.0975157211	15.6049842
3	6	15.6049842	-.0962854491	15.411191
4	8	15.411191	-.0950633612	15.2198424
5	10	15.2198424	-.0938487739	15.0309387
6	12	15.0309387	-.0926424914	14.8444473
7	14	14.8444473	-.0914443055	14.6603688
8	16	14.6603688	-.0902542383	14.4786701
9	18	14.4786701	-.0890724193	14.2993519
10	20	14.2993519	-.0878987795	14.1223804
11	22	14.1223804	-.0867334783	13.9477568
12	24	13.9477568	-.0855764431	13.7754465
13	26	13.7754465	-.0844278382	13.605451
14	28	13.605451	-.0832875823	13.4377351
15	30	13.4377351	-.0821558445	13.2723007
16	32	13.2723007	-.0810325368	13.1091118
17	34	13.1091118	-.0799178321	12.9481705
18	36	12.9481705	-.0788116357	12.7894404
19	38	12.7894404	-.0777141242	12.632924
20	40	12.632924	-.0766251959	12.4785839

```
5  PR# 1: PRINT "";
10  REM   METODO DE NYSTROM  (EC 5.6)
15  DIM H(21)
20  READ DT
30  DATA 2
35 D1 = 2 * DT
40  REM  SE DEFINE LA FUNCION F(T,H)
50  DEF  FN F(H) = 0.1 - 0.05 *  SQR (H)
60  REM  DATOS INICIALES
65 TN = 0:H(0) = 16
67  REM  DATO TOMADO DEL METODO DE EULER MODIFICADO
68 H(1) = 15.8012539
70  PRINT "METODO DE NYSTROM (DT= "DT" )"
71  PRINT " "
72  PRINT "N                TN                H(N-1)           F(TN,HN)
       H(N)"
74  PRINT " "
80  REM  EMPIEZA EL METODO
90  FOR N = 1 TO 20
100 F =  FN F(H(N))
110 H(N + 1) = H(N - 1) + D1 * F
120 TN = TN + DT
130  PRINT N,TN,H(N - 1),F,H(N)
140  NEXT N
150  END
```

de h) $h_{n+1} = h_n + \Delta t\ T^{(1)}$ $TP = T^{(1)} = f_n + f'_n \frac{\Delta t}{2}$; $f_n = 0.1 = 0.05\sqrt{h_n}$; $f'_n = -0.025\ \frac{1}{\sqrt{h_n}}$

METODO DE TAYLOR (DT= 1)

N	TN	H(N)	TP	H(N+1)
0	1	16	-.103125	15.896875
1	2	15.896875	-.102489547	15.7943855
2	3	15.7943855	-.101856029	15.6925294
3	4	15.6925294	-.101224453	15.591305
4	5	15.591305	-.100594828	15.4907101
5	6	15.4907101	-.0999671525	15.390743
6	7	15.390743	-.09934144	15.2914016
7	8	15.2914016	-.0987176944	15.1926839
8	9	15.1926839	-.0980959218	15.0945879
9	10	15.0945879	-.0974761294	14.9971118
10	11	14.9971118	-.0968583202	14.9002535
11	12	14.9002535	-.0962425034	14.804011
12	13	14.804011	-.0956286842	14.7083823
13	14	14.7083823	-.0950168684	14.6133654
14	15	14.6133654	-.0944070623	14.5189584
15	16	14.5189584	-.0937992718	14.4251591
16	17	14.4251591	-.0931935029	14.3319656
17	18	14.3319656	-.0925897615	14.2393759
18	19	14.2393759	-.0919880536	14.1473878
19	20	14.1473878	-.0913883851	14.0559994
20	21	14.0559994	-.0907907618	13.9652086

METODO DE TAYLOR (DT= 2)

N	TN	H(N)	TP	H(N+1)
0	2	16	-.1025	15.7975
1	4	15.7975	-.1049593[illegible]	15.5775813
2	6	15.5775813	-.1036763[illegible]	15.3702285
3	8	15.3702285	-.102401172	15.1654262
4	10	15.1654262	-.10113373	14.9631587
5	12	14.9631587	-.0998741252	14.7634104
6	14	14.7634104	-.0986224099	14.5661656
7	16	14.5661656	-.0973786253	14.3714084
8	18	14.3714084	-.0961428089	14.1791227
9	20	14.1791227	-.094915073	13.9892926
10	22	13.9892926	-.0936954013	13.8019018
11	24	13.8019018	-.0924838644	13.616934
12	26	13.616934	-.0912805109	13.434373
13	28	13.434373	-.0900853893	13.2542022
14	30	13.2542022	-.0888985454	13.0764051
15	32	13.0764051	-.0877200279	12.9009651
16	34	12.9009651	-.0865498827	12.7278653
17	36	12.7278653	-.0853881555	12.557089
18	38	12.557089	-.0842348914	12.3886192
19	40	12.3886192	-.0830901353	12.222439
20	42	12.222439	-.0819539306	12.0585311

```
5  PR# 1: PRINT "";
10  REM   METODO DE TAYLOR (EC 5 10)
15  DIM H(21)
20  READ DT
30  DATA 2
35 D1 = 0.5 * DT
40  REM  SE DEFINE LA FUNCION F(T,H)
50  DEF  FN F(H) = 0.1 - 0.05 *  SQR (H)
55  REM  SE DEFINE LA DERIVADA DE LA FUNCION F(T,H)
58  DEF  FN FP(H) =  - 0.025 /  SQR (H)
60  REM  DATOS INICIALES
65 TN = 0:H(0) = 16
70  PRINT "METODO DE TAYLOR (DT= "DT" )"
71  PRINT " "
72  PRINT "N                TN                H(N)             TP
        H(N+1)"
74  PRINT " "
80  REM  EMPEIEZA EL METODO
90  FOR N = 0 TO 20
100 TP =  FN F(H(N)) + D1 *  FN FP(H(N))
110 H(N + 1) = H(N) + DT * TP
120 TN = TN + DT
130  PRINT N;TN;H(N);TP;H(N + 1)
140  NEXT N
150  END
```

e) Solución mediante el método de Runge-Kutta de tercer orden

METODO DE RUNGE-KUTTA DE TERCER ORDEN (DT= 2)

N	H(N)	K1	K2	K3
0	16	-.1	-.0993740204	-.0987618168
1	15.8012474	-.0987539143	-.0981318593	-.0975235921
2	15.6049791	-.0975156897	-.0968975836	-.0962932762
3	15.4111793	-.0962853745	-.0956712417	-.0950709171
4	15.2198322	-.0950630168	-.0944528811	-.0938565624
5	15.0309218	-.0938486641	-.0932425495	-.0926502593
6	14.8444321	-.0926423638	-.0920402937	-.0914520546
7	14.6603469	-.0914441624	-.0908461604	-.0902619945
8	14.47865	-.0902541064	-.0896601955	-.0890801248
9	14.299325	-.0890722414	-.0884824445	-.0879064907
10	14.1223555	-.087898613	-.0873129526	-.0867411371
11	13.947725	-.0867330655	-.0861517641	-.085584108
12	13.7754168	-.0855762433	-.0849989229	-.0844354469
13	13.6054144	-.08442759	-.0838544722	-.0832951967
14	13.4377008	-.0832873483	-.0827184546	-.0821633998
15	13.2722593	-.0821555605	-.081590912	-.0810400975
16	13.1090729	-.0810322682	-.0804719855	-.0799253307
17	12.9481245	-.0799175122	-.0793614156	-.0788191394
18	12.789397	-.0788113324	-.0782595418	-.0777215627
19	12.6328733	-.0777137682	-.0771663031	-.0766326391
20	12.4785361	-.0766248577	-.0760817373	-.0755524061

METODO DE RUNGE-KUTTA DE TERCER ORDEN (DT= 4)

N	H(N)	K1	K2	K3
0	16	-.1	-.0987460692	-.0975476617
1	15.6049787	-.0975156875	-.0962775289	-.0950949848
2	15.2198315	-.0950630124	-.0938408206	-.0926743172
3	14.8444311	-.0926423572	-.0914363236	-.0902860344
4	14.4786487	-.0902540975	-.08906441	-.0879305042
5	14.1223538	-.0879006018	-.0867254443	-.0856080863
6	13.7754148	-.08557623	-.0844197819	-.083319131
7	13.4376985	-.0832873327	-.0821477687	-.0810639784
8	13.1090703	-.0810322504	-.0799097397	-.0788429573
9	12.7893942	-.0788113122	-.077706018	-.0766563844
10	12.478533	-.0766248353	-.0755369143	-.074504563
11	12.1763483	-.0744731231	-.0734027251	-.072387782
12	11.8827004	-.0723564649	-.0713037322	-.0703063152
13	11.5974486	-.0702751346	-.0692402015	-.0682604196
14	11.320451	-.0682293897	-.0672123818	-.0662503348
15	11.0515648	-.0662194698	-.0652205035	-.0642762818
16	10.7906463	-.0642455962	-.0632647784	-.062338462
17	10.5375509	-.0623079703	-.0613453978	-.0604370562
18	10.2921331	-.0604067732	-.0594625317	-.0585722232
19	10.0542471	-.0585421637	-.057616328	-.0567440995
20	9.82374602	-.0567142784	-.0558069115	-.054952798

```
5  PR# 1: PRINT "";
10  REM   METODO DE RUNGE-KUTTA (ECS  5 24)
15  DIM H(21)
20  READ DT
30  DATA   4
40  REM  SE DEFINE LA FUNCION F(T,H)
50  DEF  FN F(H) = 0.1 - 0 05 *  SQR (H)
60  REM  DATOS INICIALES
65 TN = 0:H(0) = 16
70  PRINT "METODO DE RUNGE-KUTTA DE TERCER ORDEN (DT= "DT" )"
71  PRINT " "
72  PRINT "N                  H(N)          K1                  K2
          K3"
74  PRINT " "
80  REM  EMPIEZA EL METODO
90  FOR N = 0 TO 20
100 K1 =  FN F(H,N))
110 K2 =  FN F(H(N) + DT * K1 + 0,5)
120 K3 =  FN F(H(N) + DT * (- K1 + 2 * K2))
130 H(N + 1) = H(N) + DT * (K1 + 4 + K2 + K3) / 6
140 TN = TN + DT
150  PRINT N,H(N),K1,K2,K3
160  NEXT N
170  END
```

f) Solución mediante el método de Runge-Kutta de orden cuatro

METODO DE RUNGE-KUTTA DE CUARTO ORDEN (DT= 2)

N	H(N)	K1	K2	K3	K4
0	16	- 1	-. 0993740204	-. 099377945	-. 0987538939
1	15. 8012474	-. 0987539144	-. 0981318594	-. 0981357839	-. 0975156602
2	15. 6049791	-. 09751569	-. 0968975839	-. 0969015079	-. 0962853541
3	15. 4111794	-. 0962853749	-. 0956712421	-. 0956751653	-. 0950629965
4	15. 2198223	-. 0950630174	-. 0944528818	-. 0944568038	-. 0938486429
5	15. 030922	-. 093848665	-. 0932425503	-. 093246471	-. 0926423436
6	14. 8444323	-. 0926423647	-. 0920402947	-. 0920442136	-. 0914441424
7	14. 6603471	-. 0914441636	-. 0908461615	-. 0908500783	-. 0902540863
8	14. 4786502	-. 0902541077	-. 0896601967	-. 089664111	-. 0890722215
9	14. 2993252	-. 0890722429	-. 0884824459	-. 0884863573	-. 087898592
10	14. 1223557	-. 0878986145	-. 0873129541	-. 0873168624	-. 0867332456
11	13. 9477252	-. 0867332672	-. 0861517658	-. 0861556706	-. 0855762236
12	13. 7754171	-. 0855762453	-. 0849989248	-. 0850028256	-. 0844275707
13	13. 6054147	-. 0844427592	-. 0838544743	-. 0838583708	-. 0832872287
14	13. 4377011	-. 0832873505	-. 0827184569	-. 0827223486	-. 0821555489
15	13. 2722596	-. 0821555629	-. 0815908144	-. 0815948011	-. 0810328489
16	13. 1090732	-. 0810322707	-. 0804718881	-. 0804757693	-. 0799174969
17	12. 9481249	-. 0799175148	-. 0793614193	-. 0793652938	-. 0788113131
18	12. 7893974	-. 0788113252	-. 0782595447	-. 0782634139	-. 0777137489
19	12. 6328738	-. 0777137711	-. 0771663061	-. 0771701687	-. 0766248284
20	12. 4785366	-. 0766248608	-. 0760817405	-. 076085596	-. 0755446197

METODO DE RUNGE-KUTTA DE CUARTO ORDEN (DT= 4)

N	H(N)	K1	K2	K3	K4
0	16	- 1	-. 0987460698	-. 0987618415	-. 0975155614
1	15. 6049791	-. 09751569	-. 0962775314	-. 0962933013	-. 0950628481
2	15. 2198223	-. 0950630174	-. 0939408256	-. 0938565878	-. 0926421815
3	14. 8444323	-. 0926423648	-. 0914363311	-. 0914520993	-. 0902538781
4	14. 4786502	-. 0902541077	-. 0890644202	-. 0890801508	-. 0878984497
5	14. 1223557	-. 0878986146	-. 0867254563	-. 0867411833	-. 0855760441
6	13. 7754171	-. 0855762453	-. 0844197971	-. 0844354733	-. 0832871737
7	13. 4377011	-. 0832873505	-. 0821477865	-. 0821634265	-. 0810320528
8	13. 1090732	-. 0810322708	-. 07990976	-. 0799253575	-. 078811156
9	12. 7893975	-. 0788113353	-. 0777060409	-. 0777215897	-. 0766246287
10	12. 4785366	-. 076624861	-. 0755369398	-. 0755524333	-. 0744729694
11	12. 1763522	-. 0744731514	-. 0734027531	-. 0734181848	-. 0723563113
12	11. 8827847	-. 0723564959	-. 0713037629	-. 0713191258	-. 0702749847
13	11. 5974532	-. 0702751682	-. 0692402347	-. 069255522	-. 0682292415
14	11. 3204559	-. 0682294258	-. 0672124175	-. 0672276222	-. 0662193234
15	11. 05157	-. 0662195084	-. 0652205418	-. 0652356568	-. 0642454517
16	10. 7906518	-. 0642456374	-. 0632648193	-. 0632798374	-. 062307825
17	10. 5375566	-. 0623080141	-. 0613454411	-. 0613603551	-. 0604066329
18	10. 2921391	-. 0604068193	-. 0594625774	-. 0594773802	-. 0585420256
19	10. 0542532	-. 0585422123	-. 0576163761	-. 0576310604	-. 0567141425
20	9. 82375242	-. 0567143294	-. 0558069619	-. 0558215204	-. 0549230965

```
5  PR# 1: PRINT "";
10  REM    METODO DE RUNGE-KUTTA (EDS. 5.25)
15  DIM H(21)
20  READ DT
30  DATA 2
40  REM  SE DEFINE LA FUNCION F(T,H)
50  DEF  FN F(H) = 0.1 - 0.05 *  SQR (H)
60  REM  DATOS INICIALES
65 TN = 0:H(0) = 16
70  PRINT "METODO DE RUNGE-KUTTA DE CUARTO ORDEN (DT= "DT" )"
71  PRINT " "
72  PRINT "N                    H(N )           K1                    K2              K3
        K4"
74  PRINT " "
80  REM  EMPIEZA EL METODO
90  FOR N = 0 TO 20
100 K1 =  FN F(H(N))
110 K2 =  FN F(H(N) + DT * K1 * 0.5)
120 K3 =  FN F(H(N) + DT * K2 * 0.5)
125 K4 =  FN F(H(N) + DT * K3)
130 H(N + 1) = H(N) + DT * (K1 + 2 * K2 + 2 * K3 + K4) / 6
140 TN = TN + DT
150  PRINT N,H(N),K1,K2,K3,K4
160  NEXT N
170  END
```

g) Solución mediante la fórmula cerrada de cuarto orden de Adams

METODO DE ADAMS (DT= 2)

N	TN	H(N)	F(N)
3	6	15.4111794	-.0962853752
4	8	15.2198323	-.0950630177
5	10	15.030923	-.0938486652
6	12	14.8444323	-.092642365
7	14	14.6603471	-.0914441638
8	16	14.4786502	-.0902541079
9	18	14.2993252	-.0890722431
10	20	14.1223553	-.0878986147
11	22	13.9477253	-.0867332674
12	24	13.7754171	-.0855762454
13	26	13.6054147	-.0844275922
14	28	13.4377012	-.0832873507
15	30	13.2722597	-.0821555631
16	32	13.1090732	-.0810322709
17	34	12.9481249	-.079917515
18	36	12.7893975	-.0788113354
19	38	12.6328736	-.0777137713
20	40	12.4785369	-.076624861

METODO DE ADAMS (DT= 4)

N	TN	H(N)	F(N)
3	12	14.8444323	-.0926423649
4	16	14.4786502	-.0902541079
5	20	14.1223557	-.0878986146
6	24	13.7754171	-.0855762452
7	28	13.4377011	-.0832873503
8	32	13.1090732	-.0810322705
9	36	12.7893974	-.0788113349
10	40	12.4785365	-.0766248604
11	44	12.1763521	-.0744731507
12	48	11.8827045	-.072356495
13	52	11.597453	-.0702751672
14	56	11.3204557	-.0682294246
15	60	11.0515696	-.066219507
16	64	10.7906515	-.0642456358
17	68	10.5375563	-.0623080123
18	72	10.2921388	-.0604068174
19	76	10.054253	-.0585422101
20	80	9.82375211	-.056714327

```
5  PR# 1: PRINT "";
10  REM    METODO DE ADAMS (EC 5.37)
15  DIM H(21)
20  READ DT
30  DATA 4
32  REM  CONSTANTES
33 A0 = 55 / 24:A1 =  - 59 / 24
35 A2 = 37 / 24:A3 =  - 9 / 24
40  REM  SE DEFINE LA FUNCION F(T,H)
50  DEF  FN F(H) = 0.1 - 0.05 *  SQR (H)
60  REM  DATOS INICIALES
65 TN = 0:H(0) = 1
67  REM  DATOS TOMADOS DEL METODO DE RUNGE-KUTTA DE CUARTO ORDEN
68 H(1) = 15.6049791:H(2) = 15.2198323
69 H(3) = 14.8444323
70  REM  CALCULOS PREVIOS
71 F3 =  FN F(H(0)):F2 =  FN F(H(1))
72 F1 =  FN F(H(2))
77  PRINT "METODO DE ADAMS (DT= "DT" )"
78  PRINT " "
79  PRINT "N                TN                   H(N)             F(N)"
80  PRINT " "
88  REM EMPIEZA EL METODO
89 TN = 2 * DT
90  FOR N = 3 TO 20
95 F0 =  FN F(H(N))
100 H(N + 1) = H(N) + DT * (A0 * F0 + A1 * F1 + A2 * F2 + A3 * F3)
110 TN = TN + DT
120 PRINT N,TN,H(N),F0
130 F3 = F2:F2 = F1:F1 = F0
140  NEXT N
150  END
```

h) Solución mediante el método predictor-modificador-corrector de Hamming

METODO DE HAMMING (DT= 2)

N	TN	H(N)	F(N)
4	8	15.2198323	-.0950630174
5	10	15.030922	-.0938486649
6	12	14.8444323	-.0926423647
7	14	14.6603471	-.0914441635
8	16	14.4786582	-.0902541076
9	18	14.2993252	-.0890722428
10	20	14.1223557	-.0878986143
11	22	13.9477252	-.0867332671
12	24	13.7754171	-.085576245
13	26	13.6054146	-.0844275918
14	28	13.4377011	-.0832873503
15	30	13.2722596	-.0821555627
16	32	13.1090732	-.0810322706
17	34	12.9481248	-.0799175143
18	36	12.7893974	-.0788113351
19	38	12.6328738	-.0777137709
20	40	12.4785366	-.0766248609

METODO DE HAMMING (DT= 4)

N	TN	H(N)	F(N)
4	16	14.4786582	-.0902541078
5	20	14.1223557	-.0878986146
6	24	13.7754171	-.0855762453
7	28	13.4377011	-.0832873505
8	32	13.1090732	-.0810322702
9	36	12.7893974	-.0788113353
10	40	12.4785366	-.0766248609
11	44	12.1763522	-.0744731513
12	48	11.8827046	-.0723564958
13	52	11.5974531	-.0702751681
14	56	11.3204559	-.0682294256
15	60	11.05157	-.0662195083
16	64	10.7906517	-.0642456371
17	68	10.5375565	-.0623080139
18	72	10.2921391	-.0604068191
19	76	10.0542532	-.0585422122
20	80	9.82375239	-.0567143293

```
5  PR# 1: PRINT "";
10  REM METODO DE HAMMING (ECS 5.48 A 5.50)
15  DIM H(21)
20  READ DT
30  DATA 4
40  REM  SE DEFINE LA FUNCION F(T,H)
50  DEF  FN F(H) = 0.1 - 0.05 *  SQR (H)
60  REM  DATOS INICIALES
65 TN = 0:H(0) = 16
67  REM  DATOS TOMADOS DEL METODO DE RUNGE-KUTTA DE CUARTO ORDEN
68 H(1) = 15.6049791:H(2) = 15.2198323
69 H(3) = 14.8444323:H(4) = 14.4786502
70  REM  CALCULOS PREVIOS
71 F3 =  FN F(H(1)):F2 =  FN F(H(2))
72 F1 =  FN F(H(3))
77  PRINT "METODO DE HAMMING (DT= "DT" )"
78  PRINT " "
79  PRINT "N                TN                   H(N)              F(N)"
80  PRINT " "
88  REM EMPIEZA EL METODO
89 TN = 3 * DT
90  FOR N = 4 TO 20
92 F0 =  FN F(H(N))
93 H0 = H(N - 4) + 1.33333 * DT * (2 * F1 - F2 + 2 * F3)
95 H1 = H(N - 3) + 1.33333 * DT * (2 * F0 - F1 + 2 * F2)
98 H = H1 + (H(N) - H0) * (112 / 121)
100 PH =  FN F(H)
105 H(N + 1) = 0.125 * (9 * H(N) - H(N - 2)) + 0.375 * DT * (PH + 2 * F0 -
    F1)
110 TN = TN + DT
120  PRINT N,TN,H(N),F0
130 F3 = F2:F2 = F1:F1 = F0
140  NEXT N
150  END
```

DIVISION DE EDUCACION CONTINUA
FACULTAD DE INGENIERIA U.N.A.M.

INGENIERIA MARITIMA, MODULO:
METODOS NUMERICOS EN HIDRAULICA
DEL 1o. DE JULIO AL 6 DE SEPTIEMBRE DE 1985.
MEXICO, D.F.

ECUACIONES DIFERENCIALES PARCIALES

M. EN I. OSCAR FUENTES MARILES
M. EN I. POLIOPTRO MARTINEZ A.

5. ECUACIONES DIFERENCIALES PARCIALES

Una ecuación diferencial parcial es aquella en la cual aparecen derivadas parciales de una función desconocida con respecto a dos o mas variables independientes. Un ejemplo de una ecuación de este tipo es:

$$a^2 \frac{\partial^2 h}{\partial x^2} = \frac{\partial h}{\partial t} \qquad (5.1)$$

La solución de una ecuación diferencial parcial en una región R es la función definida en la región R, al igual que todas sus derivadas parciales incluidas en la ecuación diferencial parcial, y tal función reduce la ecuación a una identidad en cada punto de la región R. Se dice que esta función satisface la ecuación en R.

El orden de una ecuación diferencial parcial corresponde al de la derivada de mayor orden envuelta en ella.

Una importante propiedad de las ecuaciones diferenciales parciales es si posee o no linealidad. Por definición, una ecuación diferencial parcial es lineal para u(x,y) si tiene la forma:

$$\sum_{n=0}^{N} \sum_{m=0}^{M} a_{nm}(x,y) \frac{\partial^{n+m} u}{\partial x^n \partial y^m} = g(x,y) \qquad (5.2)$$

donde $a_{nm}(x,y)$ y $g(x,y)$ son funciones conocidas de x y y N,M son constantes enteras positivas. (Se acepta que $\frac{\partial^0 u}{\partial x^0 \partial y^0} = u$) cuando $g(x,y) = 0$ se dice que también la ec. 5.2 es homogenea. Como en el caso de ecuaciones diferenciales ordinarias, el principio de superposición de 5.2, sea el conjunto de funciones $u_1(x,y)$, $u_2(x,y)$,..., $u_p(x.y)$ soluciones también de 5.2 entonces se debe satisfacer 5.2 para cualquiera de ellas, así:

$$\sum_{n=0}^{N}\sum_{m=0}^{M} a_{nm}(x,y)\,\frac{\partial^{n+m}u_j}{\partial x^n \partial y^m} = 0 \qquad j = 1,2,..,p$$

entonces si $a_1, a_2,...,ap$ son constantes cualesquiera, por sustitución directa en 5.2 se demuestra que:

$$u = u + a_1u_1 + a_2u_2 + ... + a_pu_p$$

tambien es solución de 5.2.

Sean las siguientes ecuaciones diferenciales parciales

$$x^2u_{xx} + u_{xy} - \pi^2 u_{yy} + 3u_x - u = e^{x+y} \qquad (5.3)$$

$$\frac{\partial^2\phi}{\partial x^2} + \frac{\partial^2\phi}{\partial y^2} = 0 \qquad (5.4)$$

$$\phi_{tt} - k\phi = 0 \qquad (5.5)$$

$$u_{xyy} - u_x - (x+y)^2 u = e^x \qquad (5.6)$$

$$u_{xyy} - u_x - (x+y)u^2 = e^x \qquad (5.7)$$

$$\phi_{xx} - \frac{1}{c^2}\,\phi_{tt} = 0 \qquad (5.8)$$

$$\phi_{xx} - \left[\phi_x\right]^2 = \phi_y \qquad (5.9)$$

Son ecuaciones diferenciales parciales lineales, 5.3, 5.4, 5.5 5.6 y 5.8 de estas, son homogéneas 5.4, 5.5 y 5.8. Nótese que 5.7 es no lineal porque la función $u(x,y)$ aparece a una potencia diferente de la unidad y que la 5.9 es no lineal porque tiene una derivada al cuadrado.

Raras veces se puede hacer un progreso formal en las ecuaciones no lineales, afortunadamente, muchas de las ecuaciones de interés práctico son lineales o casi lineales.

Una clase de ecuaciones que es frecuente encontrar son del tipo:

$$A\frac{\partial^2\phi}{\partial x^2} + B\frac{\partial^2\phi}{\partial x\partial y} + C\frac{\partial^2\phi}{\partial y^2} + D\frac{\partial\phi}{\partial x} + E\frac{\partial\phi}{\partial y} + F\,\phi + G = 0 \qquad (5.10)$$

Cuando A,B,C,D,E,F, y G son funciones de x y y, corresponde al caso de una ecuación diferencial parcial lineal de segundo orden, como se dijo antes. Pero si estas funciones dependen de $x, y, \phi, \frac{\partial\phi}{\partial x}, \frac{\partial\phi}{\partial y}$ se dice quasi-lineal.

La ec. 5.10 se dice que puede ser de tres tipos según resulte ser B^2-4AC respecto a cero, esto es, cuando:

$B^2 - 4AC < 0$ es elíptica

$B^2 - 4AC = 0$ es parabólica

$B^2 - 4AC > 0$ es hiporbólica

en atención a esta clasificación la búsqueda de su solución puede ser diferente como se verá adelante.

De acuerdo con esto, la ec. 5.4 es elíptica (es la ec. conocida como de Laplace), la ec. 5.1 es parabólica (ec. de calor o difusión) y la 5.8 es hiporbólica (ec. de la onda).

En la práctica ingenieril muchas de las ecuaciones diferenciales parciales son difíciles de resolver o bien no estan resueltas por métodos analíticos. Algunas veces se ha encontrado su solución para condiciones iniciales o de orilla particulares; sin embargo, algunas de estas soluciones no son útiles en la práctica. Para resolver las ecuaciones diferenciales parciales se puede recurrir a varios métodos numéricos.

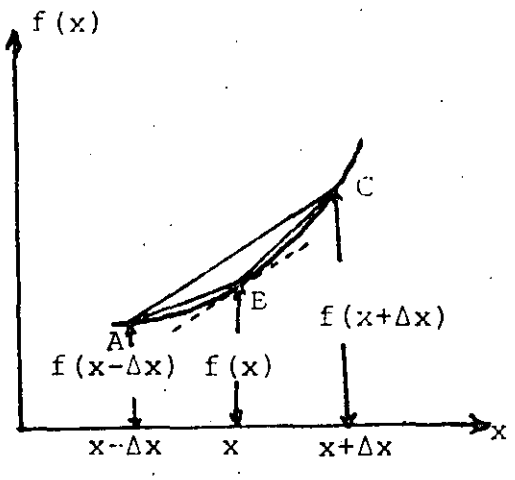


Fig. 5.1 Aproximaciones a

Cuando se trata con deri
ciones anteriores también t
que tener cuidado en variar
bles, esto significa:

Sea f(x,t), las difere a
x y t serían:

$$\left(\frac{\partial f}{\partial x}\right)_t = \frac{f(x+\Delta x,t)-f(x,} \qquad 20)$$

$$\left(\frac{\partial f}{\partial t}\right)_x = \frac{f(x,t+\Delta t)-f(x} \qquad .21)$$

de igual manera se pod n-
tral o hacia atrás.

DERIVADAS DE ORDEN DOS EN ADELANTE

En función de los operadores involucrados en las ecs. 5.17, 5.18 y 5.19 pueden plantearse las aproximaciones a las derivadas de orden dos en adelante, sin embargo, por el momento se va a preferir utilizar la serie de Taylor.

Para mostrar como se obtienen las aproximaciones de las derivadas se planteará encontrar $\frac{\partial^2 f}{\partial x^2}$ y $\frac{d^3 f}{dx^3}$.

Si se suman las ecs. 5.12 y 5.13 se tiene:

$$f(x+\Delta x)+f(x\Delta x)=2f(x)+f''(x)\Delta x^2 + O_3(\Delta x^4) \qquad (5.22)$$

donde

$$O_3(\Delta x^4) = O_1(\Delta x^4) + O_2(\Delta x^4) \qquad (5.23)$$

De la 5.12 se tiene:

$$f''(x) = \frac{f(x-\Delta x)-2f(x)+f(x+\Delta x)}{\Delta x^2} + O_3(\Delta x^4) \qquad (5.24)$$

si se desprecia el error $O_3(\Delta x^4)$ y se considera que f = f(x,t), con base en la ec. 5.24 se tiene:

$$\left(\frac{\partial^2 f}{\partial x^2}\right)_t = \frac{f(x-\Delta x,t)-2f(x,t)+f(x+\Delta x,t)}{\Delta x^2} \qquad (5.25)$$

Por lo que se refiere a la otra derivada de interés.

Al restar la ec. 5.13 a la 5.12 se tiene:

$$f(x+\Delta x)-f(x-\Delta x) = 2f'(x)\Delta x + f'''(x)\frac{\Delta x^3}{3} + O_4(\Delta x^4) \qquad (5.26)$$

con

$$O_4(\Delta x^4) = O_1(\Delta x^4) - O_2(\Delta x^4)$$

Si se sutituye la ec. 5.14 en la ec. 5.26

$$f(x+\Delta x) - f(x-\Delta x) = 2\ f(x+\Delta x)-2f(x)+f'''(x)\frac{\Delta x^3}{3} + O_4(\Delta x^4)-$$

$$-O_1(\Delta x^2) \qquad (5.27)$$

Al despejar f'''(x)

$$f'''(x) = \frac{-\frac{1}{3}f(x-\Delta x)+\frac{2}{3}f(x)-\frac{1}{3}f(x+\Delta x)}{\Delta x^3} + \frac{O_1(\Delta x^2)-O_4(\Delta x^4)}{3\Delta x^3}$$

$$(5.28)$$

si se desprecian los errores

$$f'''(x) = \frac{-\frac{1}{2}f(x\Delta x)+\frac{2}{3}f(x)-\frac{1}{2}f(x+\Delta x)}{\Delta x^3} \qquad (5.29)$$

Las ecs. 5.25 y 5.29 son las aproximaciones buscadas. Nótese que al igual con las derivadas de primer orden pue de plantearse varias aproximaciones a las derivadas, como hubiese sido si en lugar de sustituir la ec. 5.14 en la ec. 5.26 se sustituye la ec. 5.15 en la ec. 5.26. Desde luego, aquí tambien habrá mejores aproximaciones que otras, dado que mientras mas chico sea el error de truncado la diferencia entre la derivada y el cociente de diferencias será menor y se hará una adecuada aproximación (desde el punto de vista de discretización como se discutirá después).

5.1 Ecuaciones diferenciales parciales parabólicas

Sea el problema de flujo con potencial mostrado en la fig.6.2. Se desea calcular la línea de saturación Q'PQ en el transcurso del tiempo.

La ecuación diferencial que describe el problema es:

$$\boxed{\frac{\partial h}{\partial t} = a^2\frac{\partial^2 h}{\partial x^2}} \qquad (5.30)$$

siendo $a^2 = \frac{k\bar{h}}{S}$ (5.31)

donde

K coeficiente de permeabilidad

S rendimiento específico (cociente, volumen que se puede drenar entre el volumen total de la muestra).

$\bar{h}$ nivel promedio en el espacio y el tiempo de h.

Se puede ver que se trata de una ecuación diferencial parcial lineal de segundo orden parabólica, ya que al comparar la ec. 5.30 con la ec. 5.10 se tiene:

$$A = a^2$$
$$B = 0$$
$$C = 0$$
$$B^2 - 4AC = 0^2 - 4a^2(0) = 0$$

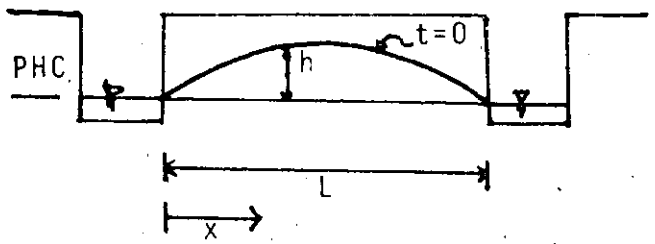


Fig. 5.2 Problema de flujo con potencial

Se ocurre sustituir las derivadas de la ec. 5.30 por diferencias finitas. Si se representa a $\frac{\partial h}{\partial t}$ en función de una derivada hacia adelante y a $\frac{\partial^2 h}{\partial x^2}$ por una ec. similar a la 5.24, se obtiene:

$$\frac{h(x,t+\Delta t)-h(x,t)}{\Delta t} = a^2 \frac{h(x-\Delta x,t)-2h(x,t)+h(x+\Delta x,t)}{\Delta x^2} \qquad (5.32)$$

Con objeto de simplificar la escritura se propone la siguiente notación:

$$x = m\Delta x \qquad x+\Delta x = (m+1)\Delta x \qquad x-\Delta x = (m-1)\Delta x$$
$$t = p\Delta t \qquad t+\Delta t = (p+1)\Delta t$$
$$h(x,t)=h(m\Delta x,p\Delta t) = h_{m,p}$$

de modo que la ec. 5.32 se puede escribir:

$$\frac{h_{m,p+1}-h_{m,p}}{\Delta t} = a^2 \frac{h_{m-1,p}-2h_{m,p}+h_{m+1,p}}{\Delta x^2} \qquad (5.33)$$

o bien:

$$\boxed{h_{m,p+1} = \lambda h_{m-1,p} + (1-2\lambda) h_{m,p} + \lambda h_{m+1,p}} \qquad (5.34)$$

siendo:

$$\lambda = \frac{a^2 \Delta t}{\Delta x^2} \qquad (5.35)$$

Para cálculos con calculadora o la escritura de un programa de cómputo es útil representar la ec. 5.31 por la "molécula" de la fig. 5.3. En la molécula aparecen en los "átomos" los coeficientes de los distintos valores de la función.

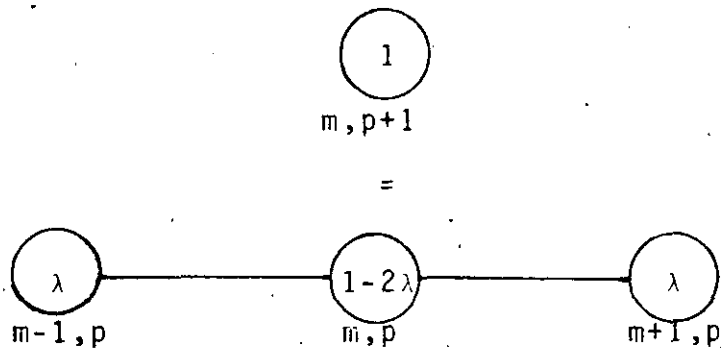


Fig. 5.3 "Molécula" de cálculo según la ec. 5.33.

Ejemplo 5.2

Para la aplicación de la ec. 5.34 considérese que en la fig. 5.2 la longitud L se divide en cuatro tramos de longitud Δx, en $x=0$ y $x=L$ la carga h es nula en todo tiempo y las cargas en el tiempo $t=0$ son en $x=\Delta x$, $h=1.5$; en $x=2\Delta x$, $h=2$ y en $x=3\Delta t$ $h=1.5$. $\Delta t=1$ S y $a^2/\Delta x^2 = 0.5$ 1/s.

Así las cosas se tiene $\lambda = 0.5$ y

$$h_{m,p+1} = 0.5\, h_{m-1,p} + 0.5 h_{m+1,p}$$

Para m=1,p=1

$$h_{1,2} = 0.5h_{0,1} + 0.5h_{2,1} = 0(0.5)+(0.5)2=1$$

Para m=2,p=1

$$h_{2,2} = 0.5_{h1,1}+0.5h_{3,1} = 0.5(1.5)+0.5(1.5)=1.5$$

Para m=3,p=1

$$h_{3,2} = 0.5h_{2,1}+ 0.5h_{4,1}=0.5(2)+0.5(0)=1$$

De esta manera, se han obtenido los valores consignados en la siguiente tabla:

t	p	m=0	m=1	m=2	m=3	m=4
0	0	0	1.5	2	1.5	0
1	1	0	1	1.5	1.	0
2	2	0	0.75	1.0	0.75	0
3	3	0	0.50	0.75	0.50	0
4	4	0	0.375	0.5	0.375	0

Así, en el tiempo t=4s, las cargas serían:

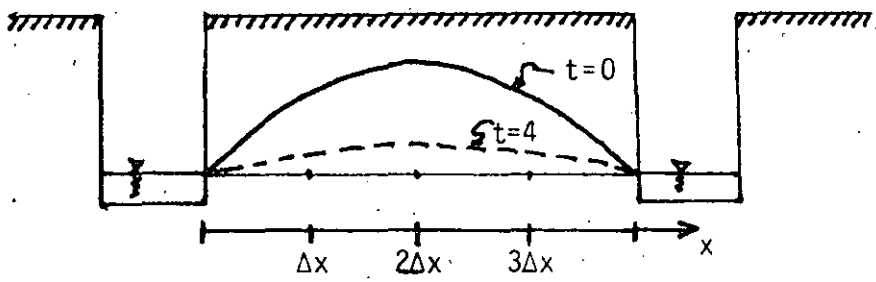


Fig. 5.4

Hasta aquí, todo parece sencillo y parece demasiado bueno para ser cierto, sin embargo, cuando se escoge un intervalo de tiempo de $\Delta t=2s$ a efecto de llegar con menos cálculos al tiempo t=4s, sucede que $\lambda=1$ y entonces:

$$h_{m,p+1} = h_{m,p-1} - h_{m,p} + h_{m,p+1}$$

y por tanto

t	p	m=0	m=1	m=2	m=3	m=4
0	0	0	1.5	2	1.5	0
2	1	0	0.5	1	0.5	0
4	2	0	0.5	0	0.5	0
6	3	0	-0.5	0.5	-0.5	0
8	4	0	1	-1.5	1	0

Se aprecia que las cargas como que "bailan", suben y bajan, " como que no" corresponden al problema físico. Por tanto, SE DEBE TENER CUIDADO EN LA SELECCION DEL INTERVALO DE TIEMPO, pues a veces no se esta encontrando una solución aproximada a la solución de la ec. diferencial.

La ec. 5.34 recibe el nombre de esquema de diferencias EXPLICITO, en la ecuación solo aparece una incognita y para valuarla no se necesita resolver algún sistema de ecuaciones en mas de una incognita.

Siguiendo con la ec. 5.30, si ahora la segunda derivada respecto a la distancia se representa como:

$$\left(\frac{\partial^2 h}{\partial x^2}\right)_{t=t+\Delta t} \doteq \frac{h_{m-1,p+1} - 2h_{m,p+1} + h_{m+1,p+1}}{\Delta x^2} \qquad (5.36)$$

y la derivada respecto al tiempo se aproxi como antes, otra versión en diferencias finitas de 5.30 es:

$$\frac{h_{m,p+1} - h_{m,p}}{\Delta t} = a^2 \frac{h_{m-1,p+1} - 2h_{m,p+1} + h_{m+1,p+1}}{\Delta x^2}$$

o bien:

$$\boxed{\lambda h_{m-1,p+1} - (1+2\lambda)\, h_{m,p+1} + \lambda h_{m+1,p+1} = -h_{m,p}} \qquad (5.37)$$

siendo:

$$\lambda = \frac{a^2 \Delta t}{\Delta x^2} \tag{5.38}$$

Ejemplo 5.3

Para la fig. 5.4, tambien con $L=4\Delta x$ y las mismas condiciones de frontera del ejemplo 5.2; el esquema 5.37 se plantea de la siguiente manera:

Para $p = 0$

$$\begin{aligned} m=1 \quad & \lambda h_{0,1} - (1+2\lambda) h_{1,1} + \lambda h_{2,1} = -h_{1,0} \\ m=2 \quad & \lambda h_{1,1} - (1+2\lambda) h_{2,1} + \lambda h_{3,1} = -h_{2,0} \\ m=3 \quad & \lambda h_{2,1} - (1+2\lambda) h_{3,1} + \lambda h_{4,1} = -h_{3,0} \end{aligned} \tag{5.39}$$

como por las condiciones de frontera, $h_{0,1}=h_{4,1}=0$, las ecs. 5.39 forman un sistema de ecuaciones lineales con incognitas $h_{1,1}$, $h_{2,1}$ y $h_{3,1}$, para conocer su valor se necesita resolver tal sistema.

Los esquemas de diferencias finitas como el de la ec. 5.37 se llaman IMPLICITOS porque en ellos aparece mas de una incognita y por lo tanto para conocer a estas se necesita resolver un sistema de ecuaciones.

Para la ec. 5.37 la molécula de cálculo aparece en la fig. 5.4

$$\begin{matrix} \lambda \\ m-1,p+1 \end{matrix} \qquad \begin{matrix} -1(1-2\lambda) \\ m,p+1 \end{matrix} \qquad \begin{matrix} \lambda \\ m+1,p+1 \end{matrix}$$

$$=$$

$$-1$$

$$m,p$$

Fig. 5.4 Molécula de cálculo de la ec. 5.37

La ec. 5.39 se puede escribir en forma matricial como:

$$\begin{bmatrix} -(1+2\lambda) & \lambda & 0 \\ \lambda & -(1+2\lambda) & \lambda \\ 0 & \lambda & -(1+2\lambda) \end{bmatrix} \begin{bmatrix} h_{1,1} \\ h_{2,1} \\ h_{3,1} \end{bmatrix} = \begin{bmatrix} -h_{1,0} \\ -h_{2,0} \\ -h_{3,0} \end{bmatrix}$$

o en forma abreviada:

$$[A]\,[h_1] = -[h_0]$$

al resolver el sistema anterior se conoce $[h_1]$; para el siguiente intervalo:

$$[A]\,[h_2] = -[h_1]$$

nuevamente al resolver el sistema, se obtiene $[h_2]$, y así se puede proseguir hasta el tiempo de interés.

Dos comentarios adicionales sobre este esquema son:

a) Este problema se puede simplificar si se toma en cuenta la simetría y se resuelve solo la mitad de $x=0$ a $x=\frac{L}{2}$ con lo cual se reducen las incognitas.

b) El sistema de ecuaciones anteriores se puede resolver mediante el método de la matriz inversa, así:

$$h_1 = [A]^{-1}\,[-h_0]$$
$$h_2 = [A]^{-1}\,[-h_1]$$
$$\vdots$$
$$h_{p+1} = [A]^{-1}[-h_p]$$

y entonces la matriz inversa se CALCULA SOLO UNA VEZ, para $p=0$, y luego para $p=1,2,3,\ldots$ ya sólo se efectua el producto de ella por el vector obtenido previamente.

Ejemplo 5.5

Para drenar un terreno de cultivo se han hecho dos canales paralelos a una distancia de L=16m. El suelo tiene un rendimiento específico de 0.10 y una permeabilidad de .2m/hr. El nivel freático promedio es de 0.5m. Se sabe que h=0 en x=0 y x=L y que las condiciones iniciales son $h=4x(L-x)/L^2$ si t=0. Obtener el cambio en el tiempo de la superficie freática.

Solución

La ecuación por resolver es:

$$\frac{\partial h}{\partial t} = a^2 \frac{\partial^2 h}{\partial x^2}$$

dividiendo la distancia en 10 tramos de longitud Δx= 1.6m se tiene:

$$a^2 = \frac{(0.5)(0.2)}{0.1} = 1 \quad \frac{m^2}{hr}$$

$$\lambda = \frac{a^2 \Delta t}{\Delta x} = \frac{1 \Delta t}{1.6}$$

escogiendo Δt=0.42667hr, λ = 1/6.

Por simetría se resolverá sólo la mitad, así la ec. 5.34 resulta:

$$h_{m,p+1} = \frac{1}{6} h_{m-1,p} + \left(\frac{2}{3}\right) h_{m,p} + \frac{1}{6} h_{m+1,p}$$

Para p=0

si m=1

$$h_{1,1} = \frac{1}{6} (0) + \left(\frac{2}{3}\right) (0.36) + \frac{1}{6} (0.64) = 0.34667$$

si m=2

$$h_{2,1} = \frac{1}{6}(0.36)+(\frac{2}{3})(0.64)+\frac{1}{6}(0.84) = 0.62667$$

si m=3

$$h_{3,1} = \frac{1}{6}(0.64)+(\frac{2}{3})(0.84)+\frac{1}{6}(0.96) = 0.82667$$

y así sucesivamente se obtiene

Tiempo (horas)	p	m						
		0	1	2	3	4	5	6
0	0	0	0.36	0.64	0.84	0.96	1	0.84
0.427	1	0	0.34667	0.62667	0.82667	0.94667	0.98667	0.94667
0.853	2	0	0.33556	0.61333	0.81333	0.93333	0.97333	0.93333
1.280	3	0	0.32593	0.60037	0.8000	0.92000	0.96000	0.92000
	⋮							
15.360	36	0	0.17653	0.33568	0.46184	0.54276	0.57061	0.54276
		NOTA: $h_6=h_4$, $h_7=h_3$, $h_8=h_2$, $h_9=h_1$						

La ecuación diferencial del ejemplo, para las condiciones de frontera e iniciales tiene como solución a:

$$h_{m,p} = \frac{32}{\pi^3} \sum_{i=1}^{\infty} \frac{1}{(2i+1)^3} e^{\frac{-(2i+1)^2\pi^2 p}{6m^2}} \operatorname{sen} \frac{(2i+1)\, m\Delta x}{L}$$

Se ha resuelto el ejemplo para diferentes valores de Δt usando los esquemas explícito e implícito para un tiempo t=15.36 horas y se comparan los resultados con la solución exacta.

h_1	h_2	h_3	h_4	h_5	Número de cálculos (p)	λ	Método
0.17655	0.33570	0.46187	0.54278	0.57065	36	1/6	explícito
0.17362	0.33398	0.45957	0.54014	0.56700	18	1/3	explícito
0.17480	0.33203	0.45752	0.53711	0.56543	12	1/2	explícito
0.07684	0.11140	0.20124	0.18005	0.24875	28	0.55	explícito
0.17793	0.33826	0.46526	0.54664	0.57465	72	1/12	implícito
0.17839	0.33911	0.46637	0.54790	0.57596	36	1/6	implícito
0.17931	0.34079	0.46838	0.55039	0.57853	18	1/3	implícito
0.18073	0.34246	0.47075	0.55282	0.58104	12	1/2	implícito
0.18093	0.34736	0.47706	0.55985	0.58628	6	1	implícito
0.18824	0.35672	0.48888	0.57284	0.60159	3	2	implícito
0.20682	0.38811	0.52741	0.61462	0.64426	1	6	implícito
0.17655	0.33541	0.46188	0.54280	0.57066	—	—	Solución exacta

De los resultados anteriores se observa que en el esquema explícito requiere que el intervalo de tiempo sea pequeño y que puede dar valores ilógicos para cierto tamaño (λ=0.55). El esquema implícito no presenta este último problema pero si a medida de que se aumenta Δt se encuentran resultados mas apartados del exacto.

Adelante se verá, que el esquema explícito requiere que:

$$\lambda \leq \frac{1}{2}$$

y que el implícito:

$$\lambda > 0$$

o sea que prácticamente no existe restricción, aunque si es conveniente que sea λ pequeño a efecto de no tener una diferencia grande respecto a la solución exacta (luego se discute el porque).

APROXIMACION DE PROMEDIOS PESADOS

Aunque el método explícito es simple de usar tiene el inconveniente de requerir un paso de tiempo Δt muy pequeño. A efecto de reducir el volumen de cálculo se sugiere una mas general aproximación de diferencias finitas a la ec. 6.1 mediante:

$$\frac{\partial h}{\partial t} \doteq \frac{h_{m,p+1}-h_{m,p}}{\Delta t} \qquad (5.40)$$

$$\frac{\partial^2 h}{\partial x^2} = \frac{1}{\Delta x^2}\left[\theta(h_{m-1,p+1}-2h_{m,p+1}+h_{m+1,p+1}) + (1-\theta)(h_{m-1,p}-2h_{m,p}+h_{m+1,p})\right] \qquad (5.41)$$

donde

$$0 \leq \theta \leq 1$$

Se observa que esta última ecuación corresponde a una aproximación a la derivada segunda respecto a x (ec. 5.24) tanto para $t=p\Delta t$ como para $t=(p+1)\Delta t$ y que ambas se promedian de acuerdo al valor del factor θ. Si se desea darle mayor importancia a la aproximación en $t=(p+1)\Delta t$, θ adquiere un valor mayor a 1/2 y menor o igual a 1.

El esquema de diferencias finitas resulta ser:

$$h_{m,p+1}-h_{m,p}=\lambda[\theta(h_{m-1,p+1}-2h_{m,p+1}+h_{m+1,p+1})+(1-\theta)(h_{m-1,p}-2h_{m,p}+h_{m+1,p})] \qquad (5.42)$$

siendo

$$\lambda = a^2\Delta t/\Delta x^2$$

El esquema anterior representa tambien tres casos de interés, dos de los cuales ya se comentaron, los cuales son a saber:

a) El método explícito (ec. 5.34) cuando $\theta = 0$

b) El método implícito (ec. 5.37) cuando $\theta = 1$

c) El método propuesto por Crank y Nicolson en 1947 si $\theta = \frac{1}{2}$

En la fig. 5.5 se muestra la molécula de cálculo del esquema dado por 5.42:

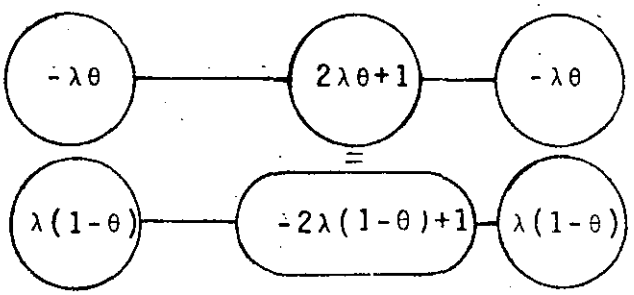


Fig. 5.5 Molécula de cálculo del esquema de promedios pesados.

5.2 Convergencia, estabilidad y consistencia

Para que la solución de las ecuaciones de diferencias finitas tengan una razonable aproximación a la solución de la correspondiente ecuación diferencial parcial parabólica o hiperbólica se deben cumplir algunas condiciones; estas estan asociadas con dos problemas interrelacionados, el primero se refiere a la tendencia a parecerse la solución del esquema de diferencias a la solución exacta, el segundo esta relacionado al decaimiento controlado o crecimiento desproporcionado de cualquier error asociado con la solución de diferencias finitas.

5.2.1 Convergencia

Sea H la solución exacta de la ec. 5.10 y h la solución exacta del esquema de diferencias finitas usado para aproximar a la ec. 5.10. La ecuación de diferencias finitas se dice CONVERGENTE cuando h tiende a H en un punto fijo o a lo largo de un nivel y cuando Δx y Δt ambos tienden a cero.

Lo anterior significa que el error de discretización o truncado (ec. 4.78) tiende a cero a medida que también lo hacen Δt y Δx.

Ejemplo 5.6

Sea $\varepsilon = H-h$ y $\frac{\partial h}{\partial t} = a^2\frac{\partial^2 h}{\partial x^2}$, $0<x<1$, y

H es conocida para:

$0<x\leq 1$ cuando $t=0$ y en $x=0$ y 1 cuando $t\geq 0$.

Considerando el esquema explícito y que

$$h_{m,p} = H_{m,p} - \varepsilon_{m,p};\ h_{m+1,p} = H_{m+1,p} - \varepsilon_{m+1,p},\cdots$$

De acuerdo con la definición de error de truncado de la serie de Taylor (ec. 5.11) resulta:

$$\theta(\Delta z^n) = \frac{\partial^n f(z+\theta\Delta z)}{\partial z^n}\bigg|_{máx} \quad \text{donde } 0 \leq \theta \leq 1 \qquad (5.43)$$

lo anterior significa que θ se debe escoger entre 0 y 1 de modo tal que la derivada de f de orden n respecto a z sea máxima entre z y $z+\Delta z$.

Con base en lo anterior se tiene según la serie de Taylor:

$$H(x,t+\Delta t) = H(x,t) + \frac{\partial H}{\partial t}(x,t+\theta_1\Delta t) \qquad (5.44)$$

de donde:

$$H(x,t+\Delta t)-H(x,t) = \frac{\partial H}{\partial t}(x,t+\theta_1\Delta t)\Delta t \qquad (5.45)$$

ahora desarrollando en serie de Taylor según x

$$H(x+\Delta x,t) = H(x,t) + \frac{\partial H}{\partial x}(x,t)\,\Delta x + \frac{\partial^2 H}{\partial x^2}(x+\theta_2\Delta x,t)\,\frac{\Delta x^2}{2}$$

y

$$H(x-\Delta x,t) = H(x,t) - \frac{\partial H}{\partial x}(x,t)\Delta x + \frac{\partial^2 H}{\partial x^2}(x+\theta_3\Delta x,t)\,\frac{\Delta x^2}{2}$$

sumando y arreglando

$$H(x-\Delta x,t)-2H(x,t)+H(x+\Delta x,t) = \frac{\partial^2}{\partial x^2}H(x+\theta_4\Delta x,t) \qquad (5.46)$$

Usando la notación con índices en 5.45 y 5.46

$$H_{m,p+1} - H_{m,p} = \frac{\partial H}{\partial t}(m\Delta x,p\Delta t+\theta_1\Delta t) \qquad (5.47)$$

$$H_{m-1,p}-2H_{m,p} + H_{m+1,p} = \frac{\partial^2}{\partial x^2}\,H(m\Delta x+\theta_4\Delta x,p\Delta t) \qquad (5.48)$$

Por otra parte, segun el esquema de diferencias y la definición de error:

$$H_{m,p+1}-\varepsilon_{m,p+1} = \lambda(H_{m-1,p}-\varepsilon_{m-1,p})+(1-2\lambda)(H_{m,p}-\varepsilon_{m,p}) +$$

$$\lambda(H_{m+1,p}-\varepsilon_{m+1,p})$$

ordenando

$$-\varepsilon m,p+1 +\lambda\varepsilon m-1,p+(1-2\lambda)\varepsilon\lambda m,p+\lambda\varepsilon m+1,p +\lambda(Hm,p+1-Hm,p)-$$

$$\lambda(H_{m-1,p}-2Hm,p+ H_{m+1,p}) = 0 \qquad (5.49)$$

sustituyendo 5.46 y 5.48 en 5.49

$$\varepsilon_{m,p+1} = \lambda\,\varepsilon_{m-1,p} + (1-2\lambda)\,\varepsilon_{m,p} + \lambda\varepsilon_{m+1,p} + M\,\Delta t \qquad (5.50)$$

donde

$$M = \frac{\partial H}{\partial t}(m\Delta x, p\Delta t+\theta_1\Delta t) - \lambda\frac{\partial^2 H}{\partial x^2}(m\Delta x+\theta_4\Delta x, p\Delta t) \qquad (5.51)$$

Para asegurar que los errores tengan signo positivo y continuar con el control del error en este análisis, se tomará valor absoluto de 5.50 y así:

$$|\varepsilon m,p+1| \leq \lambda|\varepsilon m-1,p| + (1-2\lambda)|\varepsilon m,p|+\lambda|\varepsilon m+1,p| + /M/\Delta t \qquad (5.52)$$

esta ecuación es cierta cuando los coeficientes son positivos o iguales a cero. Si el mayor de los errores para cualquier m en el instante p es Ep, al asignar:

$$Ep \rightarrow \varepsilon m-1,p$$
$$Ep \rightarrow \varepsilon m,p$$
$$Ep \rightarrow \varepsilon m+1,p$$

La ec. 5.52 sigue del lado conservador (pues se esta considerando más error que el real(y se transforma a:

$$E_{p+1} \leq \lambda E_p + (1-2\lambda)\, E_p + \lambda E_p + |M|\Delta t$$

$$E_{p+1} \leq E_p + \Delta t|M| \qquad (5.53)$$

así, si M se considera constate

$$E_1 \leq E_0+\Delta t|M|$$
$$E_2 \leq E_1+\Delta t|M|, \; E_2 \leq E_0+\Delta t|M|$$
$$E_3 \leq E_2+\Delta t|M|, \; E_3 \leq E_0+3\Delta t|M|$$
$$E_{p+1} \leq E_0+(p+1)\Delta t|M|$$

como al tiempo cero no existe error $E_0 = 0$ y por tanto

$$E_{p+1} \leq (p+1)\Delta t|M| \qquad (5.54)$$

si $\Delta t \to 0$ $\lambda \to 0$ y la ec. 6.51 $M \to 0$

por tanto 5.54 tambien tiende a cero. Y POR LO TANTO LA SOLUCION h CONVERGE A LA H. Para esto se requiere que los coeficientes de 6.52 sean positivos o iguales a cero como se apuntó antes, para que lo sean

$$\boxed{0 \leq \lambda \leq \tfrac{1}{2}}$$

lo anterior constituye la llamada condición de convergencia.

5.2.2 Estabilidad

Si ahora se estudia el comportamiento exclusivamente del error de redondeo (ec. 4.78), se afirma que, cuando a medida de que se utiliza un esquema de diferencias finitas para una ec. del tipo de la 4.10 y a medida de que se avanza en "y" el error (de redondeo) acumultado tiene a un límite de magnificación el esquema de diferencias finitas es ESTABLE.

Los textos sobre este tema, al tratar estabilidad se refieren al error de redondeo y señalan que es difícil estimarlo, ya que depende del tipo de computadora usada; en realidad el error que se analiza puede ser de un tipo distinto al de redondeo, como sería aquel debido a una equivocación y si tal error llega a estar limitado después de un gran número de pasos en "y", el esquema de diferencias es estable; cuando el error crece sin tender a un valor definido se dice inestable, y la mayoría de las veces adquiere valores que oscilan y difieren cada vez mas.

Para tratar la estabilidad se puede aplicar el método de von Neumann o de la serie de Fourier, el método consiste en expresar el error en un punto fijo $x=m\Delta x$ y $y=p\Delta y$ como:

$$E_{m,p} = \sum_{n=0}^{N} A_n \; e^{i\beta nm\Delta x} \; e^{\alpha p\Delta t} \qquad (5.55)$$

donde

$$\beta_n = n\Pi/N\Delta x \;, \quad N\Delta x = L \quad e \quad i = \sqrt{-1}$$

Se considera que analizando un término de la serie de Fourier, se conoce el comportamiento de toda ella, y que los coeficientes A_n pueden ser despreciados; así el error se toma como

$$E_{m,p} = e^{i\beta m\Delta x} \; e^{\alpha p\Delta y} = e^{i\beta m\Delta x} \; \xi^p \qquad (5.56)$$

y se deduce que el error no crecerá al aumentar p cuando

$$|\xi| \leq 1 \qquad (5.57)$$

Este criterio permite establecer las condiciones que se deben cumplir para que un esquema sea estable. En la práctica, el representar de este modo el error es útil aunque no está completamente justificado el método que a continuación se enuncia.

Ejemplo 5.7

Realizar el análisis de estabilidad del esquema explícito (ec. 5.34).

Sea el error (de redondeo)

$$E = h - N \qquad (5.58)$$

siendo h la solución del esquema de diferencias finitas con todas sus cifras decimales y N la solución del esquema con un número limitado de cifras.

Exprcsando el error (de redondeo) en términos de los subíndices m y p se tiene

$$N_{m,p} = h_{m,p} - E_{m,p},\ N_{m+1,p} = h_{m+1,p} - E_{m+1,p}, \qquad (5.59)$$

$$N_{m-1,p} = h_{m-1,p} - E_{m-1,p}, \ldots$$

Según el esquema dado por 5.34

$$N_{m,p+1} = \lambda N_{m-1,p} + (1-2\lambda)\, N_{m,p} + \lambda N_{m+1,p} \qquad (5.60)$$

Al sustituir las ec. 5.59 en 5.60

$$h_{m,p+1} - E_{m,p+1} = \lambda(h_{m-1,p} - E_{m-1,p}) + (1-2\lambda)\,(h_{m,p} - E_{m,p}) + \lambda(h_{m+1,p} - E_{m+1,p})$$

arreglando términos

$$h_{m,p+1} - \lambda h_{m-1,p} - (1-2\lambda)\, h_{m,p} - h_{m+1,p} = E_{m,p+1} - \lambda E_{m-1,p} - 2\lambda E_{m,p} - \lambda E_{m+1,p}$$

Ya que h es la solución del esquema de diferencias finitas, $h_{m,p+1} = \lambda h_{m-1,p} + (1-2\lambda)\, h_{m,p} + \lambda h_{m+1,p}$, y por lo tanto TODO el miembro izquierdo de la ec. 5.61 es nulo. Así

$$E_{m,p+1} = \lambda E_{m-1,p} + (1-2\lambda)\, E_{m,p} + \lambda E_{m+1,p} \qquad (5.61)$$

Nótese que la ec. 5.61 tiene exactamente la forma del esquema de diferencias finitas.

Ahora si se expresan los errores E como en 5.56

$$e^{i\beta m\Delta x}\,\xi^{p+1} = \lambda e^{i\beta(m-1)\Delta x}\,\xi^{p} + (1-2\lambda)e^{i\beta m\Delta x}\,\xi^{p} + \lambda e^{i\beta(m+1)\Delta x}\,\xi^{p}$$ al dividir entre $e^{i\beta m\Delta x}\xi^{p}$

$$\xi = \lambda e^{-i\beta\Delta x} + (1-2\lambda) + \lambda e^{i\beta\Delta x}$$

$$\xi = \lambda(e^{i\beta\Delta x} + e^{-i\beta\Delta x}) + (1-2\lambda)$$

$$\xi = \lambda(e^{i\beta\Delta x} + e^{-i\beta\Delta x} - 2) + 1 \qquad (5.62)$$

Por otra parte, de las propiedades de las funciones seno, y coseno tiene que $\frac{1}{2}(e^{i\theta} + e^{-i\theta}) = \cos\theta$

así

$$\frac{1}{2}(e^{i2\theta} + e^{-i2\theta}) = \cos 2\theta$$

tambien

$$\operatorname{sen}^2\theta = \frac{1 - \cos 2\theta}{2}, \text{ por lo tanto}$$

$$\operatorname{sen}^2\theta = \frac{1 - \frac{1}{2}(e^{i2\theta} + e^{-i2\theta})}{2}$$

o bien

$$4\operatorname{sen}^2\theta = 2 - e^{i2\theta} - e^{-i2\theta}$$

si

$$\text{si } 2\theta = \beta\Delta x$$

$$4\operatorname{sen}^2\frac{\beta\Delta x}{2} = 2 - e^{i\beta\Delta x} - e^{-i\beta\Delta x} \qquad (5.63)$$

Al considerar la identidad 5.63 en 5.62

$$\xi = 1-(\lambda 4\operatorname{sen}^2\frac{\beta\Delta x}{2})$$

ya que para que no aumenten los errores

$$|\xi| < 1 \text{ ó sea } -1 \leq \xi \leq 1$$

entonces

$$-1 \leq 1- \lambda 4\operatorname{sen}^2\frac{\beta\Delta x}{2} \leq 1$$

caso 1

$$1-\lambda 4\operatorname{sen}^2\frac{\beta\Delta x}{2} \leq 1$$

simplificando

$$-\lambda 4 \operatorname{sen}^2 \frac{\beta \Delta x}{2} \leq 0$$

lo cual se cumple para $\lambda \geq 0$, pues $\operatorname{sen}^2 \frac{\beta \Delta x}{2}$ es una cantidad positiva.

caso 2

$$-1 \leq 1 - \lambda 4 \operatorname{sen}^2 \frac{\beta \Delta x}{2}$$

reduciendo

$$-2 \leq -\lambda 4 \operatorname{sen}^2 \frac{\beta \Delta x}{2}$$

$$\tfrac{1}{2} \geq \lambda \operatorname{sen}^2 \frac{\beta \Delta x}{2}$$

como el valor mas grande de $\operatorname{sen}^2 \frac{\beta \Delta x}{2}$ es uno.

$$\tfrac{1}{2} \geq \lambda$$

por lo tanto

$$\boxed{0 \leq \lambda \leq \tfrac{1}{2}}$$

lo cual constituye la condición de estabilidad del esquema.

El lector interesado puede realizar un análisis de estabilidad para el esquema.

$$\frac{h_{m,p+1} - h_{m,p-1}}{2\Delta t} = \frac{a^2}{\Delta x^2} \left[h_{m-1,p} - 2h_{m,p} + h_{m+1,p}\right] \qquad (5.64)$$

el cual es muy parecido al explícito; la diferencia consiste en que se ha aproximado la parcial respecto al tiempo, por una diferencia central.

El análisis de estabilidad lleva a

$$|\xi| > 1 + 4\lambda \operatorname{sen}^2 \frac{\beta \Delta x}{2}$$

lo cual significa que la ec. 5.64 es siempre inestable. Este resultado resalta la importancia de atender las propiedades matemáticas de las ecuaciones de diferencias.

Ejemplo 5.8

Plantear un esquema de diferencias finitas para resolver

$$\frac{\partial^2 h}{\partial x^2} = \frac{\partial h}{\partial t} + c\,\frac{\partial^2 h}{\partial t^2} \qquad (5.65)$$

Escogiendo

$$\frac{\partial^2 h}{\partial x^2} = \frac{h_{m-1,p} - 2h_{m,p} + h_{m+1,p}}{\Delta x^2} \qquad (5.66)$$

$$\frac{\partial h}{\partial t} = \frac{h_{m,p+1} - h_{m,p-1}}{2\Delta t} \qquad (5.67)$$

$$\frac{\partial^2 h}{\partial t^2} = \frac{h_{m,p-1} - 2h_{m,p} + h_{m,p+1}}{\Delta t^2} \qquad (5.68)$$

Al sustituir 5.66, 5.67 y 5.68 en 5.65 y después de varios pasos algebráicos se llega a

$$\boxed{h_{m,p+1} = \frac{(\frac{1}{2} - \frac{c}{\Delta t})h_{m,p-1} + \lambda h_{m-1,p} + (-2\lambda + \frac{2c}{\Delta t})h_{m,p} + \lambda h_{m+1,p}}{\frac{1}{2} + \frac{c}{\Delta t}}} \qquad (5.69)$$

donde

$$\lambda = \frac{\Delta t}{\Delta x^2} \qquad (5.70)$$

Si por alguna razón se escoge un valor para Δt igual a $2c$ la ec. 5.69 deviene en

$$h_{m,p+1} = \lambda h_{m-1,p} + (1-2\lambda)\, h_{m,p} + \lambda h_{m+1,p} \qquad (5.71)$$

Pero la ec. 5.71 es el esquema de diferencias finitas de la ec.

$$\frac{\partial h}{\partial t} = a^2 \frac{\partial^2 h}{\partial x^2}$$

CUANDO EL ESQUEMA RESUELVE UNA ECUACION DIFERENCIAL DISTINTA A LA QUE SE PRETENDE RESOLVER SE DICE QUE EL ESQUEMA DE DIFERENCIAS ES INCONSISTENTE.

5.2.3 Consistencia

Para encontrar el error de discretización o truncado en de un paso a otro de Δx se refiere a la serie de Taylor, para aclarar esto, cosidere que interesa valuar el error de discretización local del esquema explícito (ec. 5.34). Entonces se requieren estos desarrollos en series de Taylor.

$$H(x+\Delta x,t) = H(x,t) + \frac{\partial}{\partial x}H(x,t)\Delta x + \frac{\partial^2}{\partial x}H(x,t)\frac{\Delta x^2}{2!} + \ldots + \frac{\partial^n}{\partial x^n} H(x,t)\frac{\Delta x^n}{n!}$$

así

$$H_{m+1,p} = H_{m,p} + \left(\frac{\partial H}{\partial x}\right)_{mp}\Delta x + \left(\frac{\partial^2 H}{\partial x^2}\right)\frac{\Delta x^2}{2} + \ldots + \left(\frac{\partial^n H}{\partial x^n}\right)_{m,p} \frac{\Delta x^n}{n!}$$

$$H_{m-1,p} = H_{m-1,p} - \left(\frac{\partial H}{\partial x}\right)_{m,p}\Delta x + \left(\frac{\partial^2 H}{\partial x^2}\right)_{m,p}\frac{\Delta x^2}{2} + \ldots + (-1)^n \left(\frac{\partial^n H}{\partial x^n}\right)\frac{\Delta x^n}{n!}$$

sumando y ordenando

$$\frac{H_{m-1,p} - 2H_{m,p} + H_{m+1,p}}{\Delta x^2} = \left(\frac{\partial^2 H}{\partial x^2}\right)_{m,p} + \left(\frac{\partial^4 H}{\partial x^4}\right)\frac{\Delta x^2}{12} + \left(\frac{\partial^6 H}{\partial x^6}\right)_{m,p} \frac{\Delta x^4}{360} + \ldots \qquad (5.72)$$

Ahora, por lo que se refiere al tiempo

$$H_{m,p+1} = H_{m,p} + \left(\frac{\partial H}{\partial t}\right)_{m,p} \Delta t + \left(\frac{\partial^2 H}{\partial t^2}\right)_{m,p} \frac{\Delta t^2}{2} + \left(\frac{\partial^3 H}{\partial t^3}\right)_{m,p} \frac{\Delta t^3}{3!} + \ldots$$

De esta exprexión se llega a

$$\frac{H_{m,p+1} - H_{m,p}}{\Delta t} = \left(\frac{\partial H}{\partial t}\right)_{m,p} + \left(\frac{\partial^2 H}{\partial t^2}\right)_{m,p} \frac{\Delta t}{2!} + \left(\frac{\partial^3 H}{\partial t^3}\right)_{m,p} \frac{\Delta t^2}{3!} + \left(\frac{\partial^4 H}{\partial t^4}\right)_{m,p} \frac{\Delta t^3}{4!} + \ldots \tag{5.73}$$

Al sustituir 5.72 y 5.73 en 5.33 y después de ordenar

$$\left(\frac{\partial H}{\partial t} - a^2 \frac{\partial^2 H}{\partial x^2}\right)_{m,p} + \left(\frac{\partial^2 H}{\partial t^2}\right)_{m,p} \frac{\Delta t}{2!} + \left(\frac{\partial^3 H}{\partial t^3}\right)_{m,p} \frac{\Delta t^2}{3!} + \ldots - \left(a^2 \frac{\partial^4 H}{\partial x^4}\right)_{m,p} \frac{\Delta x^2}{12} -$$

$$- \left(a^2 \frac{\partial^6 H}{\partial x^6}\right)_{m,p} \frac{\Delta x^4}{360} - \ldots = 0 \tag{5.74}$$

por la ecuación diferencial por resolver, el primer paréntesis es nulo, lo què queda se conoce como error de truncado $T(H)_{m,p}$, esto es

$$T(H)_{m,p} = \left(\frac{\partial^2 H}{\partial t^2}\right)_{m,p} \frac{\Delta t}{2} + \left(\frac{\partial^3 H}{\partial t^3}\right)_{m,p} \frac{\Delta t^2}{6} + \ldots \left(a^4 \frac{\partial^4 H}{\partial x^4}\right)_{m,p} \frac{\Delta x^4}{12} -$$

$$\left(a^6 \frac{\partial^6 H}{\partial x^6}\right)_{m,p} \frac{\Delta x^4}{360} - \ldots \tag{5.75}$$

si $\Delta t \to 0$ y $\Delta x \to 0$ se observa que $T(H)_{m,p} \to 0$.

Definición

Si el error de truncado local tiende a cero cuando las diferencias discretas Δt y Δx tienden a cero, el esquema de diferencias finitas es CONSISTENTE

Por otra parte, la ec. 5.75 se puede escribir como

$$T(H)_{m,p} = \frac{\Delta t}{2} \left(\frac{\partial^2 H}{\partial t^2} - a^4 \frac{\partial^4 H}{\partial x^4} \frac{\Delta x^2}{6 \Delta t}\right)_{m,p} + \left(\frac{\partial^3 H}{\partial t^3}\right)_{m,p} \frac{\Delta t^2}{6} -$$

$$- \left(a^6 \frac{\partial^6 H}{\partial x^6}\right)_{m,p} \frac{\Delta x^4}{360} + \ldots . \tag{5.76}$$

al considerar

$$\frac{\Delta x^2}{6\Delta t} = a^2 \qquad (5.77)$$

resulta que el primer paréntesis es

$$\frac{\partial^2 H}{\partial t^2} - a^4 \frac{\partial^4 H}{\partial x^4}$$

como

$$\frac{\partial H}{\partial t} = a^2 \frac{\partial^2 H}{\partial x^2} \qquad (5.78)$$

se tiene que al derivarla respecto al tiempo

$$\frac{\partial^2 H}{\partial t^2} = a^2 \frac{\partial^3 H}{\partial x^2 \partial t} \qquad (5.79)$$

y que al derivar 5.78 dos veces respecto a x

$$\frac{\partial^3 H}{\partial t \partial x^2} = a^2 \frac{\partial^4 H}{\partial x^4} \qquad (5.80)$$

como se puede intercarbiar el order de derivación, al igualar 5.79 con 5.80

$$\frac{\partial^2 H}{\partial t^2} - a^4 \frac{\partial^4 H}{\partial x^4} = 0$$

así

$$T(H)_{m,p} = \overline{\left(\frac{\partial^3 H}{\partial t^3}\right)_{m,p}} \frac{\Delta t^2}{12} - \left(a^6 \frac{\partial^6 H}{\partial x^6}\right)_{m,p} \frac{\Delta x^4}{360} + \ldots \quad \text{si} \quad \frac{\Delta x^2}{6\Delta t} = a^2$$

o sea que se involucra a errores en términos de Δt^2 y Δx^4, es decir $\theta(\Delta t^2, \Delta x^4)$, por lo tanto es mejor pues así el error de truncado es mas pequeño. Esto significa que

para $\lambda = \frac{1}{6}$ en el esquema explícito es conveniente pues reduce el error de truncado, de otro modo, $\lambda \neq \frac{1}{6}$ implica $\theta(\Delta t, \Delta x^2)$.

La extensión de diferencias finitas en la Hidráulica es muy amplia y así puede plantearse la solución de las ecuaciones de flujo no permanente entre otras.

5.3 Método de las características

Recuérdese que la diferencial total de una función $u(x,y)$ es

$$du = \frac{\partial u}{\partial x} dx + \frac{\partial u}{\partial y} dy \qquad (5.81)$$

sin embargo, se puede pensar en un método que simplifique el cálculo de du, para ilustrar esto considere el siguiente caso particular.

Ejemplo 5.9

Encontrar la función $u(x,y)$ que satisface a

$$\frac{\partial u}{\partial x} + 2x \frac{\partial u}{\partial y} = y \qquad (5.82)$$

sujeta a la condición $u(0,y) = 1+y^2$ para $1 < y < 2$

Sea la derivada total

$$\frac{du}{dx} = \frac{\partial u}{\partial x} + \frac{\partial u}{\partial y} \frac{dy}{dx} \qquad (5.83)$$

comparando 5.83 y 5.82 resultan

$$2x = \frac{dy}{dx} \qquad (5.84)$$

$$y = \frac{du}{dx} \qquad (5.85)$$

De 5.84

$$y = x^2 + c_1 \qquad (5.86)$$

Sustituyendo 5.86 en 5.85 e integrando

$$\frac{x^3}{3} + c_1 x + c_2 = u \qquad (5.87)$$

Si la ec. 5.87 pasa por (x_0, y_0)

$$y_0 = x_0^2 + C_1$$

así 5.86 queda

$$y = x^2 + y_0 - x_0^2 \qquad (5.88)$$

Para encontrar a C_2, se sabe de la restricción que

$$u = 1 + y^2 \text{ en } x = 0$$

así, según 5.88 y 5.87

$$u = 1+(0^2+y_0-x_0^2) = \frac{0^3}{3} + C_1(0)+C_2$$

por lo que

$$C_2 = 1+ (y_0-x_0^2)^2$$

de este modo

$$u = \frac{1}{3} x^3 + \underbrace{(y_0 - x_0^2)}_{C_1} x + \underbrace{[1+(y_0-x_0^2)^2]}_{C_2}$$

como (x_0, y_0) puede ser punto de la región sombreada de la fig. 5.6, se tiene

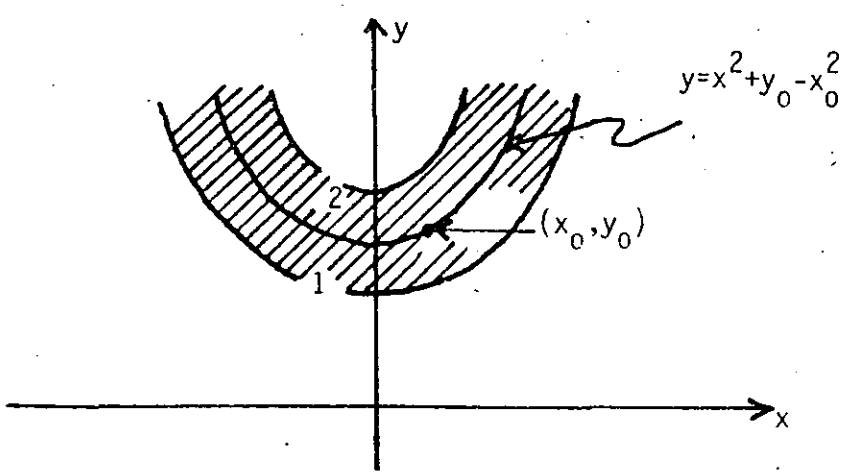


Fig. 5.6

$$\boxed{u = \frac{1}{3}x^3 + (y-x^2)x + 1+(y-x^2)^7}$$

u únicamente esta determinada en la zona sombreada, no se sabe nada de ella fuera de tal área.

Observaciones sobre el ejemplo

a) En lugar de resolver una ec. diferencial parcial se resolvieron dos ec. diferenciales ordinarias.

b) La ec. 5.84 permitió conocer una función a lo largo de la cual u varía independientemente de $\frac{\partial u}{\partial x}$ y $\frac{\partial u}{\partial y}$.

c) Como u varía según la ec. 5.84 se pudo determinar u.

5.3.1 Método de características para una ec. de segundo orden

Sea la ec. diferencial paracial lineal de segundo orden

$$\boxed{A\frac{\partial^2\phi}{\partial x^2} + B\cdot\frac{\partial^2\phi}{\partial x\partial y} + C\frac{\partial^2\phi}{\partial y^2} + E = 0} \qquad (5.90)$$

si

$$p = \frac{\partial\phi}{\partial x} \qquad (5.91)$$

$$q = \frac{\partial\phi}{\partial y} \qquad (5.92)$$

$$s = \frac{\partial\phi}{\partial x\partial y} = \frac{\partial\phi}{\partial x\partial y} = \frac{\partial p}{\partial y} = \frac{\partial q}{\partial x} \qquad (5.93)$$

Sustituyendo 5.91, 5.92 y 5.93 en 5.90

$$A\frac{\partial p}{\partial x} + Bs + C\frac{\partial q}{\partial t} + E = 0 \qquad (5.94)$$

Por otro lado se sabe

$$\frac{dp}{dx} = \frac{\partial p}{\partial x} + \frac{\partial p}{\partial y}\frac{dy}{dx}$$

$$\frac{dq}{dy} = \frac{\partial q}{\partial x}\frac{dx}{dy} + \frac{\partial q}{\partial y}$$

por lo que al considerar 5.93

$$\frac{\partial p}{\partial x} = \frac{dp}{dx} - s\frac{dy}{dx} \qquad (5.95)$$

$$\frac{\partial q}{\partial y} = \frac{dq}{dy} - s\frac{dx}{dy} \qquad (5.96)$$

Al sustituir 5.95 y 5.96 en 5.94 y ordenando

$$s\left[-A\frac{dy}{dx} + B - C\frac{dx}{dy}\right] + A\frac{dp}{dx} + C\frac{dq}{dy} + E = 0$$

Multiplicando por $\frac{dy}{dx}$

$$s\left[-A\left(\frac{dy}{dx}\right)^2 + B\frac{dy}{dx} - C\right] + A\frac{dp}{dx}\frac{dy}{dx} + C\frac{dq}{dy}\frac{dy}{dx} + E\frac{dy}{dx} = 0 \qquad (5.97)$$

como s corresponde a la derivada mixta y para simplificar la ec. anterior conviene hacer nulo el primer término ésta ecuación, de modo que

$$-A\left(\frac{dy}{dx}\right)^2 + B\frac{dy}{dx} - C = 0 \qquad (5.98)$$

Al resolver esta ecuación de segundo orden

$$\frac{dy}{dx} = \frac{-B \pm \sqrt{B^2 - 4AC}}{-2A} \qquad (5.99)$$

A efecto de garantizar dos raices reales distintas

$$B^2 - 4AC > 0 \qquad (5.100)$$

Llamando

$$\boxed{\frac{dy}{dx} = \frac{-B + \sqrt{B^2 - 4AC}}{-2A} = f} \qquad (5.101)$$

$$\boxed{\frac{dy}{dx} = \frac{-B - \sqrt{B^2 - 4AC}}{-2A} = g} \qquad (5.102)$$

(f y g se llaman características).

La otra parte de la ec. 5.97

$$A \frac{dp}{dx} \frac{dy}{dx} + C \frac{dq}{dy} \frac{dy}{dx} + E \frac{dy}{dx} = 0$$

al multiplicar por dx

$$A\, dp \frac{dy}{dx} + C\, dq + E dy = 0 \qquad (5.103)$$

Si se considera 5.101

$$\boxed{A\, dp\, f + c\, dq + E dy = C} \qquad (5.104)$$

y si ahora se toma en cuenta en 5.103 a 5.102

$$\boxed{A\, dp\, g + C\, dq + E dy = 0} \qquad (5.105)$$

Esto significa que la ec. diferencial parcial de segundo orden puede transformarse (cuando es hiperbólica, pues $B^2-4AC>o$) a cuatro ecuaciones diferenciales ordinarias (5.101, 5.102, 5.104 y 5.105).

Para la solución de 5.94 se considera

$$d\phi = \frac{\partial \phi}{\partial x} dx + \frac{\partial \phi}{\partial y} dy$$

o sea

$$\boxed{d\phi = pdx + q\, dy} \qquad (5.106)$$

El método de solución consiste en

1. Encontrar f y g según 5.101 y 5.102
2. Sustituir f y g en 5.104 y 5.105, con lo cual se forma un sistema de ecuaciones con incógnitas p y q, al resolverlo se conocen estas.
3. Conocidas p y q se sustituyen en 5.106 y se integra para conocer ϕ.

En la fig. 5.7 se ilustra este procedimiento

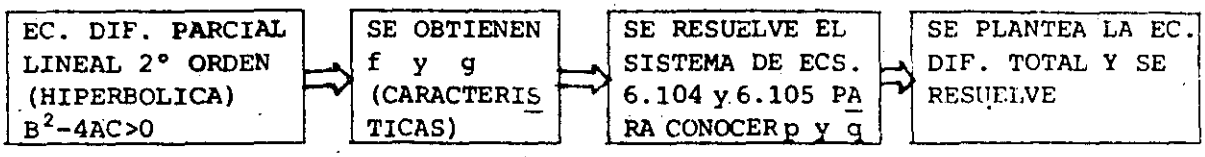


Fig. 5.7

Cuando las ecuaciones diferenciales ordinarias son difíciles de resolver o no tienen solución mediante métodos analíticos, el plateamiento anterior se puede realizar en diferencias finitas, para ello considérese las aproximaciones a las ecs. 5.101, 5.102 5.104, 5.105 y 5.106. Para ello, tome se en cuenta a la fig. 5.8

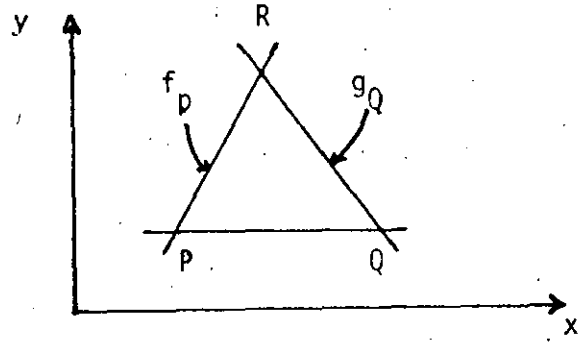


Fig. 5.8

En los punto P y Q se conoce el valor de x,y,f,g, p y q, mientras que el punto R no se conoce ninguna.

De las ecs. 5.101 y 5.102, al integrar se definen dos rectas, características, cada una de las cuales tienen pendiente f y g. En la intersección de ellas, se define el punto R. Cuando ellas se definen en términos de valores conocidos (del paso Δy anterior) se tiene, para f

$$f_p = \frac{y_R - y_p}{x_R - x_p} \qquad (5.107)$$

y para g

$$g_Q = \frac{y_R - y_Q}{x_R - x_Q} \qquad (5.108)$$

Al resolver simultáneamente 5.107 y 5.108 se encuentran x_R y y_R.

Por otra parte, la ec. 5.104 en diferencias puede ser:

$$A_p (p_R - p_P) f_p + C_p (q_R - q_P) + E_p (y_R - y_P) = 0 \qquad (5.109)$$

y la ec. 5.105 es

$$A_Q (p_R - p_Q) g_Q + C_Q (q_R - q_Q) + E_Q (y_R - y_Q) = 0 \qquad (5.110)$$

por último, la ec. 5.106, sería apoyandose en P (podría ser en Q)

$$\phi_r - \phi_P = p_P (x_R - x_P) + q_P (y_R - y_P)$$

se considera una mejor aproximación a

$$\phi_R - \phi_P = \frac{(p_P + p_Q)}{2} (x_R - x_P) + \frac{(q_P + q_Q)}{2} (y_R - y_P) \qquad (5.111)$$

El procedimiento de cálculo consiste en

a) Calcular f_P y g_Q según 5.101 y 5.102. Esto es

$$f_P = \frac{-B_P + \sqrt{B^2_P - 4A_P C_P}}{-2A_P}$$

$$g_Q = \frac{-B_Q + \sqrt{B^2_Q - 4A_Q C_Q}}{-2A_Q}$$

b) Encontrar x_R y y_R al resolver 5.107 y 5.108

c) Definir p_R y q_R de la solución del sistema de ecuaciones formado por 5.109 y 5.110.

d) Encontrar ϕ_R de la ec. 5.111

Ejemplo 5.10

Escribir las ecuaciones necesarias para resolver la ec. diferencial de onda

$$\frac{\partial^2 h}{\partial t^2} = g_1 d \frac{\partial^2 h}{\partial x^2} \quad ,$$

donde g_1 es la aceleración de la gravedad, d la profunidad media. En ella aparece implícita la ec. $u_t + u\, u_x = -gh_x$ donde u es velocidad y los índices señalan respecto a cual variable aparece derivada de la función. Se consideran como condiciones de frontera a u=0 en x=L y h= asen wt en x=0. Las condiciones iniciales son h(x,0) = m(x). En la fig. 5.9 se muestra el problema de interés. Se considera un ancho unitario

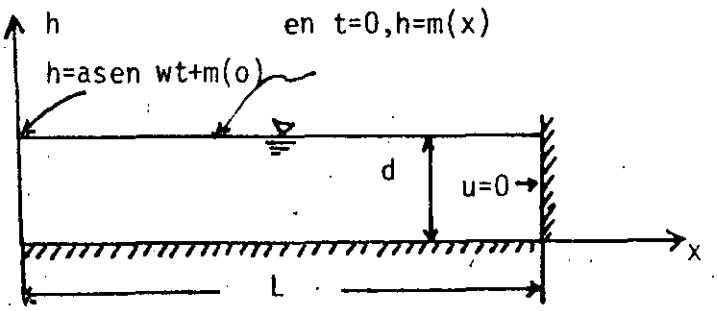


Fig. 5.9

Se considera que para la frontera izquierda con u=0, u_t=0 y $u_x \neq 0$ por lo que $u_t + u\, u_x = -g_1 hx$ implica que $0 + 0\, u_x = -g_1 hx$, o sea $h_x = 0$.

Al comparar la ec. de onda con respecto a la 5.90, se tiene (se toma t como "y" y ϕ como h)

$A = g_1 d = c$
$B = 0$
$C = -1$
$E = 0$

así $B^2-4AC = 0 -4(gd)(-1) = 4gd$

como $4gd>0$ se trata de una ec. hiporbólica y por ello se puede aplicar el método de las características.

así

$$\frac{dt}{dx} = \frac{0 \pm 0^2+ 4dg}{-2gd_1} = \pm \frac{1}{\sqrt{gd_1}} = \pm \frac{1}{C}$$

$$f = \frac{1}{C}$$

$$g = -\frac{1}{C}$$

como c es constante, las ecs. 5.107 y 5.108 resultan ser

$$C(t_R - t_P) = x_R-x_P \qquad (5.112)$$

$$-C(t_R - t_Q) = x_R-x_Q \qquad (5.113)$$

por lo tanto, al sumar

$$\boxed{x_R = \frac{C(t_Q-t_P) + x_P+x_Q}{2}} \qquad (5.114)$$

conocida x_R, de 5.113

$$\boxed{t_R = \frac{x_Q-x_R}{C} + t_Q} \qquad (5.115)$$

Las ecs. 5.109 y 5.110 para los valores de A, B, C y E definidos antes quedan

$$C^2(p_R-p_P) - (q_R-q_P) = 0 \qquad (5.116)$$

$$-C^2(p_R-p_Q) - (q_R-q_Q) = 0 \qquad (5.117)$$

así al restar 5.116 a 5.115

$$p_R = \frac{q_Q-q_P + C^2 (p_P+q_Q)}{2C^2} \qquad (5.118)$$

una vez valuda p_R, entonces de 5.116

$$\boxed{q_R = q_Q - C^2(p_R - p_Q)}$$

Así entonces se tienen definidos los valores de interés para los puntos interiores.

Para la frontera, se tiene

a) frontera izquierda (en x=0)

$$q = \frac{\partial h}{\partial t} = a\ w \cos wt \qquad (5.120)$$

$$p = \frac{\partial h}{\partial x} = 0 \qquad (5.121)$$

por lo tanto

$$\boxed{x_R = 0} \qquad (5.122)$$

y como a el llega la característica negativa (fig 5.8), al considerar x_R=0 en 5.113, se obtiene

$$\boxed{t_R = \frac{x_Q}{C} + t_Q} \qquad (5.123)$$

por 5.120 y 5.121

$$\boxed{q_R = a w \cos wt_R} \qquad (5.124)$$

$$\boxed{p_R = C} \qquad (5.125)$$

b) frontera derecha (en x=L) como $\frac{\partial h}{\partial x} = 0$ se tiene

$$\boxed{x_R = L} \qquad (5.126)$$

al llegar la característica positiva (fig. 5.8), al considerar x_R=L en 5.112

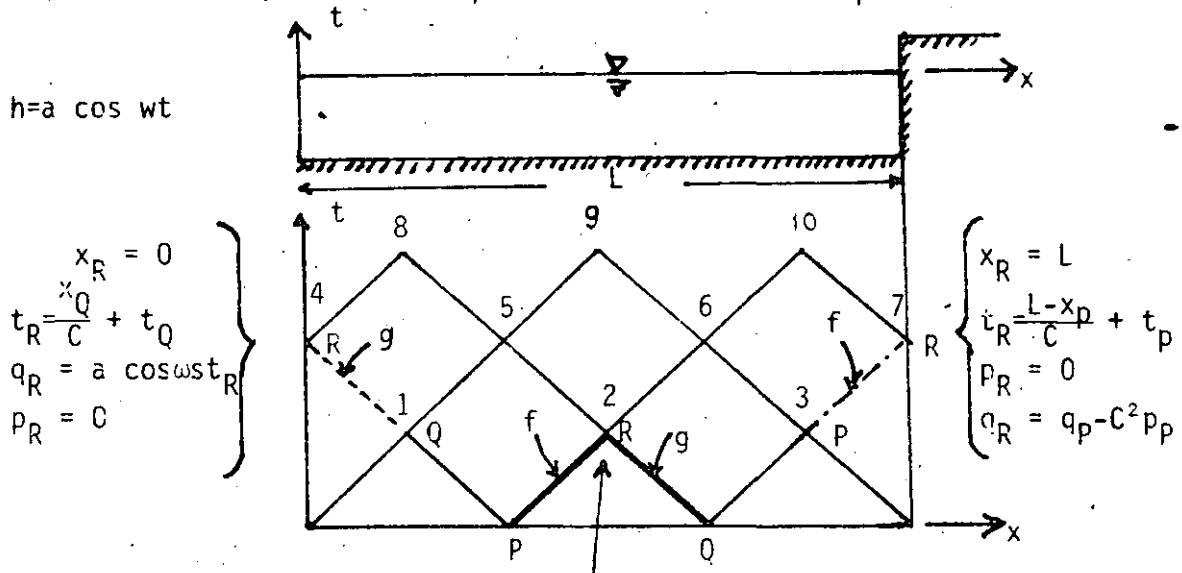
$$\boxed{t_R = \frac{L - x_p}{C} + t_p} \qquad (5.127)$$

como $\frac{\partial h}{\partial x} = 0$, $\boxed{p_R = 0}$ (5.128)

al sustituir 5.128 en la ec. asociada a la característica positiva (la que tiene f), la ec. 5.116 se tiene

$$\boxed{q_R = q_P - C^2 p_P} \qquad (5.129)$$

En la fig. 5.10 se resumen las ecuaciones encontradas. Nótese que en este caso, se pueden usar los índices i-1 para q, i para R e i+1 para Q; y escribir las ecuaciones de una forma sencilla de manejar en un programa. Todo ello porque las características siempre tienen la misma pendiente.



puntos interiores
frontera izquierda
frontera derecha

Para cualquier punto R (interior o frontera)

$$h_R = h_P + \tfrac{1}{2}(p_R + p_P)(x_R - x_P) + \tfrac{1}{2}(q_R + q_P)(t_R - t_P)$$

$$x_R = \frac{C(t_Q - t_P) + x_P + x_Q}{2}$$

$$t_R = \frac{x_Q - x_R}{C} + t_Q$$

$$p_R = \frac{q_Q - q_P + C^2(p_P + p_Q)}{2C^2}$$

$$q_R = q_Q - C^2(p_R - p_Q)$$

Fig. 5.10

El método consiste en encontrar h_R en el plano x,t operandose de este modo

a) Calcular x_R, t_R, p_R q_R y h_R para los puntos interiores (con ello se definen h_R en los puntos 1, 2 y 3 de la fig 5.10).

b) Con la condición de frontera izquierda, valuar t_R, q_R h_R (así se define h_R en 4)

c) Repetir el paso a. para definir h_R en 5 y 6

d) Usar las ecs. de frontera derecha para encontrar t_R, q_R y h_R (así se define h_R en 7)

Este procedimiento se continua hasta donde se desee.

5.3.2 Método de las características para dos ecuaciones diferenciales parciales.

El método de las características corresponde a una técnica donde el problema de resolver dos ecuaciones diferenciales parciales simultáneas es reemplazado por otro donde se resuelven cuatro ecuaciones diferenciales ordinarias. Ello implica una situación de continuidad y de definición para todas las derivadas.

Sean las ecuaciones fundamentales de la hidráulica en una dimensión

$$\frac{\partial D}{\partial t} + v \frac{\partial D}{\partial x} + D \frac{\partial v}{\partial x} = 0 \qquad (5.130)$$

$$\frac{\partial v}{\partial t} + v \frac{\partial v}{\partial x} + g \frac{\partial D}{\partial x} = g(s_o - s_f) \qquad (5.131)$$

multiplicando por g la ec. 5.130 y llamando $C^2 = gD$ (celeridad de aguas profundas) se tiene

$$\frac{\partial gD}{\partial t} + v \frac{\partial gD}{\partial x} + gD \frac{\partial v}{\partial x} = 0$$

o sea

$$\frac{\partial c^2}{\partial t} + v\frac{\partial c^2}{\partial x} + c^2\frac{\partial v}{\partial x} = 0$$

al derivar

$$2c\frac{\partial c}{\partial t} + v2c\frac{\partial c}{\partial x} + c^2\frac{\partial v}{\partial x} = 0$$

al dividir entre c e introducir el 2 en las derivadas

$$\frac{\partial 2c}{\partial t} + v\frac{\partial 2c}{\partial x} + c\frac{\partial v}{\partial x} = 0 \qquad (5.132)$$

la ec. 5.131 se puede escribir como

$$\frac{\partial v}{\partial t} + v\frac{\partial v}{\partial x} + c\frac{\partial 2c}{\partial x} = g(s_o - s_f) \qquad (5.133)$$

Al sumar 5.132 y 5.133

$$\frac{\partial 2c}{\partial t} + \frac{\partial v}{\partial t} + v\frac{\partial 2c}{\partial x} + v\frac{\partial v}{\partial x} + c\frac{\partial v}{\partial x} + c\frac{\partial 2c}{\partial x} = 0 + g(s_o - s_f)$$

(ec. 5.132)

lo anterior se escribe tambien como

$$\underset{1}{\frac{\partial}{\partial t}(v+2c)} + v\underset{2}{\left(\frac{\partial v}{\partial x} + \frac{\partial 2c}{\partial x}\right)} + c\underset{3}{\left(\frac{\partial v}{\partial x} + \frac{\partial 2c}{\partial x}\right)} = g(s_o - s_f)$$

o bien

$$\frac{\partial}{\partial t}(v+2c) + (v+c)\frac{\partial}{\partial x}(v+2c) = g(s_o - s_f) \qquad (5.134)$$

de manera similar al restar 5.133 a la 5.132 se llega a

$$\frac{\partial}{\partial t}(v-2c) + (v-c)\frac{\partial}{\partial x}(v-2c) = g(s_o - s_f) \qquad (5.135)$$

Por otra parte, la derivada total respecto al tiempo de cualquier cantidad q(x,t) es

$$\frac{dq}{dt} = \frac{\partial q}{\partial x}\frac{dx}{dt} + \frac{\partial q}{\partial t} \qquad (5.136)$$

al comparar 5.134 con 5.136, se tiene

$$q = v + 2c$$

$$\frac{dx}{dt} = v + c \qquad (5.137)$$

$$\frac{dq}{dt} = g(s_o - s_f)$$

y de esta última

$$d(v+2c) = g(s_o - s_f)\, dt \qquad (5.138)$$

y al comparar 5.135 con 5.136, se observa que

$$d(v-2c) = g(s_o - s_f)\, dt \qquad (5.139)$$

siempre que

$$\frac{dx}{dt} = v-c \qquad (5.140)$$

Las direcciones en el plano x,t definidas por 5.37 y 5.140 son llamadas DIRECCIONES CARACTERISTICAS y las ecs. 5.138 y 5.139 se conocen como INVARIANTES DE RIEMANN. Las ecs. 5.137 y 5.140 son cuatro ecuaciones diferenciales ordinarias que reemplazan a las dos ec. parciales 5.130 y 5.131.

Dos casos de interés por analizar son

a) Flujo subcrítico ($F < 1$)

El número de Froude $F = \frac{v}{c}$ es menor que 1, por ello

$\frac{v}{c} < 1$ y $v < c$ (si $c > 0$)

de acuerdo a esto;

$v + c > 0$ por lo tanto (según 5.137) $\frac{dx}{dt} > 0$

$v - c < 0$ por lo tanto (según 5.139) $\frac{dx}{dt} < 0$

y la pendiente de las características son positiva y negativa.

b) Flujo supercrítico (F > 1)

como $F > 1$, $\frac{v}{c} > 1$ y $v > c$ (si $c > 0$)

así

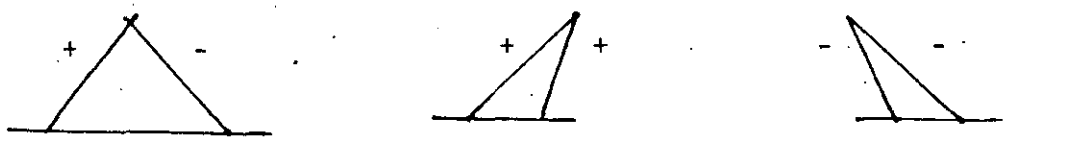
$v + c > 0$ por lo tanto (según 5.137) $\frac{dx}{dt} > 0$

$v - c < 0$ por lo tanto (según 5.139) $\frac{dx}{dt} > 0$

y la pendiente de las características son del mismo signo.

En la fig 5.11 se ilustra lo anterior.

Flujo subcrítico Flujo supercrítico (hay 2 casos posibles)



cuando $v > 0$ cuando $v < 0$

Fig. 5.11 Características según el tipo de flujo.

Ejemplo 5.11

Expresar las ecuaciones de aproximación para resolver mediante el método de características las ecs. 5.130 y 5.131. Considere flujo subcrítico y que las condiciones de frontera son en x=0 el gasto Q es conocido en todo tiempo t y que en x=L el tirante D es constante. En x=0 se conoce una curva gastos contra tirantes.

Las ecuaciones de interés son (ecs. 5.137 y 5.140)

$$\frac{dx}{dt} = v + c \rightarrow d(v+2c) = g(x_o - s_f)\, dt$$

$$\frac{dx}{dt} = v-c \rightarrow d(v-2c) = g(s_o-s_f)\,dt$$

por lo tanto expresando 5.137 y 5.140 en diferencias

$$\frac{x_R-x_P}{t_R-t_P} = v_P + c_P \qquad (5.141)$$

$$\frac{x_R-x_Q}{t_R-t_Q} = V_Q - c_Q \qquad (5.142)$$

al resolver par x_R y t_R

$$\boxed{t_R = \frac{(v_P+c_P)tp - (v_Q-c_Q)t_Q - x_P+x_Q}{(v_P+c_P) - (v_Q-c_Q)}} \qquad (5.143)$$

conocida t_R

$$\boxed{x_R = x_P + (v_P+c_P)\,(t_R-t_P)} \qquad (5.144)$$

Ahora, al escribir 5.138 y 5.139 en diferencias

$$(v_R+2c_R) - (v_P+2c_P) = g(s_o-s_f)_P\,(t_R-t_P) \qquad (5.145)$$

$$(v_R-2c_R) - (v_Q-2c_Q) = g(s_o-s_f)_Q\,(t_R-t_Q) \qquad (5.146)$$

al despejar a v_R se encuentra

$$\boxed{v_R = \frac{g[s_o-s_f)_P(t_R-t_P)+(s_o-s_f)_Q(t_R-t_Q)] + (v_P+2c_P)+(v_Q-2c_Q)}{2}} \qquad (5.147)$$

conocida v_R

$$\boxed{c_R = \frac{(v_P+2c_P) + g(s_o-s_f)P\,(t_R-t_P) - v_R}{2}} \qquad (5.148)$$

El procedimiento anterior tiene la desventaja de que las intersecciones de las características (coordenadas del punto R) no quedan igualmente especiadas ni en x ni en t por lo que si interesan conocer algunas propiedades del flujo a distancias y tiempos regulares (fig. 5.12) se necesita hacer algunas interpolaciones lineales. Existen métodos donde se fija la intersección de las características y lo que se determina son los punto P y Q, con lo cual se obtienen las propiedades del flujo en mallas regulares. En ambos procedimientos aparece un error por la interpolación y dependerá de las aproximaciones empleadas al decidir cual de las dos versiones es mejor.

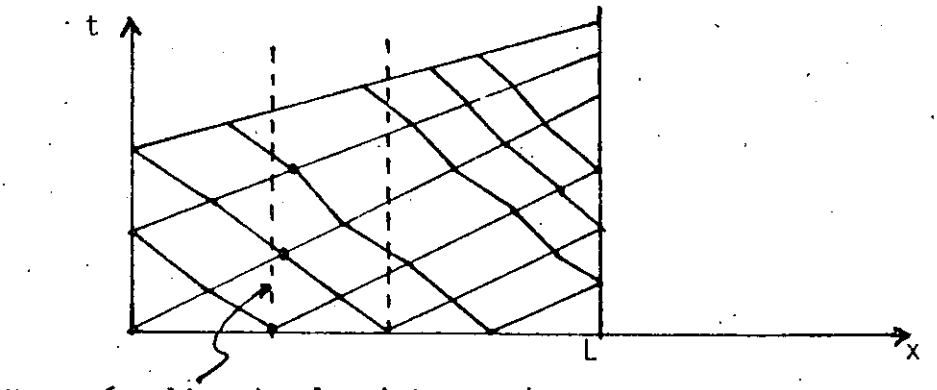


Fig. 5.12

5.4 Método de elemento finito

En el subcapítulo 4.8.1 se comentó brevemente sobre el cálculo de variaciones, y se discutió la ec. de Euler-Lagrange. Ahora se puede extender esa idea al caso de un funcional.

$$J = \iint_A F(x,y,z, \frac{\partial z}{\partial x}, \frac{\partial z}{\partial y})\, dxdy \qquad (5.157)$$

siendo su ecuación diferencial asociada

$$F_z - \frac{\partial}{\partial x} F_p - \frac{\partial}{\partial y} F_q = 0 \qquad (5.158)$$

donde:

$$p = \frac{\partial z}{\partial x} \quad y \quad q = \frac{\partial z}{\partial y} \tag{5.159}$$

Ejemplo 5.12

Encontrar la ec. de Euler del funcional

$$J = \iint \left[\left(\frac{\partial \phi}{\partial x}\right)^2 + \left(\frac{\partial \phi}{\partial y}\right)^2 \right] dxdy$$

Según 5.159

$$F = p^2 + q^2$$

$$\frac{\partial F}{\partial \phi} = 0 \quad ; \quad F_p = 2p \quad ; \quad F_q = 2q$$

así

$$F_\phi - \frac{\partial}{\partial x} F_p - \frac{\partial}{\partial y} F_q = 0 - \frac{\partial}{\partial x} 2p - \frac{\partial}{\partial y} 2q = 0$$

$$= -2\frac{\partial}{\partial x}\left(\frac{\partial \phi}{\partial x}\right) - 2\frac{\partial}{\partial y}\left(\frac{\partial \phi}{\partial y}\right) = 0$$

$$\therefore \boxed{\frac{\partial^2 \phi}{\partial x^2} + \frac{\partial^2 \phi}{\partial y^2} = 0} \tag{5.160}$$

Por otra parte, tratando de aproximar lo mejor posible una área irregular como la mostrada en la fig. 5.13 se observa que al dividirla en triángulos se cubre el área con mayor detalle que con una serie de cuadrados.

De una manera similar al método de Ritz se hará una introducción al del elemento finito, por lo que sólo se tratará el enfoque del cálculo variacional y no el de Galerkin. Conviene aclarar que se preferiran los triángulos únicamente por la razón señalada antes y que en el método de elemento finito se pueden tratar triángulos, cuadrados, rectángulos, etc.

Solución

1. La ec. por resolver es

$$\frac{\partial^2 \phi}{\partial x^2} + \frac{\partial^2 \phi}{\partial y^2} = 0 \qquad (5.165)$$

con las condiciones de frontera dadas por los valores conocidos de ϕ.

2. Se considera la fig. 5.14

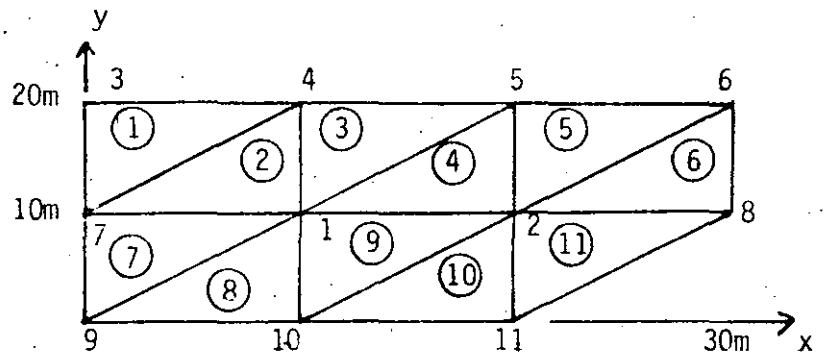


Fig. 5.14

3. Según el ejemplo 5.12, el funcional de la ec. diferencial es

$$X = \iint_A \left[\left(\frac{\partial \phi}{\partial x}\right)^2 + \left(\frac{\partial \phi}{\partial y}\right)^2 \right] dx\,dy \qquad (5.166)$$

4. Para la fig. 5.15

$$\phi_i = P_\ell + Q_\ell x_i + R y_i \qquad (5.167a)$$
$$\phi_j = P_\ell + Q_\ell x_j + R y_j \qquad (5.167b)$$
$$\phi_R = P_\ell + Q_\ell x_R + R y_R \qquad (5.167c)$$

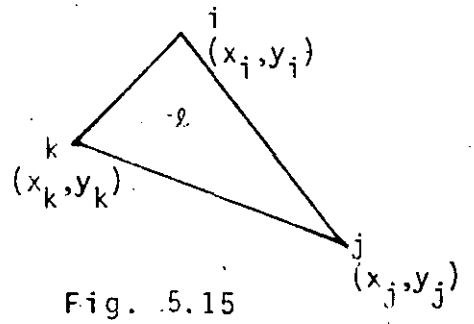


Fig. 5.15

5. Al sustituir la ec. 5.162 en 5.166

$$X = \int\int_A [Q^2 + R^2]\, dydx$$

$$X = \sum_{\ell=1}^{N} (Q_\ell^2 + R_\ell^2)\, A_\ell \qquad (5.168)$$

siendo N el número de triángulos y A_ℓ el área del triángulo ℓ.

6. Para minimizar 5.168

$$\frac{\partial X}{\partial \phi_v} = \frac{\partial}{\partial \phi_v} \sum_{\ell}^{N} (Q_\ell^2 + R_\ell^2)A$$

$$= 2 \sum_{\ell=1}^{N} \left(Q_\ell \frac{\partial Q_\ell}{\partial \phi_v} + R_\ell \frac{\partial R_\ell}{\partial \phi_v}\right)A_\ell$$

al igualar a cero

$$\frac{\partial X}{\partial \phi_v} = \sum_{\ell=1}^{N} \left(Q_\ell \frac{\partial Q_\ell}{\partial \phi_v} + R_\ell \frac{\partial R_\ell}{\partial \phi_v}\right)A_\ell \qquad (5.169)$$

lo anterior se puede escribir

$$\frac{\partial X}{\partial \phi_v} = \sum_{\ell=1}^{N} \{c(\ell,v,i)\phi_i + c(\ell,v,j)\phi_v + c(\ell,v,k)\phi_R\} = 0 \qquad (5.170)$$

Para el caso particular de la fig. 5.14 se tiene

TABLA 5.1

TRIANGULO	VERTICES i j R	APARECE V = 1	APARECE V = 2
1	3,4,7		
2	4,1,7	✓	
3	4,5,1	✓	
4	5,2,1	✓	✓
5	5,6,2		✓
6	6,8,2		✓
7	7,1,9	✓	
8	1,10,9	✓	
9	1,2,10	✓	✓
10	2,11,10		✓
11	2,8,11		✓

si $\phi_v = \phi_1$ sólo se considera en 5.170 a los triángulos donde aparece el vértice 1, o sea los triángulos 2,3, 4,7,8 y 9.

Si $\phi_v = \phi_2$, se toman en cuenta en la suma 5.170 a los triángulos donde aparece precisamente 2 o sea en 4,5, 6,9,10 y 11.

De este modo las sumas son

$$\frac{\partial X}{\partial \phi_1} = \{c(2,1,4)\phi_4 + c(2,1,1)\phi_1 + c(2,1,7)\phi_7\} + \{c(3,1,4)\phi_4 + c(3,1,5)\phi_5 + c(3,1,1)\phi_1\} + \{c(4,1,5)\phi_5 + c(4,1,2)\phi_2 + c(4,1,1)\phi_1\} + \{c(7,1,7)\phi_7 + c(7,1,1)\phi_1 + c(7,1,9)\phi_9\} + \{c(8,1,1)\phi_1 + c(8,1,10)\phi_{10} + c(8,1,9)\phi_9\} + \{c(9,1,1)\phi_1 + c(9,1,2)\phi_2 + c(9,1,10)\phi_{10} = 0 \qquad (5.171)$$

$$\frac{\partial X}{\partial \phi_2} = \{c(4,2,5)\phi_5 + c(4,2,2)\phi_2 + c(4,2,1)\phi_1\} + \{c(5,2,5)\phi_5 + c(5,2,6)\phi_6 + c(5,2,2)\phi_2\} + \{c(6,2,6)\phi_6 + c(6,2,8)\phi_8 + c(6,2,2)\phi_2\} + \{c(9,2,1)\phi_1 + c(9,2,2)\phi_2 + c(9,2,10)\phi_{10}\} + c(10,2,2)\phi_2 + c(10,2,11)\phi_{11} + c(10,2,10)\phi_{10}\} + \{c(11,2,2)\phi_2 + c(11,2,8)\phi_8 + c(11,2,11)\phi_{11}\} = 0 \qquad (5.172)$$

Como en las ecuaciones anterioes sólo son incógnitas ϕ_1 y ϕ_2 al factorizar se tiene:

$$a_{11}\ \phi_1 + a_{12}\ \phi_2 = b_1 \qquad (5.171')$$

$$a_{21}\ \phi_1 + a_{22}\ \phi_2 = b_2 \qquad (5.172')$$

De 5.171

·Los términos que multiplican a ϕ_1

$$a_{11} = c(2,1,1)+c(3,1,1)+c(4,1,1)+c(7,1,1)+c(8,1,1)+c(9,1,1) \qquad (5.173)$$

·Los coeficientes de ϕ_2 son

$$a_{12} = c(4,1,2)+c(9,1,2) \qquad (5.174)$$

·Los términos independientes

$$b_1 = -\{c(2,1,4)-c(3,1,4)\}\phi_4 - \{c(3,1,5)+c(4,1,5)\}\phi_5 + -\{c(2,1,7)+c(7,1,7)\}\phi_7 - \{c(7,1,9)+c(8,1,9)\}\phi_9 + -\{c(c,1,10)+c(9,1,10)\}\phi_{10} \quad (5.175)$$

A partir de la ec. 5.172 se tiene

·Coeficientes de ϕ_1

$$a_{21} = c(4,2,1) + c(9,2,1) \quad (5.176)$$

·Coeficientes de ϕ_2

$$a_{22} = c(4,2,2)+c(5,2,2)+c(6,2,2)+c(9,2,2)+c(10,2,2)+ c(11,2,2) \quad (5.177)$$

·Y los productos donde no aparecen ni ϕ_1 y ϕ_2

$$b_2 = -\{c(4,2,5)+c(5,2,5)\}\phi_5 - \{c(5,2,6)+c(6,2,6)\}\phi_6 + -\{c(6,2,8)+c(11,2,8)\phi_8 - c(9,2,10)+c(10,2,10)\phi_{10} + -\{c(10,2,11)+c(11,2,11)\}\phi_{11} \quad (5.178)$$

Se aprecia que conocidos los coeficientes c, se calculan a_{11}, a_{12}, b_1, a_{21}, a_{22} y b_2 y se resuelve el sistema de ecuaciones lineales formado por 5.167' y 6.168' y así se definen ϕ_1 y ϕ_2 concluyendo el ejercicio.

Sin embargo, surge la pregunta ¿Cómo calcular los coeficientes c?

Cálculo de los coeficientes c(ℓ,v,i)

Como para cada triángulo ℓ se conocen las coordenadas de los vértices, resulta que a través de las e-

cuaciones 5.167 se pueden obtener P_ℓ, Q_ℓ y R_ℓ. En efecto, si su solución se plantea mediante la regla de Cramer.

Determinante de la matriz de coeficientes

$$\Delta = \begin{vmatrix} 1 & x_i & y_i \\ 1 & x_j & y_j \\ 1 & x_k & y_k \end{vmatrix}$$

en geometría analítica se demuestra que

$$\Delta = 2A_\ell \qquad (5.179)$$

siendo A_ℓ el área del triángulo ℓ

Determinante asociado a la incógnita P_ℓ

$$\Delta_p \begin{vmatrix} \phi_i & x_i & y_i \\ \phi_j & x_j & y_j \\ \phi_k & x_k & y_k \end{vmatrix} = \phi_i(x_j y_k - x_k y_j) + \phi_j(x_k y_i - x_i y_k) + \phi_k(x_i y_i - x_j y_j) \qquad (5.177)$$

Determinante asociado a la incógnita Q_ℓ

$$\Delta_a = \begin{vmatrix} 1 & \phi_i & y_i \\ 1 & \phi_j & y_i \\ 1 & \phi_k & y_k \end{vmatrix} = \phi_i(y_j - y_k) + \phi_j(y_k - y_i) + \phi_k(y_i - y_j) \qquad (5.178)$$

Determinante asociado a la incógnita R_ℓ

$$\Delta_R = \begin{vmatrix} 1 & x_i & \phi_i \\ 1 & x_j & \phi_j \\ 1 & x_k & \phi_k \end{vmatrix} = \phi_i(x_k - x_j) + \phi_j(x_i - x_k) + \phi_k(x_j - x_i) \qquad (5.179)$$

Por lo que resulta ser

$$P\ell = \frac{\Delta p}{2A_\ell} \qquad (5.180)$$

$$Q_\ell = \frac{\Delta_Q}{2A_\ell} \qquad (5.181)$$

$$R_\ell = \frac{\Delta_R}{2A_\ell} \qquad (5.182)$$

Al sustituir 5.181 en 5.169

$$\frac{\partial X}{\partial \phi_v} = \sum_{\ell=1}^{N} \left\{ \frac{\Delta_Q}{(2A_\ell)^2} \frac{\partial \Delta_Q}{\partial \phi_v} + \frac{\Delta}{(2A_\ell)^2} \frac{\partial \Delta_R}{\partial \phi_v} \right\} A_\ell = 0$$

$$= \sum_{\ell=1}^{N} \left\{ \Delta_Q \frac{\partial \Delta_Q}{\partial \phi_v} + \Delta_R \frac{\partial \Delta_R}{\partial \phi_v} \right\} \frac{A_\ell}{4} = 0 \qquad (5.183)$$

Al sustituir 5.178 y 5.179 en 5.183

$$\frac{\partial X}{\partial \phi_v} \sum_{\ell=1}^{N} [\phi_i(y_j-y_k)+\phi_j(y_k-y_i)+\phi_k(y_i-y_j)] [(y_j-y_k) \frac{\partial \phi i}{\partial \phi_v} +$$

$$+ (y_k-y_i) \frac{\partial \phi j}{\partial \phi_v} + (y_i-y_j) \frac{\partial \phi k}{\partial \phi_v}] \frac{A_\ell}{4} +$$

$$+[\phi_i(x_K-x_j)+\phi_j(x_j-x_k)+\phi_k(x_j-x_i)][(x_k-x_j) \frac{\partial \phi i}{\partial \phi_v} +$$

$$+(x_i-x_k) \frac{\partial \phi j}{\partial \phi_v} +(x_j-x_i) \frac{\partial \phi k}{\partial \phi_v}] \frac{A_\ell}{4} = 0$$

Al factorizar ϕ_i, ϕ_j y ϕ_k

$$\frac{\partial x}{\partial \phi_v} = \sum_{\ell=1}^{N} \{[(y_j-y_k)(y_j-y_k)+(x_k-x_j)(x_k-x_j)] \frac{\partial \phi i}{\partial \phi_v} +$$

$$+[(y_j-y_k)(y_k-y_j)+(x_k-x_j)(x_j-x_k)] \frac{\partial \phi j}{\partial \phi_v} +$$

$$+[(y_j-y_k)(y_i-y_j)(x_k-x_j)(x_j-x_i)] \frac{\partial \phi k}{\partial \phi_v}]\} \frac{A_\ell}{4} \phi_i +$$

$$+[(y_k-y_i)(y_j-y_k)+(x_i-x_k)(x_k-x_j)] \frac{\partial \phi i}{\partial \phi_v} +$$

$$+[(y_k-y_i)(y_k-y_i)+(x_i-x_k)(x_i-x_k)]\frac{\partial\phi j}{\partial\phi_v} +$$

$$+[(y_k-y_i)(y_i-y_j)+(x_i-x_k(x_j-x_i)\frac{\partial\phi k}{\partial\phi_v}]\ \frac{A_\ell}{4}\}\phi j +$$

$$+[(y_i-y_j)(y_j-y_k)+(x_j-x_i)(x_k-x_j)]\frac{\partial\phi i}{\partial\phi_v} +$$

$$+[(y_i-y_j)(y_k-y_i)+(x_j-x_i)(x_i-x_k)]\frac{\partial\phi j}{\partial\phi_v} +$$

$$+[(y_i-y_j)(y_i-y_j)+(x_j-x_i)(x_j-x_i)]\frac{\partial\phi k}{\partial\phi_v}\ \frac{A_\ell}{4}\ \phi k = 0$$

Obsérvese que $\frac{\partial\phi i}{\partial\phi_v}$, $\frac{\partial\phi j}{\partial\phi_v}$ y $\frac{\partial\phi k}{\partial\phi_v}$ son iguales a cero o a uno según sea v.

Al comparar 5.184 con 5.170 se tiene

a) si v = i

$$c(\ell,i,i) = [(y_j-y_k)(y_j-y_k)+(x_k-x_j)(x_k-x_j)]\frac{A_\ell}{4} \quad (6.185)$$

$$c(\ell,i,j) = [(y_k-y_i)(y_j-y_k)+(x_i-x_k)(x_k-x_j)]\frac{A_\ell}{4} \quad (6.186)$$

$$c(\ell,i,k) = [(y_i-y_j)(y_j-y_k)+(x_j-x_i)(x_k-x_j)]\frac{A_\ell}{4} \quad (6.187)$$

Nota: En 5.184 se consideró: $\frac{\partial\phi i}{\partial\phi_v} = 1$; $\frac{\partial\phi j}{\partial\phi_v} = \frac{\partial\phi k}{\partial\phi_v} = 0$

b) si v = j

$$c(\ell,j,i) = [(y_j-y_k)(y_k-y_i)+(x_k-x_j)(x_i-x_k)]\frac{A_\ell}{4} \quad (6.188)$$

$$c(\ell,j,j) = [(y_k-y_i)(y_k-y_i)+(x_i-x_k)(x_i-x_k)]\frac{A_\ell}{4} \quad (6.189)$$

$$c(\ell,j,k) = [(y_i-y_j)(y_k-y_i)+(x_j-x_i)(x_i-x_k)]\frac{A_\ell}{4} \quad (6.190)$$

Nota: En 5.184 se consideró: $\frac{\partial\phi j}{\partial\phi_v} = 1$; $\frac{\partial\phi i}{\partial\phi_v} = \frac{\partial\phi k}{\partial\phi_v} = 0$

c) si v=k

$$c(\ell,k,i) = [(y_j-y_k)(y_i-y_j)+(x_k-x_j)(x_j-x_i)]\frac{A_\ell}{4} \quad (5.191)$$

$$c(\ell,k,j) = [(y_k-y_i)(y_i-y_j)+(x_i-x_k)(x_j-x_i)]\frac{A_\ell}{4} \quad (6.192)$$

$$c(\ell,k,k) = [(y_i-y_j)(y_i-y_j)+(x_j-x_i)(x_j-x_i)]\frac{A_\ell}{4} \quad (6.193)$$

Nota:
En 5.184 se consideró: $\frac{\partial \phi k}{\partial \phi_v} = 1$ y $\frac{\partial \phi j}{\partial \phi_v} = \frac{\partial \phi i}{\partial \phi_v} = 0$

Los coeficientes anteriores dependen solamente de las coordenadas de los vértices de los triángulos, porque conocidas estas (paso 2) ya se pueden valuar ellos.

Para el caso particular de interés se calcularan algunos de los coeficientes c.

a) $c(2,1,1)$ como corresponde al triángulo 2 y ahí i=4, j=1 y k=7, se usará la ec. 5.189.

$$c(2,1,1) = [(y_7-y_1)(y_7-y_1)+(x_4-x_7)(x_4-x_7)]\frac{A_2}{4}$$

de la fig. 5.14.

$$c(2,1,1) = [(10-10)(10-10)+(10-0)(10-0)]\frac{A_2}{4}$$

$$c(2,1,1) = 100\,\frac{A_2}{4}$$

b) $c(5,2,6)$ como corresponde al triángulo 5 y de la tabla 5.1 se ve que i=5, j=6 y k=2; se usará la ec. 5.192

$$c(5,2,6) = [(y_2-y_5)(y_5-y_6)+(x_5-x_2)(x_6-x_5)]\frac{A_5}{4}$$

de la fig. 5.14.

$$c(5,2,6) = [(10-20)(20-20)+(20-20)(30-20)]\frac{A_5}{4}$$

$$c(5,2,6) = 0$$

A efecto de disminuir el número de cálculos se toma en cuenta que:

según 5.186 y 5.188, $c(\ell,i,j) = c(\ell,j,i)$
de 5.187 y 5.191, $c(\ell,i,k) = c(\ell,k,i)$
de 5.190 y 5.192, $c(\ell,j,k) = c(\ell,k,j)$

Como se observa en las ecs. anteriores el método de elemento finito implica una cantidad considerable de cálculos pero ellos son muy simples y fáciles de incluir en un programa de cómputo.

OSCAR FUENTES MARILES
NOVIEMBRE 1984

METODOS NUMERICOS EN HIDRAULICA

1.- Burden R. L., Faires J. D. and Reynolds A.C., Numerical Analysis, U.S.A.; Prindle, Weber and Schmidt, 1978, 579 pp.

2.- Smith G. D., Numerical Solution of Partial Differential Equations: Finite Difference Methods, Second edition, Great Britain, Oxford University Press, 1978, Oxford Applied Mathematics and Computing Science Series, 304 pp.

3.- Sánchez B. J. L., Doce Algorítmos para Resolver Problemas de Hidráulica, México, Instituto de Ingeniería, UNAM, 1978, Series de Instituto de Ingeniería UNAM, No. 412, 376 pp.

4.- Desai, C.S., Elementary Finite Element Method, U.S.A., Prentice-Hall, 1979, Civil Engineering and Engineering Mechanics Series, 434 pp.

5.- Arden B.W. and Astill K. N., Numerical Algorithms: Origins and Applications, U.S.A., Addison-Wesley Publishing Company, 1970, 308 pp.

6.- Hornbeck R. W., Numerical Methods, U.S.A., Quantum Publishers, 1975 QPI Series, 310 pp.

7.- Dahlquist G. and Björck A., Numerical Methods, Translated by Anderson N., U.S.A., Prentice-Hall 1974, Series in Automatic Computation, 573 pp.

8.- Ralston A., Introducción al Análisis Numérico, Traducción por Cervantes, G.C.E., México, Limusa, 1978, 629 pp.

9.- Milne W.E., Numerical Solution of Differential Equations, Second Edition, U.S.A., Dover Publications, 1970, 359 pp.

"INGENIERIA MARITIMA MODULO: INGENIERIA PORTUARIA"
DEL 1º DE JULIO AL 6 DE SEPTIEMBRE.
MEXICO, D.F.

INSTALACIONES PARA EL MANEJO DE CARGA TERMINALES DE CONTENEDORES.

ING. JULIO PINDTER VEGA.
SEPTIEMBRE 1985.

TERMINAL DE CONTENEDORES.

ANTECEDENTES:

Los costos elevados en el manejo de carga general fraccionada por los puertos de los países industrializados, principalmente por los altos salarios de estibadores, aunado al gran número de movimientos (25 en promedio a nivel mundial), entre la zona de producción y de consumo, que repercuten en los precios de venta de las mercancias, propiciaron la modificación del sistema tradicional por medio de la unitarización de la carga.

La unitarización de la carga se logró con la adpción de cajas con dimensiones compatibles con los diversos modos de transporte. Aunque los contenedores se bienen usando desde mediados del siglo pasado, por medio de la utilización de cajas de madera de diferentes tamaños para la unitarización de cargas específicas que por su precio y densidad lo ameritan.

Así en 1960 se inicia la utilización de los contenedores en los E.U. las compañias , Sea Trein, Sea Land y Matson, inician el transporte de contenedores de 8' de ancho por 8' de alto y 32, 35 y 24' de largo que cumplian con las normas de transportación via F.C. y carretera. En 1968 la ISO de la ONU fija como contenedores standar los de 20 y 40' de largo de 8' de ancho y 4, 8, 8',6" y 9' 6" de alto con las dimensiones estandar se logra captar una gran cantidad de la carga general fraccionada sucep tible a unitarizarse bajo este sistema, y permiten cumplir con las normas de carreteras y ferrocarriles de la mayoria de los países.

Con el empleo de los contenedores se logró dar un gran impulso al -- transporte intermodal internacional al permitir estandarizar una "unidad de carga" compatible con los diversos modos de transporte, redundando en beneficio de distribución de comercialización de mercancias. El hecho de iniciarse el manejo de contenedores por los puertos, el sistema repercute en el interior del país ya que se requerirá establecer "centros de carga" con el equipo adecuado para la carga/descarga de los contenedores.

La alta tecnología y valor de los equipos de carga/descarga y de los barcos especializados que implica la contenedorización, proporcionan la formación de monopolios mundiales, alejando a los países en vias de desarrollo en la participación directa del sistema, beneficiando indirectamente al país que lo adopta por la reducción de costos de transporte en el proceso de distribución y comercialización de la producción y del consumo.

A continuación se muestra la evolución cronológica del transporte intermodal.

EVOLUCION CRONOLOGICA DEL TRANSPORTE INTERMODAL:

1830: En Inglaterra con la introducción de las plataformas de F.C., se inicia el transporte de carga en contenedores de madera para prestar un servicio de estación a estación.

1847: En E.U. se inicia el manejo de contenedores de madera en plataformas de F.C. dando servicio de estación a estación.

1930: Se inicia en forma incipiente el servicio de Plooyback, es decir camiones sobre plataformas de F.C. para prestar servicio puerta a puerta

1940: En E.U. se inicia el empleo de las paletas ó pallets y con ello los montacargas.

1950: Con el desarrollo del autotransporte con trailer, tuvo un gran impulso el sistema Piggyback, al permitir transportar unicamente las cajas de los trailers sobre plataformas de F.C. proporcionando de ésta manera un servicio puerta a puerta.

1957: Se transportan las primeras cajas de trailers en la cubierta y en celdas practicadas en las bodegas de barcos.

1958: Sea Train, Sea Land y Matson, inician el manejo de contenedores en barcos transformados, de 8 X 8' de sección transversal y 32; 35 y 24' de largo respectivamente. Se inicia el manejo de contenedores en patio con grúas "U" (Straddle Carrier), considerados de la. generación con 6 ruedas y transmisión de cadena y de motores hidrostaticos, con dos alturas de estiba.

1960: Se contruyen las primeras grúas especializadas para manejo de contenedores, entre las costa oeste de E.U. y Hawai, dando inicio el manejo de contenedores como sistema.

1965: Matzon , opera el primer barco especializado para contenedores con sistema Lift/on - Lift/off (Lo/lo).

1966: Se inicia el manejo de contenedores en Europa.

1967: Se inicia el manejo de contenedores en Japón.

1968: La ISO de la ONU, fija como contenedores estandard los de 20' y 40' de largo, por 8' de ancho y 4, 8, 8' 6" de alto.

La compañía Sea-Land continúa con sus contenedores de 35' - de largo.

1970: Se inicia el empleo de grúas portico de patio, sobre neumaticos y/o sobre rieles (Transteiner; Trave Lift; Rubber, Rail Gantry Crane), para almacenamiento de contenedores en patio.

1977: Se introduce la 2a. generación de Straddle Carrier, con 8 ruedas y transmisión mecánica por flecha y tres alturas de estiba.

1978: Se inicia el empleo de grúas hidráulicas con pluma telescopica con movimiento en un plano vertical para manejo de contenedores en patio.

1980: Se inicia el manejo de contenedores en México por el puerto - de Veracruz, Ver.

1981: Se Establece la Empresa Mexicana de Transporte Intermodal.

1982: Se instalan las primeras grúas portacontenedores en los puertos de Veracruz, Ver., y Lázaro Cárdenas, Mich.

1983: La Compañía Americana President Line, introduce los contenedores de 45' de largo para tráficos especificos entre E.U. y Oriente, permitiendo un aumento del 25% en la capacidad de - carga respecto al de 40', para cargas de alto valor y baja densidad. Se introduce Straddle Carrier, con 10 ruedas, transmisión mecánica por flecha y 4 alturas de estiba, que algunos denominan de 3er generación. Este tipo movera equipo para manejo de contenedores, es el que mayor modificaciones a sufrido desde su implantación.

1984: En los E.U. se inicia el agrupamiento de carga en bodegas de consolidación, para formar bloques del total de la capacidad - del contenedor, los cuales son introducidos al contenedor por medio de rieles, reduciendo notablemente el costo de consolidación de carga.

Para la implantación de la contenedorización se tendrá que tomar en cuenta:

- La reducción de mano de obra, en una terminal de contenedores respecto a una terminal convencional de carga general, que varia aproximadamente de 4 a 1 dependiendo del grado de mecanización.
- Una terminal de contenedores tiene un rendimiento en el manejo de carga del orden 5 veces, respecto a una terminal de carga general con una inversión tres veces mayor. Por lo que el costo por tonelada manejada por efecto de las inversiones realizadas y los rendimientos, equivale del orden de la mitad.

Repercución directa e indirecta en la producción y consumo.

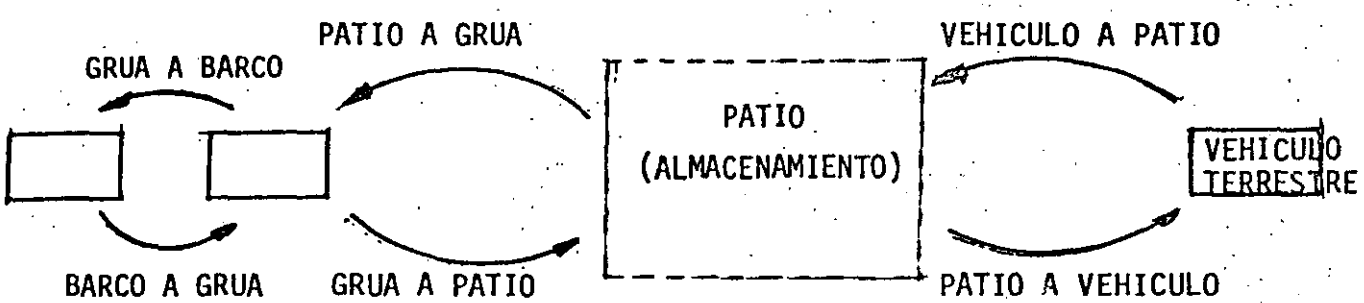
(6)

Planeación de una terminal de contenedores.

1.- Flujo de carga. Actividades de la terminal.

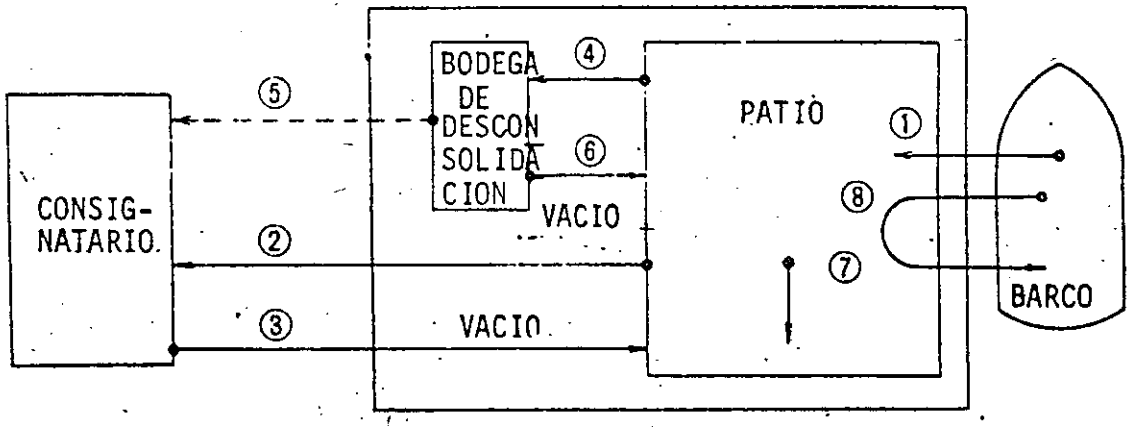
- Carga/descarga de contenedores.
- Recepción y despacho de contenedores vía transporte terrestre.
- Almacenamiento en patio.
- Consolidación y desconsolidación de contenedores.
- Mantenimiento y conservación de contenedores, vehículos y equipos de manipulación de carga.

En la siguiente figura se muestra el flujo de la carga en una terminal.



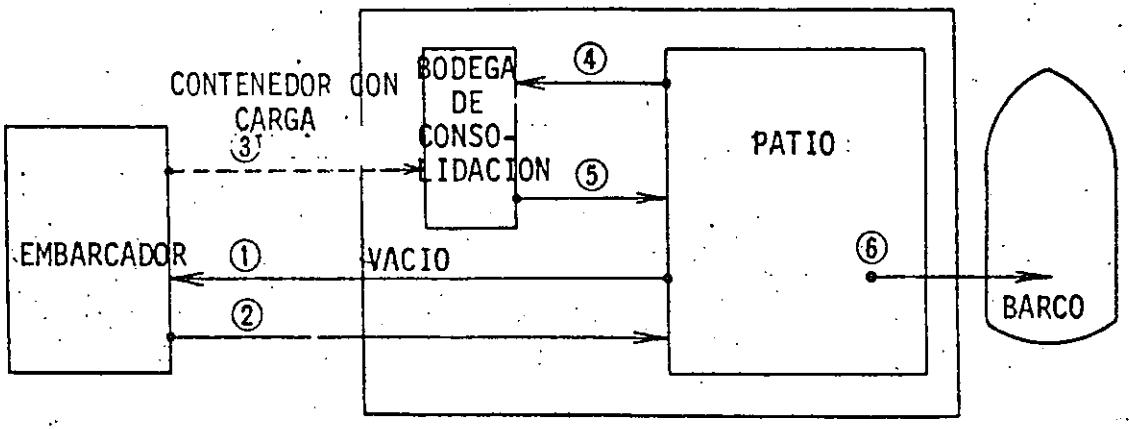
Las figuras muestran las líneas de flujo de contenedores de exportación y de importación.

No.	CONCEPTO	CONSIGNATARIO	BODEGA DE DESCONSOLIDACION	PATIO	BARCO
1	DESCARGA				
2	ENTREGA				
3	REGRESO CONTENEDOR VACIO				
4	TRANSLADO CONTENEDOR CARGADO				
5	REGRESO CONTENEDOR VACIO				
6	TRANSLADO POR CUARENTENA				
7					
8	REUBICACION DE CARGA				



FLUJO DE IMPORTACION DE CONTENEDORES

No.	CONCEPTO	EMBARCADOR	BODEGA DE CONSOLIDACION	PATIO	BARCO
1	ENVIO DE CONTENEDOR VACIO				
2	RECEPCION CON CARGA				
3	RECEPCION EN BODEGA				
4	A BODEGA DE CONSOLIDACION				
5	RECEPCION EN PATIO				
6	CARGA				



FLUJO DE CONTENEDORES DE EXPORTACION

2.- Aspectos Generales.

- La terminal se proyectará de tal manera que los bar-- cos porta-contenedores no tengan estadias prolongadas en espera de muelle.
- Que las operaciones de carga/descarga se puedan efectuar las 24 horas del día y durante todo el año.
- Disponer de amplias zonas de de almacenamiento, dotadas de acceso carretero y ferroviario.

3.- Localización.

- El volúmen previsto de tráfico determinará la longitud de atraque y la extención de los patios de almacena--- miento de contenedores.

NOTA.- Con frecuencia, la importancia de las areas de almacenamiento de contenedores, impide la utilización de los muelles convencionales de carga general, por sus dimensiones reducidas.

- Las condiciones fisicas influyen en la localización, - por lo que la zona elegida debe estar protegida de la agitación ya que el manejo de contenedores requiere una posicion estable del barco (altura máxima de la ola de 0.75 m.). La calidad del suelo es importante por las - grandes descargas producidas por los contenedores apilados y el equipo de manejo.
- Es deseable que la localización de la terminal no provoque largos trayectos del barco entre la bocana del -

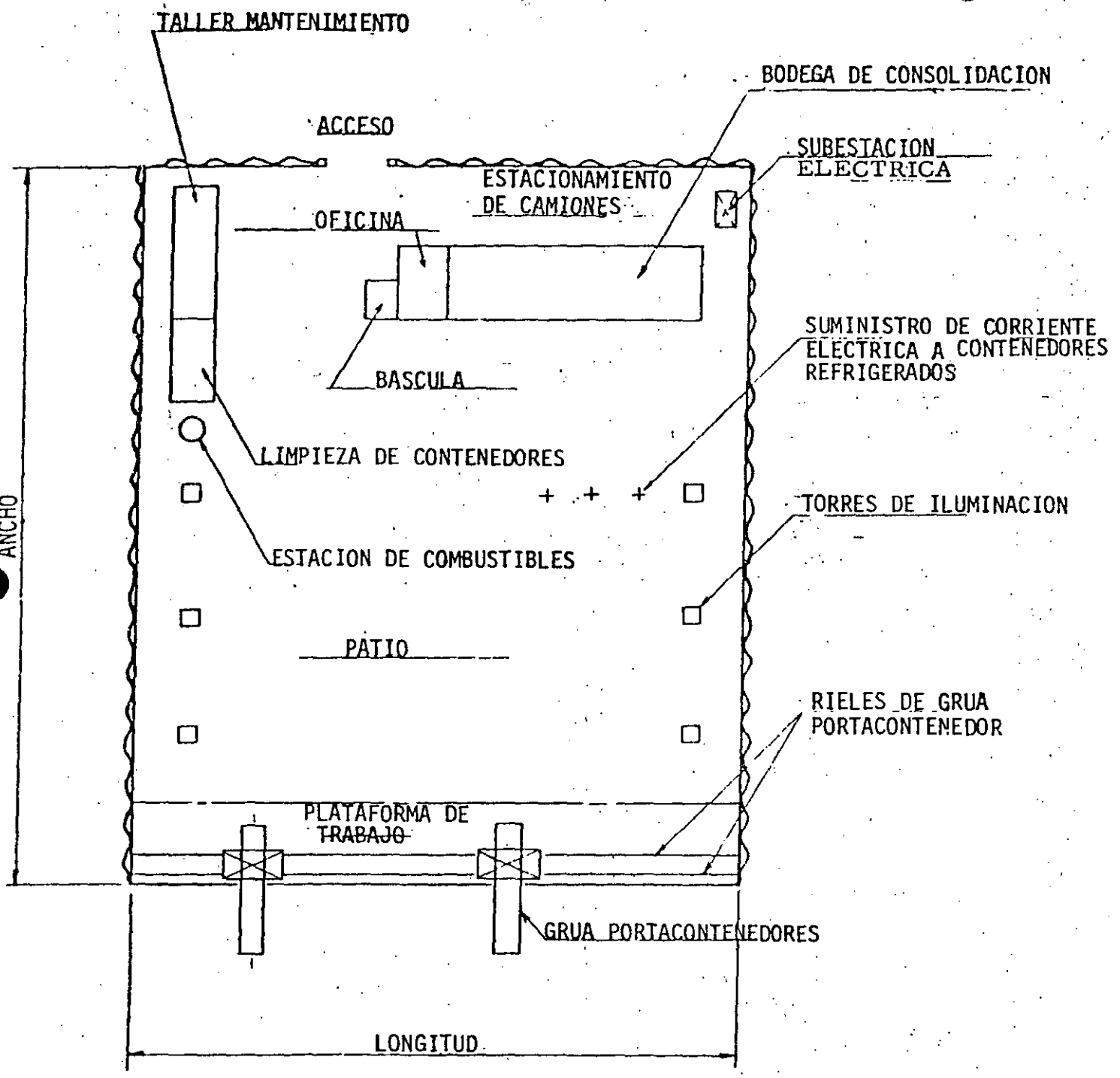
puerto y la terminal para reducir el timepo en puerto.

- Se deberá contar con reserva territorial para ampliación de patio y prever el aumento en el tamaño de los barcos.

La siguiente figura muestra una distribución general de una terminal.

4.- Muelles.

- Si la predicción del tráfico indica la posibilidad del envio de embarcaciones de la 3a. generación, deberá preverse una profundidad de 13 a 14 m.
- Los barcos de la 2a. generación requieren 11 m. de profundidad.
- El muelle deberá contar con una vía para la grúa porta-contenedores, cuyo peso fluctua entre 500 - 800 tons., y cuya altura es de hasta 80 m. con el brazo de carga elevado.
- La longitud media de un atraque varia de 250 a 300 m. para los barcos de 2a y 3a. generación. En el caso de requerirse varios tramos de atraque, estos deberán tener el mismo alineamiento para poder desplazar las grúas porta-contenedores de un tramo a otro.
- Para el empleo de barcos porta-contenedores alimentadores que comuniquen puertos pequeños con grandes terminales, es conveniente prever atraques de menores dimensiones, sin interferencia en su manejo.



DISTRIBUCION GENERAL DE UNA TERMINAL DE CONTENEDORES

- La utilización cada vez mayor a nivel mundial de barcos mixtos, es decir Lo/Lo y Ro/Ro, en donde el autotransporte juega un papel preponderante no obliga a prever rampas fijas en un extremo de la terminal ó - bien el uso de rampas flotantes moviles.

5.- Patios:

- Una de las características del sistema de transporte por contenedores es la gran extensión de terreno necesaria para almacenamiento.
- Cuando se inician las operaciones en una terminal y - hasta 20 000 TEU. se requieren del orden de 300 m. de ancho, llegando a 500 m. para un manejo de 100,000 -- teu/año por terminal.
- Cuando existe un gran movimiento de contenedores va-- cios, las experiencias en otras partes del mundo fijan a 600 m. el ancho del patio.
- Un patio de contenedores, se compone de tres partes - pricipales:

A.- Zona de preparación del plan de carga (instalaciones de control)

B.- Zona de almacenamiento de contenedores.

C.- Circulación de vehículos y equipo.

Las diversas áreas de la terminal se definiran en función de los volumenes previstos de contenedores de -- importación y explotación, con carga y vacios para -- contenedores de 20 ó 40', refrigerados o con cargas - peligrosas, áreas para los que requieren reparación y fundamentalmente el tipo de equipo para la transferencia y estiva de contenedores.

Por regla general a una mayor densidad de almacena-- miento de contenedores, se requiere una administra-- ción rigurosa y un mayor valor del equipo para la estiva a gran altura.

El conjunto de patios debe proyectarse de manera uniforme para poder modificar los limites de las diversas áreas, de acuerdo con la demanda de los flujos - de los tipos de contenedores que se manejen.

Es importante proyectar adecuadamente los patios para obtener un dren de aguas pluviales eficiente y alumbrado general que permita el trabajo nocturno con seguridad y eficiencia. Estos conceptos representan del orden del 30% del costo de los patios, y los patios tienen un costo en su totalidad de aproximadamente similar al del muelle.

Los patios deberán proyectarse a nivel por la gran economía que representa el ahorro de energia (el 2% de pendiente representa el doble de consumos de energia).

La eficiencia en las operaciones de carga/descarga y almacena-

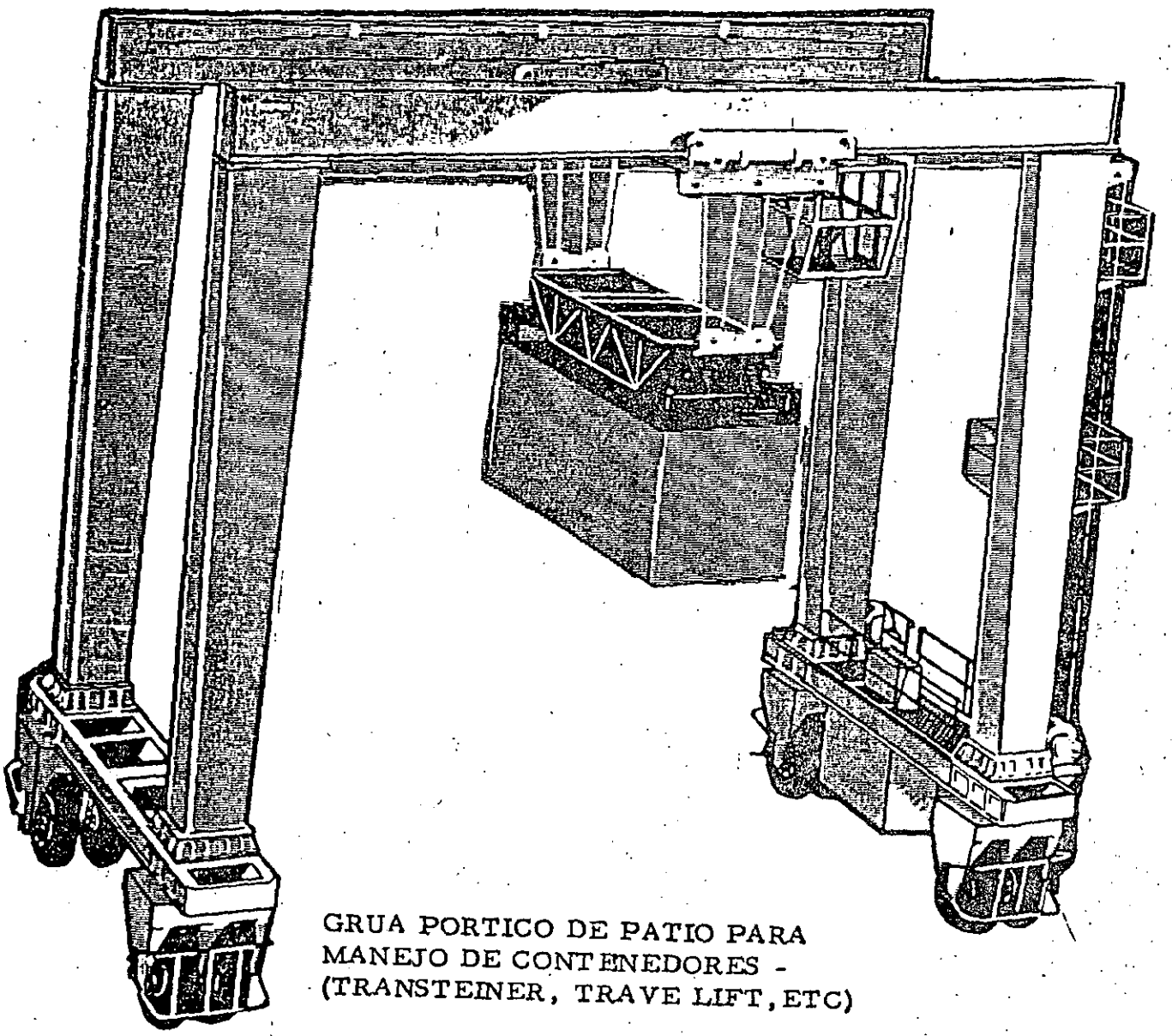
miento de contenedores en patio, deberá ser igual o mayor que la del equipo de carga/descarga en muelle para obtener su máxima eficiencia en la operación.

6.- Comunicaciones terrestres.

Dado que el ritmo del transbordo del sistema de transporte terrestre es menor que la carga/descarga de barcos, la terminal deber contar con una vialidad expedita y con estacionamientos de vehículos terrestres para evitar congestionamientos.

El dimencionamiento de la vialidad, tanto carretero como ferroviario estará en función del volúmen de carga del tráfico marítimo.

El proyecto detallará la operación ferroviaria, la cual formará por tres vias, equipadas con gruas sobre rieles que permita la carga/descarga de vagones. Las vias pueden instalarse ya sea perpendicularmente o paralelas al muelle, lo cual dependerá de la dirección de ampliación de la terminal, dado que es deseable no cortar los patios de almacenamiento con vias ferreas, generalemente se localizan éstas, al fondo de la terminal, es decir - en el extremo contrario a la dirección de ampliación de patios.

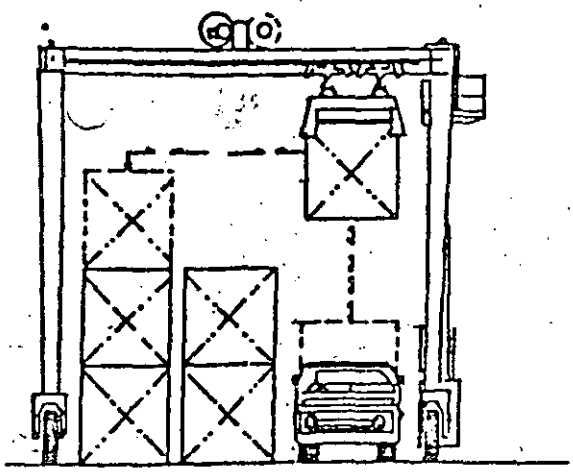


GRUA PORTICO DE PATIO PARA
MANEJO DE CONTENEDORES -
(TRANSTEINER, TRAVE LIFT, ETC)

EQUIPO PARA EL MANEJO DE CONTENEDORES

15

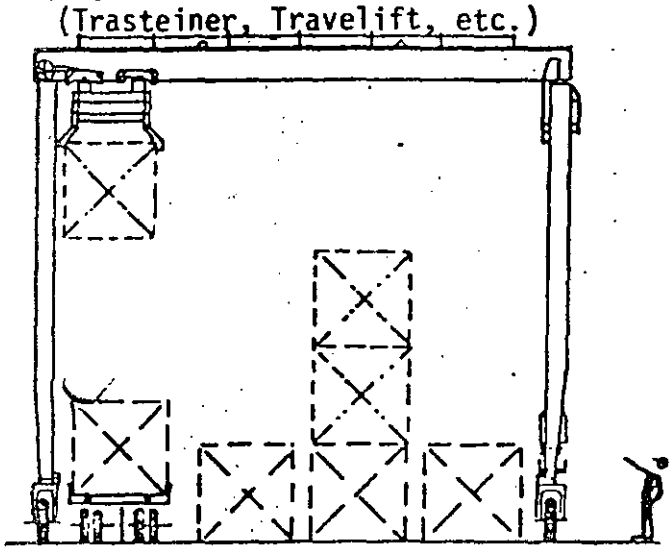
15 15



) Grúa de Pátio sobre llantas: ancho 2+1/

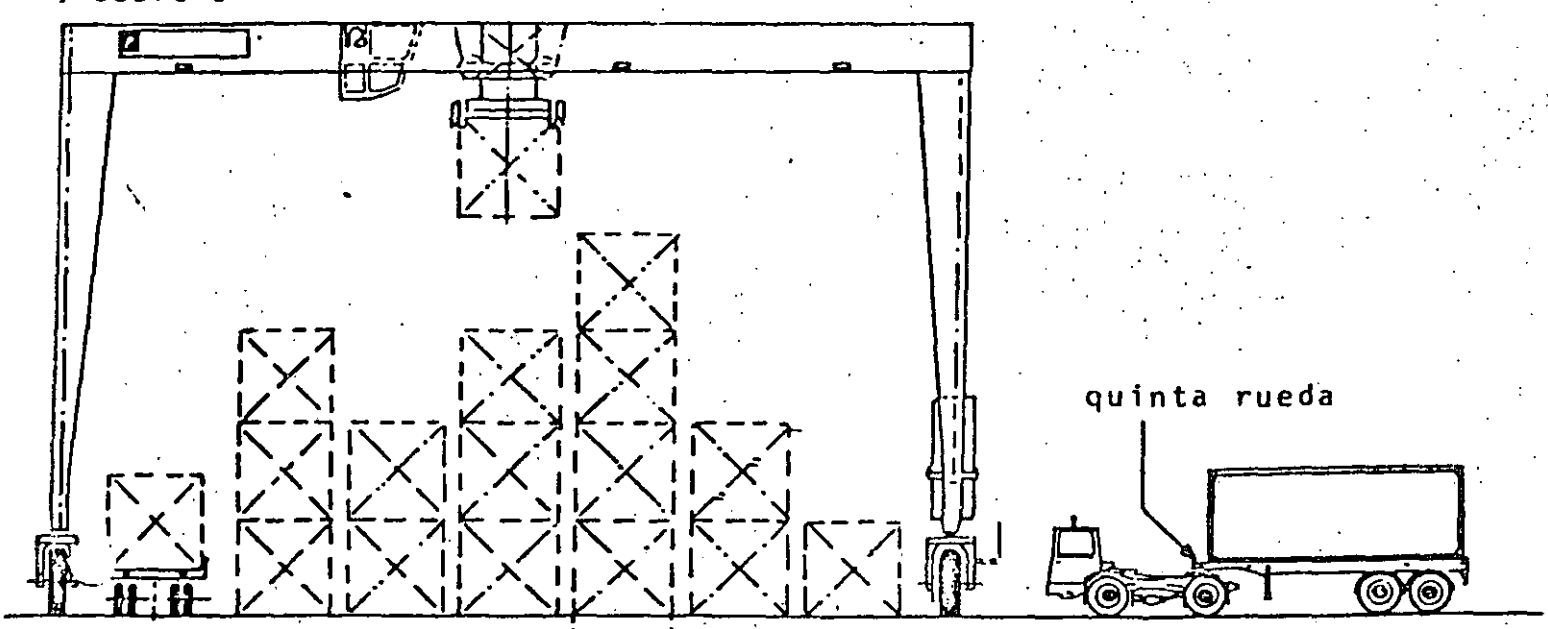
Altura 1 sobre 2

(Trasteiner, Travelift, etc.)

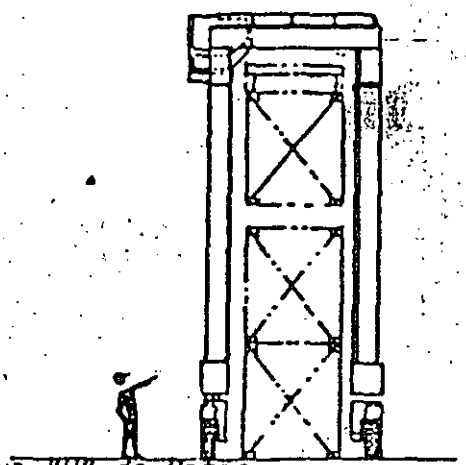


) Grúa de patio sobre llantas: 3+1/

1 sobre 3

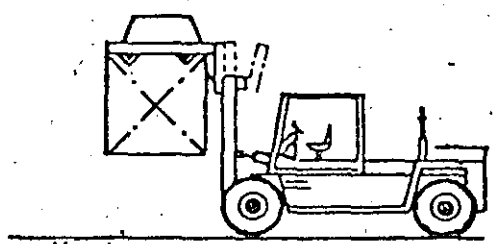


c) Grúa de patio sobre llantas: 6+1/1 sobre 4



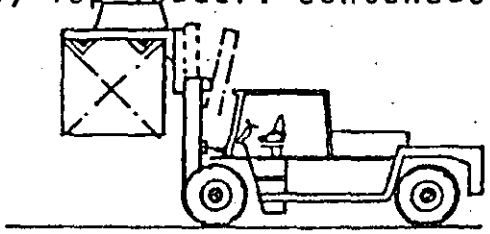
Grua "U" de Patio

d) Straddle carrier: 1 sobre 2



Montacargas

e) Top loader: contenedor de 20 pie



Montacargas

f) Top loader: contenedor de 40 pie

g)Tractor + Chasis

DATOS ESTADISTICOS ANUALES DE MANEJO DE CONTENEDORES PARA FINES DE DIAGNOSTICO DE OPERACION DE UNA TERMINAL

CONCEPTO	20'		40'	
	IMPORTACION	EXPORTACION	IMPORTACION	EXPORTACION
SERVICIO PUERTA A PUERTA				
SERVICIO PUERTA A PUERTO				
SERVICIO PUERTO A PUERTO				
CONTENEDOR CON CARGA COMPLETA				
CONTENEDOR CON CARGA MIXTA (DIVERSOS EMBARCADORES)				
VACIOS				
REFRIGERADOS				
CARGA PELIGROSA				
EN CUARENTENA				
DAÑADOS				
LIMPIEZA CONTENEDORES				
RECEPCION CAMION FF.CC.				
ENVIO CAMION FF.CC.				
TOTAL NUMERO DE BARCOS				
TIPO DE BARCOS 1a. GENERACION 2a. GENERACION 3a. GENERACION				

MANEJO ANUAL DE CONTENEDORES CON DISTRIBUCION MENSUAL

CONCEPTO	ENE		FEB		MAR		ABR		MAY		JUN		JUL		AGO		SEP		OCT		NOV.		DIC	
	I	E	I	E	I	E	I	E	I	E	I	E	I	E	I	E	I	E	I	E	I	E	I	E
CONTENEDORES:																								
C/CARGA 20'																								
C/CARGA 40'																								
VACIOS 20'																								
VACIOS 40'																								
CONT. REFRIGERADOS 20'																								
" " 40'																								
Nº DE CONTENEDORES POR BARCO																								

I = IMPORTACION
E = EXPORTACION

EQUIPAMIENTO DE UNA TERMINAL DE CONTENEDORES. — 18

Una vez concluidos los estudios económicos a nivel nacional y regional que determinen la necesidad de contar con una -- terminal de contenedores, su desarrollo puede ser por esta-- pas.

La primera comprende la planeación general de la terminal,- incluyendo largo y profundidad del muelle, extensión de --- áreas de tierra y los accesos terrestres. El muelle de refe rencia debera estar con la preparación para los rieles de - transito de una grúa de portico portacontenedores, los patios para almacenamiento de contenedores y la bodega de consolida ción y desconsolidación de contenedores. En esta étapa se - pueden utilizar las grúas del barco, una movil sobre camión y el equipo para transferencia y estiba.

Lo anterior obedece a que la grúa porta-contenedores tiene un costo del orden de $ 700 millones (1983), la cual se jus- tifica económicamente a partir de los 20,000 TEU/año.

La segunda etapa consiste en que una vez logrado el manejo mínimo de contenedores por año para ser rentable la grúa, se analise al sistema de equipamiento total mas adecuado.

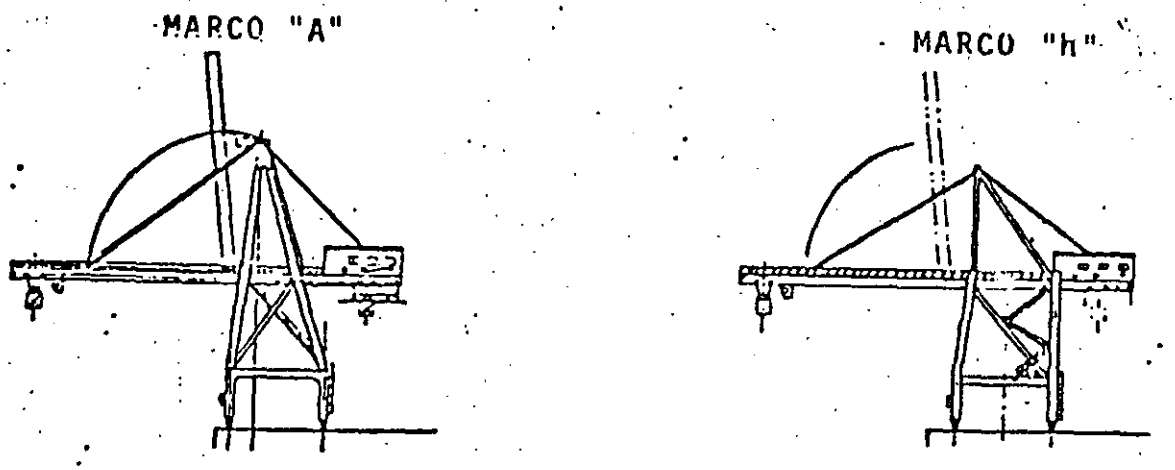
Una grua porta-contenedores de portico puede manejar un pro medio de 20 a 30 contenedores por hora y aproximadamente --- 40,000 contenedores al año.

La selección de las dimensiones de grúa porta-contenedores depende principalmente, del tamaño de los barcos a los que servirá la carga útil, varia de 30 a 40 tons., el alcance va de 25 m. para barcos de pequeños hasta 40 m. para barcos de la 2a y 3a generación.

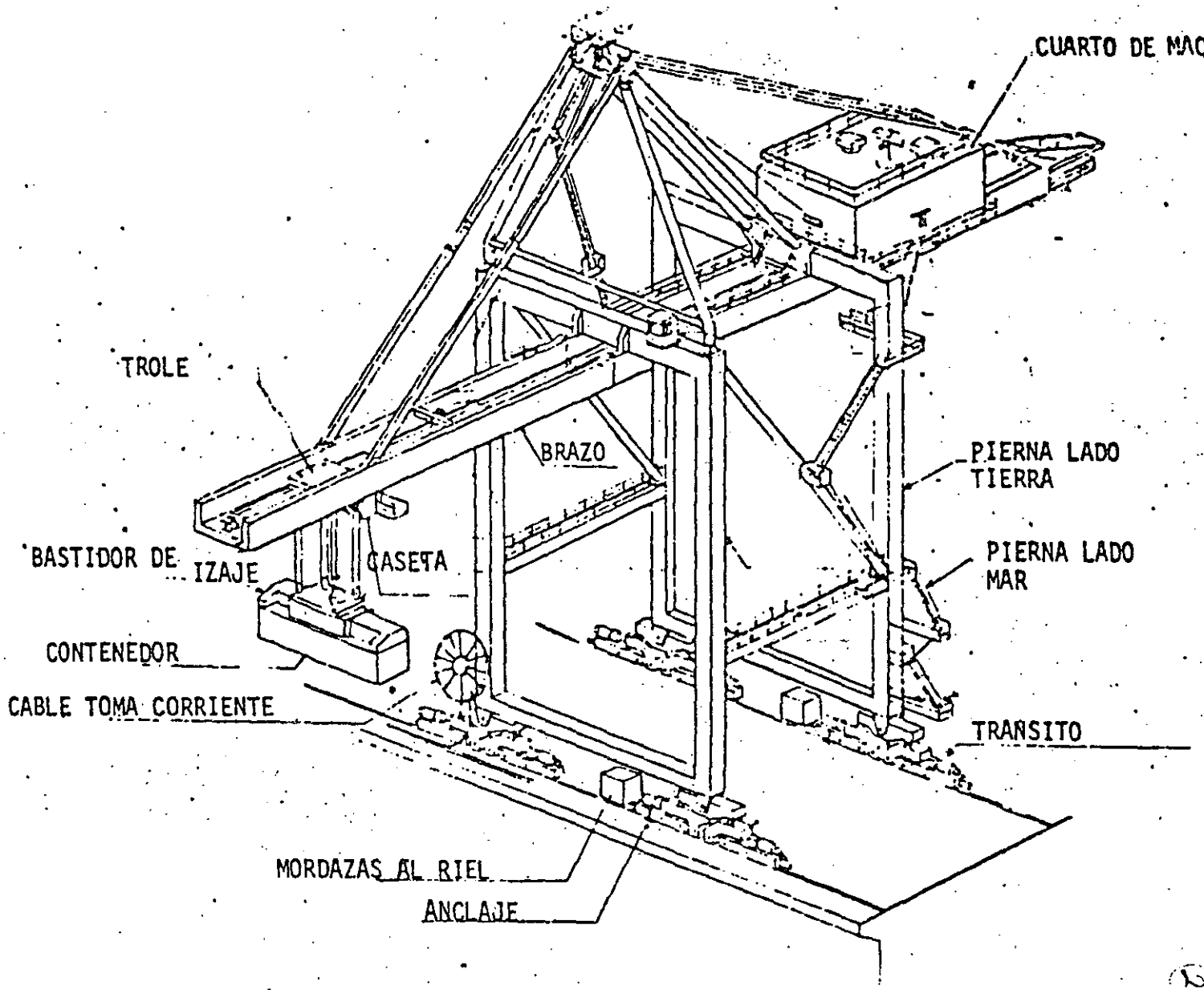
Las condiciones de operación fijan separación entre rieles que dependen de la desición de pasar vias bajo el portico. Dicha separación varia entre 15 y 20 m.

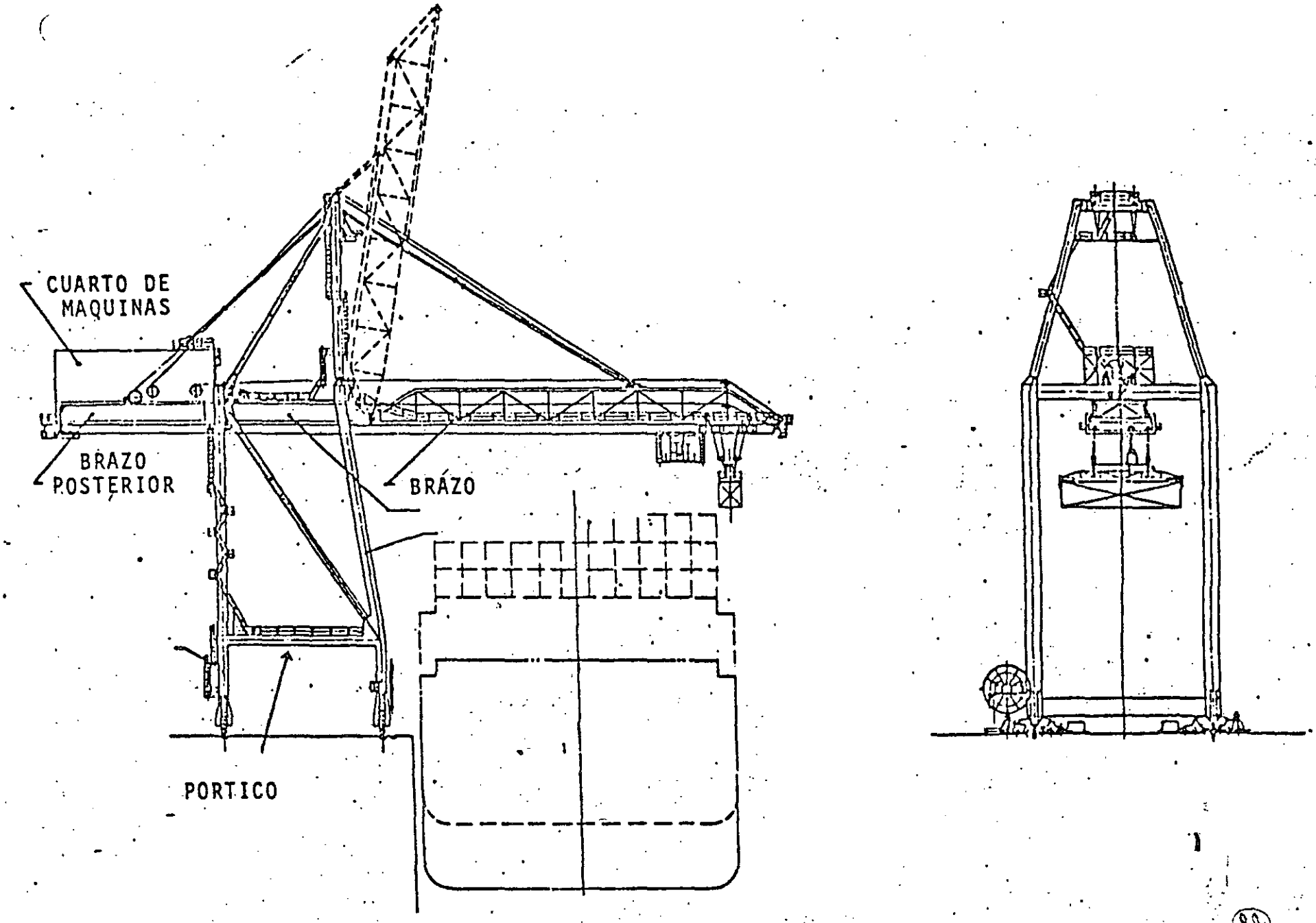
El número de grúas de portico depende del tráfico que se reciba, y es proporcionalmente mas elevado para un número reducido de tramos de atraque. En general es necesario una grúa más que el número de tramos de atraque, es decir, dos grúas para un tramo, tres grúas para dos tramos, etc.

Las siguientes figuras, muestran dimensiones; detales de construcción y operación de una grúa portacontenedores.

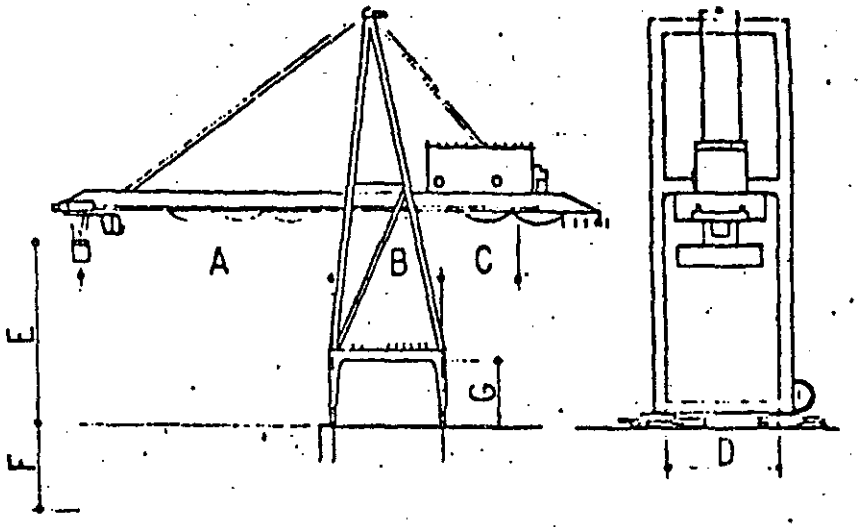


TIPO DE GRUAS PORTACONTENEDORES


CUARTO DE MAQUINAS
TROLE
BRAZO
PIERNA LADO TIERRA
PIERNA LADO MAR
BASTIDOR DE IZAJE
CASETA
CONTENEDOR
CABLE TOMA CORRIENTE
TRANSITO
MORDAZAS AL RIEL
ANCLAJE



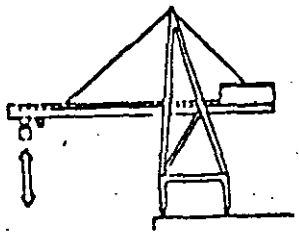
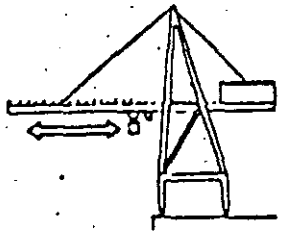
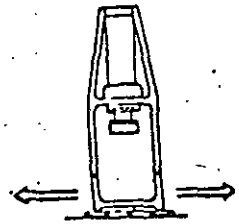
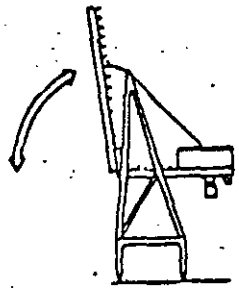
VISTA GENERAL DE UNA GRUA PORTACONTENEDORES



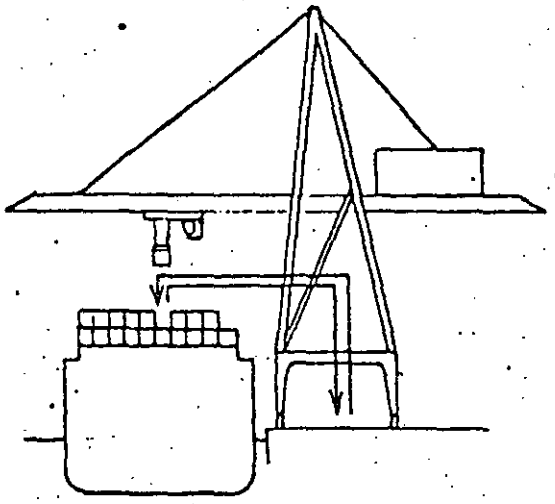
	CONCEPTO	FACTORES	DIMENSIONES	
A	BRAZO	° MANGA	500 TEU	: 27.5 m
			1000 "	: 30.5 m
			2000 "	: 35.5 m
B	SEPARACION RIELES	° ESTABILIDAD	2 LINEAS	: 13 m
		° Nº CARRILES DE EQUIPO DE TRANSFERENCIA	3 "	: 18.5 m
C	BRAZO POSTERIOR	°	1 LINEA	: 4 m
			2 LINEAS	: 9.5 m
D	ANCHO	° LARGO DEL CONTENEDOR	40'	: 14.5 m
E	ALTURA SOBRE MUELLE	° CALADO	500 TEU	: 21 m
			1000 "	: 22 m
			2000 "	: 25 m
F	ALTURA ELEVACION BAJO EL MUELLE	° CALADO	500 TEU.	: 9 m
			1000 "	: 10.5 m
			2000 "	: 12 m
G	BALIBO	° ALTURA DEL EQUIPO DE TRANSFERENCIA	Straddle carrier DE 3	ALTURA DE CONTENEDORE
			(for 8'6" CONTENEDORES	: 10 m
			(for 9'6" " " "	: 11 m

DIMENSI :S DE UNA GRUA POTA-CONTENEDORES

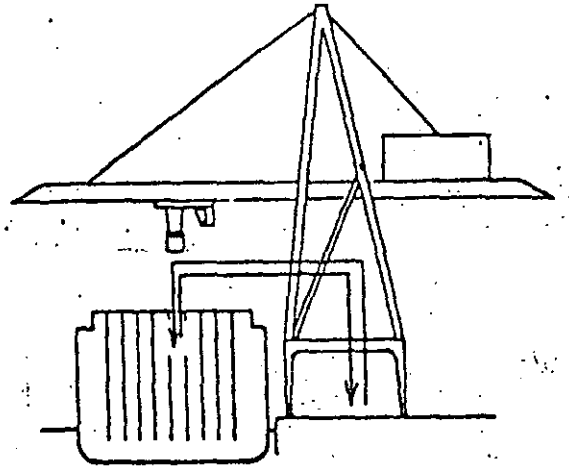
(22)

		VELOCIDAD		REQUERIMIENTOS DE ENERGIA ELECTRICA
		VACIO	CARGADO	(APROXIMADAMENTE)
	IZAJE	70-120 m/min.	35-50 m/min.	330 kW
	TRANSLACION CONTENEDOR		120-150 m/min.	75 kW
	TRANSLACION GRUA		abt. 45 m/min.	8 x 12.5 kW
	BRAZO		7-9 min./cycle	75 kW

VELOCIDADES DE OPERACION



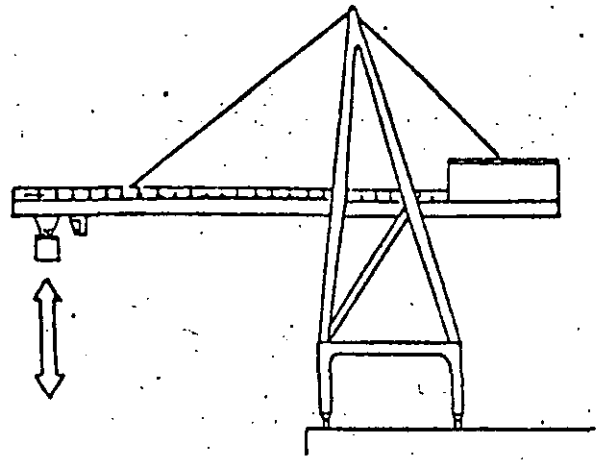
CICLO DE DESCARGA EN CUBIERTA DEL BARCO ..... 110 sec.
(32 units/hour)



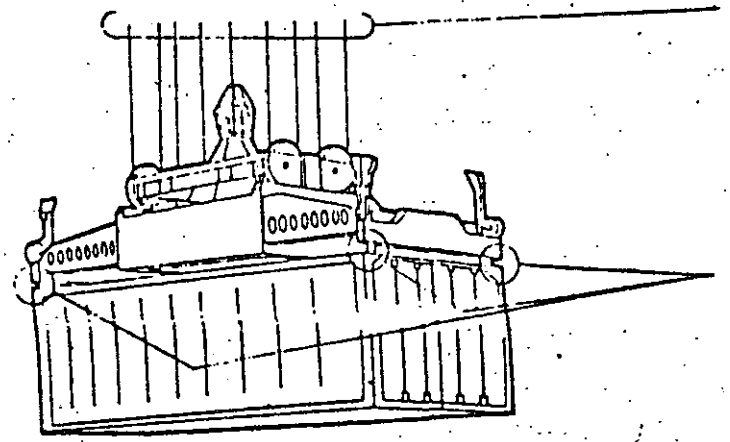
CICLO DE CARGA EN BODEGA DEL BARCO ..... 150 sec.
(24 units/hour)

CONDICIONES DE VIENTO

VELOCIDADES DE VIENTO $\leqq$ 16 m/sec.



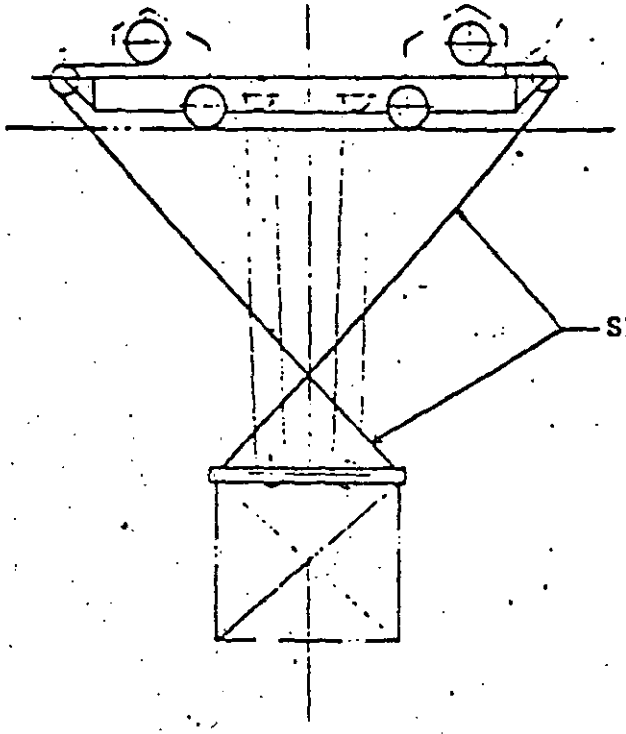
EN OPERACION



CARGA DE IZAJE

CARGA DE IZAJE
= PESO BASTIDOR (8 - 10 ton)
+ PESO CABLES
+MARGEN DE SEGURIDAD

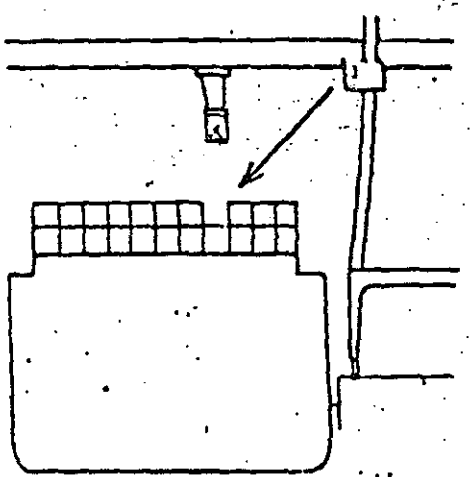
CARGAS ABAJO BASTIDOR DE CARGA
= PESO CONTENEDOR + MARGEN DE SEGURIDAD



SISTEMA ANTI-PENDULO

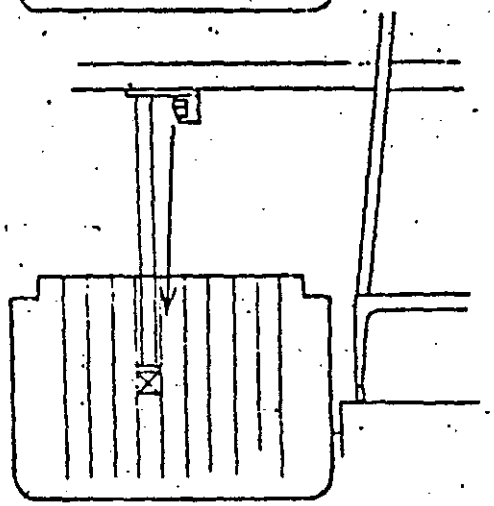
FIJO EN LA PIERNA LADO MAR DE LA GRUA

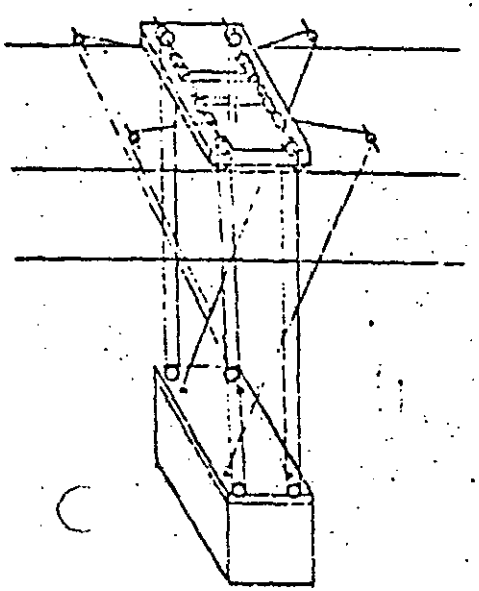
- BUENA VISIBILIDAD EN CUBIERTA DE BARCO



FIJA EN EL TROLE

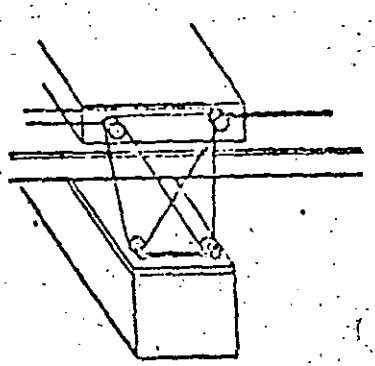
- BUENA VISIBILIDAD EN BODEGA DEL BARCO
- SISTEMA MAS EMPLEADO

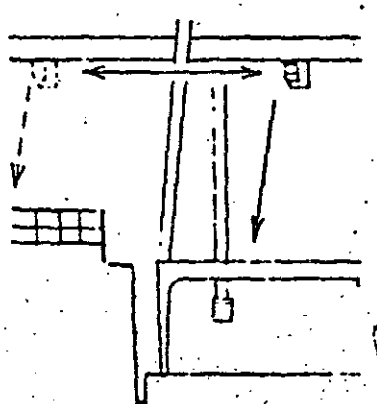




MOVIL

- SE ELIGE LA MEJOR VISIBILIDAD
- ALTO COSTO

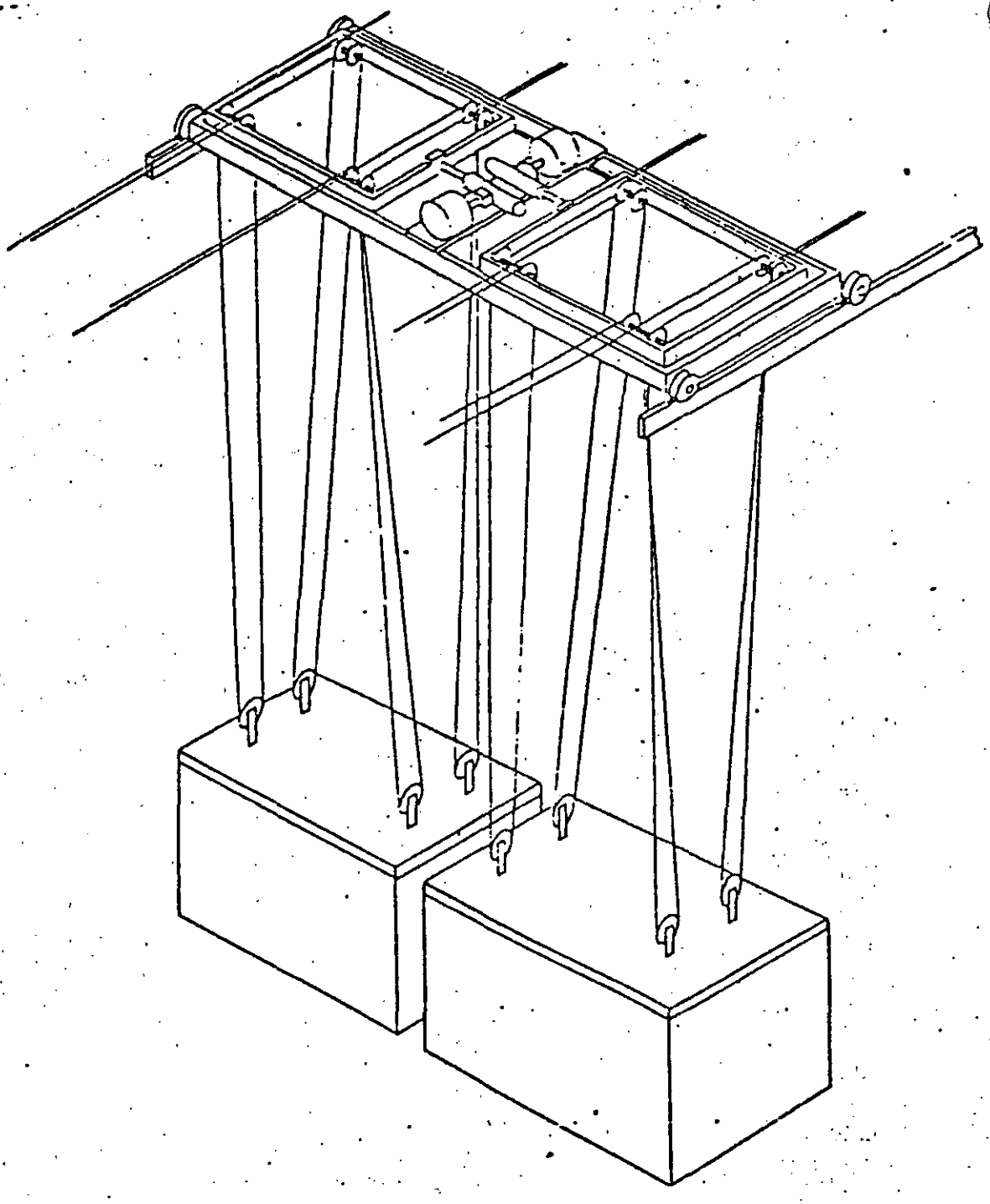




LOCALIZACION CASETA DE OPERACION

92

27



IZAJE DOBLE

SISTEMA DE MANEJO DE CONTENEDORES EN PATIO.

Los sistemas de manejo, se pueden dividir en:

A.- Sistema de chasis.

B.- Sistema de grúas tracto-apiladora (Straddle carrier)

C.- Sistema de montacargas.

D.- Sistema de Grúa Portico de patio sobre neumaticos (Transteiner, travelift, etc).

E.- Sistema de Grúa de patio sobre rieles. (Transteiner, travelift, etc.).

A continuación se describen los diversos sistemas:

A.- Sistema de Chasis.

La grúa portacontenedores deposita el contenedor sobre el chasis que un tractor transporta al patio, el cual es almacenado sobre el chasis. Este sistema es el empleado por la Compañía - SEA-LAND y presenta las siguientes ventajas.

1.- Es el sistema ideal para el servicio puerta a puerta.

2.- Los contenedores se manejan con mayor facilidad y rapidez que con cualquier sistema. El manejo de contenedores por - año es del orden de 2 a 3 veces el de los otros sistemas.

3.- Se reduce la frecuencia de movimientos directos de los contenedores, por lo que se reducen a un mínimo los daños.

4.- Dado que no existen vehiculos pesados, la superficie de rodamiento no demanda una pavimentación para servicio pesado.

Desventajas:

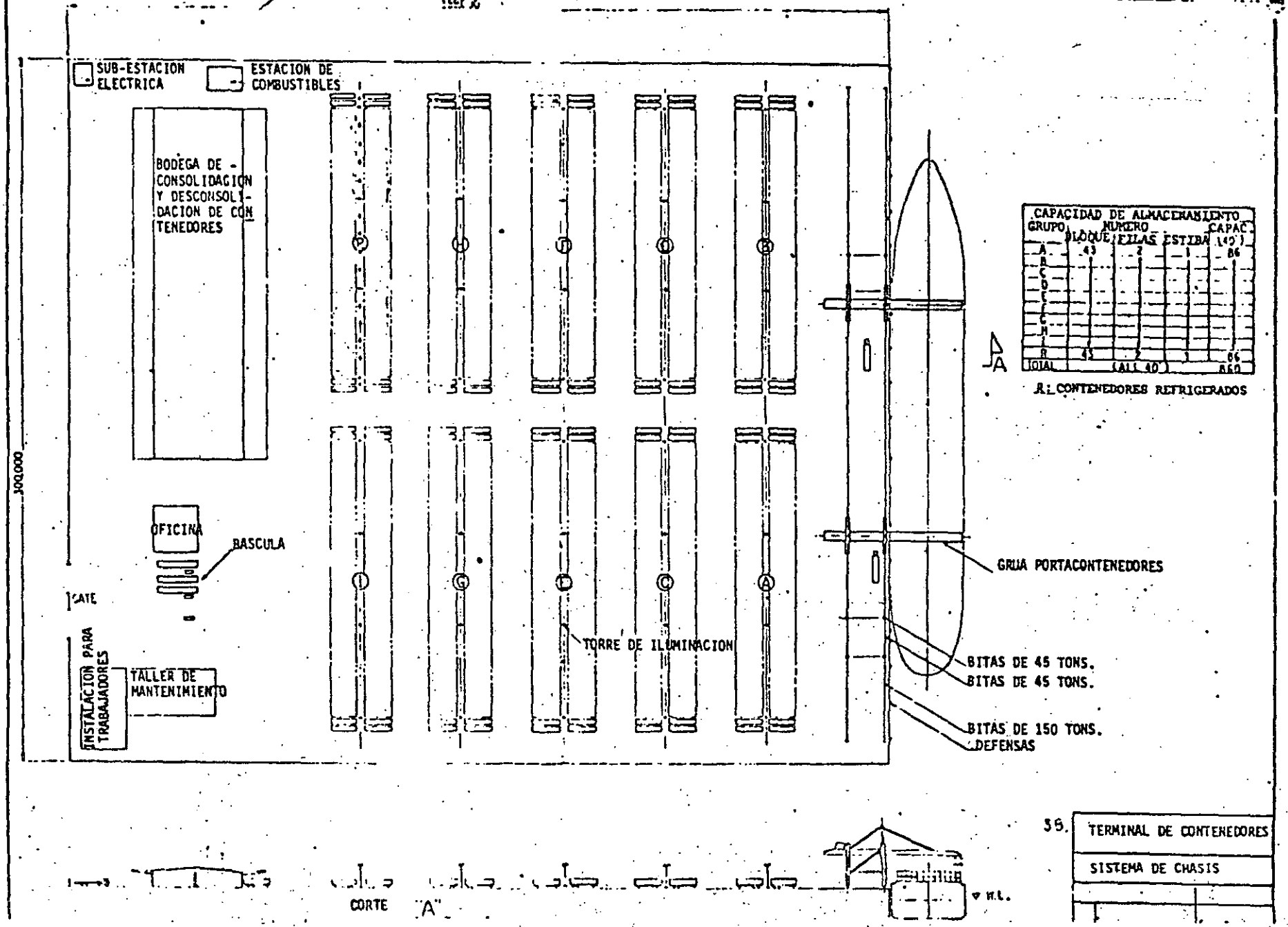
1.- Se requieren tantos chasis como contenedores en -- Patio, lo que elevara el valor inicial del equipamiento.

2.- Dado que los contenedores no pueden apilarse en capas multiples, los patios son de gran amplitud. Lo cual aumenta la inversión en intalaciones y servicios en tierra.

3.- Los chasis no solo se utilizan internamente en los patios, sino también fuera del mismo, por lo que - requieren ser chasis de carretera con alto valor y costo de mantenimiento. (SERVICIO PUERTA A PUERTA)

Este sistema requere de 40 m2./TEU de patios.

En la siguiente figura se muestra uan terminal operada bajo el sistema de chasises.



35.	TERMINAL DE CONTENEDORES
	SISTEMA DE CHASIS

B.- Sistema de grúas tracto-apiladoras.
(Straddle Carrier)

No obstante que este tipo de equipos utilizaron desde el incio de la contenedorización en 1960, es el equipo que a sufrido un mayor número de modificaciones en su sistema de transmisión. La primera generación contaba con 6 ruedas (neumaticos), trasmisiones mecánica por cadena e hidráulicos.

Los de trasmisión hidráulica a la fecha no se perfecciona causando trastornos en las zonas de circulación por las fugas de aceite.

La 2a. generación, con ocho ruedas, de trasmisión mecanica por flecha y motores eléctricos han mejorado su funcionamiento.

Estos equipos diseñados para el transporte y almacenamiento (uno sobre dos alturas de estiba) de contenedores en patio requieren gran habilidad de los operadores ya que con frecuencia se --- presentan daños en los contenedores y en el propio equipo; por los pequeños espacios libres disponibles a ambos lados del contenedor, su velocidad de transito es de 15 km/hr. ; cuando el número de contenedores por embarque ocupa una gran parte del bloque de contenedores de exportación es posible estibar a tres alturas para los de - importación a dos alturas por la necesidad de hacer entregas par-- ciales a través del autotransporte.

Actualmente existen del orden de 500 terminales de contenedores en el mundo, de las cual es el 40% utilizan este sistema.

El sistema presenta las siguientes ventajas y desventajas:

Ventajas:

1.- Es flexible para hacer frente a las modificaciones de la distribución de contenedores en los patios .

2.- Baja utilización de personal en la operación de transporte y almacenamiento de contenedores en patio.

3.- Es posible despachar rapidamente los contenedores.

4.- Dado que los contenedores pueden apilarse en capas multiples, se utilizan en forma eficiente los patios.

5.- Aumenta la utilización de la grúa portacontenedores al colocar el contenedor directamente en el muelle.

Desventajas:

1.- El pavimento de los patios deberá diseñarse para soportar una mayor carga.

2.- Dado que la mayoria son de accionamiento hidráulico, por lo general presentan gran número de escapes de liquido que dificultan el transito de otro tipo de vehículos y personal.

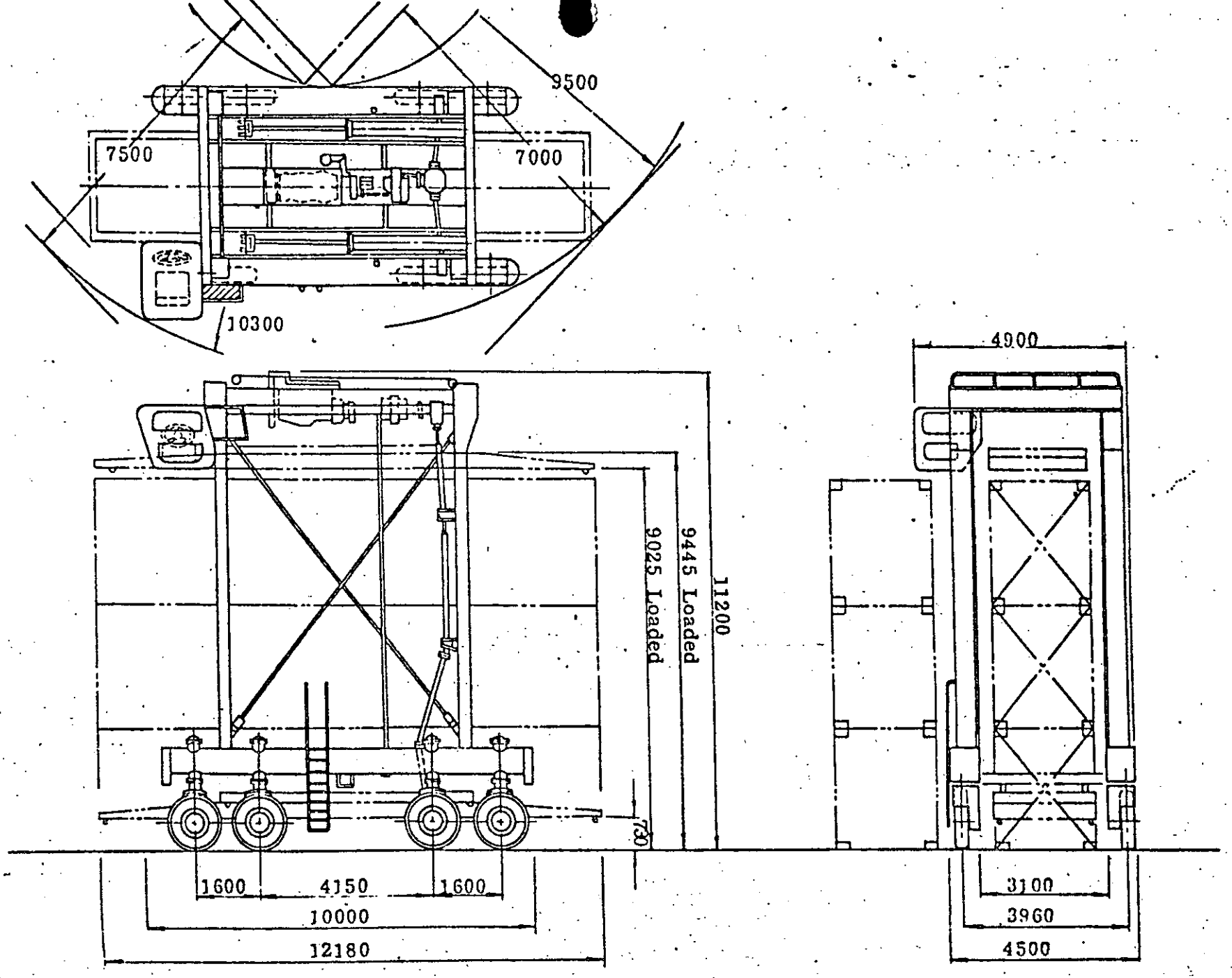
3.- Requieren de un alto costo de mantenimiento y alta habilidad para operarlos, que los demás métodos para el transporte y almacenamiento de contenedores en patio.

4.- No es posible la carga/descarga de contenedores en plataformas de ferrocarril.

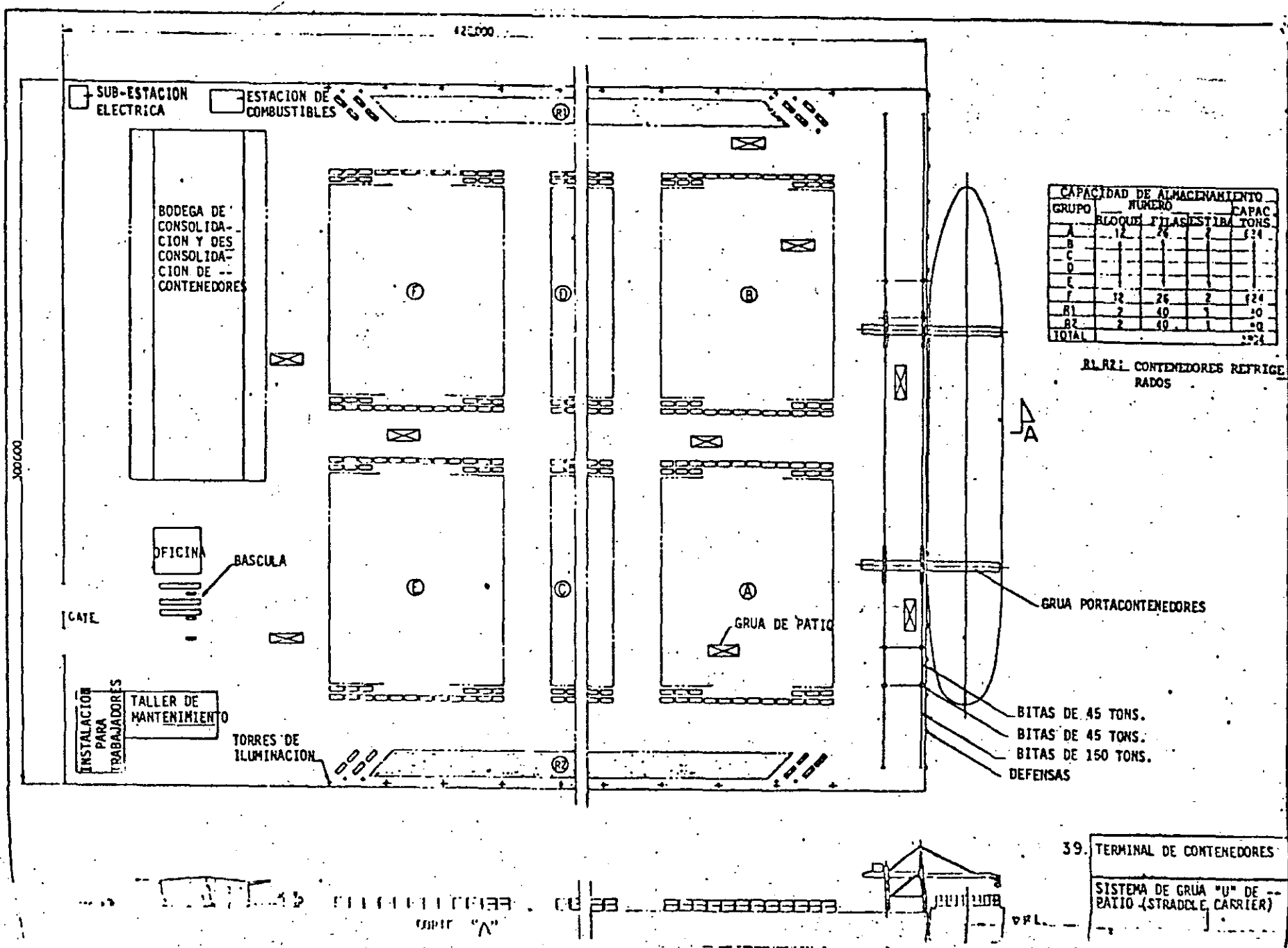
5.- Baja disponibilidad, (del orden del 80%) por fallas y trabajos de mantenimiento.

Este sistema requiere del orden de 25 m2./TEU de patios para - dos alturas de estiba.

En la siguiente figura se muestra una terminal manejada con el sistema de grúas "tracto-apiladoras"



GRUA "U" DE PATIO



CAPACIDAD DE ALMACENAMIENTO

GRUPO	NUMERO BLOQUES	FILAS	ESTIBA	CAPAC TONS
A	12	26	2	624
B				
C				
D				
E				
F	12	26	2	624
R1	2	40	1	80
R2	2	40	1	80
TOTAL				1408

R1,R2: CONTENEDORES REFRIGERADOS

39. TERMINAL DE CONTENEDORES

SISTEMA DE GRUA "U" DE PATIO (STRADDLE CARRIER)

C.- Sistema de Montacargas ,

Los montacargas pueden estibar los contenedores a dos alturas (los proveedores recomiendan 3 alturas, pero se reduce la eficiencia) y dos hileras .

Estos equipos operan en los patios unicamente estibando contenedores, siendo alimentados por equipos de transporte como tractores y chasises.

Dado que los contenedores de 20' en un 95% cuentan con perforaciones para las horquillas, se pueden utilizar montacargas.

Los contenedores de 40' estan diseñados para izarse por las cuatro esquinas superiores verticalmente, por lo que los montacargas requieren bastidor de izaje de contenedores. El 50% de estos contenedores (aproximadamente) cuentan con perforaciones para las horquillas del montacargas.

Este sistema requiere del orden de 40 a 50 m2./TEU. de patio en promedio.

Este tipo de equipo es el adecuado para la carga/descarga de contenedores transportados por barcos Ro/Ro.

Una variante de este tipo de equipos, lo forma la grúa hidráulica con pluma telescopica que permite la estiba a tres alturas y hasta cuatro hileras de contenedores, dado el alcance de su pluma.

D.- Sistema de Grúa Portico de patio sobre neumáticos. (37)
(Rubber tire transfer crane Transteiner).

Con este sistema de grúas se pueden almacenar hasta 7 hileras y apilar 4 contenedores.

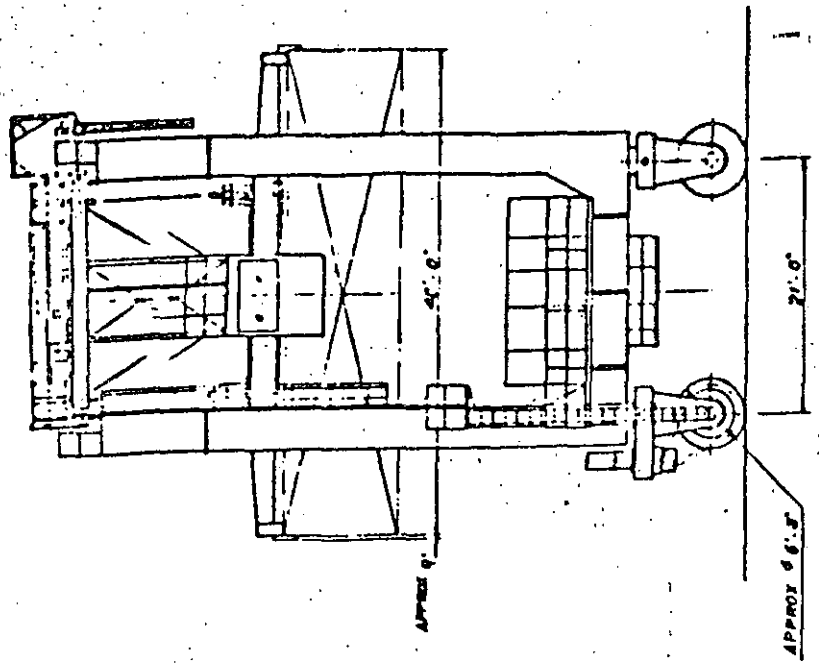
El tamaño mínimo de este tipo de grúas es para 3 hileras más un carril de transito, y 3 alturas de estiba.

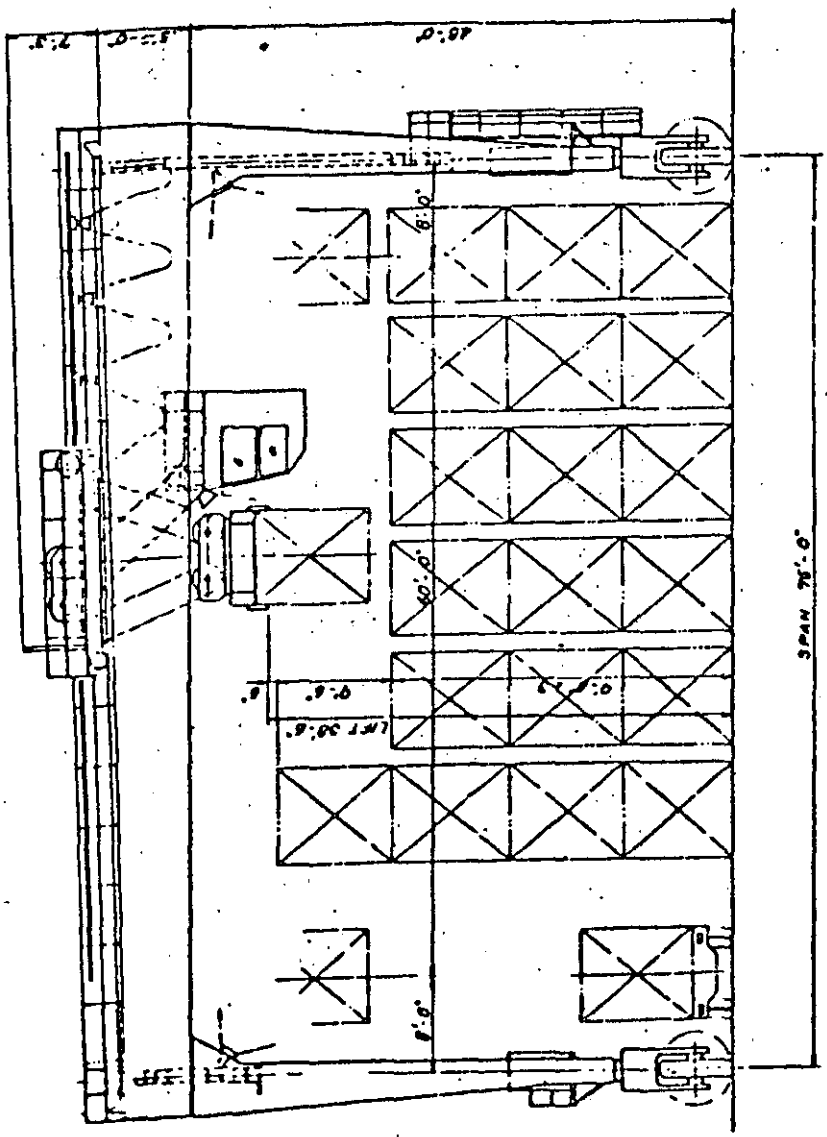
La velocidad de transito varia de 100 a 150 m/min. y la transversal del bastidor de izaje de 40 a 70 m/min.

Este tipo de grúas requiere de un pavimento para servicio pesado, sin embargo dado que esta limitada la ruta de transito, se requiere reforzar el pavimento en el área de desplazamiento.

Este sistema requiere del orden de 10-15 m2./TEU de patios, dependiendo de su capacidad de almacenamiento estático, ya que varian de 3 X 3 a 6 X 4 .

En la siguiente figura se muestra una grúa portico de patio.





GRUA DE PATIO SOBRE NEUMATICOS

SUB-ESTACION ELECTRICA

ESTACION DE COMBUSTIBLES

BODEGA DE CONSOLIDACION Y DESCONSOLIDACION DE CONTENEDORES

OFICINA

BASCULA

CAJE

TORRES DE ILUMINACION

INSTALACION PARA TRABAJADORES

TALLER DE MANTENIMIENTO

GRUA PORTACONTENEDORES

GRUA PORTICO DE PATIO S/NEUMATICOS

BITAS DE 45 TONS.

BITAS DE 45 TONS.

BITAS DE 150 TONS.

DEFENSAS

CAPACIDAD DE ALMACENAMIENTO

GRUPO	NUMERO BLOQUE	FILAS	ESTIBA	CAPAC
A	19	6	3	342
B				
C				
D				
E				
F				
G				
H				
I				
J				
K	19		1	342
L	5	1	3	180
R	12	6*	2	144
TOTAL				4086

R: CONTENEDORES REFRIGERADOS

40

TERMINAL DE CONTENEDORES

SISTEMA DE GRUA DE PATIO SOBRE NEUMATICOS

F.- Sistema de Grúas Portico de patio sobre rieles.
(Reil Monted transfer Crane)

Estas grúas, denominadas "del mañana", permiten el máximo de almacenamiento de contenedores en la menor -- área de patios, permitiendo la total automatización del sistema.

La translación y movimiento del bastidor de izaje son a base de motores electricos con toma corriente paralela a uno de los rieles.

Aunque, en patio se pueden apilar hasta 8 contenedores, seguridad, ante todo por sismo y viento, de acuerdo con la norma 150 1496 (provisional), sección uno, fija en 5 la altura máxima de contenedores cargados del mismo tamaño.

Por lo anterior, para este tipo de grúas la altura máxima de apilamiento es de 4 y 5 contenedores.

Las dimensiones máximas en la práctica son: 35 a 45 m. de claro con dos volados en uno ó ambos extremos de -- 5 a 7 m.

En su interior pueden almacenar hasta 15 hileras de contenedores con 5 alturas de apilamiento, uno delos volados es para la carga-descarga de plataformas de F.C. ó

tracto-camiones; en el otro volado se almacenan de 2 a 3 hileras de contenedores.

La velocidad de transito es 100 a 150 m./min. y la de desplazamiento transversal del bastidor de izaje de 80 a 100 m./min.

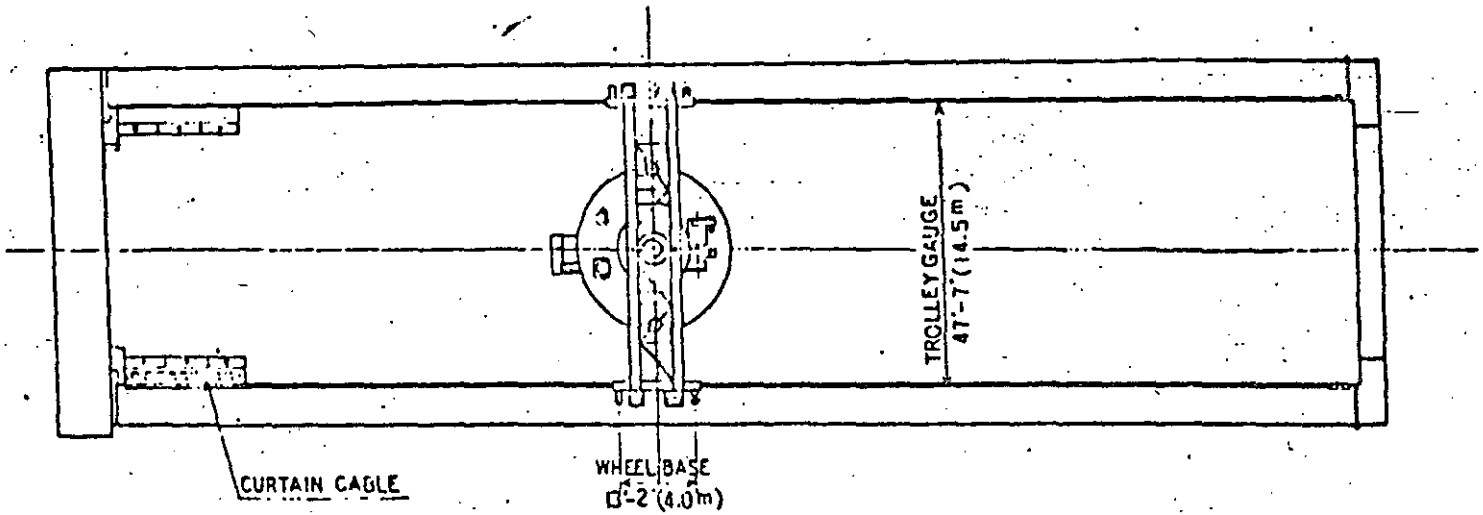
Esta grúa requiere de un diseño especial de cimentación por las grandes descargas que provoca.

Este sistema es adecuado para un alto manejo de contenedores sobre todo transportador con barcos de 3a. generación, donde se requiere una total automatización.

La transferencia de contenedores de la grúa a los patios se realiza por medio de tractores y chasis. Respecto a los tractores, se prevé para un futuro sercano la transferencia de mas de 4 a la vez.

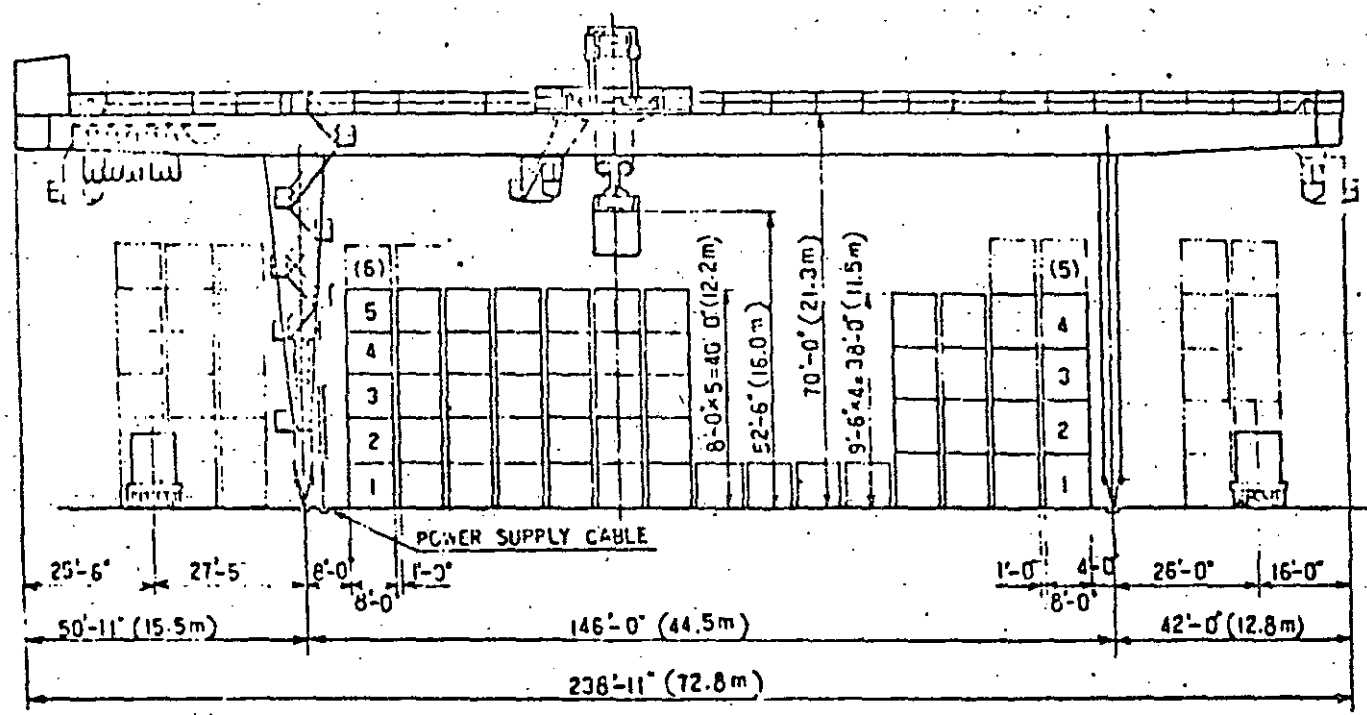
Este tipo de grúas es la que requiere menos costo de mantenimiento y de operación con respecto a los otros sistemas. En patio se requiere del orden de 9 m2/TEU.

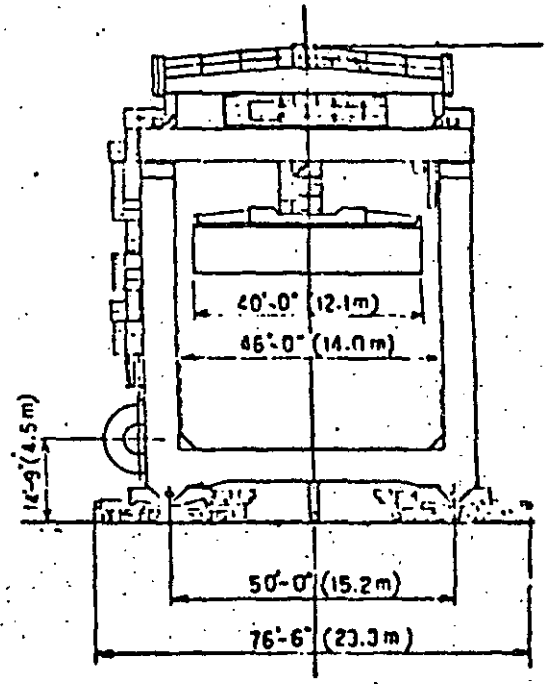
En la siguiente figura se muestra una grúa de este tipo.



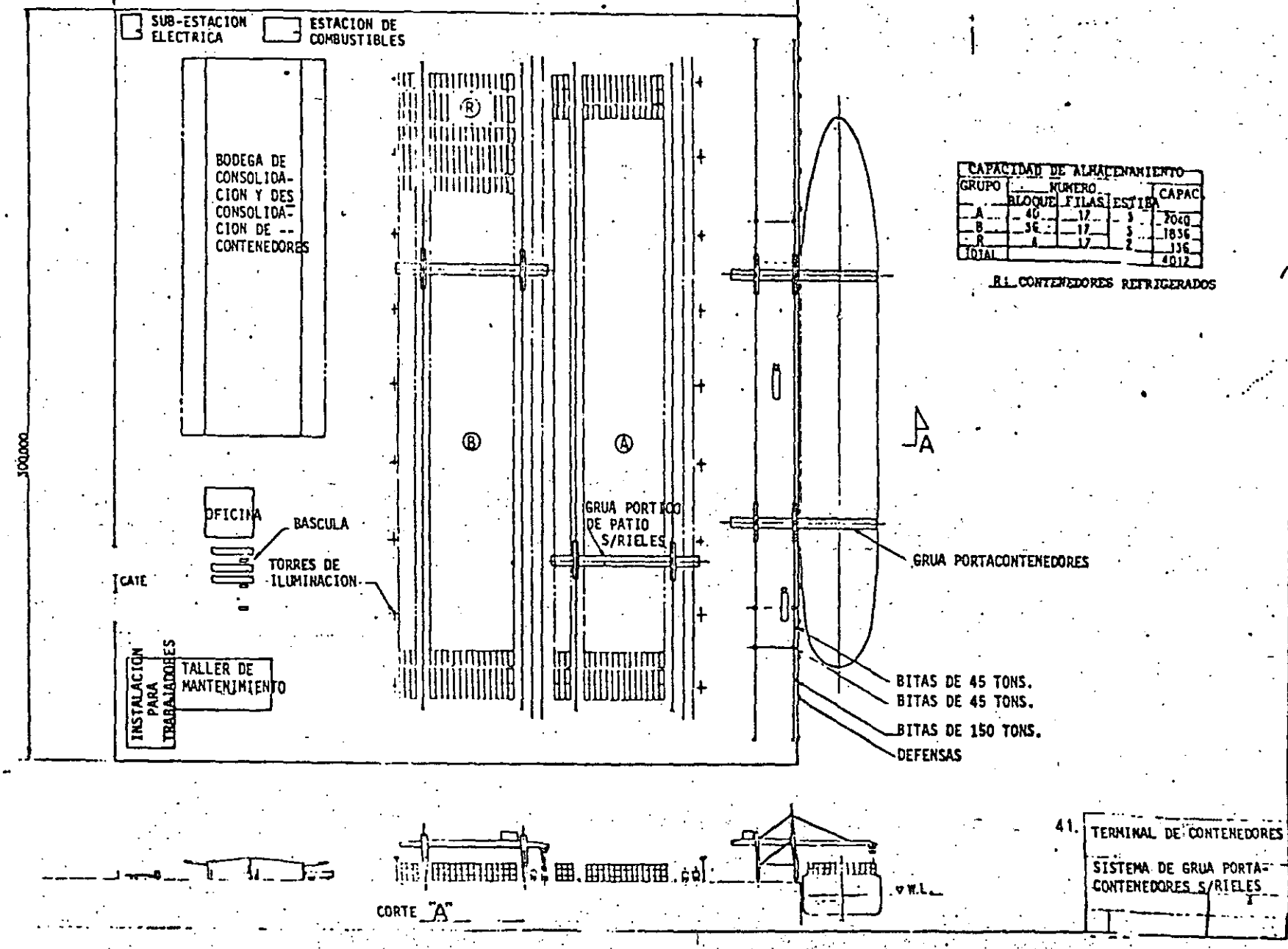
CARACTERISTICAS

RATED LOAD	35LT
RAIL SPAN	144'-0" (43.89m)
HEIGHT OF LIFT	52'-6" (16.00m)
HOIST / LOWER	LOADED 150ft/min (45.72 m/min)
HOIST / LOWER	UNLOADED 300ft/min (91.44 m/min)
TROLLEY TRAVEL	400ft/min (121.92 m/min)
GANTRY TRAVEL	300ft/min (91.44m/min)
TROLLEY ROTATION	3/4 r/min
GANTRY RAIL	73 kg/m
POWER SUPPLY	AC 3300V 50Hz





GRUA DE PATIO SOBRE RIELES



41. TERMINAL DE CONTENEDORES

SISTEMA DE GRUA PORTA-CONTENEDORES S/RIELES

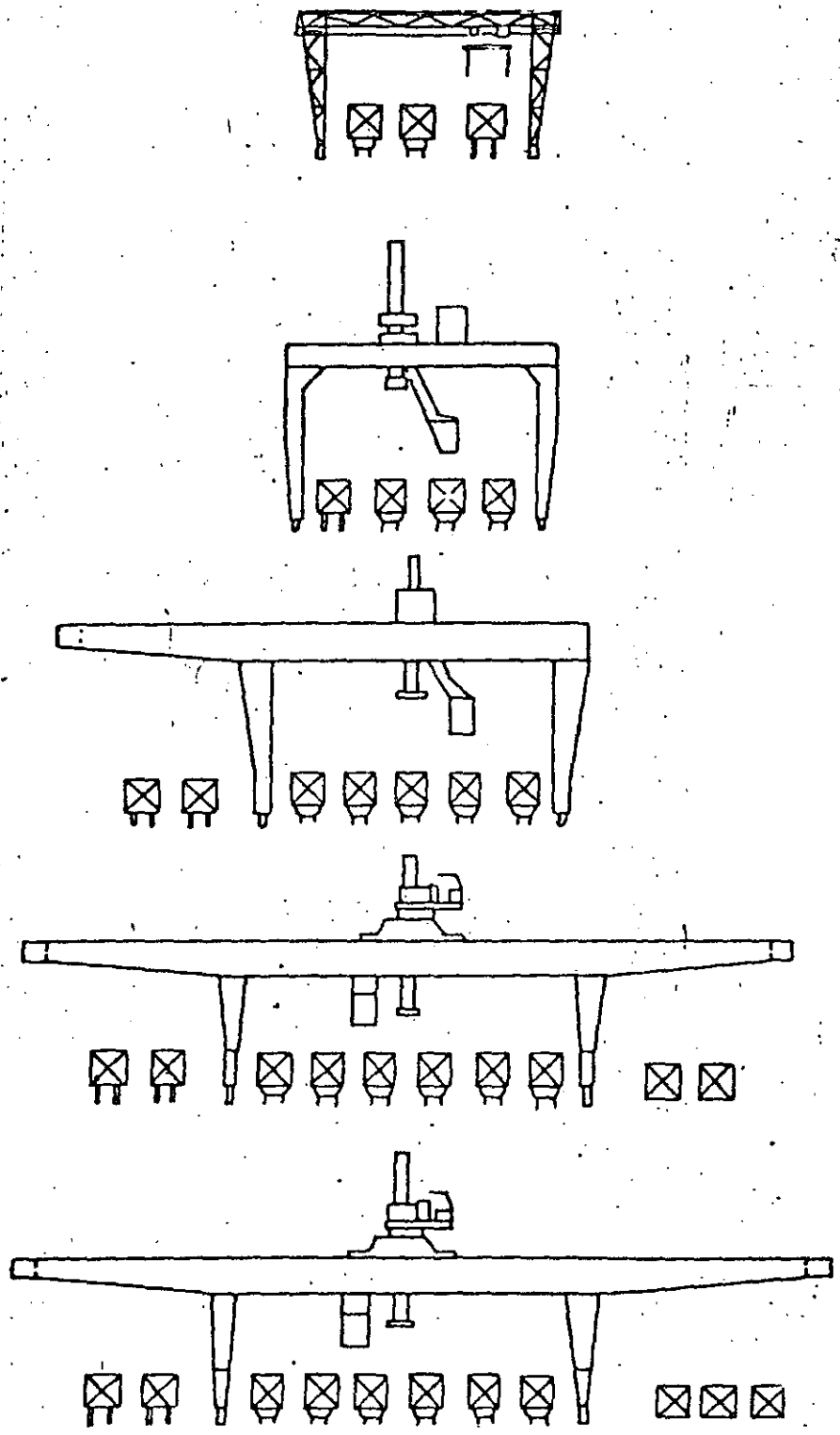
030

040

250

262

263



VISTA DE FRENTE DE GRUAS PORTICO S/RIELES
PARA CARGA/DESCARGA DE FF.CC. Y CAMIONES

VISTA DE FRENTE ~~GRUAS~~ /RIELES

PARA CARGA/DESCARGA DE FC Y CAMIONES

COMPARACION DE LOS DIVERSOS SISTEMAS DE MANEJO DE CONTENEDORES EN PATIO.

Dado el alto coosto de los equipos, esconveniente realizar - estudios de la demanda en en campo de los contenedores, para elejir desdee el punto de vista económico a largo plazo - el sistema adecuado.

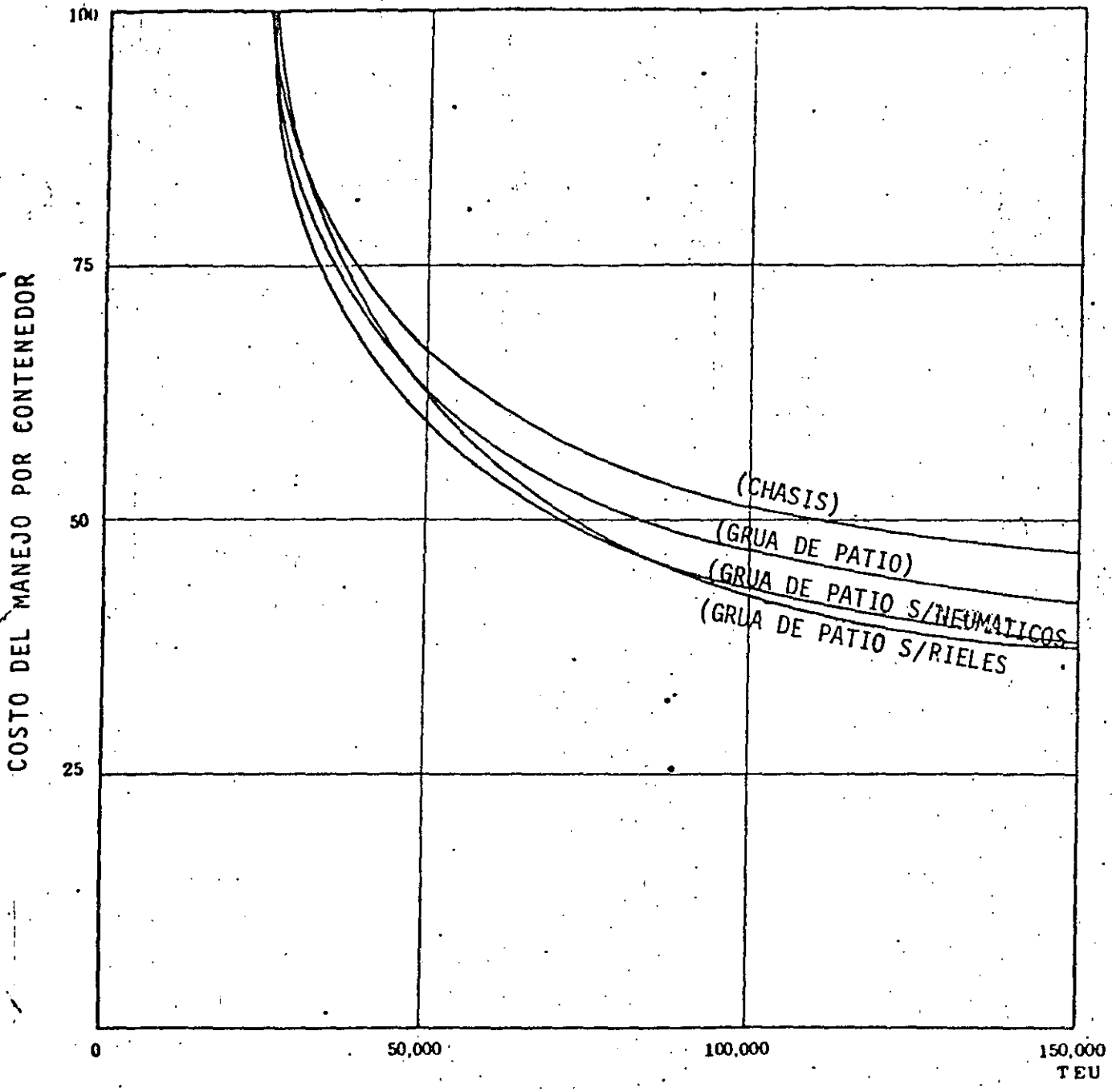
Los equipos mayores, tienen vidas económicas que varian de 15 a 25 años por lo que la decisión debe incluir horizontes de análisis del mismo orden.

A continuación se presentan una tabla mostrando el equipo - y áreas necesarias en cada sistema. Una comparación de costos en función del sistema empleado y el número de contenedores manejados por año y por último un diagrama mostrando el costo relativo entre valor de terreno, el muelle y patios, los servicios y el equipamiento en terminales de contenedores del Japon. y una comparación cualitativa de los sistemas.

COMPARACION CUANTITATIVA DE LOS DIVERSOS SISTEMAS PARA MANEJO DE CONTENEDORES.

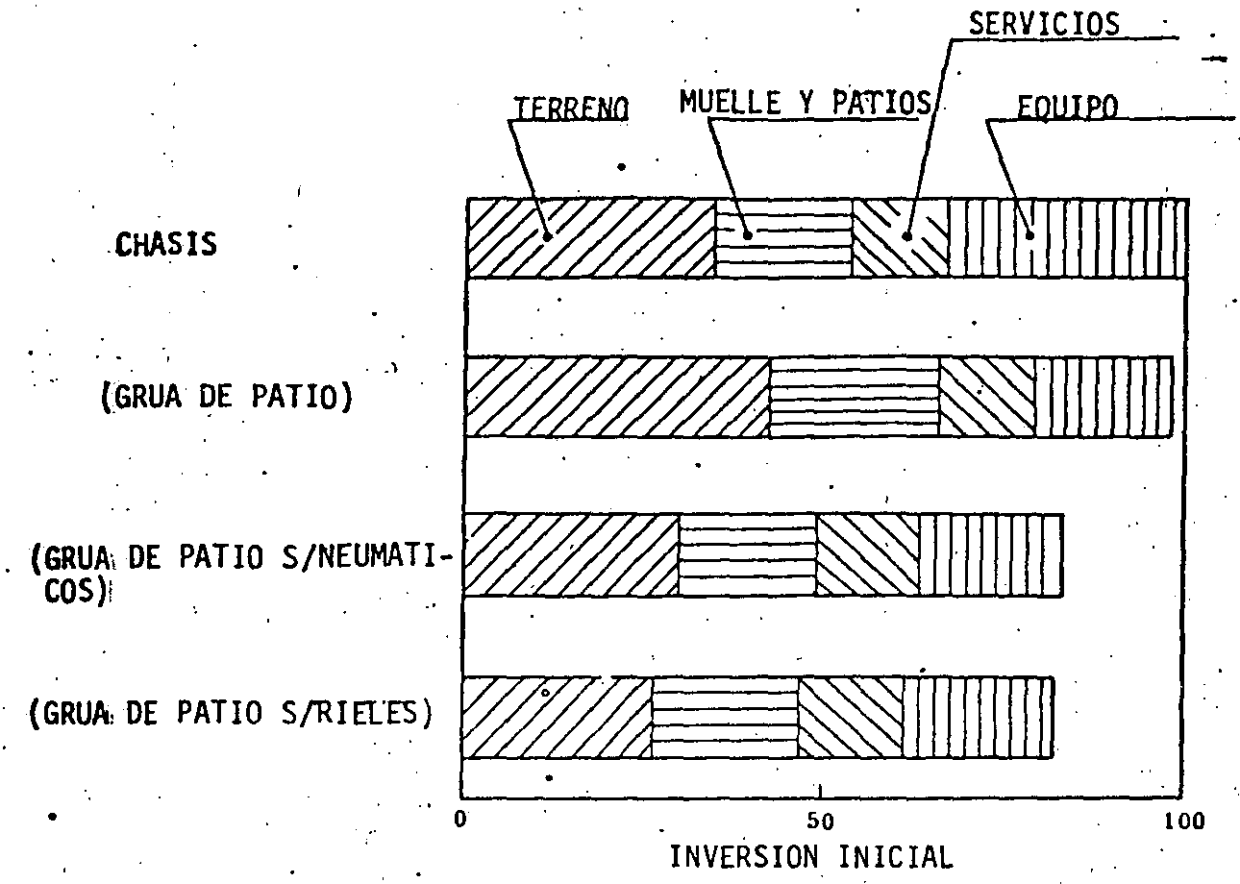
(CANTIDAD = 100,000 TEU)

SISTEMA	EQUIPO DE MANEJO						AREA EN TIERRA (HA)	CAPACIDAD DE PATIOS (TEU)
	GRUA PORTACONTENEDORES	CHASIS	GRUA DE PATIO	GRUA DE PATIO S/ NEUMATICOS	GRUA DE PATIO S/RIELES	TRACTOR		
CHASIS	2	720 + 30				18 + 2	10.77	1720
GRUA "U" DE PATIO	2		9				12.60	2904
GRUA PORTICO DEPATIO S/ NEUMATICOS	2	10 + 30		5		10 + 2	9.60	4086
GRUA PORTICO DE PATIO S/RIELES	2	10 + 30			2	10 + 2	8.37	4012



MANEJO DE CONTENEDORES POR AÑO

COSTO DEL MANEJO POR CONTENEDOR



COSTO RELATIVO A LOS DIVERSOS SISTEMAS DE MANEJO DE CONTENEDORES EN EL JAPON

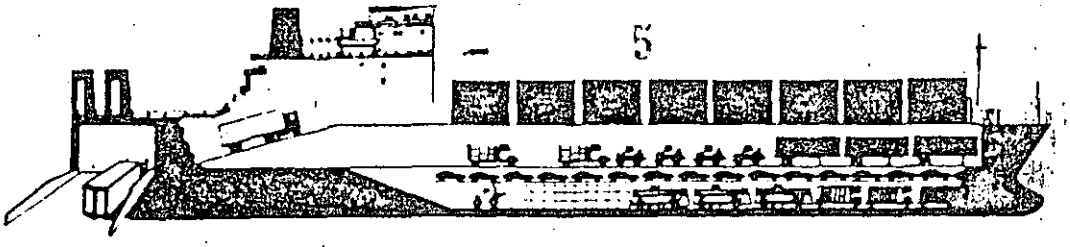
COMPARACION CUALITATIVA DE SISTEMAS PARA MANEJO DE CONTENEDORES

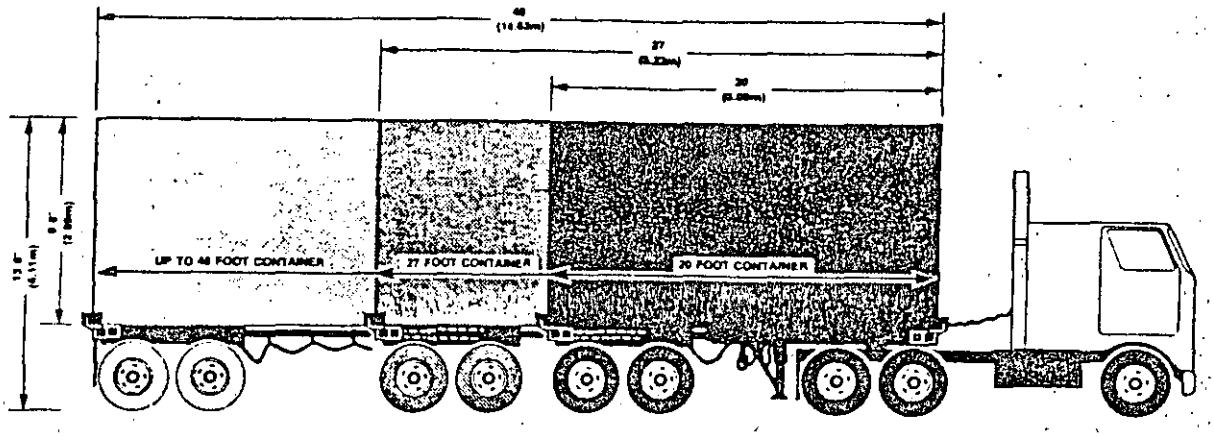
CONCEPTO \ SISTEMA	TRACTOR CON CHASIS	GRUA DE PATIO	GRUA DE PATIO S/ NEUMATICOS	GRUA DE PATIO S/RIELES
CAPACIDAD DE ALMACENAMIENTO	△	○	◎	◎
COSTO INICIAL	△	○	◎	◎
SIMPLICIDAD	◎	○	△	△
EFICIENCIA EN LA TRANSFERENCIA	◎	○	△	○
EFICIENCIA EN MUELLE	△	◎	△	△
FLEXIBILIDAD DE LAS OPERACIONES	◎	○	△	△
SEGURIDAD DE LOS CONTENEDORES	◎	△	○	○
COSTO DE MANTENIMIENTO DE EQUIPOS	○	△	◎	◎
FLEXIBILIDAD DE LAS OPERACIONES	◎	◎	△	△
POSIBILIDAD DE AMPLIACION DE AREAS	◎	○	△	△
ADAPTACION A LA AUTOMATIZACION	△	△	○	◎
CARGA /DESCARGA A F.C.	△	△	○	◎

◎ EXCELENTE ○ BUENO △ CUESTIONABLE

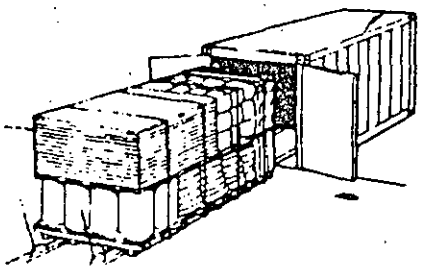
NOMENCLATURA DE EQUIPOS PARA MANEJO DE CONTENEDORES

Equipo	Denominación Generica en Ingles	Nombres Comerciales
1) Tracto apiladora	Straddle Carrier	Straddle Carrier Container Carrier
2) Grúa portacontenedores	Container grane Ship-To - Shore Container Gantry Crane	Portainer Container Crane
3) Bastidor de izaje	Spreader	Spreader
4) Montacargas lateral	Side Loader	Side Loader
5). Montacargas	Fork Lift truck Frontend Loader	Fork lift
6) Grúa Pórtico de Patio sobre neumaticos y/o rieles.	Rubber tire Gantry Crane Rail Gantry Crane	- Transteiner - Shifter - Straddle Hoist - Stacker Crane - Straddle Crane
7) Grúa Hidráulica con - pluma telescopica		Hidraulic Crane
8) Silo para contenedores	Container silo Storage System	Silo Cont.
9) Tractor Ferroviario		Track Movil

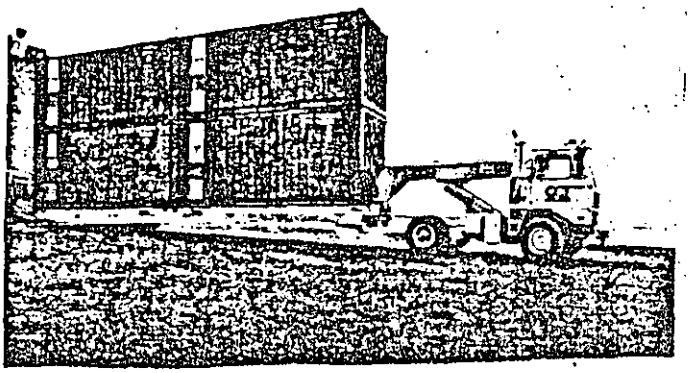


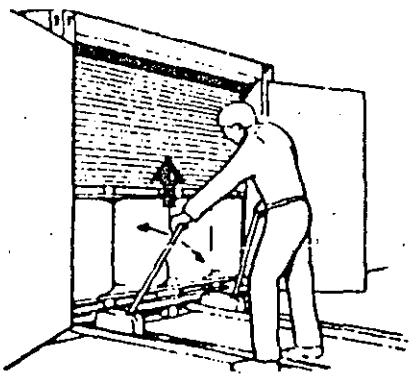


A new telescoping trailer is revolutionizing container transportation by making it possible for one trailer to handle all container sizes, including the hi-bulk, 9'6" containers. Contact: John Lee (415) 986-3868.



The conveyors are pumped up raising the pallet clear of the container floor.





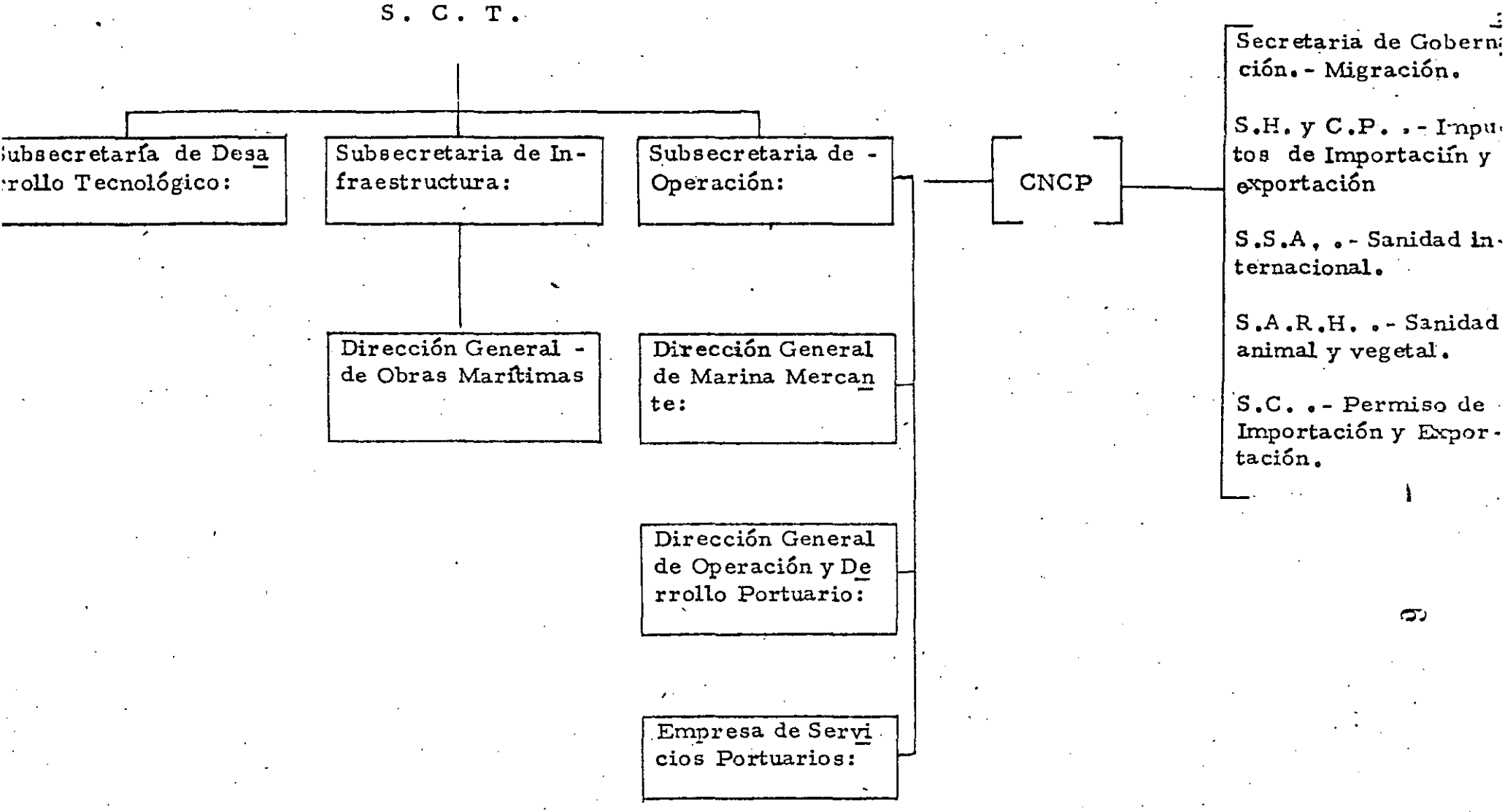
The whole load is rolled clear of the container.

INOVACIONES TECNOLOGICAS

EN MANEJO DE CONTENEDO-

RES

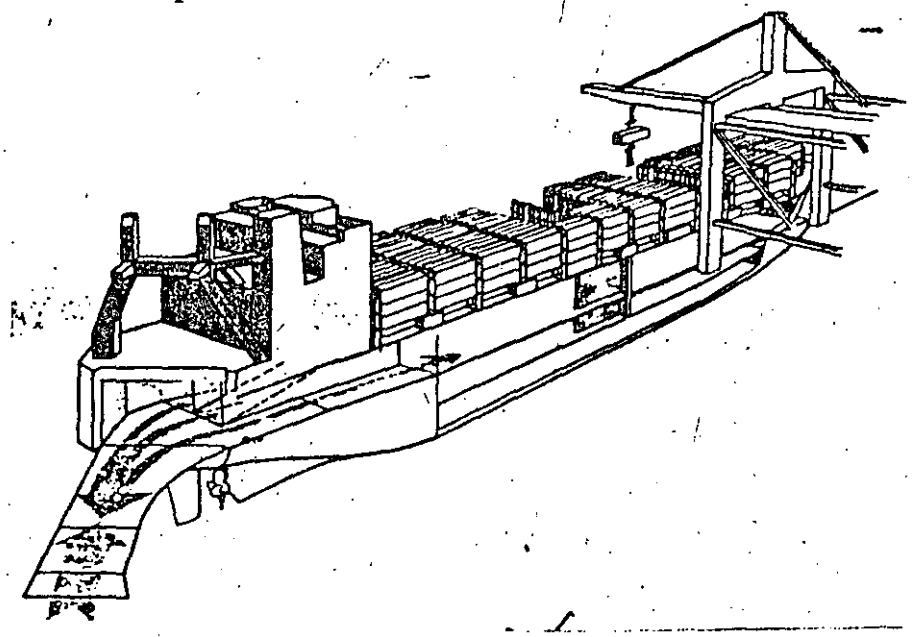
For descriptive brochure contact:
McQuade-Cormany Associates
26 Broadway, Suite 741
New York, NY 10004
(212) 425-6828



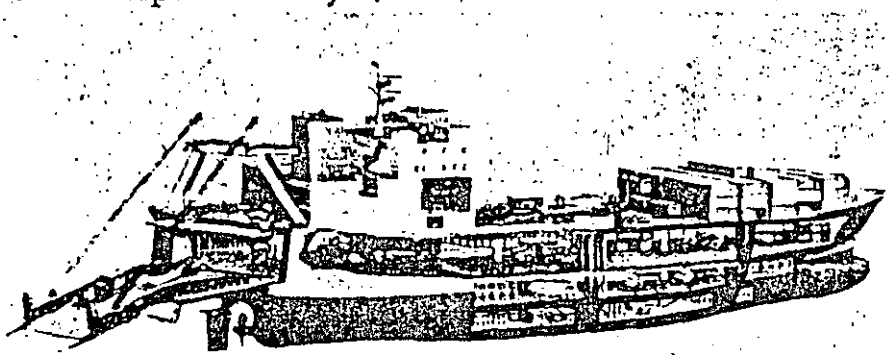
P/ Abril 2, 1984.

Barco mixto para el transporte de contenedores (Lift on/Lift off Lo/Lo) y transbordo de carga por rodadura (Roll on/Roll off - Ro/Ro), paralelamente al inicio del manejo de contenedores con equipo de tierra y barcos especializados (1960) surgio la necesidad de contar con una embarcación capaz de realizar el transbordo de la carga con el mínimo de equipo, para atender a los puertos que carecian de él. Estos barcos denominados "de transbordo por rodadura" (Ro/Ro) existen de la primera generación, o transbordadores de la 2a. , con rampas externas normalmente con un sentido de circulación de vehículos y los de la 3a. generación denominados con/Ro, por rampa con doble circulación para carga Ro/Ro bajo cubierta y sobre cubierta, transportar contenedores con el sistema Lo/Lo. Estos barcos por su alto costo, son empleados por los países Industrializados que manejan grandes volúmenes de carga de alta densidad económica en viaje redondo ó en países contituidos por infinidad de islas. Sin embargo se considera el barco del futuro por su alta eficiencia. A continuación se muestra este tipo de barcos.

Barco Tipo 1500 y 1300



Barco Tipo 1100 y 900



Tipo	1500	1300	1100	900
Eslora total	198.00 m.	182.26 m.	166.53 m.	150.80 m.
Eslora - P.P.	183.00 "	167.26 "	151.53 "	135.00 "
Manga	28.00 "	28.00 "	28.00 "	28.00 "
Puntal	17.70 "	17.70 "	17.70 "	17.70 "
Calado	9.30 "	9.40 "	9.50 "	9.50 "
TPM	23,900 t	21,100 t	18,200 t	14,800 t
Area de carga	12,210 m2.	10,690 m2.	9,170 m2.	7,650 m2.
N° de contenedores	1,340 TEU	1,161 TEU	999 TEU	840 TEU
Velocidad	16.3 km.	16.8 km.	16.6 km.	17.1 km.

FUENTE: WWS/WORLD PORTS

53

ANEXO 1

CLASIFICACION Y DIMENSIONES DE CONTENEDORES

VALOR DE CONTENEDORES (A DICIEMBRE DE 1983)

Contenedor standar nuevo de 20' , puesto en México.

$ 20,000.00 U.S. usado 1/3 de nuevo

Contenedor nuevo de 40', puesto en México

$ 30,000.00 usado 1/3 del valor de nuevo

* El contenedor usado, requiere mantenimiento menor.

Los contenedores refrigerados (autonomos) tienen un valor aproximado del doble respecto a los contenedores estandar nuevos.

JP/ 26/III/84.

LOS CONTENEDORES ESTAN DISEÑADOS PARA SER APILADOS HASTA -- SEIS ALTURAS EN PATIOS DE ALMACENAMIENTO (EN LA PRACTICA -- CINCO ALTURAS, DEPENDIENDO DE LA VELOCIDAD DE LOS VIENTOS - DOMINANTES Y REINANTES). A BORDO DE LAS EMBARCACIONES,LA ALTURA DE ESTIBA EN BODEGA ES DE HASTA NUEVA CONTENEDORES Y - SOBRE CUBIERTA Y TAPA ESCOTILLAS DEL 25 AL 35% DE LA ESTÍBA EN BODEGA, O SEA DE TRES A CUATRO CONTENEDORES,NORMALMENTE VACIOS.

CONTENEDORES FUERA DE LAS NORMAS 150:

LOS CONTENEDORES DE 20' ESTAN DISEÑADOS PARA OPERAR CON CARGA BRUTA DE 20,320 KG., SIN EMBARGO EN ALGUNAS RUTAS SE MANEJAN DE 24,000 KG.

LOS CONTENEDORES-TANQUE GENERALMENTE ESTAN DISEÑADOS PARA - 24,000 KG., DE PESO BRUTO, PERO EXISTEN DE 25,000 KG.. EN - LAS RUTAS MARITIMAS DE AMERICA DEL NORTE SE UTILIZAN CONTENEDORES DE 40' X 8' X 6" (2900 mm.). LOS CUALES NO SON UTILIZADOS EN PAISES CON LIMITACIONES DE DESCARGA POR EJE Y GALIBO DE PUENTES.

CONTENEDORES SEALAND (35 PIES DE LARGO)

EN ESTE TIPO DE CONTENEDORES LOS PUNTOS DE IZAJE DE LAS ESQUINAS ES DIFERENTE A LOS CONTENEDORES-150, POR LO QUE HAY QUE PREVER ESTO EN EL DISEÑO DEL BASTIDOR DE IZAJE DE LOS - CONTENEDORES. SEA-LAND INTRODUJO UN NUEVO CONTENEDOR DE 40' CON DOBLE SISTEMA DE IZAJE.

CONTENEDOR DE 45 PIES DE LARGO.

EN LAS RUTAS ENTRE FILIPINAS, JAPON Y EE.UU. SE INICIO LA - UTILIZACION DE ESTE TIPO DE CONTENEDORES, POR LO QUE SE TENDRA QUE ESTUDIAR SU POSIBLE UTILIZACION EN UN FUTURO EN --- NUESTRO PAIS.

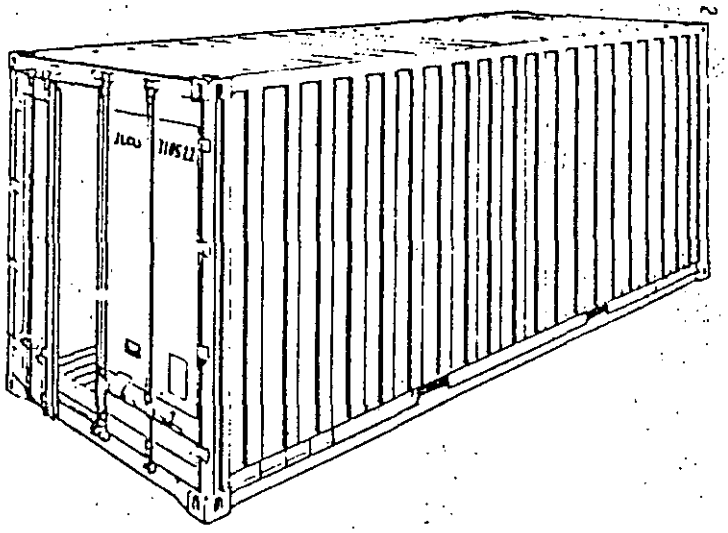
TIPOS DE CONTENEDORES 150.

1.- CONTENEDORES DE CARGA GENERAL.

COMPRENDE LOS CONTENEDORES CERRADOS CON PUERTAS EN UN EXTREMO Y EN LAS PAREDES LATERALES; LOS DE TECHO ABIERTO; LOS DE PAREDES LATERALES ABIERTAS; LOS DE PAREDES Y TECHO ABIERTO, PLATAFORMAS, MEDIA ALTURA Y LOS VENTILADOS (NO ISOTERMOS).

20'x8'x8' Dry Cargo

CARGA SECA

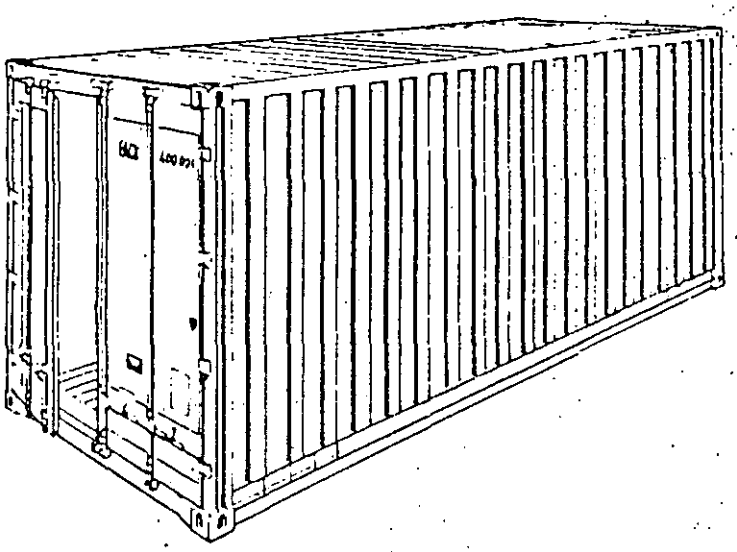


Manufactured according to ISO and ASA recommendations and standards.
JLCU containers approved by Germanische Lloyd and/or American Bureau of Shipping.
BSLU/EACU containers approved by Lloyds Register of Shipping.
Certified for inland transport under customs seal.

Fittings:
fork pockets
● fumigation point

20'x8'x8'6" Dry Cargo

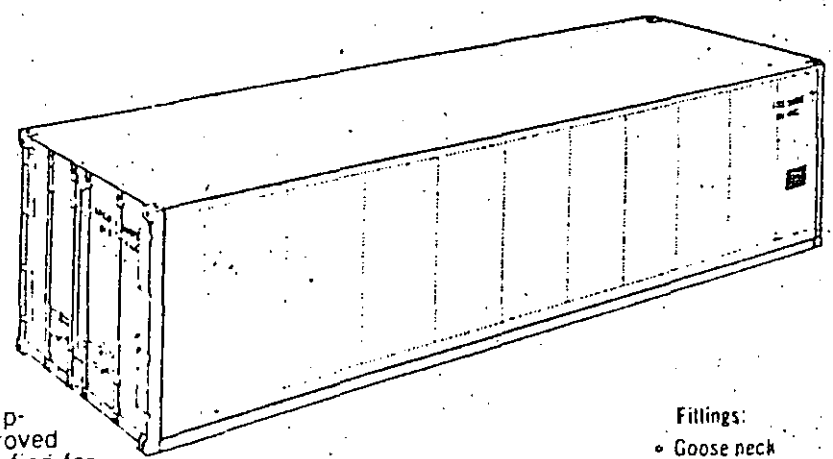
CARGA SECA



Manufactured according to ISO and ASA recommendations and standards.
Approved by Lloyds Register of Shipping.
Certified for inland transport under customs seal.

40'x8'x8'6" Dry Cargo

CARGA SECA



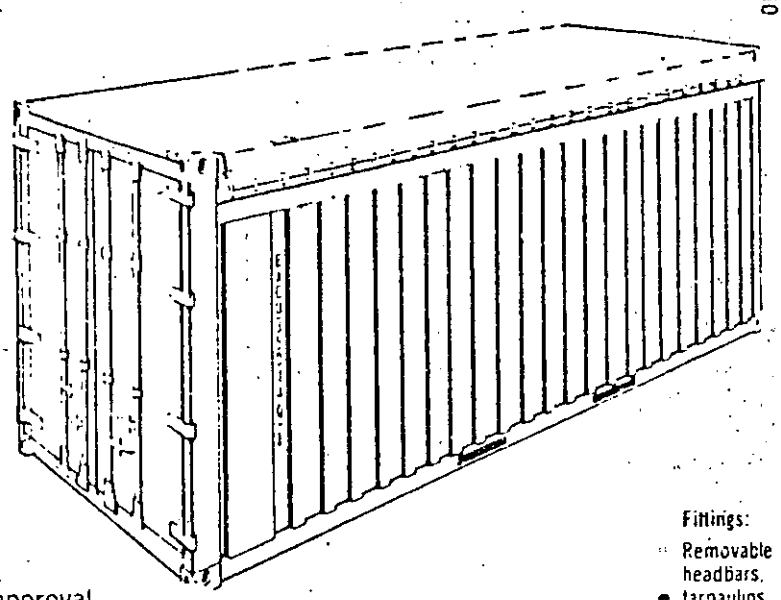
Manufactured according to ISO and ASA recommendations and standards. JLCU containers approved by Germanische Lloyd and/or American Bureau of Shipping. BSLU/EACU containers approved by Lloyds Register of Shipping. Certified for inland transport under TIR approval.

Fittings:
○ Goose neck
● Fumigation point

(62)

20'x8'x8' OpenTop

SIN TECHO

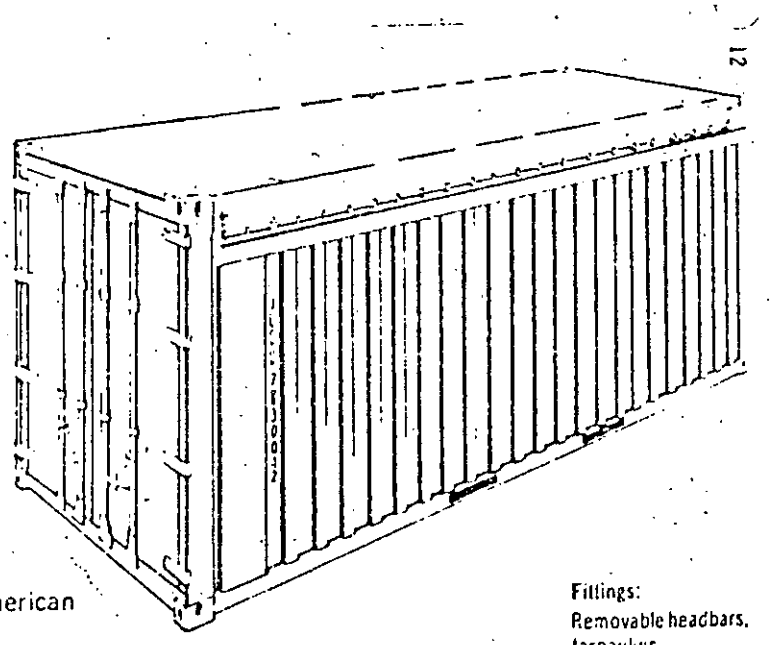


Manufactured according to ISO and ASA recommendations and standards.
Approved by Lloyds Register of Shipping.
Certified for inland transport under TIR approval.

Fittings:
- Removable headbars,
- tarpaulins

20'x8'x8'6" OpenTop

SIN TECHO

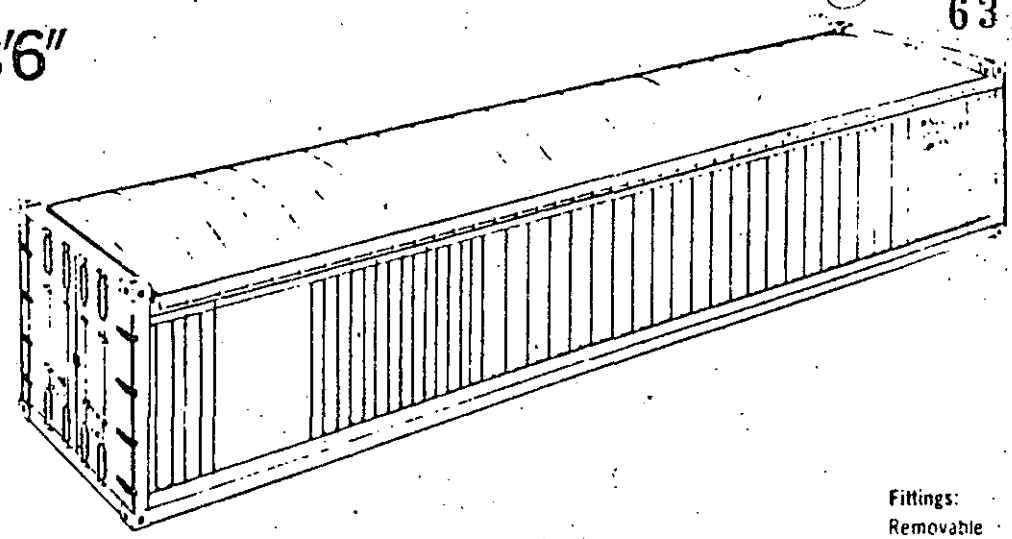


Manufactured according to ISO and ASA recommendations and standards.
Approved by Germanische Lloyd and/or American Bureau of Shipping.
Certified for inland transport under TIR approval.

Fittings:
Removable headbars,
tarpaulins

40'×8'×8'6" OpenTop

SIN TECHO

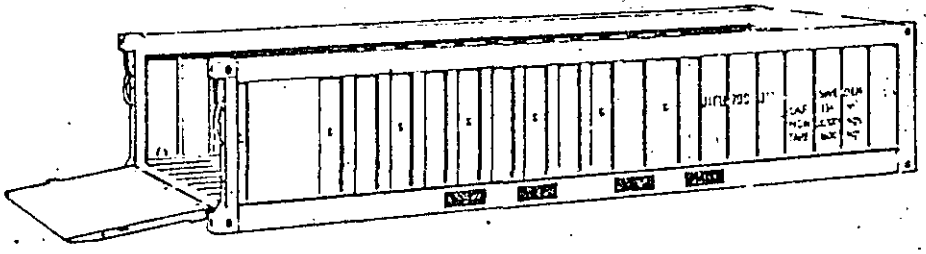


Fittings:
Removable headbars, tarpaulins and lift out bows

Manufactured according to ISO and ASA recommendations and standards.
Approved by Lloyds Register of Shipping.
Certified for inland transport under TIR approval.

20'x8'x4' Bin

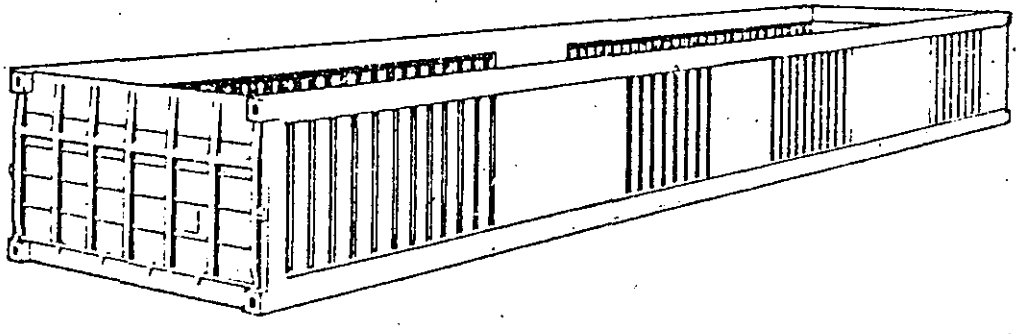
MEDIO CONTENEDOR



Manufactured according to ISO and ASA recommendations and standards.
JLCU containers approved by Germanische Lloyd and/or American Bureau of Shipping.
BSLU and EACU containers approved by Lloyds Register of Shipping.
Not certified for inland transport under TIR approval.

40'x8'x4' Bin

MEDIO CONTENEDOR



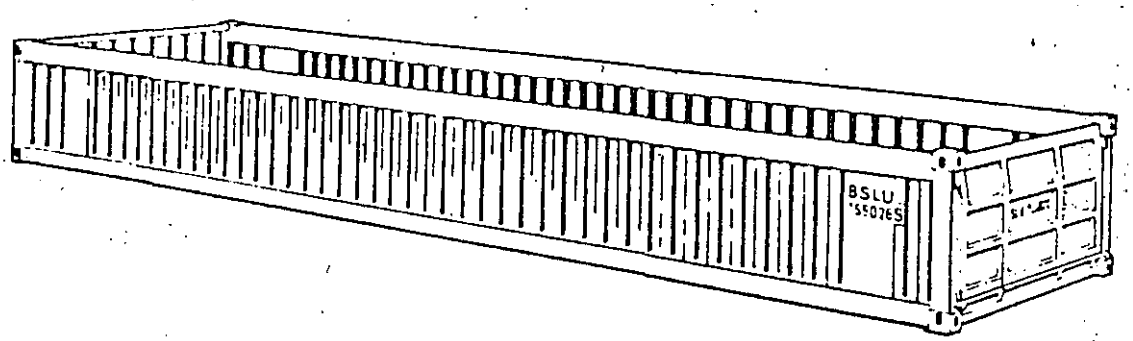
Manufactured according to ISO and ASA recommendations and standards.
Approved by Germanische Lloyd and or American Bureau of Shipping.
Not certified for inland transport under TIR approval.

64

40'x8'x4'3"
Bin

MEDIO

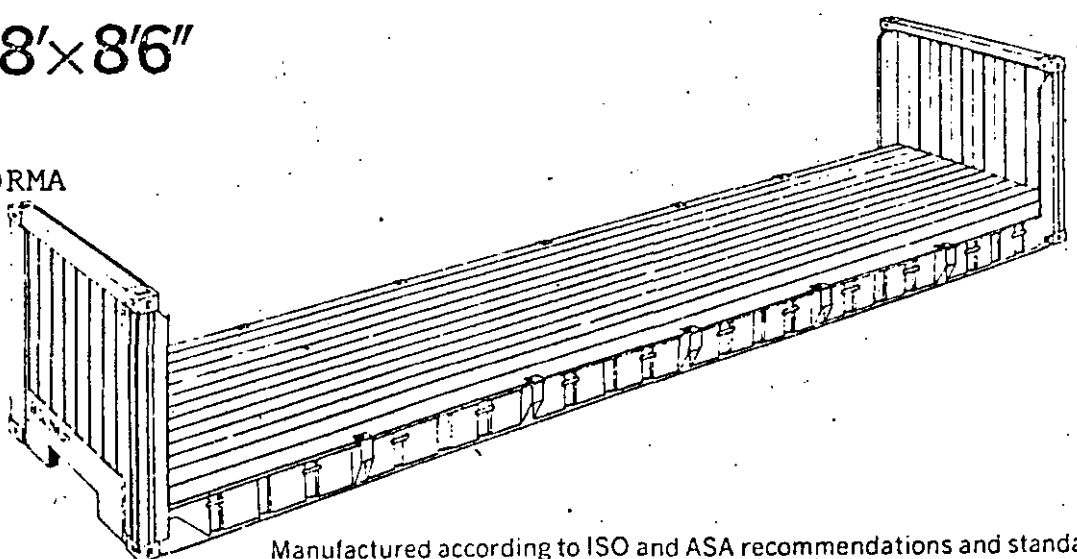
ONTENEDOR



Manufactured according to ISO and ASA recommendations and standards.
Approved by Lloyds Register of Shipping.
Not certified for inland transport under TIR approval.

40'×8'×8'6"
Flats

PLATAFORMA



Manufactured according to ISO and ASA recommendations and standards.
Approved by Lloyds Register of Shipping.
Not certified for inland transport under TIR approval.

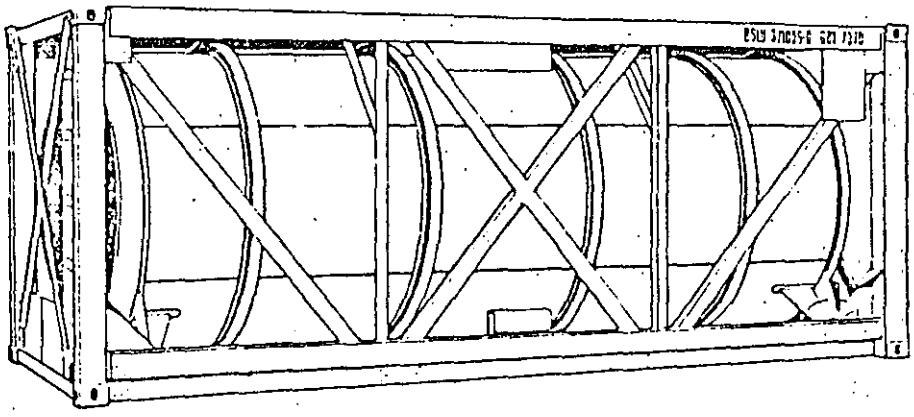
(65)

2.- CONTENEDORES CISTERNA .- PARA EL TRANSPORTE DE LIQUIDOS A GRANEL Y DE GAS COMPRIMIDO.

3.- CONTENEDORES TANQUE .- PARA CARGA SECA A GRANEL, DE DESCARGA A GRAVEDAD O POR PRESION.

20'x8'x8' Tank

TANQUE

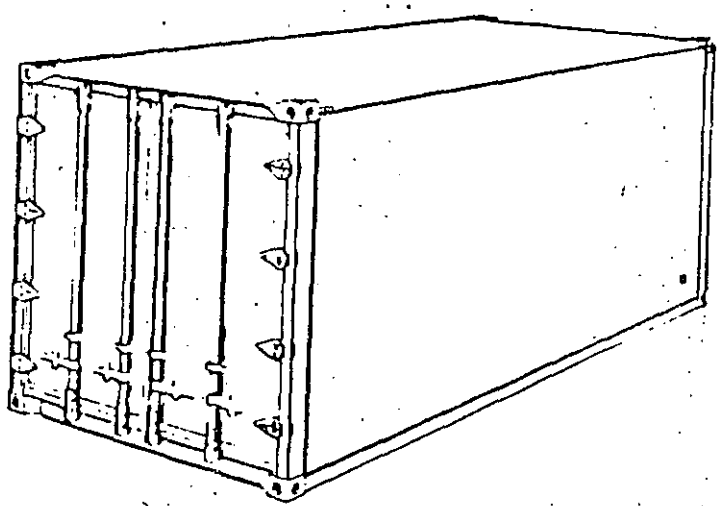


Manufactured according to ISO and ASA recommendations and standards.
JLCU containers approved by Germanische Lloyd and American Bureau of Shipping.
BSLU - EACU containers approved by Lloyds Register of Shipping.
Certified for inland transport under TIR approval.
DOT certificates: JLCU Nos. 6253, 6858. BSLU EACU No. 6500.

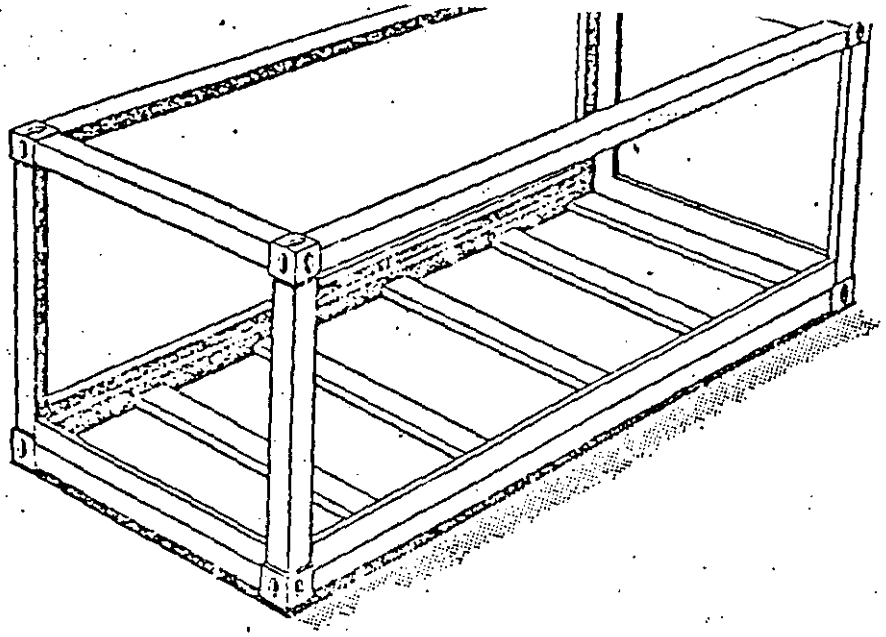
4.- CONTENEDORES ISOTERMO.

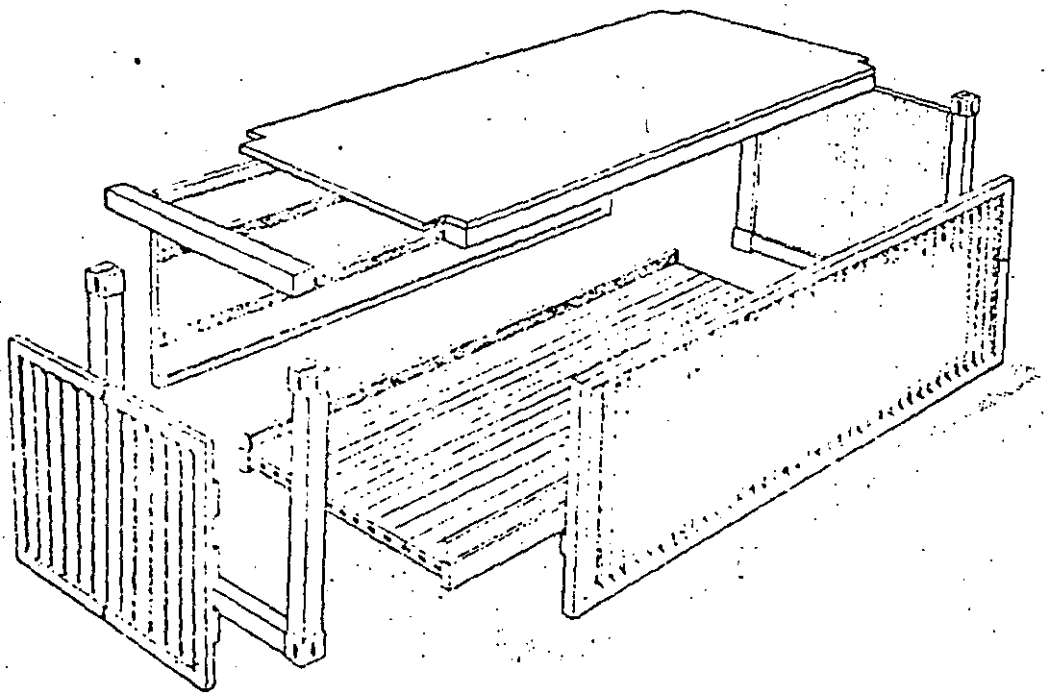
20'x8'x8' Insulated

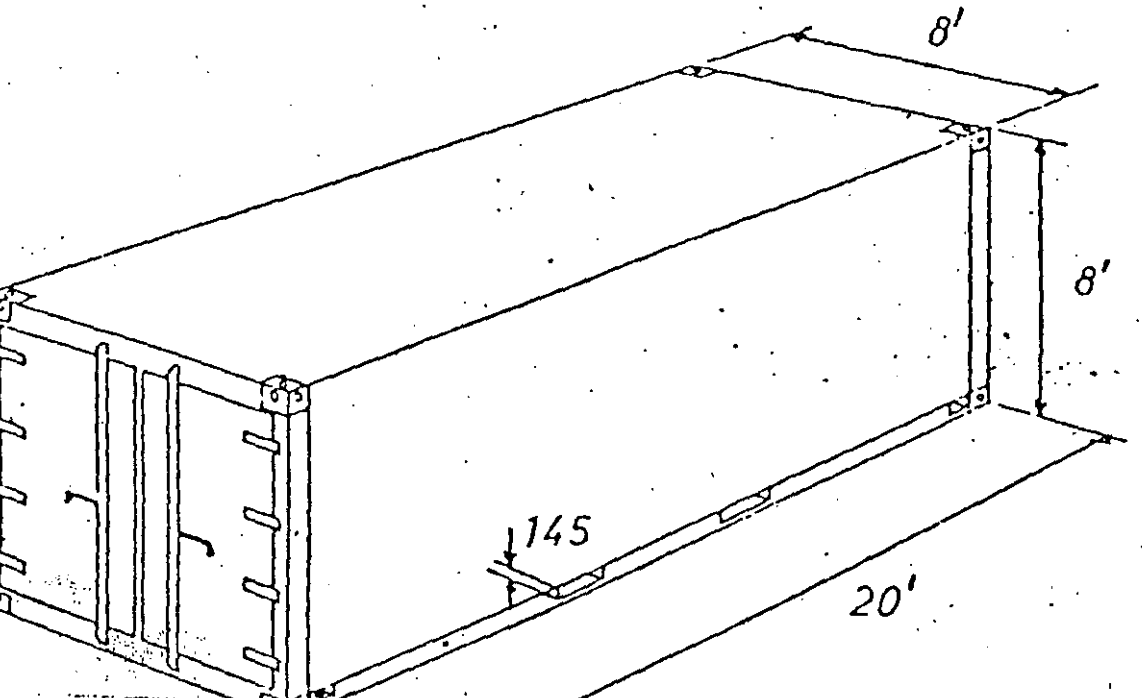
150 TERMOS



Manufactured according to ISO and ASA recommendations and standards.
JLCU containers approved by American Bureau of Shipping.
BSLU. EACU containers approved by Lloyds Register of Shipping.
Certified for inland transport under customs seal.

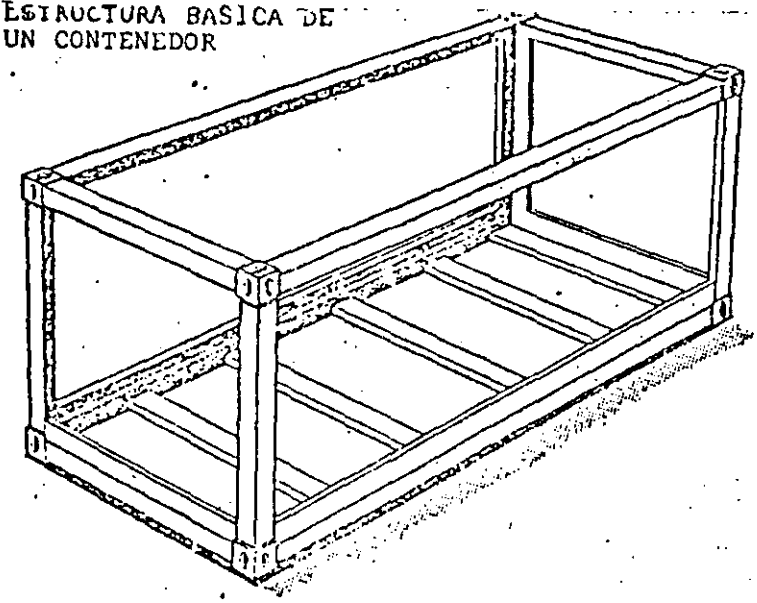


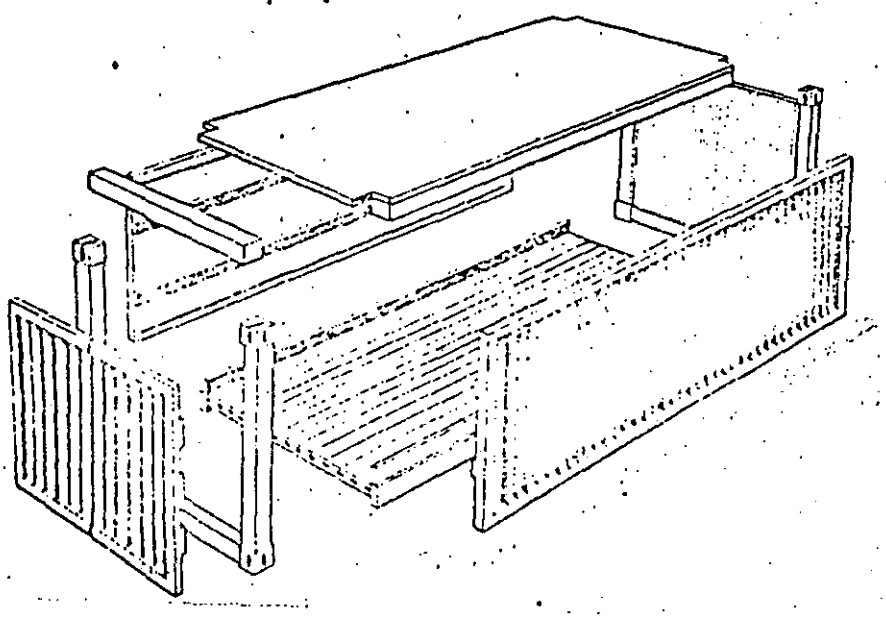


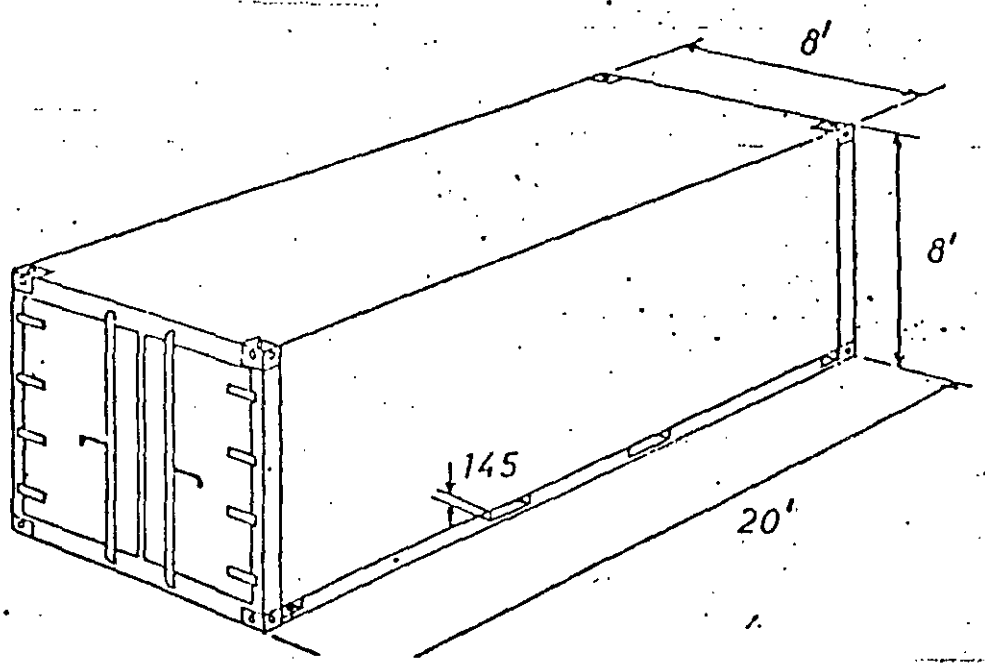

8'
8'
145
20'

5.- CONTENEDORES ESPECIALES.- PLEGABLES, PARA GANADO Y CON PERFORACIONES PARA PIERNAS DE SOPORTE.

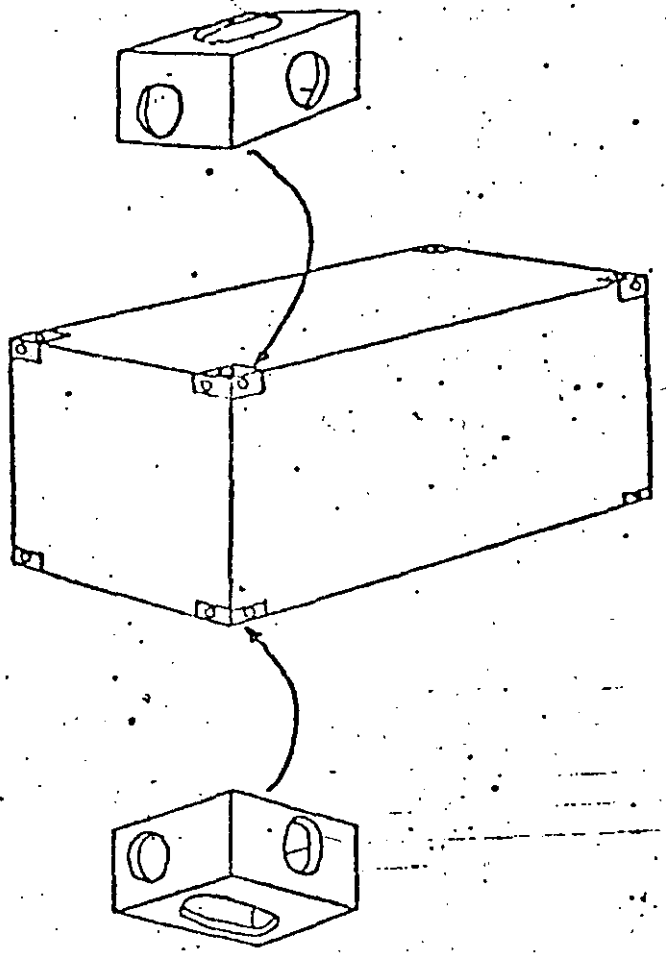
ESTRUCTURA BASICA DE UN CONTENEDOR

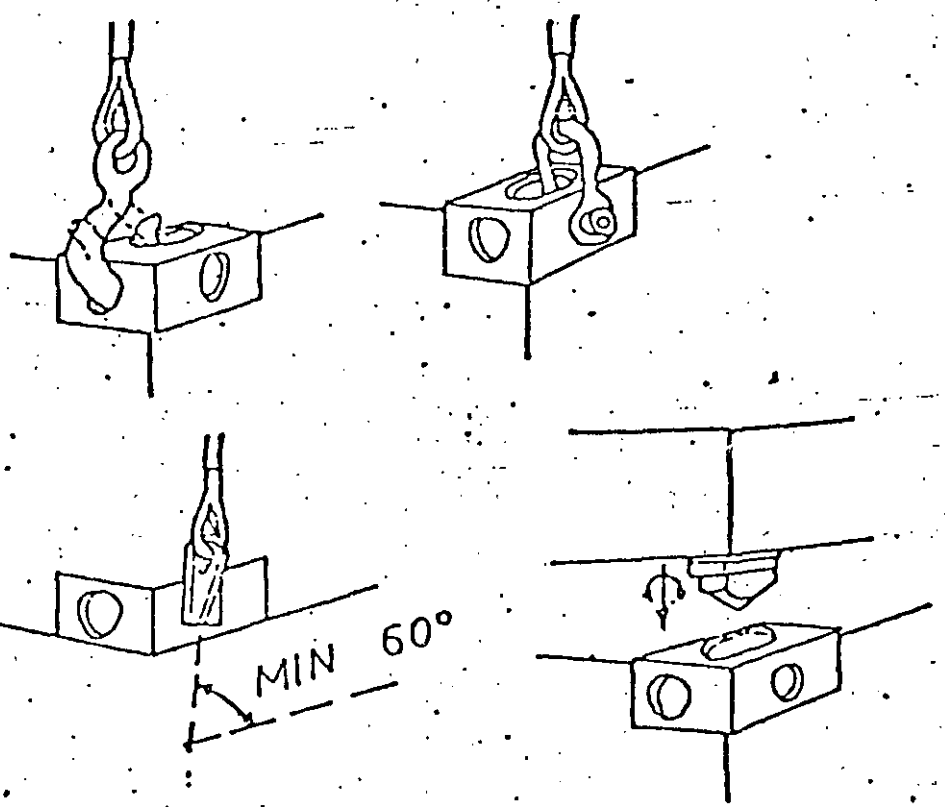




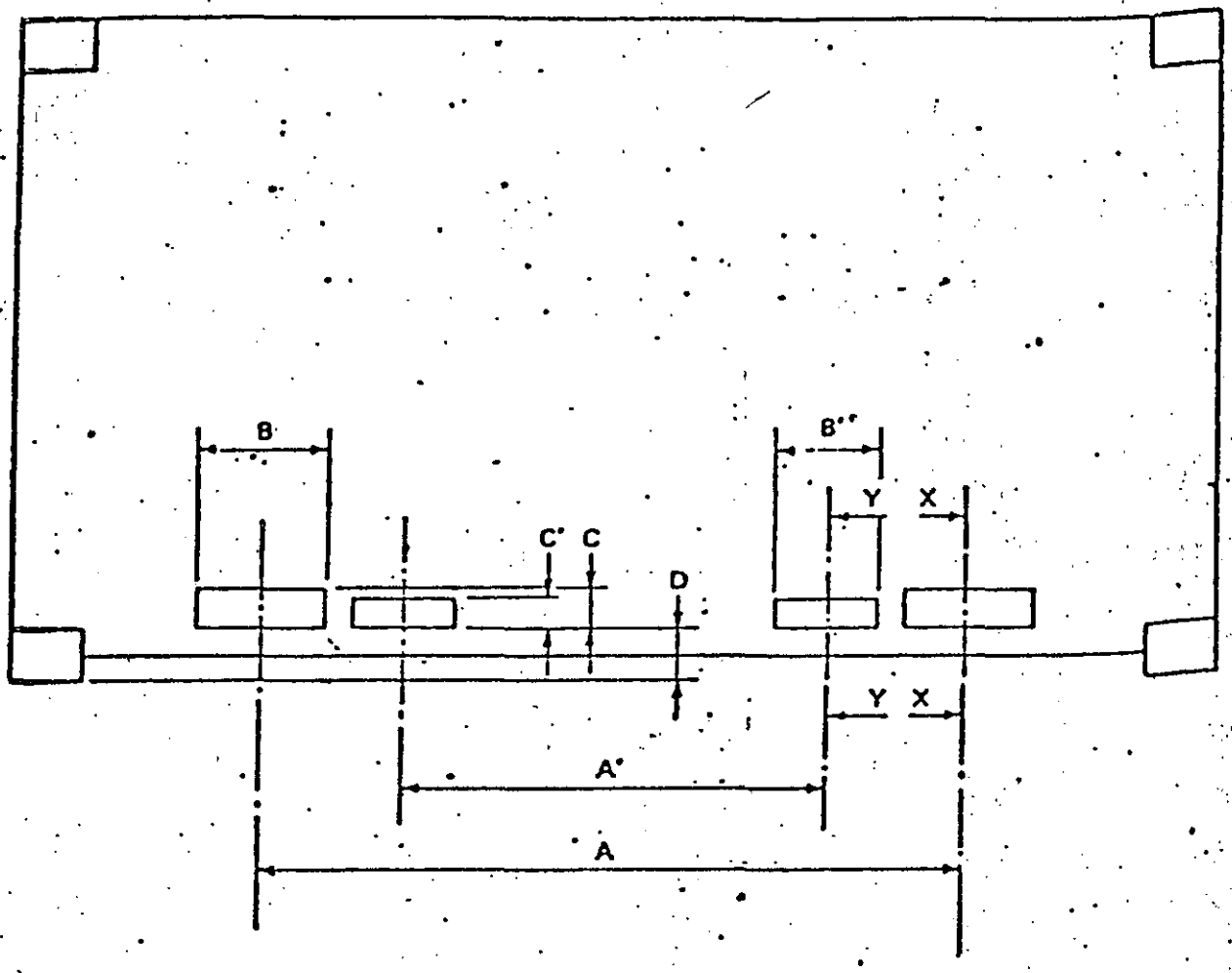


SISTEMA PARA IZAJE Y FIJACION DE CONTENEDORES EN CUBIERTA DEL BARCO.





DIMENSIONES DE TUNELES PARA HORQUILLAS DE MONTACARGAS



SECCION SECCION

CONTENE-DOR	DIMENSIONES													
	TUNELES PARA CARGA/DESCARGA DE CON TENEDORES CARGADOS								TUNELES PARA CARGA/DESCARGA DE CONTENEDORES VACIOS					
	mm				in				mm			in		
	A	B	C	D	A	B	C	D	A'	B'	C	A'	B'	C'
12C 12	2 050 ± 50	355 min.	115 min.	20 min.	81 ± 2	14 min	4½ min.	0 8 min.	900 ± 50	305 min.	102 min.	35½ ± 2	12 min	4 min.
10	900 ± 50	305 min.	102 min.	20 min	35½ ± 2	12 min	4 min.	0 8 min.						

NOTE - C = ALTURA LIBRE

Distr.
GENERAL

TD/B/AC.15/19/Corr.1
8 de diciembre de 1975

ESPAÑOL
Original: INGLES

Conferencia de las Naciones Unidas sobre Comercio y Desarrollo

JUNTA DE COMERCIO Y DESARROLLO
Grupo Preparatorio Intergubernamental para un Convenio sobre el transporte intermodal internacional
Tercer período de sesiones
Ginebra, 16 de febrero de 1976
Tema 2 a) del programa provisional

ASPECTOS TECNICOS Y FINANCIEROS DE LAS TECNOLOGIAS MODERNAS DE TRANSPORTE UTILIZADAS EN LAS OPERACIONES DE TRANSPORTE MULTIMODAL

Informe de la secretaría de la UNCTAD

Corrección

Anexo

Página 5, pregunta 10, cuadro del texto (1972)

Donde dice "2 400 TEU" debe decir "3 000 TEU".

Página 10, pregunta 19, último párrafo

Donde dice "Deben manipularse" debe decir "Pueden manipularse".

Página 19, pregunta 44, línea 1

Donde dice "por metro cuadrado útil" debe decir "por metro útil".

Página 35, pregunta 68, línea 5

Suprímase "parcialmente".

Distr.
GENERAL

TD/B/AC.35/15
24 de septiembre de 1975

ESPAÑOL
Original: INGLES

Conferencia de las Naciones Unidas sobre Comercio y Desarrollo

JUNTA DE COMERCIO Y DESARROLLO
Grupo Preparatorio Intergubernamental para un Convenio sobre el transporte intermodal internacional
Tercer período de sesiones
Ginebra, 16 de febrero de 1976
Tema 2 a) del programa provisional

ASPECTOS TECNICOS Y FINANCIEROS DE LAS TECNOLOGIAS MODERNAS DE TRANSPORTE UTILIZADAS EN LAS OPERACIONES DE TRANSPORTE MULTIMODAL

Informe de la secretaría de la UNCTAD

1. En los períodos de sesiones primero y segundo del Grupo Preparatorio Intergubernamental para un convenio sobre el transporte intermodal internacional se hizo patente que la falta de información y de datos adecuados sobre los aspectos técnicos y financieros de las tecnologías modernas de transporte utilizadas en las operaciones de transporte multimodal constituía para los países en desarrollo un problema de capital importancia. El presente informe ha sido preparado para ayudar a los países en desarrollo a tomar decisiones, en relación con los trabajos del Grupo y en el orden práctico, con respecto a la unitarización en sus puertos.

2. Para que la presentación de la información y los datos pertinentes fuera lo más clara y sucinta posible, la secretaría de la UNCTAD consideró que lo más apropiado sería presentar un texto redactado en forma de una serie de preguntas y respuestas. Por consiguiente, la secretaría encargó al Sr. A. Behnam, del Irak, que preparase las respuestas a una serie de preguntas relativas a los tipos principales y el costo del material utilizado en las técnicas modernas de transporte aplicadas en las operaciones de transporte multimodal.

3. La lista de preguntas y respuestas figura en el anexo al presente informe y abarca el transporte marítimo -incluidos los buques- y los puertos, así como el transporte por ferrocarril, por carretera, por vías de navegación interior y por vía aérea. Las preguntas giran en torno a los distintos métodos de unitarización de la carga que, si bien en principio no son necesarios para que pueda haber transporte multimodal, contribuyen a que se obtengan las máximas ventajas de esa forma de transporte.

Anexo I

PREGUNTAS Y RESPUESTAS SOBRE LOS ASPECTOS TECNICOS Y FINANCIEROS DE LAS TECNOLOGIAS MODERNAS DE TRANSPORTE UTILIZADAS EN LAS OPERACIONES DE TRANSPORTE MULTIMODAL

INDICE

SIGLAS Y ABREVIATURAS

ANSI	American National Standards Institute
CCC	carga completa de contenedor
CPC	carga parcial de contenedor
DUC	dispositivo de unidad de carga
ECC	Estación de carga de contenedores
FAK	flete uniforme para toda clase de mercancías
ISO	Organización Internacional de Normalización
LASH	buque portagabarras del tipo LASH
OTM	operador de transporte multimodal
Ro-Ro	buque de transbordo por rodadura
Seabee	buque portagabarras de tipo Seabee
T	toneladas
TM	transporte multimodal
TPM	toneladas de peso muerto
TEU	unidad equivalente a 20 pies (medida uniforme para expresar la cantidad de contenedores de diferentes tamaños)
UIC	Unión Internacional de Ferrocarriles (código)

6. Gran parte de la información suministrada en el anexo tiene un carácter puramente técnico y en rigor quizás no sea pertinente para los trabajos del Grupo Preparatorio Intergubernamental. Se incluye porque puede ser útil para los servicios técnicos de las entidades y organizaciones que se ocupan del transporte de los países en desarrollo cuando, en la práctica, tomen decisiones sobre unitarización.

7. Los datos de orden financiero proporcionados en el anexo al presente informe se refieren al costo por unidad de los principales tipos de materiales necesarios para la explotación de un servicio multimodal y no al costo global de todo el equipo. Este último variará según el número de unidades que se precisen, el cual, a su vez, dependerá de la escala de operaciones en cada país. Tampoco se facilita en el anexo información acerca del costo total de la nueva infraestructura o de las mejoras de la infraestructura existente que requieren los puertos y los sistemas de transporte interior, costo que también variará según los países, pero se proporciona información acerca de los niveles mínimos que es menester establecer.

8. De todos modos, la información que figura en el anexo pone claramente de manifiesto que la introducción de sistemas de unidad de carga, especialmente la contenedorización en un sistema de puerta a puerta, requiere cuantiosas inversiones de capital[1/]. Las inversiones de tal magnitud tendrían que ser costeadas en divisas y, sin duda, plantearían graves problemas financieros a la mayoría de los países en desarrollo. Además, tendrían que competir con las inversiones en los sectores agrícola e industrial y en otros sectores prioritarios de la economía de esos países.

9. Las observaciones que anteceden ponen de manifiesto la necesidad de que los países en desarrollo procedan con cautela al tomar una decisión acerca de la introducción de esos sistemas de unidad de carga más complejos que permiten llevar a cabo operaciones multimodales. De estas observaciones se desprende además la necesidad de que los gobiernos y las instituciones financieras de los países desarrollados, así como los organismos financieros internacionales, incluidos el Banco Mundial y los bancos regionales de desarrollo, faciliten a los países en desarrollo ayuda o créditos en condiciones preferenciales en todos aquellos casos en que se considere que en un país es oportuno implantar operaciones multimodales que ofrecerán ventajas económicas generales a todos los países participantes en una corriente de tráfico determinada.

1/ A este respecto, conviene señalar que a esos desembolsos de capital hay que añadir los gastos de explotación y, en especial, los dependientes por concepto de conservación, reparaciones y repuestos, que, habida cuenta de la fragilidad de la mayor parte del material descrito en el anexo, es probable que asciendan a sumas considerables.

A. TECNOLOGIAS MODERNAS DE TRANSPORTE Y OPERACIONES MULTIMODALES

1. ¿Qué es el transporte multimodal internacional?

El transporte multimodal internacional consiste en el porte de mercancías de un país a otro por dos o más modos de transporte en virtud de un contrato único -el documento o contrato de transporte multimodal (TM)- expedido por la persona o la empresa que organiza ese servicio -el operador de transporte multimodal (OTM)- y que asume además la responsabilidad de la ejecución de toda la operación. El sistema contrapuesto al transporte multimodal es el transporte fragmentado.

2. ¿Cuáles son las características principales del transporte multimodal?

Lo que caracteriza al transporte multimodal es la naturaleza de la relación contractual entre el OTM y el cargador. El OTM actúa como entidad con personalidad jurídica propia que ofrece a los cargadores un contrato único para el porte de mercancías por más de un modo de transporte.

3. ¿Qué se entiende por transporte de puerta a puerta?

El porte completo de la mercancía desde el local del expedidor hasta el local del destinatario.

4. ¿Ha de ser el transporte multimodal necesariamente de puerta a puerta?

No. Puede ser de muelle a puerta, de muelle a almacén en el interior, de puerta a muelle, de puerta a almacén en el interior y de almacén en el interior a almacén en el interior. Ahora bien, las ventajas pueden ser mayores si toda la operación de transporte se planea y realiza como una operación única de puerta a puerta.

5. ¿Qué modos de transporte pueden utilizarse en el transporte multimodal?

Todos los modos de transporte, es decir, por vía marítima, por carretera, por ferrocarril, por vía aérea o por vías de navegación interior, que sean compatibles con las unidades de mercancías acarreadas.

6. ¿El transporte multimodal, ¿supone automáticamente la unitarización de la carga?

No. En teoría, el transporte multimodal es compatible también con la carga fraccionada. Sin embargo, como el movimiento de carga fraccionada entraña operaciones de manipulación de la carga en cada punto de transbordo de la cadena de transporte, suele ser difícil que un solo operador asuma la responsabilidad de toda la operación y expida un documento único de transporte directo. Es ahí que el transporte multimodal internacional ofrezca su máxima utilidad en relación con el transporte unitarizado; de las diversas formas de unitarización, el sistema de contenedores es el que mejor se presta al transporte multimodal internacional, ya que permite la máxima protección contra daños y hurtos.

B. OPERACIONES MULTIMODALES Y TRANSPORTE MARITIMO

7. ¿Qué tipos de buques se utilizan en el transporte multimodal?

Como el transporte multimodal es compatible, al menos en principio, con todas las tecnologías de transporte, se pueden utilizar buques de tipo tradicional o buques para carga unitarizada. De estos últimos, los tipos principales son los siguientes:

1. Buques portacontenedores:
   a) exclusivamente celulares
   b) parcialmente celulares
   c) Ro-Ro celulares
   d) portagabarras con bodegas celulares
   e) celulares frigoríficos
2. Buques de transbordo por rodadura (buques Ro-Ro)
3. Buques portagabarras
4. Buques portapaletas
5. Buques de carga fraccionada transformados
6. Buques graneleros
7. Buques de carga general polivalentes

8. ¿Qué son los buques celulares?

Buques especialmente destinados al transporte de contenedores y cuyas bodegas están provistas de células permanentes dotadas de guías verticales por las que pueden deslizarse los contenedores de modo que éstos queden firmemente apilados y se encuentren sujetos en todas las esquinas.

9. ¿Qué es el buque de transbordo por rodadura?

Es un tipo de buque en el que la mercancía se traslada por rodadura y no por elevación. La carga puede estibarse en las cubiertas del buque o permanecer sobre material rodante durante la travesía.

10. ¿Cuáles son las principales dimensiones de algunos buques portacontenedores?

Las principales dimensiones de los buques portacontenedores han variado con el tiempo conforme ha ido evolucionando la tecnología. Hay, pues, varias "generaciones" de buques portacontenedores:

Buques celulares

Año de entrada en servicio	Capacidad de transporte de contenedores	Peso muerto	Eslora total	Manga total	Calado
Primera generación (1963)	750 TEU	12 000 T	175 m	25 m	8 m
Segunda generación (1970)	1 500 TEU	30 000 T	225 m	29 m	11 m
Tercera generación (1972)	2 400 TEU	35 000 T-50 000 T	275 m	32 m	12

Velocidad: 20-25 nudos (algunos alcanzan los 33 nudos)

Buques típicos de carga general y contenedores

Capacidad de transporte de contenedores	Eslora total	Manga total	Calado
792 TEU	202 m	23 m	10 m
233 TEU	171 m	23 m	9
247 TEU	150 m	22 m	8 m

Velocidad: 16-20 nudos

Buques LASH (buques portadores)a/
(Clase Delta Mar)

Capacidad de transporte de contenedores o gabarras	Peso muerto	Eslora total	Manga total	Calado
288 TEU y 74 gabarras LASH	41 000 T	272 m	30,5 m	11,62 m

Velocidad: 22,5 nudos

---

a/ El buque portagabarras tipo LASH tiene una popa ancha en forma de U provista de unas estructuras que sobresalen del casco y sirven para las operaciones de carga y descarga. Para descargar las gabarras se utiliza una grúa-pórtico de 510 toneladas montada sobre railes. El buque es del tipo de cubierta única con la superestructura a proa.

Buques Seabee[b/]

Nº de gabarras	Peso muerto	Eslora total	Manga	Calado
38	39 000 T	267 m	32,36 m	11,90 m

Velocidad: 20 nudos

Ro-Ro (pequeño)

Capacidad de transporte	Eslora	Manga	Calado (cargado)	Toneladas de registro neto
5 TEU cubierta de vehículos 24 TEU cubierta superior	110 m	17 m	4 m	1 072

Ro-Ro con rampa en la popa

Capacidad de transporte de contenedores	Peso muerto	Eslora	Manga	Calado
1 200 TEU	20 650 T	199 m	287 m	9,6 m

Velocidad: 22 nudos

Buque celular Ro-Ro
(Clase Atlantic Champagne)

Capacidad de transporte	Eslora	Manga	Calado	Peso muerto
845 TEU y 990 vehículos	212 m	28 m	9,3 m	18 850 T

Velocidad: 24 nudos

---

b/ El buque portagabarras tipo Seabee tiene tres cubiertas de carga continuas. Para izar las gabarras se utiliza una plataforma sumergible, con una capacidad de 2.000 toneladas, que se hace descender dentro del agua. Pueden estibarse 12 gabarras en la cubierta inferior, otras 12 en la cubierta principal y 14 en la cubierta superior.

11. ¿Cuáles son las instalaciones de manipulación a bordo de los buques portacontenedores?

Algunos buques portacontenedores no van provistos de aparejos de carga. En los demás, es posible cualquier combinación del siguiente material de manipulación:

Pescantes de carga giratorios de gran potencia: pescantes de carga moderno provisto de dos amantillos gemelos que permiten un movimiento giratorio y basculante sin tensiones excesivas.

Grúas de cubierta gemelas: grúas montadas sobre una plataforma giratoria común de tal manera que pueden funcionar juntas o independientemente.

Grúa-pórtico: aparato de elevación móvil cuyos pies se desplazan a lo largo de cabas bandas del buque. La viga entre ambos pies sostiene un carrillo que lleva el material de izada y que está dotado de un movimiento de traslación por el través del buque.

Montacargas (buques portagabarras): mecanismo de elevación hidráulico situado en la popa del buque que permite la carga o descarga de los contenedores (en gabarras).

Estiba mecanizada: dispositivos automatizados de las bodegas para facilitar el desplazamiento horizontal y la estiba de los contenedores.

Vehículos de plataforma baja: tipo de remolque que permite la máxima utilización del espacio en los buques Ro-Ro.

Instalaciones de acceso: tales como rampas en la popa o la aleta de los buques Ro-Ro que permiten el acceso del equipo de transbordo.

Carretillas-pórtico de silueta baja: equipo utilizado en los buques Ro-Ro para embarcar y desembarcar contenedores, grandes plataformas y otras grandes unidades.

12. ¿Cuál es el costo indicativo de algunos de esos buques? c/

Buque portacontenedor para el transporte de 800 contenedores de carga seca y 400 contenedores frigoríficos (25.000 TPM-22 nudos)	50,6 millones de dólares

c/ Datos obtenidos por cortesía de los editores del Fairplay Journal; nivel de precios a mediados de 1975 o según las estimaciones publicadas en la sección "Venta de buques" del Fairplay Journal, números de febrero a junio de 1975. Los precios de las gabarras LASH y Seabee han sido tomados de Containerization International Yearbook, 1974.

Buque de carga general polivalente con una capacidad de transporte de contenedores de 300 TEU (14.000 a 15.000 TPM-16 nudos)	14,4 millones de dólares
Buque de línea regular de segunda mano apto para su transformación y cuyo año de construcción se sitúa entre fines del decenio de 1950 y fines del decenio de 1960 (13.000 a 20.000 TPM)	4 a 9 "
Buq. LASH	50 millones de dólares, aproximadamente
Gabarras (LASH)	75.000 dólares por unidad
Buque Seabee (Astilleros Vallmet Helsinki) contratado a principios de 1975	53,1 millones de dólares
Gabarras Seabee	100.000 dólares, aproximadamente, por unidad

13. ¿Es posible transformar los buques de carga general corrientes para el transporte de carga unitarizada?

Sí. Los buques de carga general corrientes pueden ser adaptados para el transporte de contenedores, paletas y unidades preeslingadas o convertidos en buques Ro-Ro. Para estas transformaciones, los buques más adecuados son quizás los buques de línea de 10.000 a 13.000 TPM construidos a mediados del decenio de 1960.

14. ¿Cuál es el costo de una transformación de este tipo?

El costo de la transformación depende de múltiples factores, entre otros de la clase de buque, su edad y estado, el lugar en que se efectúe la transformación, el momento en que se lleve a cabo y las modificaciones accesorias realizadas.

15. ¿Qué es una paleta?

Una paleta es una tarima o plataforma de alrededor de 32 x 48 ó 40 x 48 pulgadas, generalmente de madera, en la que se puede colocar cierto número de bultos para formar una unidad de carga con objeto de proceder a su transporte, manipulación o apilamiento por medios mecánicos. La paleta puede tener o no una superestructura y, según sus

características, puede utilizarse repetidamente o ser desechable. El término paletas comprende las paletas planas, las paletas-caja y las paletas con montantes:

- La paleta plana es una simple base de uno o dos pisos sin ninguna superestructura;
- La paleta-caja es una paleta plana con una estructura en forma de caja, cuyos lados pueden ser de red o macizos y que puede tener o no una cubierta;
- La paleta con montantes es una paleta plana con montantes de esquina ajustables y largueros desmontables.

16. ¿Qué paletas normalizadas se utilizan en el transporte multimodal?

Las recomendaciones de la ISO[d/] sobre las dimensiones normales de las paletas de doble piso para el transporte directo de mercancías y de paletas de mayor tamaño para los mismos fines, son las siguientes:

Pulgadas	Milímetros
32 x 48	800 x 1 200
40 x 48	1 000 x 1 200
32 x 40	800 x 1 000
48 x 69	1 200 x 1 600
48 x 72	1 200 x 1 800

17. ¿Qué es un contenedor?

Se ha dado del contenedor la siguiente definición internacional[e/]:

un elemento del equipo de transporte

a) de carácter permanente y por tanto suficientemente resistente para permitir su uso repetido;
b) especialmente ideado para facilitar el porte de mercancías por uno o varios modos de transporte, sin manipulación intermedia de la carga;
c) provisto de dispositivos que permitan su fácil manejo y, en particular, su transbordo de un modo de transporte a otro;
d) diseñado de manera que sea fácil de llenar y vaciar;
e) de un volumen interior de un metro cúbico (35,3 pies cúbicos), por lo menos.

El término contenedor no comprende los vehículos ni los embalajes de tipo corriente

---

d/ ISO 198 1961 y R.339 1963

e/ ISO 668.

18. ¿Cuáles son las características materiales de los contenedores?

Si bien los contenedores deben ser de construcción rígida, algunos son plegables, o pueden ser desmontados y luego ser montados nuevamente, en tanto que otros están montados de modo permanente. Pueden ser de acero, aluminio, madera contrachapada o fibra de vidrio o de una combinación de esos materiales. El contenedor puede tener una puerta en un extremo o en una pared lateral o aberturas en su parte superior para su llenado y vaciado. Los principales tipos de contenedores que se emplean actualmente son los de 20 pies, con un peso bruto máximo de 20 toneladas y los de 40 pies, con un peso bruto máximo de 40 toneladas. Debido a su estanqueidad, los contenedores protegen la carga de la intemperie.

19. ¿Cuáles son los diferentes tipos de contenedores?

Los contenedores pueden clasificarse en seis tipos principales:

Contenedores de carga general	-	comprenden los contenedores cerrados con puertas en un extremo y en las paredes laterales; los de techo abierto; los de paredes laterales abiertas; los de paredes laterales y techo abiertos; los de paredes laterales, techo y extremo abiertos; los de media altura y los ventilados (no isotermos);
Isotermos	-	contenedores aislantes, refrigerados o con calefacción;
Contenedores cisterna	-	para el transporte de líquidos a granel y de gas comprimido;
Contenedores para carga seca a granel	-	de descarga por gravedad o descarga a presión;
Contenedores plataforma	-	esencialmente contenedores planos sin superestructura que no forman parte de los sistemas plenamente automatizados de transporte en contenedores y que carecen de parte superior por donde puedan ser izados con su carga;
Contenedores especiales	-	contenedores para ganado y contenedores plegables;
Contenedores planos	-	los contenedores planos con en realidad grandes paletas, con paredes o sin ellas. Deben manipularse con grúas y están dotados de bordes especiales para tal fin. Los contenedores planos con paredes plegables pueden estibarse cuando son desmontados o sin carga.

20. ¿Cuál es el costo aproximado de los contenedores? f/

Contenedor de acero	20 pies	1 900 a 2 500 dólares
Contenedor de acero	40 pies	3 900 a 4 500 dólares
Contenedor refrigerado	20 pies	4 500 a 7 000 dólares
Contenedor refrigerado	40 pies	11 000 a 14 000 dólares

21. ¿Pueden transportarse los contenedores por todos los modos de transporte?

En general, sí. No obstante, los contenedores transportados por buques se ajustan generalmente a las normas de la ISO y no son compatibles con el transporte aéreo.

22. ¿Cuáles son los diversos sistemas de manipulación de los contenedores?

1. Sistema ordinario de manipulación
2. Sistema de transbordo por elevación (lift-on/lift-off)
3. Sistema de transbordo por rodadura (roll-on/roll-off)
4. Sistema de transbordo por flotación (float-on/float-off)

¿En qué consiste el sistema ordinario de manipulación de contenedores?

Este sistema se suele emplear cuando el movimiento de contenedores es limitado, especialmente cuando éstos son transportados en buques de línea corrientes. Los contenedores pueden cargarse en el buque empleando los aparejos propios de éste, y la descarga en el muelle se efectúa con estos aparejos o con grúas instaladas en tierra. Los contenedores pueden ser transbordados directamente del buque a vagones de ferrocarril o vehículos de carretera, o bien trasladados con camiones plataforma o remolques de muelle a una zona de almacenamiento cercana. La carga y descarga de los vagones de ferrocarril y vehículos de carretera puede efectuarse con grúas móviles, grúas pórtico, carretillas de horquilla elevadora y otro material análogo.

24. ¿Qué es un sistema de transbordo por elevación?

Se trata de un sistema usual cuando el tráfico de contenedores es regular y considerable. Los contenedores se descargan del buque utilizando los aparejos de éste o bien grúas de gran potencia instaladas en tierra; en la mayor parte de los casos los contenedores se colocan directamente en un camión plataforma, en un chasis semirremolque arrastrado por un tractor de carretera ordinario, o en un tren de chasis consistente en un tractor que arrastra cierto número de contenedores hasta la zona de

f/ Cotizaciones de fabricantes de la República Federal de Alemania, sujetas a fluctuaciones (febrero de 1975).

almacenamiento. Una vez en ella los contenedores se dejan sobre los chasis y pueden ser retirados directamente por los tractores de carretera para su remolque hasta el punto de destino definitivo; también pueden ser trasladados a una zona de maniobra donde, por medio de una grúa móvil, sea posible despacharlos en vagones de ferrocarril, gabarras o servicio de enlace de cabotaje.

Con este sistema también se pueden descargar los contenedores directamente sobre el muelle, donde una carretilla-pórtico los recoge uno por uno y los transporta a la zona de almacenamiento, desde donde son trasladados luego, por medio de carretillas-pórtico o de una grúa-pórtico, a camiones o vagones de ferrocarril. Para la manipulación en el muelle cabe emplear carretillas-pórtico, carretillas de elevación lateral, carretillas de horquilla elevadora o carretillas en U (la carretilla en U se emplea para levantar y transportar contenedores: éstos quedan sujetos a los brazos en forma de U de la carretilla por medio de un sistema de enganche especialmente adaptado a las piezas de esquina ordinarias de los contenedores).

25. ¿Qué es el sistema de transbordo por rodadura?

En este sistema se emplean buques especialmente diseñados que permiten efectuar las operaciones de carga y descarga conduciendo directamente a bordo o a tierra, a través de portalones laterales, de proa o de popa, y de las rampas del buque, el equipo de tracción o equipo de manipulación de poca altura, como por ejemplo carretillas de horquilla elevadora. Para el transbordo de la carga (contenedores) del buque al muelle, es posible utilizar también todos los sistemas de manipulación empleados en el transbordo por elevación. Los contenedores montados en semirremolques pueden conducirse directamente a bordo del buque o a tierra. Para el transbordo de contenedores con este sistema cabe emplear también carretillas de elevación lateral provista de bastidores de suspensión u horquillas. El material de manipulación suele ser de silueta baja debido a las limitaciones de altura en los buques.

(El bastidor de suspensión está suspendido de una grúa. El bastidor se apoya en la parte superior de un contenedor y por medio de un sistema de cerrojos giratorios sujeta al contenedor para izarlo.)

26. ¿Qué es el sistema de transbordo por flotación?

En este sistema un buque portagabarras transporta gabarras de 100 a 500 TPM. El buque puede izar las gabarras a bordo o descenderlas con una grúa o una plataforma elevadora. Las gabarras descendidas desde el buque pueden ser arrastradas por remolcadores. Los contenedores transportados en gabarras se transbordan de éstas al muelle con grúas; su acarreo en el terminal suele hacerse con el mismo equipo que se utiliza

27. ¿Cuáles son los diferentes tipos de remolques que se emplean en la manipulación y el transporte de contenedores?

Remolque plataforma:	Un chasis con una plataforma de madera y generalmente con un tablero delantero.
Remolque de chasis sencillo:	Bastidor sobre ruedas en el que se colocan los contenedores para transportarlos por carretera.
Semirremolque:	Remolques sin ruedas delanteras. La parte delantera se apoya sobre la unidad de tracción o, cuando está separada de ésta, en patas de soporte.
Remolque con ejes en tándem:	Remolque o semirremolque dotado de ejes dobles.
Chasis plano:	Chasis remolque con superficie totalmente plana, sin los rebordes de los remolques corrientes.
Chasis de carretera:	El remolque especial en que se colocan los contenedores para ser transportados por un vehículo o una unidad de tracción de carretera.
Remolque bajo:	Remolque de silueta baja, empleado en los buques Ro-Ro.
Semirremolque de toldo:	Semirremolque de costados bajos y pared posterior abatible. Provisto de anillas metálicas y lona impermeabilizada.
Semirremolque extensible:	Semirremolque que puede extenderse, en general de 40 a 60 pies; tiene ranuras y clavijas de cierre simétricas.
Semirremolque de volquete:	Semirremolque con dispositivo neumático de descarga. Levantando un extremo del contenedor, permite volcar la carga a granel en silos, sin necesidad de emplear otro equipo de manipulación en los puntos de entrega.
Bogie:	Juego de ruedas especialmente destinadas a emplearse como ruedas traseras bajo los contenedores o chasis.
Tren delantero:	Juego de ruedas que puede colocarse en la parte delantera de un semirremolque para convertirlo en un remolque de cuatro ruedas.

Existen modelos adaptados a los diversos números y longitudes de contenedores del sistema modular de la ISO.

Los semirremolques especiales para el transporte de contenedores corresponden en general a una de dos categorías principales: los destinados a transportar contenedores en el perímetro de un terminal o en las operaciones de transbordo de un buque Ro-Ro, y los empleados para transportar contenedores por carretera.

28. ¿Cuáles son las características materiales de los chasis portacontenedores y de los chasis remolque portacontenedores para transbordo por rodadura?

Chasis portacontenedores

Tamaño	Altura (milímetros)	Peso bruto (milímetros)	Peso muerto (toneladas)	Carga útil (toneladas)	Carga por eje (toneladas)
Contenedor de 20 pies	1 350	24,3	3,75	20 551	16
Contenedor de 40 pies	1 360	31	4,63	26 371	20

Chasis remolque portacontenedores para transbordo por rodadura

Longitud	Ancho	Altura de la plataforma	Capacidad
6,055 a 12,19 m	2,435 m	0,515 a 0,826 m	20/39/48/55 T

29. ¿Cuál es el costo aproximado de esos remolques? E/

Chasis portacontenedores	20 pies	9.500 dólares aproximadamente
Chasis portacontenedores	40 pies	10.000 dólares aproximadamente
Remolque para transbordo por rodadura	20 a 25 pies	3.000 dólares aproximadamente
Remolque para transbordo por rodadura	40 a 60 pies	5.500 dólares aproximadamente

¿Qué es la agrupación de la carga?

Es un servicio que permite reunir en un solo contenedor, para su transporte, envíos pequeños o diferentes que separadamente no alcanzan a llenar un contenedor. Así, pues, las cargas que se agrupan son "cargas parciales de contenedor" (CPC).

31. ¿Qué son las "cargas parciales de contenedor" (CPC)?

Son los envíos de carga expedida a más de un destinatario o por más de un cargador o usuario que se transportan en un mismo contenedor. El contenedor puede llenarse con CPC en una estación de carga de contenedores para que sea entregada como carga completa de contenedor (CCC) a un destinatario. Es posible también que un cargador empaque la mercancía como CCC para ser entregada como CPC, o que la empaque como CPC para que sea entregada como CCC.

E/ Precios basados en cotizaciones solicitadas a fabricantes de la República Federal de Alemania, febrero de 1975. Las cotizaciones sólo indican un orden de magnitud y están sujetas a fluctuaciones.

32. ¿Debe el cargador o usuario pagar por una carga completa de contenedor aunque envíe sólo una carga parcial?

No, en tal caso el cargador paga un flete basado en un peso o volumen mínimo de utilización.

33. ¿Qué es una "carga completa de contenedor" (CCC)?

Es la enviada en un contenedor, generalmente cargado por un cargador o un agente de grupaje, a un destinatario, y por cuyo transporte se paga la tarifa de contenedor completo.

34. ¿Cómo se estiban los contenedores en los buques?

La estiba de los contenedores varía según el tipo de buque. En los buques corrientes de carga fraccionada, los contenedores se colocan unos junto a otros en sentido longitudinal, en la cubierta y sobre las tapas de las escotillas y se afianzan mediante dispositivos inmovilizadores. En los buques portacontenedores celulares, se apilan verticalmente en células con capacidad de hasta nueve contenedores. Las células están provistas de guías especiales que permiten el fácil deslizamiento de los contenedores. En la mayoría de los buques portacontenedores se apila del 25 al 35% de los contenedores sobre las tapas de las escotillas, que tienen dispositivos de anclaje adecuados. En los buques portacontenedores más pequeños que carecen de células, los contenedores se llevan en las bodegas sujetos con barras.

En los buques Ro-Ro, los contenedores se estiban sobre cubierta y se afianzan con dispositivos de amarre, o permanecen sobre remolques de diversos tipos durante el transporte.

C. LAS OPERACIONES MULTIMODALES Y LOS PUERTOS

35. ¿Qué supone para los puertos el transporte multimodal?

Un movimiento más rápido de las mercancías gracias a la unitarización. Ahora bien, el transporte multimodal exige inversiones suplementarias en las instalaciones que se necesitan para manipular la carga unitarizada y una reorganización de la estructura y las operaciones del puerto, sobre todo en lo que respecta a la mano de obra.

36. ¿Qué se necesita en los puertos para manipular carga unitarizada?

Las necesidades de los puertos en este aspecto, difieren mucho según el tipo de unitarización que se adopte y según las instalaciones de manipulación y control que en ellos existan ya. Entre las diversas formas de unitarización, la más compatible con las instalaciones portuarias tradicionales es la paletización, siempre que se disponga de un número suficiente de carretillas elevadoras, que éstas tengan vía libre para maniobrar y que se cuente en el terminal con una explanada provista de un número suficiente de pistas.

El sistema de gabarras tampoco exige muchos cambios en la disposición del puerto siempre que el muelle tenga profundidad suficiente para gabarras de escaso calado y se disponga de una zona para maniobra. Los buques portagabarras requieren asimismo un fondeadero de amplitud suficiente que tenga de uno a cuatro puntos de amarre.

El movimiento de contenedores, si el tráfico es reducido, puede hacerse también en un muelle de tipo corriente siempre que se disponga de ciertos elementos materiales tales como grúas y remolques para manipular la carga fraccionada ordinaria de los contenedores. Ahora bien, contrariamente a lo que ocurre en el caso de los muelles para paletas, el terminal especialmente construido para contenedores es muy distinto de un muelle para carga fraccionada en cuanto a las condiciones materiales que debe reunir y a la organización interna. Es preciso contar con equipo de muelle idóneo para la manipulación, como, por ejemplo, grúas-pórtico y carretillas-pórtico, así como una zona extensa de muelles reforzados para clasificar y estacionar los contenedores. El terminal debe estar dotado además de tinglados, servicios de conservación, talleres de reparación y parques de remolques.

37. ¿Cómo adaptar un puerto a la manipulación del tráfico unitarizado?

La mayoría de los puertos están en condiciones de manipular las unidades de carga de tipo más sencillo, pero la manipulación de unidades más complejas y voluminosas exige nuevos cambios.

Ante la perspectiva de una unitarización deben tenerse en cuenta los puntos de transbordo y la infraestructura que el puerto puede ofrecer. Hay que determinar, además, hasta qué punto los cargadores o usuarios se percatan de lo que es la unitarización y están en condiciones de organizar sus operaciones a base de unidades de carga. La elecci del método de unitarización, que influirá en definitiva en las inversiones destinadas a los servicios portuarios, dependerá también del tipo de mercancías que han de pasar por el puerto.

Debe dotarse al puerto de equipo adecuado para la manipulación o modificar el equipo existente. También deben efectuarse en la disposición material del puerto los cambios que sean necesarios para el movimiento de las unidades de carga que, según se prevé, han de pasar por el puerto. Por ejemplo, habilitar lugares para estacionamiento, prepara una zona de almacenamiento de contenedores o construir muelles adecuados.

38. ¿El hecho de que el movimiento de mercancías se haga por métodos multimodales supone un cambio en las operaciones del puerto?

Sí. Cuando el movimiento de mercancías se hace por procedimientos multimodales, operaciones del puerto son principalmente de tránsito. La manipulación de la carga y la estiba se reducen al mínimo. Sin embargo, en un puerto de tipo corriente se plantearían diversos problemas nuevos, sobre todo en el caso de la entrega directa. Habrá que proceder a una vasta reorganización del puerto de modo que se reúnan las condiciones necesarias para las operaciones de entrega directa a vehículos de carretera, trenes o gabarras. Por ejemplo, la maniobra de los vagones de ferrocarril constituirá un problema importante a menos que se creen las instalaciones necesarias. La regulación del tráfico por carretera es otro ejemplo de ello; la planificación de ese tráfico y la medida en que pueda satisfacer las exigencias del transporte multimodal influirá en las operaciones del puerto. Las operaciones requerirán menor densidad de mano de obra y exigirán más espacio abierto del que se dispone en la mayoría de los puertos de tipo corriente. Habrá que mejorar los métodos en materia de documentación y de información. El movimiento multimodal de las mercancías supone una comunicación más activa con los clientes y los destinatarios.

39. ¿Harán falta en los puertos muelles especiales para manipular la carga unitarizada?

Esto dependerá de la escala utilizada y del sistema de unitarización. Cuando el movimiento de contenedores no es grande la manipulación puede hacerse en los muelles corrientes de carga fraccionada que suelen consistir en un espacio o andén angosto

que sólo son apropiados para los buques portacontenedores más pequeños o para los buques corrientes de carga fraccionada que transportan contenedores. Debe disponerse de espacio suficiente para almacenamiento detrás del muelle; de otro modo, los contenedores han de ser trasladados directamente de los buques a vagones de ferrocarril o a vehículos de carretera. En todo caso, los patios de los muelles corrientes de carga fraccionada no suelen ser apropiados para la buena marcha de las operaciones.

Así, pues, para poder implantar en gran escala el método del transbordo de contenedores por elevación, los muelles tendrán que ser más reforzados y las aguas junto a ellos más profundas que cuando sólo han de prestar servicios a los buques corrientes de carga fraccionada, y sus servicios de manipulación habrán de ser más complejos.

40. ¿Requiere un muelle para contenedores mayor espacio que un muelle para carga fraccionada?

Sí. En vista del mayor movimiento de mercancías que la contenerización lleva consigo así como del ritmo más acelerado de las operaciones de manipulación, del uso de equipo de manipulación de contenedores suplementario o especializado, y de las mayores dimensiones de los buques portacontenedores de tercera generación en comparación con los buques corrientes de carga fraccionada, es evidente que los puestos de atraque habrán de tener mayor amplitud.

41. ¿Será necesario que en un muelle de contenedores la zona de descarga y almacenamiento tenga una pavimentación especial? ¿Cuáles serían los costos? h/

Sí. Los costos dependerán del tipo de pavimentación, que a su vez dependerán del volumen y la densidad del tráfico. De todos modos, como cifras indicativas, puede calcularse 35 dólares de los EE.UU. por metro cuadrado de pavimentación de la zona del muelle destinada a las operaciones de carretillas-pórtico, y 30 dólares por metro cuadrado para la pavimentación de los tinglados, a base de los precios registrados en junio de 1975 al Reino Unido. El costo de construcción de los tinglados será de unos 50 dólares por metro cuadrado, a base de los precios registrados en junio de 1975 en Reino Unido.

---

h/ Las cotizaciones fueron comunicadas por una empresa de consultores de ingeniería del Reino Unido y deben considerarse únicamente como órdenes de magnitud.

D. TERMINALES DE CONTENEDORES, DEPOSITOS E INSTALACIONES CONEXAS

12. ¿Cuál es la diferencia entre un muelle de contenedores y un terminal de contenedores?

Las expresiones "muelle" y "terminal" pueden denotar idénticas instalaciones. Ahora bien, desde un punto de vista técnico el muelle es un andén donde atracan los buques y donde los contenedores son cargados a los buques o descargados de ellos, mientras que la palabra "terminal" puede denotar el muelle y la zona, por lo general contigua al muelle, donde los contenedores son almacenados, llenados, vaciados y trasladados a trenes o vehículos de carretera. Pueden establecerse también terminales de contenedores en puntos situados en el interior.

13. ¿En qué consiste la explanada en un terminal para contenedores?

Por lo general el término "explanada" se refiere al andén paralelo al muro del muelle y adyacente a éste donde se clasifican los contenedores para proceder a su carga o descarga con grúas.

14. ¿Cuáles son los costos aproximados de construcción del muro del muelle y de la explanada? 1/

a) Muro del muelle: por metro cuadrado útil para grúa de contenedores ..... 4.000 a 20.000 dólares EE.UU.
Relleno de explanada por metro cuadrado ..... 28 a 42 dólares EE.UU
Pavimento de hormigón o asfalto, por metro cuadrado ..... 7 dólares EE.UU.

15. ¿Cuáles son las partes integrantes de una terminal de contenedores?

a) La zona de clasificación que va del muro del muelle a la zona de apilamiento.
b) Las zonas de apilamiento y el espacio reservado para contenedores vacíos y averiados.
c) Los parques de remolques.

1/ Las cifras están basadas en datos publicados en "Aspectos técnicos de los sistemas de transporte de grandes contenedores" (ST/ECA/170) Nueva York 1970, págs. 106 a 112, y reajustados con arreglo a los precios registrados en junio de 1975 en el Reino Unido, conforme a un índice proporcionado por una firma de consultores en ingeniería de ese país.

d) Los andenes de recepción y entrega.

e) El tinglado de agrupación de la carga (que no se encuentra necesariamente en el terminal).

f) El taller de conservación y reparaciones.

g) Espacio para oficinas.

h) Otras instalaciones tales como la torre de control de operaciones, el puente-báscula, enchufes para contenedores refrigerados, las instalaciones de lavado de contenedores, etc.

46. ¿Cuáles son las funciones de un terminal de contenedores?

Las operaciones que se realizan en un terminal de contenedores son la carga y descarga de contenedores, su almacenamiento y agrupación y su traslado a vehículos de carretera y trenes. El terminal de contenedores puede dividirse en los elementos siguientes:

a) Zona de maniobra y embalaje para recepción de la carga y servicios de transportes.

b) Zona de maniobra y clasificación para la manipulación del equipo.

c) Zona de almacenamiento para contenedores cargados.

d) Zona de almacenamiento para contenedores vacíos.

e) Zona para llenar y vaciar contenedores.

f) Sector destinado a servicios de conservación del equipo y a administración.

g) Zonas de estacionamiento para remolques de chasis bastidor vacíos, vagones de ferrocarril, vagones plataforma y camiones de carretera.

47. ¿A qué nivel de tráfico se hace económicamente viable un terminal de contenedores?

Basándose en las prácticas actuales en materia de precios, se calcula que dada una inversión de 18 millones de dólares en un terminal de contenedores para grandes buques tendría que haber un movimiento de 30.000 a 35.000 cajones por año, para que los ingresos compensaran los gastos.

48. ¿Qué son las estaciones de carga de contenedores?

En estas estaciones de carga, llamadas también "centros de grupaje" o "centros de agrupación" de la carga se reciben los bultos o mercancías, que son agrupados y cargados en contenedores, descargados de ellos y distribuidos. Los depósitos de

contenedores ofrecen una fórmula eficaz, en sustitución del movimiento de contenedores de puerta a puerta, para los cargadores o usuarios que tienen cargas parciales o cargas completas de contenedores pero no cuentan con medios apropiados para efectuar ellos la carga.

49. ¿Quién puede organizar una estación de carga de contenedores?

Las estaciones de carga de contenedores pueden ser organizados por las empresas navieras, los transportistas por carretera, los comerciantes al por menor que operan con mercancías homogéneas o bien por el Estado en forma de una red nacional de estaciones de carga auspiciadas por él, fórmula ésta que puede ser más ventajosa que la de un gran número de centros de grupaje administrados por entidades particulares.

50. ¿Son necesarias las estaciones de carga de contenedores en un puerto donde sólo se mueve un pequeño número de contenedores?

No. Sobre todo si los contenedores pueden trasladarse directamente afuera del puerto o a un tinglado corriente de carga fraccionada donde sea posible efectuar su llenado vaciado, siempre que las dimensiones materiales de los tinglados permitan la maniobra sin obstáculos del equipo de traslación.

51. ¿Es posible instalar los servicios conexos, es decir, el tinglado de agrupación de la carga, la zona de almacenamiento y apilamiento, las zonas de entrega y conservación de contenedores en un lugar que no sea contiguo al muelle?

Aunque es posible contestar afirmativamente a esta pregunta, debe señalarse que, en caso de que el tráfico aumente considerablemente, esta solución impondrá ciertas limitaciones a las operaciones de manipulación de contenedores. Asimismo, al establecer los servicios conexos lejos de los servicios de grúas aumentarán los gastos de capital en equipo de traslación, disminuirá el ritmo de manipulación de la carga y se perderá flexibilidad, sobre todo en lo que se refiere a la elección del equipo de traslación. Por ejemplo las carreterillas-pórtico no son apropiadas para la traslación a distancia.

52. ¿Cómo pueden utilizarse los tinglados de carga fraccionada en muelles comunes para contenedores y carga fraccionada en caso de que aumente el tráfico de contenedores?

Si se transforma el muelle para destinarlo exclusivamente al movimiento de contenedores, los tinglados pueden utilizarse para el llenado y vaciado de contenedores siempre que esos tinglados, por su anchura y altura, permitan la maniobra del equipo de

traslación. La zona de tinglados habrá de tener mayor o menor amplitud según las necesidades del tráfico. Su extensión puede determinarse atendiendo al número de toneladas que han de manipularse en el tinglado al año.

Si los tinglados de carga fraccionada no son adecuados para las operaciones de llenado y vaciado, habrá que desmontarlos y utilizar el espacio para almacenar contenedores sobre todo en reserva.

Si en algunos países en desarrollo se construyen terminales que estén expresamente destinadas a atender a la vez el tráfico de contenedores y el de carga fraccionada, s tinglados que se instalen no deben ser edificios de carácter permanente, sino estructuras desmontables constituidas de preferencia por elementos prefabricados.

E. UNITARIZACION Y MATERIAL DE MANIPULACION DE LA CARGA

53. ¿Cuál es el material necesario para manipular la carga unitarizada en el puerto? 1/

El material diferirá de un puerto a otro según las distintas variedades de unidades de carga acarreadas a través del puerto y según la configuración física de éste. Sería imposible hacer una lista completa de todo el equipo de manipulación para todas las formas de unitarización. Ahora bien, el equipo para manipular la carga unitarizada en los puertos puede consistir en cualquier combinación de los elementos siguientes:

a) Grúas martillo, grúas de torre, grúas para la manipulación de contenedores, grúas-puente, grúas móviles, material para transporte sobre vagones plataforma (sistema "canguro"), etc.

b) Carretillas de carga lateral, dispositivos de elevación lateral y unidades de transbordo lateral.

c) Carretillas elevadoras provistas de una variedad de dispositivos: mecanismo de desplazamiento lateral, horquillas extensibles, soportes con vástago, dispositivo portabarriles, brazo con gancho, dispositivo portabidones, prensores, etc. k/.

d) Gatos rodantes, gatos accionados mecánicamente.

e) Material de acceso para los contenedores: rampas de andén de carga, rampas móviles, rampas levadizas, plataformas elevadoras, etc.

f) Carretillas-pórtico.

g) Material de tracción: remolques, chasis, tractores, sistemas rodantes neumáticos o hidráulicos portátiles, "multifits", etc.

1/ El sistema de transbordo por rodadura y, hasta cierto punto, el sistema de transbordo por flotación, no necesita el mismo material de elevación de gran potencia que el sistema de contenedores. Sin embargo, el material utilizado para el transbordo y la traslación de las unidades de carga en todos los sistemas no es físicamente diferente. Aunque esta pregunta ha sido formulada con respecto a los puertos, parte del material enumerado es necesario para manipular cargas unitarizadas en cualquier punto de transbordo y en los terminales de contenedores.

k/ Los soportes con vástago se utilizan para izar alfombras, tubos, etc.

El dispositivo portabarriles consiste en dos pares de brazos superpuestos que ciñen el barril.

El brazo con gancho es un accesorio que habilita la carretilla elevadora para funcionar como grúa móvil.

El dispositivo portabidones es un aparato de elevación con movimiento vertical y un gancho para izar y manipular uno o dos bidones simultáneamente a modo de una eslinga.

Los prensores son un dispositivo consistente en dos brazos colocados en lugar de las horquillas que permiten asir, sujetándola por los lados, una unidad de carga o un solo artículo sin ayuda de paletas.

"b) Accesorios de manipulación de la carga: bastidores de suspensión, aparejo de estabilización automática de la carga, dispositivos de rotación de la carga, pulpos, dispositivos de vástago y sujetador[1/]

i) Eslingas y estrobos.

ii) Transpaletas, eléctricas o manuales.

54. ¿Qué es una grúa-pórtico y cuál es su costo aproximado?[m/]

La grúa-pórtico es una grúa para la manipulación de contenedores y carga fraccionada. Cuando está instalada a bordo, abarca toda la anchura del buque. Existen también muchos tipos de grúas-pórtico de muelle para la manipulación de contenedores. Las grúas-pórtico para el transbordo ferrocarril/carretera, por ejemplo, pueden ser automotoras y desplazarse con ruedas que corren sobre rieles o con ruedas de caucho.

Las dimensiones de una grúa-pórtico varían según su diseño y capacidad. Una grúa típica de 20 a 30 toneladas con un mecanismo de izada auxiliar de 10 toneladas puede manipular contenedores de 20 a 40 pies. La velocidad de izada es de 45 m/minuto y 30 minuto, respectivamente.

Una grúa pórtico de terminal típica para la manipulación de contenedores tiene una capacidad de elevación en bastidor de suspensión de 30 toneladas. La altura de izada es de 10,1 m y la velocidad de izada de 9 m/minuto. La velocidad de traslación de la grúa es de 120 m/minuto.

La luz de las grúas-pórtico de mayor tamaño puede variar entre 10 y 25 a 30 m. La capacidad de carga oscila entre 30 y 40 toneladas.

1/ El dispositivo de rotación de la carga es un aparato que permite hacer girar la carga, una vez izada por la grúa, 360° en ambas direcciones por medio de un motor eléctrico conectado con una transmisión hidráulica situada dentro del soporte giratorio.

El pulpo es un dispositivo adaptado a las eslingas de la grúa y destinado a la carga y descarga de balas de pasta de madera. Se compone de varios elementos prensores de sección rectangular y brazos ajustables.

El dispositivo de vástago y sujetador es un accesorio especial para la manipulación del papel en bobinas. El vástago se inserta en el alma de la bobina y unos sujetadores se cierran mecánicamente para la izada.

m/ Los precios indicados en las respuestas a las preguntas 54 a 62 son precios de fábrica aproximados, suministrados por varios fabricantes del Reino Unido, la República Federal de Alemania y el Canadá, y están sujetos a variaciones.

El costo aproximado de una grúa-pórtico es el siguiente:

Grúa-pórtico económica para la manipulación de contenedores, 30 toneladas, contenedor ISO 20'/30'/40', utilizado en barcos destinados a travesías cortas	500.000 dólares EE.UU.
Grúa-pórtico de 40 toneladas para la manipulación de contenedores, de un alcance útil de 115 pies con bastidor de suspensión, según el grado de automatización	2 a 3 millones de dólares EE.UU.

55. ¿Qué es una grúa móvil y cuál es su costo aproximado?

Una grúa móvil es una grúa automotriz destinada a funcionar sobre una superficie vial. Tiene múltiples aplicaciones y suele estar provista de eslingas y brazos de suspensión. Las grúas móviles pueden utilizarse para la manipulación de contenedores cuando el número de éstos es pequeño.

El costo de las grúas móviles, según su capacidad, que oscila entre 11 y 60 toneladas, y la longitud del aguilón, se sitúa entre 43.000 y 215.000 dólares.

56. ¿Qué es una carretilla de toma lateral y cuál es su costo aproximado?

Una carretilla de toma lateral es una carretilla elevadora cuyo mecanismo de izada está situado en un lado del vehículo para la manipulación de contenedores. Las carretillas de toma lateral provistas de bastidores de suspensión y horquillas se utilizan también para la carga y descarga de buques Ro-Ro y para el transbordo de contenedores desde las zonas de almacenamiento y hasta ellas. El mástil de elevación puede desplazarse de modo que el contenedor descanse sobre el chasis de la carretilla durante el transporte. Sus especificaciones técnicas son las siguientes:

Capacidad de elevación: 20 a 35 toneladas.

Apilamiento: 2 a 3 alturas de contenedores de 20 x 8 x 8 ó 40 x 8 x 8.

Velocidad de izada: 12 m/minuto a 15 m/minuto, cargada.

Una carretilla de carga lateral con una capacidad de elevación de 35 toneladas permite manipular por término medio de 12 a 15 contenedores por hora, en la carga o descarga de vagones de ferrocarril o camiones.

El costo aproximado de una carretilla de carga lateral es el siguiente:

Con una capacidad de 20 y 25 toneladas para contenedores 20'/30'/40'	100.000 a 140.000 dólares EE.UU.
Con una capacidad de 35 toneladas para manipular contenedores de 20'/30'/40' por medio de un bastidor de suspensión hidráulico	221.000 dólares EE.UU.

57. ¿Qué es una carretilla elevadora y cuál es su costo aproximado?

La carretilla elevadora es el más versátil de todos los materiales de manipulación y consiste en un vehículo automotor provisto de horquillas de acero afiladas para manipular cargas paletizadas sobre patines y de horquillas de gran potencia para manipular contenedores. La carretilla elevadora puede estar provista de una serie de accesorios diversos para operaciones especializadas. Puede ir equipado con llantas de caucho de tipo macizo o neumático. Con una carretilla elevadora de gran potencia es posible manipular los contenedores al costado del buque y transportarlos a la zona de almacenamiento o cargarlos en camiones de carretera o en vagones de ferrocarril. Las especificaciones técnicas de las carretillas elevadoras son las siguientes:

La capacidad de elevación oscila entre 0,5 y 50 toneladas. La altura de izada normal oscila entre 2 y más de 5 metros. Una carretilla elevadora con una capacidad de elevación de 25 toneladas permite manipular por término medio de 15 a 20 contenedores por hora, en la carga y descarga de camiones y vagones de ferrocarril. Una carretilla elevadora de gran potencia dotada de un bastidor de elevación de contenedores intercambiable de 20 a 30 pies permite apilar tres contenedores.

Una carretilla elevadora con una capacidad de carga de 25 toneladas puede recoger o descargar un contenedor en un tiempo de 1 a 2 minutos. La velocidad de traslación es de 360 m/minuto.

El costo aproximado de una carretilla elevadora es el siguiente:

3 a 5 toneladas	25.000 dólares
7 a 12 toneladas	50.000 a 70.000 dólares
20 a 25 toneladas	130.000 a 160.000 dólares

5. ¿Qué es un dispositivo de elevación lateral y cuál es su costo aproximado?

Los dispositivos de elevación lateral son mecanismos de manipulación de contenedores montados en camiones, remolques y semirremolques de transporte por carretera y consisten en dos brazos elevadores accionados hidráulicamente con eslingas o bastidores de suspensión montados en cada extremo de un chasis. Estos elevadores laterales pueden alzar un contenedor del suelo y depositarlo sobre su propio chasis o sobre el de un camión o un vagón de ferrocarril. La capacidad y las especificaciones técnicas de los dispositivos de elevación lateral son las siguientes:

Un elemento de izada lateral de 20/40 pies montado sobre un semirremolque extensible equipado para contenedores ISO de 20/40 pies permite el transbordo de contenedores desde el suelo, un vagón de ferrocarril o un remolque sólo por un lado y apilar dos contenedores de 8' 6". Un elevador lateral, con una capacidad de elevación de 30 toneladas, permite manipular por término medio de 6 a 7 contenedores por hora, en la carga o descarga de vehículos de carretera o vagones de ferrocarril.

El costo aproximado de un dispositivo de elevación lateral es de 80.000 dólares.

¿Qué son las unidades de transbordo lateral y cuál es su costo aproximado?

Las unidades de transbordo lateral consisten en brazos accionados hidráulicamente que desplazan lateralmente los contenedores empujándolos o arrastrándolos. Están provistas además de gatos nivaladores hidráulicos, dos en cada extremo de chasis, lo que permite el transbordo intermodal en superficies a diferente nivel. Este sistema se utiliza para el transbordo de contenedores entre camiones de carretera y vagones de ferrocarril y para el almacenamiento intermedio sobre patas o soportes. El costo aproximado del material de transbordo lateral es de 60.000 dólares por unidad.

6. ¿Qué es un gato rodante y cuál es su costo aproximado?

Los gatos rodantes se componen de dos ejes independientes provistos de sendos mecanismos hidráulicos de elevación y bastideros. Los dos bastideros tienen a su vez dispositivos de cierre para su acoplamiento a las cantoneras de los contenedores de tipo corriente. La elevación se efectúa con una bomba de mano. Una vez izado el contenedor, se remolca toda la unidad desde el costado del buque hasta la zona de almacenamiento con un tractor ordinario. Las especificaciones técnicas de un gato rodante son las siguientes:

Capacidad de elevación: 20 a 25 toneladas.

Altura de izada: 1,70 m (66,9 pulgadas).

El costo de un gato rodante oscila entre 5.600 y 6.300 dólares.

61. ¿Qué es una rampa y cuál es su costo aproximado?

Las rampas son una clase de material que se utiliza para salvar un espacio horizontal y/o vertical entre el piso de un andén de carga y el piso de un contenedor o de un camión o vagón de ferrocarril. Las rampas son de diversos tipos y tamaños. Las rampas móviles para la manipulación de contenedores permiten que las carretillas elevadoras introduzcan directamente las cargas en los contenedores, los camiones o los vagones de ferrocarril de puerta ancha.

Rampas de maniobra móviles

Capacidad: 7,26 a 11,34 toneladas.

Longitud: 9,15 a 10,97 m.

Anchura: 1,77 m.

Peso estimado: 1,6 a 1,96 toneladas.

Rampas de maniobra pequeñas:

Capacidad: 7,26 a 9,10 toneladas.

Longitud: 6,10 a 7,93 m.

Anchura: 1,77 m.

Peso estimado: 1,19 a 1,52 toneladas.

Hay dos tipos de rampas de esta clase:

La rampa de carga de paletas para la entrada horizontal en el contenedor que se está llenando y la rampa de carga directa que permite llegar sin interrupción hasta la cola del vehículo en que descansa el contenedor.

El costo de una rampa oscila entre 4.300 y 7.000 dólares. El de una rampa de paletas para salvar desniveles oscila entre 750 y 1.100 dólares.

62. ¿Qué es una carretilla-pórtico y cuál es su costo aproximado?

Las carretillas-pórtico son máquinas automotoras consistentes en un pórtico con patas que enmarca el contenedor. El pórtico lleva un dispositivo de elevación que puede ser un bastidor de suspensión; en algunos casos la carretilla está también provista de brazos elevadores que izan los contenedores por su parte inferior. La

carretilla-pórtico recoge los contenedores y los transporta a la zona de almacenamiento y puede incluso cargarlos en camiones de carretera. Las especificaciones técnicas de las carretillas-pórtico son las siguientes:

Las destinadas a la manipulación de contenedores de dimensiones normalizadas (20, 30 y 40 pies) tienen una velocidad de traslación de 23 km/hora y una capacidad de elevación de 30 toneladas. Pueden subir pendientes de un desnivel máximo del 3%. Su altura de izada es superior a los 5 metros y pueden apilar hasta tres alturas de contenedores.

La carretilla-pórtico puede maniobrar hacia adelante, hacia atrás y de lado. Su velocidad de traslación con una capacidad de carga de 30 toneladas es de 400 m/minuto. 24 km/h

El costo aproximado de una carretilla-pórtico es el siguiente:

Carretilla-pórtico capaz de apilar tres contenedores ISO de 20'/30'/40'	180.000 a 200.000 dólares EE.UU.
Tractor "Tugmaster" para contenedores ISO 20'/30'/40' sobre remolque, utilizado principalmente en el sistema de transbordo por rodadura	17.000 dólares 37.000 dólares

El costo de los bastidores de suspensión es el siguiente:

Grúa de 40 toneladas para la manipulación de contenedores	60.000 dólares
Grúa de 40 toneladas de desplazamiento horizontal de la carga	57.000 dólares.

F. OPERACIONES MULTIMODALES Y TRANSPORTE POR CARRETERA

63. ¿Qué condiciones debe reunir el transporte por carretera para las operaciones de transporte multimodal?

La mayoría de las carreteras se han construido para responder a las necesidades locales y atendiendo a las características geográficas de la región. Puede ser que tales carreteras no satisfagan las exigencias del transporte multimodal y en particular del transporte de contenedores y otras grandes unidades de carga. Por este motivo es posible que sea necesario examinar las exigencias materiales del transporte por vehículos de carretera y la reglamentación aplicable, en lo que respecta a la idoneidad de las carreteras, así como al peso, la carga y las dimensiones de los vehículos que transportan las unidades de carga. Por ejemplo, para los efectos del transporte de los contenedores de mayor tamaño, bastarán cargas por eje sencillo de 10 toneladas y cargas por ejes en tándem de 16 toneladas como máximo. A fin de aumentar al máximo las cargas por eje autorizadas, es posible utilizar remolques de plataforma baja que puedan transportar contenedores apilados, siempre que no existan limitaciones en cuanto a la altura. Los transportistas por carretera tendrán que reorganizar sus servicios e instalar depósitos. La red de carreteras debe proporcionar un acceso adecuado a los puertos, los terminales de contenedores y las estaciones de carga de contenedores. Ahora bien, a fin de sacar el máximo provecho de las ventajas del transporte multimodal, la infraestructura de carreteras debería desarrollarse de tal manera que, a la larga, permita efectuar entregas puerta a puerta.

64. ¿Cuáles son las características físicas y técnicas que deben reunir las carreteras para el transporte multimodal?

Un requisito mínimo de la infraestructura vial es que la capacidad de carga de la red de carreteras sea por lo menos igual a las cargas por eje sencillo y por ejes en tándem y al peso bruto de los vehículos de carretera que transportan un contenedor cargado al máximo. Para el transporte de contenedores de 20, 30 y 40 pies, las carreteras deben reunir los requisitos siguientes:

Anchura del carril:	3 metros
Curva horizontal mínima:	30 metros
Pendiente máxima:	10%
Visibilidad mínima de parada:	25 metros
Altura libre mínima:	4 metros

NORMAS SOBRE CARRETERAS PARA 30 Km/hora

estas cifras se basan en un diseño que permite una velocidad de 30 km por hora. Por otra parte, en muchos países, debido a las limitaciones del sistema de carreteras, la anchura máxima autorizada es de 2,40 m y la altura máxima de 1,80 m.

Las carreteras deben construirse con capacidad suficiente para soportar cargas por eje hasta de 13 toneladas si se prevé un tráfico de todos los tipos de contenedores ISO.

El número de carriles depende de la intensidad del tráfico. En zonas de poco tráfico, pueden bastar las carreteras de un solo carril, siempre que estén bien trazadas y se construyan zonas de cruce. Hay que determinar además la resistencia de los puentes de carretera. Si la resistencia del puente sólo permite que pase un vehículo con contenedor cada vez, habrá que regular el tráfico en consecuencia. Si el puente no tiene resistencia bastante para soportar el peso de un contenedor, habrá que hacer una desviación, a menos que se pueda reforzar el puente o construir uno nuevo.

65. ¿Cuáles son los costos de construcción y mejora de las carreteras y de la infraestructura de carreteras?

a) Carreteras[n/]

Pavimentación de una carretera de grava. Ensanchamiento de la plataforma de 8 a 11 metros. Mejoras secundarias del trazado. Costo total unos 138.000 dólares por km.

Construcción de una carretera principal de dos carriles para velocidades de 100 km/hora. Anchura del pavimento, 8 metros. Costo total, unos 340.500 dólares por km.

Construcción de una carretera con fines de desarrollo, incluidos varios puentes pequeños. Terreno quebrado. Ancho del pavimento, 6 metros. Costo total unos 48.000 dólares por km.

Al construir carreteras nuevas, el costo suplementario de aumentar la carga por eje será relativamente reducido.

---

n/ Los costos se refieren al Pakistán y se basan en cifras publicadas en "Aspectos técnicos de los sistemas de transporte de grandes contenedores" (ST/ECA/170), página 117, ajustadas a los precios de 1975 aplicando un índice basado en estimaciones de una empresa de consultores de ingeniería del Reino Unido. Las cifras de costos se indican órdenes de magnitud aproximados.

Costos marginales de construcción al aumentarse la carga total

Carga por eje autorizada	8 toneladas	10 toneladas	13 toneladas
Costo marginal de una carretera de dos carriles (en dólares por km)	0	5 400	10 800
Costo de obras de refuerzo (en dólares por km)			
Carga por eje actualmente autorizada	Refuerzo hasta 10 toneladas de carga por eje		Refuerzo hasta 13 toneladas de carga por eje
10 toneladas 8 toneladas menos de 8 toneladas	10 500 18 000		12 000 16 500 28 500

b) Puentes a/

Se supone inicialmente que los puentes de nueva construcción fueran calculados para resistir una carga por eje de por lo menos 13 toneladas, o el correspondiente peso por eje y peso total. El costo marginal para aumentar la carga por eje de 8 a 10 ó 13 toneladas puede ser módico. En algunos casos hay la posibilidad de reforzar los puentes existentes en vez de recurrir a nuevas construcciones.

Ejemplo de mejoras en puentes para aumentar a 13 toneladas la carga por eje autorizada

Tipo de puente	Mejora	Costo (en dólares por unidad)
Tablero de madera	Viga de acero suplementaria	300 - 450 por tonelada
Viga de acero y tablero de hormigón	Tablero nuevo	45 - 75 por m^2
Armazón de riostras	Refuerzo del armazón	2 250 - 22 500 por puente

a/ Los costos para Europa se basan en las cifras publicadas en "Aspectos técnicos de los sistemas de transporte de grandes contenedores" (ST/ECA/170) ajustadas a los precios de 1975 aplicando un índice basado en estimaciones de una empresa de consultores de ingeniería del Reino Unido. Las cifras de costos sólo indican órdenes de magnitud aproximados.

Costo de construcción de puentes

Tipo de puente	Costo (en dólares por m^2)
Puente colgante	450 - 900
Puente de celosía	300 - 375
Otros tipos	150 - 375

Costo de construcción de túneles por kilómetro cuadrado para una carretera de dos carriles: 750.000 - 1.500.000 dólares.

C. OPERACIONES MULTIMODALES Y TRANSPORTE FERROVIARIO

6. ¿Qué condiciones deben reunir los servicios ferroviarios para el transporte multimodal?

La adaptación del equipo existente o la adquisición de material móvil nuevo para que las unidades sean intercambiables, ya se trate de contenedores o paletas, de modo que se cuente con la capacidad de acarreo y la capacidad cúbica que requiera el transporte de unidades de carga. Es indispensable contar con terminales provistos de equipo de transbordo suficiente y situados en las inmediaciones de los terminales de contenedores. Al planificar redes nuevas o la ampliación de las ya existentes debe tenerse en cuenta la posibilidad de agrupar los envíos de mercancía en contenedores o paletas. Las reformas no tienen por qué ser muy ambiciosas ni llevarse a efecto con demasiado apresuramiento. Procediendo a una adaptación por etapas se podrán allanar obstáculos. Tal vez sea necesario adoptar un sistema de trenes-bloque. Un tren-bloque se compone de una serie de vagones y locomotoras permanentemente enganchados que funcionan formando una unidad en una determinada ruta, que el tren recorre en ambos sentidos con breves paradas en los puntos de transbordo. De esta manera es posible reducir al mínimo el tiempo de espera en los terminales, sobre todo si se trata de trayectos cortos o medianos. También es aconsejable adoptar el sistema de transporte "canguro"[p/] Este sistema permite trasladar rápidamente los remolques por ferrocarril de un terminal a otro y combina la flexibilidad de los camiones remolques para las operaciones de recogida y entrega con el transporte en masa que es la ventaja de los ferrocarriles. Puede ser necesario disponer de vagones plataforma provistos de bogies para gran velocidad que permitan transportar los contenedores ISO más grandes. Es recomendable el uso de vagones con bastidor de acero que no llevan la plataforma corriente y son más ligeros. Los contenedores se transportan directamente sobre el bastidor de acero. El peso que se ahorra con un vagón bastidor, en comparación con los vagones corrientes, es de 2.000 a 3.000 kg. Tal vez sea necesario revisar la organización, la política de precios y la infraestructura de la red ferroviaria con objeto de responder a las nuevas necesidades.

---

p/ El sistema "canguro" consiste en el transporte de camiones-remolque, cargados o vacíos, sobre vagones plataforma. Una designación más oficial es la de transporte de remolque sobre vagón plataforma (TCRV). No es sino una modalidad de transporte por contenedores, que se convierten en remolques cuando se les acopla un juego de ruedas traseras (bogies).

67. ¿Cuáles son las normas en cuanto a las características materiales y técnicas de una red ferroviaria utilizada en el transporte multimodal?

Más del 60% de las líneas ferroviarias de todo el mundo tienen un ancho de vía de 1,435 mm. Otros anchos de vías son 1,675 mm y 1,524 mm para dos tipos de vía ancha, 1.067 mm para las vías de El Cabo y 1,000 mm para la vía estrecha. En las curvas de pequeño radio se necesita un cierto ensanchamiento de la vía para que puedan pasar los vagones de más de dos ejes. Para una velocidad máxima de 70 km/h, el radio mínimo de la curvatura debe ser de 5.000 m y para velocidades máximas del orden de 75 a 100 km/h, el radio de curvatura mínimo debe ser de 10.000 m. Debe tenerse en cuenta los movimientos laterales y el hecho de que el peralte de los carriles y la fuerza centrífuga harán que los vagones sobresalgan más de los carriles cuando el tren está en movimiento. Por consiguiente, el espacio libre deberá ser mayor que la sección transversal máxima de los vagones junto con sus cargas. Los perfiles de carga deben diseñarse teniendo en cuenta las dimensiones del espacio libre, los movimientos laterales del tren sobre los carriles, las posiciones extremas de los vagones en las curvas y los márgenes de seguridad necesarios para túneles. Los ferrocarriles con perfiles de carga de la UIC pueden transportar asimismo todos los contenedores ISO de la serie 1. En los ferrocarriles de menor perfil de carga, el problema de la altura podrá resolverse utilizando vagones especiales con plataforma baja o bien vagones con ruedas de pequeño diámetro.

68. ¿Cuáles son las normas en cuanto a las características materiales y técnicas de los vagones de ferrocarril utilizados en el transporte multimodal?

Hay muchos tipos de vagones plataforma para el transporte de contenedores ISO. En los vagones plataforma ordinarios puede instalarse un dispositivo de sujeción que permita transportar dos contenedores de 20 pies en cada vagón. Puede ser necesario hacer ciertas modificaciones en los vagones plataforma de uso general, de manera que los contenedores queden parcialmente sujetos. Los vagones portacontenedores especiales tienen simples bastidores de hierro en vez de plataformas de tipo corriente; esto permite economías importantes en la tara, y por tanto, una mayor carga útil. Algunos vagones especiales están provistos de dispositivos de sujeción de contenedores que se accionan pulsando un botón y de señales luminosas que indican si los contenedores están firmemente sujetos a los vagones. Algunos están provistos de placas de guarda retráctiles. Un vagón equipado con 12 de estas placas, que permiten un desplazamiento

de 14 pulgadas en cada dirección, el contenedor queda afianzado y protegido contra los movimientos violentos. Los vagones deben estar dotados de dispositivos amortiguadores especiales para evitar los daños durante las maniobras.

69. ¿Es posible el transporte multimodal en redes ferroviarias con distintos anchos de vía?

Sí, en la medida en que este problema pueda superarse de modo que se evite el transbordo a base de cambiar los ejes en los puntos de empalme de vías de distinto ancho. Se considera que el nuevo tipo de vagones de ajuste automático suprimirá en breve los problemas que plantean los diferentes anchos de vía.

H. OPERACIONES MULTIMODALES Y VIAS DE NAVEGACION INTERIOR

70. ¿Qué condiciones deben reunir las vías de navegación interior para las operaciones de transporte multimodal?

Las vías de navegación interior deben estar coordinadas con los otros modos de transporte. Esto puede exigir inversiones en diversos programas relativos a las embarcaciones y el equipo. Tal vez sea necesario un servicio constante para mantener el calado las vías de navegación de modo que éstas puedan ser utilizadas todo el año. En algunos países en desarrollo las vías de navegación interior pueden habilitarse para el movimiento de gabarras LASH y Seabee. Es indispensable, pues, que las vías de navegación interior tengan ciertas dimensiones mínimas que permitan el paso de gabarras LASH y de otro tipo que puedan transportar unidades de carga. En vías poco profundas dotadas de un canal de navegación cuya anchura y radio de curvatura sean suficientes, el transporte se puede efectuar eficazmente en grupos y "trenes" de gabarras. El "remolque de empujo" ofrece ventajas sobre los métodos tradicionales de desplazar las gabarras sin propulsión propia. Los remolcadores de empuje dirigen mejor las gabarras, sobre todo a los canales estrechos y sinuosos de poca profundidad. Las esclusas deben tener la suficiente profundidad, anchura y longitud para que quepa todo el tren de gabarras en el cuenco. En los puntos de distribución del interior debe instalarse equipo de manipulación de carga, tales como grúas de tierra firme, junto a los ríos o canales. En algunos casos pueden utilizarse pontones grúa. Puede ser necesario además techar ciertas zonas para proteger de la intemperie las operaciones de carga y descarga de las gabarras.

I. OPERACIONES MULTIMODALES Y TRANSPORTE AEREO

1. ¿Qué condiciones debe reunir el transporte aéreo para las operaciones del transporte multimodal?

Los contenedores para el transporte aéreo difieren en tamaño, construcción y capacidad de los utilizados en otros modos de transporte. Los contenedores aéreos y los dispositivos de unidad de carga DUC, contenedores iglú o contenedores de cubierta inferior aprobados por la IATA, son mucho más ligeros que los contenedores ISO, pues responden a la preocupación principal de los portadores aéreos que es evitar todo daño al avión y aligerar el peso del contenedor. Los contenedores aéreos no tienen cantoneras ni montantes pesados ni están provistos de una estructura que permita el transporte integrado por mar y no podrían resistir el transporte marítimo ni la manipulación que éste entraña. Además la forma de iglú no es compatible con las características que exigen los materiales de manipulación de contenedores ISO tales como las grúas pórtico o grúas de contenedores y sus bastidores de suspensión. Los dispositivos de unidad de carga DUC son compatibles con las operaciones de transporte por carretera. Pero todavía hace falta diseñar un contenedor intermodal que sea apropiado tanto para el transporte aéreo como para los demás modos de transporte. Los aviones de gran fuselaje de la nueva generación (B-747, DC-10, L-1011), que pueden transportar contenedores intermodales aire/carretera de 20 pies, tienen un servicio de transporte bimodal y hasta cierto punto hacen innecesaria la maniobra suplementaria en terminales aéreos. Se trata, sin embargo, de aviones que suponen cuantiosas inversiones de capital. Tal vez sea posible en el futuro que los contenedores utilizados por las líneas aéreas en el transporte multimodal tengan las dimensiones señaladas por la ISO y vayan provistos de cantoneras normalizadas. Esos contenedores habrían de tener resistencia suficiente para soportar el transporte por aire y por tierra. Es indispensable, sin embargo, que su peso sea mucho menor que el de los usados en el transporte de superficie. Mientras tanto, a fin de lograr un transporte más integrado, hay que usar los contenedores proscritos por la IATA, útiles para cien viajes por lo menos, que están construidos con paneles formados por un alma de madera de balsa en sección transversal recubierta por ambos lados con láminas de una aleación de aluminio. Hay que contar con materiales especiales de manipulación tales como tractores, pinzas de izado, aparatos de elevación, dispositivos transportadores, etc.

72. ¿Cuáles son los procedimientos de manipulación en el transporte de un contenedor por vía aérea?

En un viaje normal habrán las operaciones siguientes:

a) Carga de la mercancía por el cargador;
b) Traslado al almacén provisional;
c) Traslado del almacén al camión;
d) Traslado del camión al transportador de rodillos de la línea aérea;
e) Traslado del transportador de rodillos al almacén provisional del terminal;
f) Traslado del almacén del terminal a la paleta de avión;
g) Traslado al avión;
h) Estiba a bordo del avión;

Estas operaciones se harán a la inversa cuando el avión llegue a su destino.

J. LAS OPERACIONES MULTIMODALES Y LOS PAISES SIN LITORAL

3. ¿Es la unitarización de la carga beneficiosa para los países sin litoral?

Sí. Sobre todo porque la unitarización supone una mayor rapidez del tránsito, a simplificación de los trámites de aduanas, la reducción de la documentación, y la ainución de los hurtos y averías.

4. ¿Es posible el transporte multimodal para los países sin litoral?

Depende de la infraestructura que comunica al país sin litoral con los puertos o otros países. Puesto que el denominador común en los principales servicios de ánsporte multimodal es el uso de contenedores, el transporte multimodal será posible rincipalmente si los contenedores pueden transportarse por una red de ferrocarriles, rreteras y vías de navegación interior que comunique a los países sin litoral con s puertos.

"INGENIERIA MARITIMA MODULO: PROGRAMACION Y CONTROL DE OBRAS". DEL 1º DE JULIO AL 6 DE SEPTIEMBRE 1985.

D O C U M E N T O S .

ING. PORFIRIO LEON RUIZ.

3

DOCUMENTO

ESCRITO PROPOSICION

DOCUMENTO CT-3 2

ESCRITO PROPOSICION

México, D.F., a de de 198

C.
Secretario de
P r é s e n t e

En atención a la convocatoria número - de fecha - de de 198 que se publicó el día de del año en curso, por medio de la cual esa Dependencia invita a participar en el concurso número relativo a:

comunico a usted que esta empresa cumplió con los requisitos establecidos, quedando inscrita para participar en dicho concurso y por lo tanto como Representante de _____

_____

_____

manifiesto a usted lo siguiente:

Que oportunamente se recogió el Pliego de Requisitos y sus apéndices relativos al concurso de que se trata y se ha tomado debida nota de las bases a que se sujetará dicho concurso y conforme las cuales se llevará a cabo la obra; se aceptan íntegramente las condiciones en el citado pliego y sus apéndices, los que para tal efecto se devuelven debidamente firmados por el suscrito en los términos de la SEXTA CONDICION del documento citado.

Asímismo, expreso que se conoce la Ley de Obras Públicas y demás disposiciones administrativas para la contratación y ejecución de obras públicas, las especificaciones y Normas de Construcción que tiene en vigor esa Secretaría, y que se acepta que tales documentos rijan, en lo conducente, respecto al concurso indicado y demás actos que de él se deriven.

Igualmente comunico a usted, que se conocen los planos del proyecto que nos fueron proporcionados y conforme a los cuales se realizará la construcción de la obra.

De conformidad con lo anterior, se presenta la proposición respectiva conteniendo los documentos que a continuación se detallan y que se encuentran integrados progresivamente, de acuerdo con el orden establecido por esa Secretaría en la SEXTA CONDICION del pliego respectivo.

_____
Firma del postor

A. - a). -DOCUMENTO CT-4, Garantía para sostener el cumplimiento de la proposición con un importe de $

a favor de la Secretaría de Comunicaciones y Transportes.

b). -DOCUMENTO CT-6a, Constancia de visita al sitio de la obra, - o manifestación escrita de conocer elsitio de los trabajos.

B. - Debidamente constestados, llenados y firmados en todas sus hojas, los documentos siguientes:

a). -DOCUMENTO CT-3, el presente Escrito Proposición

b). -DOCUMENTO CT-7, Programa de trabajo y montos mensuales - de obra.

c). -DOCUMENTO CT-8, Relación de equipo que se empleará en la obra.

d). -DOCUMENTO CT-8a, Calendario de utilización del equipo.

e). -DOCUMENTO CT-9a, Relación de conceptos para los que debe - rán presentar análisis detallados de Precios Unitarios.

f). -DOCUMENTO CT-9c, Análisis para la determinación del cargo indirecto,de acuerdo con el ejemplo contenido en el documento - CT-9b.

g). -DOCUMENTO CT-9e, Costo-horario de maquinaria, de acuerdo con el ejemplo contenido en el Documento CT-9b.

h). -DOCUMENTO CT-11, Catálogo de conceptos y cantidades de obra para proposición de precios unitarios y monto total de la propo - sición.

i). -DOCUMENTO CT-15, Procedimientos de construcción con sus anexos correspondientes.

C. - Formulados por el suscrito en papel membretado con nuestra razón - social y debidamente firmados en todas sus hojas.

a). -DOCUMENTO CT-9c-1, Programa de utilización del personal en - cargado de la dirección, supervisión y administración de los tra bajos.

Firma del postor

b). -DOCUMENTO CT-9c-2, Análisis del costo financiero.

c). -DOCUMENTO CT-9d, Datos básicos relativos a mano de obra, materiales y equipo que sirven de base al postor para elaborar análisis de costo directo.

d). -DOCUMENTO CT-9f, Análisis detallado de precios unitarios.

D. - Debidamente firmados en todas sus hojas.

a). -DOCUMENTO CT-6, Pliego de Requisitos y Especificaciones Particulares.

b). -DOCUMENTO CT-9b, Modelo para análisis detallado de precios unitarios.

c). -DOCUMENTO CT-10, Modelo de Contrato

d). -DOCUMENTO CT-13, Circular(es) Aclaratoria(s) en su caso.

Finalmente manifiesto a usted que nuestro representante técnico en la obra con amplias facultades para tomar decisiones en nombre de la empresa, será el Ing.
con cédula profesional No. expedida por la Dirección General de Profesiones de la Secretaría de Educación Pública.

Anexos: Los que se detallan en hojas.

ATENTAMENTE

Nombre y firma del postor o su representante autorizado.

4

DOCUMENTO

GARANTIA PARA EL SOSTENIMIENTO DE LA PROPOSICION

SECRETARIA DE

SUBSECRETARIA DE

DIRECCION GENERAL

FORMA CT-4

GARANTIA

QUE OTORGA EL POSTOR:

_____

nombre o razón social

engrapar aquí el cheque o certificado.

(en caso de certificado, doblar hacia el frente la parte superior de la hoja para que sea visible el título "garantía").

TIPO DE DOCUMENTO _____

NUMERO _____ DE FECHA: _____

IMPORTE $ _____ CON CARGO A _____

_____

PARA PARTICIPAR EN EL CONCURSO _____

número

RELATIVO A _____

descripción

_____

_____

6

DOCUMENTO

PLIEGO DE REQUISITOS

SECRETARIA DE
SUBSECRETARIA DE
DIRECCION GENERAL DE

DOCUMENTO CT-6 "PLIEGO DE REQUISITOS"

Concurso No.

Para los fines del presente Pliego de Requisitos, en lo sucesivo se denominará "SECRETARIA" a la Secretaría de Comunicaciones y Transportes, "POSTORES" a los Contratistas invitados a participar en el Concurso y "CONTRATISTA" al que resulte favorecido con la adjudicación del Contrato.

La Secretaría y los postores aceptan que para la presentación de este concurso y demás actos que de él se deriven, rijan las siguientes:

C O N D I C I O N E S:

PRIMERA. -El acto de presentación y apertura de las proposiciones con toda su documentación debidamente firmada en todos sus tantos, y colocada dentro del sobre cerrado en forma inviolable (de preferencia sellado con cinta adhesiva), se efectuará exactamente a las ..... horas ..... minutos del día ..... del mes de ..... de 19..... en la Sala de..... sita en: ..... y será presidido por el C. Funcionario que designe la Secretaría, quien de inmediato iniciará el acto y sólo permitirá la participación de los postores que se encuentren presentes a la hora citada, aun cuando no asistan el Representante de la Secretaría de la Contraloría General de la Federación y el invitado de la Cámara Nacional de la Industria de la Construcción, pasando lista de asistencia se procederá a la apertura y revisión de los sobres que presenten los postores, verificándose que los documentos estén completos y que satisfagan en principio los requisitos establecidos para el concurso.

Se levantará el acta primera que será firmada por todos los participantes e invitados al acto y se informará del lugar, fecha y hora en que se dará a conocer el fallo, dentro de un plazo que no excederá de (20) veinte días hábiles, contados a partir de la celebración del acto de apertura de las proposiciones.

SEGUNDA. -La obra deberá iniciarse a más tardar el día ..... de..... de 198 ... y la fecha límite para la conclusión de la misma será el..... .....

Firma del postor

TERCERA. - Para asegurar la seriedad de su proposición, deberá acompañarse con el DOCUMENTO Nº CT-4 un cheque cruzado por la cantidad de: $ .....
.....
.....
expedido a nombre de la SECRETARIA DE .....
por la cantidad mencionada con cargo a cualquier institución de banca y crédito debidamente autorizada. Todos los documentos relativos y las garantías otorgadas por los POSTORES serán conservados por la SECRETARIA hasta el acto del fallo o adjudicación y su devolución se hará conforme a lo dispuesto en los siguientes incisos:

a) Las garantías otorgadas por los POSTORES serán devueltas por la SECRETARIA en el acto en que se dé a conocer el fallo, a cambio del recibo original expedido por la misma, excepto aquélla que corresponda a la proposición del CONTRATISTA a quien se haya adjudicado el Contrato, la que retendrá la SECRETARIA hasta el momento en que dicho CONTRATISTA firme el Contrato respectivo, constituya la Fianza correspondiente al mismo y entregue la totalidad de los análisis de Precios Unitarios y el Programa de Ejecución de los Trabajos detallado por conceptos.

b) Si transcurrido el plazo señalado para dar a conocer el fallo (dentro de los veinte días hábiles siguientes al de la apertura de las proposiciones), la SECRETARIA no lo emitiera, lo hará del conocimiento de los POSTORES fijando la nueva fecha y hora en que dará a conocer el fallo definitivo, esta nueva fecha quedará comprendida dentro de los veinte días hábiles siguientes, contados a partir de la fecha fijada en primer término.

CUARTA. - La SECRETARIA mostrará a los invitados al concurso por una sola vez el (los) día (s) ..... del mes ..... de 198.. a las ..... horas ..... minutos, el sitio donde se construirán las obras: deberán ocurrir al .....
.....
.....
.....
debiendo presentarse con el C. .....
.....
quien acompañará a los POSTORES en la visita por parte de la SECRETARIA, debiendo recabar la constancia respectiva.

Firma del Postor.

Las dudas planteadas al Repdresentante de la SECRETARIA por los POSTORES en la visista a la obra deberán confirmarse por escrito, en los términos de la QUINTA CONDICION; en caso contrario se darán por no habidas tanto las preguntas como las respuestas.

QUINTA. - Para aclarar cualquier duda por escrito en relación con la documentación del concurso, durante el plazo disponible para la preparación de la proposición, los POSTORES podrán ocurrir al [illegible], como límite OCHO días antes de la fecha señalada para la presentación y apertura de las proposiciones, en la PRIMERA CONDICION de este Pliego. Las dudas serán aclaradas por escrito a más tardar TRES días antes de la celebración del acto de apertura, con copia a los demas POSTORES.

SEXTA. - Para formular la proposición se utilizará el DOCUMENTO Nº CT-3 el cual deberá estar debidamente firmado por el POSTOR y tendrá todos los anexos que a continuación se mencionan, exigidos por la SECRETARIA y ordenados en forma progresiva, precisamente después de cada una de las hojas índice que se proporcionan para tal objeto.

DOCUMENTO Nº CT-3:	Escrito Proposición.
DOCUMENTO Nº CT-4:	Garantía de cumplimiento de la proposición (engrapada en la hoja respectiva), legalmente expedida y que satisfaga los requisitos exigidos por la SECRETARIA en la TERCERA CONDICION de este Pliego.
DOCUMENTO Nº CT-6:	El presente Pliego de Requisitos que deberá tener anexas las especificaciones particulares.
DOCUMENTO Nº CT-6a:	Constancia de visita al sitio de la obra, o manifestación escrita de conocer el sitio de los trabajos.
DOCUMENTO Nº CT-7:	Programa de Trabajo y Montos mensuales de Obra.
DOCUMENTO Nº CT-8:	Relación de equipo que se empleará en la obra.
DOCUMENTO Nº CT-8a:	Calendario de Utilización del Equipo
DOCUMENTO Nº CT-9a:	Relación de Conceptos para los que deberán presentar análisis detallados de Precios Unitarios.

Firma del postor.

DOCUMENTO Nº CT-9b:	Modelo para análisis detallado de Precios Unitarios.
DOCUMENTO Nº CT-9c:	Análisis para la determinación del cargo indirecto de acuerdo con el ejemplo contenido en el Documento Nº CT-9b.
DOCUMENTO Nº CT-9c-1:	Programa de utilización del personal encargado de la dirección, supervisión y administración de los trabajos.
DOCUMENTO Nº CT-9c-2:	Análisis del costo financiero.
DOCUMENTO Nº CT-9d:	Datos básicos relativos a mano de obra, materiales y equipo que sirven de base al postor para elaborar análisis de costo directo, de cada uno de los conceptos de trabajos de acuerdo con lo establecido en el Documento Nº CT-9b.
DOCUMENTO Nº CT-9e:	Costo horario de maquinaria, de acuerdo con el ejemplo contenido en el Documento Nº CT-9b.
DOCUMENTO Nº CT-9f:	Análisis detallado de precios unitarios.
DOCUMENTO Nº CT-10:	Modelo de contrato de obras a base de precios unitarios, debidamente firmado en todas sus hojas.
DOCUMENTO Nº CT-11:	Catálogo de conceptos y cantidades de obra para proposición de precios unitarios y monto total de la proposición.
DOCUMENTO Nº CT-13:	Circular(es) Aclaratoria (s), en su caso.
DOCUMENTO Nº CT-15:	Procedimientos de construcción que el postor deberá presentar. Estos procedimientos deberán estar claramente explicados y el equipo en ellos propuesto será congruente con la relación contenida en el documento No. CT-8 "Relación de equipo que se empleará en la obra". Deberá contener los anexos solicitados en el documento CT-15.

Firma del postor.

SEPTIMA. - Para satisfacer todos los requisitos a que se refiere la SEXTA CONDICION, además de los anexos que se soliciten y que sean proporcionados por el POSTOR se deberán utilizar precisamente las Formas e impresos proporcionados por la SECRETARIA que integran la carpeta correspondiente al presente Concurso. Para presentar los documentos cuyos formatos no se proporcionan, el Postor utilizará papel membretado de su empresa y llenará estos documentos en la forma que considere más conveniente.

IMPORTANTE. - Será condición indispensable para que la propuesta sea aceptada que se incluyan los 19 (diecinueve) documentos descritos en la SEXTA CONDICION de este pliego.

Los documentos que forman la proposición deberán estar integrados y ordenados en la forma anotada en el primer párrafo de la SEXTA CONDICION, después de cada hoja índice.

Deberá evitarse engrapar o encuadernar todos los documentos de la proposición en un solo juego.

Para presentar la proposición se utilizará el sobre que para ese objeto entrega la SECRETARIA, anotando en el lugar correspondiente, el número del Concurso, la fecha, el nombre de la obra y el nombre y cargo de la persona que asiste al Concurso.

OCTAVA. - Al formular la proposición se aceptará por los POSTORES lo siguiente:

1. - Que la obra se llevará a cabo con sujeción a: Las Especificaciones Particulares y complementarias de la Dirección General de Obras Marítimas, las cuales forman parte de esta carpeta de concurso y las Normas de Construcción de la S.C.T., el Programa de Trabajo y Montos Mensuales de Obra, los Precios Unitarios anotados en el Catálogo de Conceptos y Cantidades de obra, conforme a las Cláusulas del Modelo de Contrato de Obras a Base de Precios Unitarios y las condiciones de este Pliego de Requisitos.

2. - Que tomaron en consideración las condiciones meteorológicas(lluvias), topográficas y geológicas de la región, así como las vías de comunicación existentes, compenetrándose de las condiciones generales y especiales del lugar específico de las obras, y que el desconocimiento de las condiciones anteriores, en ningún caso servirá posteriormente para aducir justificación por incumplimiento del Contrato y del Programa de Trabajo, o para solicitar bonificaciones a los Precios Unitarios Consignados en la proposición.

3. - Que habrán juzgado y tomado en cuenta todas las condiciones que pue

den influir en los precios unitarios, independiente de lo que dichos precios incluyan por razón del costo directo, del indirecto, de la utilidad y de los cargos adicionales establecidos y que el pago de los diversos conceptos se hará al precio unitario que se fije en el Contrato para la unidad de obra terminada, a satisfacción de la SECRETARIA.

4.- Que las diferencias que pudieran resultar en las cantidades de obra anotadas por la SECRETARIA en el DOCUMENTO No. CT-11 durante la ejecución de la obra, ya sean aumentos o reducciones, no justificará reclamación alguna del CONTRATISTA en relación con los precios unitarios respectivos.

5.- Que deberán elaborar y presentar anexo a su proposición el DOCUMENTO No. CT-7, Programa de Trabajo, que además les servirá para deducir los montos mensuales de obra.

6.- Que se propondrán precios unitarios únicamente para los conceptos contenidos en el DOCUMENTO No. CT-11 y que no deberán proponerse alternativas que modifiquen lo establecido en este Pliego de Requisitos.

7.- Que deberán presentar análisis detallados de Precios Unitarios para todos y cada uno de los conceptos contenidos en el DOCUMENTO No. CT-9a.

8.- Que la ejecución de las obras se ajustará a las inversiones que cada año autorice el C. Presidente de la República por conducto de la Secretaría de Programación y Presupuesto, y en caso de que la inversión para el año en curso no alcance a cubrir el monto de la adjudicación, se tramitarán inversiones adicionales por los importes comprometidos.

9.- Que en ningún caso se considerará que las modificaciones al DOCUMENTO No. CT-7, Programa de Trabajo y Montos Mensuales de Obra, formulado por el POSTOR, motivadas por causas ajenas a la SECRETARIA o derivadas de la incorrecta e ineficiente operación de sus equipos e instalaciones, implicará cambio alguno en los precios unitarios cotizados por él mismo, para los conceptos contenidos en el Catálogo correspondiente.

NOVENA.- EL DOCUMENTO No. CT-11, Catálogo de Conceptos y Cantidades de Obra para proposición de precios unitarios, se formulará procediendo de acuerdo con lo siguiente:

1.- Se llenará a máquina preferentemente, y de ser manuscrito se usará tinta, escribiendo con letra fácilmente legible; en ambos casos, dicho Catálogo deberá presentarse sin correcciones, raspaduras ni enmendaduras.

2.- Se anotarán los precios unitarios con número y letra expresándolos en moneda nacional aproximados a un centésimo. Si hubiere discrepancias entre precios unitarios anotados con número y los anotados con letra, serán estos últimos los que se tomarán en cuenta al establecer el monto corregido de la proposición.

3. - En caso de encontrarse errores en las operaciones aritméticas, se reconocerá como correcto el producto de las cantidades de obra anotadas por la SECRETARIA y los precios unitarios anotados con letra por el POSTOR

4. - Deberá anotarse, el importe total de cada una de las partidas contenidas en el DOCUMENTO No. CT-11, inmediatamente después del último concepto de la misma: este importe total deberá ser la suma de los importes de todos los conceptos que intervienen en la partida correspondiente, en su caso.

La suma de los importes totales de todas las partidas, representará el Monto Total de la Proposición la que deberá anotarse con número y letra en la última hoja del DOCUMENTO No. CT-11, en el lugar asignado para ello.

5. - De acuerdo con las correcciones que en su caso se hagan, se modificarán los importes de los conceptos y el monto total de la proposición que resulte, será el MONTO CORREGIDO para efectos de la adjudicación.

DECIMA. - Deberá señalarse en la proposición el nombre del Técnico titulado que sea el representante del POSTOR en la obra objeto de este concurso, el cual deberá estar registrado en la Dirección General de Profesiones de la Secretaría de la Educación Pública y además, deberá tener suficiente experiencia en las obras de la índole de la que se llevará a cabo: tendrá también obligación de conocer ampliamente las especificaciones generales de construcción, el proyecto y las especificaciones complementarias en su caso.

DECIMA PRIMERA. - Serán rechazadas por la SECRETARIA en el acto de presentación y apertura, las proposiciones:

1. - Cuando el sobre no esté cerrado en forma inviolable.

2. - Cuando el Postor no presente uno o más documentos en los términos de este Pliego.

DECIMA SEGUNDA. - La SECRETARIA se reserva el derecho de descalificar, posteriormente al acto de apertura y durante el estudio de las mismas, aquellas proposiciones de las personas físicas y/o morales.:

1. - Que contengan uno o varios precios unitarios no remunerativos.

2. - Que en los análisis detallados de Precios Unitarios hagan intervenir destajos o lotes por concepto de mano de obra, materiales y equipo.

3. - En las que no coincidan los precios unitarios analizados detalladamente con los anotados con letra en el DOCUMENTO No. CT-11.

4. - Que omitan en uno o más conceptos la cotización con letra y número en el DOCUMENTO No. CT-11

Firma del postor.

5. - Que no contengan completos los datos básicos relativos a salarios del personal obrero, precios unitarios de todos los materiales que intervengan en la obra y costo horario de la maquinaria y equipo DOCUMENTOS Nos. CT-9d, CT-9e y CT-9f.

6. - En las que no presenten completos los datos requeridos, en el análisis para la determinación del COSTO INDIRECTO, en los términos del modelo respectivo.

7. - Que no contengan la totalidad de los análisis detallados de precios unitarios solicitados por la SECRETARIA en el DOCUMENTO No. CT-9a.

8. - Que propongan alternativas que modifiquen las condiciones establecidas por la SECRETARIA en este Pliego, y conforme a las cuales se desarrollará el Concurso y la Obra.

9. - Cuando no contengan los documentos requeridos completos o que hayan omitido algún requisito.

10. - Cuando el POSTOR o su representante no firmen alguno o algunos de los documentos básicos que integran la proposición.

11. - Que se les haya rescindido o cancelado por incumplimiento un contrato.

12. - Que haya cometido hecho ilícitos en perjuicio de la SECRETARIA.

13. - Que se encuentren sujetas a suspensión de pagos o declaradas en estado de quiebra, con posterioridad de la apertura del Concurso.

14. - Que estén al servicio de la SECRETARIA percibiendo una retribución por nómina u honorarios.

15. - Que no satisfagan los requisitos de forma o de fondo, determinados en este Pliego y sus Apéndices.

16. - Que contengan información que, al ser verificada por la Secretaría, no coincida con la indicada en la propuesta.

DECIMA TERCERA. - LA SECRETARIA se reserva el derecho de declarar desierto el Concurso.:

1. - Cuando en el acto de presentación y apertura de las proposiciones se presenten menos de tres POSTORES.

2. - Cuando al ser estudiadas las proposiciones posteriormente, sean aceptadas o resulten satisfactorias menos de tres proposiciones.

E. P. 2

CONCURSO Nº :

1.5.- ACTUALIZACION DE PLANOS.

Una vez terminada la obra o suspendida por - cualquier motivo, antes de hacer la entrega, el contratista deberá ejecutar con todo detalle el levantamiento de todos los trabajos - realizados, vaciando los datos en los planos correspondientes, sin costo alguno para la - Secretaría, revisados y autorizados por el - Representante.

CONCURSO Nº

1.6.- GENERALIDADES:

1.6.1.- El contratista en la visita de inspección al sitio de la obra deberá obtener todos los datos que considere necesarios con relación al abastecimiento de agua potable, --- energía eléctrica, accesos y servicios necesarios durante el período de ejecución de los trabajos; además deberá obtener de las autoridades correspondientes los importes de las conexiones que requiera, debiendo asímismo -- considerar sus consumos durante el tiempo que duren los trabajos, ya que estas actividades no le serán pagadas - en forma adicional debiendo por lo tanto considerarlas - dentro de sus precios unitarios.

1.6.2.- El Representante entregará al contratista al inicio de - los trabajos y éste se dará por recibido de los trazos y bancos de nivel necesarios para la ejecución de la obra, debiendo tomar las medidas que requiera para tener referencias de comprobación de trazos y niveles.

Durante el proceso de toda la obra, el contratista ejecutará los trazos y nivelaciones que se requieran basado en los datos que proporcione el Representante, de acuerdo con el proyecto.

El Representante solicitará al contratista cada vez que lo considere necesario verificar trazos y niveles de la obra y este suministrará lo necesario y proporcionará -- los datos requeridos.

La localización general, coordenadas, distancias y tra-- zos de los ejes principales de la obra aparecen en los - planos de proyecto. Con base en los mismos, el contra-- tista deberá proceder a su ejecución haciéndose responsables de cualquier futuro error causado por negligencia - o pérdidas de una o varias referencias y/o por interpretaciones erróneas.

1.6.3.- Los volúmenes de obra contenidos en el catálogo de con-- ceptos son aproximados.

1.6.4.- El contratista al elaborar el precio unitario de cada -- concepto del documento CT-11 "CATALOGO DE CONCEPTOS...", deberá tener en cuenta las especificaciones particulares las especificaciones complementarias y las normas de --- construcción en el entendido que los números de las normas de construcción que se marcan en los conceptos del - catálogo antes mencionado, sólo son indicativos pero no limitativos; por lo tanto, si existen algunas especificaciones dentro de las normas, aplicables al concepto en - cuestión cuyo número no aparezca, dichas especificaciones deberán ser consideradas para analizar el precio y/o efectuar los trabajos correspondientes.

CONCURSO Nº

1.6.5.- Si las normas de construcción de esta Secretaría , se contraponen a las especificaciones complementarias, serán estas últimas las que rijan, asimismo, las especificaciones particulares regirán sobre las anteriores.

1.6.6.- En los precios unitarios que presente el postor deberá incluir todas las erogaciones que tenga que efectuar para poder realizar las actividades en cuestión de acuerdo con el proyecto.

1.6.7.- Si por necesidades propias de la obra fuera precisoque de uno a varios conceptos de trabajo se efectuara un volumen mayor o menor al indicado en el catálogo de conceptos de obra, el contratista realizará el volumen que resulte, no teniendo por este motivo ningún derecho al cambio de los precios unitarios originalmente aprobados, y únicamente se pagarán los trabajos realmente ejecutados.

1.6.8.- El contratista que ejecute la obra, objeto del presente concurso se obliga a realizar la misma, a los precios unitarios propuestos por él aun cuando pornecesidades de última hora de la Secretaría o por dificultades del sitio originalmente escogido sea necesario modificar la localización del lugar de lostrabajos, si las condiciones del nuevo sitio son iguales o semejantes a los del original.

1.6.9.- Si el contratista ganador de este concurso, no hizo las debidas consideraciones en sus análisis de precios unitarios de acuerdo con las normas y especificaciones del proyecto, el hecho de otorgarsele el contrato no significa que la Secretaría lo exima del estricto cumplimiento de todas las condiciones originalmente establecidas.

1.6.10.- El contratista ganador del presente concurso se obliga a presentar al Departamento de Concursos y Contratos los originales de los análisis de precios unitarios que complementan la totalidad del catálogo proporcionado en un plazo no mayor de diez días hábiles contados a partir de la fecha del fallo.

1.6.11.- El contratista ganador del presente concurso se obliga a tener en el sitio de la obra, un técnico residente ingeniero civil titulado con amplia experiencia en el tipo de obra que se va a ejecutar, el cual tendrá poder amplio -----

CONCURSO Nº

y bastante para actuar en nombre del contratista y cualquier orden dada por el Representante al mismo se considerá como transmitida al propio contratista.

Dicho ingeniero será a quien se dirija el Representante para tratar asuntos relacionados con la obra, motivo por el cual el lugar de residencia del ingeniero debe ser, debiendo proporcionar a esta Dirección General su nombre, dirección y número telefónico con el objeto de ser localizado cuando así se requiera; en igual forma estará autorizado para firmar las estimaciones por parte del contratista y deberá permanecer al frente de las obras durante todo el lapso que duren las mismas. En caso de ausencia temporal o total del técnico representante del contratista, éste se obliga a dar aviso a la Secretaría con la anticipación debida y por escrito del nombramiento legal del nuevo representanté técnico, el cual deberá cumplir con los requisitos y obligaciones del anterior.

1.6.12.- El contratista estará obligado a atender cualquier llamado del Representante, cuando su presencia sea requerida en la obra por motivos de trabajo.

1.6.13.- El contratista ganador de este concurso, se obliga a cumplir con el plazo de ejecución de la obra que se menciona en la cláusula segunda del Pliego de requisitos.

En caso de violación o falta de cumplimiento parcial o total del inciso anterior, el contratista se hará acreedor a la sanción estipulada en la cláusula quinta del contrato.

1.6.14.- Si el equipo propuesto no es el adecuado para los trabajos por ejecutar, el contratista se obliga a sustituirlo, sin que este hecho motive cambio en los precios unitarios propuestos, ni que se computen tiempos perdidos por dicha causa.

1.6.15.- Todos los materiales que se utilicen en la construcción serán nuevos y de primera calidad, a menos que se indique lo contrario y pasarán satisfactoriamente las pruebas que indique la Secretaría.

1.6.16.- Al término de los trabajos, deberán quedar completamente limpias las obras, por lo que el contratista tomará las providencias necesarias ya que dichos trabajos no le serán pagados por separado. El contratista se abstendrá de arrojar al fondo marino cualquier tipo de objetos, cuando así sucediera la limpieza respectiva será por cuenta y cargo.

1.6.17.- Será requisito indispensable, para firmar la última estimación y levantar el acta de recepción de obra, que el contratista efectué el levantamiento de los trabajos realizados, vaciando los datos en los planos correspondientes, revisados y autorizados por el Representante.

1.6.18.- En el documento CT-8 "RELACION DEL EQUIPO QUE SE EMPLEARA EN LA OBRA....", el contratista describirá con todo detalle, las características de dicho equipo, como son modelo, capacidad, etc.

1.6.19.- La medición del acero de refuerzo para concreto hidráulico, se efectuará cuantificando la longitud geométrica que indica el proyecto, sin incluir desperdicios, ganchos, traslapes, dobleces, ---- alambre de amarre y/o soldadura. Los pesos unitarios se tomarán de la tabla 11-1, Tomo VIII de las Normas de Construccion de la S.C.T. (Pag. 193).

1.6.20.- El hecho de que la Secretaría otorgue el contrato objeto de este concurso, a un determinado contratista, no quiere decir que acepta todas las condiciones que éste indica en su proposicion y estará en el derecho de exigirle, el estricto cumplimiento de las Normas de Construccion y especificaciones correspondientes.

1.6.21.- Cualquier trabajo no considerado en este concurso, sólo podrá ejecutarse previa autorizacion por escrito del Representente, debiendo quedar asentada en la bitácora oficial.

1.6.22.- Cuando algún siniestro debidamente catalogado como tal por la Secretaría, afecte obras en proceso o terminadas, pero no recibidas por la Secretaría, únicamente tendrá derecho el contratista al pago de daños causados por el siniestro, cuando su avance de obra esté de acuerdo con el calendario y que

las mismas hayan sido efectuadas de acuerdo --- al proyecto, las Normas de Construcción, especificaciones y a satisfacción del Representante.

1.6.23.- El control de calidad se hará de acuerdo con las indicaciones que gire al respecto la Secretaría.

1.6.24.- El contratista se obliga a mantener limpia y ordenada el área de trabajo, durante todo el tiempo que dure la obra.

1.6.25.- A la firma del contrato correspondiente el contratista ganador entregará el programa de ejecución de los trabajos detallados por conceptos -- consignando por períodos las cantidades por ejecutar e importes correspondientes.

1.6.26. Cuando en la cotización se propongan productos o o materiales de calidad similar a la especificada en el proyecto, estos deberán ser descritos - detalladamente en la proposición a fín de que se pueda comparar la calidad de los mismos.

E. P. 8

CONCURSO Nº .

1.7.- SUMINISTRO Y COLOCACION DE PIEDRA NATURAL PARA CONSTRUCCION DE ESCOLLERA.

1.7.1.- La piedra para la construcción de la escollera la obtendrá el contratista del banco denominado ---- , el cual lo mostrará el Representante el día de la visita; se encuentra --- situado a una distancia aproximada de km del lugar de la obra.

El banco lo explotará el contratista de acuerdo - al método que proponga y que acepte la Secretaría; en caso de usar explosivos, su adquisición, transporte, almacenamiento y uso estará sujeto al cumplimiento por parte del contratista de lo que indiquen los reglamentos en vigor de la Secretaría - de la Defensa Nacional, en caso de accidente y/o - daños a terceros el contratista será el único responsable, debiendo hacer las reparaciones necesarias por su cuenta y cargo.

Estudiará cuidadosamente el contratista el banco - en lo que respecta a brechas interiores que se encuentran muy pronunciadas a fin de definir las -- maniobras y lugares donde deberá clasificar y almacenar la piedra, para los pesos correspondientes a núcleo, capa secundaria, coraza en tal forma que - faciliten las maniobras de carga.

El día de la visita, el contratista se documentará en todo lo relativo al transporte de la piedra desde el banco hasta el sitio definitivo de su colocación, contando el tipo y capacidad de los vehículos que puedan circular por los caminos a recorrer, permisos que se necesiten, y en general todo lo necesario para transportar la piedra.

Unicamente se cargarán los vehículos con piedra -- que reúna los requisitos de proyecto, no permitiendose que en el mismo vehículo se cargue piedra que sirva para dos o más capas diferentes.

Las operaciones para colocar la piedra en la escollera podrá efectuarse con grúa de tierra y/o con grúa flotante según requiera, debiendo tener presente el contratista que únicamente en la construcción del núcleo podrá ser a volteo, completando los taludes con grúa, En la formación de los taludes de la escollera, así como en la capa secundaria y coraza se requiere de acomodo de las piedras, con grúa equipada con almeja, gancho, red de cable, charola o cualquier otro equipo similar.

A fin de evitar daños durante la construcción de la escollera, no se permitirá dejar tramos de núcleo sin protección, una vez terminado un tramo de núcleo a las líneas y niveles de proyecto de una longitud tal que estará determinada por el alcance de la pluma de la grúa empleada, se procederá a cubrir las áreas expuestas con piedra de capa secundaria, cuando esta capa este terminada se procederá a recubrir su área expuesta con piedra de coraza, y así sucesivamente, hasta terminar la construcción de la escollera.

El contratista tomará en cuenta todo lo anterior, ya que será responsabilidad de el exclusivamente, el evitar que durante la construcción de la escollera se produzcan degradaciones en la misma, por la acción ordinaria del oleaje.

En caso que se produzcan estas degradaciones el contratista tendrá que reparar los daños por su cuenta incluyendo el valor de los materiales, sin que por esta circunstancia se le releve del cumplimiento del programa de obra, aprobado por la Secretaría y de las especificaciones respectivas.

1.7.2.- M E D I C I O N.

La unidad de medición será la tonelada. El Representante computara el peso neto de la piedra que efectivamente se coloque en la escollera de acuerdo con el proyecto. El destare de los vehículos se efectuará pesándolos sin carga, se pesarán en las mismas condiciones que cuando esten cargados.

E. P. 10 — 26

CONCURSO Nº

1.7.3.- BASE PARA EL PAGO.

En el precio unitario respectivo, el contratista incluirá despalme del banco y retiro del material producto, acondicionamiento de áreas para almacenaje de piedra, explotación del banco, selección acopio, carga a camión, acarreo al primer kilómetro, descarga y colocación de piedra en la escollera, desperdicios, brechas de acceso al banco y a la obra y su mantenimiento, durante el tiempo que dure la misma, capa de rodamiento que construya el contratista en la corona de la escollera para operar su equipo de construcción, retornos y en general todo lo necesario para efectuar el concepto de trabajo.

E. P. 11

CONCURSO Nº

1.8.- ACARREO DE PIEDRA PARA ESCOLLERA EN KILOMETROS SUBSECUENTES AL PRIMERO.

1.8.1.- El día de la visita el Representante indicará -- la ruta que seguirán los vehículos para acarrear la piedra desde el banco al lugar de la obra.

La distancia de recorrido para fines de pago se redondeará al kilómetro, cuando haya fracción y esta exceda a un hectometro, se tomará el kilometro subsecuente.

1.8.2.- M E D I C I O N.

La unidad de medición será la tonelada-Kilómetro cuantificada teniendo en cuenta el producto que resulte al multiplicar el peso neto de la piedra por la distancia de recorrido previamente señalada por el Representante.

1.8.3.- BASE PARA EL PAGO.

En el precio unitario respectivo el contratista incluirá el acarreo en kilometros subsecuentes al primero de roca, teniendo en cuenta, mano de obra, equipo, permisos y todo lo necesario para efectuar el concepto de trabajo.

E. P. 12 28

CONCURSO Nº

1.9.- INSTALACION DE BASCULA.

1.9.1.- La báscula que se usará para pesar la piedra para la construcción de la escollera, se encuentra actualmente en [illegible]. el día de la visita el Representante mostrará el sitio donde se localiza la báscula, y condiciones de funcionamiento en que se encuentra la misma.

El contratista instalará la báscula en el sitio donde indique el Representante, apegándose a las recomendaciones y especificaciones del fabricante; ya instalada, se obtendrá el visto bueno de la Dirección General de Normas para poder operarla.

Una vez terminada la obra el contratista efectuará el traslado de la báscula a [illegible], colocándola en el lugar que señale el Representante, quedando correctamente embalada. El contratista deberá tener en cuenta que durante el traslado de la báscula no deberá sufrir daños, de ser así el contratista será el único responsable, debiendo hacer las reparaciones necesarias por su cuenta y cargo.

1.9.2.- M E D I C I O N.

La unidad de medición será el Lote, cuantificado directamente en obra.

1.9.3.- BASE PARA EL PAGO.

En el precio unitario el contratista incluirá el transporte al lugar de la obra, construcción de obra civil, instalación de báscula, dotación de boletos foliados, mantenimiento, limpieza, embalaje de la misma y transporte al lugar de [illegible], que indique el Representante y en general todo lo necesario para efectuar el concepto de trabajo, además considerará en su caso la reparación de la báscula.

E. P. 13

CONCURSO Nº

El Representante efectuará la estimación de la siguiente forma: El 80% del valor del -- concepto al estar instalada, verificada por la Dirección General de Normas, y funcionando la báscula; el 20% restante al término de la obra y que la báscula haya sido embalada y depositada en el lugar que indique -- el Representante.

1.10.- CAMPAMENTO. CONCURSO Nº

ESTRUCTURACION DE CONCRETO

1.10.1.- TRABE PERIMETRAL, CASTILLOS Y DALA DE REMATE.

1.10.1.1. El Contratista retirará la basura que se encuentre en el área de construcción del campamento, despalmará dicha área, depositando basura y producto del despalme al lado de la obra.

El día de la visita se documentará en relación altipo de material que será necesario excavar, antes de colar las trabes perimetrales de cimentación se apisonará el fondo de la excavación.

La cimbra y obra falsa podrán ser de madera o de metal, debiendo tener la suficiente rigidez para no sufrir deformaciones durante el colado y vibrado del concreto, antes del colado el acero de refuerzo estará limpio, libre de óxido, grasa y polvo.

El Concreto se elaborará con cemento tipoII, agregados y agua limpios, libres de materia orgánica.

1.10.1.2. MUROS Y TECHOS DE MULTY-PANEL

Los muros y techos que se usen serán a base de láminas de la marca indicada en el proyecto o de marca similar en calidad y características.

En La instalación tanto de muros como de techos, el Contratista se sujetará a las especificaciones, recomendaciones y medidas del fabricante, pero en todos los casos, el Contratista será el único responsable de la calidad de los materiales y trabajos efectuados.

1.10.1.3.- RELLENOS, FIRME DE CONCRETO Y PISO DE CEMENTO PULIDO

Para efectuar el relleno, se usará el material, producto de excavación de cepas y de materiales obtenido en banco de préstamo lateral. Todo el material que se use estará limpio, libre de materia orgánica y se efectuará el relleno en capas compactadas al 90% de su P. V. S. M. con espesor máximo de 20 cm. el concreto del firma se elaborará con cemento tipo I, el piso de cemento pulido será del color que indique el Representante.

1.10.1.4.- PUERTA, VENTANAS Y CANCELES DE ALUMINIO ANODIZADO

Las cotas anotadas en el proyecto, serán verificadas en obra antes de fabricar las piezas, empleando para ello el perfil adecuado, debiendo tener en cuenta en puertas o ventanas y/o canceles el costo de vidrios y/o mosquiteros y/o multipanel y sello de silicón, las puertas interiores llevarán cerradura de intercomunicación y la puerta de entrada principal llevará cerradura de doble cilindro todas de la marca Schalogo de marca similar en características y calidad.

1.10.1.5.- INSTALACION DE GAS

El Contratista efectuará la instalación de gas tanto a la estufa como al calentador de agua usando tubería de cobre, debiendo proporcionar dos tanques portatiles cada uno con capacidad de 30 kg y proveer durante todo el tiempo que dure la obra del gas necesario para estufa y calentador de agua del campamento.

1.10.1.6.- INSTALACION HIDRAULICA Y SANITARIA

Además de lo indicado en le proyecto, se incluirá una cisterna con capacidad de 5,000 L y bomba con motor eléctrico de 1/2 HP, el agua necesaria para llenar la cisterna la proporcionará el Contratista durante todo el tiempo que dure la obra, usando una pipa para el suministro y acarreo respectivo.

En caso de no encontrar en la zona de la obra tabique rojo recocido para construir registros y pozo de absorción se usará tabicón o block.

La fosa séptica será marca Sanimex o de marca similar en calidad y características..

Los muebles de baño y accesorios respectivos serán marca Ideal Standard o similar de color, WC Mod-zafiro, Lavabo modelo-Veracruz, regadera marca Helvex o similar con chapetón y brazo, fregadero blanco marca Cinsa o similar.

El día de la visita se documentará el Contratista de las fuentes de aprovisionamiento de agua potable, así como la cantidad que deberá suministrar diariamente para cubrir las necesidades del campamento.

1.10.1.7.- INSTALACION ELECTRICA.

En las oficinas, cocineta, baño y pasillos deberá hacerse la instalación eléctrica para una lámpara y un contacto, en los talleres habrá dos lámparas y cuatro contactos, en igual forma se construirá la instalación eléctrica para operar la bomba de la cisterna.

La instalación eléctrica deberá quedar funcionando por lo tanto el Contratista considerará gastos de pago de derechos de interconexión a la línea que indique la Comisión Federal de Electricidad y pago de permisos respectivos.

El Contratista deberá entregar el campamento terminado y en condiciones tales que sea posible hacer uso de este como tal -- antes de empezar a construir el rompeolas.

1.10.2.- M E D I C I O N.

La unidad de medición será el lote.

1.10.3.- BASE PARA EL PAGO.

En el precio unitario el Contratista incluirá todo lo necesario para construir el campamento conforme a proyecto teniendo en - - cuenta excavación, trabes de cimentación, castillos, dalas, muros y techos de Multy-Panel, rellenos, firme de concreto, piso de cemento pulido, puertas, ventanas y canceles de aluminio anodizado, instalación de gas, dos tanques portátiles para gas con capacidad cada uno de 30 kg., instalación hidráulica y sanitaria, cisterna con capacidad de 5,000 litros, bomba con motor eléc-- trico de 1/2 H.P., fosa séptica, pozo de absorción, muebles de baño y accesorios, calentador de agua automático de 40 litros y fregadero, instalación eléctrica de oficinas, cocineta, baño, pasillos, talleres, bomba, limpieza y en general todo lo necesario para efectuar el concepto de trabajo. El Contratista tendrá en cuenta que tanto el Campamento como las instalaciones respectivas deberán estar terminadas funcionando perfectamente antes de iniciar el rompeolas. El Representante efectuará la estimación al terminar el concepto de trabajo, de acuerdo al proyecto y a su entera satisfacción.

Asimismo, deberá incluir dos unidades de ventana climatisadora de aire de 12200 BTU/h, marca York o similar.

E. P. 17

CONCURSO Nº

R E L A C I O N D E P L A N O S.

N U M E R O .	R U B R O .	F E C H A .
	PROYECTO OBRA EXTERIOR	
.-	CAMPAMENTO TIPO.- PLANTA, FACHADAS, CORTES Y ESPECIFICACIONES.	
	CAMPAMENTO TIPO.- INSTALACION HIDRAULICA Y SANITARIA.	

RELACION DE ESPECIFICACIONES COMPLEMENTARIAS.

S/N.-	R O M P E O L A S .

E. P. 18 — 34

CONCURSO N°.

A N E X O 1

SIGNIFICADO DE LAS ABREVIATURAS MAS USUALES EMPLEADAS EN EL PROYECTO.

SCT.	-Secretaría de Comunicaciones y Transportes.
D.G.O.M.	-Dirección Genral de Obras Marítimas.
C.F.E.	- Comisión Federal de Electricidad.
E.P.	-Especificaciones Particulares.
E.C.	-Especificaciones Complementarias.
N.C.	-Normas de Construcción.
N.B.M.I.	-Nivel de bajamar media inferior.
N.B.M.	-Nivel de bajamar medio.
N.P.T.	-Nivel de piso terminado.
P.V.S.M.	-Peso volumétrico seco máximo.
C.R.S.	-Capacidad relativa de soporte respecto al peso. volumétrico seco máximo.
V.R.S.	- Valor relativo de soporte.
P.V.C.	-Cloruro de polivinilo.
c.a.c.	-Centro a Centro.
fo. fo.	-Fierro fundido.
fo. ga.	-Fierro galvanizado.

E.P. 19

CONCURSO Nº

A N E X O II.

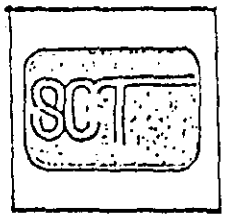
El Contratista ganador del presente concurso, deberá incluir entre sus indirectos el costo de un letrero que se debera colocar y permanecer todo el tiempo que dure la obra, en el lugar que indique el Representante.

ESPECIFICACIONES PARA LA CONSTRUCCION Y COLOCACION DEL LETRERO DE IDENTIFICACION DE LA OBRA:

1.- El tamaño del letrero, será el que se marca en el croquis adjunto.

2.- El letrero se podrá construir de madera, metal o cualquier otro material que proponga el Contratista y apruebe el Representante, deberá ser fabricado de tal manera que resista a las acciones meteorológicas de la zona.

3.- El letrero se deberá colocar en la zona de la obra, y de tal manera que sea visible desde cualquier punto de la misma.



SCT — SEPESCA

REALIZAN OBRA DE _____

CON UN MONTO DE _____

FECHA DE INICIO _____

FECHA DE TERMINACION _____

CONTRATISTA _____

LAS OBRAS SE REALIZARAN CON FINANCIAMIENTO PARCIAL DEL BANCO INTERAMERICANO DE DESARROLLO A TRAVES DEL BANCO NACIONAL PESQUERO Y PORTUARIO.

E. P. 20 — CONCURSO Nº. — M — 36

2.70 m.

Esc.

CONCURSO Nº .

INFORMACION ADICIONAL.

Por tratarse de una obra que forma parte del - PROGRAMA DE DESARROLLO REGIONAL PESQUERO MEXICO-BID, PRESTAMO 72/I.C-ME, se incluyen en este documento (CT-6 " PLIEGO DE REQUISITOS ") los siguientes anexos que el Postor deberá considerar para hacer su Proposición.

ANEXO Nº II. - NORMAS SOBRE NACIONALIDAD.

ANEXO Nº III. - ORIGEN DE LOS BIENES.

ANEXO Nº IV. - MARGEN DE PREFERENCIAS.

ANEXO Nº V. - ESPECIFICACIONES PARA LA CONSTRUCCION - Y COLOCACION DEL LETRERO DE IDENTIFI- - CACION DE LA OBRA:

E. P22 — 38

CONCURSO Nº

ANEXO III

NORMAS SOBRE NACIONALIDAD.

Contratistas de Construcción. El Banco, para determinar la nacionalidad de una firma Constructora y su elegibilidad para participar en la licitación de contratos financiados por el Banco, utilizará las siguientes normas:

- Que esté constituida u organizada de otra manera, en un país elegible;

- Que tenga la sede principal de sus negocios en un país elegible;

- (i) Que más del 50% de su capital sea propiedad de una empresa o empresas en uno o más países elegibles (dicha empresa o empresas también deberán calificar en cuanto a su nacionalidad) y/o de ciudadanos o residentes "bona-fide" de esos países elegibles, y (ii) que constituya una parte integral de la economía del país elegible en que está domiciliada;

- Que no exista arreglo alguno en virtud del cual una parte sustancial de las utilidades netas o de otros beneficios tangibles de las empresas sean acreditados o pagados a personas que no sean ciudadanos o residentes "bona fide" de los países elegibles;

- Que por lo menos el 80% de todas las personas que presten servicios conforme al contrato de construcción en el país donde ésta se lleve a cabo y estén empleadas directamente por el Contratista o por un subcontratista, sean ciudadanos de un país elegible. Para los efectos de este cómputo y respecto de una firma proveniente de un país que no sea el de la localidad de la construcción, no se tendrá en cuenta ciudadanos o residentes permanentes del país donde se lleve a cabo la construcción.

- Las normas anteriores se aplicarán a cada uno de los miembros de un "joint venture" o consorcio (asociación de dos o más empresas) y a cada empresa que se proponga para subcontratar parte del trabajo.

- El banco tiene reglamentos detallados relativos a los requisitos de nacionalidad que se proporcionan a pedido.

Dichos reglamentos estarán también disponibles de parte del prestatario en los documentos de precalificación.

E. P. 23

CONCURSO N°

ANEXO IV

ORIGEN DE LOS BIENES.

Respecto de los materiales y/o equipos por adquirir, se entenderá que su "origen" es el país en que dichos materiales y/o equipos se han extraído, cultivado o producido, ya sea mediante manufactura, procesamiento o montaje. El origen de un artículo "producido" es necesariamente el país en el cual, por conducto de dicha manufactura, procesamiento o montaje, se elabora otro producto, comercialmente reconocido, que difiere sustancialmente en sus características básicas, finalidad o utilidad de cualquiera de sus componentes importados, para determinar el origen de esos bienes y equipos, es indiferente la nacionalidad de la firma que produce o vende los bienes o equipos.

CONCURSO Nº

ANEXO

MARGEN DE PREFERENCIA.

El Banco autoriza un determinado margen de preferencia para los bienes producidos en el país prestatario, en las adquisiciones realizadas mediante licitaciones internacionales, sobre la base de las siguientes reglas:

(a) Un bien se considerará de origen local cuando el costo de los materiales, mano de obra y servicios locales empleados en su fabricación representan no menos del 40% del producto terminado.

(b) En la comparación de las ofertas locales y extranjeras, el precio propuesto u ofrecido de artículos de origen nacional será el precio de entrega en el sitio del proyecto, una vez deducidos: (i) los derechos de importación pagados sobre materias primas principales o componentes manufacturados y; (ii) los impuestos nacionales sobre ventas, al consumo y al valor agregado, incorporados al costo del artículo o artículos que se ofrezcan. El proponente local proporcionará la prueba de las cantidades a deducir de conformidad con los incisos (i) y (ii) que anteceden. El precio propuesto u ofrecido del extranjero será el precio CIF (excluyendo los derechos de importación, los consulares y los portuarios) al cual se agregarán los gastos de manipuleo en el puerto y el transporte local del puerto o frontera al sitio del proyecto.

En la adjudicación de licitaciones, el prestatario podrá agregar un determinado margen de preferencia o el derecho aduanero real, según cual sea menos, al precio CIF de las ofertas extranjeras expresadas en el equivalente de su moneda nacional. Para los contratos de construcción no se autoriza un margen de preferencia en favor de los Contratistas nacionales que compitan con Contratistas extranjeros.

ESPECIFICACIONES COMPLEMENTARIAS PARA LA CONSTRUCCION DE ROMPEOLAS, ESCOLLERAS, ESPIGONES Y PEDRAPLENES.

1.- LOCALIZACION DEL PROYECTO.

2.- DESCRIPCION DEL PROYECTO.

3.- PLANOS DEL PROYECTO.

4.- SUMINISTRO DE MATERIALES.

5.- ESTRUCTURA DE ENROCAMIENTOS.

6.- EQUIPO REQUERIDO.

7.- PLAZOS PARA INICIAR Y TERMINAR LA OBRA.

8.- PRECIOS UNITARIOS.

9.- LIQUIDACIONES PARCIALES Y FINAL DE OBRA.

10.- SANCIONES

11.- PERSONAL DEL CONTRATISTA.

12.- CONCLUSIONES.

ESPECIFICACIONES COMPLEMENTARIAS PARA LA CONSTRUCCION DE ROMPEOLAS, ESPIGONES Y PEDRAPLENES.

D E F I N I C I O N.

Las especificaciones complementarias para la construcción de rompeolas, escolleras, espigones y pedraplenes, regirán en todos los contratos que al respecto la Secretaría de Comunicaciones y Transportes Subsecretaría de Infraestructura, por conducto de la Dirección General de Obras Marítimas. Para cada proyecto se establecerán " Especificaciones Particulares" estas contendrán toda la información y requerimientos relativos, las que justamente con estas "Especificaciones Complementarias " y planos formarán parte integrante del proyecto. Cuando estas " Especificaciones Complementarias " , se contrapongan en alguna parte a las " Especificaciones Particulares", serán estas últimas las que rijan.

Son estructuras formadas por capas de piedras naturales de diferentes tamaños, protegidas por una o más cubiertas de piedra de un mayor peso colocadas cuidadosamente para obtener secciones de forma trapecial.

La cubierta mencionada anteriormente puede estar formada también por elementos precolados como tetrápodos, bloques de concreto, doms, stabilits, etc.

1.- LOCALIZACION DEL PROYECTO.

La Secretaría proporcionará todo lo relativo a la localización del proyecto, así también una información de los medios de comunicación, servicios públicos y servicios sociales con los que cuente la localidad donde tendrá verificativo la obra.

2.- DESCRIPCION DEL PROYECTO.

2.1. Trabajos por ejecutar.- La Secretaría proporcionará una información lo más completa posible, referente a la construcción de rompeolas, escolleras, espigones y pedraplenes de acuerdo con el proyecto de que se trate.

Será definido todo lo relativo a las operaciones de extracción, clasificación, carga, transporte y colocación de materiales pétreos y/o elementos fabricados que intervengan transversales la disposición de los materiales pétreos y/o elementos fabricados que integren el proyecto, definiendo el rango de pesos de roca admisibles en las diferentes capas que constituyen el enrocamiento. Se especificará todo lo relativo a la construcción y aplicación de los elementos fabricados que contengan el proyecto, entendiéndose que la fabricación de ellos estará sujeta al control de calidad, por parte de la Secretaría.

2.2.- Tipos de materiales.-

Los materiales más usados en la construcción de rompeolas, escolleras, espigones y pedraplenes son:

a).- Fragmentos de roca de diferentes tamaños, extraídos por medio de explotación de canteras.

b).- Depósitos de roca natural fragmentada --- (piedra de pepena).

c).- Elementos manufacturados tales como blocks de concreto, tetrápodos, doms. stabilits, dolos, etc.

2.3.- Datos proporcionados por la Secretaría.

El Representante indicará la localidad de las formaciones rocosas, canteras y/o bancos de agregados de donde podrá extraerse todo el material para la construcción de los enrocamientos a que se refiere el proyecto, indicando todo lo relativo a la potencia que se espera obtener de ello, densidad, dureza, absorción, resistencia al desgaste y resistencia al intemperismo.

El contratista tomará en cuenta todos los aspectos que presentan las operaciones para la extracción, clacificación, carga, transporte de materiales y colocación de ellos, para la formación de los enrocamientos del proyecto y si asi lo considera conveniente, podrá efectuar por su cuenta las exploraciones, sondeos, calas en general estudios y pruebas que juzgue necesario realizar, a fin de verificar los datos que proporcione la Secretaría en relación a calidad y potencia de las formaciones rocosas, en el entendido que la Secretaría al proporcionar datos de formaciones rocosas para extracción de material para el proyecto, lo hace como guía para la formación de precios unitarios. Podrá haber cambios posteriormente en relación con la localización de dichas zonas rocosas, si es que estos se justifican.

La Secretaría no garantiza la exactitud de los sondeos geológicos entregados al Contratista, si no como anteriormente se dijo los proporciona únicamente como guía.

2.4.- Trazos y niveles.-

En todos los casos de estas especificaciones que se indiquen trazos y niveles del proyecto, estó significará:

a).-Las líneas, niveles, acotaciones y en general todas las indicaciones que aparecen en los planos.

b).- Lo contenido en el inciso anterior, para con las modificaciones que en su caso haga el Representante.

c).- Las líneas, niveles o indicaciones que establezcan el Representante directamente, sin estar anotadas en ningún plano, los que serán proporcionados al Contratista y registrado en la bitácora de obra.

2.5.- Programa de trabajo.-

El Contratista estará obligado a proponer el programa de trabajo, el que se deberá sujetar al plazo establecido por la Secretaría, de tal forma que implique todas las operaciones, para realizar el proyecto, el que contendrá el suministro y secuencia de colocación del material en el enrocamiento, atendiendo a los volúmenes correspondientes de los diferentes tipos de roca que requieran las capas que integran el proyecto, y así mismo cuando se utilicen elementos fabricados para la coraza, para el cuerpo y/o para el morro del enrocamiento. El programa será revisado por la Secretaría sin que exista por parte de esta última, obligación de aceptarlo, pudiendo en su caso establecer y aprobar el programa de operaciones que más convenga a sus intereses.

La Secretaría está facultada para ordenar la alteración del programa de operaciones aprobado al Contratista en lo concerniente a suministro de materiales y colocación en la obra, cuando así lo considere necesario.

El Contratista podrá sugerir alguna modificación estructural al proyecto, pero no procederá a efectuar cambio alguno hasta obtener la autorización escrita de la Secretaría.

3.- PLANOS.-

3.1.- planos del proyecto.

La Secretaría proporcionará al Contratista todos los planos que contengan la localización y estructuras de los enrocamientos a construir. En la estructura de los enrocamientos estarán definidas las líneas y niveles de las capas integrantes de los mismos, indicando los tipos de roca que corresponden a cada capa. Las estructuras de los enrocamientos estarán referidas a la cota 0.00 según el N.B.M.I. (Nivel de Bajamar Media Inferior), en el Océano Pacífico. N.B.M. (Nivel de Bajamar Media), en el Golfo de México y Mar Caribe.

En planos complementarios, estarán definidos, si es el caso los elementos manufacturados que deberán ser empleados para formar la capa final de protección (coraza) cuando no se disponga de roca adecuada para formarla. Sus elementos fabricados pueden ser bloques de concreto ciclópeo, tetrápodos, doms, stabilits, dolos, etc., quedará definido el sistema de suspensión usado para cada caso en particular, para manejo y colocación de estos elementos.

Las secciones tipo de los enrocamientos estarán referidas a los niveles antes citados según sea el caso donde se localice el sitio de la obra y contendrá la batimería contígua, indicando el tipo de suelo, donde serán desplantados los mismos.

Se proporcionará también el regimen de los vientos de la localidad, pudiendo el Contratista tomar las medidas necesarias, para definir los períodos del mejor tiempo aprovechable para el desarrollo de los trabajos, debiendo tomarlos en cuenta para la elaboración de su programa de trabajo.

4.- SUMINISTRO DE MATERIALES.-

4.1.-datos informativos.-

El Contratista deberá verificar de inmediato y antes de indicar la operación para la extracción de roca, todos los datos relativos a la calidad y potencia de las zonas rocosas que establece el proyecto para la extracción de materiales petreos, así como tendrá en cuenta lo siguiente:

a).- Los fragmentos de roca para aplicarse al proyecto en su estado natural pueden ser:

1.- Igneas In-trusivas.....Granito.
(Grano Grueso).....Diorita.

2.- Ingenias Extrusivas.....Riolita
(Grano fino).....Andesita.
Basalto.
Toba
Brecha Volcánica.

3.- Sedimentarias.....Calizas.
Travertino.
Arenisca
Conglomerado
Brecha.

4.- Metamórficas.....Gnesis.

Las normas que deberán satisfacer las rocas naturales que se apliquen al proyecto, son las siguientes:

1.-Resistencia a la Compresión en estado húmedo.....
.....150 kg/cm² minimo.

2.- Resistencia a la compresión en estado húmedo aplicando la carga paralelamente a los planos de estratificación, cuando los haya.....100 kg/cm2 minimo.

3.- Absorción en por ciento.....4 maximo.

4.- Densidad.....2.3 minimo

5.- Resistencia al intemperismo acelerado (sanidad), porciento de pérdida en peso 10 máximo.

6.- Resistencia al desgaste determinado - por la prueba de los angeles, en porciento..... 40 máximo.

b(.-La determinación de la potencia de las zonas rocosas que indique el proyecto a fin de comprobar si satisfacen los requerimientos establecidos.

4.2.- Características de los bancos.-

Las formaciones pétreas pueden presentar frentes de ataque naturales, que por su altura, longitud y aspectos que ofrecen en si mismo, faciliten definir la conveniencia de su - explotación. En otros casos el Contratista deberá realizar pruebas según se requiera, para establecer la conveniencia de su explotación, cuando así lo ordene la Secretaría.

El Contratista podrá iniciar la explotación de bancos rocosos, una vez obtenida la autorización del Representante -- sin que esta autorización releve al Contratista de sus responsabilidades en relación con las operaciones y equipo - que emplee para la extracción del material.

Cuando el contratista obtenga el consentimiento del Representante para iniciar la explotación de una formación rocosa, procederá con su personal a efectuar el levantamiento topográfico del banco, en presencia del personal de la Secretaría que indique el Representante. El levantamiento - topográfico se utilizará también para fines de pago de desmontes y despalmes necesarios en su caso. El seccionamiento de bancos y los sondeos geológicos, podrán proporcionar -- aproximadamente la potencia del banco. Durante todo el proceso de extracción se continuará el control topográfico, a fin de determinar el volúmen de desperdicio pudiendo establecer la relación entre el volúmen de roca fija media en la cantera y el volúmen geométrico del rompeolas, El Representante indicará al Contratista la secuela que deberá seguir en los desmontes y despalmes de los bancos.

4.3.- Extracción del Material.-

La explotación de la cantera la debe realizar el Contratista, en forma tal, que la obtención de materiales concuerde con los requerimientos de la secuencia establecida para la construcción de las diferentes capas de roca que integran la estructura del enrocamiento, logrando el mínimo de material de desperdicio.

Los procedimientos clásicos de extracción de materiales por medio de " explotación por bancos ", aún se considerán vigentes para obtener resultados satisfactorios, con un mínimo de desperdicios. Aún cuando el proyecto requiera grandes volúmenes de roca, no se recomienda emplear sistemas de extracción que produzcan grandes masas de roca derrumbada en una sola voladura tales como el de túneles de coyote que -- emplea grandes cargas de explosivos depositadas en ellos, o cualquier otro sistema similar ya que estos aumentos considerablemente, con el consiguiente incremento del material no utilizado.

La fragmentación de una sección tronco piramidal a lo largo del frente de atraque y en la base del mismo, de dimensiones convensionales, puede proporcionar resultados satisfactorios, en una voladura.

El Contratista deberá seleccionar la piedra que debe ser moneada, la cual será aproximadamente el 7.5% del volúmen extraído .

Cada formación rocosa presenta condiciones típicas propias y en consecuencia no se considera conveniente delimitar sistemas de extracción de material a emplear y solamente se establece lo anterior con carácter informativo; queda a la experiencia del Contratista elegir los sistemas óptimos para obtener el material requerido, pero se establece claramente que si por el empleo de operaciones inadecuadas en la extracción de roca, el Contratista llegará a fisurar la formación rocosa en explotación, en forma tal que inutilizare el banco, para seguirlo explotando eficientemente, será por su -- cuenta el importe de los trabajos que resulten necesarios, para localizar acondicionar y abrir una nueva formación rocosa que resulte aceptable, así como todos los importes que resulten de acondicionamiento de caminos nuevos de acceso y obras complementarias. Si el nuevo banco de encuentra a mayor distancia del deteriorado, la Secretaría no pagará el - aumento que resulte por tonelada kilómetro de acarreo.

Esta circunstancia no justificará ninguna modificación a la fecha de terminación aprobada.

4.4.- Clásificación del material en el Banco.-

El Contratista después de cada tronada y antes de efectuar operaciones de carga debe clásificar la roca derrumbada de acuerdo con los diferentes rangos que marque el proyecto y de acuerdo con el inciso 5.4 "Tolerancias" de estas especificaciones. El Contratista podrá cargar el material de - roca para transportarlo a la obra solamente de los almace-

El Contratista deberá contar con el equipo adecuado para la carga de los elementos de la coraza.

Si la transportación del material fuera por carretera, - El Contratista proporcionará el equipo de transporte requerido, para cumplir con el programa de trabajo. El Contratista construirá, acondicionará, reparará y conservará los caminos de acceso del banco y/o bancos, a los lugares de descarga. Si utilizarase caminos federales, estatales o vecinales, gestionará con quién corresponda la - autorización de servicio, teniendo en cuenta que deberá repararlos y conservarlos, mie tras lo use igualmente -- deberá obtener la información sobre el tipo de vehículo - máximo que podrá transitar sobre el camino a recorrer. Si el caso lo requiere, el Contratista construirá las líneas, ramales y espuelas de ferrocarril, para el transporte del material debiéndo el Contratista efectuar los convenios para estos servicios con los organismos que correspondan, teniendo en cuenta el cumplimiento del programa - de trabajo.

4.10.- Elementos Precolados.-

Cuando no pueda obtenerse en la localidad roca del peso - requerido para la coraza de protección y/o morro, podrá - ser sustituido por elementos precolados, tales como bloques de concreto, tetrápodos, stabilits, doms, etc., los que quedarán definidos según lo requiera el proyecto, en todo lo relativo a su fabricación, almacenamiento, carga, acarreo y colocación en la obra. Para la fabricación del concreto ciclópeo, concreto armado y concreto simple de - de los elementos de referencia; el Contratista se sujetará a "Las Normas de Construcción de esta Secretaría", y a las Especificaciones Particulares. El Contratista acondicionará las mesas de colado que se requieran para fabricar los elementos precolados del proyecto. Los moldes empleados estarán debidamente dimensionados y estructurados en tal forma, que eviten deformaciones y/o escurrimientos al efectuarse el colado.

Para cada caso en particular la Secretaría indicará el tipo de cemento a utilizar en la construcción de los elementos precolados.

5.- ESTRUCTURA DE ENROCAMIENTOS.-

5.1.- Materiales Constitutivos.-

La estructura de los enrocamientos para rompeolas, escolleras, espigones y pedraplenes generalmente tienen la forma trapecial, y están constituidos por un núcleo formado por material pétreo muy fragmentado cuyo rango de peso pueda considerarse entre P/200 a P/600, siendo P el peso de roca considerada para la coraza. Al núcleo lo cubre una capa secundaria de protección, las rocas que forman esta capa pueden tener un rango de peso de P/10 a P/15. A la capa secundaria la cubre la capa principal o sea coraza, constituida por roca de peso según proyecto. La base superior del trapecio, según lo especifique el proyecto, puede rematar en un coronamiento de dimensión y tipo estructural que quedará definido en el mismo, los rompeolas y escolleras pueden terminar al final de su desarrollo en su ensanchamiento (morro).

5.2.- Secuencia en la Colocación del Material.-

El primer paso para la construcción del enrocamiento es la formación del núcleo, se pueden presentar dos casos cuando el nivel de la corona del núcleo esté colado al nivel (0.00) o más bajo y cuando está se encuentre arriba del nivel (0.00).

La elevación 0.00 estará referida al Nivel de Marea Baja Media Inferior en el océano Pacífico y al Nivel de Marea Baja Media en el Golfo de México y Mar Caribe.

En el primer caso para acomodar el material a sus líneas y niveles, es necesario que el Contratista utilice chalanes de volteo manejadas con grúa, o cualquier otro sistema similar que pueda depositar el material a las líneas niveles bajo el agua.

En el segundo caso, el núcleo podrá construirse en una parte con camiones a volteo, pero los taludes finalmente deberán ser terminados colocando la piedra, utilizando charolas de volteo manejadas con grúa o cualquier otro sistema similar.

Para evitar que la acción del oleaje desaloje el material de los taludes y/o corona del núcleo, al ser terminado un cierto tramo de éste, el contratista procederá de inmediato a cubrirlo con la roca de la capa secundaria que le corresponde y cubrirá toda el área terminada del mismo, de manera tal, que la roca de la capa secundaria no permita se sea desalojado el material del ----

tramo de núcleo en cuestión, llevándose a término totalmente el espesor del tramo de capa secundaria en término totalmente el espesor del tramo de capa secundaria en la zona de que se trate. Si la construcción del enrocamiento se está realizando con un cierto grado de agitación del mar, deberá procederse a recubrir de inmediato la capa secundaria con la capa de roca de coraza, según lo indique el Representante.

El Representante determinará en cada caso y dependiendo de la violencia del oleaje, la longitud de núcleo que pueda permanecer sujeta a esta sección, sin ser removida por ella, antes de ser protegida por la capa secundaria. El Contratista cumplirá estrictamente con lo especificado por el Representante a este respecto.

Antes de continuar con la construcción de un nuevo tramo de núcleo, se cubrirá con la roca de la capa de coraza toda el área terminada de la capa secundaria y una vez cubierta esta, como lo indique el proyecto , se continuará la construcción de un nuevo tramo de núcleo, repitiendo la secuencia ya descrita.

En la colocación de las capas que forman el enrocamiento del proyecto, siempre se empleará la grúa adecuada que garantice que estas sean construídas a las líneas y niveles de proyecto. El uso de tractores en la colocación de roca, serán únicamente empleados en acciones complementarias de las operaciones fundamentales de las grúas. Todas las capas del enrocamiento siempre deberán construirse colocando los elementos que las forman del pié del talud hacia la corona y nunca empujando éstos hacia los taludes. Se deberá evitar siempre toda operación que en alguna forma tienda a degradar alguna capa. En algunos casos será necesario el empleo complementario de grúas flotantes, el contratista deberá tener en cuenta estas circunstancias.

La parte superior del rompeolas no terminado puede usarse como acceso del equipo de transporte, teniéndose encuenta que antes de que se coloque piedra adicional, los materiales usados en el acceso anterior deberán removerse dejando limpio el mismo. Esta remoción permitirá tener superficies de piedra limpia, de tal manera que una capa nueva quede colocada directamente sobre la anterior amarrada perfectamente.

A fin de disponer de un ancho mayor para las maniobras podrá permitirse que el núcleo se construya en dos etapas de acuerdo con las condiciones reinantes en el mar y las

ordenes del Representante, pudiendose dejar de tramo en tramo retornos para facilitar las maniobras del equipo de acarreo y colocación, pero invariablemente deberá protegerse con la capa secuandaria y la coraza.

Si la colocación de piedra de núcleo se realiza en dos etapas las capas de protección de éste se llevarán hasta la misma altura.

Si fuera indispensable suspender temporalmente la obra, deberá protegerse lo suficiente el área frontal expuesta del nucleo por medio de la capa secundaria y esta a su vez con roca de coraza, según lo determine el Representante con el fin de evitar que la acción del oleaje, mientras dure la suspensión, desaloje los materiales del núcleo y capa secundaria; al reanudarse la obra, se continuará la construcción siguiendo la misma secuela aquí descrita.

Si por la falta de observación de parte del contratista a la secuencia de operaciones aquí establecida motiva que se computen valores a la tolerancia que se fije para el tonelaje total de roca empleada para el proyecto, la diferencia no será pagada al contratista.

Cuando se trate de la construcción de rompeolas que no arranquen de tierra, sino que sean paralelos a la costa o formen ciertos ángulos con ella, deberá el contratista contemplar la misma secuencia aquí descrita, para la construcción de enrocamientos.

5.3.- Consolidación de la coraza.-

Cuando se dispone en la localidad de roca lo suficientemente pesada, requerida para la construcción de la coraza, un sistema de consolidación de estas, consiste en rellenar los espacios que quedan entre las rocas al construirlas, con una mezcla asfáltica, similar a la empleada para la construcción de carpetas asfálticas de carreteras a mayor temperatura que le permita fluir dentro del agua y rellenar los vacios entre las rocas de la coraza.

Después de colar la primera capa de piedra de coraza, se procederá a colocar la mezcla asfáltica hirviente utilizando cajas metálicas de volteo, cubos de fondo compuerta, manejando con grúa o cualquier otro sistema similar. Previamente serán señalados los lugares de descarga del material en la zona de coraza bajo el agua, en la zona fuera del agua estarán a la vista los lugares de descarga si se requiere una segunda capa de coraza se procederá en forma similar.

Las temperaturas a que se colocará la mezcla variarán de 180°C. a 230°C. según el clima y la profundidad de colocación de la mezcla.

El proporcionamiento de la mezcla a reserva que el proyecto indique lo contrario.será:

75% de arena fina, 8% de gravilla con tamaño máximo de 6 mm. y 17% de asfalto No. 7 o también: 48% de arena para concreto hidráulico, 33% de grava triturada, 10% de gravilla menor de 6 mm. y 9% de asfalto. En cada caso la Secretaría indicará - el proporcionamiento de la mezcla.

La consolidación de enrocamiento con la mezcla asfáltica se empleará también por ser un método sencillo y económico, para reparar rompeolas, insuficientes o deterioradas.

5.4.- Cuantificación de proyecto.-

Para cuantificar el volúmen aproximado de roca fija que requiere la construcción del proyecto, se procederá como sigue:

Solo cuando se omita las cantidades de obra del "Catálogo de Conceptos..." Se indicará el criterio de cuantificación que - el Contratista empleará.

El obtener el volúmen total de roca fija aplicada al proyecto -- dará un índice de la potencia que debe tener el banco.

5.5.- CUBICACION DE LA ROCA.

La cubicación de la obra se realizará para efectos de medición y pago, cuantificando el tonelaje total de material útil pesado en báscula, aplicado al proyecto; para el caso cuando así se establesca el Contratista proporcionará, instalará y - pondrá en operaciones la báscula y/o las básculas del tipo y capacidad requerida para pesar la roca en los equipos aprobados para el transporte del material a la obra, las básculas - ya instaladas deberán ser verificadas por las autoridades correspondientes de la Secretaría de Comercio y Fomento Industrial, en presencia del Representante quién en su caso, ordenará el uso de ellas. El Representante podrá ordenar al Contratista, cuantas veces lo requiera durante el desarrollo de los trabajos, proceda a verificar el correcto funcionamiento de las básculas. Independientemente de la verificación que -- realice el Contratista, este debe afectar por medio de los -- técnicos fabricantes de la báscula, revisiones y ajustes que se requieran a la misma con lapsos de 2 meses entre cada revisión, o con la periodicidad que recomiendan los fabricantes. En caso de que el Representante encontrará funcionamiento defectuoso de alguna báscula suspenderá de inmediato su operación, hasta no ser reparada o ajustada por el Contratista a - satisfacción del Representante de acuerdo con los requerimientos. El Contratista se sujetará a todas y cada una de las partes del instructivo para el control de pesaje, que se adjunta a estas Especificaciones.

5.6.- Localización de básculas.-

La báscula y/o básculas requeridas para registrar el peso del material útil aplicado al proyecto. deberán instalarse lo más cercano posible a la localización del mismo y en la ruta de los vehículos que transporten el material a los enrocamientos.

La marca y tipo de básculas que proporcione el Contratista deberán ser aprobadas por la Secretaría, para el caso, el Contratista proporcionará a la misma con la debida anticipación los planos y especificaciones de estos equipos, y una vez aprobados por la Secretaría procederá el Contratista a su adquisición e instalación en el sitio de la obra de acuerdo como lo indique el Representante. Las básculas estarán capacitadas para registrar en una sola pesada las unidades de mayor peso que transporten el material para el proyecto. Podrán ser si asi lo indica el proyecto básculas de ejes. Estarán provistas si asi lo indica el proyecto de caseta de control para el personal tanto de la Secretaría como del Contratista, con todo el equipo y material necesario para el registro de pesos (Taras de verificación) Las taras de verificación deberán ser como minimo de la vigésima parte de la capacidad máxima de la báscula.

5.7.- La instalación de las básculas.-

Las efectuará el Contratista de acuerdo con los planos, especificaciones y recomendaciones que establezca la casa proveedora de estos equipos. El Representante supervisará los trabajos de instalación de básculas, pero su presencia no eximirá al Contratista de la responsabilidad de eficiencia de las instalaciones. El fabricante o vendedor supervisará la instalación de la báscula, así tambien cualquier cambio de lugar de ella.

5.8.- Verificación.-

Una vez terminada la instalación de la o las básculas, el Contratista deberá solicitar a la Secretaría de Comercio y Fomento Industrial la verificación de éstas, debiendo obtener el documento oficial de aprobación proporcionando el original al Representante.

Las recomendaciones que haga el proveedor para efectuar el mantenimiento de estos equipos, los deberá llevar a cabo estrictamente el Contratista a satisfacción del Representante.

Siempre que el Representante lo considere conveniente ordenará al Contratista realice la verificación oficial de las básculas si encontrare alguna irregularidad procederá como lo indica la presente especificación en el párrafo 5.5 CUBICACION DE LA ROCA.

5.9.- Peso de Materiales.-

5.9.1.- Personal.-

La Secretaría mantendrá personal que nombrará el Representante para cada jornada de trabajo, debiendo permanecer en la caseta de control del pesaje, todo el tiempo correspondiente quién efectuará la operación de pesaje de las unidades que contengan la roca destinada al proyecto; ningun pesaje será considerado si no se realiza en presencia y autorizado por el auxiliar del Representante. El Contratista mantendrá un Representante autorizado por él para que esté presente en el pesado del material dando su conformidad al registro de peso del material que se pese para ser colado en la obra.

5.9.2.- Control de peso.-

Dependiendo de cada caso en particular, el Representante le indicará al Contratista en que forma se realizará el control de las unidades vacías, ya sea efectuando el peso de las mismas después de descargar cada viaje o sacando un promedio de los pesos del vehículo en el primero y último viaje que realice cada día. Estos datos quedarán registrados en una libreta especial que el auxiliar del Representante tendrá permanentemente en la caseta de control.

Cada unidad debe pesarse, cuando esta se encuentre estacionada totalmente dentro de la plataforma para pesaje o centrando cada eje si la báscula por sus dimensiones no admite el vehículo completo.

El auxiliar del Representante antes de que una unidad entre a ser pesada, se cerciorará que el fiel de la báscula se encuentre en su posición de equilibrio y ajustandolo si se hace necesario, una vez logrado esto, dara la orden de entrada para ser pesada, a la plataforma de la báscula. El registro de peso deberá ser marcado cuando la unidad se encuentre totalmente estacionada y toda ella se encuentre dentro de la plataforma de peso y sensiblemente centrada, a no ser que la báscula sea de ejes en cuyo caso se pesará cada uno de estos.

Las tarjetas de registro para el peso, estarán foliadas y el Contratista entregará al Representante por medio de un escrito la cantidad de tarjetas suficientes para el consumo de un mes de trabajo, en el escrito indicará la cantidad de tarjetas y el rango de folio correspondiente; la entrega de tarjetas la deberá hacer el contratista con toda puntualidad, ya que si llegaren a faltar tarjetas de registro en la caseta de control, se suspendería el pesaje de material, con el consiguiente perjuicio en el retraso de las obras, no justificable para el Contratista.

Las tarjetas de pesaje tendrán original y copia, al marcarse la tarjeta, quedará impresa la cantidad que indica el peso, en el original y copia; en los renglones correspondientes el inspector del Representante escribirá con tinta la fecha, hora, número de la unidad, tipo de roca que contiene, peso correspondiente a la unidad vacía, cálculo de peso correspondiente al material, al deducir del peso bruto al peso de la unidad vacía y su firma; estos mismos datos estarán contenidos en la copia de la tarjeta, el original de la tarjeta quedará en poder del inspector del Representante y la copia de ella en poder del Representante del Contratista.

Diariamente se registrará en la Bitácora el total de roca en las secciones transversales correspondientes, mismo que fue acarreado a la obra.

Si la báscula es de ejes, el vehículo deberá encontrarse sensiblemente en posición horizontal, al pesar cada uno de ellos sobre la báscula, en caso de no ser asi el Representante no permitirá se efectue el pesaje de las unidades.

El inspector registrará todas las tarjetas que se hubiesen operado en su turno, en la libreta de control, incluyendo todos los datos correspondientes de cada tarjeta, y así mismo elaborará una relación, vaciando todos los datos de las tarjetas en una forma especial, en la cual tendrá una columna para registrar el peso total de material y tipo de roca, colocada en la obra durante el turno; esta relación deberá estar debidamente firmada por el inspector del Representante y por el Representante autorizado del Contratista, éste último expresará su conformidad con los datos en la relación. Una vez terminado el turno, el inspector entregará al Representante la relación del pesaje del turno, debidamente firmada, adjuntando todos los originales de las tarjetas de pesaje registrados en el turno.

En Contratista se obliga a proporcionar la papelería requerida para el control del peso de material, como lo indique el Representante.

Cuando sean empleados elementos precolados porque así lo indique el proyecto, el Representante decidirá si deben o no ser pesados estos elementos, en caso afirmativo, él mismo reglamentará lo necesario para el control de peso.

Los elementos manufacturados sólo despúes de 28 días de fabricados se podrán manejar, transportar y colocar en el sitio de la obra.

En el caso de elementos precolados por lo general su medición será por metro cúbico o por pieza prefabricada, para lo cuál el Contratista deberá considerar todas las partes que integran el colado de la pieza en cuestión así como su colocación en el sitio de la obra.

Una vez efectuada la estimación mensual el Representante archivará cuidadosamente todos los comprobantes y la relación de ellos que servirán para la formulación de la estimación respectiva, esta documentación podrá utilizarse para efectuar cualquier verificación que fuera necesario realizar en cualquier tiempo.

El Contratista se compromete a realizar la obra, con la cuantificación que finalmente resulte, y ya sea de aumento o disminución de los volúmenes de obra, sin que tenga derecho a modificación alguna a los precios unitarios aprobados.

5.10.- Tolerancias.-

La tolerancia en los perfiles del núcleo no será mayor en más o menos de dos veces el diámetro mayor de la roca de más peso empleada para esta capa, en las áreas bajo agua y en zonas aisladas. En las áreas fuera del agua será de una y media veces el diametro indicado en más o en menos y en zonas aisladas únicamente.

La tolerancia en los perfiles de la capa secundaria de protección no será mayor en más o menos a una vez el diametro mayor de la roca de más peso especificada para esa capa, en zonas aisladas y en las áreas dentro y fuera del agua.

La tolerancia en los perfiles finales de la capa de coraza no será mayor en más o en menos la mitad del diametro mayor de la roca de más peso especificado para esta capa, tanto en áreas dentro como fuera del agua, pero en zonas aisladas únicamente.

Con el fin de lograr una mayor trabazón del material se admitirán las siguientes tolerancias máximas:

a).- Que el material para el núcleo no contenga más del 20% en peso del mismo, en tamaños menores de los que abarquen el rango, teniendose como límite inferior piedra de la mitad del peso del valor mínimo del rango.

b).- Que el material para la capa secundaria, no contenga más del 15% en peso del mismo, en tamaños menores de los que abarque el rango, teniendose como límites inferior piedra de la mitad del peso del valor minimo del rango.

c.). Que el material para la coraza, no contenga más del 1[illegible] en peso del mismo, en tamaños menores de los que abarque el rango, teniendose como límite inferior piedra de la mitad del peso del valor mínimo del rango.

En la coraza puede admitirse roca de mayor peso del especificado, como lo indique el Representante.

No se permitirá el uso de tierra suelta, arena y polvo de roca en cantidades mayores de 5% (cinco por ciento) en peso de los materiales pétreos.

El exceso de tonelaje de roca, del calculado para el proyecto de acuerdo con el "concepto 5.4" cuantificación del proyecto, estará comprendido entre el 5% y el 15% sobre "el" volúmen de roca útil aplicada al proyecto multiplicado por el peso especificado del material empleado. Dicho exceso de tonelaje que puede ser provocado por asentamiento del enrocamiento debido a su peso propio o al acomodamiento de los bloques de piedra, se determinará por medio de secciones transversales que se levantarán al dar por terminada la obra.

En casos muy esporádicos por ser un volúmen pequeño y para piedra chica, se podrá cubicar ésta por m3 medido a bordo de los vehículos de transporte en el sitio de la obra, a los cuales previamente se les habrán cubicado sus capacidades marcando claramente el nivel de cada uno de ellos. La secretaría indicará al contratista estos casos particulares.

Por ningun motivo se permitirá que el contratista coloque piedra de pesos diferentes a las especificaciones y fuera de las tolerancias establecidas, sobre todo durante la construcción del núcleo y la capa secundaria. El Representante podrá ordenar al contratista el retiro de la roca que no cumpla con lo establecido, quién la colocará donde se le indique, o la retirará definitivamente de la obra, si asi le fuere ordenado.

La profundidad de desplante de la roca de la coraza tanto en el talud exterior, como en el interior que marca el proyecto, podrá iniciarse perfectamente más abajo de lo indicado en el proyecto, esta sobre profundidad, no seráma yor, del valor de un diámetro de la roca empleada, en la coraza.

Para la capa de coraza tanto del cuerpo del enrocamiento como en la extremidad de este (morro), se podrá admitir roca de mayor peso de lo especificado, según lo indique el Representante, principalmente en el talud exterior.

6.- EQUIPO REQUERIDO.

6.1.- Equipo que proporciona el Contratista.

El equipo que a continuación se relaciona, formará parte de la proposición del contratista, ya sea parcial o totalmente según lo indique la Secretaría para cada caso en particular.

a).- La extracción de material.-El equipo necesario para la extracción de roca de la cantera, consistirá principalmente en máquinas perforadas con sus correspondientes fuentes de aire comprimido, con todos sus accesorios, repuestos y materiales necesarios, así tambien, los explosivos, herramientas, accesorios y equipo que satisfaga el suministro requerido.

b).- Clasificación y carga.-El equipo necesario para clasificación, almacenamiento y carga de la roca extraída de la cantera, consistirá principalmente en grúas con garra, palas mecánicas, traxcavos, tractores y equipos auxiliares.

c).- Transporte y descarga del material.-Si el transporte del material se efectuará por carretera, de la cantera al lugar del proyecto, será requerido equipo de camiones de volteo y/o de plataformas, provistos de charolas o cajas metálicas de volteo.-Si el transporte se efectuará por ferrocarril, proporcionará las charolas de volteo necesarias, para cargar las plataformas del --- equipo ferroviario en la cantera.
En la descarga y colocación del material para - formar el enrocamiento el contratista proporcionará las grúas para descarga y colocación de roca provistas de garras para el manejo de ellas, chalanes tolva, o chalanes de cubierta plana y tractores.
Para la totalidad de los casos el contratista - deberá utilizar una grúa provista de garras o - equipo similar en la colocación del material de capa secundaria y coraza.

Todo el equipo estará en buenas condiciones de trabajo y será revisado por el Representante cuando se encuentre en los lugares de trabajo, verificando su eficiencia, capacidad y estado general que deben coincidir con lo propuesto por el contratista.

Si el Representante encontrase deficiencias o faltantes en - el equipo proporcionado por el contratista le exigirá que lo complemente y/o adecúe, púes se considerá que las deficien-- cias en el equipo del contratista, no modificarán el plazo - de terminación de las obras, no lo eximirán de las responsabilidades por incumplimiento, de acuerdo con las cláusulas - correspondientes del contrato.

Cuando se trate de fabricar elementos de concreto para la formación de la coraza o de protejer con estos elementos enrocamientos ya existentes, el contratista proporcionará todo el - equipo para extracción y adecuación de materiales pétreos para la fabricación del concreto y para fabricación de estos.-

así como para transportar y colocar en el sitio de la obra los elementos precolados que sean requeridos.

Por el hecho de que la Secretaría le otorgue al contratista el contrato no adquiere responsabilidad alguna, si el equipo, para extracción, carga, transporte y colocación del material para formar el enrocamiento proporcionado por el contratista, no sea el adecuado para la realización eficiente del proyecto.

Si el proyecto tratare de construcción de rompeolas que no arranquen de tierra, sino que sean paralelos a la costa o formen cierto ángulo con ella, el contratista proporcionara adicionalmente el equipo flotante requerido para el transporte y colocación del material que podrá consistir en grúas flotantes, chalanes de volteo, chalanes de cubierta plana y remolcadores en la cantidad y capacidad que requiera el proyecto.

6.2.- Aprovisionamientos.-

El suministro o aprovisionamiento que se haga necesario para la operación de los equipos, debe ser seleccionado por el contratista; la gestión de permisos para la adquisición de explosivos y accesorios, debe realizarla de inmediato, ya que las demoras que pudieran presentarse en la obra, como consecuencia de deficiencias en el abastecimiento de combustible, lubricantes, explosivos y accesorios, partes, repuestos, y materiales en general requeridos, no pueden justificar los retrasos que puede sufrir el programa de obra.

La Secretaría auxiliará al contratista haciendo el oficio a la Secretaría de la Defensa Nacional para que está de su autorización para el consumo de explosivos, debiéndo el contratista proporcionar a la Secretaría sus requeremientos mensuales y totales, tanto de explosivos como de accesorios.

6.3.-Inspección.-

La obra se realizará bajo las directrices generales que establezca la Secretaría y será inspeccionada por el Representante y/o por los inspectores designados por él, quienes cuidarán que se cumpla estrictamente con los requerimientos del proyecto, en los términos del contrato; el Representante llevará un registro de control de las secciones del enrocamiento, donde gráficamente quede señalado el perfil de cada capa terminada hasta la capa de coraza, en el tramo de que se trate pero la presencia de los inspectores, no relevará al contratista de su responsabilidad de efectuar correctamente la obra.

6.4.- Retiro del Equipo.-

Para hacer el retiro o sustitución parcial del equipo en el avance de la obra con relación al calendario de ejecución, el contratista deberá recabar autorización escrita del Representante, previa aprobación de los adelantos registrados.

7.- PLAZOS PARA INICIAR Y TERMINAR OBRA.-

7.1.- Plazo para iniciar y fecha de terminación.-

Se fijará la fecha de iniciación de la construcción del enrocamiento, después de entregada al contratista la orden de trabajo, fijando la fecha en que debe dar por terminada la obra, con base en los rendimientos de los equipos incluídos en su oferta y apoyado en ellos presentará su Programa de Trabajo.

Una vez presentado el Programa de Trabajo por el contratista indicando la fecha en que se dará por terminada la obra, dentro del período que se le dió, la Secretaría lo estudiará y en su caso lo aprobara y será el que en definitivo se adopte. En vista de lo cual, las sanciones por retraso en el cumplimiento del programa se harán efectivas a partir de la fecha indicada en este.

Cuando el proyecto incluya la construcción de dos o más enrocamientos, la Secretaría indicará el procedimiento a seguir en la construcción de acuerdo con el programa de obra o modificación lo que convenga a sus intereses.

7.2.- Retrasos.-

Si por causas imputables al contratista sufriere retrasos el Programa de Trabajo para la ejecución de las Obras estará obligado a operar en el tiempo extraordinario sus equipos o reforzar los mismos, hasta lograr la nivelación del mismo, siendo por su cuenta las erogaciones que le ocacionen.

7.3.- Ampliación de Plazo.-

La alteración de los programas de trabajo por incremento en la longitud de los enrocamientos o reforzamiento adicional en el morro o alguna otra zona, justificará una ampliación proporcional en el plazo para terminar la obra sobre el tiempo computado originalmente para las operaciones: no siendo justificada una ampliación de tiempo en ningún caso, excepto por alguna otra causa tal como a continuación se indica.

a).- Si por causa de fuerza mayor imputable a fenómeno meteorológico extraordinario, perturbación del orden público, asonada, etc., la obra sufre retrasos en el desarrollo del calendario fijado, el retraso será computado para ampliar el plazo de terminación en el lapso que corresponda.

8.- PRECIOS UNITARIOS.-

8.1.- Consideraciones que deben tener en cuenta el Contratista.

Para fijar el Precio Unitario en la obra de construcción de enrocamiento, el contratista deberá tener en cuenta como factores primordiales.

a).- Las características de la formación pétreas que se requiera explotar para la obtención de los materiales requeridos para la construcción de rompeolas, escolleras, espigones y pedraplenes. Las características de explotación para la extracción de los agregados necesarios para la fabricación de concreto, cuando sea requerido emplear elementos precolados para la capa de la coraza,

b).- Las distancias y caminos de acceso de las formaciones rocosas a la localización de los enrocamientos, y de los bancos a los patios para la fabricación de elementos de concreto.

c).- El programa de operaciones que se establezca para coordinar el ataque de las formaciones rocosas y/o bancos en concordancia con la secuencia requerida para la construcción de los enrocamientos.

d).- Explotación y apertura de nuevos bancos que puedan ser necesarios.

e).- La explotación de una formación rocosa requiere desmontes, levantamientos topográficos, despalmes y abrir el frente y/o frentes necesarios de ataque extracción de la roca, acondicionamiento de patios de clasificación de materiales, clasificación del material.

f).- Construcción de caminos, acondicionamiento de los existentes, reparación y conservación de todas las vías de comunicación entre los bancos y la obra.
Acondicionamiento y uso de vías férreas.

g).- Programa de las operaciones a seguir en la extracción, carga, trasnportes y colocación del material rocas y/o elementos prefabricados.

8.2.- Lo que incluyen los Precios Unitarios.-

Por el Precio Unitario del contrato por tonelada de roca natural extraída, clásificada y cargada para la construcción de rompeolas, escolleras, espigones y pedraplenes se cubrirá el importe de los gastos que requiera erogar el contratista para condicionar y explotar las formaciones rocosas, clasificación y carga de material, pago de alquiler o derechos por extracción del material del banco y/o bancos.

Por el Precio Unitario del contrato para transportar el material pétreo para la ejecución del proyecto, se cubrirá el importe de los gastos que requiera erogar el contratista para transportar el material del banco y/o bancos al lugar de desarrollo del proyecto, que incluya, además del importe del acarreo en sí, los gastos para acondicionar, construir, reparar y conservar los caminos de acceso de los bancos al proyecto, así como pagos por servidumbre que fuese necesario cubrir, -- permisos, etc., cuando se trate de utilizar vía férreas existentes para transportar el material del banco a la obra, se - cubrirá al contratista el importe de los gastos que requiera erogar para construir, espuelas o laderos necesarios, fletes ferroviarrios, alquiler de plataformas, pago de fletes a - - F.F.C.C., así como levantado de vías y retiro del material, al terminar el trabajo si asi lo requiere la Secretaría.

Por el Precio Unitario de descarga del material y construcción de los enrocamientos que implique el proyecto se cubrirá al contratista el importe de los gastos que requiera erogar - para pesar el material, consistente en suministro, instala--- ción y puesta en servicio de báscula, conservación y verificación de ella durante todo tiempo el importe de los gastos de descarga y colocación del material de acuerdo con el proyecto para formar los enrocamientos.

Cuando se trate de rompeolas que no arranquen de tierra, sino sean paralelos o formen cierto ángulo con la costa se le cubrirá al contratista además de lo anterior los gastos por carga de chalanes y transporte y colocación del material en el - rompeolas.

Para cada caso en particular la Secretaría le indicará al contratista que actividades de las anteriores enumeradas para -- extraer la piedra, transportarla y colocarla deberá considerar en cada concepto de trabajo que se solicite, pudiendo -- ser uno o varios conceptos según el caso.

Cuando se requiera el uso de elementos precolados para la coraza en el cuerpo del rompeolas y/o el morro, se le cubrirá al contratista el importe de todos los gastos por la extracción, trituración, clasificación, lavado, carga, acarreo, descarga de los agregados para la fabricación de elementos así como todos los gastos para la construcción, almacenamiento, carga, transporte y colocación de ellos en la obra. El importe de los gastos que enrogue el contratista para acondicionar terrenos, construcción de mesas para colado, demolición de estas y limpieza de terrenos al terminar las obras, deberá cargarlo al Precio Unitario de fabricación de los elementos. Igualmente las actividades anteriores podrán estar involucradas en uno o varios conceptos de trabajo según sea el caso.

El acondicionamiento de la formación rocosa para su explotación incluye todos los gastos que requiera erogar el contratista para desmontar, despalmar, seccionar y abrir frente de atraque, listo para ser explotado.

Cuando el material pétreo para la formación de enrocamientos proceda de bancos de almacenamiento, los conceptos de trabajo incluirán la clasificación, carga, acarreo y su colocación de la obra.

Si durante el desarrollo de la obra, se presentasen para los enrocamientos aumentos en peso del calculado, que sobrepasen la tolerancia, como consecuencia de operaciones inadecuadas del contratista en la secuencia de construcción de estos. La Secretaría no cubrirá estos excedentes al contratista, salvo algún caso no previsto, fortuito, incontrolable a juicio de la Secretaría.

8.3.- Cambio de bancos.-

El Representante podrá ordenar al contratista, cuando el caso lo requiera, proceda a localizar nuevos bancos de roca para materiales de reserva. El contratista presentará al Representante los planos y programas que pretenda llevar a cabo para lograr el objetivo, los cuales serán aprobados por el Representante en su caso, antes de ser puestos en marcha. Los trabajos se desarrollarán en tal forma, que quede definido todo lo relativo a potencia y calidad probable de las formaciones rocosas por explotar. El pago de estos trabajos será cubierto por la Secretaría al contratista en los términos que indica el contrato.

La Secretaría podrá, si asi conveniese a sus intereses recomendar estos trabajos a terceras personas.

Si la pedrera y/o pedreras en explotación, llegasen a presentar condiciones inconvenientes, tales que fuese necesario suspender la extracción de los materiales pétreos, parcial y totalmente, el contratista deberá cambiar total o parcialmente su equipo a las -

nuevas formaciones rocosas de reserva, según se indica en el concepto 4.6 "Otros bancos", de estas especificaciones. El Contratista presentará nuevos precios unitarios para su estudio y aprobación en su caso sirviendo de base para calcularlos, las condiciones que se hicierón en su propuesta inicial.

El contratista adjuntará como parte integrante de su proposición, el precio de hora máquina de los siguientes equipos; grúas provistas de garra, palas, tractores, cargadores, compresores, camiones para transportar el material, chalanes, tolvas, grúas flotantes, remolcadores y en general todo el equipo pesado incluido para su oferta.

Si por motivo de algún cambio en el, proyecto que determine la Secretaría se generasen nuevos conceptos de obra, el contratista presentará su proposición para su estudio y aprobación en su caso de los precios unitarios nuevos, sirviendo de base las consideraciones de su oferta inicial.

9.- LIQUIDACIONES PARCIALES Y FINAL DE OBRA.-

9.1.- El cálculo de volúmen de obra, quedará establecido en el proyecto.

El volúmen de obra que se liquide al contratista mensualmente será el computo de los registros diarios del tonelaje de roca que esté coloque en el enrocamiento de acuerdo con el proyecto. Con fines informativos las estimaciones mensuales de obra deberán ir acompañadas de las secciones transversales que sirvierón de control en la construcción de las capas del enrocamiento en el tramo de que se trate. La distancia entre las secciones de control se podrá ser entre los dos y diez metros, según lo determine el Representante.

Si durante la realización del proyecto, la Secretaría, considerase conveniente modificar las estimaciones de los enrocamientos y por motivo de ello aumentar el tonelaje de roca el contratista se obliga a suministrar y colocar el material excedente en la obra, al Precio Unitario considerado en su proposición inicial..

Igualmente si por alguna razón fuera necesario ejecutar menos obra de la prevista, el contratista realizará la cantidad reducida de obra, no teniendo derecho por este motivo a pago o bonificación alguna.

9.2.- Deducciones.-

Cuando ordene el Representante el retiro de algún material indeseable de la obra y este hubiese sido ya pesado, su peso será deducido de la estimación mensual correspondiente.

9.3.- Derechos de la Secretaría.-

Queda el juicio de la Secretaría la admision y aprobación de secciones de enrocamiento, que queden fuera de las tolerancias, en cada caso resolverá finalmente si el contratista debe reparar alguna zona con sección escasa según el proyecto, lo cual será por cuenta del mismo.

9.4.- División final de la obra.-

Tan pronto como sea posible , despúes de terminar el trabajo será revisado detenidamente todo lo realizado, cuando se encuentre toda la obra en condiciones satisfactorias, será finalmente aceptada.

9.5.- Retiro de equipo.-

Al ser terminado y aceptado la totalidad del proyecto, el contratista deberá retirar todos sus equipos, incluyendo materiales sobrantes, dejando libres las áreas de trabajo, de cualquier objeto producto de sus operaciones que no tengan función alguna en el proyecto.

10.- SANCIONES.-

10.1.- Atrasos en la terminación de la Obra.-

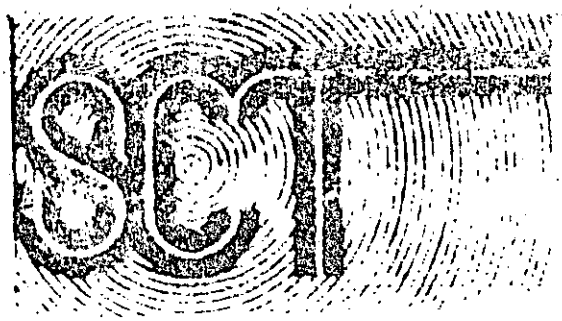
La falta de cumplimiento por parte del contratista para la conclusión de la obra, en el plazo fijado en el contrato ocasionará que se haga acreador a las sanciones que fija la cláusula correspondiente del mismo a menos que hubiese obtenido con anterioridad autorización expresa de la Secretaría de ampliación de plazo en la terminación de la obra por causas no imputables al contratista, como se indica en el párrafo 7.3 "Ampliación de Plazo",. La ampliación del plazo concedido al contratista en su caso, modificará el plazo de terminación de la obra en el lapso correspondiente.

11.- Personal del Contratista.-

11.1.- Diferentes tipos de personal.-

Los precios unitarios estipulados en el catálogo de conceptos de trabajo, deberán incluir las organizaciones por parte del contratista para sostener una planta de personal idóneo y eficiente que pueda llevar a cabo la realización del proyecto y cumplir satisfactoriamente a juicio de la Secretaría, con las siguientes actividades:

a).- Personal directivo.-Este personal estará capacitado para dirigir y manejar las actividades



SECRETARIA DE COMUNICACION
SUBSECRETARIA DE INFI
DIRECCION GENERAL DE (

CALENDARIO DE UTILIZAC

No.	ACTIVIDAD	CANTIDAD	UNIDAD	EQUIPO	RENDIMIENTO POR JORNADA DE TRABAJO

N O T A S:
En las columnas de los meses el contratista indicará el número de un

El contratista tendrá libertad para omitir o aumentar las actividade
a condición que a cada actividad le corresponda un procedimiento de

M.T.= MINIMO TEORICO
A.E.= ADICIONAL DE EMERGENCIA

PARA FORMULAR EL CALENDARIO. DEBERA ANOTARSE EL NUMERO DE UNIDADES DE LA MAQUINARIA O EQUIPO QUE SE UTILIZARA EN CADA UNO DE LOS MESES DURANTE LA EJECUCION DE LA OBRA

LUGAR Y

ES Y TRANSPORTES

RAESTRUCTURA

OBRAS MARITIMAS

CONCURSO No.

HOJA DE

OBRA

CION DE EQUIPO

DOCUMENTO CT-8a

No. DE UNIDADES			AÑO 19																					
			MESES																					
M.T.	A.E.	TOTAL																						

idades que empleará, pudiendo utilizar fracciones; no se permitirá usar barras.

s enunciadas, de acuerdo con su criterio y/o sus procedimientos constructivos,
construcción.

FECHA

NOMBRE Y FIRMA DEL CONTRATISTA

2.- Resumen de costos horarios de maquinaria.

	ACTIVO	INACTIVO
Camión de volteo de 6.00 m³.	267.58	137.52
Revolvedora de un saco	90.39	61.09
Vibrador de inmersión de 8 H.P.	66.64	52.88
Aplanadora Huber de 3 rodillos	276.22	163.44
Motoconformadora Cat. 120-B	727.85	374.83

ANALISIS DETALLADO PARA LA DETERMINACION DEL CARGO INDIRECTO.

Los cargos indirectos corresponden a los gastos generales necesarios para la ejecución de la obra, no incluidos en los cargos directos que realiza el contratista, tanto en sus oficinas centrales como en la obra, y que comprenden, entre otros, los gastos de organización, dirección técnica, vigilancia, supervisión, administración, financiamiento y prestaciones sociales correspondientes al personal directivo y administrativo y las regalías que procedan, en su caso, por el uso de patentes.

Los cargos indirectos se expresarán como un porcentaje del costo directo de cada concepto de trabajo. Dicho porcentaje se calculará sumando los importes de los gastos generales que resulten aplicables y dividiendo el resultado de esa suma entre el costo total directo de la obra de que se trate.

Para nuestro ejemplo a continuación se enlistan como ilustración, los gastos generales más frecuentes de posible aplicación, que deberán tomarse en consideración para integrar el cargo indirecto. Estos gastos son supuestos, por lo que el postor aplicará datos producto de su experiencia o investigación.

Por separado se anexa el formato para que los postores analicen detalladamente el cargo indirecto que aplicarán a los análisis de precios unitarios. Documento No. CT-9c.

	ADMON. CENTRAL	ADMON. DE OBRA
Honorarios, sueldos y prestaciones.		
1.1. Personal directivo	$ 225,000.00	XXX
1.2. Personal técnico	247,500.00	$ 226,2[illegible].00
1.3. Personal administrativo	117,000.00	[illegible]
1.4. Personal en tránsito	XXX	[illegible]
1.5. Cuota patronal del Seguro Social e impuesto adicional para remuneraciones pagadas para 1.1 a 1.4	100,981.00	[illegible]
1.6. Pasajes y viáticos	5,000.00	6,500.00
1.7. Consultores y Asesores	0.00	XXX
1.8. Estudios e Investigaciones	0.00	XX
Depreciación, mantenimiento y rentas.		
2.1. Edificios y locales	$ 38,500.00	$ 90,000.00
2.2. Campamentos	XXX	100,000.00
2.3. Talleres	XXX	120,000.00
2.4. Bodegas	XXX	300,000.00
2.5. Instalaciones generales	XXX	30,000.00
2.6. Muebles y enseres	23,750.00	[illegible].00
Servicios		
3.1. Depreciación o renta y operación de vehículos	67,500.00	135,000.00
3.2. Laboratorio de campo	XXX	12,000
Fletes y Acarreos		
4.1. De campamentos	XXX	[illegible].00
4.2. De equipo de construcción	XXX	
4.3. De plantas y elementos para instalaciones	XXX	
4.4. De mobiliario	XXX	
Gastos de oficina		
5.1. Papelería y útiles de escritorio	2,000.00	
5.2. Correos, Teléfonos, Telégramas, [illegible]	5,625.00	

5.3. Situación de fondos	XXX	0.00
5.4. Copias y duplicados	1,200.00	1,200.00
5.5. Luz, gas y otros consumos	4,000.00	3,500.00
5.6. Gastos de concurso	15,000.00	XXX
6.– Fianza y financiamientos.		
6.1. Primas por fianza	10,000.00	XXX
6.2. Intereses por financiamientos	631,944.00	XXX
7.– Trabajos previos y auxiliares.		
7.1. Construcción y conservación de caminos de acceso	XXX	15,000.00
7.2. Montajes y desmantelamientos de equipo, cuando así proceda.	XXX	0.00
SUMAS:	1'500,000.00	1'792,090.75

Administración Central:

$$\frac{1'500,000.00}{15'000,000.00} = 10.0\%$$

Administración de Obra:

$$\frac{1,792,090.75}{15'000,000.00} = 12.0\%$$

EXPRESION DE LOS CARGOS POR UTILIDAD Y ADICIONALES

A.– Cargos por utilidad.

Corresponde a la ganancia que percibe el contratista por la ejecución satisfactoria del concepto de trabajo, incluyendo el impuesto sobre la renta y quedará representada por un porcentaje sobre la suma de los cargos directos más los indirectos de dicho concepto de trabajo. En el Documento CT–9c., deberá consignarse el cargo por la utilidad correspondiente, en porcentaje.

Para este ejemplo: C. U. = 12.0%

B.– Cargos Adicionales.

Son los correspondientes a las erogaciones que realiza el contratista por estipularse expresamente en el contrato de obra como obligaciones adicionales, y que no están comprendidos dentro de los cargos directos, ni en los indirectos ni en la utilidad y se expresarán como porcentaje sobre la suma de los cargos directos más indirectos más utilidad. En el formato CT–9c deberán consignarse estos cargos.

Para este ejemplo: Derechos por inspección y vigilancia.

C. A. = 0.50%

R E S U M E N

Administración Central	10.0%
Administración de Obra	12.0%
A.– Cargo Indirecto	22.0%
B.– Cargo por Utilidad	12.0%
C.– Cargos Adicionales	0.5%

IV.– EJEMPLOS DE ANALISIS DETALLADOS DE PRECIOS UNITARIOS.

Ejemplo A:

Concepto Núm. 160.0404A010 concreto hidráulico sin incluir cimbra, de f'c = 200 kg/cm², elaborado con revolvedora y colocado manualmente, en zapatas de cimentación.

1.– Costo básico del concreto f'c 200 kg/cm².

Materiales.

Cemento	330	kg/m³	x $ 2.48/kg.		818.40
Arena	0.50	m³/m³	x 270.00/m³		135.00
Grava	1.00	m³/m³	x 270.00/m³		270.00
Agua	0.200	m³/m³	x 10.00		2.00
$\frac{\text{of. 2a.} + \text{S. P.}}{16 \text{ m}^3}$			$ 409.25 =	3.40	146.25

ACLARACIONES

D = Depreciación por hora efectiva de trabajo.

Va = Valor de adquisición de la máquina.

Vr = Valor de rescate de la máquina.

Ve = Vida económica de la máquina en horas.

Ha = Número de horas efectivas de trabajo de la máquina en un año.

I = Cargo por inversión equivalente a los intereses del capital invertido en la máquina.

i = Tasa de interés anual en vigor expresado como fracción.

S = Cargos por seguros necesarios para cubrir los riesgos a que está sujeta la maquinaria de construcción durante su vida económica por accidentes.

s = Prima anual promedio, valuada como porciento del valor de la máquina expresada como fracción.

Ka = Coeficiente en función de los costos de los locales para guardar la maquinaria, de los salarios del personal de vigilancia y del tiempo de guarda considerada.

T = Cargo por mantenimiento mayor y menor.— Comprende todas las erogaciones necesarias para conservar la maquinaria en buenas condiciones de trabajo. Incluye reposición de piezas, cambios de lubricantes y salarios devengados por el personal encargado del mismo, tanto en talleres especializados como en el campo.

Q = Coeficiente dado en base a experiencias estadísticas que incluye el mantenimiento mayor y menor.

E = Cargos por combustible que incluye los consumos de gasolina o diesel que utiliza la máquina.

c = Cantidad necesaria de combustible por hora efectiva de trabajo.

Pc = Precio unitario de combustible puesto en la máquina.

L = Lubricantes, consumos y cambios periódicos de aceite, puesto en la máquina.

a = Cantidad de aceite necesario por hora efectiva de trabajo de acuerdo con las condiciones medias de operación.

Pl = Precio unitario de los aceites puestos en la máquina.

Ll = Cargo por llantas, sólo para la maquinaria en la cual al calcular su depreciación se haya deducido el valor de las llantas del valor inicial de la misma.

Vll = Valor de adquisición de las llantas nuevas.

Hv = Vida económica de las llantas en horas.

O = Cargo por operación.— Pago de los salarios del personal encargado de la operación por hora efectiva

Sc = Salarios por turno del personal necesario para operar la máquina.

H = Horas efectivas de trabajo de la máquina dentro del turno.

10

DOCUMENTO

MODELO DE CONTRATO

SECRETARIA DE COMUNICACIONES
Y
TRANSPORTES

dencias y Entidades de la Administración Pública Federal; así como las normas para construcción e instalación, normas de calidad de materiales vigentes en "LA DEPENDENCIA" y las especificaciones de la obra, el proyecto, el programa de trabajo, los montos mensuales de obra derivados del mismo y el documento en que se consignan los precios unitarios y las cantidades de trabajo aproximadas, que, como anexos debidamente firmados por las partes, forman parte integrante de este contrato.

II.4.- Ha inspeccionado debidamente el sitio de la obra objeto de este contrato, a fin de considerar todos los factores que intervienen en su ejecución.

III.- Las partes declaran que:

III.1.- Se obligan en los términos de este contrato y del contenido de la sección correspondiente de las Reglas Generales para la Contratación y Ejecución de Obras Públicas y de Servicios relacionados con las mismas para las Dependencias y Entidades de la Administración Pública Federal, mismas que se tienen por reproducidas como parte integrante de este contrato.

C L A U S U L A S

PRIMERA.- Objeto del contrato.- "La Dependencia" encomienda a "El Contratista" y éste se obliga a realizar para ella hasta su total terminación, de conformidad con las normas de construcción e instalación, normas de calidad de materiales vigentes en "La Dependencia" y las especificaciones de la obra, el proyecto, el programa de trabajo, los montos mensuales de obra derivados del mismo y el documento en que se consignan los precios unitarios y las cantidades de trabajo aproximadas, que forman parte integrante de este contrato, una obra consistente en

ubicada en:

SECRETARIA DE COMUNICACIONES
Y TRANSPORTES

SEGUNDA.- Monto del Contrato.- El monto total del presente contrato es de: .

TERCERA.- Plazo de Ejecución.- "El Contratista" se obliga a iniciar las obras objeto de este contrato el día de de y a concluirlas a más tardar el día de de de conformidad con el programa.

CUARTA.- "La Dependencia" se obliga a tener oportunamente la disponibilidad legal y material de los lugares en que deberán ejecutarse las obras materia de este contrato.

QUINTA.- Sanciones por Incumplimiento del Programa.- "La Dependencia" tendrá la facultad de verificar si las obras objeto de este contrato se están ejecutando por "El Contratista" de acuerdo con el programa de obra aprobado, para lo cual "La Dependencia" comparará periódicamente el avance de las obras. Si como consecuencia de dicha comparación el avance de las obras es menor que lo que debió realizarse "La Dependencia" procederá a:

I.- Retener en total el 1% (uno por ciento) de la diferencia entre el importe de la obra realmente ejecutada y el importe de la que debió realizarse, multiplicado por el número de meses transcurridos desde la fecha programada para la iniciación de las obras, hasta la de la revisión. Por lo tanto, mensualmente se hará la retención o devolución que corresponda a fin de que la retención total sea la procedente.

Si al efectuarse la comparación correspondiente al último mes del programa, procede hacer alguna retención, su importe se aplicará en favor del Erario Federal, como pena convencional por el retraso en el cumplimiento de las obligaciones a cargo de "El Contratista".

SECRETARIA DE COMUNICACIONES
Y TRANSPORTES

II.- Aplicar, para el caso de que "El Contratista" no concluya la obra en la fecha señalada en el programa, una pena convencional consistente en una cantidad igual al 1% (uno por ciento) del importe de los trabajos que no se hayan realizado en la fecha de terminación señalada en el programa, que cubrirá "El Contratista", mensualmente y hasta el momento en que las obras queden concluídas y recibidas a satisfacción de "La Dependencia".

Para determinar la aplicación de las sanciones estipuladas, no se tomarán en cuenta las demoras motivadas por caso fortuito o fuerza mayor o por cualquier otra causa que a juicio de "La Dependencia", no sea imputable a "El Contratista".

Independientemente de la aplicación de las penas convencionales señaladas anteriormente, "La Dependencia" podrá optar entre exigir el cumplimiento del contrato o la rescisión del mismo.

En caso de que "La Dependencia" opte por la rescisión del contrato, en los términos de las Reglas Generales para la Contratación y Ejecución de Obras Públicas y de Servicios relacionados con las mismas para las Dependencias y Entidades de la Administración Pública Federal, aplicará a "El Contratista" una sanción consistente en un porcentaje del valor del contrato, que podrá ser, a juicio de "La Dependencia", hasta por el monto de las garantías otorgadas.

Las cantidades que resulten de la aplicación de las penas convencionales que se impongan a "El Contratista" se harán efectivas con cargo a las cantidades que le hayan sido retenidas en los términos de dichas Reglas, aplicando, si ha lugar a ello, la Fianza otorgada conforme a lo estipulado en las propias Reglas.

SEXTA.- Otras obligaciones de "El Contratista".- "El Contratista acepta que de las estimaciones que se le cubran, se deduzca el 1% (uno por ciento) del monto de los trabajos contratados para el desarrollo de los programas que "La Dependencia" tenga establecidos en materia de obras y servicios de Beneficio social.

SECRETARIA DE COMUNICACIONES
Y
TRANSPORTES

SEPTIMA.- Forma de Pago.- "El Contratista" recibirá de "La Dependencia" como pago total por la ejecución satisfactoria de los trabajos ordinarios y extraordinarios objeto del contrato, el importe que resulte de aplicar los precios unitarios a las cantidades de trabajo realizadas.

"La Dependencia" expedirá el documento que corresponde a cargo de la Tesorería de la Federación, por el monto de la asignación anual del contrato.

Las estimaciones de los trabajos ejecutados deberá elaborarse y datarse los días últimos de cada mes y en caso de que éste sea festivo, en el primer día hábil siguiente y se pagarán una vez satisfechos los requisitos establecidos para su trámite, en un plazo que no excederá de 30 (treinta) días hábiles, contados a partir de la fecha de su aprobación en la obra. La falta de pago dentro de dicho plazo dará lugar a cubrir los gastos financieros que hubiere causado su demora, en los términos del artículo 44 del Reglamento de la Ley de Obras Públicas.

En las estimaciones correspondientes, se abonará a "El Contratista", en los trabajos ordinarios y extraordinarios a base de precios unitarios, el importe que resulte de aplicar dichos precios a las cantidades de trabajo realizadas.

Las estimaciones y la liquidación aunque hayan sido pagadas no se considerarán como aceptación de los trabajos, ya que "La Dependencia", se reserva expresamente el derecho de reclamar por trabajos faltantes o mal ejecutados o por pago de lo indebido.

Si "El Contratista" estuviere inconforme con las estimaciones o la liquidación, tendrá un plazo de 30 (treinta) días calendario, a partir de la fecha en que se haya formulado la estimación o la liquidación para hacer por escrito la reclamación. Si transcurrido este plazo "El Contratista" no la efectúa, se considerará que la estimación o liquidación quedará definitivamente aceptada por él y sin derecho a ulterior reclamación.

OCTAVA.- Revisión de los Costos.- "El Contratista" cuando durante la vigencia del contrato ocurran circunstancias de orden económico no previstas, pero que de hecho y sin dolo, culpa o negligencia o ineptitud de cualquiera de las partes, determinen un aumento o reducción en un 5% (cinco por ciento) o más de los costos de los trabajos aún no ejecutados, dichos costos podrán ser revisados en los términos del artículo 46 de la Ley de Obras Públicas y 27 fracción V, 50 fracción _____ y 51 de su Reglamento.

SECRETARIA DE COMUNICACIONES Y TRANSPORTES

NOVENA.- Anticipos.

Para la iniciación de los trabajos "La Dependencia" otorgará a "El Contratista" un anticipo de $

equivalente al 10% de la asignación anual autorizada y de - - $

para la compra de equipo y materiales de instalación permanente equivalente al 20% de la citada asignación.

Para los ulteriores ejecicios "La Dependencia" otorgará los anticipos que procedan en los términos del artículo 27 del Reglamento de la Ley de Obras Públicas.

El o los anticipos que se otorguen a "El Contratista" deberán ser amortizados por éste totalmente en el ejercicio para el cual fueron otorgados, mediante deducciones proporcionales en cada una de las estimaciones que por trabajos efectuados se formulen a "El Contratista" o en su caso con la liquidación.

Si al término del ejercicio no se ha amortizado el o los anticipos, el o los saldos correspondientes deberán ser devueltos por "El Contratista" a la Tesorería de la Federación.

DECIMA.- Fianzas.

Para garantizar la correcta inversión del o los anticipos "El Contratista" dentro de los 15 (quince) días hábiles siguientes a la fecha de la firma del contrato o revalidaciones del mismo, deberán presentar a "La Dependencia" la fianza o fianzas - - - por el importe total del o los anticipos otorgada u otorgadas por institución mexicana debidamente autorizada a favor de la Tesorería de la Federación.

Contra la entrega de la o de las fianzas por parte de "El Contratista" "La Dependencia" iniciará de inmediato el trámite para la entrega del o los anticipos.

La fianza o fianzas otorgadas para garantizar la correcta inversión del o los anticipos se cancelarán cuando "El Contratista" haya amortizado el importe total del mismo. En caso de que "El Contratista" no haya devuelto el saldo del o los anticipos dentro de los 10 (diez) días posteriores al cierre del ejercicio presupuestal, se harán efectivas la o las fianzas.

SECRETARIA DE COMUNICACIONES
Y
TRANSPORTES

Para garantizar el cumplimiento de las obligaciones derivadas de este contrato, "El Contratista" se obliga a exhibir dentro de los 15 (quince) días siguientes, contados a partir de la fecha en que suscriba este contrato, una póliza de fianza por valor del 10% (diez por ciento) del importe del contrato o en su caso de la asignación anual, otorgada por institución mexicana debidamente autorizada a favor de la Tesorería de la Federación y en los términos del artículo 26 del Reglamento de la Ley de Obras Públicas y de las Reglas Generales para la Contratación y Ejecución de Obras Públicas y de Servicios relacionados con las mismas para las Dependencias y Entidades de la Administración Pública Federal.

La póliza en que sea expedida la fianza, deberá contener las siguientes declaraciones expresas de la institución que la otorgue:

a).- Que la fianza se otorga atendiendo a todas las estipulaciones contenidas en el contrato.

b).- Que en el caso de que se prorrogue el plazo establecido para la terminación de los trabajos a que se refiere la fianza o exista espera su vigencia quedará automáticamente prorrogada en concordancia con dicha prorroga o espera.

c).- Que la fianza garantiza la ejecución total de los trabajos materia del contrato, aún cuando parte de ellos se subcontraten con autorización de "La Dependencia".

d).- Que para cancelar la fianza será requisito indispensable la solicitud expresa y por escrito de la Dependencia contratante a la Tesorería de la Federación que procederá de inmediato en dicho sentido al recibir la petición respectiva de "La Dependencia".

e).- Que la institución afianzadora acepta expresamente lo preceptuado en los artículos 95 y 118 de la Ley Federal de Instituciones de Fianzas en vigor.

La fianza se cancelará cuando "El Contratista" haya cumplido con todas las obligaciones que se deriven del contrato.

DECIMA PRIMERA.- Otras Modalidades Específicas.

SECRETARIA DE COMUNICACIONES
Y
TRANSPORTES

DECLARACIONES FINALES.- NACIONALIDAD, PERSONALIDAD Y DOMICILIO.

CRETARIA DE COMUNICACIONES
Y
TRANSPORTES

"El Contratista" señala como domicilio para los fines del presente Contrato, la casa No.

y la Dependencia en

El presente contrato se firma en la Ciudad de México, D. F., a los _____ _____.

EL SUBSECRETARIO DE

EL CONTRATISTA

R E V I S A D O

EL DIRECTOR GENERAL DE

EL DIRECTOR GENERAL DE ASUNTOS JURIDICOS

11

DOCUMENTO

CATALOGO DE CONCEPTOS Y CANTIDADES DE OBRA
PARA PROPOSICION DE PRECIOS UNITARIOS

SUBSECRETARIA DE INFRAESTRUCTURA

DIRECCION GENERAL DE OBRAS MARITIMAS

OBRA :

CATALOGO DE CONCEPTOS Y CANTIDADES DE OBRA PARA PROPOSICION DE PRECIOS UNITARIOS Y MONTO TOTAL DE LA PROPOSICION

FORMA CT-11 | HOJA 1 DE 3

CONCURSO N°

LUGAR Y FECHA

CONTRATO N°

No.	CONCEPTO DESCRIPCION Y ESPECIFICACIONES		CANTIDAD	PRECIO UNITARIO LETRA	PRECIO UNITARIO NUMERO	IMPORTE
1.-	Suministro y colocación de piedra natural para núcleo de escollera , con material producto de la explotación del banco de préstamo "C. ", con un peso comprendido entre 10 y 100 kg. Incluye: desmonte, despalme, explotación, selección, acopio, carga a camión, acarreo al primer kilómetro, desperdicios, señalamiento carretero, descopete de la barra, pesado de piedra, mantenimiento del camino del banco a la obra y demás cargos correspondientes. (N C. s/n Rompeolas y Especificaciones Particulares).	ton	22,000.00			
2.-	Suministro y colocación de piedra natural para capa secundaria en escollera , con material producto de la explotación del banco de préstamo al C. , con un peso comprendido entre 450 y 1500 kg. Incluye: desmonte, despalme, explotación, selección, acopio, carga a camión, acarreo al primer kilómetro, desperdicios, pesado de piedra, señalamiento carretera, mantenimiento del camino del banco a la obra y demás cargos correspondientes. (N.C. s/n Rompeolas y Especificaciones Particulares).	ton	5,554.00			
3.-	Suministro y colocación de piedra natural para capa secundaria en escollera , con material producto de la explotación del banco de préstamo "C. ", con un peso comprendido entre 100 y 450 kg. Incluye: desmonte, despalme, explotación, selección, acopio, carga a camión, acarreo al primer kilómetro, pesado de piedra, desperdicios, señalamiento carretero, mantenimiento del camino del banco a la obra y demás cargos correspondientes. (C. s/n Rompeolas y Especificaciones Particulares).	ton	10,546.00			
4.-	Suministro y colocación de piedra natural para coraza de escollera , producto de la explotación del banco de préstamo " ", con un peso comprendido entre 1500 y 4000 kg. Incluye: desmonte, despalme, explotación, selección, acopio, car-					

EL DIRECTOR GENERAL DE

ING

NOMBRE Y FIRMA DEL POSTOR

107

SUBSECRETARIA DE INFRAESTRUCTURA
DIRECCION GENERAL DE OBRAS MARITIMAS

OBRA:

CATALOGO DE CONCEPTOS Y CANTIDADES DE OBRA PARA PROPOSICION DE PRECIOS UNITARIOS Y MONTO TOTAL DE LA PROPOSICION

FORMA CT-11 | HOJA 2 DE 3

CONCURSO N
LUGAR Y FECHA
CONTRATO N°

No.	CONCEPTO DESCRIPCION Y ESPECIFICACIONES		CANTIDAD	PRECIO UNITARIO LETRA	PRECIO UNITARIO NUMERO	IMPORTE
	ga a camión, acarreo al primer kilómetro, desperdicios, pesado de piedra, señalamiento carretero, mantenimiento del camino del banco a la obra y demás cargos correspondientes. (N.C. VI-2, 3 - E.C. s/n, Rompeolas y Esp. Part.)	ton	19,000.00			
5	Suministro y colocación de piedra natural para coraza de morro - y escollera , con material producto de la explotación del ban - co de préstamo "C ", con un peso comprendido - entre 4,000 y 7,000 kg, incluye: desmonte, despalme, explotación selección, acopio, carga a camión, acarreo al primer kilómetro pesado de piedra, desperdicios, señalamiento carretero, mante - nimiento del camino del banco a la obra y demás cargos corres - pondientes. (N.C. VI-2, 3 E.C. s/n, Rompeolas y Esp. Part.)	ton	15,028.00			
6	Acarreo terrestre en kilómetros subsecuentes al primero de pie - dra natural para núcleo, capa secundaria, coraza de cuerpo y mo - rro de escollera del banco de préstamo a la obra. (Esp. Part.)	ton km	3,028,788.00			
	IMPORTE PARCIAL.....					
7	BASCULA Transporte e instalación de una báscula con capacidad pesadora - de 50 ton. de un eje. Incluye: transporte de la báscula desde T , hasta el sitio de la obra, fabricación de fosa y - caseta para la báscula de acuerdo con las características de ésta, verificación y Vo. Bo. de la Dirección General de Normas. (Esp. Part.)	lote	1			
	IMPORTE PARCIAL					

EL DIRECTOR GENERAL DE

ING

NOMBRE Y FIRMA DEL POSTOR

105

SUBSECRETARIA DE INFRAESTRUCTURA DIRECCION GENERAL DE OBRAS MARITIMAS	CATALOGO DE CONCEPTOS Y CANTIDADES DE OBRA PARA PROPOSICION DE PRECIOS UNITARIOS Y MONTO TOTAL DE LA PROPOSICION	FORMA CT-II	HOJA 3 DE 3
OBRA: C... DA-		CONCURSO N° LUGAR Y FECHA CONTRATO N°	

No.	CONCEPTO DESCRIPCION Y ESPECIFICACIONES		CANTIDAD	PRECIO UNITARIO LETRA	PRECIO UNITARIO NUMERO	IMPORTE
8	CAMPAMENTO. Suministro, transportación y montaje de campamento tipo, fabricado a base de paredes multy-panel H y M en módulos de 86 cm con espesor de 1 1/2", techumbre multy-panel RL-80 con refuerzos y apoyos de acuerdo a normas del fabricante, incluye: firme de concreto de 10 cm. de espesor acabado afinado con color, piso de azulejo, muro de tabicón o block recubierto con azulejo, castillos y dalas de concreto armado, carpintería, ventanería de aluminio con vidrio, pintura, instalaciones eléctrica, hidraúlica, sanitaria y de gás de acuerdo al plano No. D.P. 0333, dos unidades de ventana, climatizadoras de aire de 12,200 btu/HR marca York o similar en calidad, fosa séptica con capacidad para 10 personas, pozo de absorción, cisterna para 5,000 litros, bomba de 1-2 H.P., tinaco con su base, tanques de gas y el suministro de gas y agua durante el tiempo que dure la obra, todo como lo indica el plano No. D.P. 84-174-A. (Esp. Part.)	lote	1			
	IMPORTE PARCIAL.....					
	IMPORTE DE LA PROPOSICION.....					
	I. V. A.					
	IMPORTE TOTAL DE LA PROPOSICION.....					

		NOMBRE Y FIRMA DEL POSTOR	

13

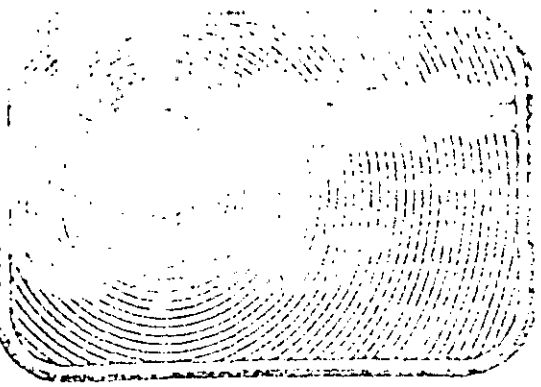
DOCUMENTO

CIRCULAR (ES) ACLARATORIA (S) EN SU CASO

15

DOCUMENTO

PROCEDIMIENTO DE CONSTRUCCION DE



SUBSECRETARIA DE
DIRECCION GENERAL DE

DOCUMENTO C T- 15
PROCEDIMIENTOS DE CONSTRUCCION QUE EL POSTOR DEBERA PRESENTAR

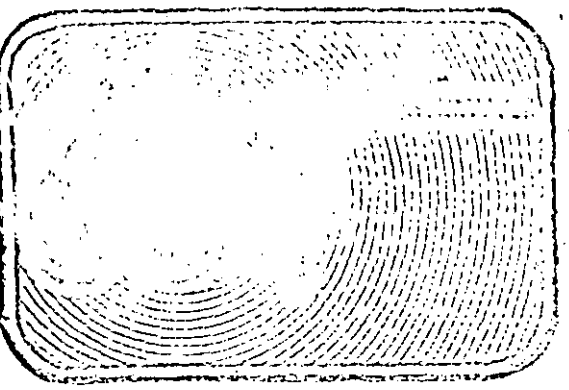
CONCURSO Nº
RUBRO: "C

EL POSTOR DEBERA PRESENTAR EN ORIGINAL Y COPIA LOS SIGUIENTES PROCEDIMIENTOS DE CONSTRUCCION:

A.- EXPLOTACION, CLASIFICACION Y ACOPIO DE MATERIAL ROCOSO. (Llenar anexo "A").

B.- COLOCACION DE PIEDRA EN NUCLEO, CAPA SECUNDARIA, CORAZA Y MORRO. (Llenar anexo "B").

NOMBRE DE LA EMPRESA:	



SUBSECRETARIA DE [illegible]
DIRECCION GENERAL DE [illegible]

ANEXO CT-15 "A"

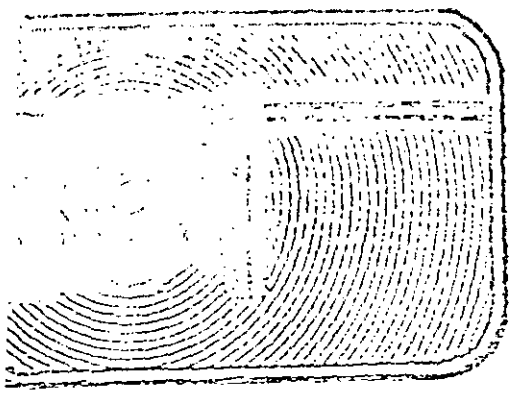
RELACION DE EQUIPO EMPLEADO EN EXPLOTACION, CLASIFICACION Y ACOPIO DE MATERIAL ROCOSO.

CONCURSO Nº
RUBRO:

Nº	NOMBRE DE LA MAQUINA O EQUIPO	M A R C A	M O D E L O	CAPACIDAD NOMINAL.

EL POSTOR DEBERA PRESENTAR EN ORIGINAL Y COPIA ESTE DOCUMENTO.

NOMBRE DE LA EMPRESA:	



SUBSECRETARIA DE
DIRECCION GENERAL DE

ANEXO CT-15 "B"

RELACION DE EQUIPO EMPLEADO EN COLOCACION DE PIEDRA EN NUCLEO, CAPA SECUNDARIA, CORAZA Y MORRO.

CONCURSO Nº
RUBRO: "C

Nº	NOMBRE DE LA MAQUINA O EQUIPO	M A R C A	M O D E L O	CAPACIDAD NOMINAL.

EL POSTOR DEBERA PRESENTAR EN ORIGINAL Y COPIA ESTE DOCUMENTO.

NOMBRE DE LA EMPRESA:

ESPECIALIDAD EN OBRAS MARITIMAS

MATERIA : "PROGRAMACION Y CONTROL DE OBRAS"
1o. DE JULIO A SEPTIEMBRE DE 1985
MEXICO, D.F

PROGRAMACION

ING. PORFIRIO LEON RUIZ
SEPTIEMBRE DE 1985

PROGRAMACION

La máxima "gobernar es prever" da una idea de la importancia que se atribuye a la previsión en el mundo de los negocios, y es cierto que si bien la previsión no es todo el gobierno, es al menos una parte esencial de él. Prever, aquí, significa a la vez calcular el porvenir y prepararlo; prever, es ya obrar.

La previsión tiene una infinidad de ocasiones y de maneras de manifestarse; su principal manifestación, su signo sensible, su instrumento más eficaz, es el programa de acción.

El programa de acción es a la vez el resultado a que se tiende, la línea de conducta que se ha de seguir, las etapas que se han de franquear, los medios que se han de emplear; es una especie de cuadro del futuro en el que los acontecimientos próximos están previstos con cierta claridad, según la idea que hemos hecho de ellos, y en el que los acontecimientos remotos aparecen cada vez más vagos; es la marcha de la empresa prevista y preparada para determinado tiempo.

El programa de acción se basa; 1° en los recursos de la empresa (inmuebles, herramientas, materias primas, capitales, personal, capacidad de construcción, mercados, relaciones públicas, etc., etc.); 2° en la naturaleza y en la importacia de las operaciones en curso y 3° en las posibilidades futuras, posibilidades que dependen,en parte de las condiciones técnicas, comerciales, financieras y otras, sometidas todas a cambios cuya importancia, ni el momento en que ocurrirán pueden determinarse de antemano.

La preparación del programa de acción es una de las operaciones más importantes y más difíciles de toda empresa; pone en juego todos los servicios y todas las funciones, y particularmente la función administrativa de la obra.

Es, en efecto, para desempeñar su función de administrador, por lo que el ingeniero toma la iniciativa del programa de acción, indica su objeto y su amplitud, fija la importancia de cada servicio en la obra común, coordina sus partes, armoniza el conjunto, y decide, finalmente, la línea de conducta que se ha de seguir. En esta línea de conducta es preciso no sólo que no haya nada que choque con los principios y las reglas de una buena administración, sino además que las disposiciones adoptadas faciliten la aplicación de esos principios y de esas reglas.

A más de las diversas capacidades técnicas, comerciales, financieras y otras, necesarias al jefe y/o ingeniero de la obra y a sus colaboradores, para establecer el programa de acción, debe poseerse una seria capacidad administrativa.

Características generales de un buen programa de acción.

Nadie discute la utilidad del programa de acción; es necesario que antes de actuar se sepa lo que se puede y lo que se quiere. Sabido es que la ausencia de programa va acompañada de titubeos, de falsas maniobras, de cambios de orientación intempestivos, que constituyen otras tantas causas de debilitamiento, cuando no de ruina para los negocios, La cuestión de la necesidad del programa de acción no se plantea siquiera, por lo tanto, y creo expresar la opinión general diciendo que el programa de acción es indispensable.

Pero hay varias clases de programas: los hay simples, complejos, sucintos, detallados, de larga o de corta duración...; los hay que han sido estudiados con una atención minuciosa, y otros tratados ligeramente; los hay buenos, medianos y malos.

¿Cómo distinguir los buenos de los otros?
Sobre el valor real de un programa, es decir sobre los servicios que puede rendir a la empresa, únicamente la experiencia puede pronunciarse de manera soberana. Y aún así hay que tener en cuenta el modo como se aplica. Hay el instrumento y el artista.

Existen, sin embargo, algunas características generales sobre las cuales es posible entenderse previamente sin esperar a que la experiencia se haya pronunciado.

La unidad de programa, por ejemplo. No puede aplicarse más que un programa a la vez: dos programas diferentes provocarían la dualidad, la confusión, el desorden.

Pero un programa puede dividirse en varias partes. En la gran empresa se encuentran, junto con el programa general, un programa técnico, un programa comercial, un programa financiero, etc., o también un programa de conjunto con un programa particular para cada servicio. Pero todos estos programas están encadenados, soldados, formando uno solo, y toda modificación introducida en uno cualquiera de ellos se traduce al punto en el programa de conjunto.

La acción directiva del programa debe ser continua. Ahora bien, los límites de la perspicacia humana limitan también forzozamente la duración de los programas. Para que no haya interrupción en la acción directiva, es preciso que un segundo programa suceda al primero sin intervalo, un tercero al segundo y así sucesivamente.

En una obra compleja, el programa anual es de un uso casi general. Otros programas, de más larga o de más corta duración, siempre estrechamente armonizados con el programa anaul, funcionan simultáneamente con éste último.

El programa debe ser lo bastante flexible para plegarse a las modificaciones que se juzgue conveniente introducir en él, ya sea bajo la presión de los acontecimientos, o por cualquiera otra razón. Antes y después, es la ley ante la cual hay que inclinarse.

Otra cualidad del programa es la de tener toda la precisión compatible con la incógnita que pesa sobre los destinos de la empresa.

Habitualmente, es posible trazar la línea de conducta próxima con un grado de precisión bastante grande.

Una simple directiva conviene a las operaciones lejanas; antes que haya llegado el momento de ejecutarlas se habrán adquirido luces que permitirán fijar mejor la línea de conducta. Cuando la parte que corresponde a la incógnita es relativamente muy grande, el programa no puede tener ninguna precisión; la obra toma entonces el nombre de aventura. Unidad, continuidad, flexibilidad y precisión, tales son los caracteres generales de un buen programa de acción.

En cuanto a las demás cualidades particulares que debe poseer y que dependen de la naturaleza, de la importancia y de las condiciones de la obra para la cual se hace, sólo es posible fijarlos de antemano por comparación con otros programas que han sido reconocidos como buenos en obras analogas.

Los buenos programas no faltan; se adivinan por la marcha de la obra, pero no se les ve con el suficiente detalle para conocerlos y juzgarlos bien. Sería, sin embargo, muy útil a todos aquellos que deben administrar, saber cómo redactan sus programas los jefes experimentados. Bastaría una docena de ejemplos bien elegidos.

ING. PORFIRIO LEON RUIZ

CONTROL

En una obra, el control consiste en comprobar si todo ocurre conforme al programa adoptado, a las órdenes dadas y a los principios admitidos.

Tiene por objeto señalar las faltas y los errores a fin de que se pueda reparar y evitar su repetición.

Se aplica a todo, a las cosas, a las personas, a los actos.

Desde el punto de vista técnico, hay que observar la marcha de las operaciones, sus resultados, sus desigualdades, el estado de conservación, el funcionamiento del personal y de las máquinas, etc.

Desde el punto de vista administrativo, hay que asegurarse de que el programa existe, de que se aplica y se mantiene al día, de que el organismo social está completo, de que los cuadros sinópticos del personal se emplean, de que la dirección se ejerce de acuerdo con los principios, de que las conferencias de coordinación se celebran, etc.

Desde el punto de vista financiero, el control se ejerce sobre los libros y la caja, sobre los recursos y las necesidades, sobre el empleo de fondos, etc.

Desde el punto de vista de la seguridad, hay que asegurarse de que los medios adoptados para proteger los bienes y las personas se encuentran en buen estado de funcionamiento.

Finalmente, desde el punto de vista de la contabilidad, hay que comprobar que los documentos necesarios llegan rápidamente, que permiten obtener una visión clara de la situación de la empresa, que el control encuentra en los libros, en las estadísticas y en los diagramas buenos elementos de verificación y que no existe ningún documento o estadística inútil.

Todas estas operaciones corresponden a la vigilancia, en la medida en que puedan ser ejecutadas por el ingeniero de la obra y sus colaboradores jerarquizados.

Pero cuando ciertas operaciones de control se hacen demasiado numerosas, o demasiado complejas, o demasiado extensas que puedan ser realizadas por los agentes ordinarios de los distintos servicios, es preciso recurrir a agentes especiales, que toman el nombre de supervisores.

Para que el control sea eficaz, es preciso que se haga en tiempo oportuno y vaya seguido de sanciones.

Es evidente que si las conclusiones de un control, incluso muy realizado, llegan demasiado tarde para que sea posible utilizarlas, el control habrá sido una operación inútil.

No es menos evidente que el control será inútil si las conclusioes prácticas que de él derivan son voluntariamente desatendidas.

Estas dos faltas son de las que una buena administración no deja cometer.

Otro peligro que hay que evitar es la ingerencia del control en la dirección y la ejecución de los servicios.

Esta invasión de atribuciones constituye la dualidad de dirección bajo su aspecto más temible: de una parte, el control irresponsable, y sin embargo provisto del poder de perjudicar a veces en amplia medida; del otro, el servicios ejecutivo que no dispone más que de débiles medios de defensa contra un control malévolo. La tendencia del control a la invasión de campos ajenos es bastante frecuente sobre todo en -- los negocios muy grandes, y puede tener las más graves consecuencias. Para combatirla es preciso ante todo definir de una manera lo más precisa posible las atribuciones del control, indicando bien los límites

que no debe franquear; es preciso después que la autoridad superior vigile el uso que el control hace de sus poderes.

Conociendo el objeto y las condiciones en que debe ejercerse el control, podemos deducir que el buen supervisor debe ser competente e imparcial.

La competencia del supervisor no necesita demostración.

Para juzgar acerca de la calidad de un objeto, del valor de un procedimiento de fabricación, de la claridad de los escritos, de los medios de mando empleados, es preciso evidentemente, en cada caso, estar provisto de la competencia adecuada.

La imparcialidad reposa sobre una conciencia recta y una completa independencia de interventor con respecto al intervenido. El control es sospechoso cuando el interventor depende en un grado cualquiera del intervenido. e incluso únicamente cuando existen entre ambos relaciones demasiado estrechas de interés, de parentesco o de camaradería.

Tales son las principales condiciones que el verificador debe llenar: la competencia, el sentimiento del deber, la independencia con respecto al intervenido, el juicio y el tacto.

Bien realizado, el control es un valioso auxiliar de la dirección, puede darle ciertas informaciones necesarias que la vigilancia jerarquizada sería a veces incapaz de suministrar.
Puede ejercerse sobre todo; depende de la dirección que su funcionamiento sea eficaz. Un buen control previene contra sorpresas enojosas que podrían degenerar en catástrofes.

Conviene poder responder siempre, a propósito de cualquier operación, a esta pregunta: "¿ Cómo se efectúa el control?"

Como se aplica a las operaciones de todo género y a los agentes de todos los niveles, el control se ejerce de mil maneras diferentes.

ING. PORFIRIO LEON RUIZ