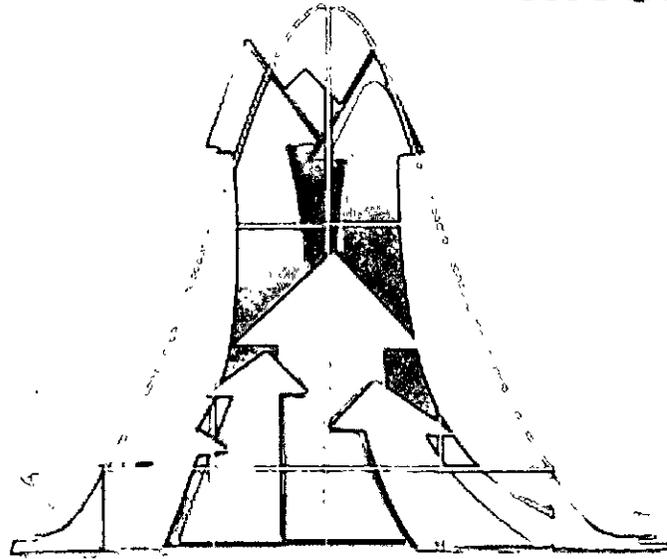
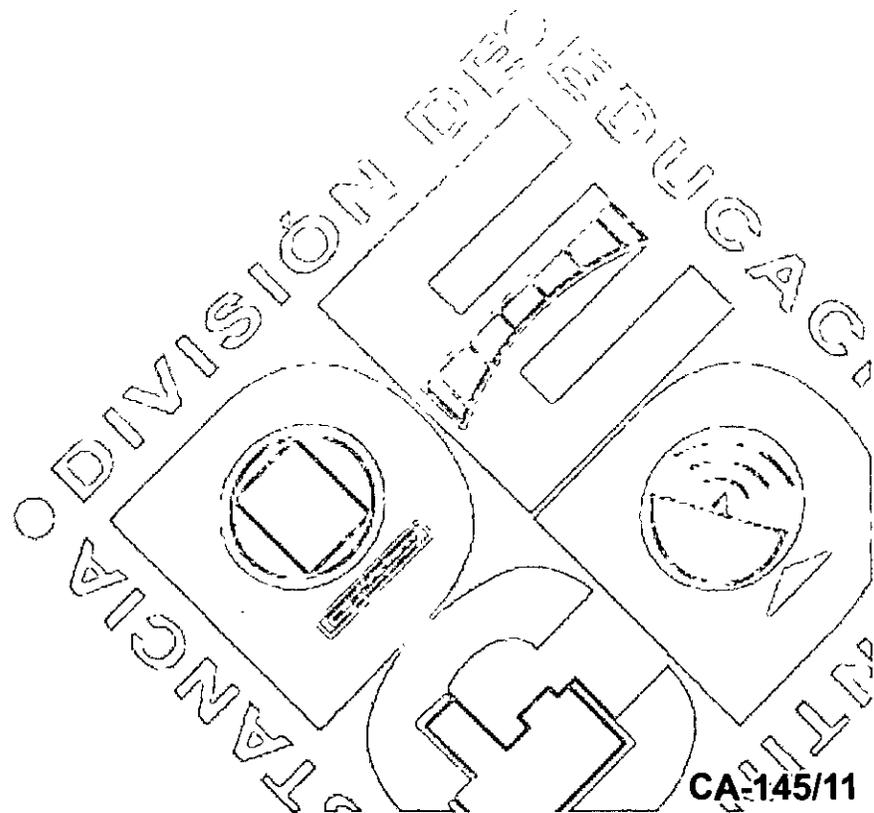


Diplomado:
Seis Sigma
nivel-Green Belt



Módulo IV: Fase - Analizar

Ing. Juan Flores Zamudio





MÓDULO IV

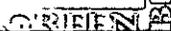
"ANALIZAR"




Fase de Análisis

OBJETIVOS:

- COMPRENDER EL PROPÓSITO DE LA FASE DE ANÁLISIS Y LA SALIDA ESPERADA EN ESTA ETAPA,
- INTERPRETAR Y APLICAR LOS SIGUIENTES CONCEPTOS:
 - Multivari-Chart
 - Five Whys (5 Porqués)
 - Diagrama de Causa y Efecto
 - AMEF (Análisis de Modo y Efecto de Falla)
 - Matriz de Causa y Efecto
 - Pruebas de Hipótesis

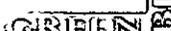



Fase de Análisis

DEFINIR
 MEDIR
 ANALIZAR
 MEJORAR
 CONTROL

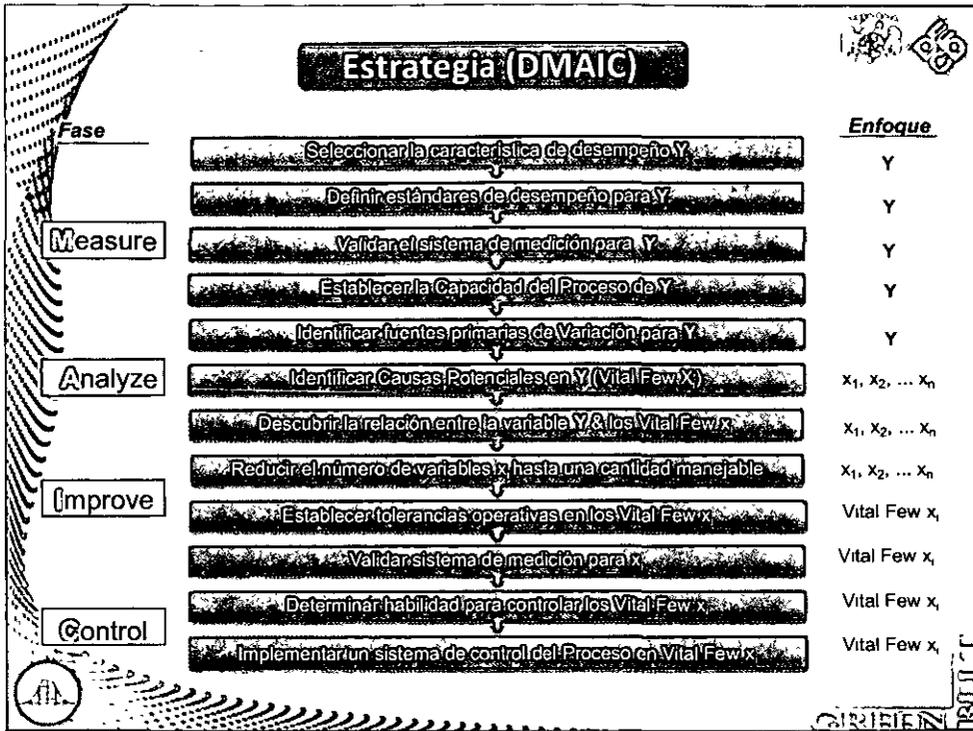
Objetivos de la fase:

- Desarrollar teorías de causas generadoras del problema
- Confirmar teorías de causas con datos
- Identificar la(s) causa(s) raíz del problema



MÓDULO IV



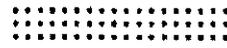
"ANALIZAR"

Identificando causas

Una vez que el problema ha sido ubicado, el siguiente paso es listar las causas potenciales que lo generan y organizarlas para visualizar relaciones entre causa y efecto, a fin de decidir sobre cuales trabajar.

La herramientas más comunes usadas para determinar esta lista son:

- Lluvia de Ideas
- Diagramas de Afinidad
- Mapa de Proceso
- Multivari-chart
- 5 Porqués
- Diagramas de Causa y Efecto
- Análisis de Modo y Efectos de Falla



Multi-Vari Chart

Multi-Vari es una herramienta gráfica utilizada para identificar dónde se genera la variación en un proceso/sistema.

En esta gráfica la variable de estudio es representada visualmente, estratificada y analizada.

Es una herramienta muy versátil ya que puede utilizarse tanto en la fase de Medición como en la fase de Análisis ya que permite también evaluar los efectos de diferentes causas "Xs" o combinación de ellas en la variable de respuesta "Y".



MÓDULO IV

"ANALIZAR"

Multi-Vari Chart

En un Análisis Multi-Vari se requiere que:

- Los estratos o grupos categóricos tengan siempre niveles de atributos asignados que son arbitrarios en naturaleza (Operador A, B, C; Nivel de satisfacción Bajo o Alto; Máquina 1, 2, 3; etc.).

BENEFICIOS

- El resultado gráfico es fácilmente entendido
- Ayuda para representar varios estratos a la vez
- Ayuda para identificar condiciones "fuera de control"
- Ayuda a analizar gráficamente datos pasivos o históricos





Cómo crear un Multi-Vari Chart

Paso 1:

Decide qué tipo de datos deberán ser recolectados para ubicar o definir con más profundidad el problema, por ejemplo: número de defectos, índice de rendimiento, tiempo de entrega, etc.

Paso 2:

Decide los estratos o grupos. Considera cómo estratificar los datos, por ejemplo: por maquinaria, por tiempo, por tipo de defecto, por operador, etc. Asegura siempre considerar estratos que reflejen variaciones dentro de grupo/pieza, entre grupos/piezas y por tiempo.

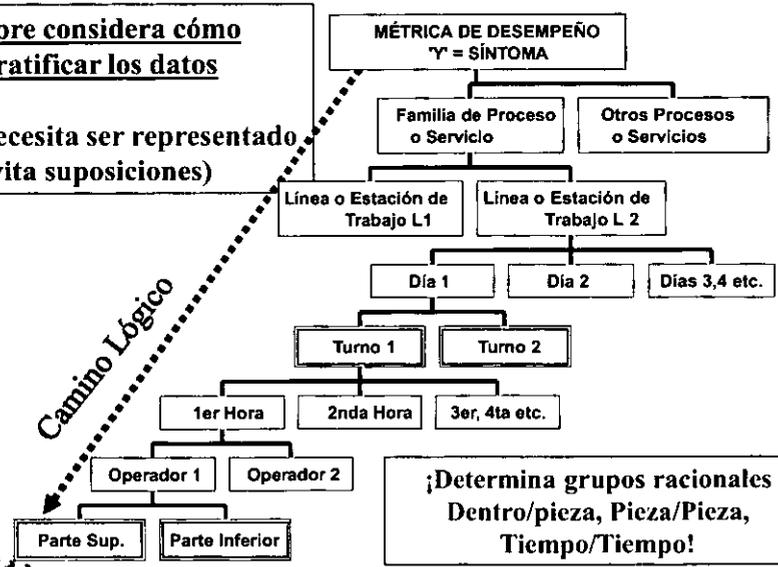
MÓDULO IV

"ANALIZAR"

Representación lógica para estratificar

Siempre considera cómo Estratificar los datos
Tu plan necesita ser representado (Evita suposiciones)

Camino Lógico



¡Determina grupos racionales Dentro/pieza, Pieza/Pieza, Tiempo/Tiempo!



Cómo crear un MULTI-VARI CHART

Paso 3:
Recolecta los datos de tal forma que las entradas para la estratificación sean reconocidas.

Recuerda asignar niveles por atributos a cada estrato.

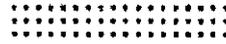
Paso 4:
Prepara el gráfico de Multi-Vari y analiza los datos. Si la estratificación resulta favorable esto es, muestra diferencias entre los estratos, será un buen indicador de dónde se localiza la mayor problemática para después medir esta zona y estudiarla con mayor profundidad en la Fase de "Análisis".

Ejemplo MULTI-VARI CHART

Una academia de Karate ha recibido diferentes reclamos de sus alumnos en cuanto a la calidad de las cintas otorgadas durante el campeonato local. La academia preocupada por esta situación decidió conducir un estudio Multi-Vari para identificar la mayor fuente de variación generadora de niveles de defectivo en las cintas y así poder reclamar a su proveedor. Muchas cintas, cortas y largas, y en cuatro colores (rojo, blanco, azul y amarillo) fueron estudiadas. Tres muestras de cada combinación fueron tomadas, para un total de 24 datos (2 x 4 x 3).

Plan de Estratificación

Color	Azul		Rojo		Blanco		Amarillo	
Longitud	Corto	Largo	Corto	Largo	Corto	Largo	Corto	Largo
Piezas	1 2 3	1 2 3	1 2 3	1 2 3	1 2 3	1 2 3	1 2 3	1 2 3



Ejemplo MULTI-VARI CHART

Disposición de los Datos recolectados



Color	Longitud	# Cinta	# Reclamos
Azul	Corto	1	14
Azul	Corto	2	13
Azul	Corto	3	12
Azul	Largo	1	23
Azul	Largo	2	22
Azul	Largo	3	21
Rojo	Corto	1	17
Rojo	Corto	2	16
Rojo	Corto	3	15
Rojo	Largo	1	26
Rojo	Largo	2	25
Rojo	Largo	3	24
Blanco	Corto	1	18
Blanco	Corto	2	19
Blanco	Corto	3	17
Blanco	Largo	1	27
Blanco	Largo	2	28
Blanco	Largo	3	26
Amarillo	Corto	1	12
Amarillo	Corto	2	11
Amarillo	Corto	3	10
Amarillo	Largo	1	21
Amarillo	Largo	2	20
Amarillo	Largo	3	19

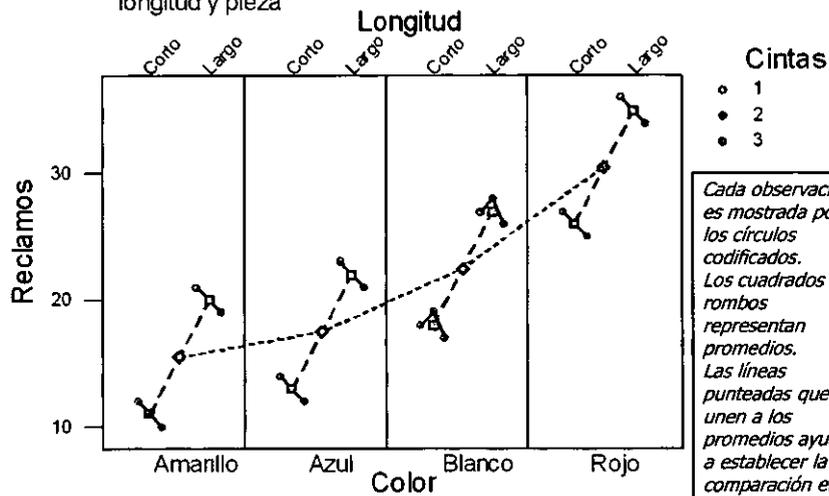
Nota: Cada Estrato debe tener asignado niveles, éstos deben tener correspondencia a un valor de la métrica de desempeño y cada combinación de niveles y estratos debe estar completa en el arreglo.

MÓDULO IV

"ANALIZAR"

Cómo luce un MULTI-VARI CHART

Multi-Vari Chart for Reclamos de Cintas por Color, longitud y pieza



¿Qué conclusión darías a la academia?



Ejercicio

- Una botella es manufacturada usando dos máquinas de inyección en dos turnos, día y noche.
- La materia prima es suministrada por dos proveedores.
- La botella sin embargo, presenta un alto % de defectivo.
- Se quiere investigar dónde está localizada la mayor fuente de variación. El índice de defectivo fue recolectado durante dos meses.

Defectivo (%)

Máquina	Mes	Materia Prima Proveedor A		Materia Prima Proveedor B	
		Turno Matutino	Turno Nocturno	Turno Matutino	Turno Nocturno
I	Julio	4.89	6.88	3.29	5.03
	Agosto	4.86	6.28	3.18	4.99
II	Julio	4.72	6.46	2.60	5.18
	Agosto	4.63	6.20	3.10	5.64



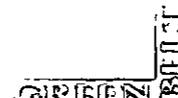
FIVE WHYS (5 Por qué)

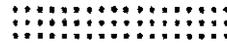
El proceso Five Why's está diseñado para encontrar las causas de un problema al someterlo a preguntas Why o ¿Por qué? y preguntando ¿Por qué? a cada una de las respuestas.

Esta serie de preguntas hará que la causa raíz sea descubierta.

Cada vez que la pregunta se realiza, la respuesta empezará a estar mejor relacionada y enfocada hasta que la pregunta no pueda ser contestada, indicando que la causa raíz ha sido encontrada.

Para desarrollar este proceso, hay que empezar con la definición del problema ENFOCADO y entonces preguntar ¿Por qué? 5 veces.





Pasos para generar un FIVE WHYS

1. Para iniciar el proceso, el problema actual necesita cambiarse a una pregunta "Por qué":

Problema: Las ventas han bajado 20% desde el año pasado

Pregunta: ¿Por qué las ventas han bajado 20% desde el año pasado?

2. Esta primera **pregunta** desencadenará una serie de respuestas :

Pregunta: ¿Por qué las ventas han bajado 20% desde el año pasado?

Respuesta #1: Porque un competidor nos ganó un cliente clave.

Respuesta #2: Porque embarcamos tarde el producto al final del año.

Respuesta #3: Porque recortamos 5% los distribuidores.



Pasos para generar un FIVE WHYS

3. Cada una de las respuestas mencionadas arriba, necesitan ser cambiadas a una pregunta "Por qué" y la cadena continuará. Tomar cada respuesta individualmente y continuar cambiando las respuestas en subsecuentes preguntas "Por qué" hasta que la pregunta no pueda ser contestada o la respuesta está fuera del control de la compañía. En este punto habrás identificado la raíz del problema que necesitará ser direccionada:

Pregunta: ¿Por qué un competidor nos ganó un cliente clave?

Respuesta #1: Nuestros competidores están ofreciendo promociones de ventas.

Respuesta #2: Nuestros productos no cumplen sus necesidades.

4. Una vez que las preguntas son contestadas, haz una lista de las diferentes causas raíz y categorízalas en grupos lógicos.

5. Determina cuál de las causas atacarás primero seleccionándolas de acuerdo a su importancia (escoger 10 o menos).





MÓDULO IV

"ANALIZAR"

Five Why's & Logic Tree

Puedes desplegar gráficamente las causas encontradas por el proceso de 5 why's haciendo uso de un Logic Tree (Árbol Lógico) para hacer más comprensible el enlace de la pregunta con sus posibles respuestas.

Problema: ¿Por qué órdenes tienen errores frecuentes en el número de Lote?

¿Por qué? ¿Por qué? ¿Por qué? ¿Por qué? ¿Por qué?

```

    graph LR
        A[Se caen las etiquetas de lote] --- B[Fallan cuando hay humedad]
        A --- C[Respuesta #2]
        A --- D[Respuesta #3]
        B --- E[Pegamento no reúne las necesidades]
        E --- F[Usamos diferentes Pegamentos]
        F --- G[No hay definida especificación]
    
```

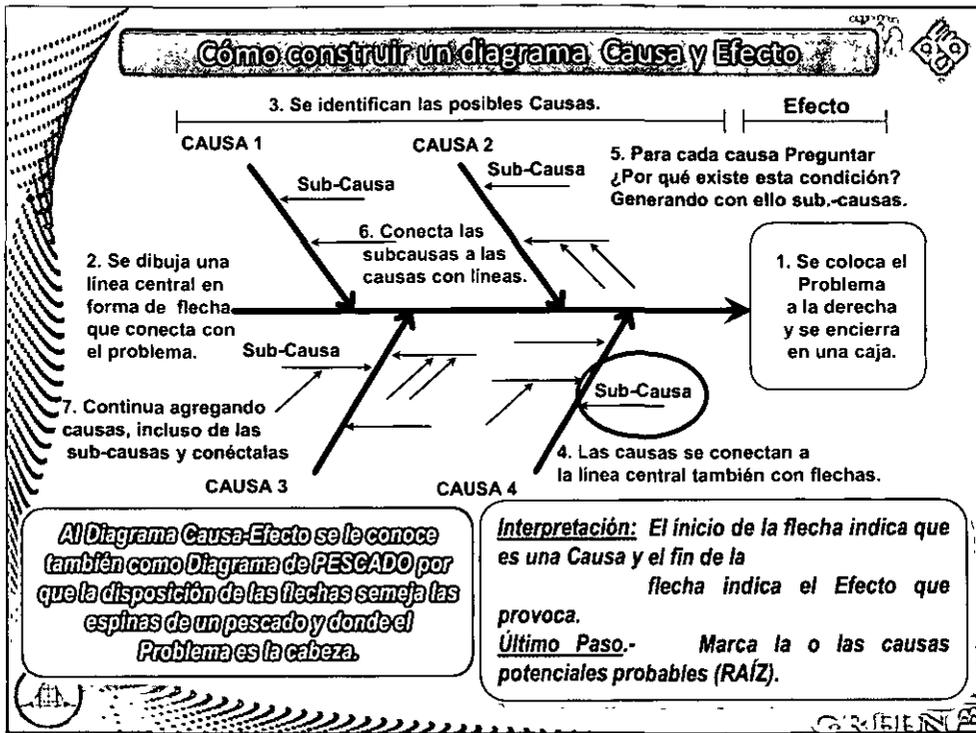
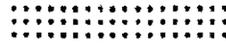
GREEN BELT

Diagrama Causa y Efecto

El diagrama Causa y efecto es una herramienta visual usada para identificar, explorar y desplegar en detalle, **todas** las posibles causas relacionadas a un problema o condición:

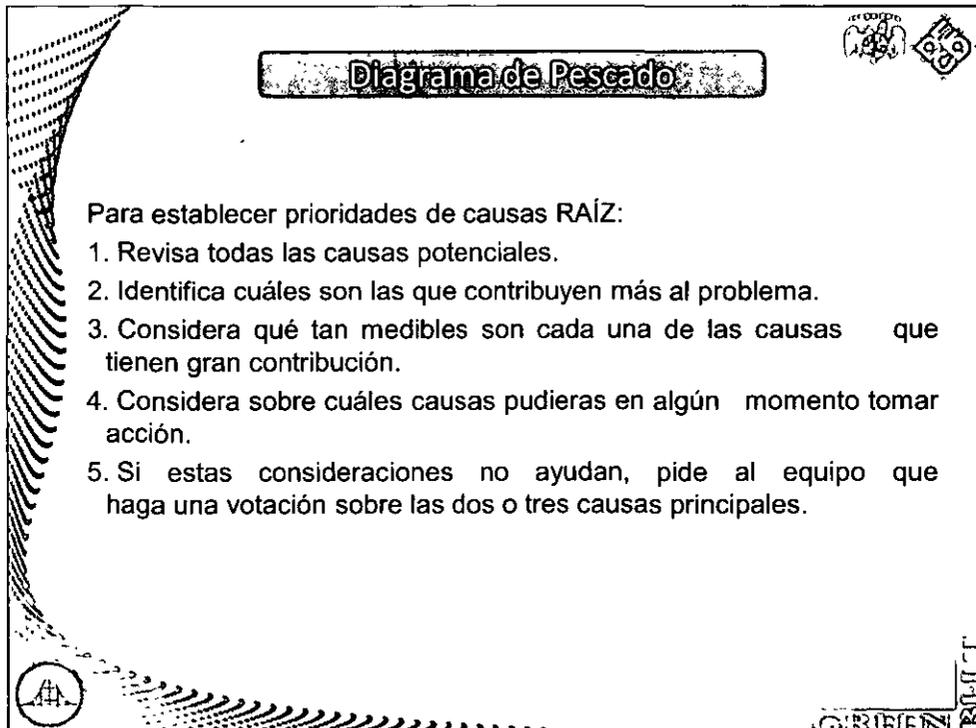
- Ayuda a enfocar al equipo en el contenido del problema, en sus causas y no en los síntomas.
- Crea una imagen instantánea del conocimiento colectivo del equipo entorno al problema.
- Crea un consenso de las causas del problema.
- Visualiza posibles **relaciones entre causas** para cualquier problema actual o futuro.
- Minimiza la tendencia natural que tiene la gente de saltar a las conclusiones de mejora sin haber estudiado antes las causas.

GREEN BELT



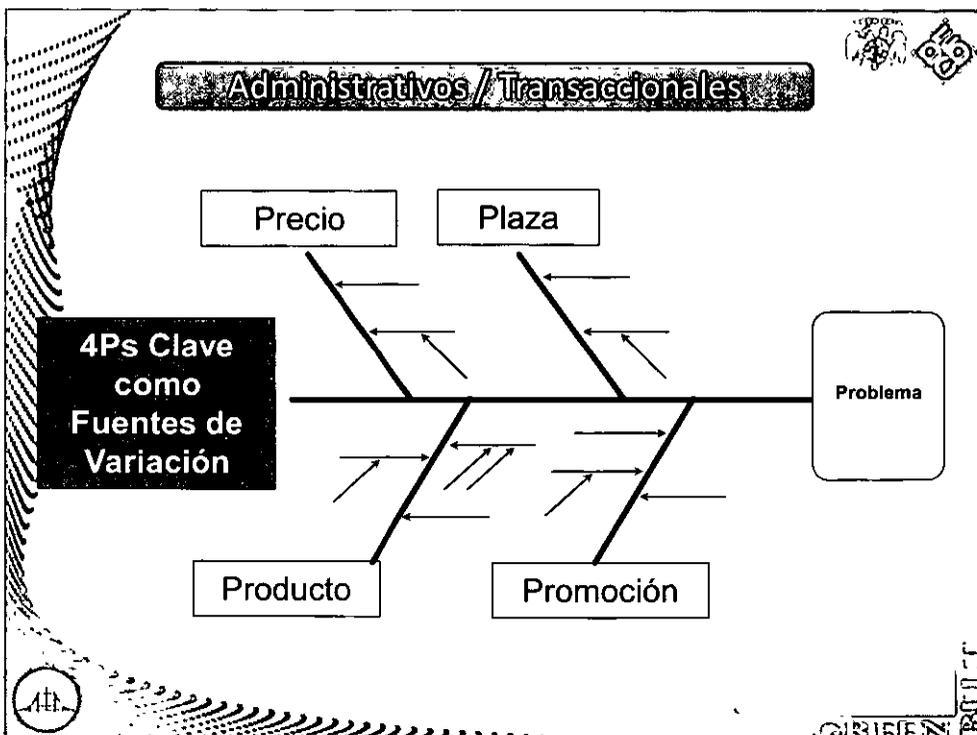
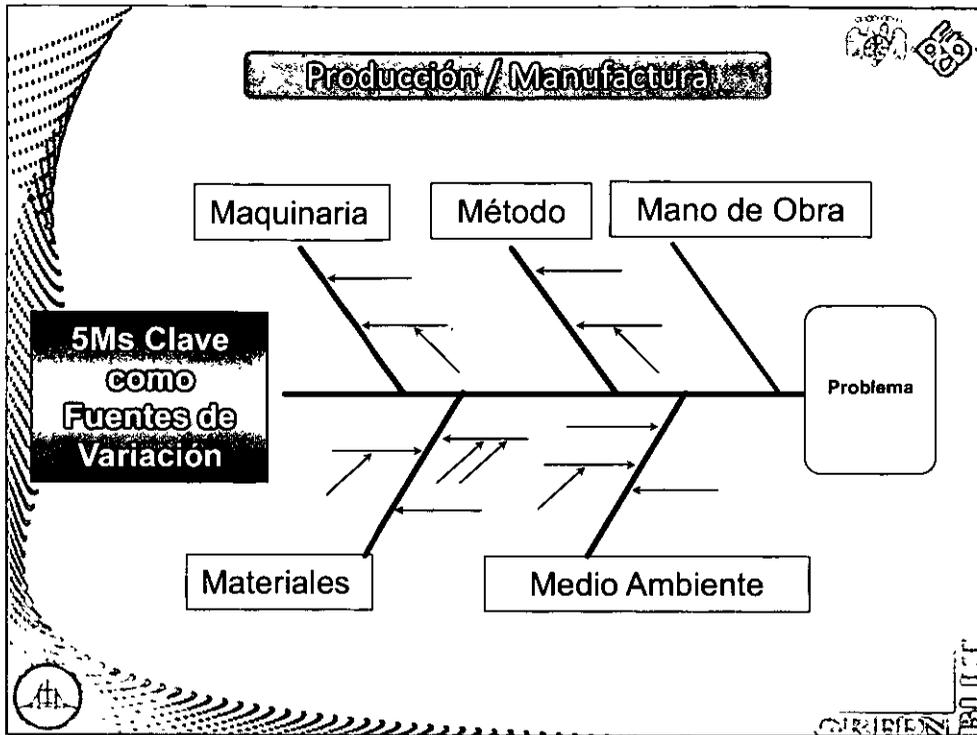
MÓDULO IV

"ANALIZAR"





MÓDULO IV



"ANALIZAR"



Matriz CAUSA-EFECTO

- La Matriz Causa & Efecto es una herramienta usada para relacionar y priorizar las entradas del proceso (Xs)-CAUSAS a los requisitos del cliente (Ys) a través de la clasificación jerárquica numérica, usando el mapa del proceso, diagramas de pescado, o de afinidades, como la fuente primaria de variables causales.
 - Las Ys (las salidas) se anotan de acuerdo a la importancia al cliente (CTQ).
 - Las Xs (las entradas) se anotan de acuerdo a la relación de cada causa con cada uno de los requerimientos.



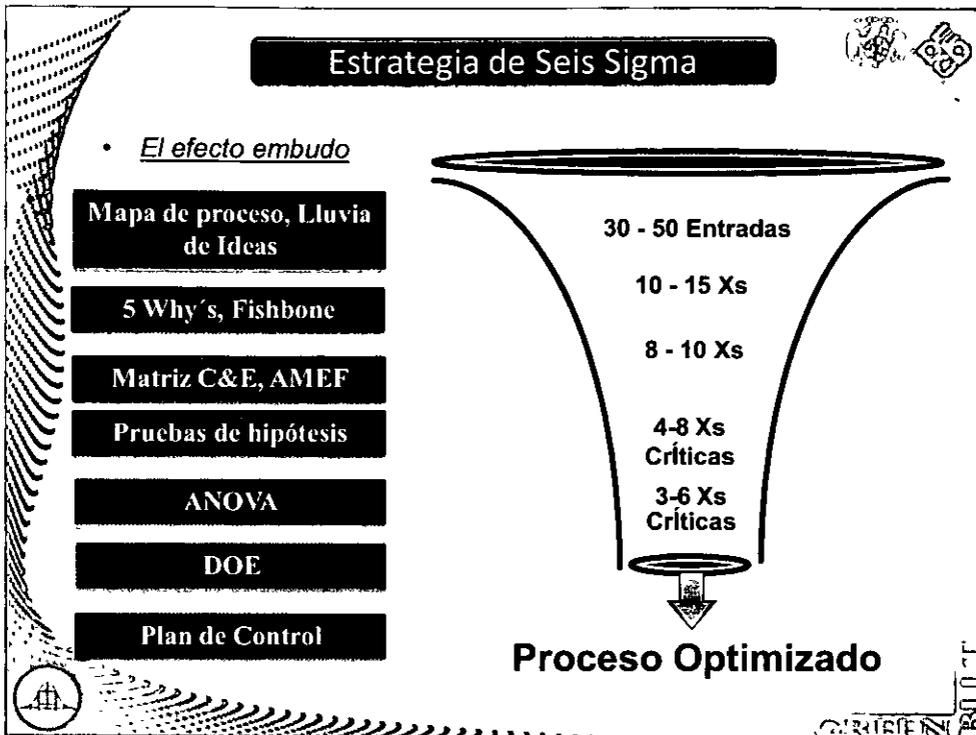
MÓDULO IV

"ANALIZAR"

Matriz CAUSA-EFECTO

- Esta herramienta permitirá a **todos** los involucrados en un proceso estar de acuerdo en las salidas que son críticas al producto y/o cliente.
- Vía la clasificación jerárquica numérica, esta matriz permitirá a su equipo asignar un nivel de importancia a cada variable de salida (Ys).
- A través de la asociación, la matriz permitirá al equipo evaluar el efecto de cada entrada numéricamente (X) en cada salida (Y).
- Esto es el primer intento de los equipos para determinar $Y=f(X)$.





MÓDULO IV

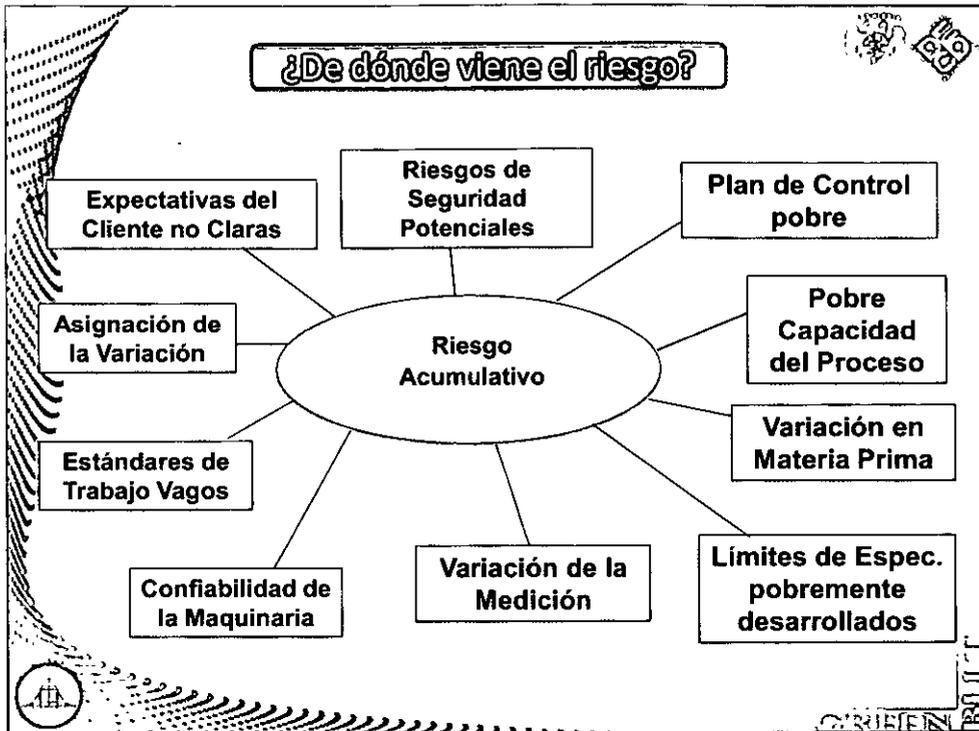
"ANALIZAR"

AMEF

El Análisis del Modo y Efecto de la Falla es también otra herramienta de EMBUDO para enfocar el esfuerzo de estudio en aquellas variables que son críticas para el proceso actual.

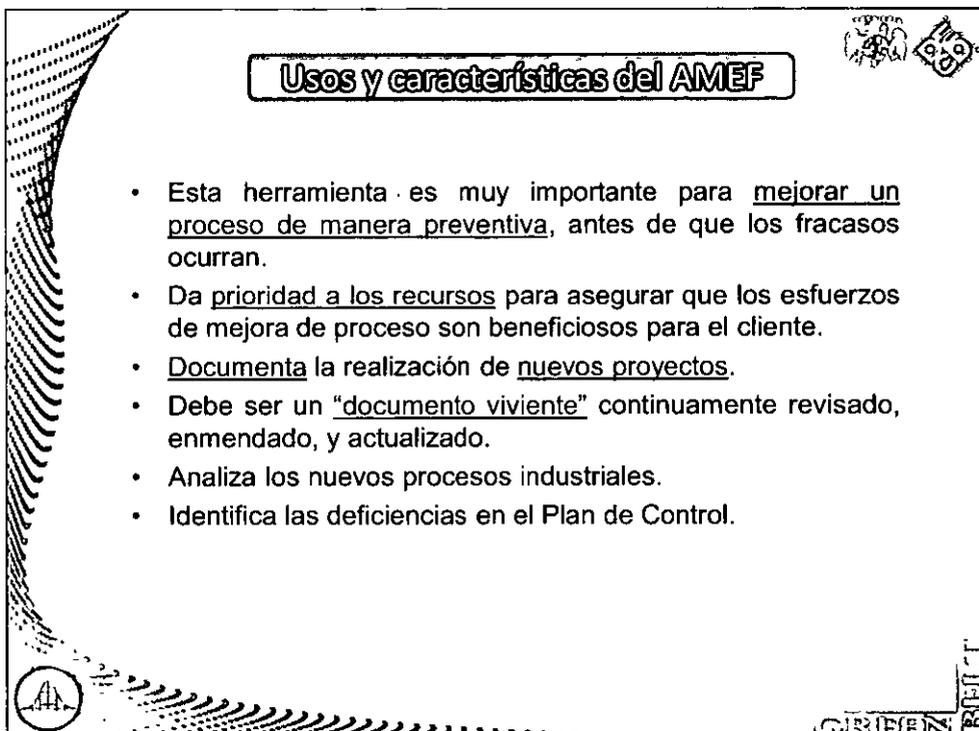
Puede ser usada tanto en la fase de Medición, o de Análisis como en la fase de Mejora.

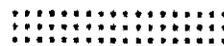
Es en si, una aproximación estructurada para identificar, estimar, priorizar y evaluar el riesgo. Esto es, ayuda para identificar las maneras en que un producto o proceso pueden fallar y da entrada para la eliminación o reducción del riesgo en relación a esas fallas para proteger al cliente.



MÓDULO IV

"ANALIZAR"





Usos y características del AMEF

- Puede establecer y dar prioridad a las acciones.
- Evalúa el riesgo de cambios del proceso.
- Identifica las variables causales potenciales a considerar
- Guía el desarrollo de nuevos procesos industriales.
- Ayuda a establecer el escenario para el descubrimiento súbito.



MÓDULO IV

"ANALIZAR"

Terminología

- **El Modo** - La manera en que una entrada específica al proceso falla, y si no es detectada, corregida o removida, causará el Efecto.
- **El Efecto** - El impacto en los requisitos del cliente. Generalmente se enfoca en el cliente externo, pero también puede incluir los procesos río abajo.
- **La Severidad** - Una valoración de la gravedad del Efecto en el cliente.
- **La Causa** - las Fuentes de variación del proceso que causa el Modo de Falla por ocurrir.





Terminología

- **La Ocurrencia** - Una valoración de la frecuencia con que ocurre una Falla.
- **Los Controles Actuales** - Métodos sistematizados / aparatos para prevenir o descubrir los Modos de Falla o Causas antes de causar los Efectos.
- **La Detección**- Una valoración de la probabilidad de que los controles actuales detectarán cuándo un Modo de Falla o la Causa ocurre.
- **Número de Prioridad de Riesgo** - $RPN = Severidad \times Ocurrencia \times la Detección$

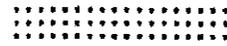
MÓDULO IV

"ANALIZAR"

Apreciación General del Formato del AMEF

MEDICIÓN			ANÁLISIS				MEJORA		
Paso de Proceso - O Entrada	Modo Potencial de Falla	Efecto Potencial de Falla	SEVERIDAD	Causas Potenciales	OCCURRENCIA	Controles Actuales	DETECTABILIDAD	RPN	Acciones Recomendadas
¿Cuál es el paso del proceso?	¿Qué puede salir mal con el paso del proceso?	¿Cuál es el efecto en las salidas?	¿Qué tan mal?	¿Cuáles son las causas?	¿Qué tan seguido?	¿Cómo podemos encontrar esto?	¿Qué tan difícil de detectar?		¿Qué puede hacerse?

(Use el formato de su organización)



Número de Prioridad de Riesgo RPN (NPR)

El resultado de un AMEF es el "Número de Prioridad de Riesgo"

El RPN es un número calculado en base a información en relación a:

- los modos de falla potencial,
- los efectos, y
- la habilidad actual del proceso de detectar las fallas antes de que lleguen al Cliente

Es calculado como el producto de tres valuaciones cuantitativas, cada una relacionada a los efectos, causas, y controles:

- **RPN = Severidad x Ocurrencia x Detección**

Efectos

Causas

Controles



GREEN BELT

MÓDULO IV

"ANALIZAR"

Anotación RPN

- Hay una variedad de escalas que pueden usarse para tasar la severidad, ocurrencias y categorías de detección de un AMEF.
- Una escala de 1 a 5 hace más fácil que los equipos decidan en las cuentas, pero falta precisión en las estimaciones.
- Una escala de 1 a 10 permite una precisión buena en la estimación, pero hace más difícil decidir en las cuentas.
- Una escala de 1, 3, 6, 9 hace más fácil decidir en las cuentas, proporciona estimaciones precisas y proporciona una variación amplia en las cuentas.



GREEN BELT



Anotación RPN (Cont.)

Ejemplo de una escala de anotación que usa 1, 3, 6, y 9; diseñada para lograr una cobertura mayor al tasar la severidad, ocurrencias y categorías de un AMEF:

Anotación	Severidad del Efecto	Posibilidad de la Ocurrencia	Habilidad de Detección
9	Defecto peligroso notado por todos los clientes	La falla es inevitable	No se detecta
6	Defectos mayores notados por mayoría de clientes	Fallas frecuentes	Posibilidad moderada de detección
3	Defecto menor notado por algunos clientes	Relativamente pocas fallas	Alta oportunidad de detección
1	No efecto	Falla poco probable	Detección segura

MÓDULO IV

"ANALIZAR"

Ejercicio de Anotación de AMEF

Llene los resultados y qué acción deberá ser tomada. Esté listo a presentar resultados.

Occ	Sev	Det	Resultado	Acciones
1	1	1		
1	1	9		
1	9	1		
1	9	9		
9	1	1		
9	1	9		
9	9	1		
9	9	9		



Respuestas

Occ	Sev	Det	Resultado	Acciones
1	1	1	Situación Ideal	No Acción
1	1	9	Dominio Seguro	No Acción
1	9	1	Falla no alcanza usuario	No Acción
1	9	9	Falla alcanza usuario	Sí
9	1	1	Fallas frec, detect, costosas	Sí
9	1	9	Fallas frec, llegan a usuario	Sí
9	9	1	Fallas frec. con impacto mayor	Sí
9	9	9	¡Problema Grande!	Sí

MÓDULO IV

"ANALIZAR"

Tips para el éxito

- Asegúrese de arreglar los problemas de seguridad.
- Las Premisas:
 - Asuma que el material entrante es perfecto y el proceso no lo es.
 - Asuma que el proceso es perfecto y el material entrante no lo es.
- Dos acercamientos, los dos son similares:
 - Empezando con la Matriz Causa & Efecto.
 - Prepare un AMEF directamente del Mapa del Proceso.



AMEF - Método

- 1) Para cada entrada del proceso, determine las maneras en que el paso del proceso puede salir mal, éstos son los Modos de Falla.
- 2) Para cada modo de falla identificado asociado con las entradas del proceso, determine los Efectos.
- 3) Identifique las Causas potenciales de cada modo de falla identificado.
- 4) Enliste los Controles Actuales para cada causa identificada.

MÓDULO IV

"ANALIZAR"

AMEF - Método cont..

- 5) Asigne las valuaciones para la Severidad, Ocurrencia y Detección.
 - Tasa de Severidad para cada uno de los Efectos
 - Tasa de Ocurrencia para la frecuencia por cada Causa.
 - Tasa de Descubrimiento para la habilidad de los Controles Actuales de detectar la Causa y/o Modo de Falla.
- 6) Calcule el Número de Prioridad de Riesgo (RPN).



AMEF – Método cont..

- 7) Determine las Acciones Recomendadas para reducir los Números de Prioridad de Riesgo altos.
- 8) Tome las acciones apropiadas y documente los resultados.
- 9) Recalcule el Número de Prioridad de Riesgo.

MÓDULO IV

"ANALIZAR"

Sumario

- El AMEF se supone ser una acción "antes de la falla", no una reacción "después del hecho".
- El AMEF es un documento "viviente" y debe ponerse al día continuamente como ocurran cambios.
- Siga la metodología exactamente.
- Incluya a todo el personal impactado para obtener información, esto incluye río arriba y abajo.
- **La importancia estadística no se ha establecido todavía, este será el siguiente paso a realizar.**



Pruebas de Hipótesis

- ¿Qué es una Hipótesis?

*Todas las suposiciones, creencias, incluso ideas locas, deben ser evaluadas y cuantificadas. Hacemos esto mediante una gama de herramientas sencillas pero muy poderosas que vamos a generalizar como: **PRUEBAS DE HIPÓTESIS.***
- Las pruebas de hipótesis nos ayudan a determinar si las causas o Xs son estadísticamente significantes para impactar en el resultado de la variable de salida Y. Esto es, si las causas tienen realmente un efecto en el síntoma.



Pruebas de Hipótesis

Para contestar la pregunta "¿tiene relación la X con la Y?" (por ejemplo: ¿Las calificaciones de los alumnos tienen relación con el método de enseñanza?), tenemos 2 opciones (sólo una prevalecerá):



Clase	Método A	Método B	Método C	Método D
1	Calificación alta	Calificación alta	Calificación alta	Calificación alta
2	Calificación alta	Calificación alta	Calificación alta	Calificación alta
3	Calificación alta	Calificación alta	Calificación alta	Calificación alta
4	Calificación alta	Calificación alta	Calificación alta	Calificación alta
5	Calificación alta	Calificación alta	Calificación alta	Calificación alta
6	Calificación alta	Calificación alta	Calificación alta	Calificación alta

- No hay relación entre las variables, no hay diferencia
no hay cambio, no pasa nada.

Hipótesis Nula (H₀)

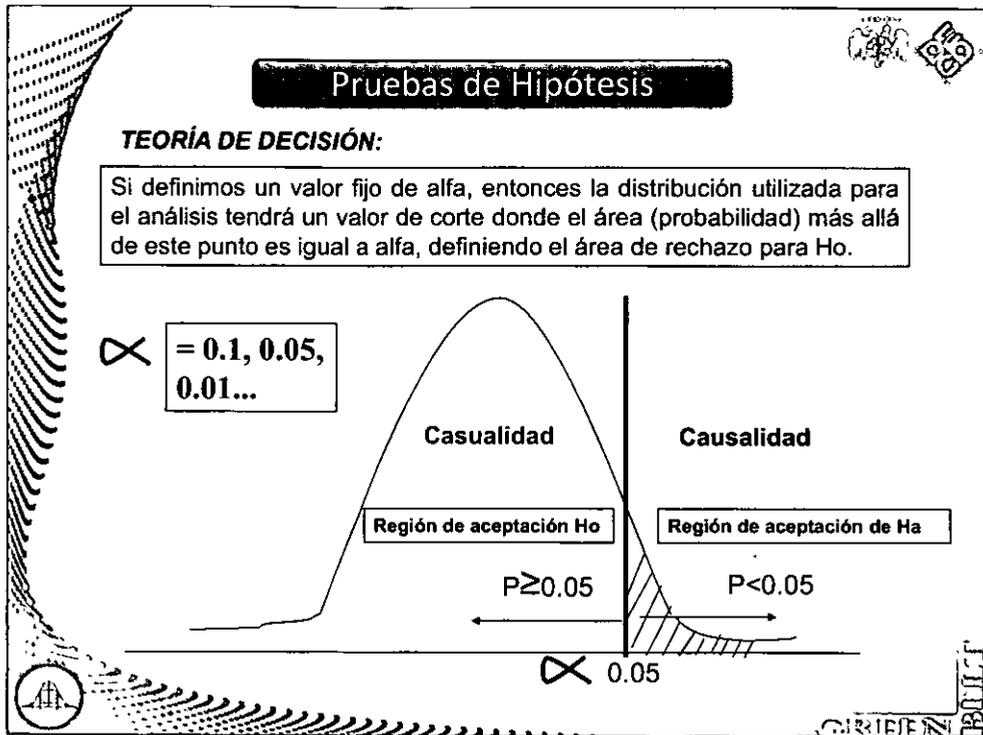
- Sí hay relación entre variables, sí hay diferencia
sí hay cambio, sí pasa algo.

Hipótesis Alternativa (H_a)



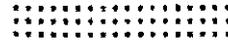
MÓDULO IV

"ANALIZAR"



Matriz de Selección de Herramientas

		FACTOR X	
		CONTINUO	DISCRETO
RESPUESTAS Y	CONTINUO	REGRESIÓN, CORRELACIÓN REGRESIÓN MULTIPLE	ANOVA PRUEBAS DE Z Y T NO PARAMÉTRICAS
	DISCRETO	REGRESIÓN LOGÍSTICA	PRUEBA DE JI CUADRADA PRUEBAS DE PROPORCIONES



Herramientas para Pruebas de Hipótesis

La calificación final del Módulo IV del Diplomado de Green Belt está determinada por la estatura de cada alumno
Herramienta:

Análisis de Regresión

La marca de cigarros vendidos depende de la zona económica
Herramienta:

Prueba de Ji-cuadrada (Chi - square)

La producción de maíz en kg depende del tipo de fertilizante usado
Herramienta:

ANOVA o prueba T

Slide content with decorative border and logos.

MÓDULO IV

"ANALIZAR"

5 Puntos para las pruebas de hipótesis

Define lo que quieres probar



Establece la "Hipótesis Nula" (H_0) y la "Hipótesis Alternativa" (H_a)



Selecciona la Herramienta Estadística



Recolecta evidencia (una muestra de la realidad) y prueba



DECIDE en función del P value:
¿Qué te sugiere la evidencia?
¿Rechazar H_0 ? o ¿No Rechazar H_0 ?

Slide content with decorative border and logos.



Ejercicios

1: En cierta comunidad suburbana se realizó una encuesta a una muestra de adolescentes para averiguar qué relación tenía el tiempo dedicado a ver televisión con el peso de los adolescentes.

Ho:

Ha:

Tipo de datos

Y: _____

X: _____

La herramienta apropiada para contestar esto es:

¿Cuál sería tu conclusión si $p = 0.1871$?



MÓDULO IV

"ANALIZAR"

Ejercicios cont..

2: ¿La memoria está relacionada con la edad en años?

Ho:

Ha:

Tipo de datos

Y: _____

X: _____

La herramienta apropiada para contestar esto es:

¿Cuál sería tu conclusión si $p = 0.051$?





MÓDULO IV

"ANALIZAR"

Ejercicios cont...

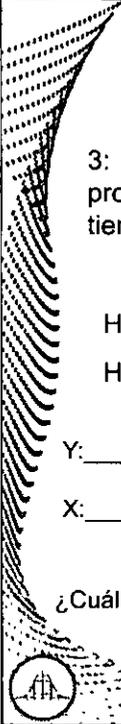
3: Arturo ha cotizado servicios de Internet de 3 diferentes proveedores, ahora debe seleccionar sólo uno con base en sus tiempos de conexión.

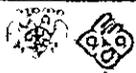
Ho: _____
 Ha: _____

Tipo de datos

Y: _____ La herramienta apropiada para contestar esto es: _____
 X: _____

¿Cuál sería tu conclusión si $p = 0.0023$?






Ejercicios cont...

4: En cierta región del país se realizó una encuesta con mujeres cuyas edades estaban entre 30 y 34 años con el objeto de determinar si existía relación entre el promedio de hijos nacidos vivos y el romanticismo de los poemas de amor recibidos durante el periodo de embarazo.

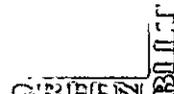
Ho: _____
 Ha: _____

Tipo de datos

Y: _____ La herramienta apropiada para contestar esto es: _____
 X: _____

¿Cuál sería tu conclusión si $p = 0.7135$?





MÓDULO IV

"ANALIZAR"

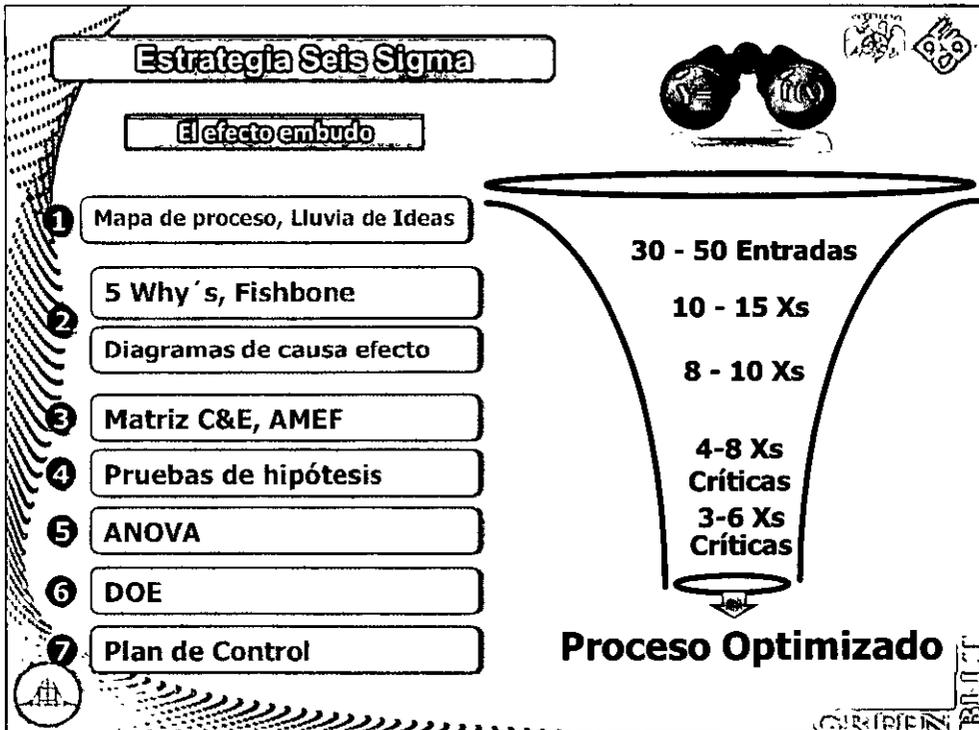


FASE DE ANÁLISIS

ENTRENAMIENTO PARA GREEN BELT

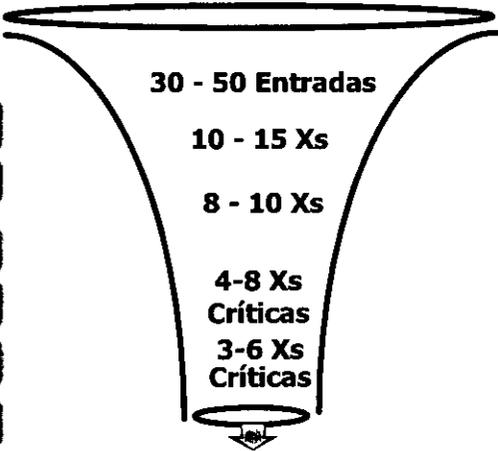


Estrategia Seis Sigma

El efecto embudo

- 1 Mapa de proceso, Lluvia de Ideas
- 2 5 Why's, Fishbone
- 3 Diagramas de causa efecto
- 4 Matriz C&E, AMEF
- 5 Pruebas de hipótesis
- 6 ANOVA
- 7 DOE
- 8 Plan de Control

30 - 50 Entradas

10 - 15 Xs

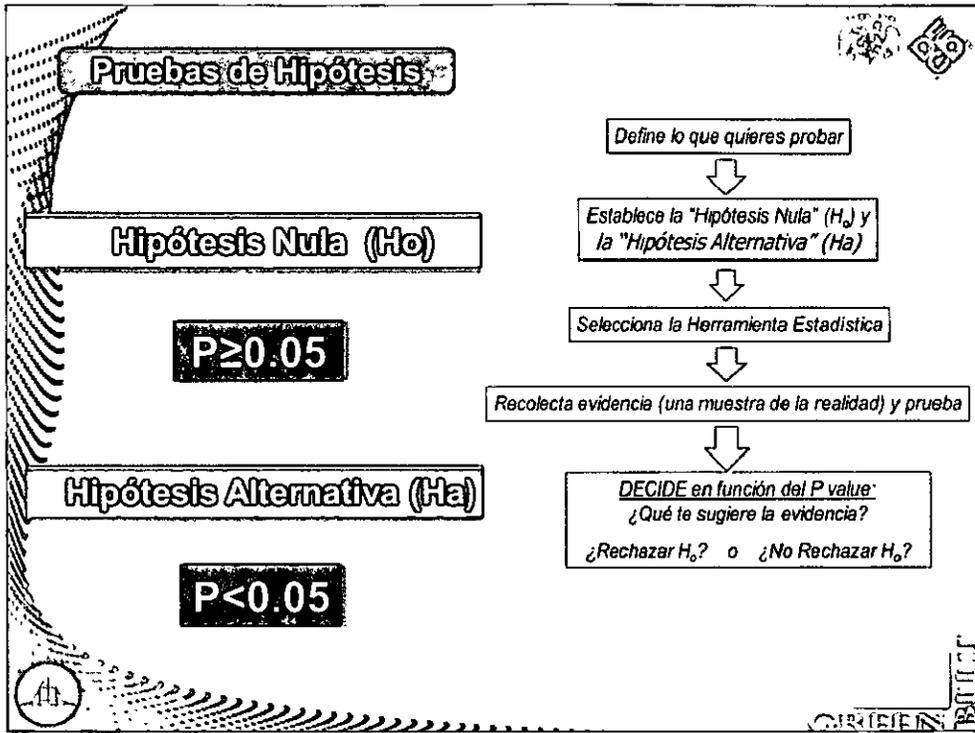
8 - 10 Xs

4-8 Xs
Críticas

3-6 Xs
Críticas

Proceso Optimizado



MÓDULO IV

"ANALIZAR"

Matriz de Selección de Herramientas

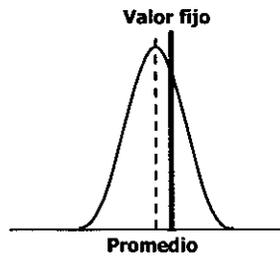
		FACTOR X	
		CONTINUO	DISCRETO
FACTOR Y	CONTINUO	REGRESIÓN, CORRELACIÓN REGRESIÓN MÚLTIPLE	ANOVA PRUEBAS DE Z Y T NO PARAMÉTRICAS
	DISCRETO	REGRESIÓN LOGÍSTICA	PRUEBA DE JI CUADRADA PRUEBA DE PROPORCIONES



Prueba Z

Al tratar con variables continuas existen 2 principales preocupaciones en el comportamiento de la distribución. El centrado y la dispersión de los datos. Es muy útil contar con un modelo específico para el análisis. Para una gran diversidad de procesos, su distribución puede ser explicada por la distribución normal.

¡No te preocupes
recuerda la
distribución Z!



¿Podemos afirmar que existe diferencia entre el promedio de la población y el valor fijo?

MÓDULO IV

"ANALIZAR"

Prueba Z

Si es normal siempre podemos estandarizar nuestros datos y buscar probabilidades (valores de p) en las tablas Z.

$$Z = \frac{\bar{X} - \mu}{\sigma}$$

← Esta es información de la POBLACIÓN

Para usarla como una herramienta para contrastar hipótesis debemos de adaptarla para comparar los datos de la muestra contra la media de la población hipotética.

$$Z = \frac{x - \mu}{s / \sqrt{n}}$$



MÓDULO IV

"ANALIZAR"

Prueba Z

$$Z = \frac{\bar{x} - \mu}{s / \sqrt{n}}$$

Media muestra

Valor de referencia

Error estándar de la Media

Con el valor de Z buscamos los valores de p para determinar si la diferencia entre las dos medias es significativa.

Prueba Z Ejemplo

Si un ingeniero de procesos afirma que cambios específicos al Lay Out reducirían el tiempo de ciclo a menos de 30 segundos. El gerente no se impresiona y afirma que la reducción se debe a variación aleatoria del proceso. "Algunos días son mejores que otros" dijo. Tu decides. ¿quién tiene la razón?.

	C. Time	C. Time	C. Time
1	37.3	16	29
2	22.5	17	30.3
3	30.3	18	25.6
4	29.5	19	28.8
5	26.7	20	29.6
6	30.5	21	25.8
7	26.3	22	29
8	29.9	23	29
9	29	24	25.6
10	25.2	25	29.1
11	27.6	26	27.2
12	28.4	27	32
13	23.6	28	29.6
14	27.8	29	30.4
15	25.5	30	27.3

$H_0: \mu = 30$ (la media del proceso es igual a 30 segundos)

$H_a: \mu < 30$ (la media es menor a 30 segundos)

$$Z = \frac{\bar{X} - \mu}{S/\sqrt{n}} = \frac{28.3 - 30}{2.8/\sqrt{30}} = -3.37$$

Z	0	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
-3.3	0.00012	0.00015	0.00019	0.00025	0.00032	0.00041	0.00051	0.00062	0.00075	0.00090

Dado que el valor de p es menor de 0.05, concluimos que **hay suficiente evidencia** de que la reducción en el tiempo ciclo es real.



MÓDULO IV

"ANALIZAR"

Prueba t

La prueba anterior es válida para tamaños de muestra grandes (>30) y si los datos están normalmente distribuidos. Es muy común que tamaños de muestra muy grandes sean difíciles de conseguir. Restricciones tales como costo, tiempo, etc. Pueden ser de consideración al tomar muestras. Cuando tenemos pequeñas muestras que sabemos que vienen de una distribución normal, utilizamos un estadístico similar para pequeñas muestras: la distribución **t**

¿cuál es la diferencia?

$$t = \frac{\bar{X} - \mu}{S/\sqrt{n}}$$

La distribución **t** tiene mayor dispersión debido a menores tamaños de muestra

Prueba t

Mismo concepto
¡¡Diferentes fórmulas!!

Región de aceptación (1-a)

Región de rechazo (a)

Valor de corte

Grados de libertad (GL)

$t_{\alpha, n-1}$

Nota: Los GL son los que afectan la dispersión de la distribución. A mayor muestra, más grados de libertad y más cercanos serán nuestros datos a la distribución de la población.



Prueba t

También hay tablas para la distribución t

df	3.00	3.05	3.10	3.15	3.20	3.25	3.30	3.35	3.40	Value for t											
										3.45	3.50	3.55	3.60	3.65	3.70	3.75	3.80	3.85	3.90	3.95	4.00
1	0.1024	0.1008	0.0993	0.0978	0.0964	0.0950	0.0937	0.0923	0.0911	0.0896	0.0886	0.0874	0.0862	0.0851	0.0840	0.0830	0.0819	0.0809	0.0799	0.0789	0.0780
2	0.0477	0.0464	0.0451	0.0439	0.0427	0.0416	0.0404	0.0394	0.0383	0.0374	0.0364	0.0356	0.0346	0.0338	0.0330	0.0322	0.0314	0.0307	0.0299	0.0293	0.0286
3	0.0288	0.0277	0.0266	0.0256	0.0247	0.0237	0.0229	0.0220	0.0212	0.0205	0.0197	0.0190	0.0184	0.0177	0.0171	0.0166	0.0160	0.0155	0.0150	0.0145	0.0140
4	0.0200	0.0190	0.0181	0.0173	0.0165	0.0157	0.0150	0.0143	0.0136	0.0130	0.0124	0.0119	0.0114	0.0109	0.0104	0.0100	0.0096	0.0092	0.0088	0.0084	0.0081
5	0.0150	0.0142	0.0134	0.0127	0.0120	0.0113	0.0107	0.0102	0.0096	0.0091	0.0086	0.0082	0.0078	0.0074	0.0070	0.0066	0.0063	0.0060	0.0057	0.0054	0.0052
6	0.0120	0.0113	0.0106	0.0099	0.0093	0.0087	0.0082	0.0077	0.0072	0.0068	0.0064	0.0060	0.0057	0.0054	0.0050	0.0048	0.0045	0.0042	0.0040	0.0038	0.0036
7	0.0100	0.0093	0.0087	0.0081	0.0075	0.0070	0.0066	0.0061	0.0057	0.0053	0.0050	0.0047	0.0044	0.0041	0.0038	0.0036	0.0034	0.0031	0.0029	0.0028	0.0026
8	0.0085	0.0079	0.0073	0.0068	0.0063	0.0059	0.0054	0.0050	0.0047	0.0043	0.0040	0.0038	0.0035	0.0032	0.0030	0.0028	0.0026	0.0024	0.0023	0.0021	0.0020
9	0.0075	0.0069	0.0064	0.0059	0.0054	0.0050	0.0046	0.0043	0.0039	0.0036	0.0034	0.0031	0.0029	0.0027	0.0025	0.0023	0.0021	0.0020	0.0018	0.0017	0.0016
10	0.0060	0.0056	0.0051	0.0047	0.0044	0.0040	0.0037	0.0034	0.0031	0.0029	0.0026	0.0024	0.0022	0.0021	0.0019	0.0017	0.0016	0.0015	0.0014	0.0013	0.0013
11	0.0050	0.0046	0.0041	0.0038	0.0034	0.0031	0.0028	0.0025	0.0022	0.0020	0.0018	0.0016	0.0015	0.0013	0.0011	0.0010	0.0009	0.0008	0.0007	0.0006	0.0006
12	0.0040	0.0036	0.0031	0.0028	0.0025	0.0022	0.0019	0.0016	0.0014	0.0012	0.0010	0.0009	0.0008	0.0007	0.0006	0.0005	0.0004	0.0003	0.0003	0.0002	0.0002
13	0.0035	0.0031	0.0026	0.0023	0.0020	0.0017	0.0014	0.0011	0.0009	0.0007	0.0006	0.0005	0.0004	0.0003	0.0002	0.0002	0.0001	0.0001	0.0000	0.0000	0.0000
14	0.0030	0.0026	0.0021	0.0018	0.0015	0.0012	0.0009	0.0006	0.0004	0.0003	0.0002	0.0001	0.0001	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
15	0.0025	0.0021	0.0016	0.0013	0.0010	0.0007	0.0004	0.0002	0.0001	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
16	0.0020	0.0016	0.0011	0.0008	0.0005	0.0002	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
17	0.0015	0.0011	0.0006	0.0003	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
18	0.0010	0.0006	0.0001	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
19	0.0005	0.0001	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
20	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
21	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
22	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
23	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
24	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
25	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
26	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
27	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
28	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
29	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000

Ahora debemos conocer los GL

MÓDULO IV

"ANALIZAR"

Prueba t

	C		C.
	Time		Time
1	37.3	16	29
2	22.5	17	30.3
3	30.3	18	25.6
4	29.5	19	28.8
5	26.7	20	29.6
6	30.5	21	25.8
7	26.3	22	29
8	29.9	23	29
9	29	24	25.6
10	25.2	25	29.1
11	27.6	26	27.2
12	28.4	27	32
13	23.6	28	29.6
14	27.8	29	30.4
15	25.5	30	27.3

Ho: $\mu = 30$ (la media del proceso es igual a 30 segundos)
 Ha: $\mu < 30$ (la media es menor a 30 segundos)

$$t = \frac{\bar{X} - \mu}{S/\sqrt{n}} = \frac{28.3 - 30}{2.8/\sqrt{30}} = -3.37$$

Dado que la t es simétrica buscamos el valor cuando t=3.37. En la tabla de 1 cola:

df	3.00	3.05	3.10	3.15	3.20	3.25	3.30	3.35	3.40
29	0.0027	0.0024	0.0021	0.0019	0.0017	0.0015	0.0013	0.0011	0.0010

Dado que el valor de p es menor de 0.05, llegamos a la misma conclusión de antes.



Prueba t Ejercicio

Con los siguientes datos determina la prueba de hipótesis que se te pide.

12.5	11.9	12.1	12.3	11.5
12.6	11.7	12.6	12.3	11.6
12.7	11.9	12.7	12.2	11.7

- La media es diferente de 12.3
- La media es mayor a 11.7
- La media es menor de 12



GRUPO
BELL

MÓDULO IV

"ANALIZAR"

Prueba t Ejercicio

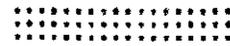
Un cliente cree que las lámparas que compró e instaló en su casa están consumiendo gran cantidad de Watts, por lo tanto, incrementa el gasto por concepto de energía eléctrica. La compañía proveedora envió una muestra de 10 lámparas a probar y confirmar esta situación. Los Watts consumidos por las lámparas fueron:

61.3	60.8	61.9	62.2	61.5
60.7	59.9	60.9	61.8	62.0

La especificación de las lámparas se supone que es de 60.5 Watts. La pregunta es: ¿el cliente tiene una queja legítima?



GRUPO
BELL



Prueba t Ejercicio

En un esfuerzo por mejorar los tiempos de respuesta, el departamento de Sistemas convirtió su sistema de atención a clientes actual a un nuevo sistema de call center para ayuda. Históricamente, el tiempo de espera ha sido de 7 minutos en promedio. Una muestra de 10 solicitantes del servicio fue monitoreado para ver si el tiempo de espera con el nuevo sistema había alcanzado su objetivo de tener un tiempo menor.

Los siguientes datos fueron recolectados durante el periodo de evaluación:

5.17	6.45	8.35	7.05	6.55
7.11	5.3	6.03	5.59	6.14

Diga si el nuevo sistema alcanzó el objetivo.

MÓDULO IV

"ANALIZAR"

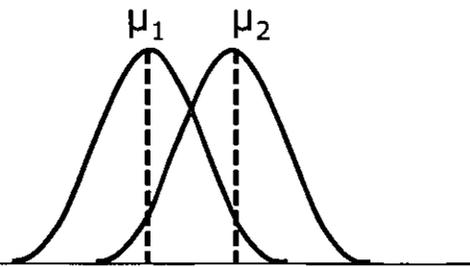
Prueba t <2 muestras>

1-sample T (Prueba de T de 1 muestra)

Propósito: Analizar las diferencias entre la media obtenida y un valor objetivo o una media histórica.

2-sample T (Prueba de T de 2 muestras)

Propósito: Analizar las diferencias entre la media obtenida de dos muestras independientes.





Prueba t <2 muestras>

- Usamos la prueba estadística llamada t-test para comparar y juzgar diferencias entre promedios de dos grupos.
- La hipótesis nula es que los promedios de los dos grupos son los mismos: $H_0 = \mu_a = \mu_b$
- La hipótesis alternativa es que los promedios son diferentes: $H_a = \mu_a \neq \mu_b$
- Al realizar la prueba y obtener el valor de p tendremos que:
Si p-value es < 0.05
 rechazamos la hipótesis nula y aceptamos la hipótesis alternativa.



GREEN BELT

MÓDULO IV

"ANALIZAR"

Prueba t <2 muestras>

○ Si **p-value es ≥ 0.05** se concluye que no hay suficiente evidencia para rechazar la hipótesis nula por que:

- Los grupos son los mismos, o
- La variación es muy grande, o
- La muestra es muy pequeña como para detectar una diferencia



GREEN BELT



Prueba t <2 muestras> Ejercicios

Servicio a clientes decide llamar a la planta para determinar si la calidad del producto que manejan está reuniendo la especificación requerida. Pide entonces al gerente de calidad datos de porcentajes de defectivo en los dos turnos que actualmente producen el producto. El gerente de calidad le envía la siguiente información:

Turno A	60.8	60.3	61.0	59.7	60.9
	59.9	59.8	60.5	60.1	60.6
Turno B	60.8	61.2	60.9	60.5	61.1
	60.7	61.0	60.9	60.8	60.5

En promedio, ¿están los dos turnos manufacturando el producto con la misma calidad?



GRUPO
SIX
SIGMA
GREEN BELT

MÓDULO IV

"ANALIZAR"

Prueba t <2 muestras> Ejercicios

La dureza de agua es medida en términos de la concentración de calcio (en millones). La dureza del agua en pipas de agua caliente y fría en un proceso de manufactura fue medida. Un técnico reclama que el agua caliente es más dura (tiene mayor concentración de calcio) que el agua fría. Las concentraciones de calcio de las muestras tomadas se muestran en la tabla.

Agua Caliente	Agua Fría
133.5	134.0
135.4	134.7
137.2	136.0
138.4	131.7
136.3	134.6
137.1	135.2
133.3	135.9
136.5	135.6
137.5	
139.4	

Establece tus hipótesis y da tus conclusiones



GRUPO
SIX
SIGMA
GREEN BELT



Prueba t <2 muestras> Ejercicios

Dos marcas de filtro de aire (A y B) para unidades de aire acondicionado son probadas. Todos los filtros fueron probados en la misma unidad y la cantidad de polvo (en gramos) filtrado en 6 horas fue medido. Se quiere determinar si el filtro B es superior al A. Los datos para ambos filtros son:

Filtro A	Filtro B
9.1	15.6
11.8	9.3
1.5	16.9
7.2	5.1
4.2	14.5
9.6	19.0
8.7	10.3
10.2	12.5
4.4	13.3
	16.1
	2.6

Establece tus hipótesis y conclusiones.

MÓDULO IV

"ANALIZAR"

Matriz de Selección de Herramientas

		FACTOR X	
		CONTINUO	DISCRETO
FACTOR Y	CONTINUO	REGRESIÓN, CORRELACIÓN	ANOVA PRUEBAS DE Z Y t
	DISCRETO	REGRESIÓN MÚLTIPLE	NO PARAMÉTRICAS
		REGRESIÓN LOGÍSTICA	PRUEBA DE Ji CUADRADA PRUEBA DE PROPORCIONES



Prueba Ji Cuadrada

Esta prueba de hipótesis es usada para comparar dos o más proporciones de grupos. Es usada cuando ambas "X" y "Y" son discretas. Los conteos son sumarizados en una tabla conocida como Tabla de contingencia.

La Ji cuadrada mide las diferencias entre los conteos observados y esperados de la siguiente forma:

$$\chi^2 = \sum \frac{(\text{Observado} - \text{Esperado})^2}{\text{Esperado}}$$



ORIENTE BELT

MÓDULO IV

"ANALIZAR"

Prueba Ji Cuadrada

Variable discreta Y:

Donde Un atributo es registrado para cada unidad

-Por ejemplo: Preciso o no preciso, Tipos de errores (direcciones o nombres equivocados)

Donde el número de unidades con cada atributo puede ser contado.

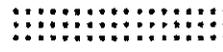
• Variable discreta X:

Datos estratificados en grupos ("por" variable)

-Por ejemplo: Localidad, método, tipo de producto, etc.



ORIENTE BELT



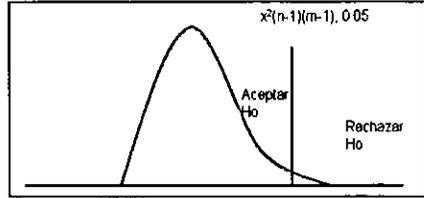
MÓDULO IV

"ANALIZAR"

Valor de Prueba de Ji Cuadrada

Consideraciones:

- La muestra es representativa de la población o proceso.
- La distribución es Binomial para datos discretos.



Valor:

- Los datos discretos son usados normalmente para analizar el desempeño de los procesos en aplicaciones de servicio.
- Conociendo si hay o no diferencias significativas entre dos o más grupos prevé el hacer cambios en el proceso sin desperdiciar tanto tiempo.
- Si las diferencias entre los grupos son identificadas vale la pena buscar causas raíces en los grupos que son diferentes.
- Recuerda siempre considerar diferencias estadísticas significantes por que son importantes para el negocio.

Prueba de Ji Cuadrada Procedimiento

- Consideremos los siguientes datos, donde se quiere demostrar que el tipo de error presentado en la facturación no difiere de la persona.

Tipos de Errores	Valores Observados				1. Totales	2. %
	Arturo	Juan	Luis	Marisol		
Fallas en la captura	3	6	5	1	15	0.16
Errores en los conceptos	7	9	5	1	22	0.24
Cálculos erróneos	9	12	11	23	55	0.60
3. Valor esperado Para fallas en La captura de Arturo	19	27	21	25	92	
	3.10	4.40	3.42	4.08	15	



MÓDULO IV

"ANALIZAR"

Prueba de Ji Cuadrada Procedimiento

Tipos de Errores	Arturo	Juan	Luis	Marisol	Total	%
Fallas en la captura	3	6	5	1	15	0.1630
Errores en los conceptos	7	9	5	1	22	0.2391
Cálculos erróneos	9	12	11	23	55	0.5978
Totales	19	27	21	25	92	
Valores esperados	3.10	4.40	3.42	4.08		
	4.54	6.46	5.02	5.98		
	11.36	16.14	12.55	14.95		

4. $\chi^2 = \sum \frac{(\text{Observado} - \text{Esperado})^2}{\text{Esperado}} = 0.003 + 0.580 + 0.726 + 2.321 + 1.328 + 1.002 + 0.000 + 4.146 + 0.490 + 1.063 + 0.192 + 4.341 = 16.191$

P-Value = 0.013

DF=Grados de libertad = (filas-1)(columnas-1)= 6

• P<0.05 por lo que se concluye que el tipo de error depende de la persona.

Prueba de Ji Cuadrada Ejercicio

- Tres locaciones están siendo analizadas para entender el impacto de la distancia (en millas) con respecto al daño de producto embarcado.
 - Locación A está a 200 millas del cliente
 - Locación B está a 500 millas del cliente
 - Locación C está a 50 millas del cliente
- Una muestra de 100 piezas fue trasladada a cada locación presentando la siguiente condición.
 - Locación A = 5 piezas dañadas
 - Locación B = 8 piezas dañadas
 - Locación C = 10 piezas dañadas
- Se quiere determinar si la cantidad de daños está relacionada a la distancia del punto de embarque al cliente.

	Locación A 200 millas	Locación B 500 millas	Locación C 50 millas
Dañadas	5	8	10
Buenas	95	92	90



Prueba de Ji Cuadrada Ejercicio

Con base en la información presentada en la parte de abajo, decide si la salida de un procedimiento quirúrgico depende del hospital usado.

Procedimiento quirúrgico	Hospital A	Hospital B	Hospital C	Hospital D	Hospital E
NI	13	5	8	21	43
PFR	18	10	36	56	29
CFR	16	16	35	51	10

Considera un mismo tamaño de muestra.

MÓDULO IV

"ANALIZAR"

Prueba de Ji Cuadrada Ejercicio

Determina si un particular escantillón crea más o menos defectos dependiendo del método de prueba.

Escantillón	Método A	Método B	Método C
No.1	37	41	44
No.2	35	72	71

Asegura un mismo tamaño de muestra.



Prueba de Proporciones

Es similar a la prueba de Chi-cuadrada, sólo que la medida de interés son las proporciones de una característica en vez de las frecuencias esperadas. Trabaja para atributos que tienen sólo dos posibles opciones (sí/no, bueno/malo, a tiempo/tarde, etc.)

La prueba busca diferencias significativas entre las proporciones de 2 poblaciones diferentes.

Ho: $p_1 = p_2$
(No hay diferencia entre las proporciones)

Ha: $p_1 \neq p_2$
(Las proporciones son diferentes)

¿Qué otras alternativas existen?

Prueba de Proporciones

Para el análisis necesitamos:
Identificar dos proporciones

$p = \frac{\text{característica contada}}{\text{Total muestreado}}$; p_1, p_2

Ejemplo: Total de defectuosos

Calcula la proporción combinada

$$P_{\text{pooled}} = \frac{x_1 + x_2}{n_1 + n_2}$$

Total de características (pointing to numerator)
Total muestreado (pointing to denominator)

Calcula el estadístico de prueba: $Z = \frac{p_1 - p_2}{se(p_1 - p_2)}$

Donde $se(p_1 - p_2) = \sqrt{P_{\text{pooled}}(1 - P_{\text{pooled}}) (1/n_1 + 1/n_2)}$



Prueba de Proporciones Ejemplo

Un estudio de Harvard esperaba determinar si el consumo de Aspirina tenía influencia en la reducción de ataques de corazón. En un periodo de 5 años, 22701 voluntarios fueron monitoreados. Los voluntarios se dividieron en 2 grupos: el grupo 1 tomó un placebo diariamente y el grupo 2 recibió aspirina en forma diaria. Los resultados se muestran en la siguiente tabla:

	Ataque	n	p
Grupo 1	239	11034	0.0217
Grupo 2	139	11037	0.0126

¿La diferencia es suficiente para determinar que la aspirina previene los ataques al corazón?

MÓDULO IV

"ANALIZAR"

Prueba de Proporciones Ejemplo

Ho: $p_1 = p_2$ (No hay diferencia entre las proporciones)
 Ha: $p_1 \neq p_2$ (Existe diferencia entre las proporciones)

$$P_{pooled} = \frac{x_1 + x_2}{n_1 + n_2} = \frac{239 + 139}{11034 + 11037} = 0.0173$$

$$se(p_1 - p_2) = \sqrt{P_{pooled} (1 - P_{pooled}) (1/n_1 + 1/n_2)} =$$

$$\sqrt{0.017(0.98)(1/11034 + 1/11037)}$$

$$se(p_1 - p_2) = 0.00175$$

$$Z = \frac{p_1 - p_2}{se(p_1 - p_2)} = \frac{0.0217 - 0.0126}{0.00175} = 5.2$$

Rechazar Ho.
 ¡concluimos que existe suficiente evidencia estadística para afirmar que la aspirina previene los ataques al corazón!

El valor de p deseado es: $2(1 - 0.99999) = 0.00002$



MÓDULO IV

"ANALIZAR"

Prueba de Proporciones Ejemplo

Se quiere determinar si un nuevo programa de computación producirá menos defectivo de formas en un proceso de compras. En la tabla se muestran datos recolectados antes y después de la implementación del software, determina si el nuevo software es una mejora.

	Programa Anterior	Programa Nuevo
Buenas	143	162
Malas	12	8

Prueba de Proporciones Ejemplo

Una compañía está revisando el número de llamadas atendidas para la satisfacción del cliente por dos diferentes equipos. ¿Un equipo hace mejor trabajo que otro?

	Llamada Contestadas OK	Llamadas Recibidas
Equipo 1	1435	1536
Equipo 2	1472	1636



MÓDULO IV

"ANALIZAR"

Proporciones VS. χ^2

1 Las pruebas de proporciones tienen mayor capacidad de detectar diferencias de la Chi-cuadrada y son más fáciles de calcular. Sin embargo, requieren de una muestra de mayor tamaño.

2 La Chi-cuadrada es más versátil que la prueba de proporciones. La prueba de proporciones sólo es útil cuando hay dos posibles salidas, (go, no-go; sí, no; etc.) y la Chi cuadrada trabaja con variables que tienen múltiples niveles.

Matriz de Selección de Herramientas

		FACTOR X	
		CONTINUO	DISCRETO
FACTOR Y	CONTINUO	REGRESIÓN, CORRELACIÓN REGRESIÓN MÚLTIPLE	ANOVA PRUEBAS DE Z Y T NO PARAMÉTRICAS
	DISCRETO	REGRESIÓN LOGÍSTICA	PRUEBA DE CHI CUADRADA PRUEBA DE PROPORCIONES



Correlación

1 La correlación es usada para determinar el grado de relación entre dos variables de proceso. Permite la comparación de una entrada con una salida, dos entradas (una en contra de la otra), o bien dos salidas (una en contra de la otra).

2 La correlación mide entonces el grado de asociación entre dos variables continuas independientes. Sin embargo, aún cuando haya un alto grado de correlación, esta herramienta no establece causalidad. Por ejemplo, el número de accidentes en esquí en Colorado está altamente correlacionado con las ventas de chamarras térmicas, pero sabemos que comprar chamarras térmicas no causa los accidentes.



GREEN BELT

MÓDULO IV

"ANALIZAR"

Correlación

3 La correlación es analizada calculando el coeficiente de correlación de Pearson.

4 Este es un coeficiente cuyo valor fluctúa entre -1 y 1.

5 Un valor positivo indica correlación positiva, que es cuando una variable aumenta su valor conforme la otra también lo amerita.

6 Un valor negativo indica correlación negativa, que es cuando una variable aumenta su valor mientras la otra disminuye.

7 Si las variables no están correlacionadas, el coeficiente de correlación se acerca a 0.



GREEN BELT



MÓDULO IV

"ANALIZAR"

Coefficiente de Correlación

$$r_{xy} = \frac{\frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y})}{S_x S_y}$$

Donde: S_x y S_y son las desviaciones estándar de la muestra

Guía para el Coeficiente de Correlación: r

-1 ← -0.8 — -0.2 — 0 — 0.2 — 0.8 → 1

Correlación Negativa fuerte	Correlación Negativa moderada	No hay Correlación O es débil	Correlación Positiva moderada	Correlación positiva fuerte
-----------------------------------	-------------------------------------	-------------------------------------	-------------------------------------	-----------------------------------

GREEN BELT

Coefficiente de Correlación

<p>No hay correlación $r = 0$</p>	<p>Correlación Moderada positiva $r = 0.6$</p>	<p>Correlación Fuerte positiva $r = 0.8$</p>
<p>Correlación Perfecta positiva $r = 1.0$</p>	<p>Correlación Fuerte negativa $r = -0.8$</p>	<p>Correlación No-lineal $r = 0$</p>

GREEN BELT



MÓDULO IV

"ANALIZAR"

CORRELACIÓN Ejercicios

Un profesor de primaria sugiere que la memoria está relacionada con la edad más que con el grado escolar, para ello, ha recolectado las variables edad (en años y décimos de años) de cada uno de los alumnos que resolvieron una prueba de memoria donde obtuvieron el puntaje mostrado en la tabla de datos.

Determine el grado de correlación entre estas variables y diga sus conclusiones.

Edad (años y décimos)	Memoria (puntos)
7.3	0.6
7.2	0.4
6.5	0.4
7.8	0.7
6.4	0.2
7.1	0.6
8.0	0.9
8.6	0.9
8.5	0.6
8.3	0.6

Edad (años y décimos)	Memoria (puntos)
9.5	1.1
9.2	0.8
8.8	0.7
9.0	0.9
9.0	1.0
9.7	1.2
10.4	1.1
10.3	1.3
9.5	0.9
10.0	0.9

Edad (años y décimos)	Memoria (puntos)
11.5	1.6
11.0	1.4
11.0	1.5
10.7	1.2
11.3	1.2
11.8	1.4
12.0	1.8
12.5	1.6
12.2	1.7
12.8	1.9

CORRELACIÓN Ejercicios

Una psicoterapeuta con muchos años de servicio le dice a uno más joven que las personas que viven en viviendas pequeñas con muchas personas tienen un nivel elevado de estrés. El terapeuta joven se muestra escéptico, por lo que decide estudiar la asociación entre las variables densidad habitacional (número de habitantes de la vivienda entre superficie de construcción de la misma) y la calificación promedio de un estudio de estrés. Determine r y diga que concluyó el terapeuta.

Calificación promedio del nivel de estrés	Densidad habitacional
9.1	0.06
7.5	0.17
8.0	0.12
7.2	0.21
8.5	0.09
6.0	0.19
6.8	0.14
6.6	0.22
8.2	0.10

Calificación promedio del nivel de estrés	Densidad habitacional
6.4	0.18
5.4	0.25
6.3	0.26
7.4	0.12
5.8	0.21
7.6	0.08
8.1	0.14
7.0	0.19
7.4	0.10



MÓDULO IV

"ANALIZAR"

CORRELACIÓN

Ejercicios

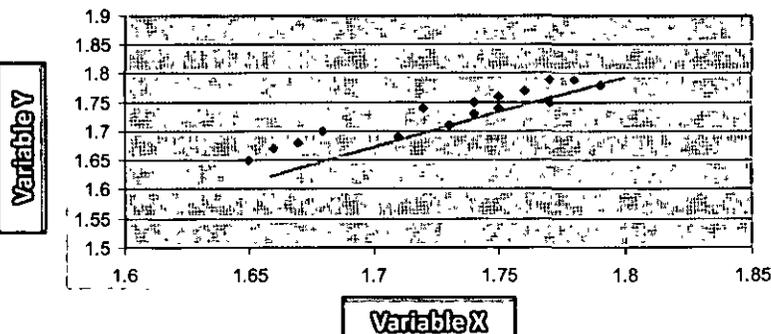
El departamento de recursos humanos de una empresa quiere averiguar cómo el ratio de trabajadores ausentes los días lunes se relaciona con la temperatura media de ese día. Una muestra aleatoria de 10 lunes se utilizó en el estudio de la empresa. En la tabla se muestran los datos que indican las variables estudiadas.

Determine el grado de asociación entre las variables y grafíquelo.

Ratio de Ausentismo (%)	Temperatura media (°C)
8	10
7	20
5	25
4	30
2	40
3	45
5	50
6	55
8	59
9	60

Análisis de Regresión

Desde un punto de vista gráfico, el análisis de regresión comienza con un diagrama de dispersión. Dos variables continuas se grafican una contra otra para definir si existe correlación entre ambas.



El Análisis de Regresión trata de ajustar una línea recta que minimice el error entre los puntos dispersos.



Análisis de Regresión

El análisis de regresión genera una línea que cuantifica la relación entre una "X" y la "Y". La línea, o ecuación de regresión, es representada como:

$$Y = mX + b$$

Donde: m = pendiente de la línea (es el cambio en "Y" por unidad incrementada en "X")

b = punto de intersección, donde la línea cruza a "Y" cuando "X" = cero

EL MODELO

¿Por qué cuantificar la relación?

El análisis de regresión puede ser usado para predecir el valor de "Y" con un valor conocido de "X".

Si la "X" siendo medida puede ser controlada, es posible producir salidas deseables al cambiar las condiciones del proceso.



MÓDULO IV

"ANALIZAR"

EL MODELO

- El modelo de regresión es bueno dependiendo de qué tan bien describe la variación entre dos variables.
- El coeficiente de correlación (r) indica qué tan pronunciada es la pendiente.
- El cuadrado de la correlación indica qué tanta variación (cambios en la Y) puede ser explicada por las variables de regresión.

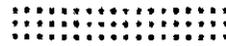
LA DECISIÓN

r^2 = Cuadrado del Coeficiente de correlación

r^2 Mide la proporción de variación que es explicada por el modelo, por la ecuación de regresión.

$r^2 = 1$	$r^2 = 1$	$r^2 = 0$

Regla de Oro: Si $r^2 \geq 0.80$ El modelo de predicción es válido.



MÓDULO IV

"ANALIZAR"

Análisis de Regresión Ejercicios

- En el curso de Green Belt se describió un experimento que consiste en observar cómo varía la longitud de un resorte, puesto verticalmente y con el extremo superior fijo, cuando varía la fuerza que se aplica al extremo inferior.
- Los datos que ahí se reportan son los mostrados en la tabla.

- Demuestre si hay relación entre las variables.
- Encuentre la recta de regresión correspondiente a estos datos, y
- Determine cuál sería la longitud del resorte si se le aplicara una fuerza de 44.82 gramos.

Fuerza aplicada (gramos)	Longitud del resorte (cm)
1.2	20
8.3	21.8
24.0	24.0
24.9	26.0
33.2	28.0
41.5	30.1
49.8	32.2
63	36.0
66.0	36.3
74.7	38.3
83.0	40.3
93.3	42.5
102	44.6
107.0	46.6
116.2	48.6
125	50.5

Análisis de Regresión Ejercicios

- Para tener más argumentos ante la Ley de No fumadores, un médico pretende demostrar al público que mientras más cigarrillos fume una persona, menor es el volumen de aire que puede emitir en una sola expiración. Para ello, toma una muestra de 12 personas a quienes pregunta cuántos cigarrillos fuman diariamente en promedio, y con cada uno realiza el siguiente experimento. Les pide que aspiren fuertemente y que inflen un globo todo lo que puedan en una sola expiración; luego anuda el globo, lo sumerge en un recipiente con agua y mide el volumen de agua desplazado. Los datos obtenidos se muestran en la tabla.

- Diga ¿A qué conclusión llegó el doctor?
- Encuentre la recta de regresión
- Prediga el valor de Y cuando: $X = 7.2$

Promedio diario de cigarrillos	Volumen de aire en el globo (cm ³)
16.9	185
10.0	293
42.3	125
8.4	220
0.0	317
32.6	186
34.1	114
49.2	81
4.7	277
28.0	126
18.8	253
28.9	225



Análisis de Regresión Ejercicios

En la Facultad de Ingeniería de la UNAM, se aplicó una encuesta en el diplomado de SS a los alumnos para conocer ciertas características tales como X = estatura y Y = peso. De la población se seleccionó una muestra aleatoria de alumnos; las tablas proporcionan la información obtenida.

Calcule la recta de regresión, interprete el resultado de r^2

Estatura (cm)	Peso (kg)
160	50
156	56
155	56
162	51
154	48
161	60
155	45
150	50
159	51
152	55
156	46
157	49

Estatura (cm)	Peso (kg)
158	55
169	61
155	51
160	54
154	55
159	49
152	58
162	63
157	49
153	49
150	52
143	46

FASE DE ANÁLISIS

1) *Lluvia De Ideas* 2) *Análisis Gráfico* 3) *Análisis Estadístico*

Técnicas como brainstorming y mapa de procesos nos permiten enfocarnos en áreas de oportunidad al compartir la experiencia "entre" los miembros del equipo, lo que nos permitirá enlistar las prioridades de mejora de acuerdo con las causas raíz más importantes encontradas.

El análisis gráfico nos permitirá:

- Confirmar las causas raíz ("lo podemos ver" y no sólo eso "pensamos que es una causa raíz").
- Enfocar en un área en específico (por ejemplo, con estratificación).

3) El análisis estadístico nos permitirá confirmar o rechazar teorías o hipótesis basadas en los factores más representativos o de mayor peso. Con esto entendemos el riesgo que existe al tomar decisiones basadas en análisis incorrectos.

Debido a que el análisis estadístico requiere de tiempo y en algunos casos es complejo, éste se elaborará después de una exhaustiva búsqueda y después de una preselección de los datos que contengan las posibles causas raíz.



SUMARIO DE LA FASE DE ANÁLISIS

Propósito:

- Priorizar las variables de entrada* que causan la variación en "Y"
- Analizar los datos para determinar causas raíz*
- Validar* las variables clave de entrada con **DATOS**

Preguntas a responder	
¿Quién?	¿Quién es el dueño del proceso?
¿Cuáles?	¿Cuáles son todas las variables clave de entrada del proceso? ¿Has encontrado alguna que se pueda mejorar? ¿Qué resistencia has experimentado o anticipas?
¿Dónde?	¿Dónde fueron recolectados los datos para las entradas?
¿Cuándo?	¿Cuándo te diste cuenta de las oportunidades para direccionar el problema? ¿Pudiste cuantificar con mayor precisión los beneficios (COPQ) del proyecto?
¿Por qué?	¿Por qué varía la salida del proceso? ¿Cuáles son las entradas más importantes?
¿Cómo?	¿Cómo has analizado los datos para identificar los factores pocos vitales que cuentan para la variación en el proceso? ¿Cómo fueron verificadas la "Xs" de tu diagrama C&E? ¿Cuáles son las "Xs" de tu diagrama C&E verificadas? ¿Cuáles son las causas raíz del problema?

SUMARIO DE LA FASE DE ANÁLISIS

Checklist

- Todas las fuentes de Variación Identificadas y Priorizadas
- Usar y Desplegar Datos para identificar y verificar los factores "pocos vitales"
- Enunciado del problema refinado reflejando el incremento del entendimiento del mismo
- Estimar las oportunidades cuantificables representadas por el problema

```

graph TD
    Define --> Measure
    Measure --> Analyze
    Analyze --> Improve
    Improve --> Control
    Control --> Define
    
```

Herramientas

- Mapa de Proceso
- Técnicas graficas (pareto, boxplot, Estudios Multivari, Fishbone, 5 why's)
- Matriz de priorización, AMEF
- Pruebas de Hipótesis
- Regresion & Correlación
- Análisis de Stakeholder, Riesgos



Establecimiento de la línea base.

Validar
el MSA

Objetivo: minimizar la variación que se podría obtener en factores controlables.

El estudio Gage R&R es un método usado para analizar un sistema de medición para determinar la cantidad y el tipo de variación (error) cuando se mide algo.



GREEN BELT

“Fase - Medir”

Módulo III

Análisis del Sistema de Medición (MSA).

Al evaluar la capacidad del Sistema de Medición nos estamos preguntando
¿Estamos midiendo adecuadamente la variable en estudio (Entrada ó Salida)?

Herramientas a utilizar:

Plan de recolección de datos

Estudio Gage R&R (Repetibilidad & Reproducibilidad)

por variables y por atributos.



GREEN BELT



Estudios Gage R&R.



El estudio Gage R&R nos permite:

- ✓ Determinar si el error de medición es pequeño y aceptable relativo a la variación del proceso ó especificaciones del producto.
- ✓ Determinar la confianza de la certeza de los datos.
- ✓ Obtener una adecuada resolución del Gage.
- ✓ Enfocar los esfuerzos de mejora si la variación de la medición es inaceptable.



Módulo III

"Fase - Medir"

Definamos algunos conceptos.



Medición: es una estimación de un objeto.

Tratamos de medir las cosas para compararlas contra una referencia (estándar) y tomar una decisión.

Al realizar las mediciones esperamos que se acerquen lo más posible a la referencia pero no siempre es así debido a la variación.

¿puedes citar algunos ejemplos?





Ejercicio...

Contabiliza cuantas vocales hay en este texto...

“La utilización incorrecta de Comparaciones Múltiples lleva a los investigadores a conclusiones erróneas que se reflejan en ambos tipos de error (tipo I y tipo II). En un trabajo de Coward (1991) sobre la utilización de las pruebas de comparaciones múltiples en Estados Unidos se detectan cuatro posibles situaciones que pueden conducir a error en la aplicación de las pruebas:”



“Fase - Medir”

Módulo III

Ejercicio...

Contabiliza cuantas vocales hay en este texto...

“La utilización incorrecta de Comparaciones Múltiples lleva a los investigadores a conclusiones erróneas que se reflejan en ambos tipos de error (tipo I y tipo II). En un trabajo de Coward (1991) sobre la utilización de las pruebas de comparaciones múltiples en Estados Unidos se detectan cuatro posibles situaciones que pueden conducir a error en la aplicación de las pruebas:”





Definamos algunos conceptos.

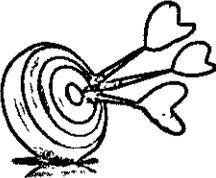
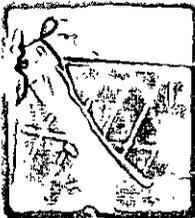
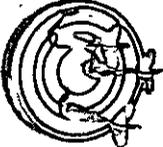
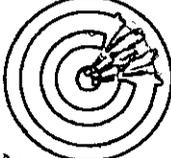
Exactitud: Minimizamos problemas con la exactitud calibrando el equipo. Nos comparamos contra un estándar y ajustamos según sea necesario.

Precisión: para problemas de precisión nos preguntamos qué tan consistentes son las mediciones de nuestro instrumento y qué tan consistentes son las mediciones de una persona y otra. (Repetibilidad y Reproducibilidad).






Exactitud vs. Precisión

	Exacto y Preciso	
	Sin Exactitud y No Preciso	
	Preciso mas no exacto	








Definamos algunos conceptos.

Estudio Gage R&R: Es una serie de evaluaciones para verificar la Repetibilidad y Reproducibilidad del sistema de medición.

¿Cómo se lleva a cabo?

Diferentes operadores repetidamente miden el objeto que está en estudio. Los resultados nos dirán cuánta variación es atribuible al operador y al sistema de medición.

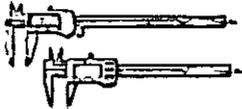




GREEN BELT

Definamos algunos conceptos.

Repetibilidad.



Consistencia con los equipos de medición.

Reproducibilidad.



Consistencia entre personas.





GREEN BELT



Definamos algunos conceptos.

Resolución: Es la diferencia más pequeña que un instrumento de medición puede registrar (L,ml).

Es la habilidad del sistema de medición para discriminar diferencias en la medición.

¿Qué equipo tiene mejor resolución?

1.01	1.02	1.03	1.04	1.05	1.06	1.07	Resolución 0.01
1.1	1.2	1.3	1.4	1.5	1.6	1.7	Resolución 0.1

El instrumento de medición deberá tener una resolución menor ó igual al 10% de la especificación ó variación del proceso.

Definamos algunos conceptos.

La variación observada de cualquier grupo de datos es la suma de la variación real de las partes más la variación del sistema de medición.

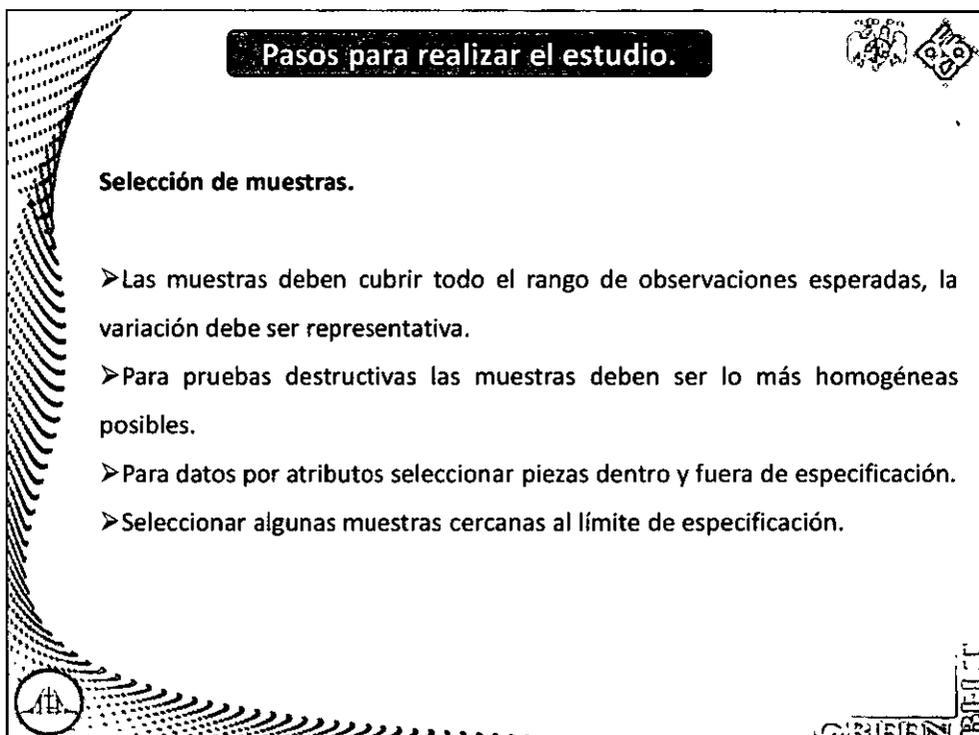
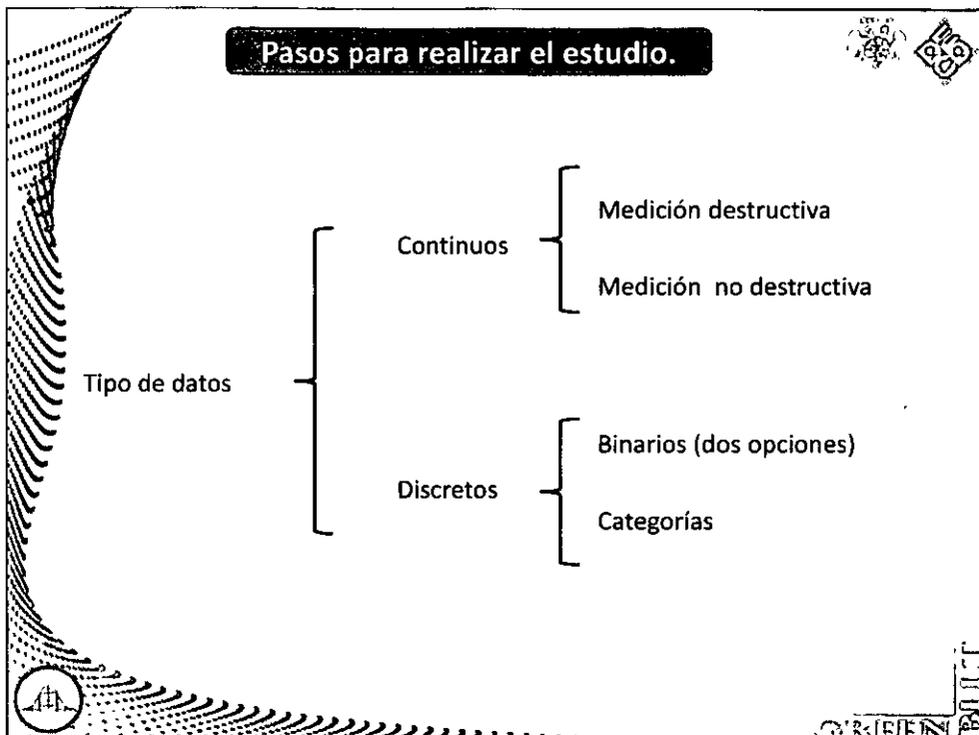
$$\sigma^2_{total} = \sigma^2_{parte a parte} + \sigma^2_{R \& R}$$

Donde:

σ^2 : variación

Parte a Parte: entre piezas

R&R: Estudio de Medición






Pasos para realizar el estudio.

Recolección de Datos.

- Asegurar que el instrumento esté calibrado y que tenga la resolución adecuada.
- Usar como mínimo dos operadores, los operadores que normalmente hacen las mediciones son los que deben desarrollar el estudio.
- Medir 10 unidades.
- Cada unidad será medida 2-3 veces por cada operador en forma aleatoria.







Pasos para realizar el estudio.

Análisis de datos:

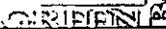
Datos Continuos.

Método Corto: Rangos.

Anova - Medias y Rangos (X-R)

Datos Discretos.

Datos Binarios (solo dos opciones) y por categorías.



Criterio de Validación del MSA Parte 1.

Inaceptable si es mayor de 30%
Condicional si está entre 20% y 30%
Aceptable si es menor al 20%

GREEN BELT

“Fase - Medir”

Módulo III

Pasos para realizar el estudio.

Método Corto: Rangos
 Proporciona un estimado rápido de la variación de la medición.
 Solo requiere 5 piezas y dos operadores.
 Los cálculos se pueden hacer a mano.
 La contribución de Repetibilidad y Reproducibilidad no pueden separarse.

Anova - Medias y Rangos (X-R).
 Muestra una mejor estimación de la variación.
 La contribución de Repetibilidad y Reproducibilidad se analizan por separado
 Requiere más recolección de datos.

GREEN BELT



Ejemplo.

Se tiene un producto de 2 pulgadas con una tolerancia de ± 0.015 para que pueda ser ensamblado, se necesita realizar un estudio R&R para determinar si el sistema de medición es adecuado, los datos obtenidos son los siguientes:

Parte	Juan	Luis
1	2.003	2.001
2	1.998	2.003
3	2.007	2.006
4	2.001	1.998
5	1.999	2.003

Realiza el ejercicio y determina si es capaz ó no el Sistema de Medición

Módulo III

"Fase - Medir"

Ejemplo.

De las fórmulas anteriores tenemos:

$$R\&R = 5.15\sigma_{\text{gage}} / \text{Tolerancia} \times 100$$

$$\sigma_{\text{gage}} = R \text{ promedio} / d^* = 0.003 / 1.19 = 0.002521$$

$$R\&R = 5.15 \times 0.002521 / (2.015 - 1.985) \times 100 = 43.3\%$$



Ejemplo en Excel.

Parte	Juan	Luis	Rango [R]
1	2.003	2.001	0.002
2	1.998	2.003	0.005
3	2.007	2.006	0.001
4	2.001	1.998	0.003
5	1.999	2.003	0.004
		Suma	0.015
		Promedio R	0.003
σ	0.00252101		
Cte	5.15		
Especificación tolerancia	2	R&R	43.3%
LSE	0.015		
LIE	2.015		

¿...y para pruebas destructivas?

Escoger muestras que minimicen la variación de las porciones.

Analizar los datos igual que con muestras no destructivas (ANOVA).

Contemplar que la variación asignada a la repetibilidad incluye variación de las porciones.

Muestra 1

1 2 3 porciones

↓

1 1 1

Muestra 2

1 2 3 porciones

↓

2 2 2

Muestra 2

1 2 3 porciones

↓

3 3 3



Estudio Método ANOVA.

Para la configuración del Estudio en Minitab se configura de la siguiente manera:

Diseño del estudio:

3 inspectores
2 intentos
10 muestras

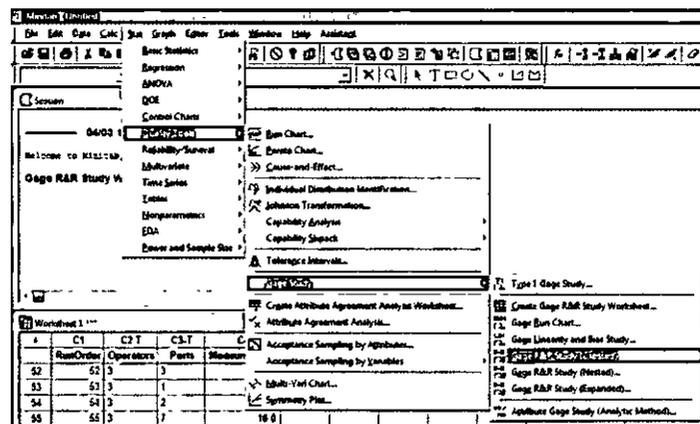
RunOrder	Operators	Parts	Measurement
1	1	9	10
2	1	10	9.8
3	1	3	9.9
4	1	4	10.1

Módulo III

“Fase - Medir”

Estudio Método ANOVA.

Realicemos el ejercicio con Minitab, sigue la ruta...



Stat>Quality Tools>Gage Study> Gage Study (Crossed)



Estudio Método ANOVA.

Realicemos el ejercicio con Minitab, sigue la ruta...

Gage R&R Study (Crossed)

Part numbers: 5.15

Operators: Operators

Method of Analysis: ANOVA

OK Cancel

Gage R&R Study (Crossed) - ANOVA Options

Study variation: 5.15 (number of standard deviations)

Process tolerance: Enter at least one specification limit

Alpha to remove interaction term: 0.25

OK Cancel

Llena los cuadros y selecciona método Anova, en la parte de "options" se escribe 5.15, da click en ok.

Para la especificación coloca

LSE: 10.4

LIE: 9.6

"Fase - Medir"

Módulo III

Estudio Método ANOVA.

¿Qué nos proporciona?

Source	VarComp	%Contribution (of VarComp)
Total Gage R&R	0.0171019	100.00
Repeatability	0.0126667	74.07
Reproducibility	0.0044352	25.93
Operators	0.0000000	0.00
Operators*Parts	0.0044352	25.93
Part-To-Part	0.0000000	0.00
Total Variation	0.0171019	100.00

Process tolerance = 0.8

¿Qué observas del % de contribución?



Estudio Método ANOVA.

¿Qué nos proporciona?

Source	StdDev. (SD)	Study Var (5.15 * SD)	%Study Var (%SV)	%Tolerance (SV/Toler)
Total Gage R&R	0.130774	0.673486	100.00	84.19
Repeatability	0.112546	0.579613	86.06	72.45
Reproducibility	0.066597	0.342976	50.93	42.87
Operators	0.000000	0.000000	0.00	0.00
Operators*Parts	0.066597	0.342976	50.93	42.87
Part-To-Part	0.000000	0.000000	0.00	0.00
Total Variation	0.130774	0.673486	100.00	84.19

Number of Distinct Categories = 4 El número de categorías debe ser 4 para una adecuada resolución

% Study Var: Estudio de variación de la muestra, se relaciona al sistema de medición

% Tolerance: se relaciona para determinar la aceptación ó rechazo del producto

Módulo III

“Fase - Medir”

Criterio de Validación del MSA Parte 2.

El MSA lo validaremos por dos criterios:

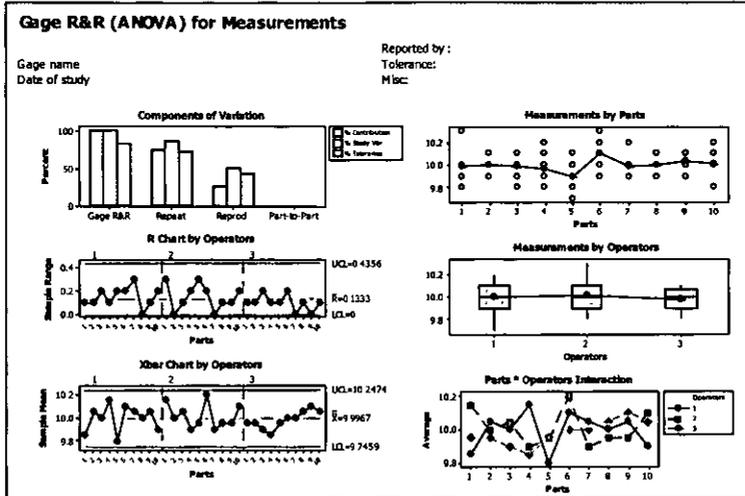
1. % de tolerancia Y % de estudio de Variación Gage
2. % de Contribución

Estatus	Método1	Método2
Inaceptable	Mayor a 30%	Mayor 8%
Condicional	Entre 20% y 30%	2% al 8%
Aceptable	Menor a 20%	Menor al 8%

¿qué conclusiones obtienes?



ANOVA – Gráficamente.



Variación de los operadores

Mediciones individuales

Módulo III

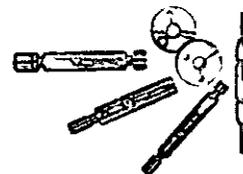
“Fase - Medir”

Estudio para datos por Atributos.

Los atributos tienen menos información que los datos variables, pero a veces es con lo único que contamos.

El problema es el mismo:

¿Puedo confiar en los datos que proporciona mi sistema de medición?





Estudio para datos por Atributos.

Selecciona un mínimo de 20 unidades del proceso. Estas unidades deben representar TODA la variación del proceso (buenas, erróneas y en límites).

Un inspector "Experto" realiza una evaluación de cada parte, clasificándola como "Pasa" o "No Pasa".

Cada persona evaluará las unidades, independientemente, en orden aleatorio y las definirá como "Pasa" o "No Pasa".



"Fase - Medir"

Módulo III

Estudio para datos por Atributos.

Inspector	Pieza	Clasificación	Experto
Hugo	1	P	P
Paco	1	P	P
Luis	1	P	P
Andres	1	P	P
José	1	P	P
Carlos	1	P	P
Hugo	2	P	P
Paco	2	P	P
Luis	2	NP	P
Andres	2	NP	P
José	2	P	P
Carlos	2	P	P
Hugo	3	NP	NP
Paco	3	NP	NP
Luis	3	NP	NP
Andres	3	NP	NP
José	3	NP	NP
Carlos	3	NP	NP
Hugo	1	P	P
Paco	1	P	P
Luis	1	P	P

Diseño del estudio:
 6 inspectores
 2 intentos
 3 muestras

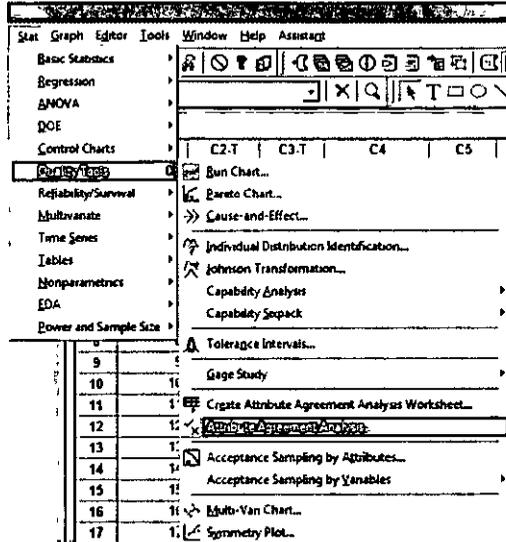




Analicemos los datos en Minitab.

Realicemos el ejercicio con Minitab, sigue la ruta...

Stat>Quality
Tools>Attribute
Agreement Analysis



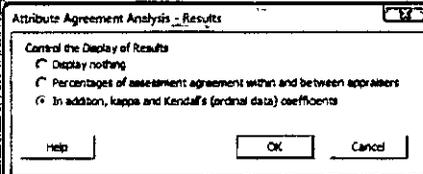
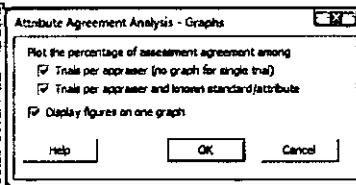
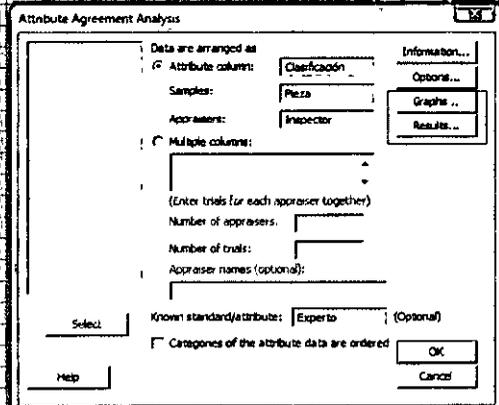
"Fase - Medir"

Módulo III

Analicemos los datos en Minitab.

Realicemos el ejercicio con Minitab, sigue la ruta...

Indica la columna correcta.



Incluye la evaluación del experto.



Estudio Método por Atributos.

¿Qué nos proporciona?

Within Appraisers

Assessment Agreement

Appraiser	# Inspected	# Matched	Percent	95% CI
Andres	3	2	66.67	(9.43, 99.16)
Carlos	3	3	100.00	(36.24, 100.00)
Hugo	3	3	100.00	(36.24, 100.00)
José	3	3	100.00	(36.24, 100.00)
Luis	3	3	100.00	(36.24, 100.00)
Paco	3	3	100.00	(36.24, 100.00)

Esto indica con qué frecuencia el evaluador da la respuesta correcta consigo mismo.

Each Appraiser vs Standard

Assessment Agreement

Appraiser	# Inspected	# Matched	Percent	95% CI
Andres	3	2	66.67	(9.43, 99.16)
Carlos	3	3	100.00	(36.24, 100.00)
Hugo	3	3	100.00	(36.24, 100.00)
José	3	3	100.00	(36.24, 100.00)
Luis	3	2	66.67	(9.43, 99.16)
Paco	3	3	100.00	(36.24, 100.00)

Esto indica con qué frecuencia el evaluador da la respuesta correcta con respecto al estándar.

Matched: Appraiser's assessment across trials agrees with the known standard.

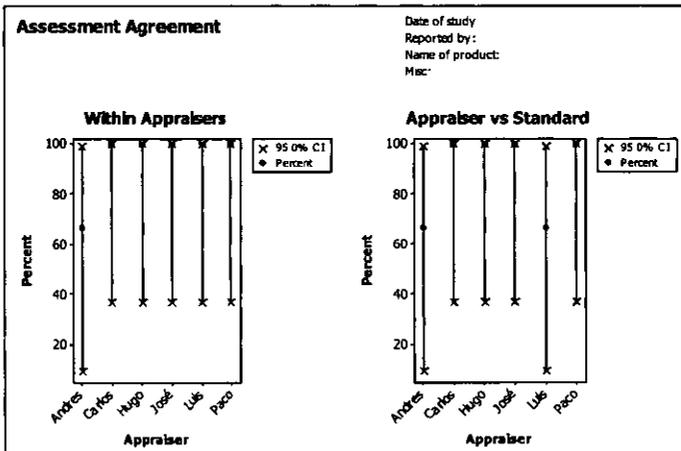
!!!Queremos el 90% ó más de acuerdos!!!



“Fase - Medir”

Módulo III

Análisis Gráfico.



Esta Gráfica muestra el % de consistencia del evaluador consigo mismo.

Esta Gráfica muestra el % de consistencia del evaluador vs. El experto.





"Fase - Medir"

Módulo III

Estudio para datos por Atributos.

- % del Evaluador es la consistencia de una persona.
- % Evaluador vs Atributo es el acuerdo entre la evaluación del operador y la del "experto".
-
- % de Efectividad de Selección es el acuerdo que existe entre los operadores.
- % de Efectividad de Selección vs. el Atributo medida general de la consistencia entre los operadores y el acuerdo con el "experto".

Estudio Método ANOVA.

¿Qué nos proporciona?

Between Appraisers

Assessment Agreement

# Inspected	# Matched	Percent	95% CI
3	2	66.67	(9.43, 99.16)

Matched: All appraisers' assessments agree with each other.

Fleiss' Kappa Statistics

Response	Kappa	SE Kappa	Z	F (vs > 0)
HP	0.719481	0.0710669	10.1240	0.0000
P	0.719481	0.0710669	10.1240	0.0000

All Appraisers vs Standard

Assessment Agreement

# Inspected	# Matched	Percent	95% CI
3	2	66.67	(9.43, 99.16)

Matched: All appraisers' assessments agree with the known standard

Indica cuántas veces dieron la respuesta correcta.

Esto indica la efectividad de los estándares de inspección.



Técnica Kappa.

- En un estudio por atributos, los números Kappa se usan para resumir el nivel de entendimiento entre evaluadores.
- Si hay un acuerdo substancial, existe la posibilidad de que las mediciones sean exactas.
 - Si el acuerdo es malo, la utilidad de los evaluadores es muy limitado.
- Requisitos para su uso:
 - Las unidades de medición son independientes de cada una.
 - Los evaluadores inspeccionan y clasifican de forma independiente.
 - Las evaluaciones de las categorías son mutuamente excluyentes.
- Minitab calcula el valor de Kappa como parte del reporte de un gage R&R para atributos.



Estudio Método ANOVA.

¿Qué nos proporciona?

Fleiss' Kappa Statistics

Response	Kappa	SE Kappa	Z	P(vs > 0)
NP	0.719481	0.0710669	10.1240	0.0000
P	0.719481	0.0710669	10.1240	0.0000

El valor general de Kappa nos representa el acuerdo del estudio.




Criterio de Validación del MSA Atributos.

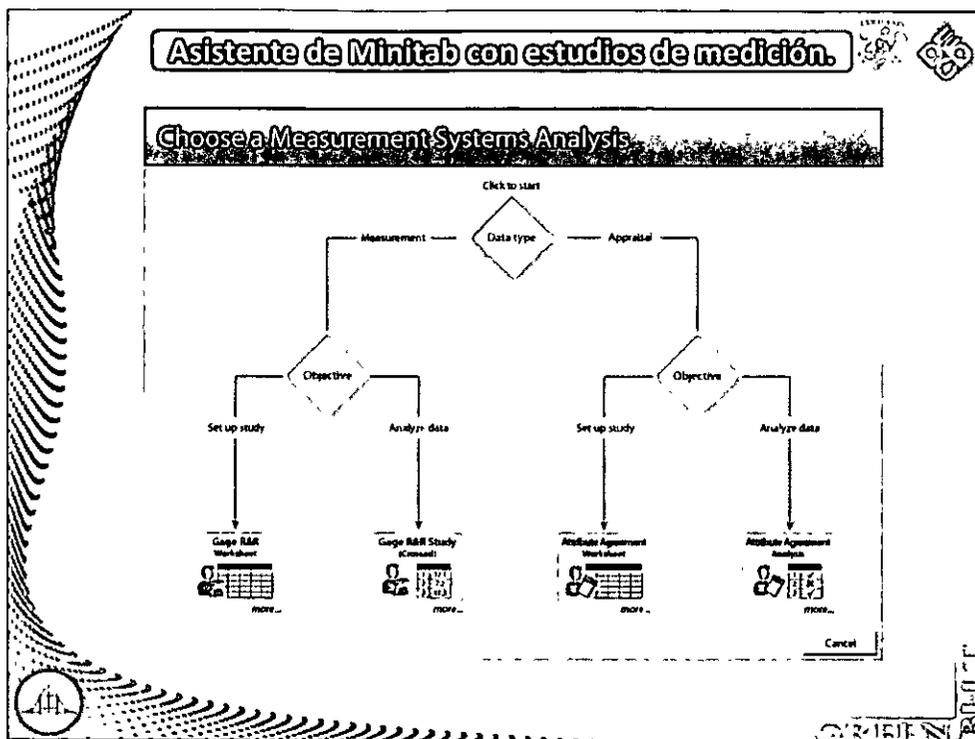
Estatus	Porcentaje
Inaceptable	Menor a 80%
Condicional	Entre 80% y 90%
Aceptable	Mayo a 90%

Kappa se mide entre -1 y 1

Estatus	Valor Kappa
Inaceptable	De 0 a 0.6
Condicional	Entre 0.6 a 0.7
Aceptable	Entre 0.7 a 0.8
Excelente	Mayor 0.9

Módulo III
"Fase - Medir"
Mejorando las Mediciones por atributos.

- Multiplicar los sentidos (dispositivos para estimular los sentidos).
- Mascaras / Plantillas.
- Check Lists.
- Automatización.
- 5 S's.
- Gestión Visual.



Módulo III

“Fase - Medir”

Establecimiento de la línea base.

Una vez validado nuestro Sistema de Medición podemos comenzar la recolección de datos, mientras más observes y realices el ejercicio tendrás mejor oportunidad en tus estudios.

Al realizar las mediciones siempre acompaña el proceso para detectar cualquier desviación.

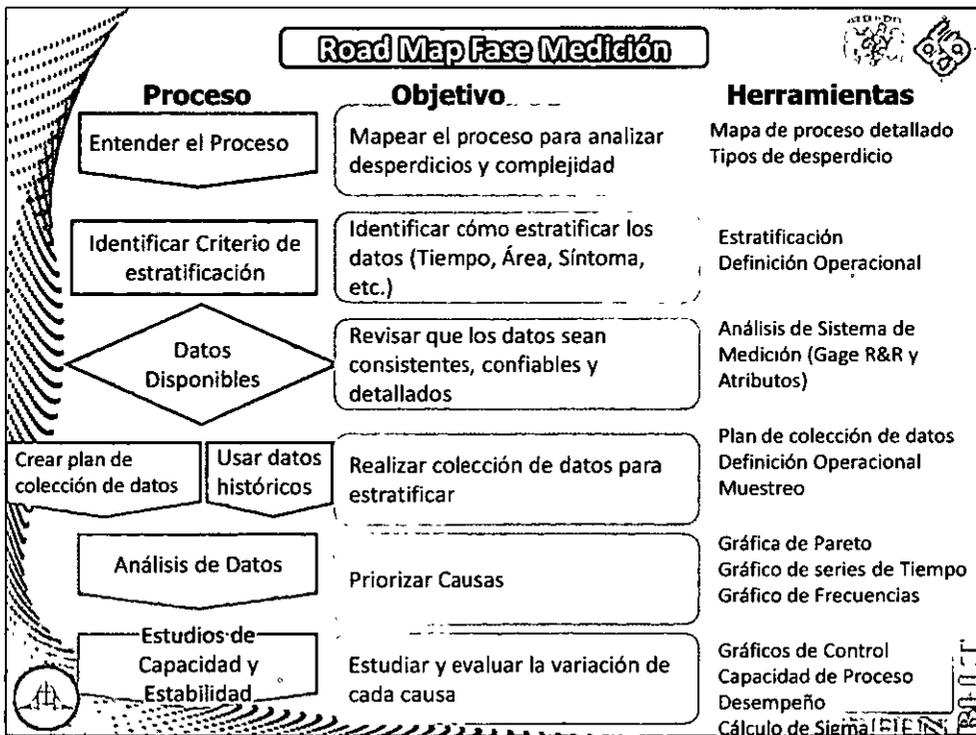


Establecimiento de la línea base.

“Mis Ojos no ven lo que mi cerebro no conoce...”

“Fase - Medir”

Módulo III





Capacidad de Proceso.

Definimos la capacidad como la habilidad de trabajar en un nivel esperado (cumplir con las especificaciones del cliente).

Límite Inferior de Especificación (LIE) Límite Superior de Especificación (LSE)

Este proceso es capaz de trabajar dentro de las especificaciones.

Capacidad de Proceso

Límite Inferior de Especificación (LIE) Límite Superior de Especificación (LSE)

¿Es capaz este proceso?
¿Qué es lo que cambió?

La capacidad se determina comparando la variación total del proceso contra la variación aceptada por el cliente.

Módulo III

"Fase - Medir"



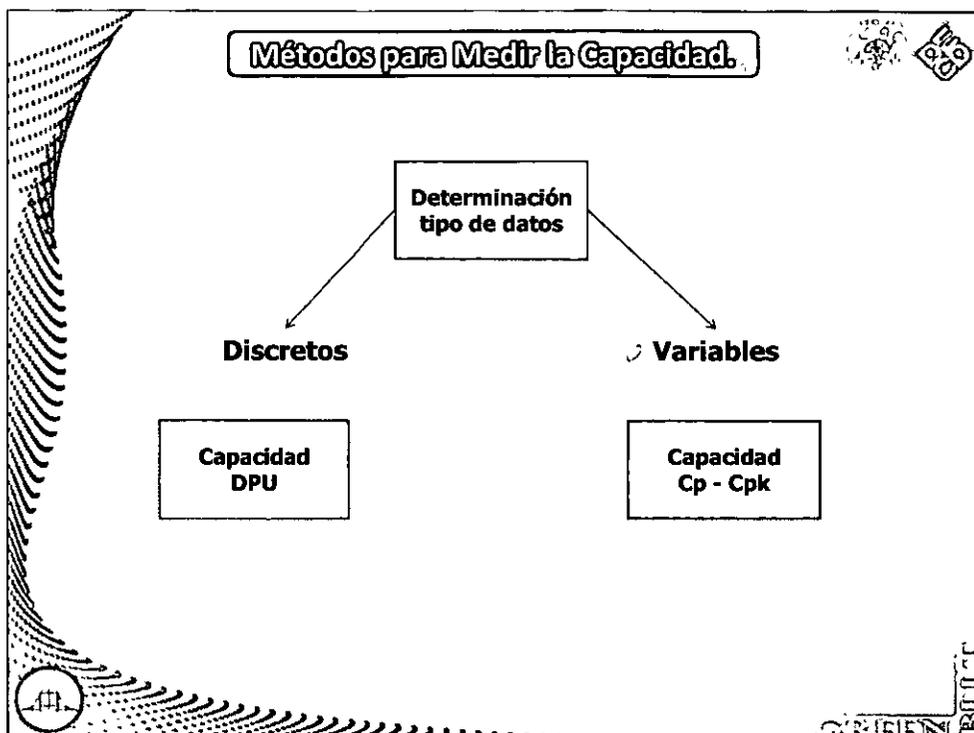
Utilización de la Capacidad de Proceso.

La capacidad del proceso nos sirve para:

- Cuantificar la naturaleza del problema del proyecto.
- Predecir los verdaderos niveles de calidad.
- Para estimar el nivel sigma inicial del proceso.

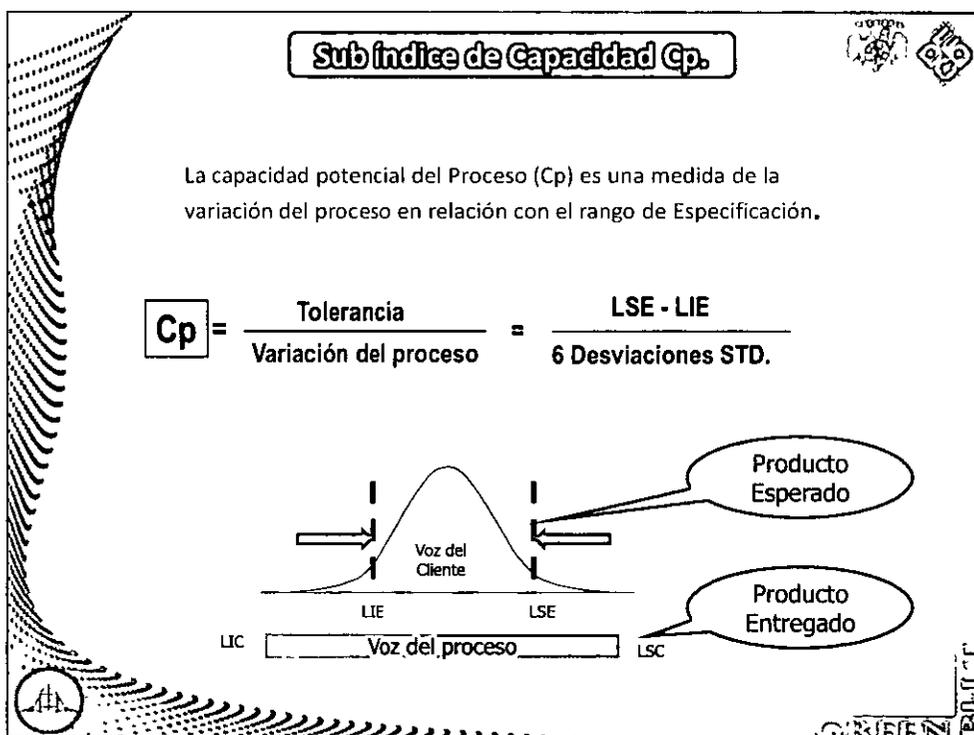
Capacidad de Proceso – Estrategia.

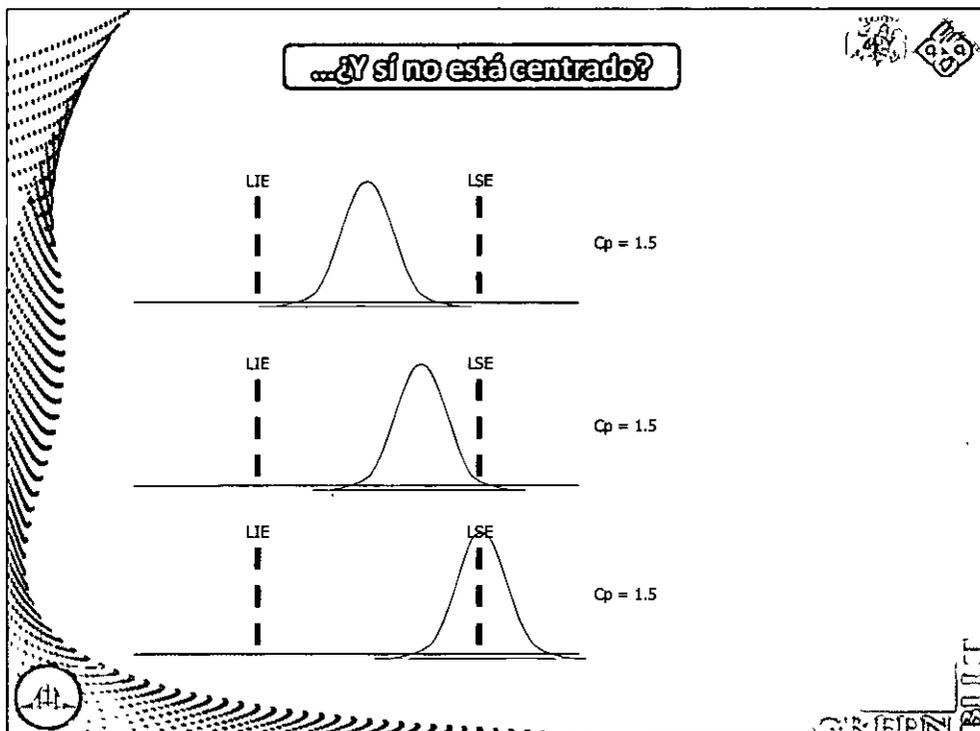
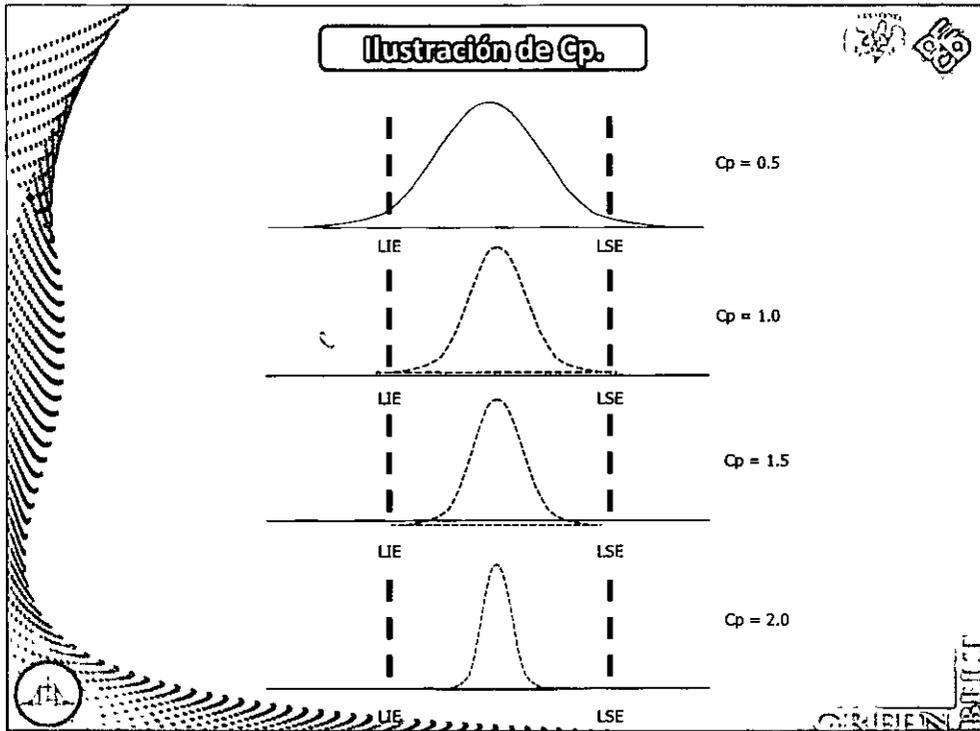
1. “Dispersión” – Reducir la variación del proceso.
2. “Centrado” – Poner el proceso en el objetivo.



Módulo III

"Fase - Medir"







Sub índice de Capacidad Cpk.

Cpk es una medida de la capacidad real del proceso en función de la posición de la media del proceso (X) en relación con los límites de especificación. Con límites bilaterales da una indicación del centrado. Es el menor de:

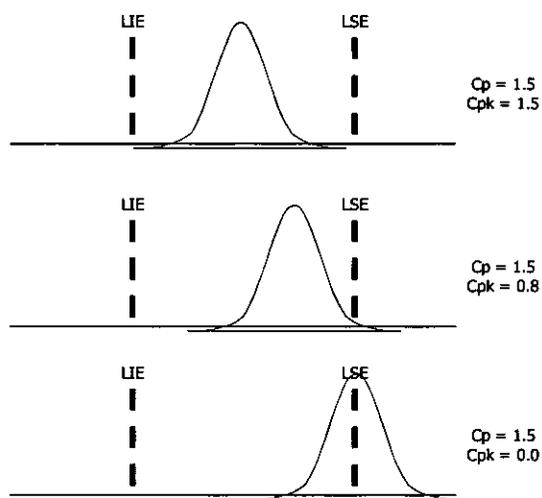
$$Cpk = \frac{LSE - \text{promedio del proceso}}{3 \text{ desviaciones STD}} \text{ y } \frac{\text{Promedio del proceso} - LIE}{3 \text{ desviaciones STD}}$$

Este indicador penaliza si estás fuera del objetivo.

Módulo III

“Fase - Medir”

Comparando Cp y Cpk.





Capacidad de Proceso.

¿Cuándo podemos decir que nuestro proceso es capaz?



$Cpk < 1.0$ (menor)

$1 \leq Cpk < 1.66$ (mayor ó igual y menor)

$Cpk \geq 1.66$ (mayor ó igual que)

GREEN BELT

Corto Plazo vs Largo Plazo.

Pensemos en esto: imagina el inicio de una línea de producción, herramientas nuevas y calibradas, operadores entrenados, proveedores certificados y entregando a un buen nivel de calidad.

¿Cómo esperas que funcione este proceso?

¿Se podrá mantener el proceso siempre?



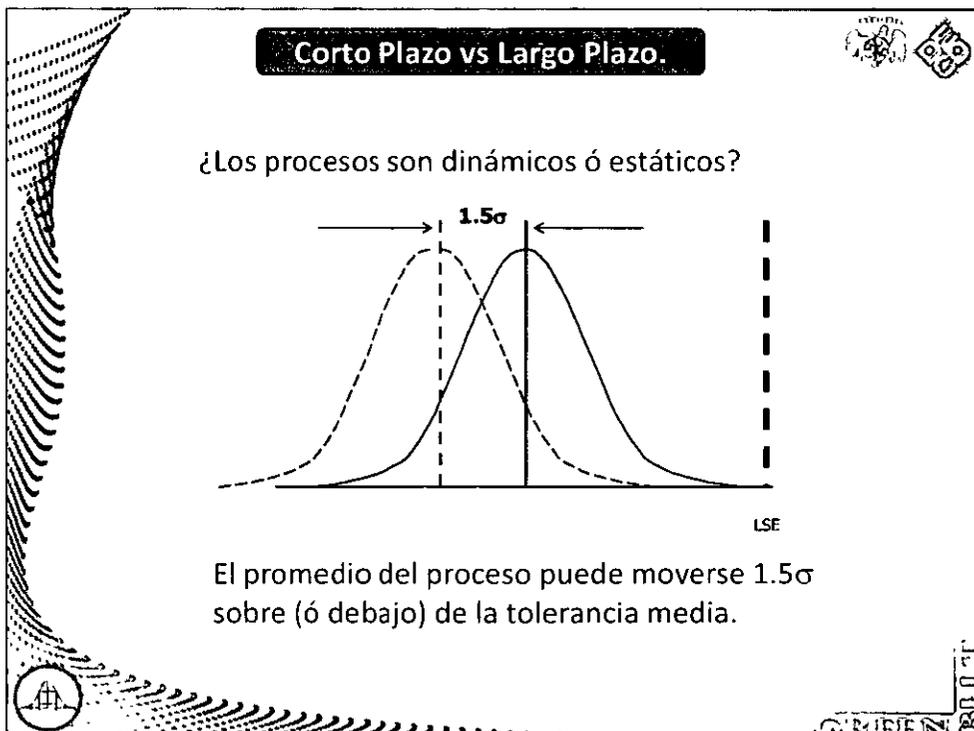
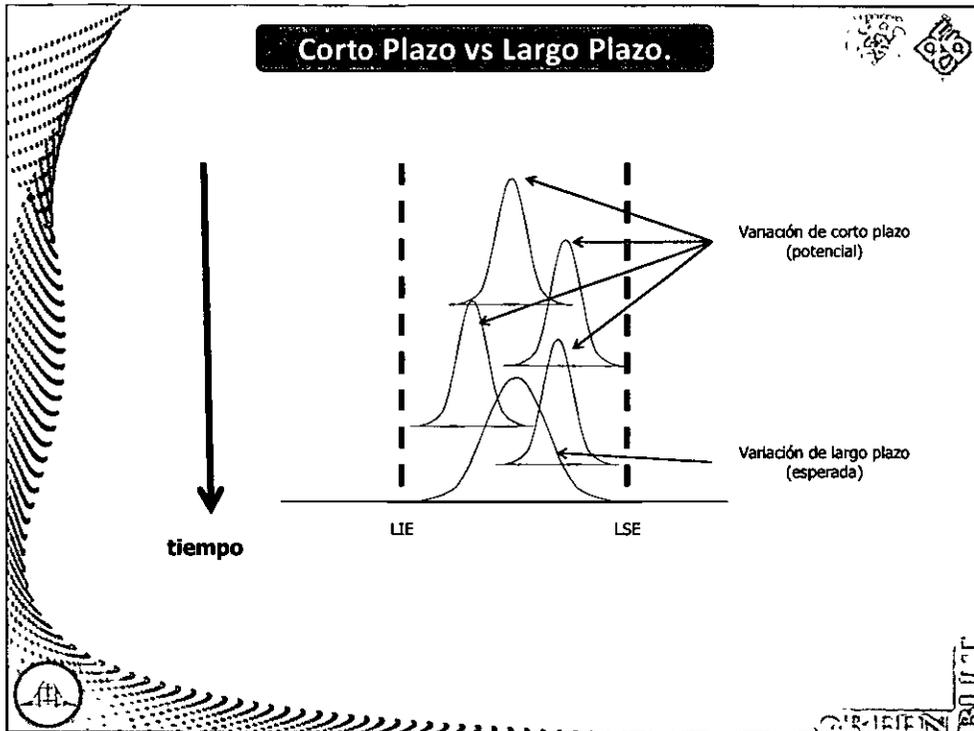


→

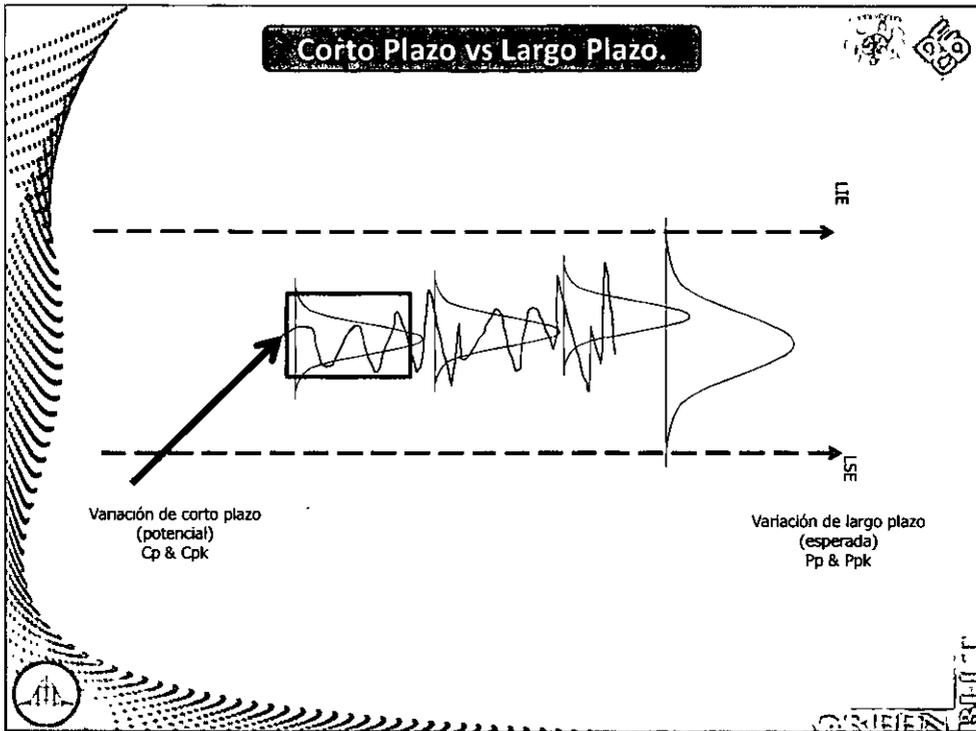
GREEN BELT



Módulo III



“Fase - Medir”



Índices de Habilidad Real y Potencial.

<div style="border: 1px solid black; padding: 5px; text-align: center; margin-bottom: 10px;"> Índices de Habilidad de Corto Plazo (Short Term) </div> <p> $Cp = (LSE - LIE) / 6 \sigma \text{ corto plazo}$ $Cpk = Z = \frac{LE - X}{3 \sigma_{st}}$ o $Cpk = \text{menor } Z_1, Z_2 / 3$ </p>	<div style="border: 1px solid black; padding: 5px; text-align: center; margin-bottom: 10px;"> Índices de Habilidad de Largo Plazo (Long Term) </div> <p> $Pp = (LSE - LIE) / 6 \sigma \text{ largo plazo}$ $Ppk = Z = \frac{LE - X}{3 \sigma_{lt}}$ o $Ppk = \text{menor } Z_1, Z_2 / 3$ </p>
---	--



La Distribución Normal.

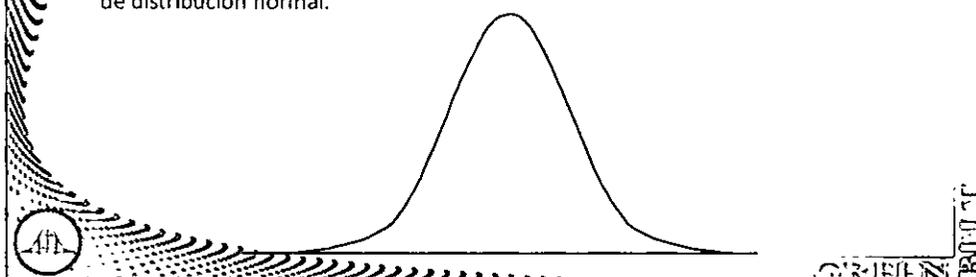
La distribución normal es una distribución de probabilidad que tiene media 0 y desviación estándar de 1.

El área bajo la curva o la probabilidad desde menos infinito a más infinito vale 1.

La distribución normal es simétrica, es decir cada mitad de curva tiene un área de 0.5.

La escala horizontal de la curva se mide en desviaciones estándar, su número se describe con Z.

Para cada valor Z se asigna una probabilidad o área bajo la curva mostrada en la Tabla de distribución normal.

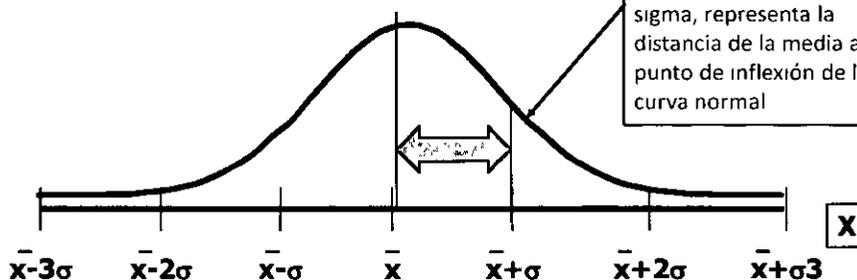


Módulo III

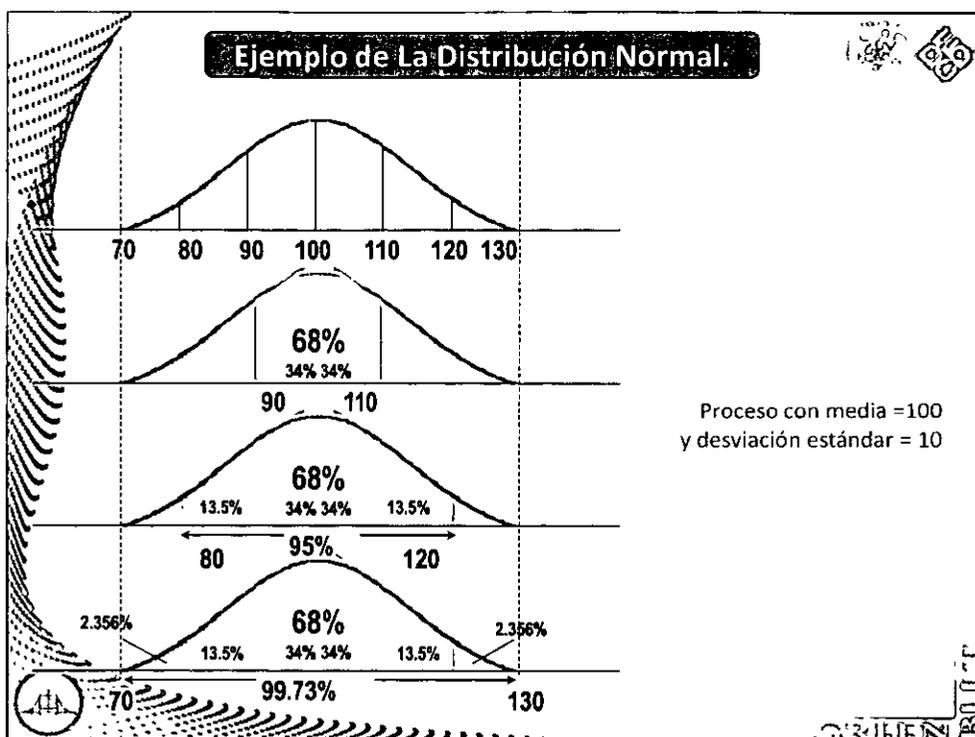
“Fase - Medir”

La Distribución Normal.

La desviación estándar ó sigma, representa la distancia de la media al punto de inflexión de la curva normal



X-Barra = Promedio
σ = desviación estándar



Módulo III

“Fase - Medir”

El Valor de Z.

Determina el número de desviaciones estándar entre algún valor x y la media de la población μ , Donde sigma es la desviación estándar de la población.

$$Z = \frac{X - \mu}{\sigma}$$



Módulo III

"Fase - Medir"

El Valor de Z como medida de capacidad.

La distancia entre un dato " X " y la media de una distribución en unidades de desviación estándar se denomina Z.

$$Z = \frac{X - \mu}{\sigma}$$

Área de Rendimiento

Límite de Rendimiento

Probabilidad de Defecto

$-\infty$ μ σ X $+\infty$

El Valor de Z como medida de capacidad.

$Z = 6$

$\pm 6\sigma$ Capacidad

$Z = 3$

$\pm 3\sigma$ Capacidad

LSE

LSE

Conforme disminuye la variación, disminuye la probabilidad de un defecto para un límite de especificación dado, de tal manera que la capacidad aumenta.



El Valor de Z como medida de capacidad.

**** Área bajo la curva = 1, el objetivo es 0 ****

Supongamos que calculamos la desviación normal estándar para un límite de rendimiento dado y encontramos que $Z = 2.76$. La pregunta sería, "¿Qué porción del área total, bajo la curva normal, se extiende más allá de un valor Z de 2.76?" Al contestar esta pregunta tendremos la probabilidad de producir un defecto. Recuerda que el valor Z es una medida de la capacidad de proceso y con frecuencia se le llama "la sigma del proceso", esto no debe confundirse con la desviación estándar del proceso.

Área Total = 1 Límite de Rendimiento Probabilidad de Defecto
Ejemplo = .00289

$X = 0$ Unidades de Medición

Z = 2.76

La tabla lista el área de la cola hacia la derecha de Z.

Tabla de Áreas bajo la Curva Normal

Z	Área	Z	Área	Z	Área	Z	Área
0.0	0.5000	0.1	0.4602	0.2	0.4243	0.3	0.3944
0.4	0.3599	0.5	0.3090	0.6	0.2743	0.7	0.2420
0.8	0.2147	0.9	0.1772	1.0	0.1454	1.1	0.1171
1.2	0.0910	1.3	0.0675	1.4	0.0479	1.5	0.0322
1.6	0.0207	1.7	0.0125	1.8	0.0072	1.9	0.0044
2.0	0.0025	2.1	0.0015	2.2	0.0009	2.3	0.0005
2.4	0.0003	2.5	0.0002	2.6	0.0001	2.7	0.0000

GRUBBS BILT

"Fase - Medir"

Módulo III

...En Excel.

Z	Z
2.79	2.67
P	P
DISTR.NORM.ESTAND 0.997364598	0.99620744
DISTR.NORM.ESTAND.INV	-2.79
	-2.67

En Excel podemos calcular la probabilidad sin tablas, sólo debemos capturar los datos adecuadamente...

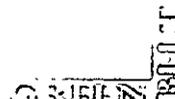
GRUBBS BILT



Conclusiones - Capacidad.

Definir la nueva capacidad del proceso, crea la base para un nuevo estándar para análisis futuros.

Comparar capacidades (antes vs. después) es una medición muy objetiva de éxito de las compañías de clase mundial. Mientras mayor sean los sigmas, menos defectos tendremos y como consecuencia, menor costo.



"Fase - Medir"

Módulo III

... ¿y si tenemos datos discretos?

Si tenemos datos por atributos (discretos) la pregunta es qué tan capaz es el proceso de no generar defectos, necesitamos definir algunos conceptos ..

Unidad : Facturas, llamadas, envíos, piezas.

Defecto: Cualquiera que no cumple los requerimientos del cliente.

DPU: Usado para medir el resultado de cuantificar defectos particulares por unidad inspeccionada.

Total de Defectos

Total de Unidades





ejemplo

Al recibir 70 facturas se encontró que 7 tenían al menos un defecto, ¿cuál es su DPU?

$$\frac{\text{Total de Defectos}}{\text{Total de Unidades}} = \frac{7}{70} = 0.10$$

¿... y si tenemos varios defectos en una factura?



Definamos algunos conceptos.

Una unidad *defectiva* es una unidad que tiene un *defecto*.

Puede haber varios *defectos* en una unidad *defectiva*.

La unidad defectiva es el resultado de los defectos, entonces:

Debemos enfocarnos en los defectos.

Trabajar en los defectivos normalmente nos lleva a retrabajos.

Ejemplo:

Una factura tiene 10 líneas potenciales de defectos, pero se contabiliza como un defectivo.



Del ejemplo anterior...

Al recibir 70 facturas se encontró que 7 tenían al menos un defecto, ¿cuál es su DPU?

$$\frac{\text{Total de Defectos}}{\text{Total de Unidades}} = \frac{7}{70} = 0.10 = \text{DPU}$$

Como se tienen 10 posibles defectos.

$$\frac{\text{DPU}}{\text{Número de Oportunidades}} = \frac{0.10}{10} \times 1,000,000 = 10,000 = \text{DPMO}$$

A esta unidad se le llama Defectos por Millón de Oportunidades (DPMO).

Módulo III

“Fase - Medir”

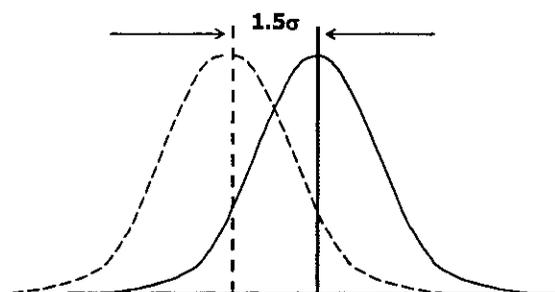
Tabla para calcular Nivel Sigma y DPMO.

1	Escribir el número de defectos posibles por unidad	O =		10
2	Escribir el número de unidades	N =		70
3	Escribir el número de defectos encontrados	D =		7
4.	Cálculo de DPU (DPO)	$DPO = \frac{D}{N \times O} =$		0.01000
5	Cálculo de Yield	$Yield = (1 - DPO) \times 100 =$		99.000%
6	Cálculo del Nivel Sigma del proceso	Sigma =		3.83

¿Recuerdas la tabla anterior de Z?



Recapitulando.



El promedio del proceso puede moverse 1.5σ sobre (ó debajo) de la tolerancia media.

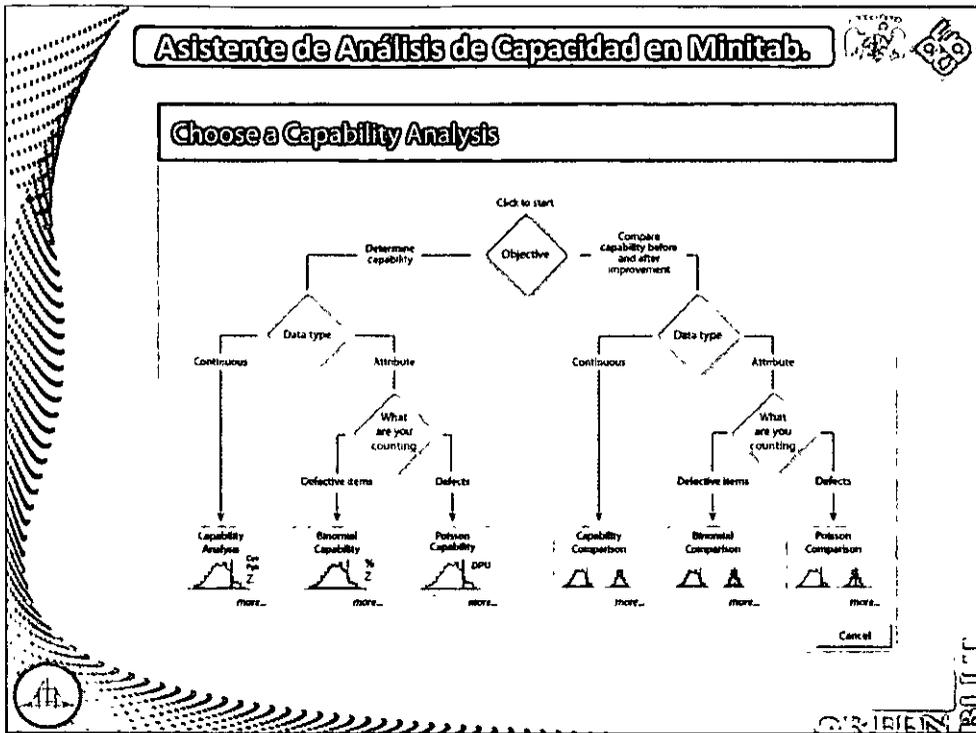
Módulo III

“Fase - Medir”

Cálculo del nivel sigma.

	Z	Z		
	2.79	2.67		
	P	P		
DISTR.NORM.ESTAND	0.997364598	0.99620744	0.0026354	0.00379256
DISTR.NORM.ESTAND.INV	-2.79	-2.67		
Nivel Sigma	-4.29	-4.17		

iiiEl nivel Sigma se calcula sumando el 1.5 del desplazamiento esperado a nuestra Z!!!



Módulo III

"Fase - Medir"

