

UNIDAD II

PRINCIPIOS ESTADÍSTICOS APLICADOS EN CONTROL DE CALIDAD

Por: Prof. Gastón A. Pérez U.

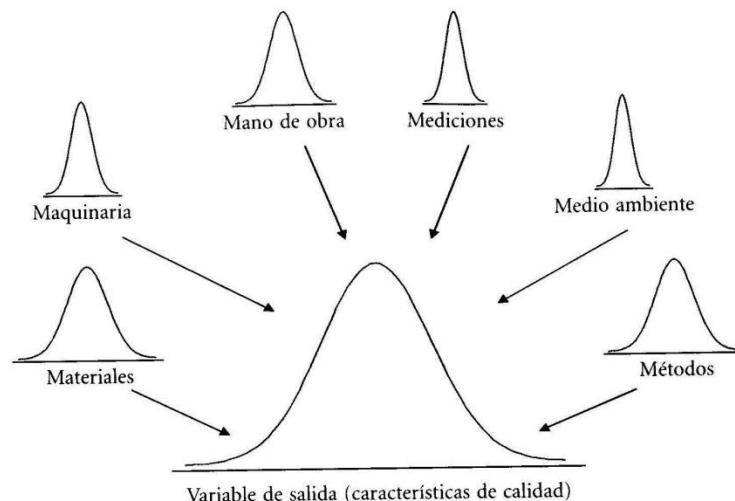
(2.1) VARIABLES ESTADÍSTICAS

(2.1.1) INTRODUCCIÓN

(a) LA VARIABILIDAD

- Cuando se desea mejorar un proceso a fin de elevar la calidad de un producto se hace necesario buscar información que pueda conducirnos al logro del objetivo; implica en consecuencia desarrollar un plan estratégico que nos conduzca a ese fin.
- Un plan de mejora de la calidad debe tomar en cuenta la variabilidad ¿Por qué?
- En un proceso interactúan factores tales como los materiales, maquinarias, mano de obra, mediciones, medio ambiente y métodos. Estos 6 elementos determinan de manera global todo el proceso, y cada uno aporta parte de la variabilidad (y de la calidad) de los resultados del proceso. Ver Fig. 5.2.

FIGURA 5.2
La variabilidad de un proceso. Cada M aporta una parte, no necesariamente igual, de la variación total observada.



- Algunos ejemplos de variabilidad son:
 - Materiales no idénticos
 - Personas con diferentes habilidades y entrenamiento.
 - Desajustes y desgastes de máquinas, etc.
- La variabilidad puede observarse incluso bajo condiciones de aparente estabilidad del proceso.
- La Norma ISO 9000 dice:
 - “El uso de técnicas estadísticas puede ser de ayuda para comprender la variabilidad y ayudar por lo tanto a las organizaciones a resolver problemas y a mejorar su eficacia y eficiencia. Asimismo estas técnicas facilitan una mejor utilización de los datos disponibles para ayudar en la toma de decisiones”
 - “Las técnicas estadísticas pueden ayudar a medir, describir, analizar, interpretar y hacer modelos de dicha variabilidad, incluso con una cantidad relativamente limitada de datos”
 - “El análisis estadístico de dichos datos puede ayudar a proporcionar un mejor entendimiento de la naturaleza, alcance y causas de la variabilidad, ayudando así a resolver e incluso prevenir los problemas que podrían derivarse de dicha variabilidad, y a promover la mejora continua”.
 - Estas técnicas son aplicables en un amplio espectro de actividades, tales como la investigación de mercado, el diseño, el desarrollo, la producción, la verificación, la instalación y el servicio.

- **ESTADÍSTICA** es la ciencia en la cual se recolectan, organizan, presentan, analizan e interpretan datos cuantitativos.
- En la Estadística hay dos fases:
 - ESTADÍSTICA DESCRIPTIVA: para describir y analizar un sujeto o un grupo.
 - ESTADÍSTICA INDUCTIVA (o inferencia estadística): es la que trata de determinar, partiendo de una cantidad limitada de datos (la muestra), una conclusión importante acerca de una cantidad mucho mayor de datos (población).

(b) VARIABLES DE UN PROCESO

En todo proceso hay variables de entrada y salida:

- **Variable de Entrada (X):** Son las que definen las condiciones de operación del proceso, y del valor de ellas depende la eficacia del proceso.
 - Ej. Temperatura, velocidad, presión, cantidad y/o características de algún insumo o material, etc.
- **Variable de Salida (Y):** Son las que reflejan los resultados obtenidos por el proceso.
 - A través de los valores que toman estas variables se evalúa la eficacia del proceso.
 - Por lo general son las características de la calidad del producto que se obtienen con el proceso.
- También pueden considerarse otras variables que no están normalmente controladas pero que influyen en el proceso, entre las cuales tenemos: humedad relativa del medio ambiente, habilidad de un operario, método de trabajo, etc.

(c) TIPOS DE VARIABLES

- **Cualitativas o de atributos:** Son aquellas en donde las características que se estudian no son numéricas.
Ej. Tipo de producto. Ensamblado/No ensamblado. Nombres de los clientes, etc.
- **Cuantitativas o continuas:** Son aquellas cuyas características pueden registrarse numéricamente.
Ej. Peso de un lote. Número de productos defectuosos. Número de clientes atendidos, etc.

Las continuas a su vez se clasifican en:

- **Discretas:** Son aquellas que pueden adquirir ciertos valores y casi siempre hay 'brechas' entre esos valores.
 - Ej.- No. de Clientes atendidos. No. de quejas. No. de artículos defectuosos por lote. No. de servicios de mantenimiento.
- **Continuas:** Son aquellas que pueden tomar cualquier valor dentro de un intervalo específico.
 - Las continuas son las que requieren de un instrumento de precisión para medir, ejemplo: peso, volumen, voltaje, longitud, resistencia, temperatura, humedad, tiempo, diversas dimensiones, etc.
 - Ej. Tiempo en que un cliente es atendido.

¿Cómo podemos cuantificar estas variables?

Debemos recurrir a los Datos:

- Los datos que se recolectan para fines de calidad se obtienen por observación directa.
- Por lo general ofrecen una información acerca de cómo está resultando el proceso en un momento determinado.
- Reflejan además una tendencia donde pueden medirse valores que reflejen una variabilidad que conduzcan a niveles de excelencia en calidad; pero también pueden ofrecer información donde la variabilidad no sea precisamente la de los estándares para el proceso.

(d) DEFINICIÓN DE POBLACIÓN Y MUESTRA

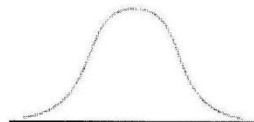
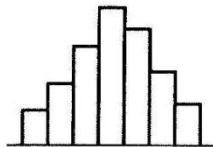
MUESTRA: Es un número limitado de artículos tomados de una fuente más grande.

POBLACIÓN: Es una fuente grande de artículos de donde se toma una muestra.

- Debemos considerar que cuando el volumen de datos es inmenso, se necesitan medios para resumirlos; verificar en qué valor o valores tienden a concentrarse y como se dispersan o extienden dichos datos.
- En Estadística existen métodos que permiten hacer cálculos de promedios, desviación estándar y otros, a partir de las muestras obtenidas y cuyos resultados se llaman estadísticos.
- Un **ESTADÍSTICO** es una cantidad calculada a partir de una muestra para estimar el parámetro de una población.
- Cuando se calculan estadísticos sobre toda la población los resultados se llaman parámetros.
- Se considera que los parámetros son valores fijos (estándares o patrones) de referencia.
- Como rara vez es posible medir toda la población, se selecciona una muestra, de manera que será necesario efectuar esta operación cuando sea imposible medir toda la población.
- Cuando se habla de promedios (\bar{X}) en una muestra en la población se habla de media y se denota con la letra griega “ μ ” (*miu*) o también como X_o = valor de referencia.
- Cuando se habla de desviación estándar (s) en una muestra, en la población se habla de desviación estándar poblacional y se denota con la palabra griega “ σ ” (sigma) o también como S_o =valor de referencia.
- La tabla 4.11 nos muestra la diferencia entre muestra y población.

TABLA 4-11 Comparación de muestra y población

MUESTRA	POBLACIÓN
Estadístico	Parámetro
\bar{X} —promedio	$\mu(\bar{X}_o)$ —media
s—desviación estándar muestral	$\sigma(s_o)$ —desviación estándar



(2.1.2) TÉCNICAS ESTADÍSTICAS

Las Técnicas Estadísticas son utilizadas como herramientas de análisis para el control de calidad.

La Norma ISO 10017 recomienda técnicas estadísticas específicas para cada apartado de la Norma ISO 9001. Ésta a su vez establece los requisitos que debe reunir un sistema de gestión de la calidad.

En general la norma recomienda las siguientes técnicas estadísticas:

- Estadísticas descriptivas
- Diseño de experimentos
- Prueba de hipótesis
- Análisis de la medición
- Análisis de la capacidad del proceso
- Análisis de regresión
- Análisis de confiabilidad
- Muestreo
- Simulación
- Gráficos de control estadístico de procesos (CEP)
- Fijación de tolerancias estadísticas

I. Análisis de series de tiempo

En este curso vamos a enfocarnos en las técnicas estadísticas que se recomiendan aplicar para un proceso donde se elabore o preste un producto o servicio. En ese sentido y de acuerdo al programa de esta asignatura vamos a concentrarnos en las siguientes técnicas que recomienda la norma así como también en otras no incluidas: Estadística descriptiva (Unidad II). Muestreo (Unidad III) y Gráficos de control estadístico de procesos (Unidad IV).

Estudiaremos en el presente capítulo una técnica de la Estadística Descriptiva llamada "Histograma" que describe la distribución de los valores de una característica de interés; además, conoceremos técnicas de análisis que se siguen cuando los histogramas reflejan problemas de calidad como por ejemplo el diagrama de Pareto.

Dice la Norma ISO 10017:

- *El término "estadística descriptiva" se refiere a procedimientos para resumir y presentar datos cuantitativos de manera que revele las características de la distribución de los datos.*
- *Las características de los datos que normalmente son de interés son su valor central (frecuentemente descrito por el promedio), y la desviación o la dispersión (normalmente medida por el rango, la desviación estándar o la varianza).*
- *Otra característica de interés es la distribución de los datos, para la cual existen medidas cuantitativas que describen la forma de la distribución (tal como el grado de "sesgo", que describe su simetría).*

(a) MEDIDAS DE TENDENCIA CENTRAL

Es un valor numérico que describe la posición central de los datos, o la forma en que los datos se tienden a acumular en el centro.

- Hay 3 medidas de uso común: Promedio. Mediana y Moda.

PROMEDIO (\bar{X}):

Es la suma de las observaciones dividida entre la cantidad de observaciones.

$$\bar{X} = \frac{\sum_{i=1}^n X_i}{n}$$

Dónde:

\bar{X} =promedio

n = # de variables observadas

X_1, X_2, \dots, X_n = Valor observado

Σ = símbolo que indica "suma de".

Ejemplo1- Los sueldos mensuales de los trabajadores en cierto departamento son los siguientes: 1100, 1300, 1000, 1500, 1600, 1100. ¿Calcule el sueldo promedio mensual del dpto.?

$$X = (1100+1300+1000+1500+1600+1100)/7 = 8400/7 = 1200$$

Nota: Si estos son todos los elementos de la población, entonces el promedio calculado es la media poblacional y se denota con la letra griega μ (mu).

MEDIANA (\tilde{x}):

- Es el valor que divide una serie de observaciones ordenadas de tal manera que la cantidad de elementos arriba de ella es igual a la cantidad de elementos debajo de ella.
- Cuando la serie es:
 - ❖ Impar: la mediana es el punto medio de los valores.
 - ❖ Par: la mediana es el promedio de los dos números a la mitad.

Ejemplo 2:

Utilizando los datos del ejemplo 1:

-Ordenamos los salarios de menor a mayor:

800, 1000, 1100, 1100, 1300, 1500, 1600

-Como la serie es impar, 1100 tiene 50% datos a la izq. Y 50% de datos a la der.

-Por lo que la mediana del ejemplo 1 es 1100.

-Suponiendo que no hubiese el sueldo de 800 ¿Cuál sería la mediana?

-Los 2 salarios centrales serían: 1100 y 1300, por lo que la mediana es:

MODA (Mo):

- Es el valor que se presenta con la máxima frecuencia.
- También es posible que no haya moda o que exista más de una moda.

-Utilizando los datos del ejemplo 1:

1100 es la Moda ya que está repetido 2 veces.

-Ej. 3, 3, 4, 5, 5, 5 y 7 → Moda = 5

-Ej. 105, 105, 105, 107, 108, 109, 109, 109, 110 y 112 → Moda = 105 y 109 (3 veces c/u)

Clasificación:

1. 1 Moda = Unimodal.
2. 2 Modas = Bimodal.
3. Más de 2 = Multimodal.

- La moda se usa principalmente como método de inspección, para determinar la tendencia central.

(b) MEDIDAS DE DISPERSIÓN

- Son las que describen la forma en que los datos se extienden o dispersan a cada lado del valor central.

- Las más utilizadas en Control de calidad son:
 1. Rango.
 2. Desviación estándar.
 3. Varianza.

RANGO:

- El rango (recorrido o intervalo) de una serie de números es la diferencia entre los valores u observaciones mayor y menor.

$$R = X_h - X_l$$

Dónde:

R = Rango

X_h = observación mayor en una serie

X_l = observación menor en una serie.

Ejemplo 3- En un departamento de ensamblaje el salario máximo es de 500 Bs. F y el salario mínimo es de 300 Bs. F ¿Determine el rango?

$$R = 500 - 300 = 200 \text{ Bs. F}$$

DESVIACIÓN ESTÁNDAR:

- Mide la tendencia de los datos a la dispersión o variabilidad.
- La desviación estándar esta expresada en las mismas unidades de medición que los datos muestrales (gr- cm- etc.)

$$S = \sqrt{\frac{\sum_{i=1}^n (X_i - \bar{X})^2}{n - 1}}$$

Dónde:

S = Desviación estándar de la muestra

X_i = Valor observado

\bar{X} = promedio

n = # de valores observados.

Ejemplo 4: Sea la siguiente tabla donde X_i representa a valores observados en kilogramos.

	X_i	$X_i - \bar{X}$	$(X_i - \bar{X})^2$
1	3.2	3.2-3.0= +0.2	0.04
2	2.9	-0.1	0.01
3	3.0	0.0	0.00
4	2.9	-0.1	0.01
5	3.1	+0.1	0.01
6	2.9	-0.1	0.01
	$\bar{X}=3.0$	$\Sigma=0$	$\Sigma= 0.08$

$$S = \sqrt{\frac{0.08}{6-1}} = 0.13 \text{ Kg.}$$

- Una desviación estándar grande indica mayor variabilidad que una pequeña.

VARIANZA

- Es el cuadrado de la desviación estándar.
- Se denota por $\rightarrow S^2$

$$S^2 = 0.13^2 = 0.016 \text{ Kg}^2$$

- Como medida de dispersión este dato no es adecuado, pero sí como medida de variabilidad en estadística superior.

(c) RELACIÓN ENTE LAS MEDIDAS DE TENDENCIA CENTRAL

PROMEDIO (valor central):

- Es la medida que más se utiliza.
- Se utiliza cuando:
 - Cuando la distribución es simétrica
 - Cuando no es marcadamente asimétrica hacia la derecha o la izquierda.
 - Cuando se calculan más estadísticos, como por ejemplo: medidas de dispersión, graficas de control, etc.
 - Contar con un valor estable en estadística inductiva.
 - Se recomienda descartar todos aquellos valores atípicos que se encuentran en la muestra cuando sea necesario determinar medidas de tendencia central.

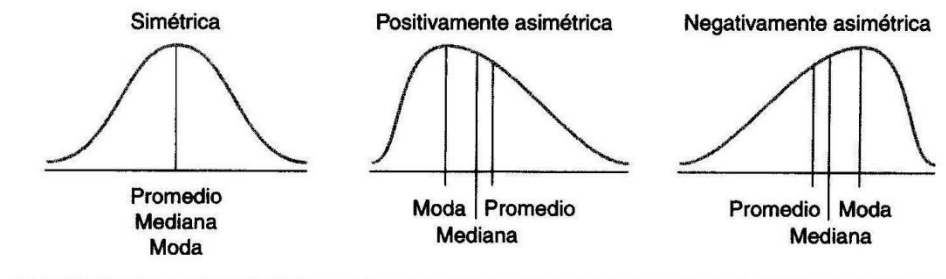


FIGURA 4-9 Relación entre promedio, mediana y moda.

- **MEDIANA:**
 - Se utiliza cuando la distribución es asimétrica positiva o negativa.
 - Determinar el punto medio exacto de una distribución.
 - Cuando la distribución tiene valores extremos (afecta al promedio pero no a la mediana).
- **MODA:**
 - Se usa cuando se desea tener una medida rápida y aproximada de la tendencia central.
 - Para describir el valor más característico de una distribución. Ej. La edad de esta sección.

(d) RELACIÓN ENTRE LAS MEDIDAS DE DISPERSIÓN

- **RANGO:**
 - Es muy común en la dispersión. Sus ventajas son:

- Proporciona el conocimiento de la extensión total de los datos y su simplicidad.
- Cuando la cantidad de datos es muy pequeña.
- Cuando los datos son demasiados dispersos para justificar el cálculo de una medida más precisa de la dispersión.
- o El rango no es una función de alguna medida de tendencia central.
- o A medida que aumenta la cantidad de observaciones, la exactitud del rango baja porque es más fácil que haya valores extremadamente altos o bajos.
- o Se sugiere limitar el uso del rango a un máximo de 10 observaciones.

• DESVIACIÓN ESTÁNDAR:

- o Es la medida más común de la dispersión.
- o Se usa cuando se desea una medida más precisa o se deben calcular otros estadísticos.
- o Cuando los datos tienen un valor extremo sea alto o bajo, es preferible la utilización de la desviación estándar al rango.
- o En la medida que la desviación estándar se hace pequeña, la calidad mejora.

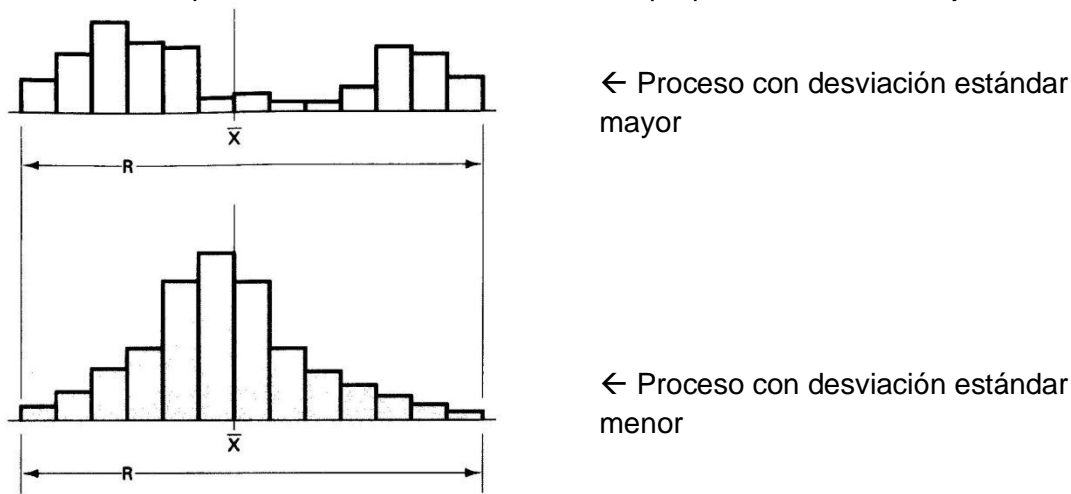


FIGURA 4-10 Comparación de dos distribuciones con igual promedio y rango.

- La grafica anterior muestra dos distribuciones con el mismo promedio (\bar{X}) y el mismo rango (R)
- La distribución inferior es mucho mejor ya que la mayoría de los datos tienden a agruparse en torno al centro.
- Además, la desviación estándar es mucho menor en comparación con la de la gráfica superior.

(e) RELACIÓN ENTRE MEDIA (\bar{X}) LA DESVIACIÓN ESTÁNDAR (S)

Existen 2 reglas que se utilizan:

- Desigualdad de Chebyshev (para cualquier tipo de comportamiento o distribución)

$$\bar{X} \pm 2S \rightarrow 75\%$$

$$\bar{X} \pm 3S \rightarrow 89\%$$

- Regla Empírica (para comportamiento similar a una campana o la distribución normal)

$$\bar{X} \pm S \rightarrow 68\%$$

$$\bar{X} \pm 2S \rightarrow 95\%$$

$$\bar{X} \pm 3S \rightarrow 99\%$$

- Los intervalos se aceptan como válidos para datos muestrales; sin embargo, si la media y la desviación se calculan a partir de la población, entonces los intervalos serán perfectamente válidos.
- Entre más grande sea la muestra aleatoria, los intervalos darán una idea aproximada de lo que ocurre en la población.
- Ambos casos ilustran muy bien la forma en que la desviación estándar mide la variabilidad en torno a la media.
- Analicemos el ejemplo 5:
- Hay que tener cuidado en la aplicación de la Regla empírica. Cuando las muestras sean pequeñas (<30) y no aleatorias, las conclusiones serán poco confiables respecto a toda la población.

(f) LÍMITES REALES O NATURALES

- Debido a las propiedades de la distribución natural, expresada por la regla empírica, ello ha dado origen a lo que se conoce como límites reales de un proceso.
- Sea μ = media estándar del proceso y σ = desviación estándar del proceso

$$\text{Límite real inferior (LRI)} = \mu - 3\sigma$$

$$\text{Límite real superior (LRS)} = \mu + 3\sigma$$
- Dentro de estos límites se ubicaran los valores de la variable de salida correspondiente.

Ejemplo 5: En una fábrica de piezas de asbesto una característica importante de calidad es el espesor o grosor de las láminas. Para cierto tipo de lámina, el espesor óptimo es de 5 mm, y se establece como discrepancia tolerable a 0.8 mm, ya que si la lámina tiene un grosor menor que 4.2 mm se considera demasiado delgada y no reunirá las condiciones de resistencia exigidas por el cliente. Pero, si la lámina tiene un grosor mayor que 5.8 mm, entonces se gastara demasiado material para su elaboración y elevarán los costos del fabricante. Por lo tanto, es de suma importancia fabricar las láminas con el grosor óptimo, y en el peor de los casos dentro de las tolerancias especificadas.

Se plantean las siguientes interrogantes:

1-¿Qué tipo de láminas en cuanto a grosor se están produciendo?

2-¿El grosor medio es adecuado?

3-¿La variabilidad del grosor es mucha o poca?

Para poder responder estas interrogantes, utilice la siguiente información: en la producción de una semana mediante muestreo sistemático^(*) se mide el espesor de 60 láminas y se obtiene las siguientes estadísticas:

$$\bar{X} \text{ (promedio)} = 4.73$$

$$X \text{ (mediana)} = 4.7$$

$$S \text{ (desviación estándar)} = 0.48$$

Solución:

1-Tipo de láminas producidas: grosor promedio no satisfactorio \rightarrow 4.73 mm (5 mm óptimo).

2-Grosor medio: la mediana indica que la mitad del lote de 60 láminas tiene un espesor de 4.7 o menos.

3-Para investigar la variabilidad y conocer si el espesor de las 60 láminas cayó dentro de la tolerancia (4.2mm-5.8mm), se aplica la Regla empírica para 99% ($X \pm 3S$):

$$4.73 + 3 \times 0.48 = 6.17 \text{ mm}$$

$$4.73 - 3 \times 0.48 = 3.29 \text{ mm}$$

$$\text{Variabilidad} \rightarrow (3.29 - 6.17)$$

Como puede observarse, hay problemas serios de fabricación ya que la tolerancia supera los límites establecidos.

En esta empresa no se están cumpliendo con los requerimientos de calidad.

Hay que trabajar primeramente en centrar el proceso (donde se hace crítico) y luego reducir la variabilidad.

(*) Nota: Muestreo sistemático es tomar las muestras al azar en intervalos fijos. El punto de partida se elige de manera aleatoria. El intervalo puede ser por ejemplo: cantidad de artículos, tiempo, longitud, área, etc.