

UNIVERSIDAD ABIERTA INTERAMERICANA
MAESTRIA EN ALTA DIRECCION DE EMPRESAS

CALIDAD Y PRODUCTIVIDAD



Profesor a cargo:

Docentes invitados:

Lic. Tomás Chahin

Lic. Marcelo Alcain

Ing. Manuel Pardo

Ing. Agr. Héctor Cappiello

CLASES 7, 8 y 9

Módulo 5

MODULO 5

El Control de la Calidad en las Operaciones



Temas de las Clases 7, 8 y 9

- Estadística aplicada a la Calidad
- Herramientas de la Calidad
- Ejercicios Prácticos

ESTADISTICA APLICADA A LA CALIDAD

- EL CONCEPTO DE CALIDAD
- HASTA 1930: solo inspección del producto final
- AÑOS 1930/1960: Control del proceso
- AÑOS 1960/1980: Competitividad y diseño
- AÑOS 1980/.....: Todas las actividades de la empresa se orientan hacia la calidad, ya sea del producto o del servicio, en todas las áreas de la organización.

LA ESTADISTICA Y LAS HERRAMIENTAS DE LA CALIDAD

- Las herramientas de la calidad
- Diagrama de flujo
- Histograma
- Diagrama de PARETO
- Hoja de inspección
- QFD – Despliegue de la función calidad
- Diagrama de Ishikawa
- Diagrama de afinidad
- Tormenta de ideas – Brainstorming
- Diagrama de árbol

INTRODUCCIÓN A LA ESTADÍSTICA

DEFINICIONES

- **Estadística:** Es una colección de métodos para planear experimentos, obtener datos y luego organizar, resumir, presentar, analizar, interpretar y llegar a conclusiones con base en esos datos.
- **Población:** Es la colección completa de todos los elementos que se van a estudiar
- **Muestra:** Es un subconjunto de elementos extraídos de una población
- Un **Parámetro** es una medición numérica que describe alguna característica de una población
- Una **Estadística/o** es una medición numérica que describe alguna característica de una muestra

DEFINICIONES

- **Estadística descriptiva:** Describe, analiza y representa un grupo de datos utilizando métodos numéricos y gráficos que resumen y presentan la información contenida en ellos.

Estadística inferencial: Apoyándose en el cálculo de probabilidades y a partir de datos muestrales, efectúa estimaciones, decisiones, predicciones u otras generalizaciones sobre un conjunto mayor de datos.

INTRODUCCIÓN A LA PROBABILIDAD

- Modelo
- Cálculo determinístico
- Cálculo probabilístico

INTRODUCCIÓN A LA PROBABILIDAD

- **Probabilidad de Laplace**

- Si un experimento cualquiera puede dar lugar a un número finito de resultados posibles, y no existe ninguna razón que privilegie unos resultados en contra de otros, se calcula la probabilidad de un suceso aleatorio **A**, según la **regla de Laplace** como el cociente entre el número de casos favorables a **A**, y el de todos los posibles resultados del experimento:

$$P[A] = \frac{\text{número de casos favorables a } A}{\text{número de casos posibles}}$$

- **Ejemplo**

- Calcular la probabilidad de que al lanzar un dado se obtenga un número impar.

- **Solución:**

- Vamos a llamar **A**, al suceso consistente en que el resultado es impar. Como no suponemos que ninguna de las caras ofrece una probabilidad de ocurrencia diferente a las demás, podemos aplicar la regla de Laplace para obtener que

-

$$\begin{aligned} P[A] &= \frac{\text{número de casos favorables a } A}{\text{número de casos posibles}} \\ &= \frac{\text{número de elementos en } A}{\text{número de elementos en } E} \\ &= \frac{\text{número de elementos en } \{1, 3, 5\}}{\text{número de elementos en } \{1, 2, 3, 4, 5, 6\}} \\ &= \frac{3}{6} = \frac{1}{2} \end{aligned}$$

DEFINICIONES

- Los **datos discretos** se obtienen de un número finito de posibles valores o bien de un número posible de valores que pueden contarse
- Los **datos contínuos** (numéricos) se obtienen de un número infinito de valores que pueden asociarse a puntos de una escala contínua, de tal manera que no haya huecos ni interrupciones
- *Cuando los datos representan conteos, son discretos; cuando representan mediciones, son contínuos*

DEFINICIONES

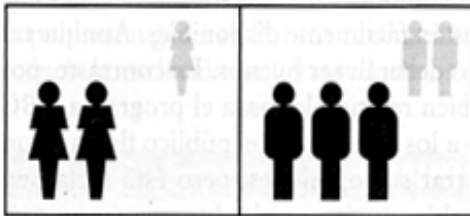
- **Muestra aleatoria** es aquella en la que seleccionamos miembros de la población de modo que cada uno tenga la misma probabilidad de ser escogido.
- **Muestreo estratificado** (una muestra de cada estrato compartiendo el mismo género)
- **Muestreo sistemático** (uno de cada enésimo)
- **Muestreo por cúmulos** (se divide en secciones y se elige en forma aleatoria)
- **Muestreo de conveniencia** (toma en cuenta resultados disponibles)

Métodos de Muestreo



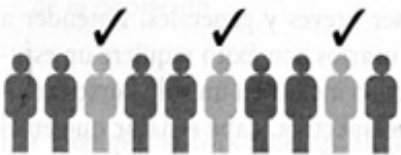
Muestreo aleatorio:

Cada miembro de la población tiene la misma probabilidad de ser seleccionado. A menudo se usan computadoras para generar números telefónicos al azar.



Muestreo estratificado:

Se clasifica la población en al menos dos estratos y se extrae una muestra de cada uno.



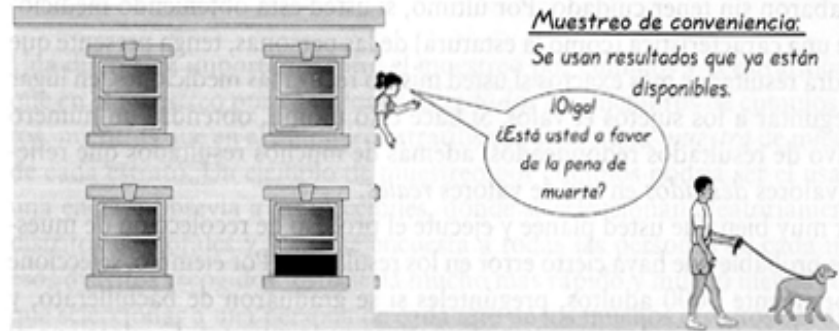
Muestreo sistemático:

Se selecciona cada k-ésimo miembro.



Muestreo por cúmulos:

Se divide el área de la población en secciones, se seleccionan al azar unas cuantas de esas secciones, y luego se escogen todos los miembros de ellas.



Muestreo de conveniencia:

Se usan resultados que ya están disponibles.

MEDIDAS DE TENDENCIA CENTRAL

LA MEDIA ARITMÉTICA

La media aritmética o promedio, de una cantidad finita de números, es igual a la suma de todos ellos dividida entre el número de sumandos. Simbólicamente se escribe como:

$$\bar{X} = \frac{X_1 + X_2 + X_3 + \dots + X_{n-1} + X_n}{n} = \frac{\sum_{i=1}^n X_i}{n}$$

MEDIA ARITMÉTICA PONDERADA

- **MEDIA ARITMÉTICA PONDERADA**

$$\bar{X}_w = \frac{\sum x_i w_i n_i}{\sum w_i n_i}$$

Ejemplo: Un estudiante realiza 3 exámenes de complejidad creciente, obteniendo los siguientes resultados: 5, 8 y 7.

El primer examen lo hizo en $\frac{1}{2}$ hora, el segundo en 1 hora y el tercero en hora y media, por lo que se les atribuye una ponderación de 1, 2 y 3 respectivamente. Se pide calcular la nota media.

EJEMPLOS

- Si calculamos la media aritmética tendremos que

$$\bar{X} = \frac{\sum x_i n_i}{n} = \frac{5 + 8 + 7}{3} = \underline{\underline{6,67}}$$

Ahora bien, si calculamos la media ponderada, obtendremos

$$\bar{X}_w = \frac{\sum x_i w_i n_i}{\sum w_i n_i}$$

$$\bar{X}_w = \frac{(5 \times 1) + (8 \times 2) + (7 \times 3)}{(1 + 2 + 3)} = \frac{5 + 16 + 21}{6} = \frac{42}{6} = \underline{\underline{7}}$$

MEDIA GEOMÉTRICA

- **Media geométrica G** : Responde a la siguiente expresión

$$G = \sqrt[n]{x_1^{n_1} x_2^{n_2} x_3^{n_2} \dots x_k^{n_k}}$$

y se la puede definir, como la raíz n-ésima del producto de todos los valores de la variable.

$$n = \sum_{i=1}^k n_i$$

MEDIA GEOMÉTRICA

- Supóngase que las utilidades obtenidas por una compañía constructora en cuatro proyectos fueron de 3, 2, 4 y 6%, respectivamente. ¿ Cual es la media geométrica de las ganancias?.

$$G = \sqrt[4]{3.2.4.6}$$

y el valor es 3.46

MEDIA ARMÓNICA

- **Media armónica. La representaremos como H:** Es la inversa de la media aritmética de las inversas de los valores de la variable, responde a la siguiente expresión:

$$H = \frac{n}{\sum \frac{n_i}{x_i}} = \frac{n}{\frac{n_1}{x_1} + \frac{n_2}{x_2} + \frac{n_3}{x_3} + \dots}$$

Entre la media aritmética la media geométrica y media armónica se da siempre la siguiente relación:

$$H \leq G \leq \bar{X}$$

MEDIDAS DE TENDENCIA CENTRAL

- **Mediana**: Punto medio de los valores después de ordenarlos de menor a mayor, o de mayor a menor. Se tiene que 50% de las observaciones se encuentran por arriba de la mediana y 50% por debajo de ella.
- **Las propiedades de la mediana son:**
 - Es única, sólo existe una mediana para un conjunto de datos.
 - No se ve afectada por valores muy grandes o muy pequeños.
 - Puede calcularse para una distribución de frecuencias con una clase de extremo abierto, si la mediana no se encuentra en una clase de tal extremo.
 - Puede obtenerse para datos de nivel de razón, de intervalo y ordinal (excepto para el nominal).

MEDIDAS DE TENDENCIA CENTRAL

- 10 29 26 28 15 DATOS
- 10 15 26 28 29 DATOS ORDENADOS
- CANT. DE DATOS 5 CIFRA IMPAR
- EL NÚMERO BUSCADO ESTÁ EXACTAMENTE EN LA MITAD
- LA MEDIANA ES 26

MEDIDAS DE TENDENCIA CENTRAL

- 500 600 800 50000 1000 500 DATOS
- 500 500 600 800 1000 50000 DATOS ORDENADOS
- CANT. PAR, SE TOMAN LOS DOS PUNTAJES QUE ESTÁN A LA MITAD. Se suman y se dividen por dos
- La mediana es 700

MEDIDAS DE TENDENCIA CENTRAL

- **Moda:** la moda de un conjunto de datos es el puntaje que ocurre con mayor frecuencia
- Ej. 5 5 5 3 1 5 1 4 3 5
- El número 5 es la moda pues es el que ocurre con mayor frecuencia

MEDIDAS DE DISPERSIÓN

- **Desviación absoluta media (DAM):**
- Equivale a la división de la sumatoria del valor absoluto de las distancias existentes entre cada dato y su media aritmética y el número total de datos.

MEDIDAS DE DISPERSIÓN

- **1.- Rango:** mide la amplitud de los valores de la muestra y se calcula por diferencia entre el valor más elevado y el valor más bajo.
- **2.- Varianza:** Mide la distancia existente entre los valores de la serie y la media. Se calcula como sumatoria de las diferencias al cuadrado entre cada valor y la media. La sumatoria obtenida se divide por el tamaño de la muestra.

$$s^2 = \frac{\sum (x_i - \bar{x})^2}{n - 1}$$

Donde:

x_i valores de la variable, x_1, x_2 , etc.

n es de la muestra

\bar{x} es la media aritmética

MEDIDAS DE DISPERSIÓN

- La varianza siempre será mayor que cero. Mientras más se aproxima a cero, más concentrados están los valores de la serie alrededor de la media. Por el contrario, mientras mayor sea la varianza, más dispersos están.
- **3.- Desviación típica:** Se calcula como raíz cuadrada de la varianza.
- **4.- Coeficiente de variación de Pearson:** se calcula como cociente entre la desviación típica y la media.

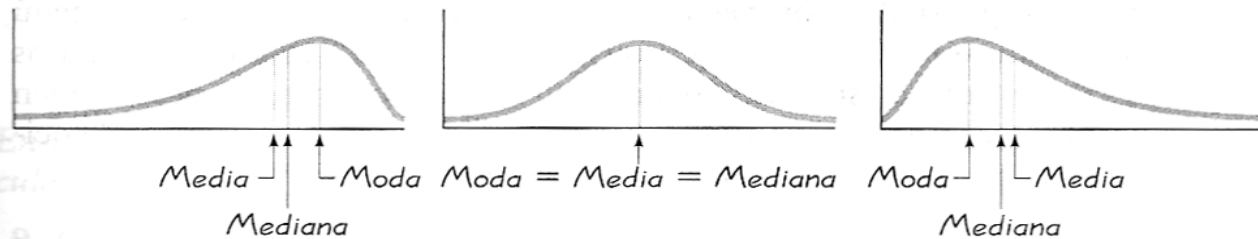
Coeficiente de variabilidad

Medida de variabilidad relativa: Se usa para comparar la variabilidad entre dos o más muestras medidas en las mismas unidades o no. Los datos que se expresan en porcentaje en la cual se compara la desviación estándar con el respectivo valor del promedio de los datos:

$$C.V. = \left(\frac{S}{\bar{x}} \right) \times 100$$

<i>Grado de variabilidad de los datos</i>	<i>Coeficiente de variabilidad</i>
Con variabilidad baja	Menos de 10%
Con variabilidad moderada	De 10% a 30%
Con alta variabilidad	Más de 30%

MEDIDAS DE FORMA



(a) **Sesgada a la izquierda** (sesgo negativo): La media y la mediana están a la *izquierda* de la moda.

(b) **Simétrica (sesgo cero)**: La media, la mediana y la moda son iguales.

(c) **Sesgada a la derecha** (sesgo positivo): La media y la mediana están a la *derecha* de la moda.

Definición

Una distribución de datos está **sesgada**, si no es simétrica y se extiende más hacia un lado que hacia otro. (Una distribución de datos es **simétrica**, si la mitad izquierda de su histograma es aproximadamente una imagen especular de su mitad derecha.)

MEDIDAS DE FORMA

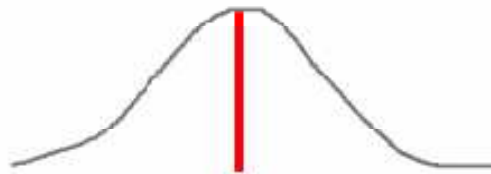
- Evalúa el grado de distorsión o inclinación que adopta la distribución de los datos respecto a su valor promedio tomado como centro de gravedad. El coeficiente de asimetría de Pearson es:

$$A_K = \frac{3(\bar{X} - M_e)}{S}$$

<i>Grado de Asimetría</i>	<i>Valor del Sesgo</i>
Simetría Perfecta	Cero. El promedio es igual a la mediana
Sesgo Positivo	Positivo. Promedio mayor que la mediana
Sesgo Negativo	Negativo. Promedio menor que mediana



Asimetría Positiva
(Promedio > Mediana)



Simétrica
Promedio = Mediana



Asimetría Negativa
Promedio < Mediana

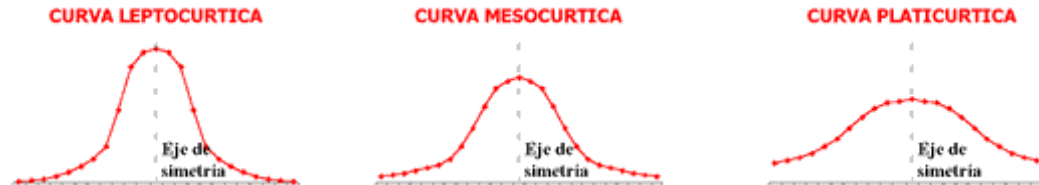
Curtosis

- El **Coefficiente de Curtosis** analiza el grado de concentración que presentan los valores alrededor de la zona central de la distribución.
- Se definen 3 tipos de distribuciones según su grado de curtosis:
- **Distribución mesocúrtica:** presenta un grado de concentración medio alrededor de los valores centrales de la variable (el mismo que presenta una distribución normal).
- **Distribución leptocúrtica:** presenta un elevado grado de concentración alrededor de los valores centrales de la variable.
- **Distribución platicúrtica:** presenta un reducido grado de concentración alrededor de los valores centrales de la variable.

Coeficiente de Curtosis

El coeficiente de curtosis está definido por:

$$\text{Coeficiente de Curtosis} = \frac{n \cdot \sum_{i=1}^n (x_i - \bar{x})^4}{\left(\sum_{i=1}^n (x_i - \bar{x})^2 \right)^2} - 3$$



Los resultados pueden ser los siguientes:

$g_2 = 0$ (distribución mesocúrtica).

$g_2 > 0$ (distribución leptocúrtica).

$g_2 < 0$ (distribución platicúrtica)

FRECUENCIA

- **Frecuencia absoluta.** Es el número de veces que aparece cualquier valor de la variable.
- **Frecuencia absoluta acumulada.** Es la suma de la frecuencia absoluta de un valor de la variable con todos los anteriores.
- **Frecuencia relativa.** Es el cociente entre la frecuencia absoluta y el número de datos (N). Al multiplicarla por 100 obtenemos el porcentaje de elementos que presentan esta característica.
- **Frecuencia relativa acumulada.** Es la suma de la frecuencia relativa de un valor de la variable con todos los anteriores. También se puede definir como el cociente entre la frecuencia absoluta acumulada y el número total de datos.

TABLA

CARGA	FRECUENCIA
• 200-209	9
• 210-219	3
• 220-229	5
• 230-239	4

CARACTERÍSTICAS DE LOS DATOS

- Límite de clase inferior 200,210,...
- Límite de clase superior 209,219...
- Frontera de clase 199,5;209,5...
- Marca de clase 204,5;214,5...
- Ancho de clase 10

EJERCICIO

- **Estudio de una variable cuantitativa discreta**
- Las notas de un examen de matemáticas de 30 alumnos de una clase son las siguientes:
- **5, 3, 4, 1, 2, 8, 9, 8, 7, 6, 6, 7, 9, 8, 7, 7, 1, 0, 1, 5, 9, 9, 8, 0, 8, 8, 8, 9, 5, 7.**
- a) Ordenar los datos y calcular las frecuencias absolutas de cada nota.
- b) Hacer un diagrama de barras de las frecuencias absolutas y dibujar el polígono de frecuencias.

EJERCICIO

a) Tabla para calcular las frecuencias.

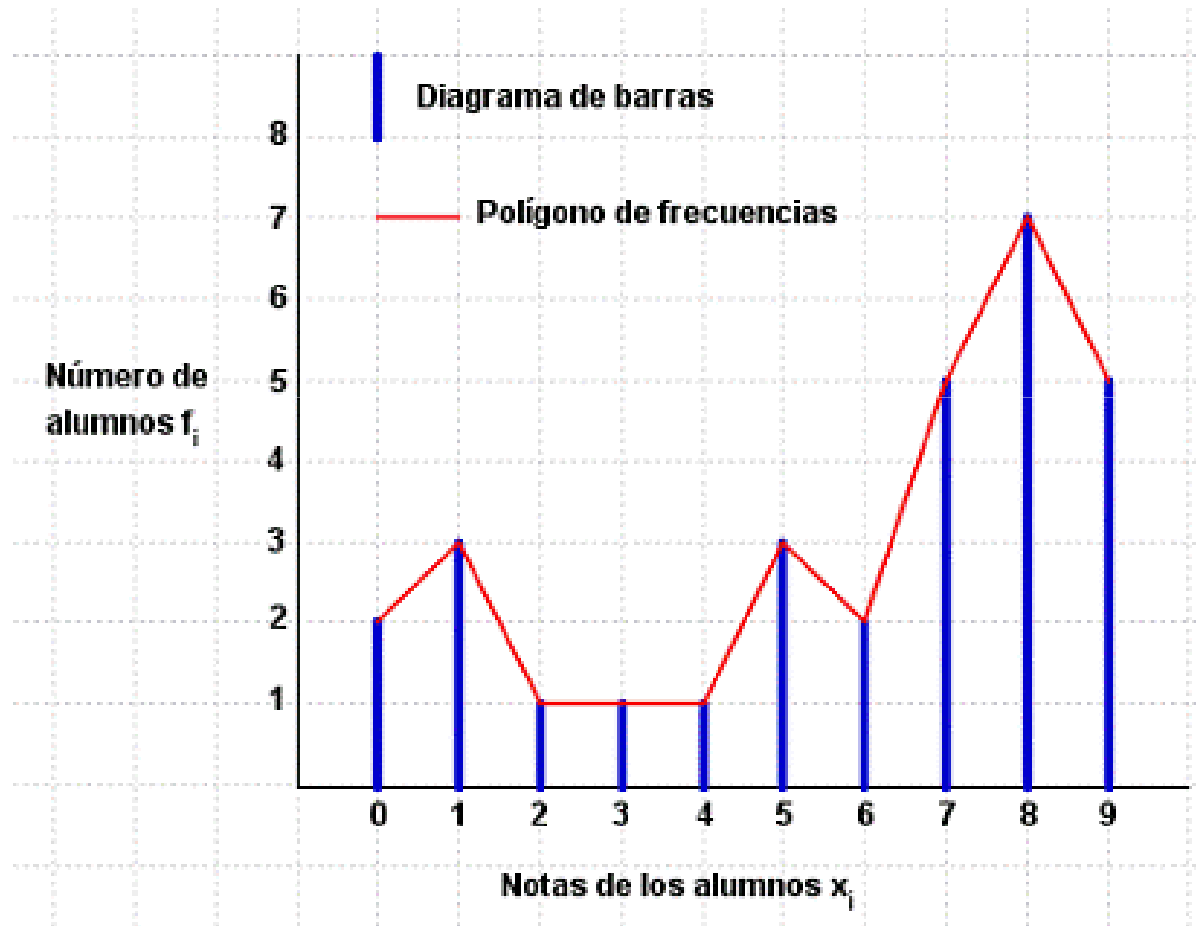
- Ordenamos los datos contando los alumnos que han sacado un **0** han sido **2**, un **1** han sido **3** y así sucesivamente. Construimos la tabla correspondiente:
- **N**: número total de datos $N = 30$.
- **xi**: variable estadística, nota del examen.
- **fi**: frecuencia absoluta, número de veces que se repite una nota. El sumatorio nos da los datos totales $N = 30$.
- **Fi**: frecuencia absoluta acumulada. Para calcularla vamos sumando los valores de la frecuencia absoluta f_i . $F_2 = f_1 + f_2 \Rightarrow 2 + 3 = 5$ $F_3 = F_2 + f_3 \Rightarrow 5 + 1 = 6$
- **hi**: frecuencia relativa. Cociente f_i / N
- **Hi**: frecuencia relativa acumulada
- Σ : sumatorio (suma de todos los datos de la columna correspondiente)

TABLA

Cálculo de frecuencias

x_i	f_i	F_i	$h_i = f_i / N$	H_i
0	2	2	2 / 30	2 / 30
1	3	5	3 / 30	5 / 30
2	1	6	1 / 30	6 / 30
3	1	7	1 / 30	7 / 30
4	1	8	1 / 30	8 / 30
5	3	11	3 / 30	11 / 30
6	2	13	2 / 30	13 / 30
7	5	18	5 / 30	18 / 30
8	7	25	7 / 30	25 / 30
9	5	30	5 / 30	30 / 30
Σ	30		1	

GRÁFICA



Estudio de una variable cuantitativa continua

Se ha controlado el peso de 50 recién nacidos, obteniéndose los siguientes resultados:

Peso (en kg)	Número de niños
[2,5 - 3)	6
[3 - 3,5)	23
[3,5 - 4)	12
[4 - 4,5)	9

- a) Formar la tabla de frecuencias.
- b) Representar gráficamente la distribución.

TABLA DE FRECUENCIAS

a) Formamos la tabla de frecuencias

Peso (en kg)	Número de niños f_i	F_i	h_i	H_i
[2,5 - 3)	6	6	0,120	0,120
[3 - 3,5)	23	29	0,460	0,580
[3,5 - 4)	12	41	0,240	0,820
[4 - 4,5)	9	50	0,180	1
Σ	50		1	

HISTOGRAMA

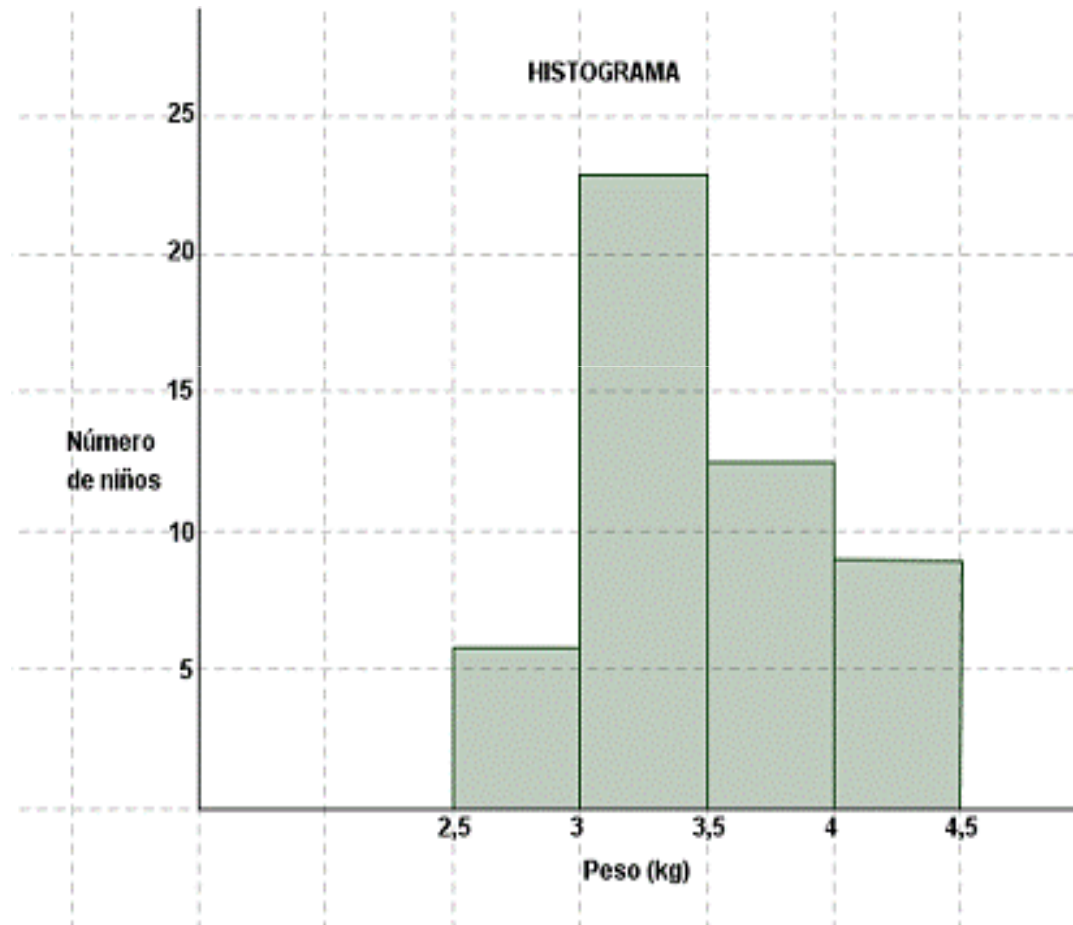
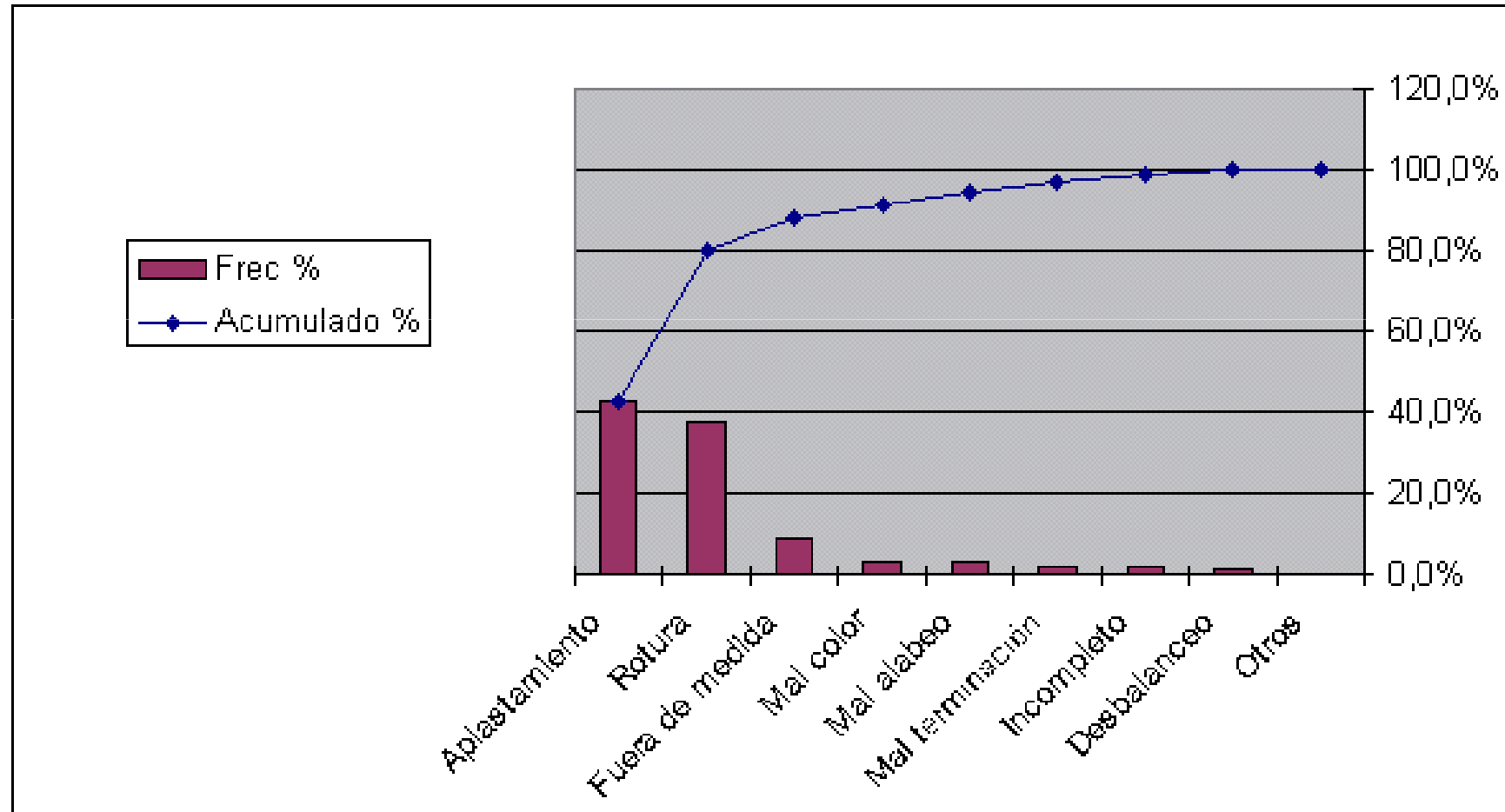


DIAGRAMA DE PARETO

- El **diagrama de Pareto**, también llamado **curva 80-20** o **Distribución A-B-C**, es una gráfica para organizar datos de forma que estos queden en orden descendente, de izquierda a derecha y separados por barras. Permite, pues, asignar un orden de prioridades.
- El diagrama permite mostrar gráficamente el [principio de Pareto](#) (pocos vitales, muchos triviales), es decir, que hay muchos problemas sin importancia frente a unos pocos graves. Mediante la gráfica colocamos los "pocos vitales" a la izquierda y los "muchos triviales" a la derecha.

DIAGRAMA DE PARETO



CLAVES PARA EL DIAGRAMA DE PARETO

- **1.** Determinar el problema o efecto a estudiar.
- **2.** Investigar los factores o causas que provocan ese problema y como recoger los datos referentes a ellos.
- **3.** Anotar la magnitud (por ejemplo: pesos, número de defectos, etc.) de cada factor. En el caso de factores cuya magnitud es muy pequeña comparada con la de los otros factores incluirlos dentro de la categoría “Otros”.

CLAVES PARA EL DIAGRAMA DE PARETO

- 4. Ordenar los factores de mayor a menor en función de la magnitud de cada uno de ellos.
- 5. Calcular la magnitud total del conjunto de factores.
- 6. Calcular el **porcentaje total** que representa cada factor, así como el **porcentaje acumulado**.
- El primero de ellos se calcula como:
- $\% = (magnitud\ del\ factor / magnitud\ total\ de\ los\ factores) \times 100$
- El porcentaje acumulado para cada uno de los factores se obtiene sumando los porcentajes de los factores anteriores de la lista más el porcentaje del propio factor del que se trate.

TIPOS DE MUESTREO

Planes por variables:

- En este tipo de planes se toma una muestra aleatoria del lote y a cada unidad de la muestra se le mide una característica de calidad aleatoria del lote (peso, longitud, etc.). Con las mediciones se calcula un estadístico, que generalmente está en función de la media y la desviación estándar muestral, y dependiendo del valor de este estadístico al compararlo con un valor permisible, se aceptará o rechazará todo el lote.

TIPOS DE MUESTREO

Planes por atributos:

- En estos planes se extrae aleatoriamente una muestra de un lote, y cada pieza de la muestra es clasificada de acuerdo con ciertos atributos como aceptable o defectuosa. Si el número de piezas defectuosas es menor o igual que un cierto número predefinido, entonces el lote es aceptado, en caso de que sea mayor el lote es rechazado. Y este a su vez se divide en:

TIPOS DE MUESTREO

- a) Sencillo o Simple:** Consiste en un tamaño de muestra n , y un número de aceptación c , ambos fijados de antemano.
- El número de unidades que se deben inspeccionar deberá ser igual al tamaño de la muestra dado por el plan.
 - Si el número de defectuosos encontrados en la muestra es igual o menor que el número de aceptación, se debe considerar como **ACEPTABLE EL LOTE o PRODUCCIÓN UNITARIA.**
 - Si el número de defectivos es igual o mayor que el número de rechazo, el **LOTE O PRODUCCIÓN SE DEBE RECHAZAR**

TIPOS DE MUESTREO

- **b) Doble:** La idea de este muestreo es tomar una primera muestra de tamaño pequeño para detectar los lotes muy buenos o lo muy malos, y si en la primera muestra no se puede decidir si aceptar o rechazar porque la cantidad de unidades defectuosas ni es muy pequeña ni es muy grande, entonces se toma una segunda muestra, para decidir si aceptar o rechazar tomando en cuenta las unidades defectuosas encontradas en las dos muestras.

TIPOS DE MUESTREO

- El número de unidades de la muestra que se inspecciona debe ser igual al primer tamaño de muestra dado por el plan.
- Cuando el número de defectivos que se encuentran en la primera muestra sea igual o menor que el primer número de aceptación, se considerará aceptable el lote o la producción unitaria.
- Si el número de defectivos en la primera muestra es igual o mayor que el primer número de rechazo, se debe de rechazar el lote o la producción.
- Si el número de defectivos en la primera muestra queda entre los primeros números de aceptación y de rechazo se toma UNA SEGUNDA MUESTRA, del tamaño dado por el plan y se inspecciona; el número de la primera y la segunda muestra se suman; si la suma es igual o menor que el segundo número de aceptación , se ACEPTA EL LOTE o PRODUCCIÓN. Si la suma de defectivos es mayor o igual que el segundo número de rechazo , el LOTE O PRODUCCIÓN SE RECHAZA.

TIPOS DE MUESTREO

c) Múltiple:

- En una inspección de muestreo múltiple, el procedimiento debe de ser , similar al descrito en el muestreo doble a excepción de que el número requerido de muestras sucesivas para llegar a una decisión, debe ser mayor de dos.

TIPOS DE DIAGRAMA DE PARETO

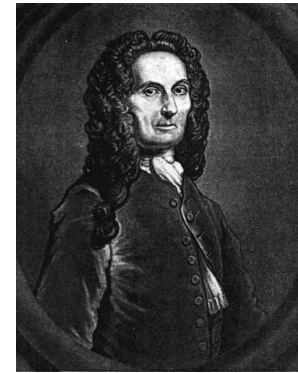
- Existen dos tipos de diagramas de Pareto:
- **Diagramas de fenómenos.** Se utilizan para determinar cuál es el principal problema que origina el resultado no deseado. Estos problemas pueden ser de calidad, coste, entrega, seguridad u otros.
- **Diagramas de causas.** Se emplean para, una vez encontrados los problemas importantes, descubrir cuáles son las causas más relevantes que los producen

CONSEJOS PARA ELABORAR Y USAR LOS DIAGRAMAS DE PARETO

- No es conveniente que la categoría de “otros” represente un porcentaje de los más altos. De ser así, se debe realizar un método diferente de clasificación.
- Es preferible representar los datos (si es posible) en valores monetarios.
- Si un factor se puede solucionar fácilmente debe afrontarse de inmediato aunque sea de poca importancia.
- Es imprescindible realizar un diagrama de causas si se quieren realizar mejoras.

La distribución normal

La distribución normal fue reconocida por primera vez por el francés Abraham de Moivre (1667-1754).



Posteriormente, Carl Friedrich Gauss (1777-1855) realizó estudios más a fondo donde formula la ecuación de la curva conocida comúnmente, como la "Campana de Gauss".



La función de distribución

- Puede tomar cualquier valor ($-\infty, +\infty$)
- Hay más probabilidad para los valores cercanos a la media μ
- Conforme nos separamos de μ , la probabilidad va decreciendo de igual forma a derecha e izquierda (es simétrica).
- Conforme nos separamos de μ , la probabilidad va decreciendo dependiendo la desviación típica σ .

La función F(x)

Función de Distribución

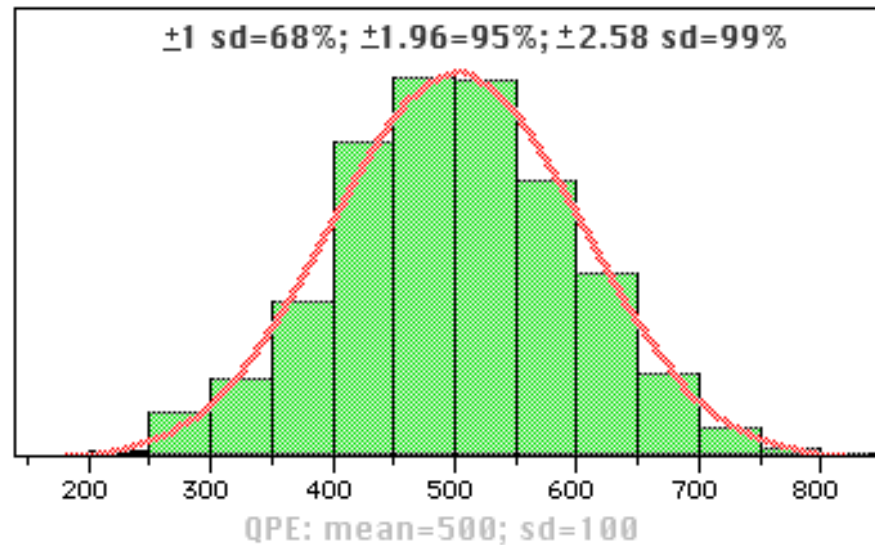
$$F(x) = \int_{-\infty}^x \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx$$

$$-\infty < x < +\infty$$

$$F(x) = P(X \leq x)$$

Propiedades de la distribución normal:

El área bajo la curva aproximado del promedio μ a más o menos una desviación estándar (1σ) es de 0.68, a más o menos 2σ es de .0 95 y a más o menos 3σ es de 0.99.



La distribución normal estándar

- Z se la denomina *variable tipificada de X*, y a la curva de su función de densidad se le conoce como la curva normal estándar.
- Es una distribución normal con promedio 0 y una desviación estándar de 1.
- Todas las variables normalmente distribuidas se pueden transformar a la distribución normal estándar utilizando la fórmula para calcular el valor Z correspondiente.

En resumen

- Podemos decir que el valor de Z es la cantidad de desviaciones estándar a la que está distanciada la variable X del promedio.
- A la variable Z se la denomina *variable tipificada de X* , y a la curva de su función de densidad se le conoce como la curva normal estándar

Características de la distribución normal estándar.

- No depende de ningún parámetro.
- Su media es 0, su varianza es 1 y su desviación estándar es 1.
- La curva $f(x)$ es simétrica respecto del eje de Y
- Tiene un máximo en el eje de Y .
- Tiene dos puntos de inflexión en $z=1$ y $z=-1$

Teorema del *Límite Central*

Nos indica que, bajo condiciones muy generales, según aumenta la cantidad de datos, la distribución de la suma de variables aleatorias tendera a seguir hacia una distribución normal.

En otras palabras el Teorema del Límite Central garantiza una distribución normal cuando el tamaño de la muestra es suficientemente grande.

Área bajo la curva normal estándar

El área bajo la curva normal estándar es útil para asignar probabilidades de ocurrencia de la variable X .

Debemos tomar en cuenta que el área total bajo la curva es igual a 1. Y que, por ser una gráfica simétrica, cada mitad tiene un área de 0.5.



Pasos para determinar el área bajo la curva normal estándar

- Paso 1 - Interpretar gráficamente el área de interés.
- Paso 2 - Determinar el valor Z
- Paso 3 - Buscar en la tabla de probabilidades.
- Paso 4 - Hacer la suma o resta de áreas para encontrar la probabilidad deseada

Ejemplos y ejercicios

Supongamos que sabemos que el peso de los/as estudiantes universitarios/as sigue una distribución aproximadamente normal, con una media de 140 libras y una desviación estándar de 20 libras.

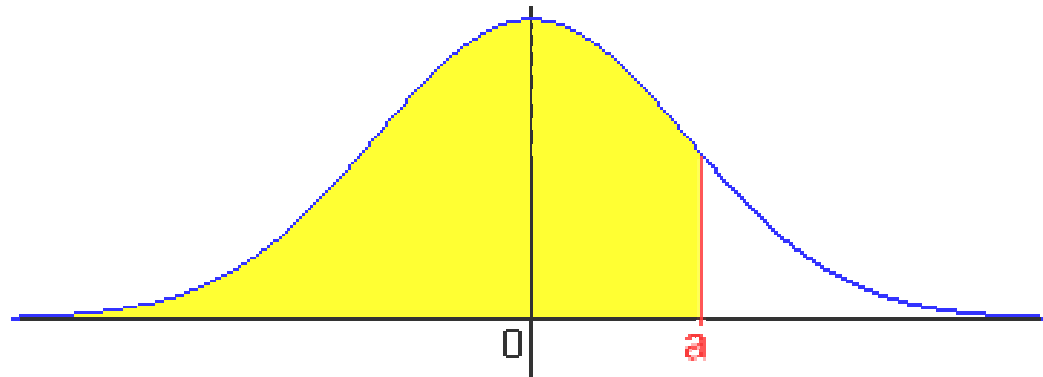
Ejemplo 1

Determine la probabilidad de que una persona tenga un peso menor o igual a 150 libras

Paso 1 Interpretar gráficamente el área de interés.

Gráficamente si decimos que $a=150$ libras, el área de la curva que nos interesa es la siguiente:

$$P(Z \leq a) \rightarrow \text{Tablas}$$



Ejemplo 1

Determine la probabilidad de que una persona tenga un peso menor o igual a 150 libras

Paso 2 - Determinar el valor Z: $Z = \frac{X - \mu}{\sigma} = \frac{150 - 140}{20} = 0.50$

Paso 3 - Buscar en la tabla de probabilidades.

Buscamos en la Tabla 1 el valor $Z=0.50$ y obtenemos el área de 0.6915

Paso 4 - Hacer la suma o resta de áreas para encontrar la probabilidad deseada. ←

En este ejemplo no es necesario realizar ningún computo adicional ya que el área es la misma que se representa en la [Tabla 1](#)

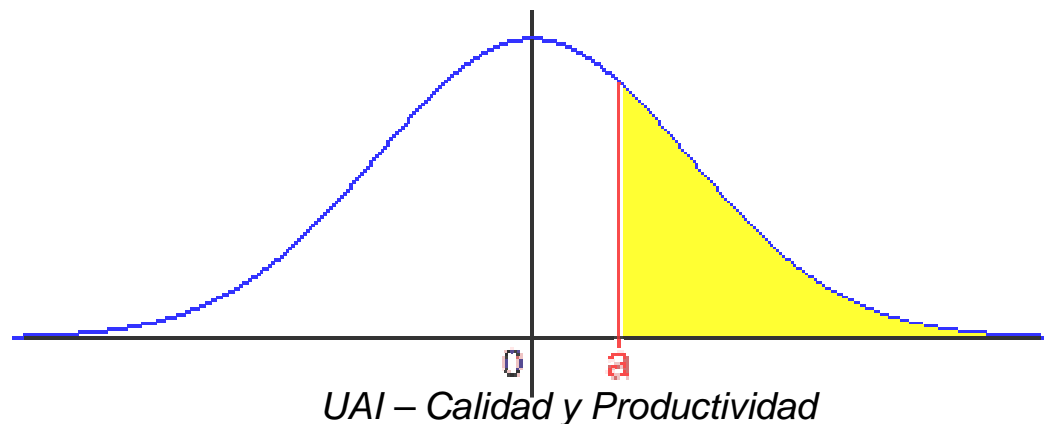
Ejemplo 2

Si deseamos la probabilidad de que una persona, elegida al azar, tenga un peso mayor o igual a 150 libras

Paso 1 Interpretar gráficamente el área de interés.

Gráficamente si decimos que $a=150$ libras, el área de la curva que nos interesa es la siguiente:

$$P(Z > a) = 1 - P(Z \leq a)$$



Ejemplo 2

Si deseamos la probabilidad de que una persona, elegida al azar, tenga un peso mayor o igual a 150 libras

Paso 2 - Determinar el valor Z:
$$Z = \frac{X - \mu}{\sigma} = \frac{150 - 140}{20} = 0.50$$

Paso 3 - Buscar en la tabla de probabilidades.

Buscamos en la Tabla 1 el valor $Z=0.50$ y obtenemos el área de 0.6915.

Paso 4 - Hacer la suma o resta de áreas para encontrar la probabilidad deseada.

En este ejemplo el área de 0.6915 no representa el área que nos interesa sino la contraria. En este caso debemos restarle 1 a la probabilidad encontrada.

$$1 - .6915 = 0.3085$$

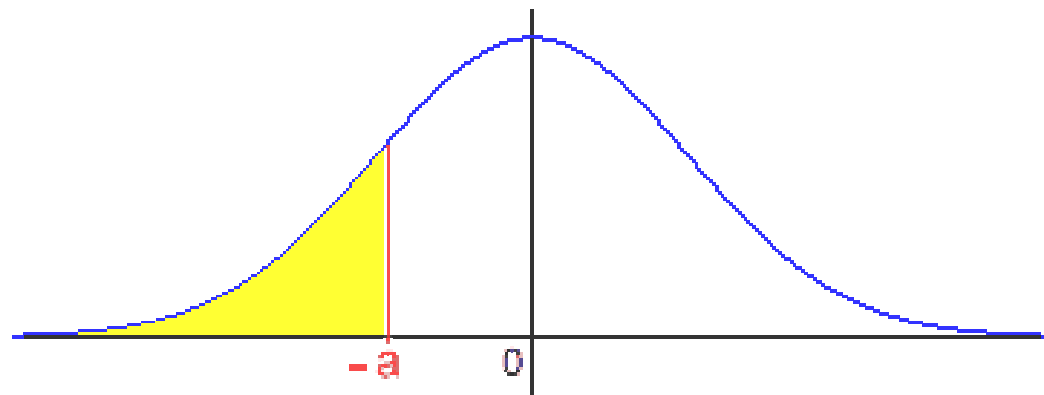
Ejemplo 3

Determine la probabilidad de que una persona, elegida al azar, tenga un peso menor o igual a 115 libras

Paso 1 Interpretar gráficamente el área de interés.

Gráficamente si decimos que $a=115$ libras, el área de la curva que nos interesa es la siguiente:

$$P(Z \leq -a) = 1 - P(Z \leq a)$$



Ejemplo 3

Determine la probabilidad de que una persona, elegida al azar, tenga un peso menor o igual a 115 libras

Paso 2 - Determinar el valor Z:
$$Z = \frac{X - \mu}{\sigma} = \frac{115 - 140}{20} = -1.25$$

Paso 3 - Buscar en la tabla de probabilidades.

Buscamos en la Tabla 1 el valor $Z=-1.25$ y obtenemos el área de 0.8944.

Paso 4 - Hacer la suma o resta de áreas para encontrar la probabilidad deseada.

En este ejemplo el área de 0.8944 no representa el área que nos interesa sino la contraria. En este caso debemos restarle 1 a la probabilidad encontrada.

$$1 - .8944 = 0.2212$$

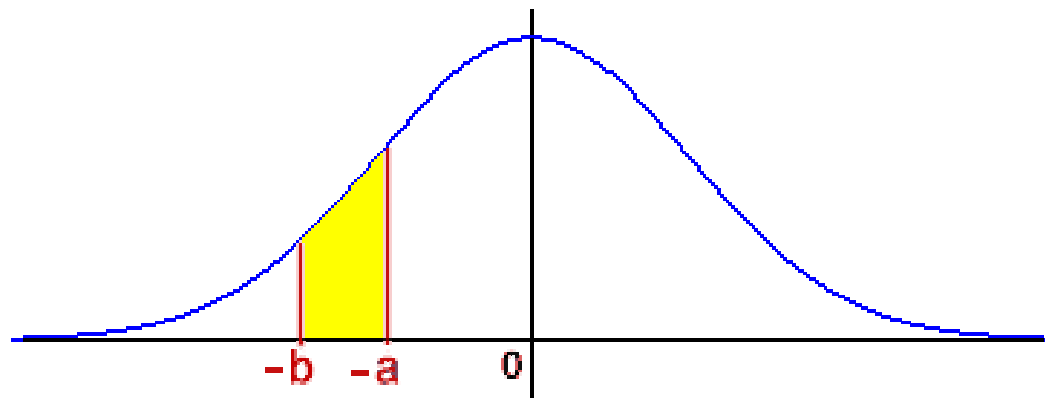
Ejemplo 4

Determine la probabilidad de elegir a una persona que pese entre 115 y 130 libras.

Paso 1 Interpretar gráficamente el área de interés.

Gráficamente si decimos que $a=115$ libras y $b= 130$ libras, el área de la curva que nos interesa es la siguiente:

$$P(-b < Z \leq -a) = P(a < Z \leq b)$$



Ejemplo 4

Determine la probabilidad de elegir a una persona que pese entre 115 y 130 libras.

Paso 2 - Determinar el valor Z

Cuando $X=115$
$$Z = \frac{X - \mu}{\sigma} = \frac{115 - 140}{20} = -1.25$$

para $X=130$
$$Z = \frac{X - \mu}{\sigma} = \frac{130 - 140}{20} = -0.50$$

Paso 3 - Buscar en la tabla de probabilidades.

Buscamos en la Tabla 1 el valor $Z=-1.25$ y obtenemos el área de $(1-0.8944.)=0.1056$

Para $Z = -.050$ el área es de $(1- 0.6915)= 0.3085$

Ejemplo 4

Determine la probabilidad de elegir a una persona que pese entre 115 y 130 libras.

Paso 4 - Hacer la suma o resta de áreas para encontrar la probabilidad deseada.

En este ejemplo el área será la diferencia de $.3085 - .1056 = .2029$.

ESTIMACIÓN PUNTUAL

La media muestral es el mejor estimador puntual con el que podemos aproximar el valor de la media de la población.

$$\bar{X} = \frac{X_1 + X_2 + X_3 + \dots + X_{n-1} + X_n}{n} = \frac{\sum_{i=1}^n X_i}{n}$$

Estimación por intervalos de confianza

Un intervalo de confianza es una gama de valores que **probablemente** contiene al verdadero valor de la media poblacional.

Intervalos de Confianza

Métodos de estimación:

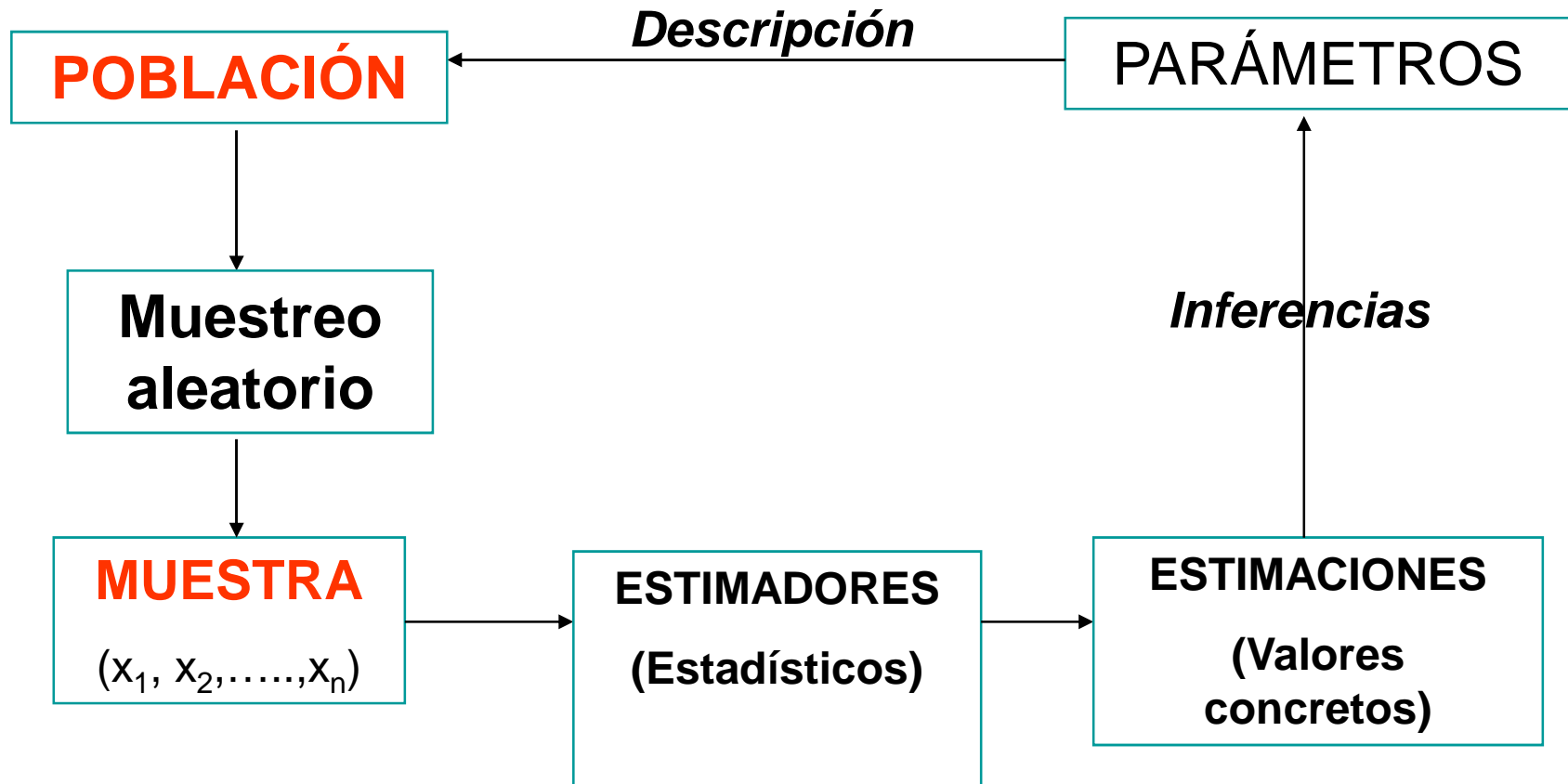
Estimación puntual:

utilización de datos de la muestra para calcular un solo número para estimar el parámetro de interés.

Estimación de intervalo:

ofrece un intervalo de valores razonables dentro del cual se pretende que esté el parámetro de interés, en este caso la media poblacional, con un cierto grado de confianza

Intervalos de Confianza



Intervalos de confianza

ESTIMADORES

$$\hat{\mu} = \sum x_i / n$$

$$\hat{\sigma}^2 = \sum (X_i - \hat{\mu})^2 / (n-1)$$

ESTIMACIONES

\bar{X}

S^2

Valores concretos

Ejemplo: distribución de tallas de neonatos

μ	σ^2	Valores desconocidos de los parámetros media y variancia de la talla de la población
$\hat{\mu} = \sum x_i / n$	$\hat{\sigma}^2 = \sum (X_i - \hat{\mu})^2 / (n-1)$	Estimadores
{46;48;51;52;52}		Muestra
$\bar{x} = \frac{46 + 49 + 51 + 52 + 52}{5} = 50$		Estimación puntual de μ
$s^2 = \frac{(46 - 50)^2 + \dots + (52 - 50)^2}{5 - 1} = 6,5$		Estimación puntual de σ^2

Intervalos de confianza bilaterales: *construcción*

Dada una variable aleatoria X con media μ
y desviación estándar σ ,
el teorema del límite central afirma que posee una distribución
normal estándar si X :

- se encuentra distribuida normalmente,
- no se encuentra distribuida normalmente y n sea suficientemente grande

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

Intervalos de confianza bilaterales: *construcción*

Para una variable normal estándar, 95% de las observaciones se ubican entre -1,96 y +1,96.

En otras palabras, la probabilidad de que Z tome un valor entre -1,96 y +1,96 es:

$$P(-1,96 \leq Z \leq 1,96) = 0,95$$

Al sustituir el valor de Z:

$$P\left(-1,96 \leq \frac{\bar{x} - \mu}{\sigma / \sqrt{n}} \leq 1,96\right) = 0,95$$

Intervalos de confianza bilaterales: *construcción*

Multiplicamos los tres términos de la desigualdad por el error estándar

$$\frac{\sigma}{\sqrt{n}}$$

Por tanto,

$$P\left(-1,96\frac{\sigma}{\sqrt{n}} \leq \bar{x} - \mu \leq 1,96\frac{\sigma}{\sqrt{n}}\right) = 0,95$$

Intervalos de confianza bilaterales: *construcción*

Al reordenar términos:

$$P\left(\bar{x} - 1,96 \frac{\sigma}{\sqrt{n}} \leq \mu \leq \bar{x} + 1,96 \frac{\sigma}{\sqrt{n}}\right) = 0,95$$

La \bar{x} ya no se localiza en el centro de la desigualdad; en lugar de eso, la afirmación de probabilística indica algo sobre μ

Cálculo del Tamaño de la Muestra para Estimar una Media

El error máximo de estimación esta dado por:

$$E = \frac{Z \sigma}{\sqrt{n}}$$

Si se eleva al cuadrado ambos lados de esta ecuación y se despeja n de la ecuación resultante, obtenemos:

$$n = \left(\frac{Z \sigma}{E} \right)^2$$

Una empresa eléctrica fabrica focos que tienen una duración aproximadamente normal con una desviación estándar de 40 horas. ¿De qué tamaño se necesita una muestra si se desea tener 96% de confianza que la media real esté dentro de 10 horas de la media real? Se necesita una muestra de 68 focos para estimar la media de la población y tener un error máximo de 10 horas.

$$n = \left(\frac{z \sigma}{E} \right)^2 = \left(\frac{(2.053)(40)}{10} \right)^2 = 67.43$$

¿Qué pasaría si en lugar de tener un error de estimación de 10 horas sólo se requiere un error de 5 horas?

Se puede observar como el tamaño de la muestra aumenta, pero esto tiene como beneficio una estimación más exacta.

$$n = \left(\frac{z \sigma}{E} \right)^2 = \left(\frac{(2.053)(40)}{5} \right)^2 = 269.74$$

Un biólogo quiere estimar el peso promedio de los ciervos cazados en el estado de Maryland. Un estudio anterior de diez ciervos cazados mostró que la desviación estándar de sus pesos es de 12.2 libras. ¿Qué tan grande debe ser una muestra para que el biólogo tenga el 95% de confianza de que el error de estimación es a lo más de 4 libras?

Solución:

$$n = \left(\frac{z \sigma}{E} \right)^2 = \left(\frac{(1.96)(12.2)}{4} \right)^2 = 35.736$$

En consecuencia, si el tamaño de la muestra es 36, se puede tener un 95% de confianza en que μ difiere en menos de 4 libras de \bar{x} .

LA CURVA NORMAL



- La curva normal tiene forma de campana con un solo pico justo en el centro de la distribución.
La media, mediana y moda de la distribución aritmética son iguales y se localizan en el pico.
La mitad del área bajo la curva está a la derecha del pico, y la otra mitad está a la izquierda.
- La distribución normal es simétrica respecto a su media.
La distribución normal es asintótica - la curva se acerca cada vez más al eje x pero en realidad nunca llega a tocarlo.



DISTRIBUCIONES DE PROBABILIDAD

- Las distribuciones de probabilidades discretas más importantes son:
- Distribución Binomial, y
- Distribución de Poisson

DISTRIBUCION BINOMIAL

La distribución binomial es una distribución de probabilidades que surge al cumplirse cinco condiciones:

Existe una serie de N ensayos,

En cada ensayo hay sólo dos posibles resultados,

En cada ensayo, los dos resultados posibles son mutuamente excluyentes,

Los resultados de cada ensayo son independientes entre si, y

La probabilidad de cada resultado posible en cualquier ensayo es la misma de un ensayo a otro.

Cuando se cumple estas condiciones, la distribución binomial proporciona cada resultado posible de los N ensayos y la probabilidad de obtener cada uno de estos resultados.

Para este tipo de distribución de probabilidad, la función matemática es la siguiente

DISTRIBUCIÓN BINOMIAL

Nº de resultados con exactamente x éxitos
de entre n ensayos

La probabilidad de x éxitos de entre n
de entre n ensayos

$$P(X) = \frac{n!}{X!(n-X)!} p^X (1-p)^{n-X}$$

EJERCICIO

- La probabilidad de que un hombre acierte en el blanco es 1/4. Si dispara 10 veces ¿cuál es la probabilidad de que acierte exactamente en tres ocasiones? ¿Cuál es la probabilidad de que acierte por lo menos en una ocasión?
- $B(10, 1/4)$ $p = 1/4$ $q = 3/4$

$$P(X = 3) = \binom{10}{3} \left(\frac{1}{4}\right)^3 \cdot \left(\frac{3}{4}\right)^7 = 0.25$$

$$P(\text{al menos uno}) = 1 - \binom{10}{0} \left(\frac{1}{4}\right)^0 \cdot \left(\frac{3}{4}\right)^{10} = 0.9437$$

DISTRIBUCION DE POISSON

- Se dice que existe un proceso de Poisson si podemos observar eventos discretos en un área de oportunidad – un intervalo continuo (de tiempo, longitud, superficie, etc.) – de tal manera que si se reduce lo suficiente el área de oportunidad o el intervalo,
- La probabilidad de observar exactamente un éxito en el intervalo es constante.
- La probabilidad de obtener más de un éxito en el intervalo es 0.
- La probabilidad de observar un éxito en cualquier intervalo es estadísticamente independiente de la de cualquier otro intervalo.
- Esta distribución se aplica en situaciones como:
- El numero de pacientes que llegan al servicio de emergencia de un hospital en un intervalo de tiempo.
- El numero de radiaciones radiactivas que se recibe en un lapso de tiempo,
- El numero de glóbulos blancos que se cuentan en una muestra dada.
- El numero de partos triples por año
- Su utilidad en el área de la salud es muy amplia.
- La expresión matemática para la distribución de Poisson para obtener **X** éxitos, dado que se esperan λ éxitos es:

DISTRIBUCIÓN DE POISSON

$$P(X) = \frac{e^{-\lambda} \lambda^X}{X!}$$

- Donde: $P(X)$ = probabilidad de X éxitos dado el valor de λ
- λ = esperanza del número de éxitos.
- e = constante matemática, con valor aproximado 2.711828
- X = número de éxitos por unidad

DISTRIBUCIÓN DE POISSON

- La distribución de Poisson se utiliza en situaciones donde los sucesos son impredecibles o de ocurrencia **aleatoria**. En otras palabras no se sabe el total de posibles resultados.
- Permite determinar la probabilidad de ocurrencia de un suceso con **resultado discreto**.
- Es muy útil cuando la muestra o segmento n es grande y la probabilidad de **éxitos** p es pequeña.
- Se utiliza cuando la probabilidad del evento que nos interesa se distribuye dentro de un **segmento** n dado como por ejemplo distancia, área, volumen o tiempo definido.

DISTRIBUCIÓN DE POISSON

La probabilidad de que un producto salga defectuoso es de 0.012.
¿Cuál es la probabilidad de que entre 800 productos ya fabricados
hayan 5 defectuosos?

En este ejemplo vemos nuevamente la probabilidad p menor que
0.1, y el producto $n * p$ menor que 10, por lo que aplicamos el
modelo de distribución de Poisson

$$P(x = 5) = e^{-9,6} * \frac{9,6^5}{5!}$$

El resultado es $P(x = 5) = 0.04602$

Por lo tanto, la probabilidad de que haya 5 productos
defectuosos entre 800 recién producidos es de 4.6%.

DISTRIBUCIÓN CHI CUADRADO

- Fórmula

$$\chi^2 = \frac{(n - 1)s^2}{\sigma^2}$$

n=tamaño de muestra

s^2 =varianza de muestra

σ^2 =varianza de población

DISTRIBUCIONES

- Hipergeométrica
- Weibull
- Exponencial

PRUEBA DE HIPÓTESIS

- La hipótesis nula (**H0**) especifica el valor de un parámetro poblacional. Se conduce un experimento para ver si el valor especificado no es razonable.
- Ejemplo :Un semillero publicita que el peso promedio de una espiga de una cierta variedad es de 180 gramos con una desví estándar de 30 gramos. Un productor de avanzada sospecha que el peso es distinto de 180 gramos, decide por lo tanto conducir un experimento. El propósito del mismo es ver si el peso de 180 gramos es incorrecto. Por lo tanto la hipótesis nula de interés es:
 - $H_0 : \mu = 180$ gramos
- La hipótesis alternativa (**H1**) da una suposición opuesta a aquella presentada en la hipótesis nula. El experimento se lleva a cabo para conocer si la hipótesis alternativa puede ser sustentada.
- En el ejemplo previo el productor sospecha que el peso medio es distinto de 180 gramos. Esta es la hipótesis a ser sustentada y así la hipótesis alternativa es:
 - $H_1 : \mu \neq 180$ gramos
- Se puede ver que las hipótesis son excluyentes. La hipótesis alternativa frecuentemente se llama **hipótesis de investigación**, porque este tipo de hipótesis expresa la teoría que el investigador o experimentador cree va a ser verdadera.

ERRORES TIPO I Y TIPO II

		Estado real de la Naturaleza	
		La hipótesis nula es verdadera.	La hipótesis nula es falsa.
Decisión	Decidimos rechazar la hipótesis nula.	Error tipo I (rechazar una hipótesis nula verdadera)	Decisión correcta
	No rechazamos la hipótesis nula.	Decisión correcta	Error tipo II (no rechazar una hipótesis nula falsa)

PRUEBA DE HIPÓTESIS

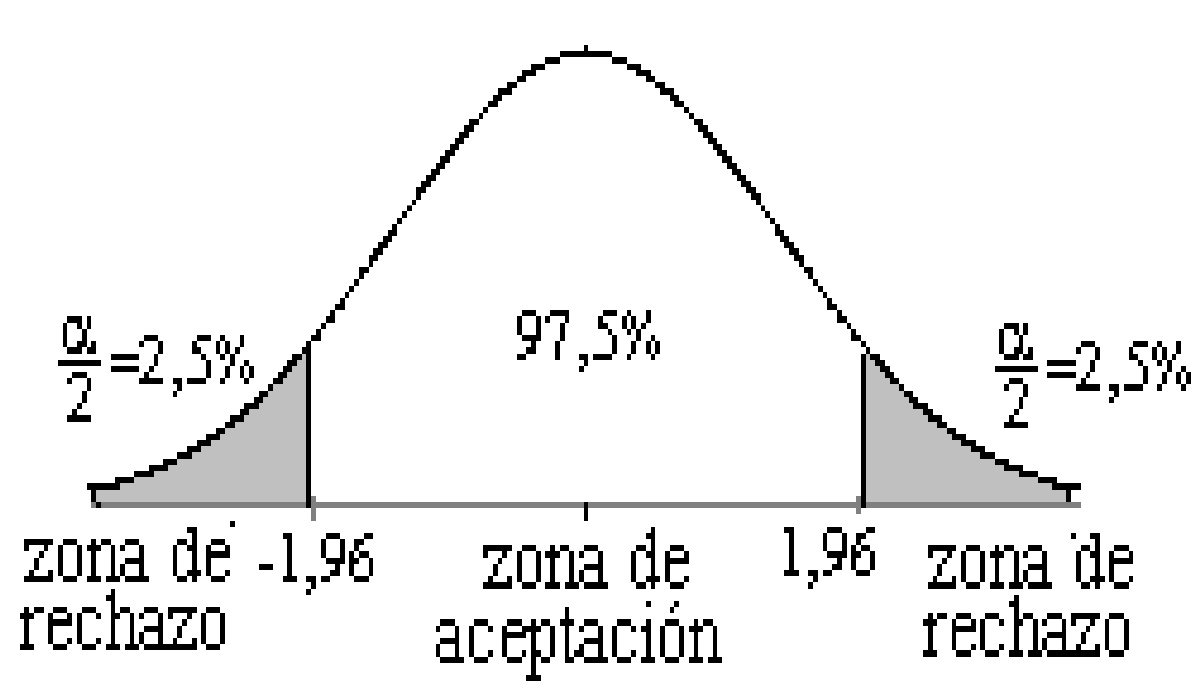
- En el ejemplo considerado el productor aceptando un error de 0,05 (5%), conocido también como **nivel de significación** y utilizando la estadística z, plantearía la hipótesis como sigue:
- $H_0 : \mu = 180$ gramos
- $H_1 : \mu \neq 180$ gramos
- Suponiendo que los resultados del experimento produjeron una media muestral de 187 gramos, el test estadístico se construiría como:

$$Z = \frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}} = \frac{187 - 180}{\frac{30}{\sqrt{50}}} = 1,65$$

- donde : 187 = media de la muestra ($\bar{x} = 187$)
- 180 = media hipotética (poblacional $\mu = 180$)
- 30 = desvío estándar poblacional (conocido) ($\sigma = 30$)
- 50 = tamaño de la muestra o repeticiones ($n=50$)

PRUEBA DE HIPÓTESIS

-



PRUEBA DE HIPÓTESIS

- como el valor de **z calculado= 1,65** es menor que 1,96 o sea cae en la **región de aceptación** , no hay evidencias suficientes como para rechazar la hipótesis de que la media de la población es igual a 180.
- Conclusión: la publicidad que hace el semillero de que el peso promedio de las espigas de una cierta variedad es de 180 gramos, es correcta, aunque podría existir una probabilidad de error tipo II, si de hecho la media de tal variedad no fuera 180 gramos

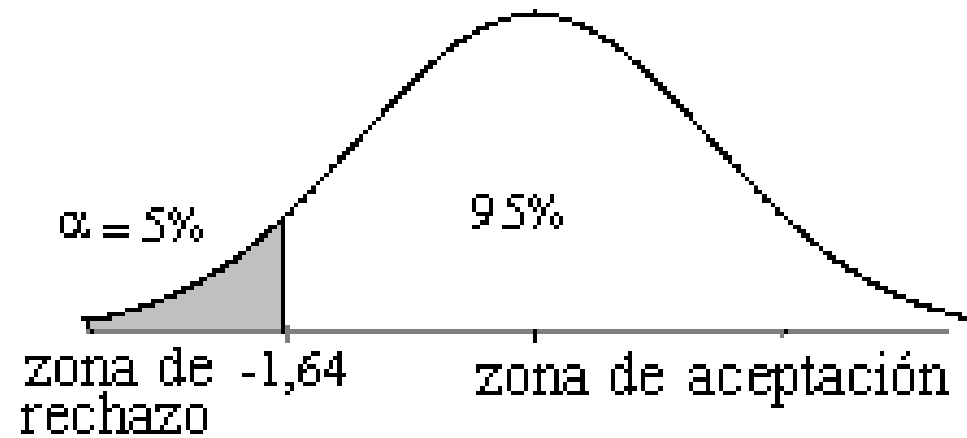
PRUEBA DE HIPÓTESIS

- Si en el mismo ejemplo, el productor, basándose en algún conocimiento de la variedad en cuestión sospechara que el peso promedio de las espigas es menor que 180, las hipótesis se plantearían como:
 - $H_0 : \mu = 180$ gramos
 -
 - $H_1 : \mu < 180$ gramos
- $\alpha = 0,05$

PRUEBA DE HIPÓTESIS

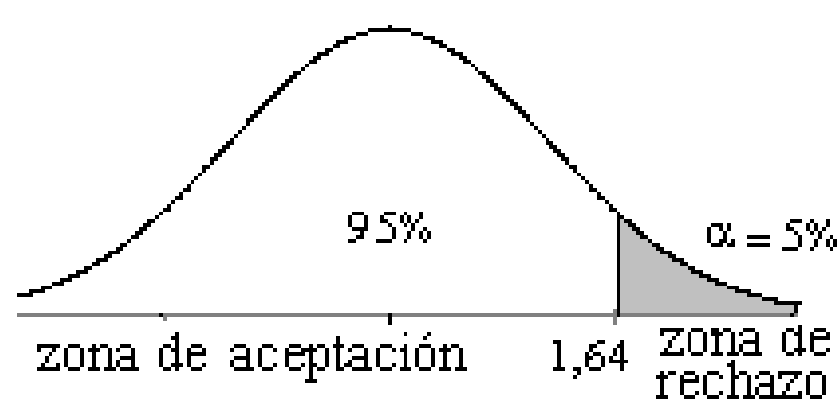
- En este caso la desigualdad de la hipótesis alternativa indica cuál sería la zona de rechazo, el valor de α ya no se particiona sino que se acumula todo hacia un solo lado, el izquierdo en este ejemplo y el valor tabulado de z se busca en la tabla con un valor de probabilidad del 95% siendo **$z = -1,64$** (el signo negativo no figura en la tabla ya que siendo la distribución normal simétrica, lo que se hace es anteponer el signo negativo al valor de z que corresponde al nivel de probabilidad especificado)

PRUEBA DE HIPÓTESIS



PRUEBA DE HIPÓTESIS

- Si por otra parte, el productor sospechara que el peso promedio es mayor que 180 gramos, la hipótesis y la zona de rechazo se plantearían como:
- $H_0: \mu \leq 180$ gramos
- $H_1: \mu > 180$ gramos
- $\alpha = 0,05$



PRUEBA DE HIPÓTESIS

- en ambas situaciones el **test estadístico** se construye como:

$$Z = \frac{(\bar{X} - \mu)}{\sigma_x} = \frac{(187 - 180)}{30/\sqrt{50}}$$

cuando σ^2 se desconoce, el **test estadístico** se construye como:

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

PRUEBA DE HIPÓTESIS

- Este valor difiere del anterior en que, en lugar de aparecer la desviación estándar de la población, nos encontramos con su estimador muestral insesgado **S**, que se distribuye, como **t de Student** ($t \sim t(n-1)$)

$$H_0 =$$

PRUEBA DE HIPÓTESIS

- El contenido de azúcar en una bebida no debe superar los 20gr/litro.
- Se toma una muestra igual a 9 botellas resultando la $X_{prom.} = 20,1$ gr/litro
- Si la Desv.Std = 0,2 gr/litro se nos pide determinar a una significación del 0,05% si el contenido de azúcar supera al aceptado.

$$H_0 : \mu = 20$$

- $H_1 : \mu > 20$

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

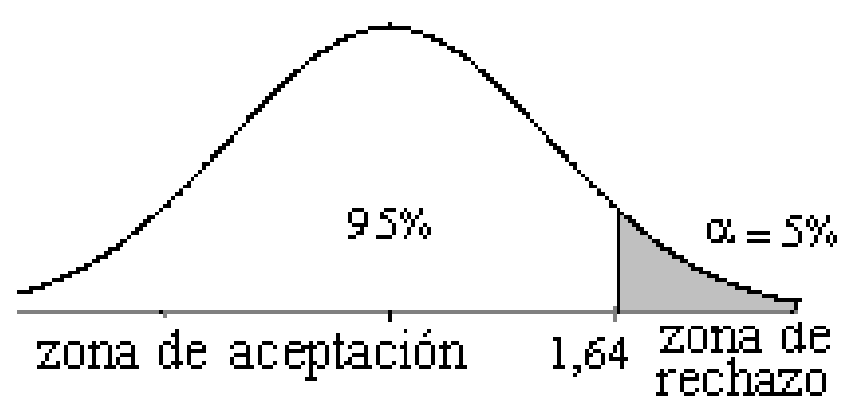
- $\bar{X} = 20,1$

- $\alpha = 0,05$

- $t = \frac{20,1 - 20,0}{0,2 / 3} = \frac{0,1 \times 3}{0,2} = 1,5$

PRUEBA DE HIPÓTESIS

- $Z_{\text{tabla}} = 1,645$ $Z_{\text{calculado}} = 1,5$



El valor calculado se encuentra en la zona de aceptación por lo que no se rechaza la Hipótesis nula

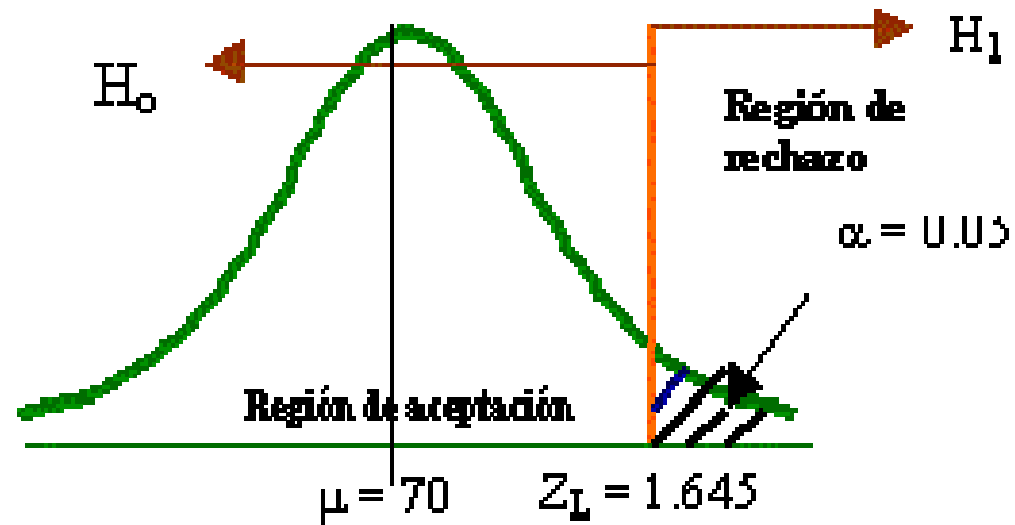
PRUEBA DE HIPÓTESIS

- Una muestra aleatoria de 100 muertes registradas en Argentina el año pasado muestra una vida promedio de 71.8 años. Suponga una desviación estándar poblacional de 8.9 años, ¿esto parece indicar que la vida media hoy en día es mayor que 70 años? Utilice un nivel de significancia de 0.05.
- *Solución:*

PRUEBA DE HIPÓTESIS

- *Se trata de una distribución muestral de medias con desviación estándar conocida.*
- *Datos:*
- $\mu = 70$ años
- $\sigma = 8.9$ años
- $\bar{X} = 71.8$ años
- $n = 100$
- $\alpha = 0.05$
 - *Ensayo de hipótesis*
- $H_0; \mu = 70$ años.
- $H_1; \mu > 70$ años.

PRUEBA DE HIPÓTESIS



Regla de decisión:

Si $z_R \leq 1.645$ no se rechaza H_0 .

Si $z_R > 1.645$ se rechaza H_0

PRUEBA DE HIPÓTESIS

- Cálculos:

$$Z_R = \frac{\bar{x}_R - \mu}{\frac{\sigma}{\sqrt{n}}} = \frac{71.8 - 70}{\frac{8.9}{\sqrt{100}}} = 2.02$$

Justificación y decisión.

Como $2.02 > 1.645$ se rechaza H_0 y se concluye con un nivel de significancia del 0.05 que la vida media hoy en día es mayor que 70 años.

Existe otra manera de resolver este ejercicio, tomando la decisión en base al estadístico real, en este caso la media de la muestra. De la formula de la distribución muestral de medias se despeja la media de la muestra:

PRUEBA DE HIPÓTESIS

- Existe otra manera de resolver este ejercicio, tomando la decisión en base al estadístico real, en este caso la media de la muestra. De la fórmula de la distribución muestral de medias se despeja la media de la muestra:

$$Z_L = \frac{\bar{X}_L - \mu}{\sigma / \sqrt{n}}$$

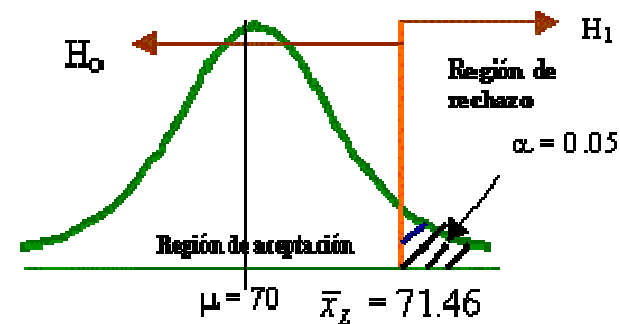
$$\bar{X}_L = \mu + \frac{Z_L \sigma}{\sqrt{n}} = 70 + \frac{(1.645)(8.9)}{\sqrt{100}} = 71.46$$

Regla de decisión:

Si $\bar{X}_R \leq 71.46$ No se rechaza H_0

Si $\bar{X}_R > 71.46$ Se rechaza H_0

Como la media de la muestral es de 71.8 años y es mayor al valor de la media muestral límite de 71.46 por lo tanto se rechaza H_0 y se llega a la misma conclusión.



PRUEBA DE HIPÓTESIS

- Prueba de Hipótesis
- La prueba de hipótesis estadística es un procedimiento de 4 pasos bien organizados
- **Paso 1:** Formular las hipótesis nula y alternativa
- **Paso 2:** Determinar el criterio de prueba o contraste
- **Paso 3:** Obtener de los datos muestrales (evidencia) y calcúlese el valor de la estadística de prueba
- **Paso 4:** Tomar una decisión e interpretarla

FIABILIDAD

- FIABILIDAD ES LA PROBABILIDAD DE QUE UN SISTEMA O COMPONENTE FUNCIONE DE ACUERDO A LAS ESPECIFICACIONES, DURANTE UN INTERVALO DADO DE TIEMPO Y EN DETERMINADAS CONDICIONES DE OPERACIÓN

FIABILIDAD

- Probabilidad de sobrevivir $P_s=R= e^{-\lambda t}$
- $e=2,718$
- t =período de operación libre de fallas
- μ =tiempo medio entre fallas (**MTBF**)
- λ =cte. de fallas

$$\lambda = \frac{1}{\mu}$$

$$R = e^{-t/\mu}$$

FIABILIDAD

- Serie $R_s = R_1 \times R_2 \times R_3$
- Paralelo $R_p = 1 - (U_1 \times U_2 \times U_3)$
- $U_1 = 1 - R_1$
- $U_2 = 1 - R_2$
- $U_3 = 1 - R_3$
- Serie/Paralelo $R_{sp} = R_1 \times R_2 \times R_3 \times R_4$

EJERCICIO

- Cual es la probabilidad de un ítem de fallar en menos de 5 horas, si tiene una constante de fallas de 1/hora

$$e = \mu^{-\lambda t}$$

- $e^{-5 \times 1}$
- $R(5) = e^{-5} = 0,00674$
- La prob. de falla es $= 1 - 0,00674 = 0,99326$

EJERCICIO

- Dado un tiempo medio de falla de 90 horas y una razón cte., cual es la prob. de que un producto sobreviva 90 horas.

$$R_{(t)} = e^{-\lambda t}$$

$$\lambda = \frac{1}{t}$$

$$\lambda = \frac{1}{90} = 0,0111$$

$$R_{(90)} = e^{-0,0111 \times 90} = e^{-0,9999} = 0,36789$$

EJERCICIO

- Cual es la fiabilidad de 3 componentes en serie si la de cada uno es de 0,40

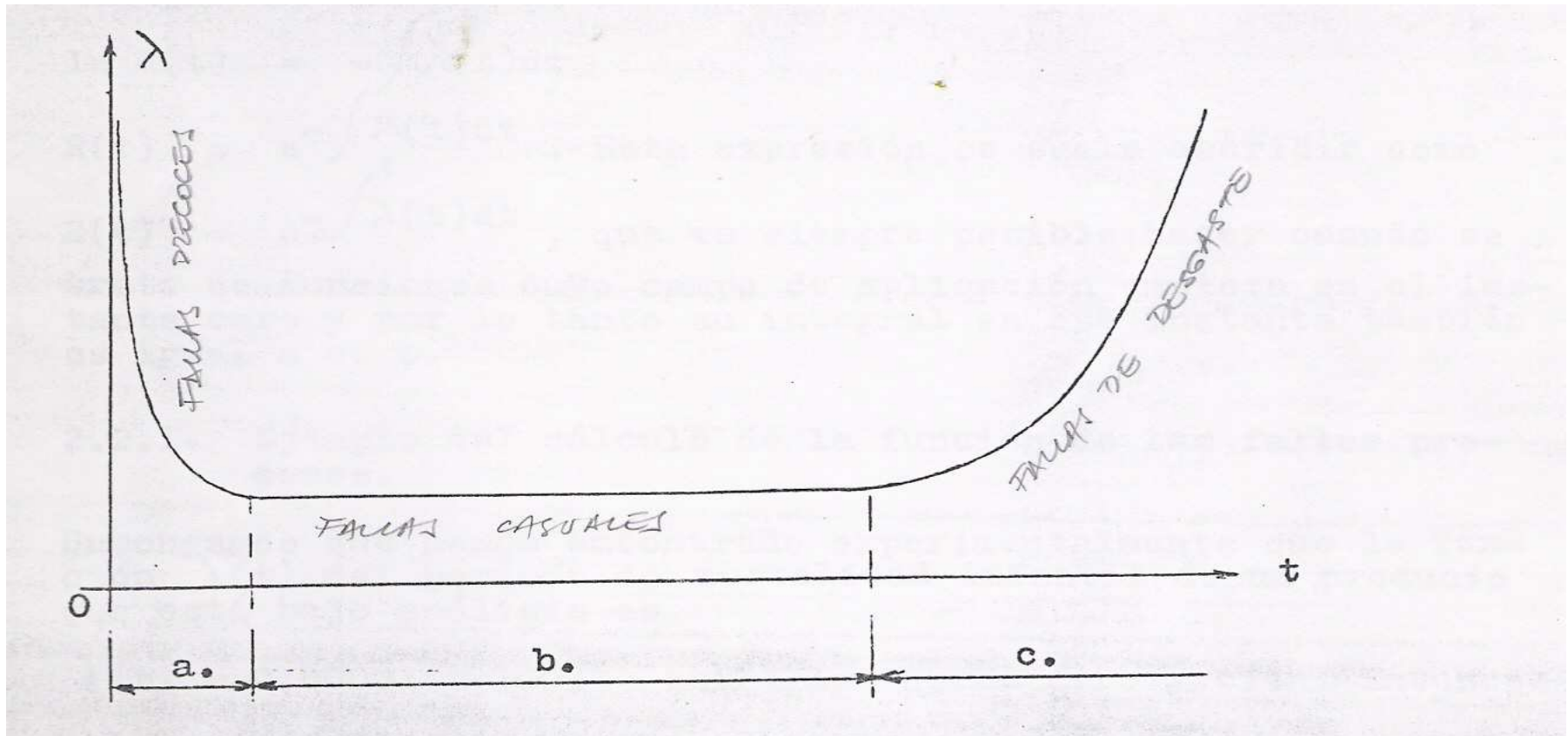
$$R_{(s)} = (0,4)(0,4)(0,4) = 0,064$$

- Cual es la fiabilidad de 6 componentes en serie si la de c/u de ellos es de 0,4

$$R_{(p)} = [(1 - 0,4)(1 - 0,4)(1 - 0,4)(1 - 0,4)(1 - 0,4)(1 - 0,4)] = 0,9533$$

Tasa de fallas variable

- THE BATHTUB CURVE



ANOVA

EJERCICIO TIPO. OK

COMP A	COMP B	COMP C	Σ
12	14	16	42
13	12	15	40
11		12	23
Σ 36	26	43	105

$N = 8$ $T = 105$
 $\Sigma x^2 = 1399$
 $FC = \frac{T^2}{N} = \frac{105^2}{8} = \frac{11025}{8} = 1378.125$
 $\Sigma x^2 - FC = 1399 - 1378.125 = 20.875$
 SUMA TOTAL DE CUADRADOS = 20.875

FUENTE	FORMULA	DESARROLLO
TOTAL	$\Sigma x^2 - FC =$	$1399 - 1378.125$
ENTRE	$\frac{36^2}{3} + \frac{26^2}{2} + \frac{43^2}{3} =$	$1386.33 - 1378.125$
	A B C	$432 + 338 + 616.33 = 1386.33$

SC	GL	CM	F _{0.05}	F _{TAB 0.05}
20.875	7	-		
8.208	2	4.104	1.62	
12.667	5	2.533		

ERROR
 * LOS GL DEL ERROR = N° DE OBSERV. - N° TRAT. = 8 - 3 = 5

Capacidad del proceso vs. Especificación del producto

- El índice de capacidad del proceso es la formula utilizada para calcular la habilidad del proceso de cumplir con las especificaciones y se expresa de

$$ICP = \frac{LSE - LIE}{6\hat{\sigma}}$$

\bar{R}

la siguiente manera

Donde: ICP: Índice de Capacidad del Proceso

LSE: Limite superior Especificado

LIE: Limite inferior Especificado

σ : Desviación estándar de los datos individuales

$$\sigma = \frac{\bar{R}}{d_2}$$

\bar{R} = Promedio de los rangos de la carta de control.
 d_2 = Constante de calculo.

ICP

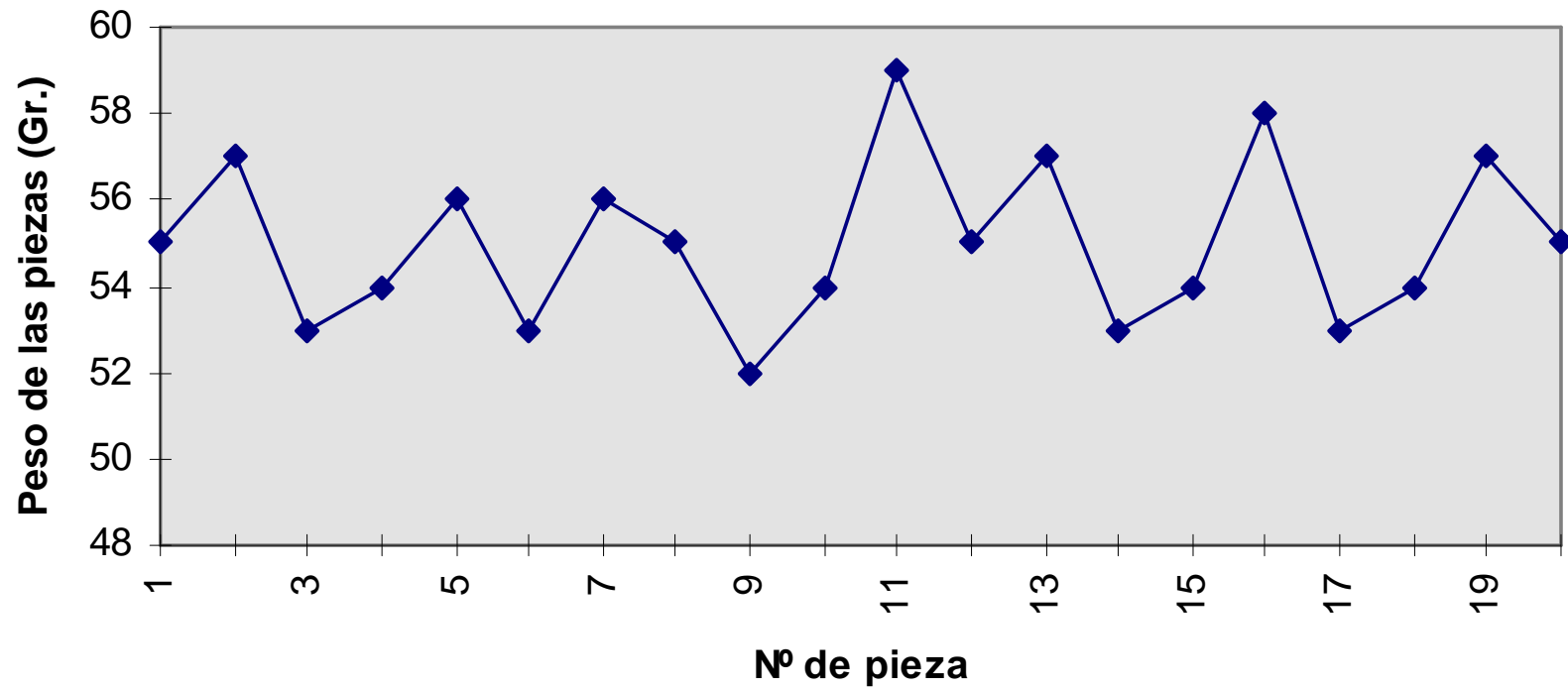
ICP	Clase de proceso	Decisión
$ICP > 1.33$	1	Más que adecuado, incluso puede exigirse más en términos de su capacidad.
$1 < ICP < 1.33$	2	Adecuado para lo que fue diseñado. Requiere control estrecho si se acerca al valor de 1
$0.67 < ICP < 1$	3	No es adecuado para cumplir con el diseño inicial.
$ICP < 0.67$	4	No es adecuado para cumplir con el diseño inicial.

Gráficos de Control

Supongamos que una serie de datos sobre el peso de piezas se registran en un gráfico de líneas en función del tiempo:

Gráficos de Control

Gráfico de las observaciones



Gráficos de Control

Cada uno de estos factores está sujeto a *variaciones* que realizan aportes más o menos significativos a la fluctuación de las características del producto, durante el proceso de fabricación.

Gráficos de Control

Causas Asignables: Son causas que pueden ser identificadas y que conviene descubrir y eliminar, por ejemplo, una falla de la máquina por desgaste de una pieza, un cambio muy notorio en la calidad del plástico, etc.

Gráficos de Control

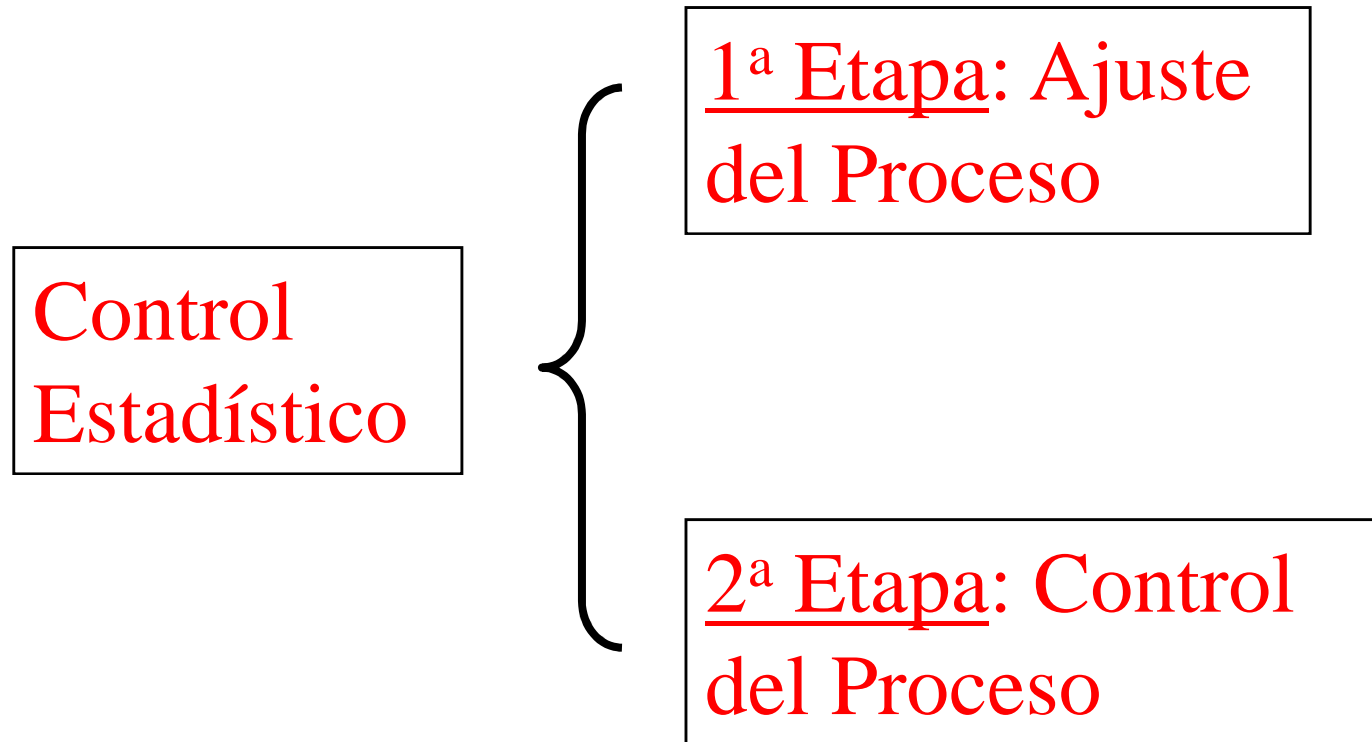
Causas No Asignables: Son una multitud de causas no identificadas, ya sea por falta de medios técnicos o porque no es económico hacerlo, cada una de las cuales ejerce un pequeño efecto en la variación total.

Gráficos de Control

Control Estadístico

¿Cómo ponerlo en marcha?

Gráficos de Control



Gráficos de Control

En la 1ª etapa se recogen unas 100-200 mediciones, con las cuales se calcula el promedio y la desviación standard:

$$\bar{X} = \frac{\sum X_i}{N}$$

$$\sigma = \sqrt{\frac{\sum (\bar{X} - X_i)^2}{N}}$$

Gráficos de Control

Luego se calculan los **Límites de Control** de la siguiente manera:

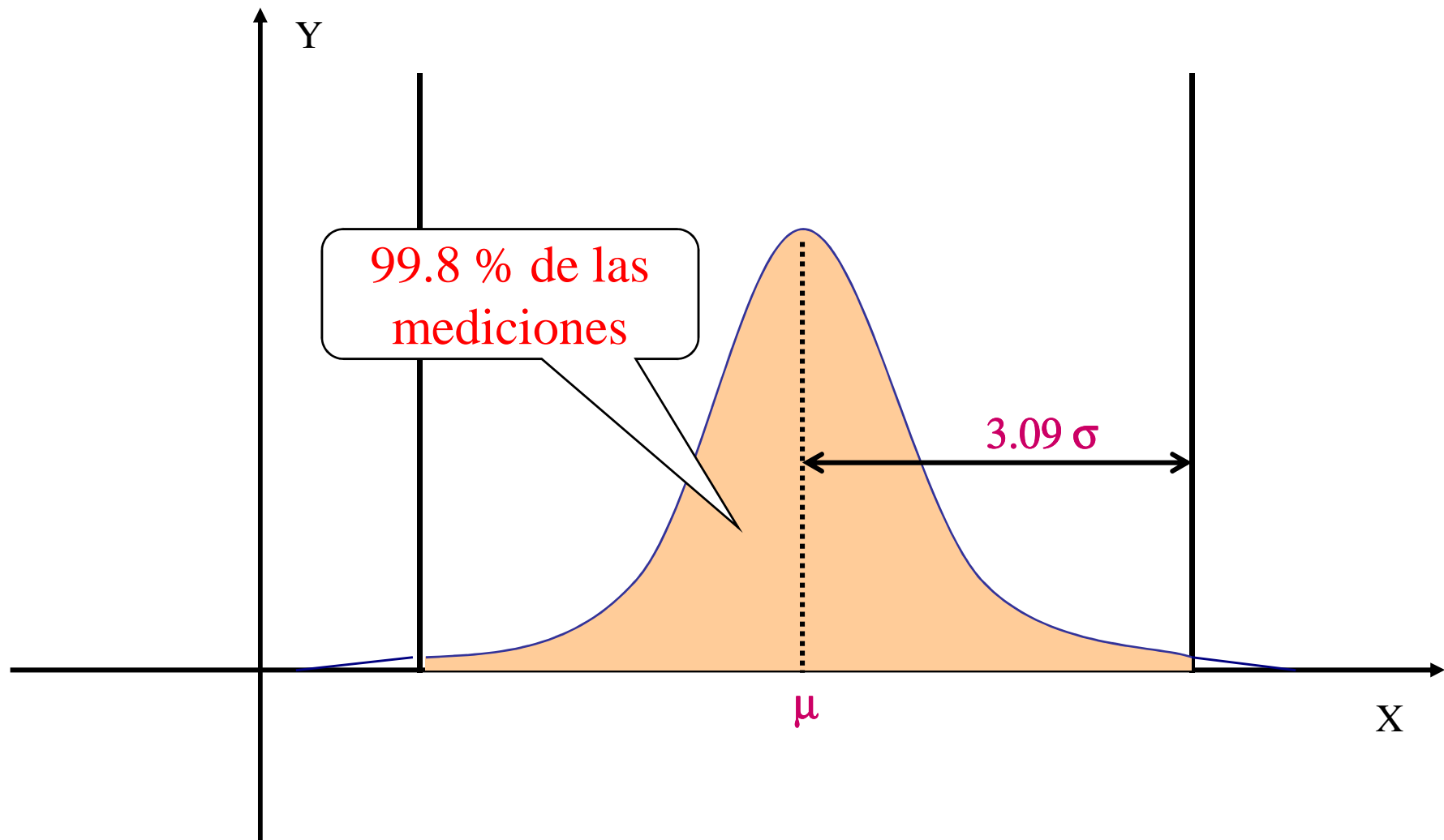
$$Lim.Superior = \bar{X} + 3.09 \cdot \sigma$$

$$Lim.Inferior = \bar{X} - 3.09 \cdot \sigma$$

Gráficos de Control

En la distribución normal, el intervalo de 3,09 sigmas alrededor del promedio corresponde a una probabilidad de 0,998.

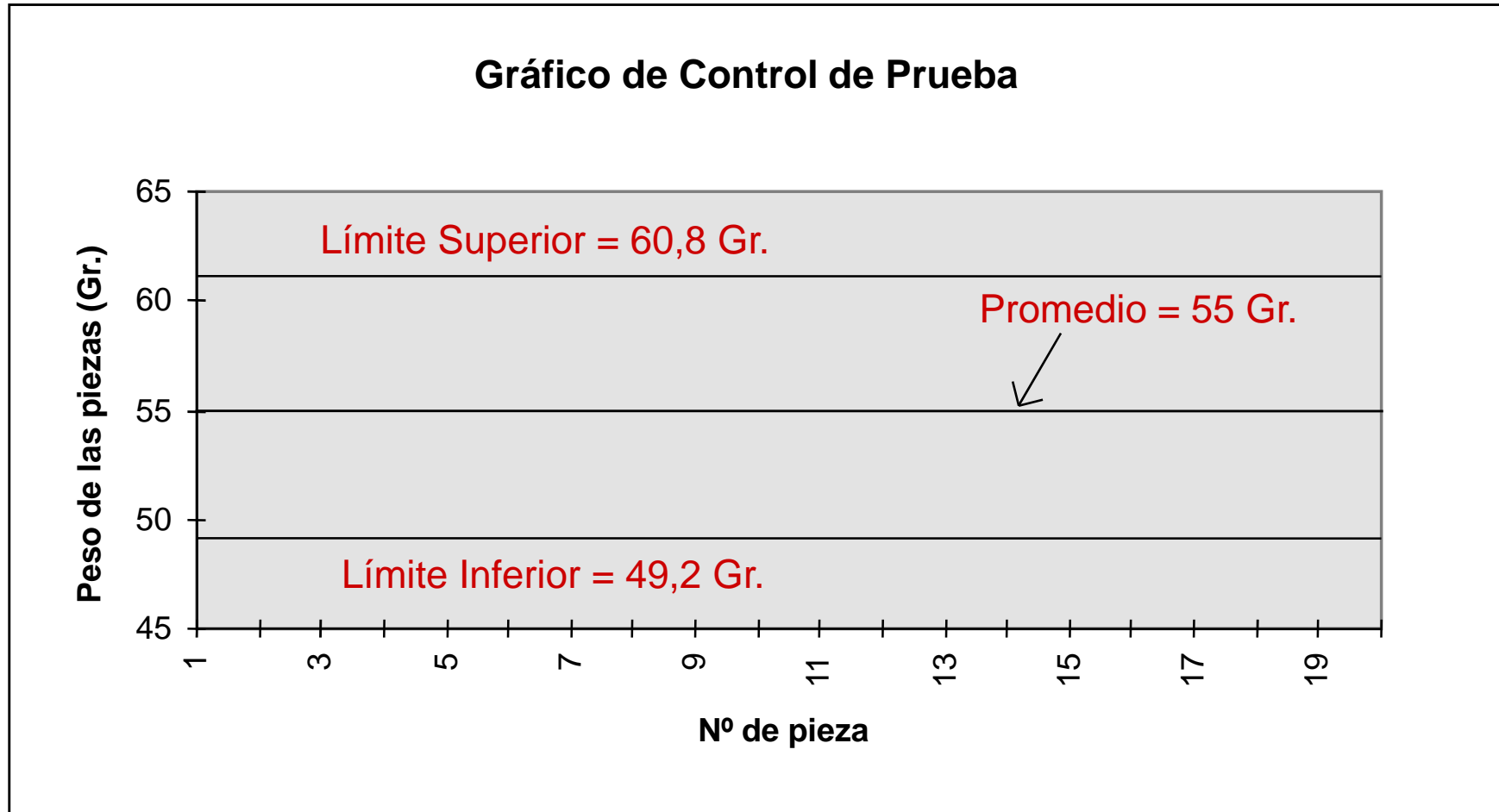
Gráficos de Control



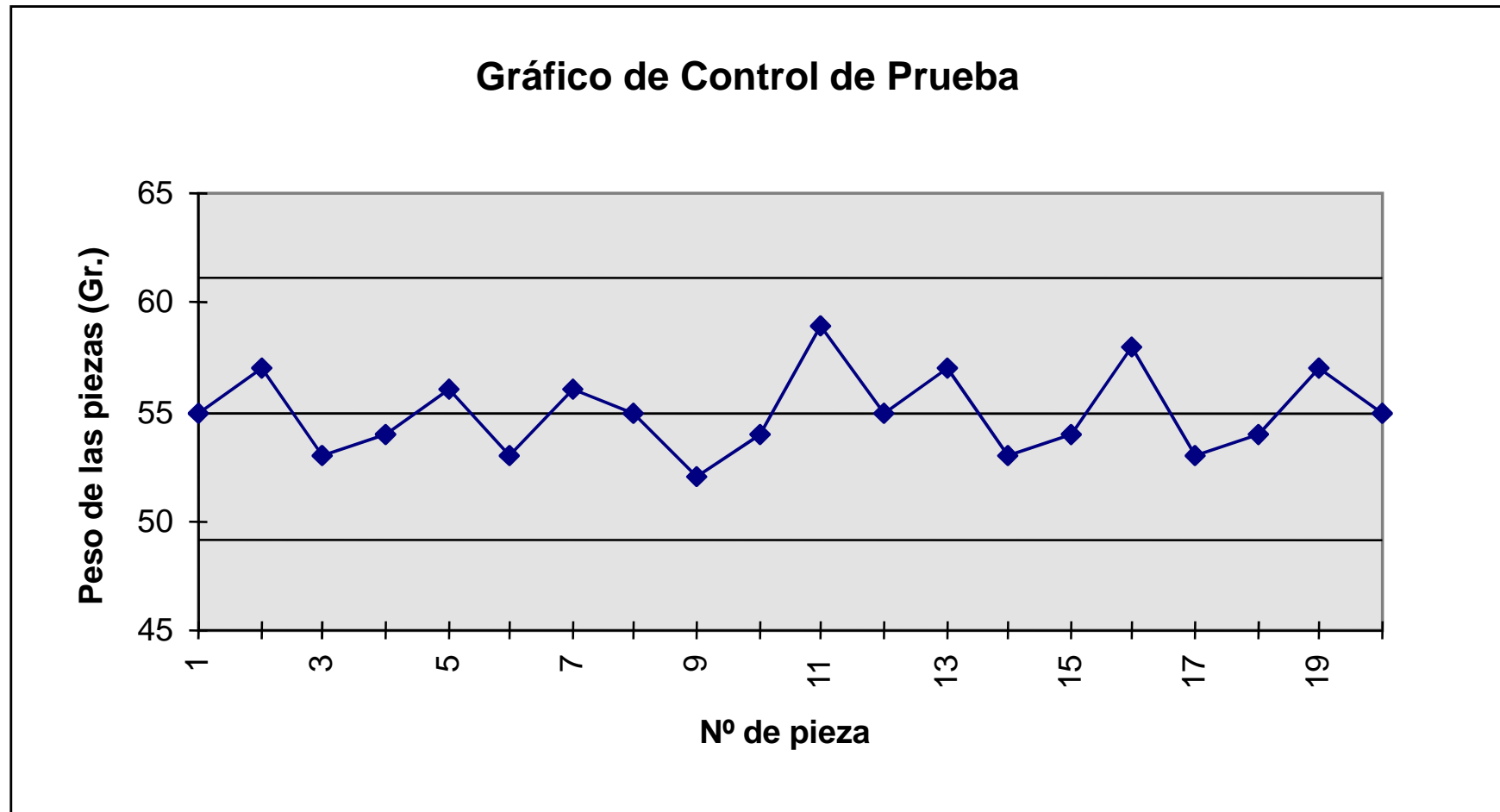
Gráficos de Control

Entonces, se construye un gráfico de prueba y se traza una línea recta a lo largo del eje de ordenadas (Eje Y), a la altura del promedio (Valor central de las observaciones) y otras dos líneas rectas a la altura de los límites de control.

Gráficos de Control



Gráficos de Control



Gráficos de Control

Este gráfico de prueba se analiza detenidamente para verificar si está de acuerdo con la hipótesis de que la **variabilidad del proceso** se debe sólo a un *sistema de causas aleatorias* o si, por el contrario, existen causas asignables de variación.

Gráficos de Control

Es difícil decir como es el gráfico de un conjunto de puntos que siguen un **patrón aleatorio de distribución normal**, pero sí es fácil darse cuenta cuando no lo es.

Gráficos de Control

Veamos algunos ejemplos de patrones
No Aleatorios:

Gráficos de Control

En la 2ª etapa, las nuevas observaciones que van surgiendo del proceso se representan en el gráfico, y se controlan verificando que estén dentro de los límites, y que no se produzcan patrones no aleatorios:

Gráficos de Control

Entonces, cuando se encuentra más de 1 punto en 500 fuera de los límites de control, esto indica que el sistema de causas aleatorias que provocaba la variabilidad habitual de las observaciones ha sido alterado por la aparición de una causa asignable que es necesario descubrir y eliminar.