FACULTAD DE INGENIERÍA UNAM
DIVISIÓN DE EDUCACIÓN CONTINUA

CURSOS INSTITUCIONALES

ESTADISTICA MEDICA

Del 13 al 24 de Marzo de 2006

APUNTES GENERALES

CI - 013

Instructor: Lic. Servando R. Martínez García
SERVICIOS DE SALUD PUBLICA DEL DISTRITO FEDERAL
MARZO DE 2006

Palacio de Minería, Calle de Tacuba No 5, Primer piso, Delegación Cuauhtémoc, CP 06000, Centro Histórico, México D.F., APDO Postal M-2285 • Tels: 5521.4021 al 24, 5623.2910 y 5623.2971 • Fax 5510.0573

CONTENIDO

OBJETIVO

CONOCER Y DESARROLLAR HABILIDADES DE DISEÑO DE INSTRUMENTOS PARA LA RECOLECCION, ORGANIZACIÓN Y ANALISIS DE DATOS, PERFILANDOLOS HACIA LA TOMA DE DECISIONES BASADA EN HECHOS REALES Y EN LA GESTION CON CALIDAD.

COMPROMISOS DE LOS PARTICIPANTES

1. REALIZAR PROPUESTAS DE MEJORA EN EL AMBITO LABORAL DEL PARTICIPANTE EN LOS TEMAS DE CALIDAD EN EL SERVICIO, DE CALIDAD DE VIDA EN EL TRABAJO, DEL MODELO DE CALIDAD DE LA INSTITUCION O DE LAS TECNICAS Y PROCEDIMIENTOS ESTADISTICOS UTILIZADOS.

2. ELABORAR UN GLOSARIO DE TERMINOS DE ESTADISTICA EN GENERAL, Y ESPECIALIZARLO CON CONCEPTOS PROPIOS DEL SECTOR SALUD.

3. FIJAR, MEDIANTE LA ELABORACIÓN DE UN MANUAL O CATALOGO, LAS TECNICAS ESTADISTICAS VISTAS EN EL CURSO/TALLER Y QUE PUEDA APOYAR LA LABOR DEL ESTADIGRAFO.

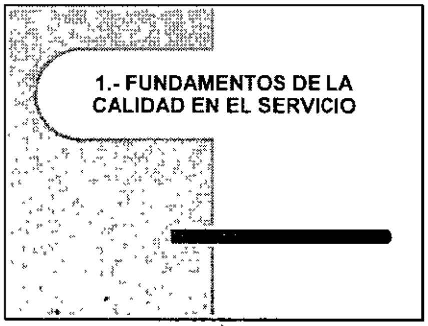
4. DESARROLLAR HABILIDADES CUANTITATIVAS Y HACERLAS EVIDENTES MEDIANTE EL ESTUDIO DE CASOS REALES Y APLICABLES A SU TRABAJO.

EVALUACIÓN

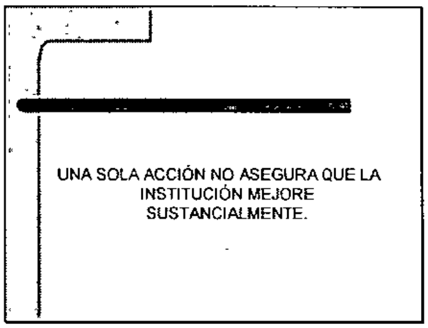
Para acreditar el curso el participante debe asistir cuando menos al 80% de las sesiones de estudio y seminarios tutoriales y obtener como mínimo 8.0 de calificación en las evaluaciones académicas, mismas que se integraran de la siguiente manera;

PARTICIPACIONES	20%
EVALUACIONES POR SESION	20%
TRABAJO FINAL	20%
EXAMEN FINAL	40%

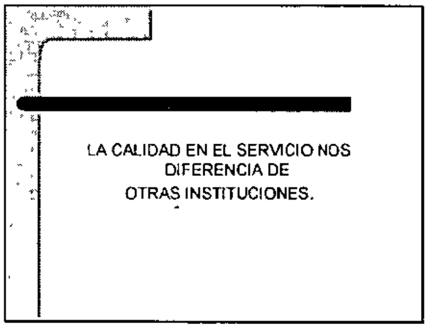
Diapositiva 1



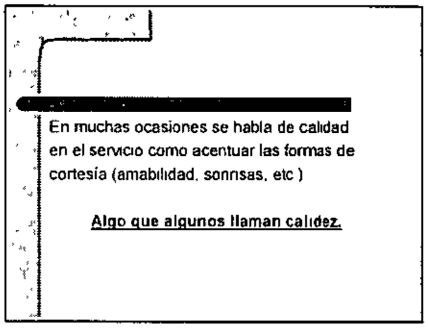
Diapositiva 2



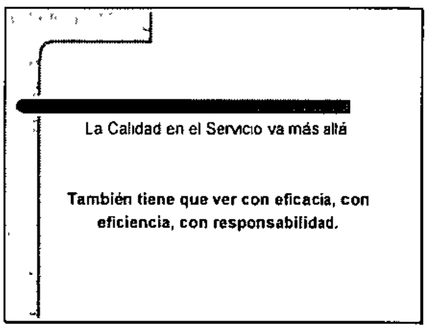
Diapositiva 3



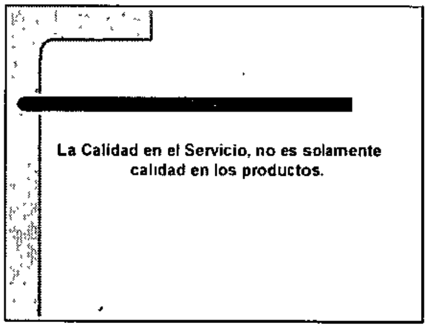
Diapositiva 4



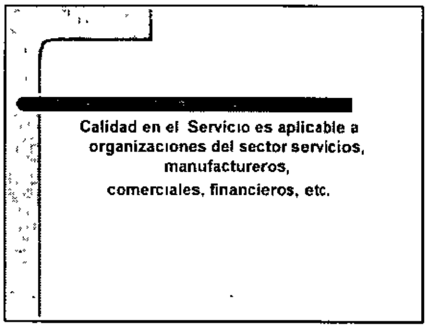
Diapositiva 5



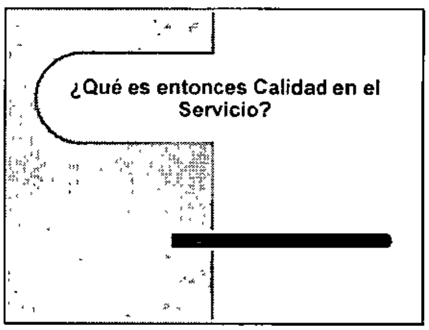
Diapositiva 6



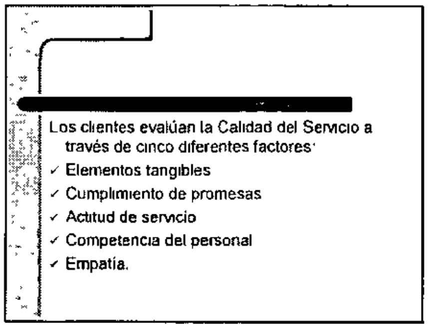
Diapositiva 7



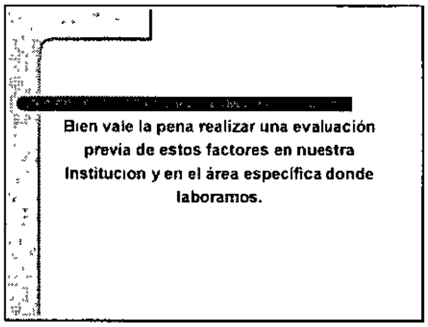
Diapositiva 8



Diapositiva 9

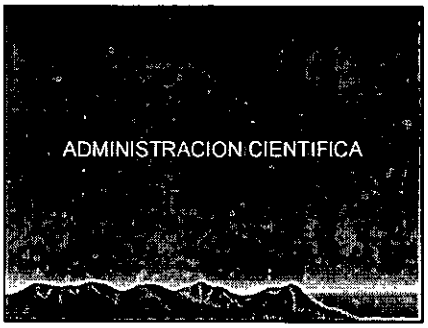


Diapositiva 10

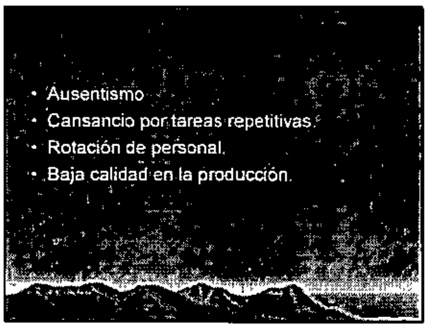


Diapositiva 1

Diapositiva 2



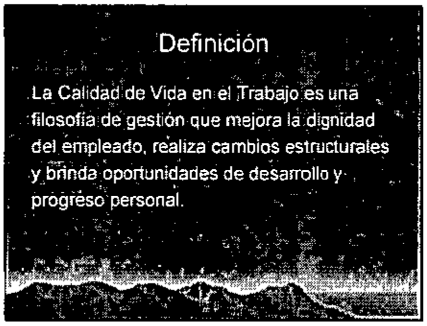
Diapositiva 3



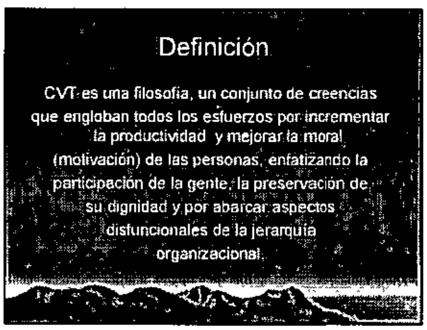
Diapositiva 4

Diapositiva 5

Diapositiva 6

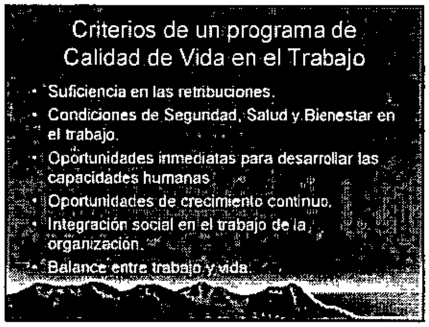


Diapositiva 7



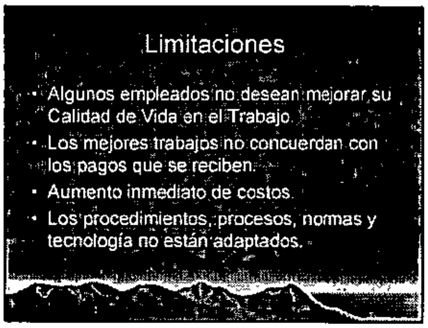
Diapositiva 8

Diapositiva 9



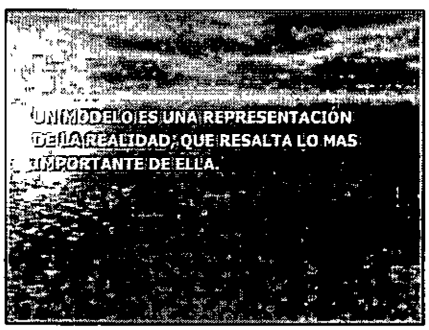
Diapositiva 10

Diapositiva 11



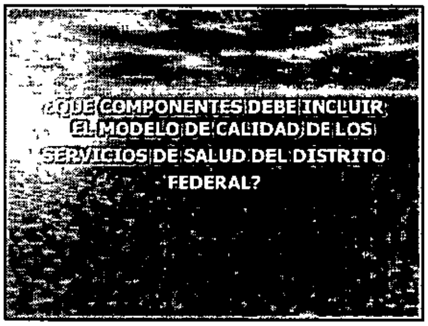
Diapositiva 1

Diapositiva 2



Diapositiva 3

Diapositiva 4



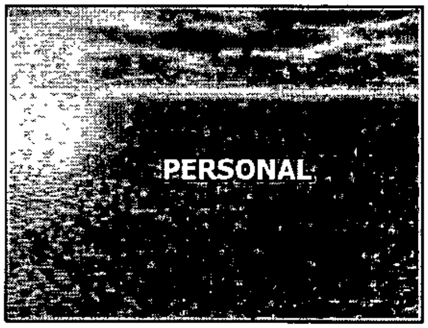
Diapositiva 5

Diapositiva 6

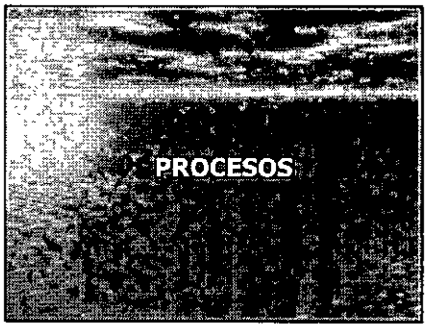
Diapositiva 7

Diapositiva 8

Diapositiva 9

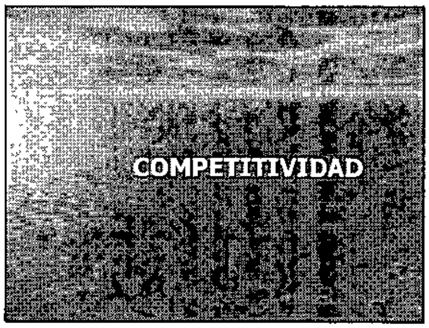


Diapositiva 10



Diapositiva 11

Diapositiva 12



Diapositiva 13

ESTADÍSTICA

ÍNDICE

I. INTRODUCCION

CLASIFICACIÓN DE LA ESTADÍSTICA

> **Estadística descriptiva.**

- Consiste en la recopilación, organización y presentación de datos.
- Permite describir los hechos relacionados con la información recopilada, a través de tablas, gráficas, cuadros e índices.

> **Inferencia estadística.**

- Es una técnica mediante la cual, a partir de los datos de una *muestra,* se obtienen conclusiones o generalizaciones acerca de las características de la *población o universo.*

POBLACIÓN Y MUESTRA

Población o universo.

- Es la totalidad de posibles observaciones o medidas que comprende una situación dada. Ejemplo, los gastos en infraestructura en todos los municipios del país.
- Puede ser:
  - > Finita, cuando la población tiene un número limitado de elementos; por Ej., el número de posibles votantes en una elección.
  - > Infinita, cuando la población consta de un número infinito de elementos; por Ej., el experimento de medir la seguridad al transitar un puente, ya que, hipotéticamente, se pueden realizar al respecto una infinidad de mediciones.

> Muestra

- Es una colección de observaciones tomadas de una población dada. Ejemplo, los gastos en infraestructura de algunos municipios seleccionados.

II. ORGANIZACIÓN DE LOS DATOS

DISTRIBUCIONES DE FRECUENCIA

> El análisis de un grupo de datos requiere, en primera instancia, el ordenamiento de los mismos.

> Una ordenación es una colocación de los datos en orden de magnitud, ya sea creciente o decreciente.

> Tal ordenación permite conocer la amplitud de los datos (la diferencia entre el valor mínimo y el valor máximo) y tener una idea general sobre su comportamiento.

> Sin embargo, si el número de datos es elevado, es necesario dividirlos en clases o categorías para poder analizarlos.

> La disposición de los datos de manera que muestren la frecuencia con que se dan los valores en cada clase se llama *distribución de frecuencia* y el número de datos pertenecientes a cada clase, es la *frecuencia de clase.*

DISTRIBUCIONES DE FRECUENCIA

^Ejemplo. **Construcción de una distribución de frecuencia.**

- **Si se tienen los datos sobre Asignación del Fondo 4, del Ramo 33 (R33) en cincuenta municipios (Mun.) en millones de pesos:** *Cuadro 1.*

Mun.	R33	Mun.	R33	Mun.	R33	Mun.	R33	Mun.	R33
1	65	11	64	21	64	31	63	41	64
2	63	12	65	22	64	32	65	42	64
3	65	13	64	23	64	33	63	43	63
4	63	14	72	24	71	34	70	44	69
5	69	15	68	25	68	35	67	45	67
6	67	16	66	26	66	36	66	46	66
7	53	17	55	27	56	37	57	47	58
8	58	18	57	28	59	38	59	48	60
9	60	19	60	29	61	39	61	49	61
10	61	20	62	30	62	40	62	50	62

Ejemplo.

Se debe primero ordenar los datos, por ejemplo, en forma ascendente: *Cuadro 2.* Ordenación de la asignación del R33 en 50 municipios

Ord.	R33	*Ord.*	R33	*Ord.*	R33	*Ord.*	R33	*Ord.*	R33
1	53	*11*	60	*21*	63	*31*	64	*41*	67
2	55	*12*	60	*22*	63	*32*	65	*42*	67
3	56	*13*	61	*23*	63	*33*	65	*43*	67
4	57	*14*	61	*24*	63	*34*	65	*44*	68
5	57	*15*	61	*25*	63	*35*	65	*45*	68
6	58	*16*	61	*26*	64	*36*	65	*46*	69
7	58	*17*	62	*27*	64	*37*	66	*47*	69
8	59	*18*	62	*28*	64	*38*	66	*48*	70
9	59	*19*	62	*29*	64	*39*	66	*49*	71
10	60	*20*	62	*30*	64	*40*	66	*50*	72

Ejemplo.

- Posteriormente, se definen los intervalos de clase.
- Los intervalos de clase no deben ser demasiado pequeños, porque cada clase podría contener pocos datos o incluso ninguno, ni demasiado grandes, ya que podrían agruparse datos con diferencias importantes en una misma clase. En general, las distribuciones contienen como mínimo 6 clases ó, como máximo, 20. Es muy común considerar 10 clases, con resultados satisfactorios.
- Los intervalos de clase deben ser de igual amplitud, para que sea posible la comparación entre clases.
- En el ejemplo, si se agrupan los datos en 8 clases, dado que la amplitud es de 19 (72-53), cada intervalo tendría una amplitud mínima de 19/8=2.375.

Ejemplo.

Sin embargo, para evitar que el primer y el último dato se ubiquen en los límites de los intervalos de clase, éste se amplía.

Se pude por ejemplo definir un intervalo de 3, e iniciar la primera clase a partir del valor 50.5.

Una vez definidas las clases, se identifica la cantidad de datos que pertenecen a cada clase, es decir, las frecuencias de clase, y con ello se tiene ya la distribución de frecuencias.

La distribución de frecuencias del ejemplo desarrollado se presenta en el ***Cuadro* 3.**

Se puede elaborar una distribución de frecuencias relativas, si en lugar de cuantificar el número de observaciones de cada clase, se calcula la participación de éstas en el total de observaciones.

Ejemplo.

***Cuadro* 3.**

Distribución de frecuencias de la asignación del R33.

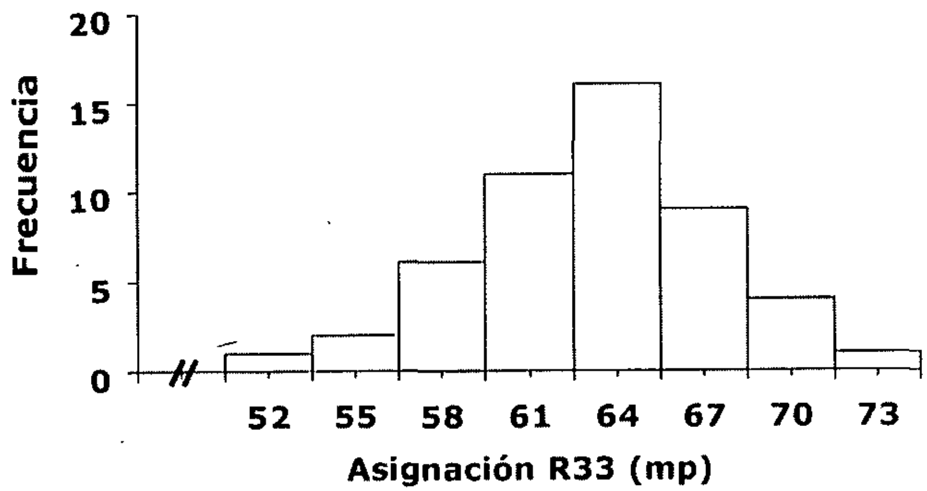
CLASES (R33, mp)	FRECUENCIA (Número de os)	Recuento	FRECUEN-ATIVA
50.5-53.5	1	/	.02
53.5-56.5	2	//	.04
56.5-59.5	6	//////	.12
59.5-62.5	11		.22
62.5-65.5	16		.32
65.5-68.5	9		.18
68.5-71.5	4		.08
71.5-74.5	1		.02
Total	50		1.00

Las distribuciones de frecuencias se pueden presentar en gráficas como el **histograma, el polígono de frecuencia, y la ojiva.**

REPRESENTACION GRAFICA DE LAS DISTRIBUCION DE FRECUENCIA

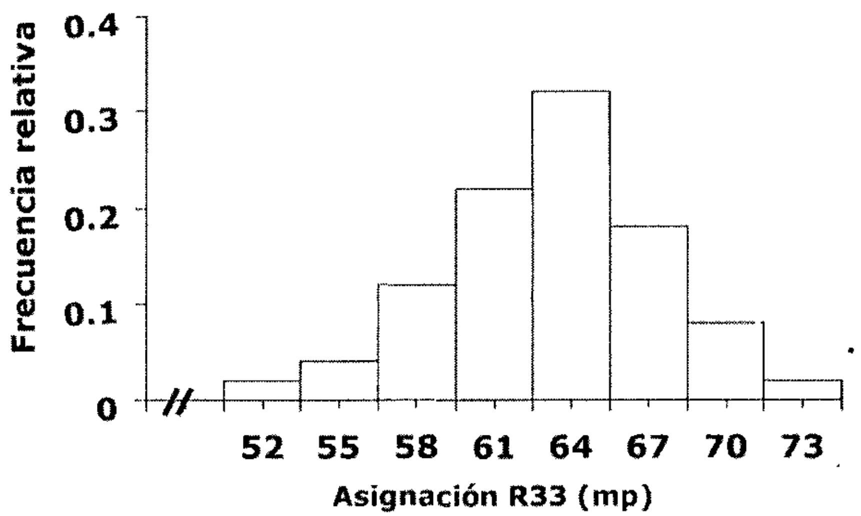
HISTOGRAMA

- Se grafican los intervalos de clase en el eje de las X, y las frecuencias (número de datos que corresponden a cada clase), en el eje de las Y. Para cada intervalo se dibuja una barra cuya altura es la frecuencia de la clase y se marca, en el eje de las X, el punto medio de cada clase (marca de clase).



Histograma

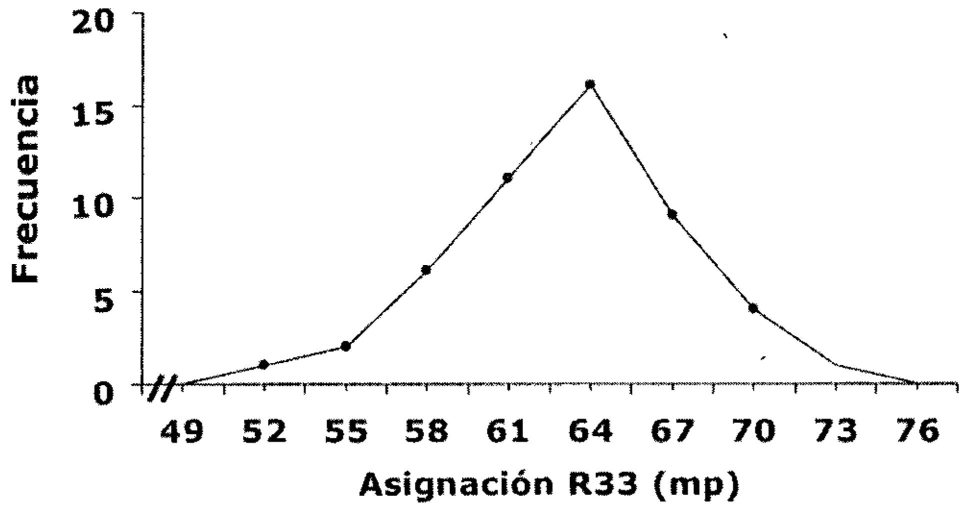
- Si en el eje de las Y se grafica la frecuencia *relativa*, s construye entonces un *Histograma de frecuencias relativas*



- De esta forma, se identifica rápidamente la importanci *relativa* de cada clase. Por ejemplo, más de una tercer parte de los municipios reciben una asignación de entr 62.5 y 65.5 millones de pesos.

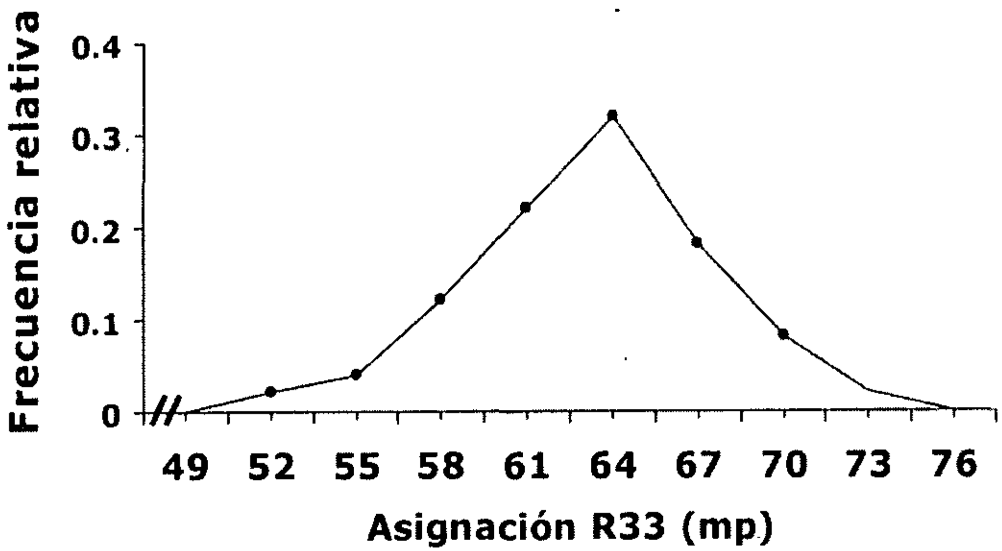
REPRESENTACION GRAFICA

- **POLÍGONO DE FRECUENCIA**
  - Se construye con los mismo ejes que el histograma, y se obtiene graficando un punto en cada marca de clase, a la altura de la frecuencia de esa clase, y uniendo los puntos correspondientes.



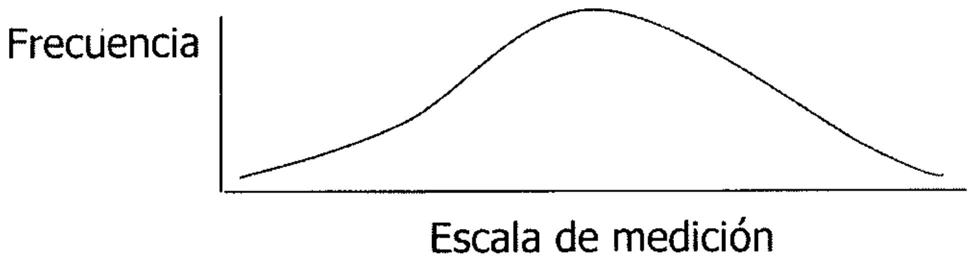
Polígono de frecuencia

- La gráfica correspondiente a las frecuencias *relativas* se denomina *Polígono de frecuencias relativas*.

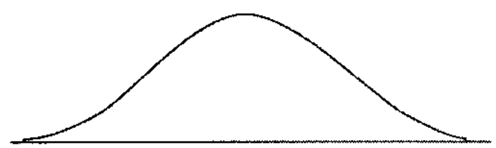


Polígono de frecuencia

- Si se construye un polígono de frecuencia suavizando la unión entre los puntos, se aproxima el aspecto que éste tendría si se tuviera un número infinito de observaciones de datos, y clases de intervalo pequeñas. Tal representación gráfica se conoce como *Curva de frecuencia*.



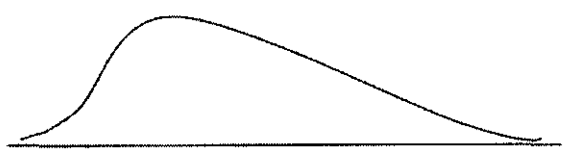
- Una curva de frecuencia puede tener una distribución:
  - <u>Simétrica</u>: la mayor frecuencia (valor modal) se encuentra en el centro de los datos (mediana).



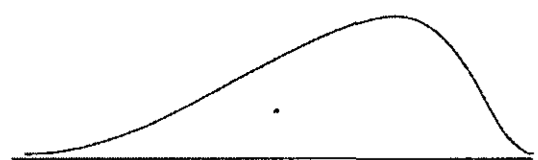
Distribución simétrica

Polígono de frecuencia

- Asimétrica o sesgada: los valores de sus distribuciones de frecuencia se concentran en el extremo inferior o superior de la escala de medición. Si la frecuencia de los valores disminuye hacia el extremo superior de la escala, se dice que la curva tiene asimetría positiva o que está sesgada a la derecha (la cola más larga de la distribución queda a la derecha). Por el contrario, si la frecuencia disminuye hacia el extremo inferior de la escala, la curva tiene asimetría negativa o está sesgada a la izquierda.



Distribución asimétrica positiva
(sesgada a la derecha)



Distribución asimétrica negativa
(sesgada a la izquierda)

REPRESENTACIÓN GRÁFICA

- **OJIVA (o polígono de frecuencia acumulada)**

> Es un gráfico que presenta las frecuencias acumuladas, es decir, cuántas observaciones se hallan por debajo de ciertos valores. Los valores que se utilizan son los límites de cada intervalo de clase.

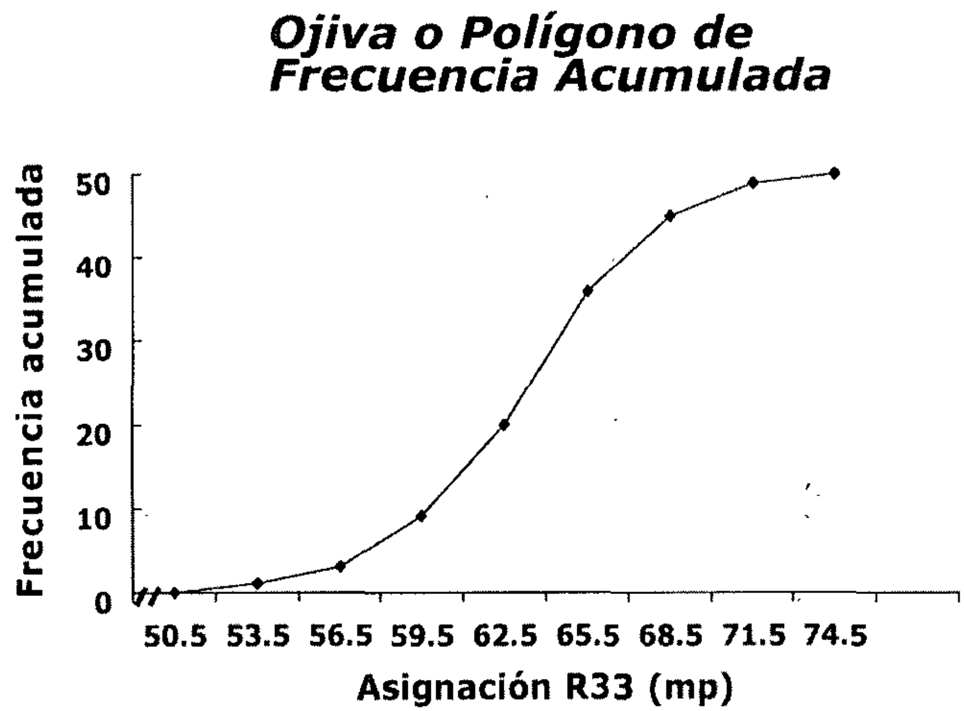
> La gráfica correspondiente se construye a partir de los datos calculados en el ***Cuadro 4.a.***

> Si en lugar de considerar el número de observaciones que se encuentran por debajo de ciertos valores, se calcula la participación de dichas observaciones en el total, se obtiene lo que se conoce como *Polígono de frecuencia relativa acumulada.* Este construye a partir de los datos del ***Cuadro 4.b.***

Ojiva

Cuadro 4.a.

R33 (mp) Menor que:	**FRECUENCIA ACUMULADA** (Número de municipios)
50.5	0
53.5	1
56.5	3
59.5	9
62.5	20
65.5	36
68.5	45
71.5	49
74.5	50

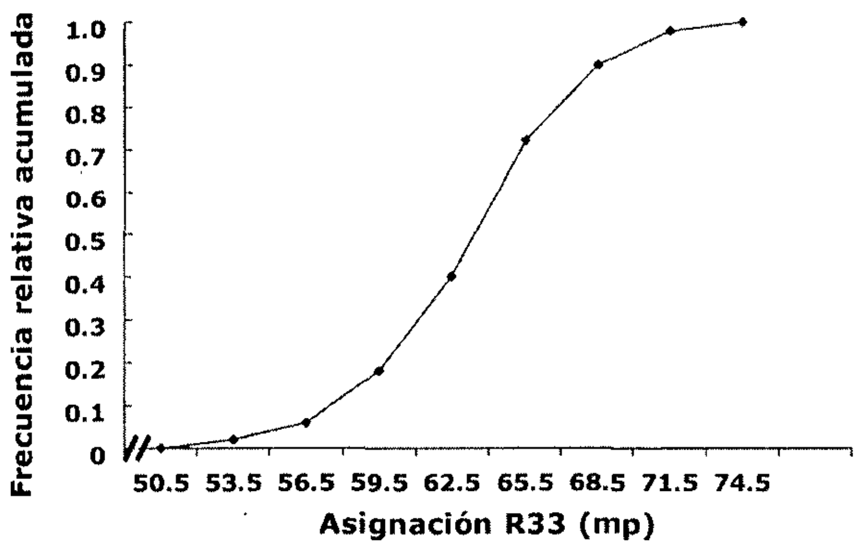


Ojiva

Cuadro 4.b.

R33 (mp) Menor que:	FRECUENCIA RELATIVA ACUMULADA
50.5	.00
53.5	.02
56.5	.06
59.5	.18
62.5	.40
65.5	.72
68.5	.90
71.5	.98
74.5	1.00

Ojiva porcentual o Polígono de Frecuencia Relativa Acumulada



- La ojiva porcentual nos permite una interpretación rápida de los datos. Por ejemplo, se observa que más del 70% de los municipios estudiados reciben aportaciones inferiores a 65.5 millones de pesos.

III. MEDIDAS DE TENDENCIA CENTRAL Y DE DISPERSION.

MEDIDAS DE TENDENCIA CENTRAL

> Al calcular un promedio, se obtiene un valor que es representativo del conjunto de datos considerado. Como los valores promedio tienen a situarse en el centro del la serie de datos ordenados según su magnitud, éstos se conocen también como medidas de tendencia central.

> Las principales medidas de tendencia central son la **media aritmética, la mediana** (y sus medidas asociadas, como cuartil, decil y percentil), y la **moda.**

MEDIDAS DE TENDENCIA CENTRAL

➢ MEDIA ARITMÉTICA

- Es la suma de los valores observados, dividida entre el número total de observaciones.
- Si $X_1, X_2, \ldots, X_n$ son los valores de una muestra, y 'n' el número total de observaciones, la media aritmética ($\overline{X}$) se calcula:

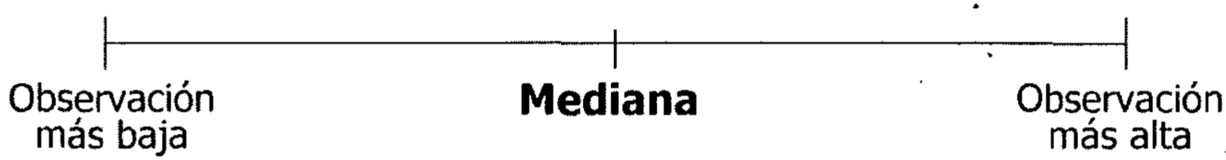
$$\overline{X} = \frac{X_1 + X_2 + \ldots + X_n}{n} = \frac{\sum_{i=1}^{n} X_i}{n}$$

- Cuando la media aritmética corresponde a la media de una población y no de una muestra, se emplea el signo μ.

MEDIDAS DE TENDENCIA CENTRAL

➢ Mediana

- La mediana es medida de posición que ubica el valor central de un grupo de datos ordenados.



- En general, si $X_1, X_2, \ldots X_n$ son números ordenados por magnitud creciente (o decreciente), y n es impar, la mediana es el número situado en el centro del conjunto de números:

$$\text{Med} = X_{(n+1)/2}$$

- Si n es par, la mediana viene dada por:

$$\text{Med} = \frac{X_{n/2} + X_{(n/2)+1}}{2}$$

Mediana.

Ejemplo. En el ejemplo desarrollado, los datos indican que n es par (n = 50), y por lo tanto la mediana se obtiene a partir del dato #25 (50/2) y del dato #26 ((50/2)+1)). Los valores correspondientes son:

$$Med= \frac{X_{25} + X_{26}}{} = \frac{63 + 64}{} = 63.5$$

A diferencia de la media, la mediana tiene la ventaja de que los valores extremos le afectan en menor medida; además, puede ser calculada para datos cualitativos, cuando estos pueden tener algún ordenamiento. Sin embargo, dado que la mediana es un promedio de posición, su cálculo requiere siempre contar con los datos ordenados.

MEDIDAS DE TENDENCIA CENTRAL

Cuartiles, deciles, percentiles

> Otras medidas asociadas a la mediana, dado que se basan también en su posición en una serie de observaciones, son los cuartiles, deciles y percentiles

> Así como la mediana es el valor que divide una serie de datos (colocados en orden de magnitud) en dos, los valores que dividen los datos en cuatro partes iguales se denominan **cuartiles.**

> Los cuartiles se representan por Q_{lr} Q_2 y Q_3 e indican el primer, segundo y tercer cuartil, respectivamente.

> El segundo cuartil, Q_2, equivale a la mediana. Para un conjunto de datos, tendríamos entonces:

Observación	ler. Cuartil	2o. Cuartil	3er. Cuartel	Observación
más baja	Q_1	Q_2	$\mathbf{Q_3}$	más alta

Cuartiles, deciles, percentil

Análogamente, los valores que dividen los datos en diez partes iguales se llaman **deciles** y se representan por D_1 D_2, ..., D_9, mientras que los valores que dividen los datos en cien partes iguales se llaman **percentiles,** y se representan por P1, P2, ..., P99.

El quinto decil y el quincuagésimo percentil corresponden a la mediana.

MEDIDAS DE TENDENCIA CENTRAL

> **Moda**

Es aquél valor que se presenta con mayor frecuencia en una serie de datos.

Una distribución que tiene una sola moda se llama unimodal; si presenta más de una, se conoce como multimodal.

Cuando se analizan datos que no son números, como por ejemplo, la ocupación de un grupo de personas, no se pueden obtener valores como la *media,* y sólo ocasionalmente se puede obtener la *mediana.* Entonces se puede utilizar la *moda* como una medida de tendencia central; en el caso de las ocupaciones, aquélla a la que se dedique el mayor número de personas sería la ocupación *modal.*

La moda, al igual que la mediana, es poco sensible a los valores extremos.

Moda.

- Generalmente, la moda se estima a partir de datos agrupados. Esto es así porque puede resultar que el valor que más se repita en un conjunto de datos no sea representativo de los mismos, en cuyo caso, la moda sería una medida poco útil para el análisis de la información.
- La estimación se realiza identificando primero lo que se conoce como *clase modal*, que es aquélla con la mayor frecuencia de la distribución, es decir, la clase que registra el mayor número de datos. Identificada la clase modal, la moda se estima con la siguiente ecuación:

$$Mo = L_{Mo} + \frac{d_1}{d_1 + d_2} w$$

donde:

LMo = límite inferior de la clase modal.

d_1 = frecuencia de la clase modal menos frecuencia de la clase premodal.

d_2 = frecuencia de la clase modal menos frecuencia de la clase postmodal.

W = amplitud del intervalo de la clase modal.

Moda.

- **Ejemplo.** De acuerdo con los datos del ***Cuadro 3***, el valor modal sería 64, ya que este es el número con mayor frecuencia (6 veces).
- Sin embargo, la estimación a partir de los datos agrupados se realiza (utilizando la información del ***Cuadro 3***) de la siguiente forma:
  - Se identifica la clase modal de la distribución como la clase 62.5 a 65.5, ya que contiene más frecuencias que cualquier otra.
  - Se calculan los valores:

    Lmo = 62.5 $d_1 = 16-11 = 5$

    w = 3 $d_2 = 16-9 = 7$
  - Se sustituyen los valores en la fórmula:

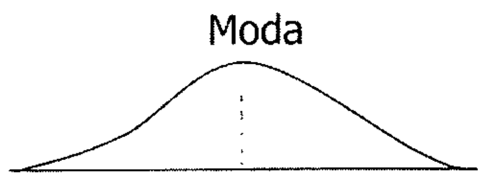
$$Mo = 62.5 + \frac{5}{5+7}3 = 62.5 + \frac{15}{12} = 62.5 + 1.25 = 63.75$$

Moda.

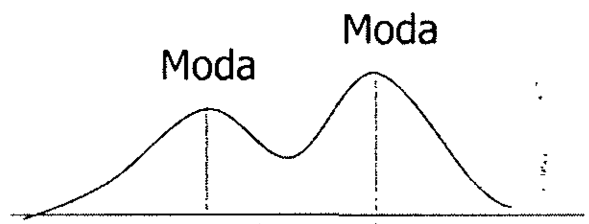
- Esta aproximación del valor modal tiene las siguientes limitaciones:
  - > El valor modal estimado en una distribución muy asimétrica está demasiado cercano a un extremo de los datos, por lo que no es un buen representante de la serie.
  - > La localización de la clase modal, y por consiguiente el valor de la moda, depende de las maneras como se hayan clasificado los datos.
- Es importante además considerar que cuando la distribución es multimodal, la moda como medida de tendencia central pierde utilidad.
- Por lo anterior, la moda es la medida de tendencia central menos utilizada.

Moda.

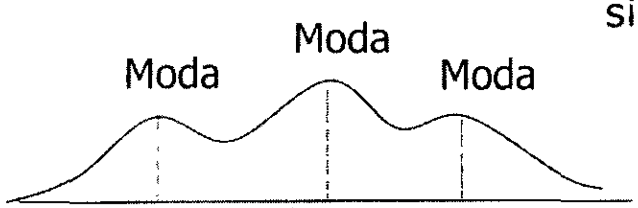
- Representación gráfica de distintas distribuciones de frecuencia, de acuerdo con el número de modas.



Distribución unimodal
(hay un solo valor modal)



Distribución bimodal
(existen dos valores que tienen significativamente mayor frecuencia que el resto de los valores)

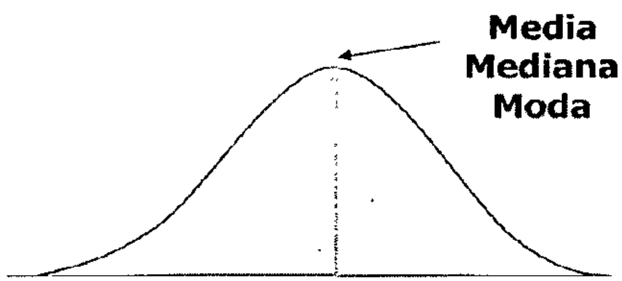


Distribución multimodal
(más de dos valores presentan una frecuencia superior a los demás valores)

MEDIDAS DE TENDENCIA CENTRAL

➢Comparación gráfica de las medidas de tendencia central.

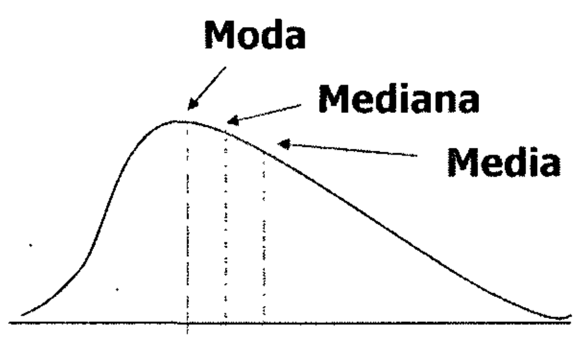
- Los valores relativos de las medidas de tendencia central, van a depender de la asimetría de la distribución.
- Si la distribución es simétrica, las tres medidas de tendencia central tienen valores idénticos. Por lo tanto, si los datos pertencen a una distribución de este tipo o se asemejan a ella, cualquiera de los promedios es útil para caracterizar los datos.



Distribución simétrica

Comparación ...

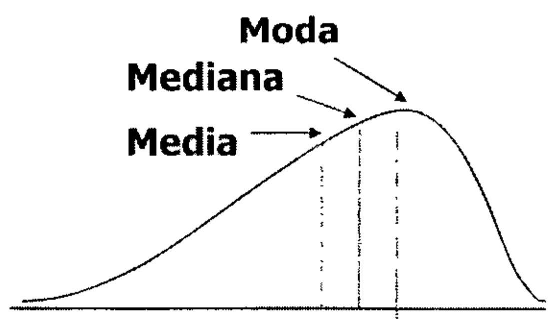
- Si la distribución es de asimetría positiva:
  - La media y la mediana se encuentran sesgadas a la izquierda, ya que existe un mayor número de datos hacia los valores inferiores de la distribución. Sin embargo, el sesgo de la media es menor, dado que ésta es más sensible a los valores extremos.
  - La moda por su parte, se ubica en el pico más alto, y a la izquierda de ambas.



Asimetría positiva

Comparación ...

- Si por el contrario, la distribución es de asimetría negativa:
  - La media y la mediana se ubican hacia la derecha de la curva, pero con menos desplazamieno de la media, y la moda corresponde al pico más alto de la curva



Asimetría negativa

- Cuando la distribución es asimétrica, puede ser más adecuado utilizar la moda o la mediana como medidas de tendencia central, ya que la media no sería un indicador representativo del conjunto de datos.

MEDIDAS DE DISPERSIÓN

> Al grado en que los datos numéricos tienden a variar respecto a un valor medio se le llama variación o dispersión de los datos.

> Una menor (mayor) dispersión de los datos indica que hay más (menos) uniformidad entre ellos.

> Las principales medidas de dispersión o variación son **rango o amplitud, desviación media, varianza y desviación estándar.**

MEDIDAS DE DISPERSIÓN

> RANGO O AMPLITUD

Es la diferencia entre el valor máximo y el valor mínimo de un conjunto de números ordenados.

Es una medida útil en algunos casos, especialmente cuando el objetivo de la investigación es conocer el alcance de las variaciones extremas (como en caso del comportamiento de una acción bursátil). Sin embargo, tiende a crecer con el tamaño de la muestra, y es muy sensible a los valores atípicos. Esto limita su utilidad como medida de variabilidad.

Rango.

➢ Ejemplo.

- Si se tienen los siguientes datos hipotéticos sobre las Participaciones (en miles de millones de pesos) para 10 estados de la zona norte del país:

Cuadro 4.

Estado	Participación (mmp)
1	3
2	3
3	5
4	5
5	5
6	7
7	7
8	8
9	8
10	9

El rango o amplitud es:

$X_{10}-X_1=9-3 = 6$

MEDIDAS DE DISPERSIÓN

➢ Desviación media

- Es el promedio de los valores absolutos de las desviaciones respecto de la media.
- Si X_1, X_2, ..., X_n forman una muestra de *n* observaciones, la desviación media se calcula:

$$DM = \frac{\sum_{i=1}^{n} \left| X_i - \overline{X} \right|}{n} = \frac{\sum_{i=1}^{n} \left| d_i \right|}{n}$$

donde *d* son las desviaciones respecto de la media y | | es el signo de valor absoluto, lo que implica que no se toman en cuenta los signos de las desviaciones.

Desviación media.

- **Ejemplo.** Cálculo de la desviación media de los datos del ***Cuadro 4***:

Estado	Part. (mmmp)	$\lvert X_i - \overline{X} \rvert = d$
1	3	$\lvert 3-6 \rvert = 3$
2	3	$\lvert 3-6 \rvert = 3$
3	5	$\lvert 5-6 \rvert = 1$
4	5	$\lvert 5-6 \rvert = 1$
5	5	$\lvert 5-6 \rvert = 1$
6	7	$\lvert 7-6 \rvert = 1$
7	7	$\lvert 7-6 \rvert = 1$
8	8	$\lvert 8-6 \rvert = 2$
9	8	$\lvert 8-6 \rvert = 2$
10	9	$\lvert 9-6 \rvert = 3$
Total	**60**	**18**

$$\overline{X} = \frac{\sum_{i=1}^{n} X_i}{n} = \frac{60}{10} = 6$$

$$DM = \frac{\sum_{i=1}^{n} \lvert d_i \rvert}{n} = \frac{18}{10} = 1.8$$

n=10

MEDIDAS DE DISPERSIÓN

La desviación media tiene la ventaja de que toma en cuenta todos los datos de la muestra, y, mediante un cálculo sencillo, proporciona un indicador sobre la dispersión de los datos.

Sin embargo, no toma en cuenta los signos de las desviaciones (de hacerlo, la z sería igual a cero), lo cual limita la utilización de la medida desde el punto de vista matemático.

MEDIDAS DE DISPERSIÓN

➢ Varianza

- Otra forma de resolver el problema del signo de las desviaciones, sin eliminarlo, es elevar al cuadrado las desviaciones.
- Al elevar las desviaciones al cuadrado, su suma ya no es cero, sino un número positivo. Esta suma de cuadrados puede considerarse como una medida de la dispersión total de la distribución. Dividiendo la suma entre *n*, número de datos de la muestra, se obtiene la media de los cuadrados de las desviaciones, medida que se conoce con el nombre de **varianza**.
- La fórmula para calcular la varianza (s^2)es:

$$s^2 = \frac{\sum_{i=1}^{n} \left(X_i - \overline{X}\right)^2}{n}$$

Varianza.

- **Ejemplo.** Cálculo de la varianza, a partir de los datos del ***Cuadro 4:***

Estado	**Part. (mmmp)**	$X_i-\overline{X}$	$(X_i-\overline{X})^2$
1	3	3 − 6 = -3	9
2	3	3 − 6 = -3	9
3	5	5 − 6 = -1	1
4	5	5 − 6 = -1	1
5	5	5 − 6 = -1	1
6	7	7 − 6 = 1	1
7	7	7 − 6 = 1	1
8	8	8 − 6 = 2	4
9	8	8 − 6 = 2	4
10	9	9 − 6 = 3	9
Total			**40**

$$s^2 = \frac{\sum_{i=1}^{n} (X_i - \overline{X})^2}{n}$$

$$s^2 = \frac{40}{10} = 4$$

n=10

Varianza

Cuando el cálculo de la varianza corresponde al de la población, ésta se representa con el

Signo (sigma al cuadrado).

La varianza presenta un problema de interpretación, debido que el resultado que se obtiene esta expresado, no en las unidades originales de los datos, sino en el cuadrado de estas unidades.

MEDIDAS DE DISPERSIÓN

➢ Desviación estándar

- Con objeto de eliminar el problema de las unidades de medida que plantea la varianza, se calcula la raíz cuadrada de ésta. El resultado es la desviación estándar (s), la cual tiene como unidades, las mismas de los datos originales.

- Por lo tanto:

$$s = \sqrt{s^2} = \sqrt{\frac{\sum_{i=1}^{n} \left(X_i - \overline{X}\right)^2}{n}}$$

- **Ejemplo.** Dado que la varianza que se obtuvo es 4:

$$s = \sqrt{s^2} = \sqrt{4} = 2$$

- La desviación estándar de la *población* corresponde por lo tanto al signo $\sigma = \sqrt{\sigma^2}$.

Desviación estándar.

La desviación estándar es de gran utilidad, ya que nos permite determinar con bastante precisión dónde se sitúan los valores de una distribución de frecuencia en relación con la media.

De acuerdo con el *Teorema de Tchebyshev,* cualquiera que sea la forma de una distribución, por lo menos 75% de los valores se ubican dentro de un intervalo u ± 2 desviaciones estándar, y un mínimo de 89% de los datos se encuentra en el intervalo u ± 3 desviaciones estándar.

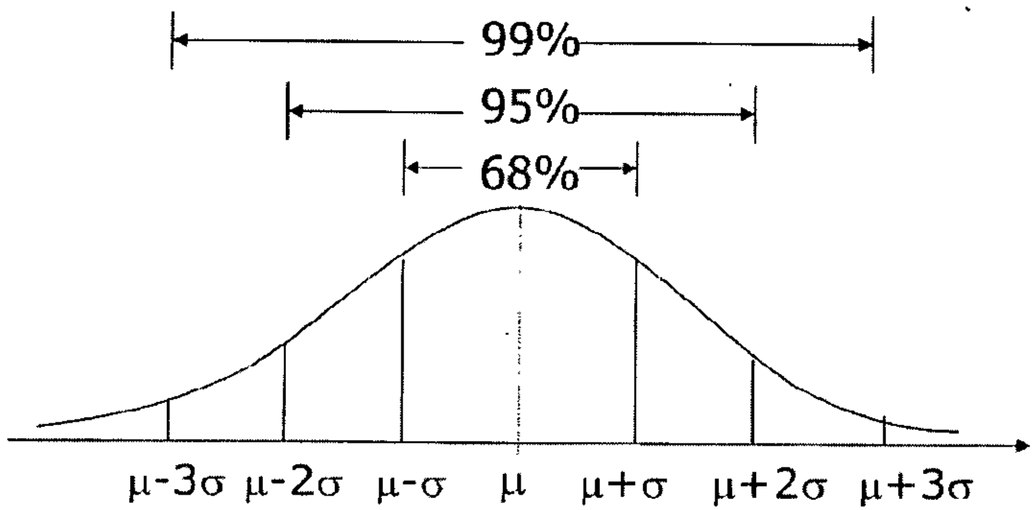
Para el caso de una distribución simétrica en forma de campana:

> Cerca del 68% de los valores de la población se ubican dentro del intervalo U ± 1 desviación estándar;

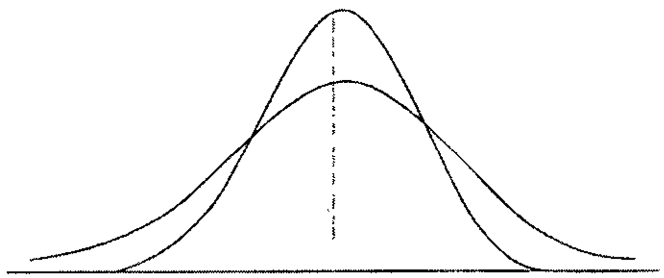
Desviación estándar.

- Cerca del 95% de los valores se encuentran en el intervalo $\mu \pm 2$ desviaciones estándar, y
- Cerca del 99% de los valores están dentro del intevalo $\mu \pm 3$ desviaciones estándar

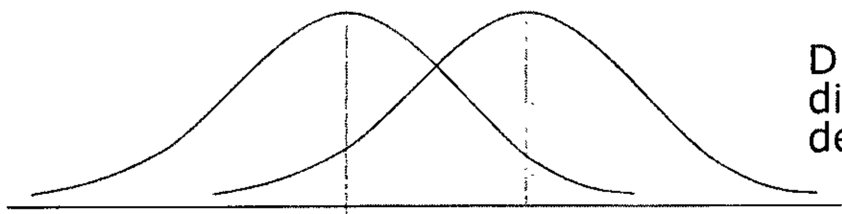
Localización de las observaciones alrededor de la media, en una distribución de frecuencia simétrica, en forma de campana.



COMPARACIÓN GRÁFICA DE LA MEDIA Y DE LA DESVIACIÓN ESTÁNDAR



Distribuciones con igual media, pero con distintas desviaciones estándar.



Distribuciones con media diferente, pero con igual desviación estándar.

DISTRIBUCIONES DE PROBABILIDAD

- La probabilidad es la posibilidad de que ocurra algo. Las probabilidades se expresan como fracciones (1/6, ½, 8/9) o como decimales (0.167,0.500,0.889) entre 0 y 1.

  0 ⟶ nunca ocurrirá

  1 ⟶ sucede siempre

  - Evento: uno o varios resultados al hacer un experimento (evento 1, sol; evento 2, águila).
  - Espacio muestral: conjunto de todos los resultados posibles de un experimento (volado = {águila, sol})
- Recordar, distribuciones de frecuencia como una forma de presentar las variaciones de los datos observados.

DISTRIBUCIONES DE PROBABILIDAD

> Distribución de probabilidad (Modelos): se puede ver como una distribución teórica de frecuencia, es decir, cómo se espera que varíen los resultados.

- Son útiles para hacer inferencia estadística y tomar decisiones en condiciones de incertidumbre.
- La distribución de probabilidad es un listado de todos los resultados que podrían presentarse de realizarse el experimento.

BIBLIOGRAFÍA BÁSICA

Archier,Georges, y Seriey, Hervé. "la Empresa del Tercer Tipo. Una nueva concepción de la Empresa" Edit. Planeta.

Beck, Ulrich, "Un Nuevo Mundo Feliz. La precariedad del trabajo en la era de la globalización" Edit. Paidós.

Buckingham, Marcus. " Ahora, descubra sus fortalezas" NORMA

Cannon, T. "La Responsabilidad de la Empresa" Edit. Folio.

Drucker, Peter F."Las Nuevas Realidades" Edit. Edhasa.

Etzioni, Amitai. "La Tercera Vía. Hacia una buena Sociedad. Propuestas desde el Comunitarismo" Edit. Trotta.

Galbraith, John K. "La Cultura de la Satisfacción" Edit. Ariel.

Giddens, Anthony. "Consecuencias de la Modernidad" Edit. Alianza,

Habermas, Jurgen. "Problemas de legitimación en el Capitalismo Tàrdío" Edit. Amorrortu.

Lipschutz, Seymour "Teoría y Problemas de Probabilidad" Mc Graw Hill.

Martin, Doris. EQ. Què es la inteligencia Emocional" EDAF

Spiegel, Murray "Estadística" Mc Graw Hill.

Touraine, Alian. "La Sociedad Post-Industrial" Edit. Ariel.