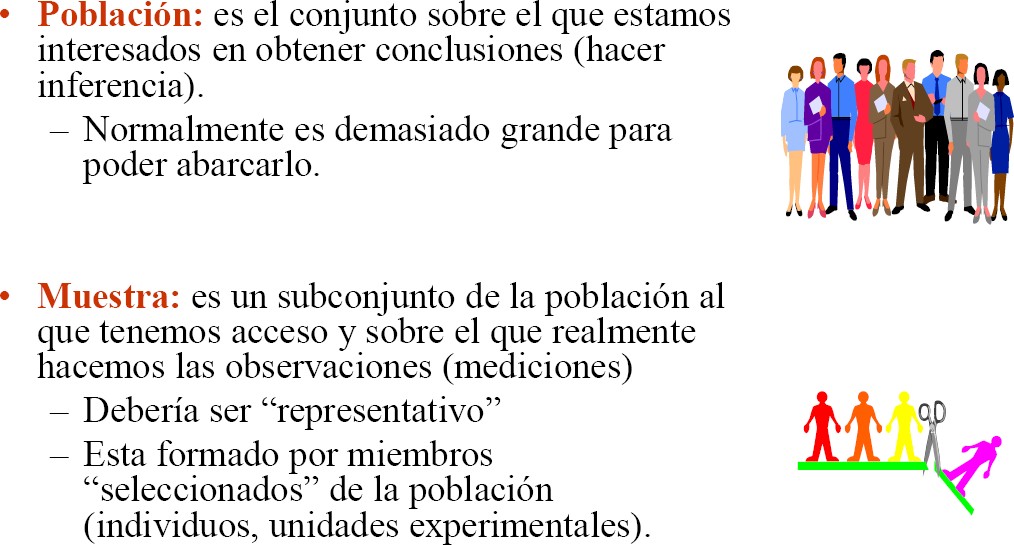
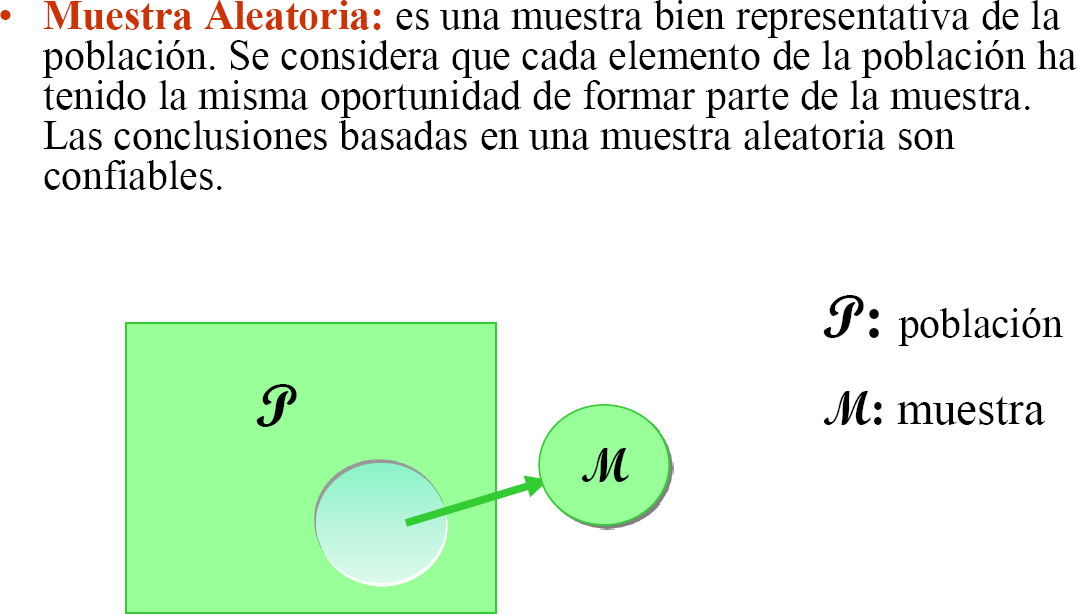
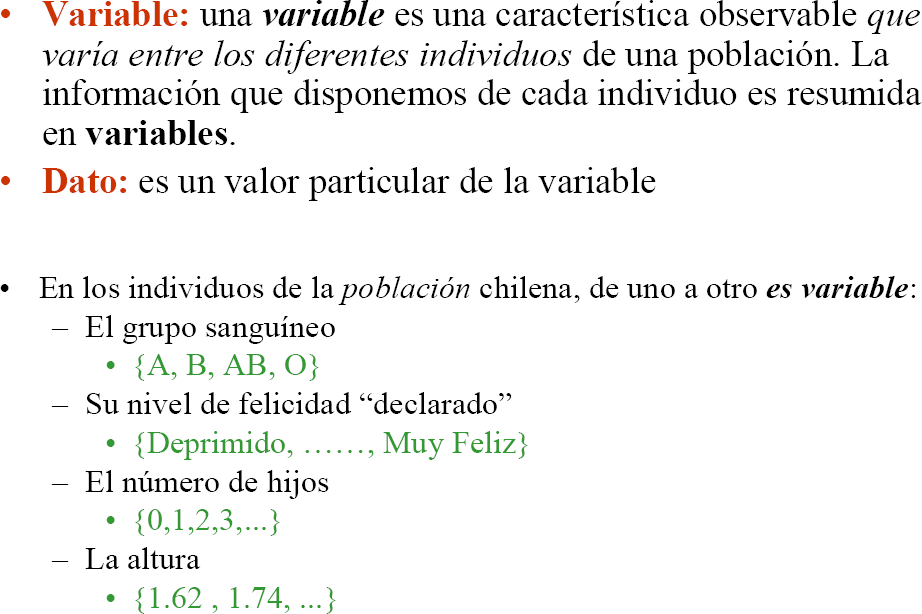
SEGUNDO MEDIO

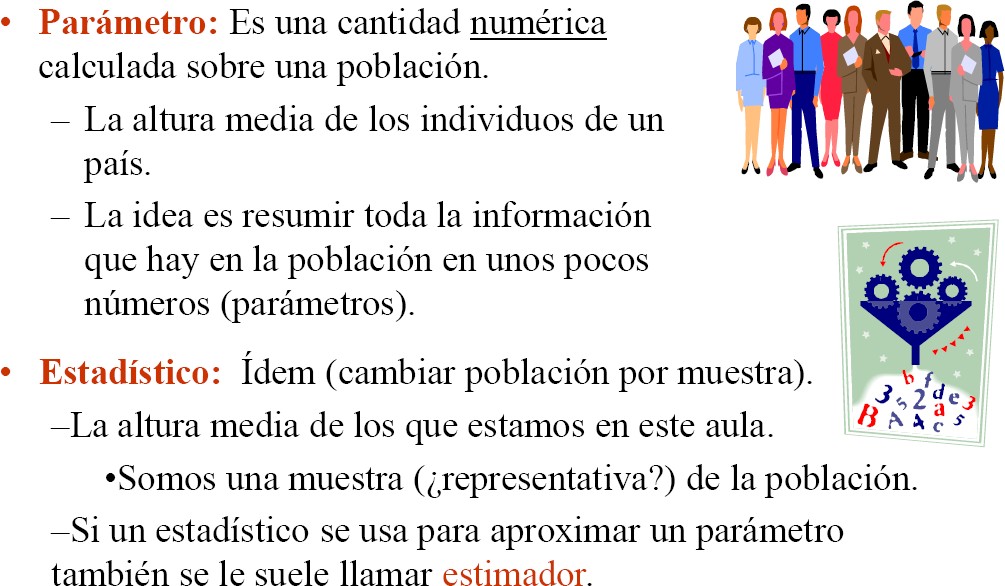
Estadística: conceptos básicos y definiciones.

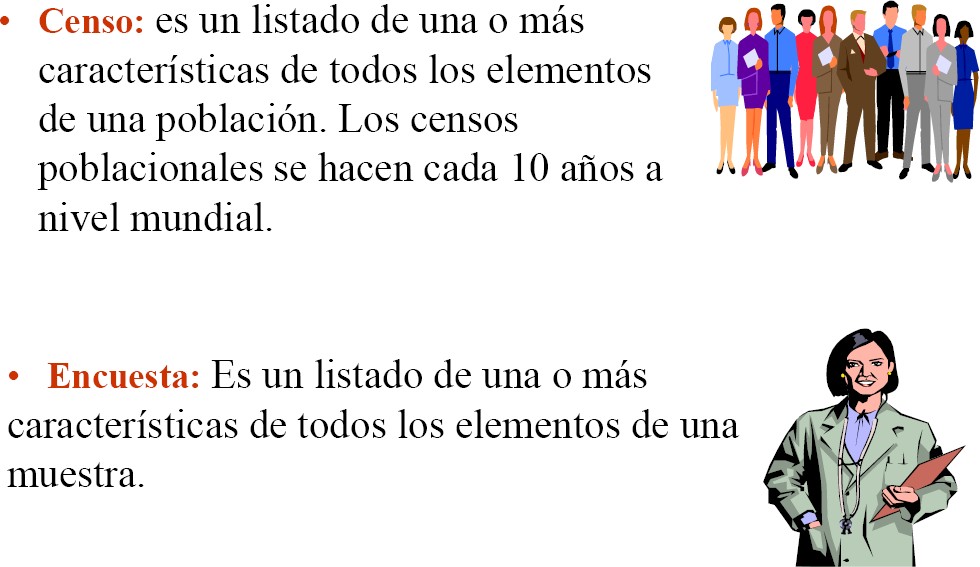
# Conceptos básicos



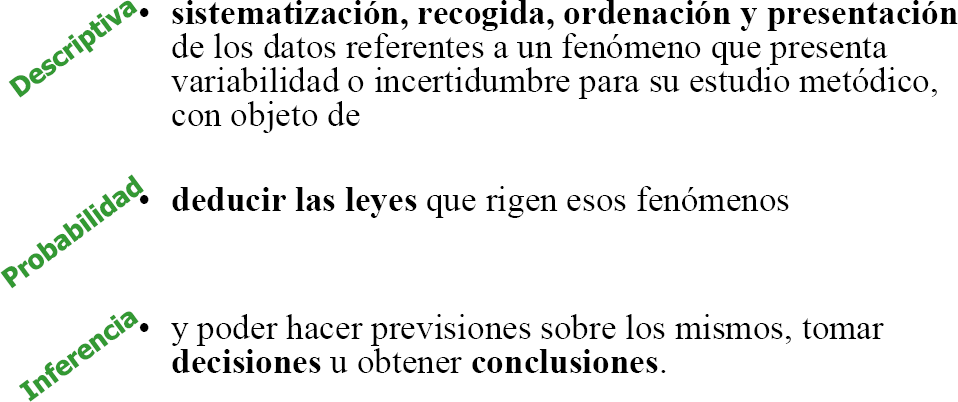


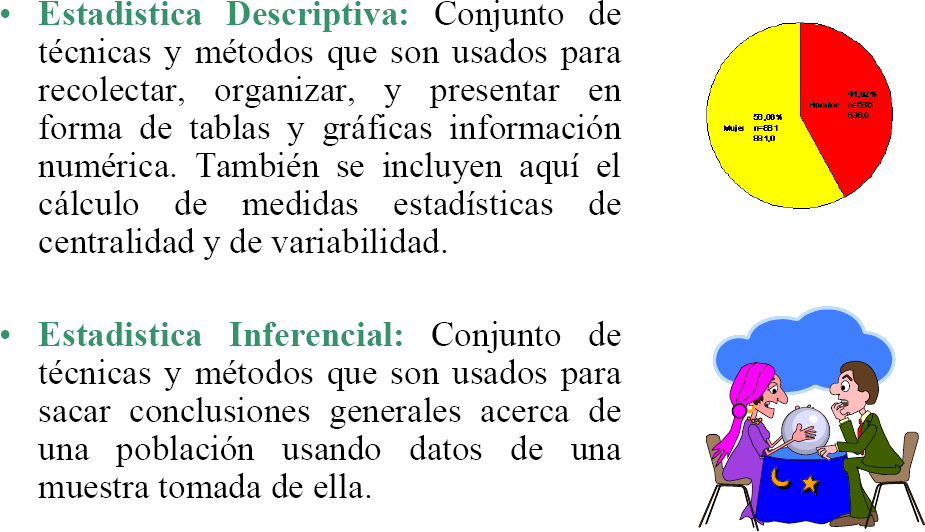


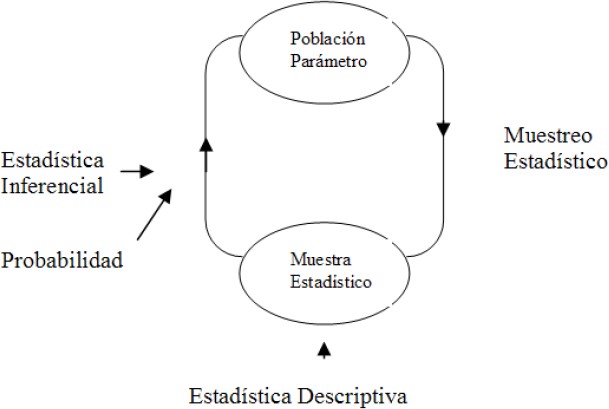


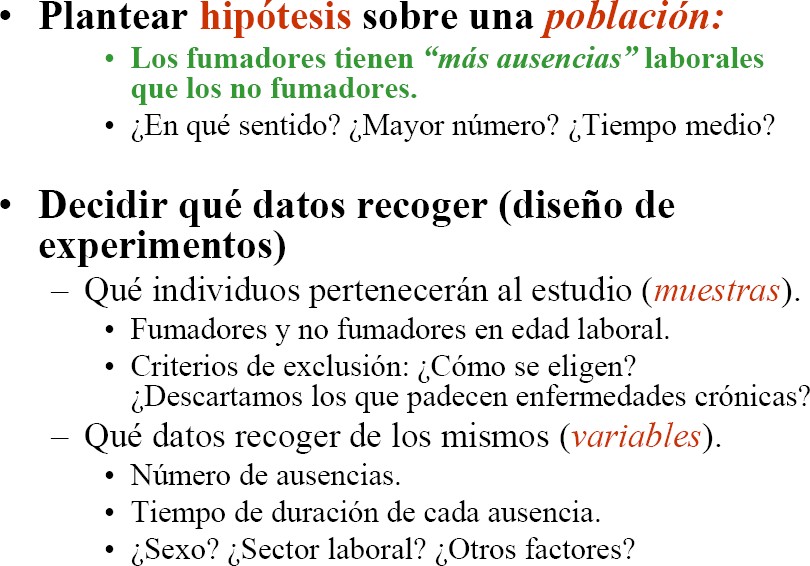


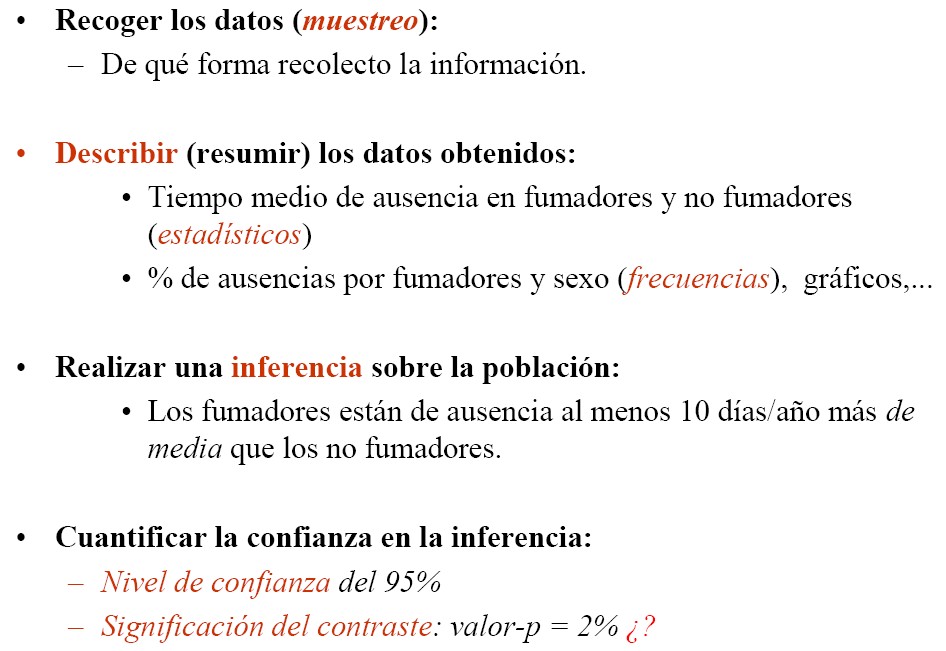
##### La estadística es la Ciencia de la





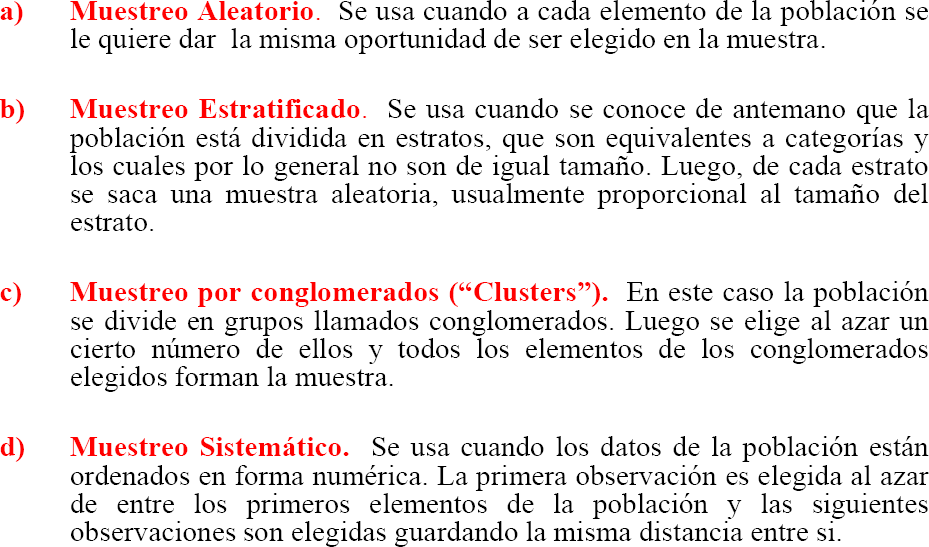




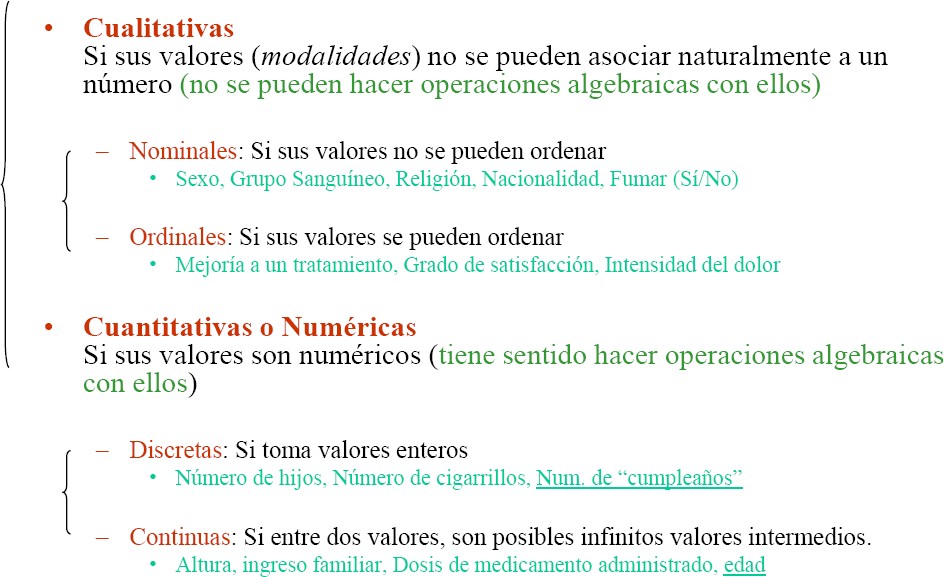
Pasos en un estudio estadístico cont.

11

# Técnicas de Muestreo



Tipo de Variables



Ejemplos:

# Tipo de variables cont.

* Es buena idea codificarlas variables como números para poder procesarlas con facilidad en un computador.
* Es conveniente asignar “etiquetas” a los valores de las variables para

recordar qué significan los códigos numéricos.

–**Género** (Cualitativa : Códigos arbitrarios) 1 : Hombre

2 : Mujer

–**Raza** (Cualitativa: Códigos arbitrarios) 1 : Blanca

2 : Negra, ...

–**Felicidad** Ordinal: Respetar un orden al codificar.

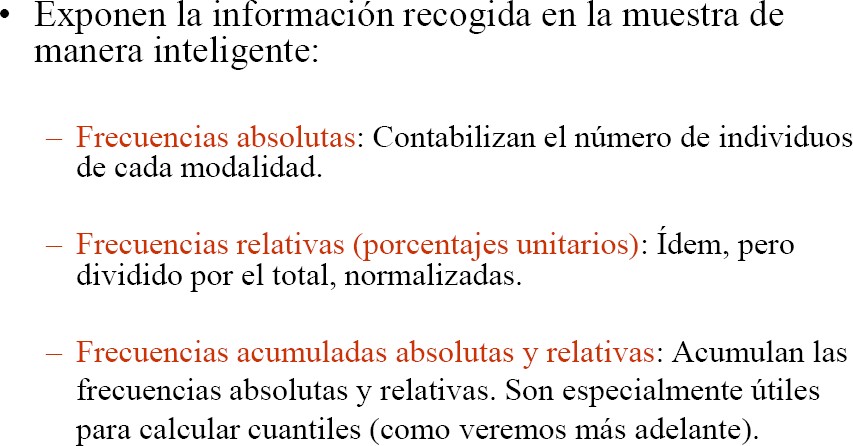
1. : Muy feliz
2. : Bastante feliz
3. : No demasiado feliz

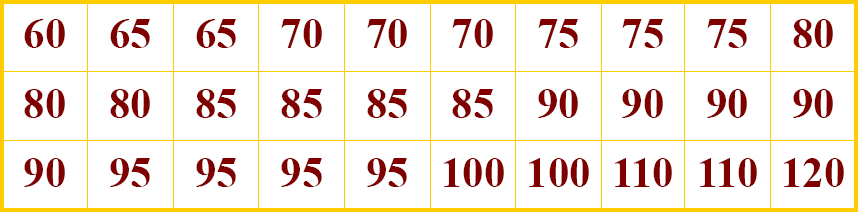
* Se pueden asignar códigos a respuestas especiales como 0 : No sabe

99 : No contesta...

# Ejemplo: Tipo de variables cont.

Tabla de Frecuencias



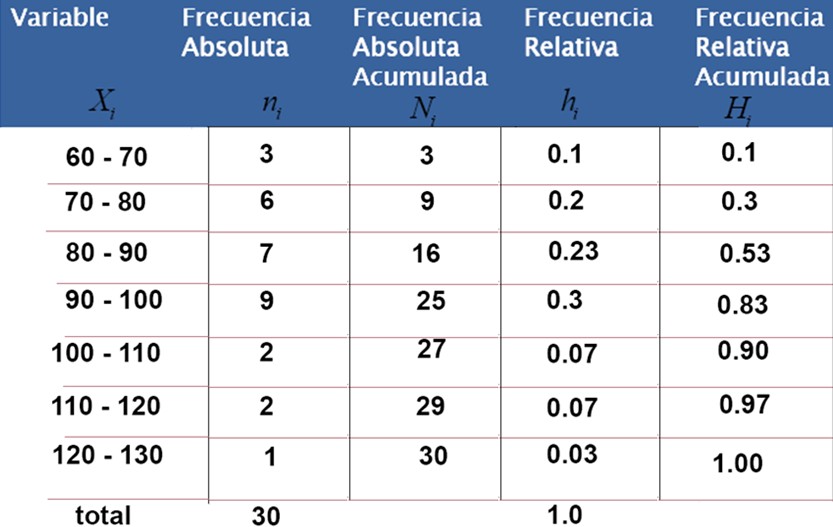
Ordenamos los datos en forma creciente:

La **amplitud** total **A = 120 –60**

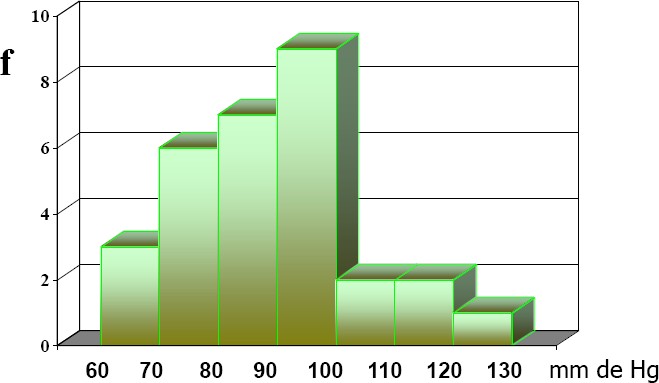
**Número de clases**: **K = 301/2 = 5.48**. Aprox. 6 clases

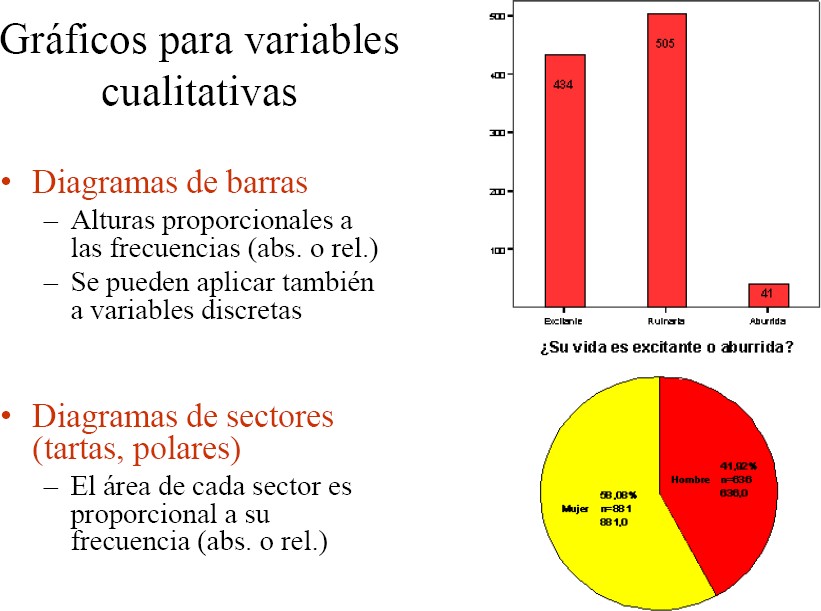
**Extensión del intervalo**: **H = A/ K = 60/6 = 10**

En este caso, entonces, la tabla de frecuencias tendrá aproximadamente 6 clases de amplitud 10 unidades en cada clase.

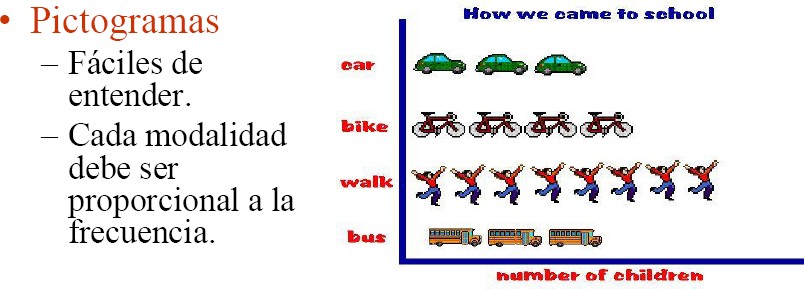


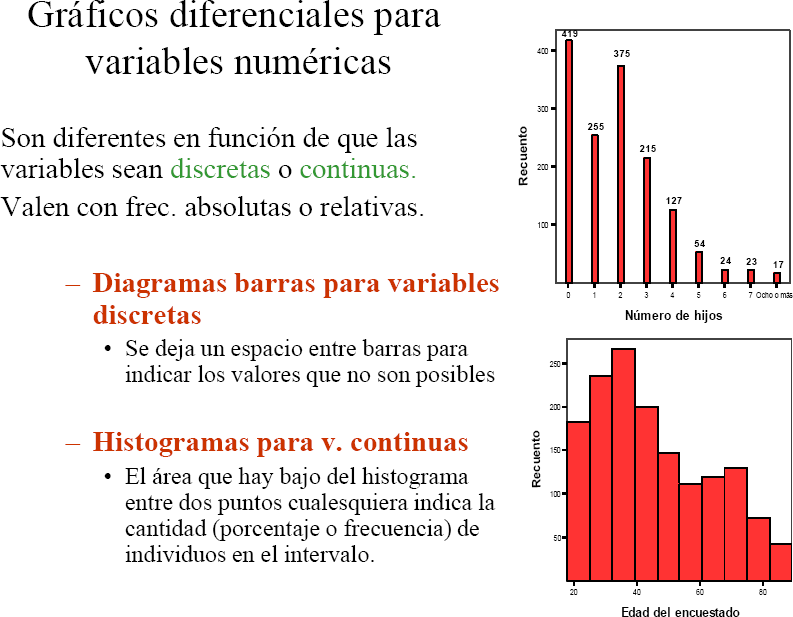
**Histograma de la distribución de presión diastólica en mm de Hg según las frecuencias absolutas:**



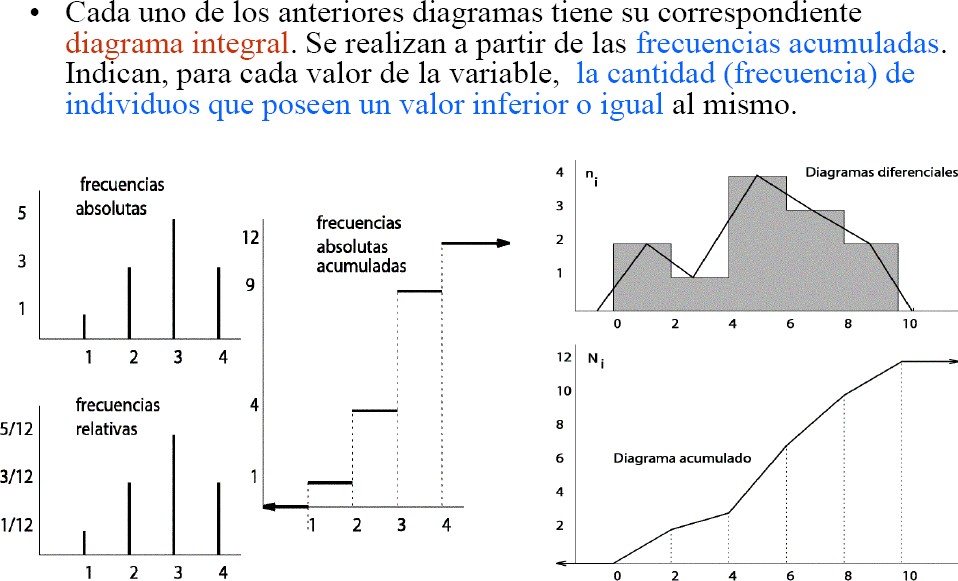


Gráficos para variables cualitativas cont.

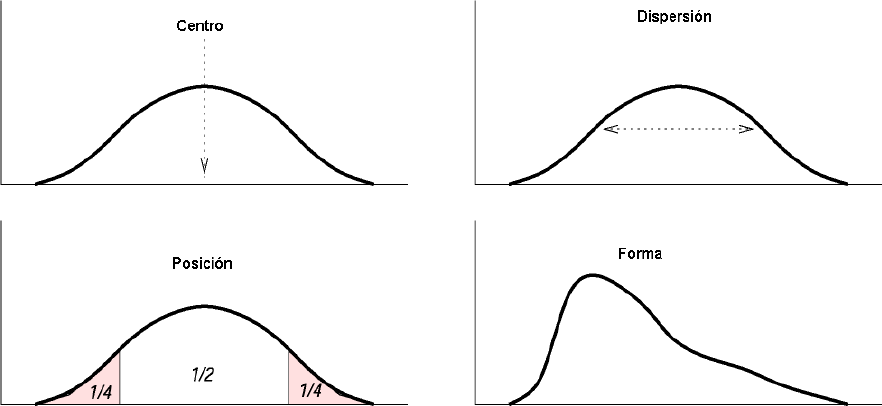




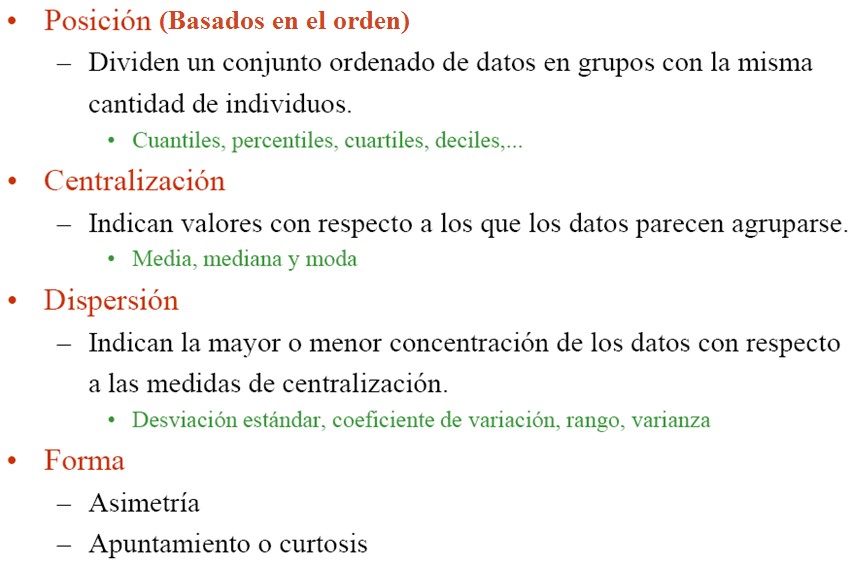
Diagramas Integrales

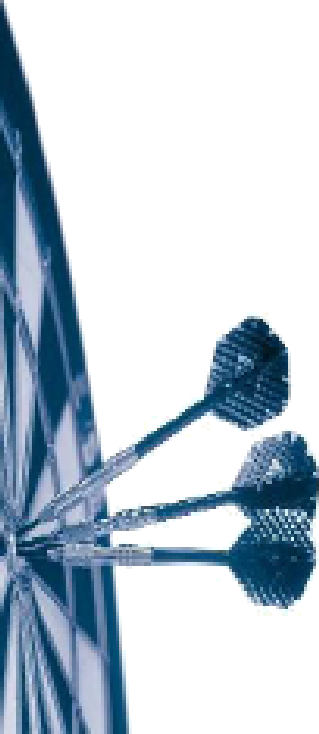


Estadísticos de forma intuitiva



Estadísticos

25



* Añaden unos cuantos casos particulares a las medidas de posicién. Son medidas que buscan posiciones (valores) con respecto a los que los datos muestran tendencia a agruparse.
* Media: es la media aritmética Qromedio) de los valores de una variable. Suma de los valores dividido por el tamano muestral.

— Media de (2, 2, 3, 7} es (2+2+3+7)/4 = 3.5

* + Conveniente cuando los datos se concentran simétricamente con respecto a ese valor. Muy sensible a valores extremos.
  + Centro de gravedad de los datos.



* Mediana. es un valor que divide a las observaciones en dos grupos con el mismo nñmero de individuos Qercentil 50). Si el nñmero de datos es par, se elige la media de los dos datos centrales.

— Mediana de 1. 2. 4, 5, 6. 6. 8 es 5

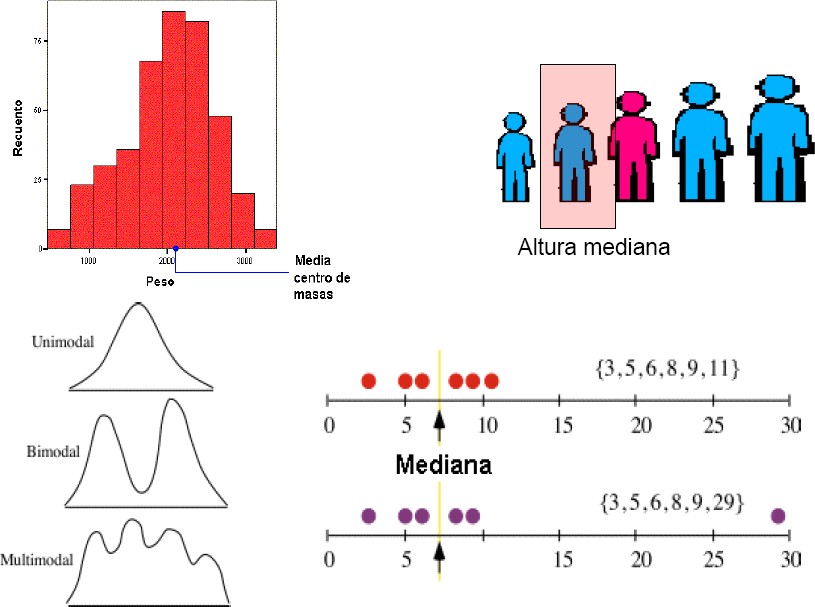
— Mediana de 1, 2, 4, 5, 6, 6, 8, 9 es (5+6)/2 = 5.5

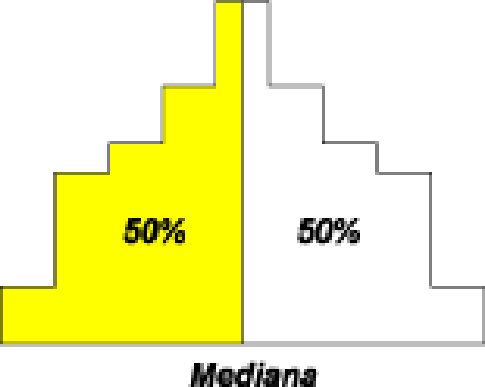
* + Es conveniente cuando los datos son asimétricos. No es sensible a valores extremos.
    - Mediana de 1, 2, 4, 5, 6, 6, 800 es 5. JLa media es

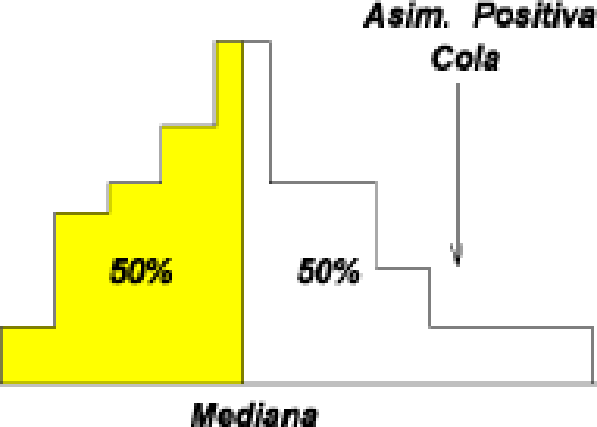
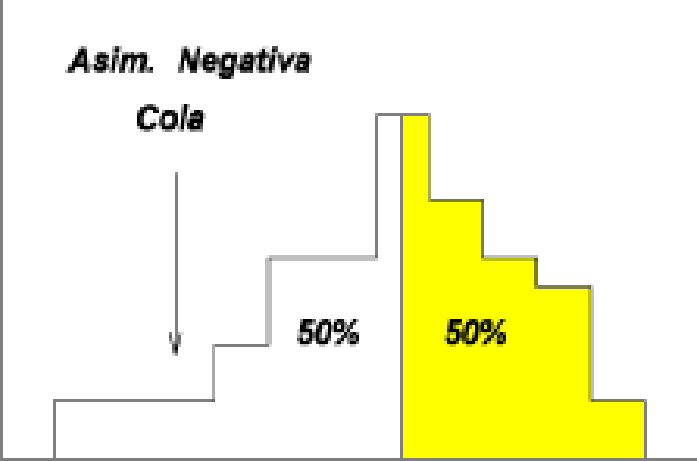
ll7.7!

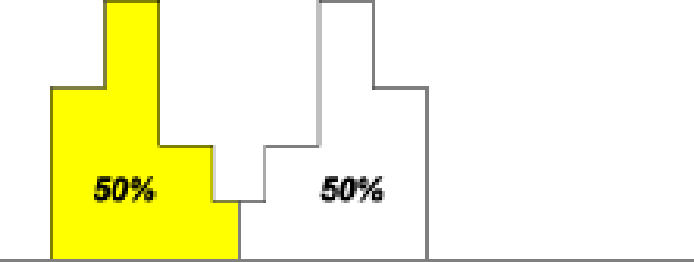
* Moda: es eldos valor/es donde la distribucién de

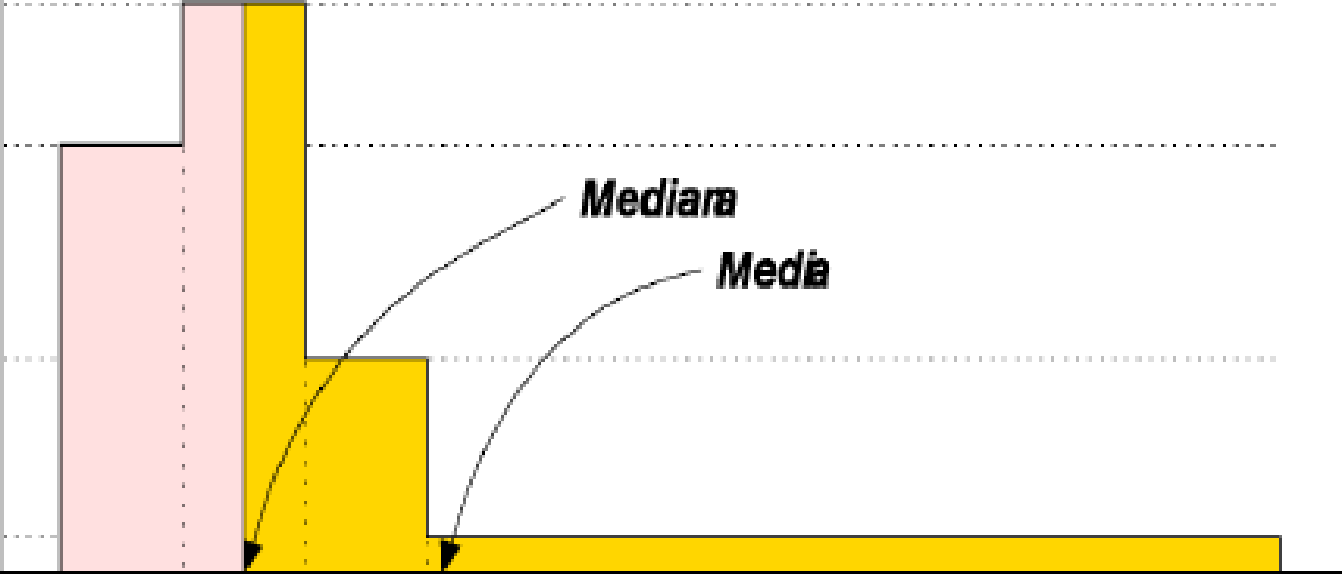
Aecuencia alcanza un mâximo.



Asimetria o sesgo



* Una disBibucién es simétrica si la mitad izquierda de su disoibucifin es la imagen especular de su mitad derecha.
* En las distribuciones siméDicas media y mediana coinciden. Si sfilo hay una moda también coincide.
* La asimetria es positiva o negativa \*°\*\* \*"\*" en funcion de a qué lado se

encuenoa la cola de la distribucifin. qg

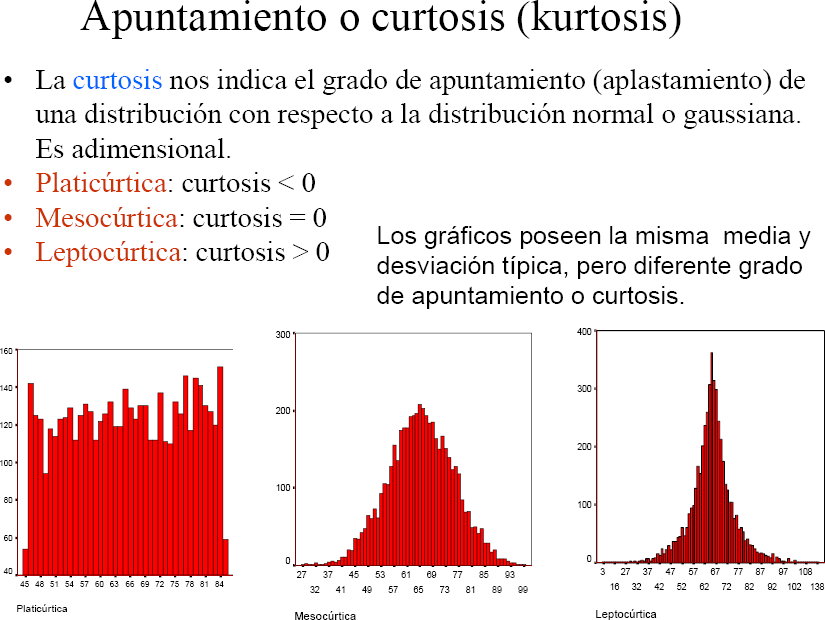
* La media tiende a desplazarse hacia °°

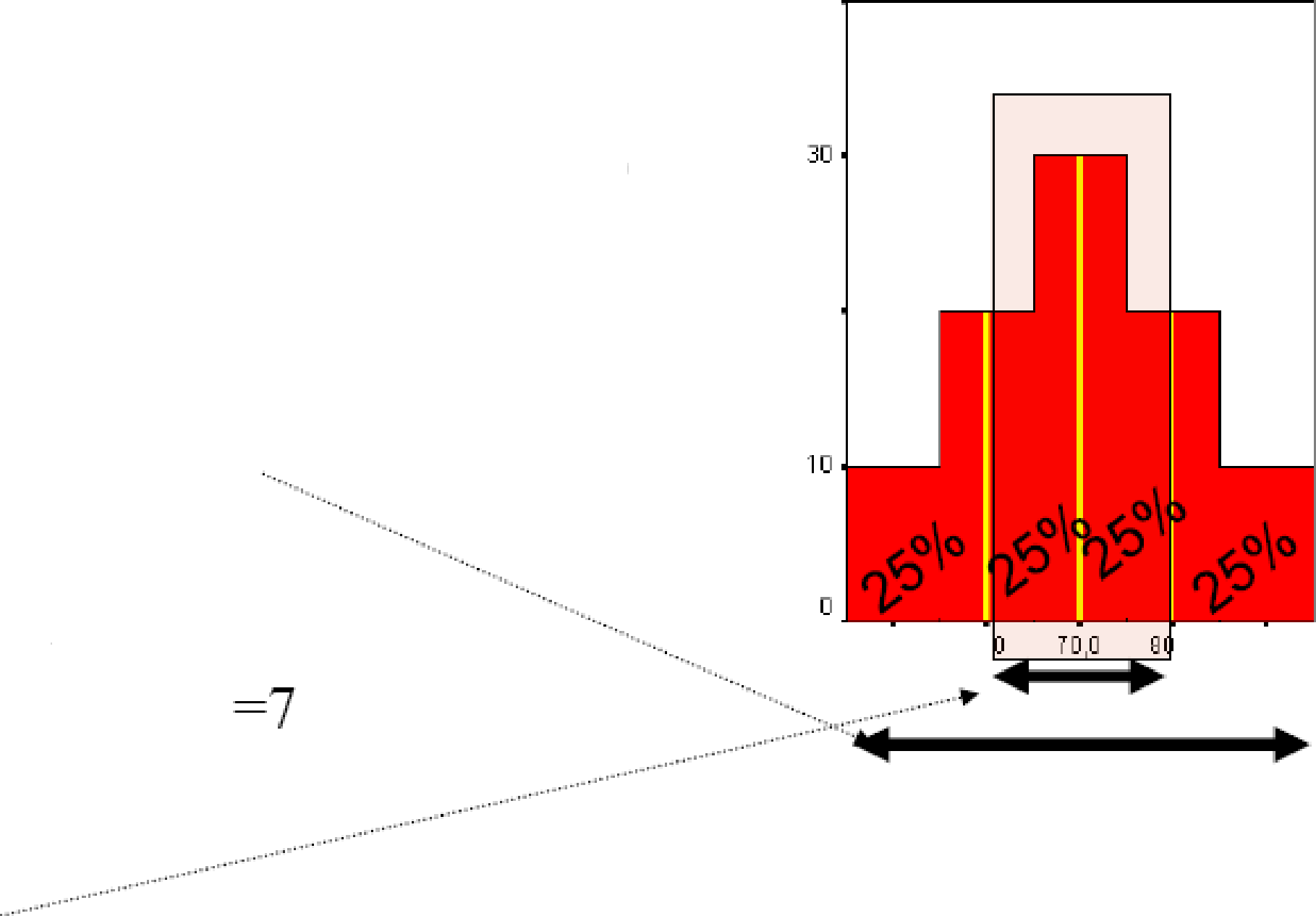
las valores extremos (colas).

* Las discrepancias entre las medidas

de cenbalizacifin son indicacién de ,

asimebia.



Medidas de dispersisn

•Miden el grado de dispersifin (variabilidad)

de los datos, independientemente de su oausa.

•AmpIitud o Rango (‘range’):

La diferenoia entre las

observaciones extremas.

— 2,J,4,3,jj,4. El rango es 8-1

* Es muy sensible a los valores

exoemos.

•Rango intercuartilico (‘interquartile range’):

* Es la distancia enbe el primer y tercer cuartil.
  + Rango intercuartilico = P 75 - Pj 5

20

Desy. tip. = T .6 Medla = r0,0

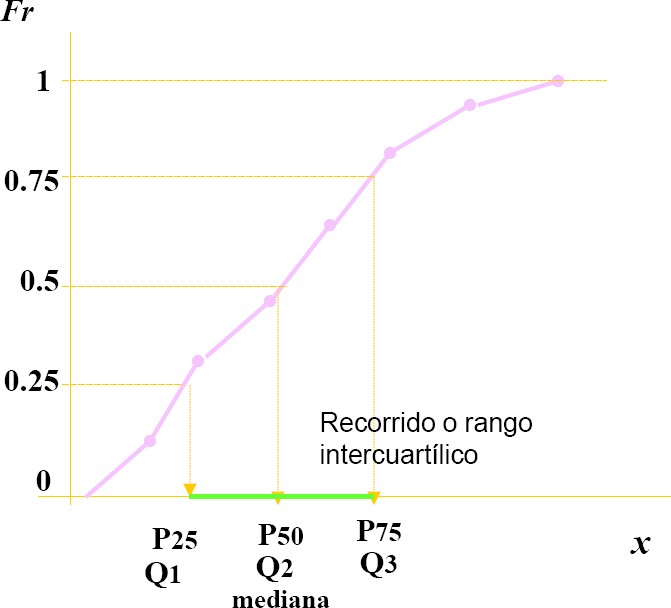
N = 9g, D0

0, 0, ,0 90, P ECO

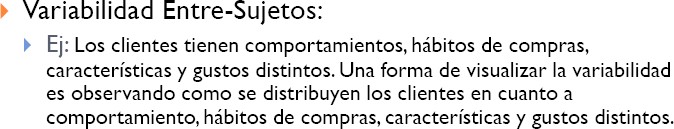
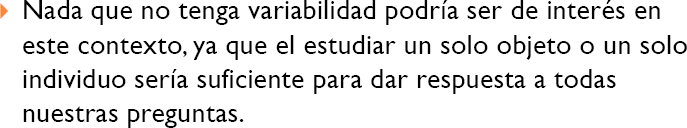
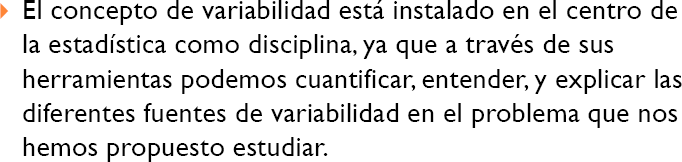
* Parecida al rango, pero eliminando las observaciones mfis extremas

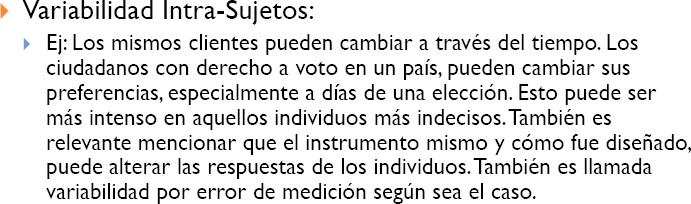
inferiores y superiores.

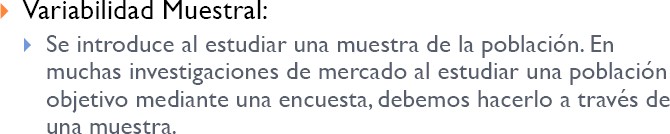
* No es tan sensible a valores extremos.

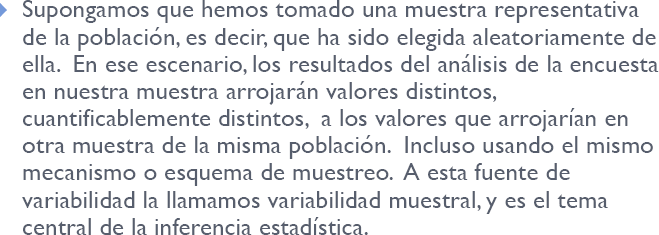


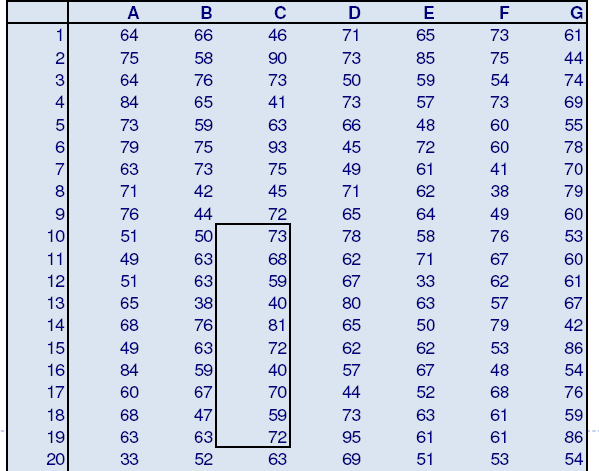
# Concepto de Variabilidad

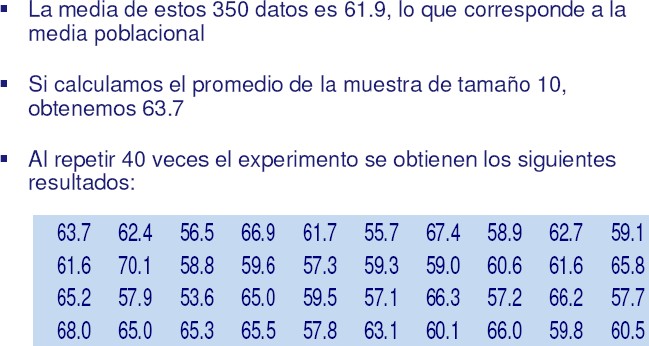


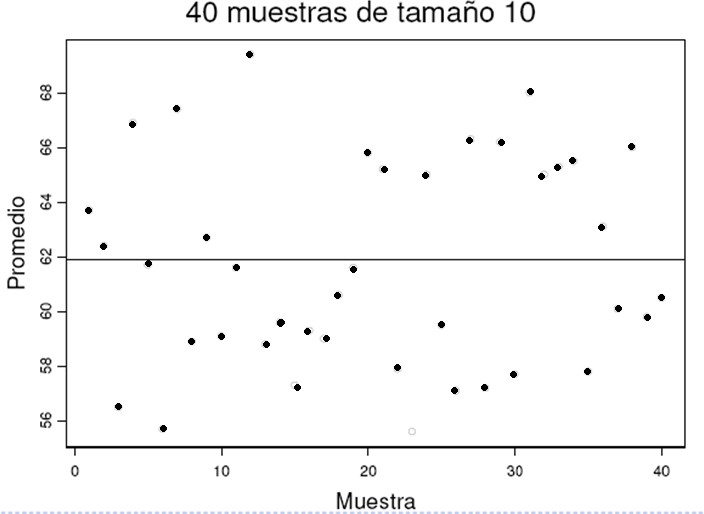


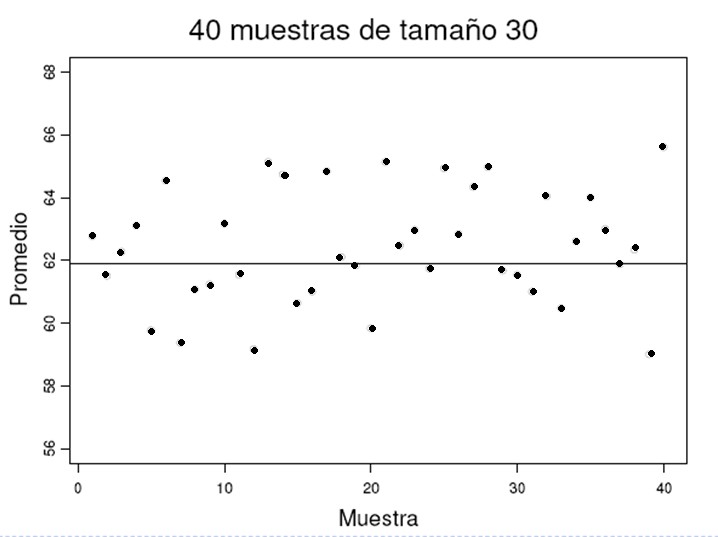


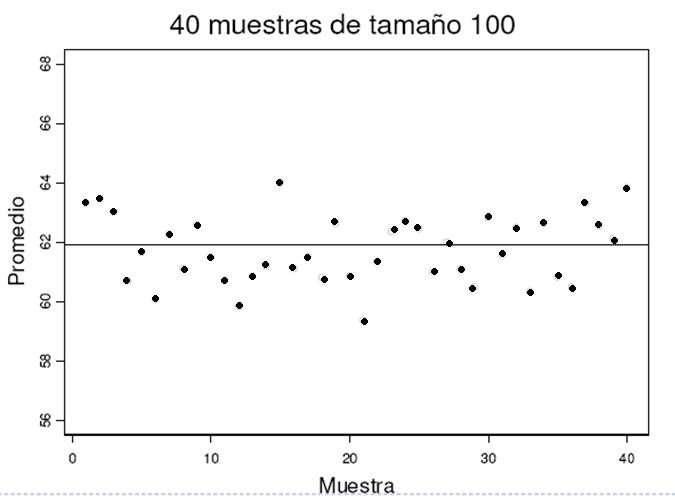




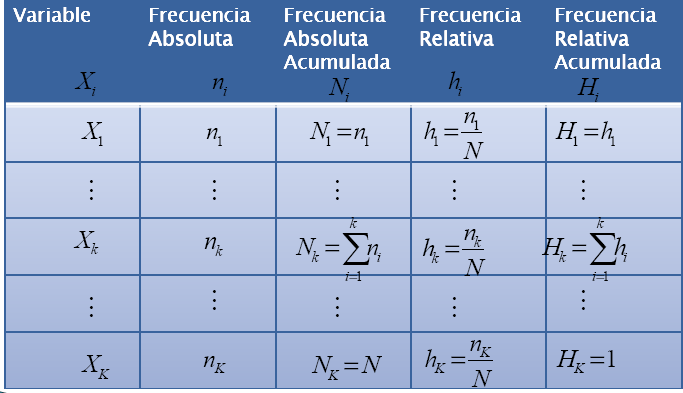




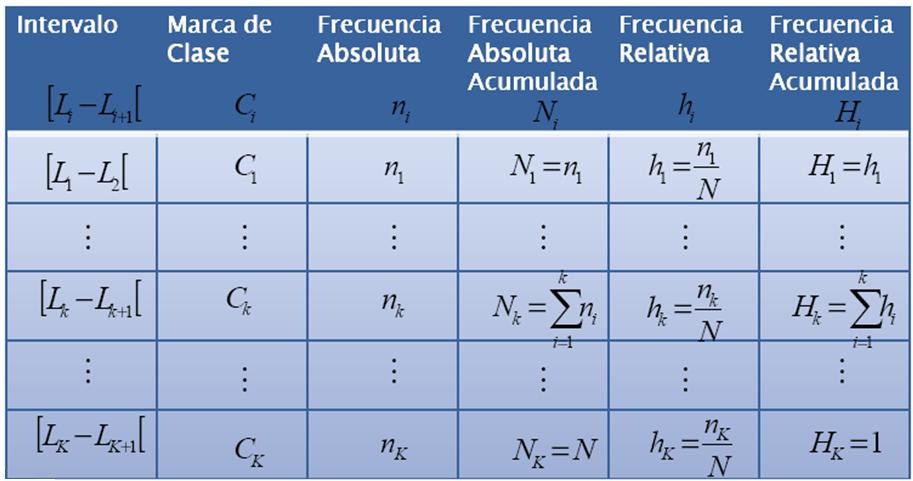
Conceptos de Variabilidad cont.

Conceptos de Variabilidad cont.

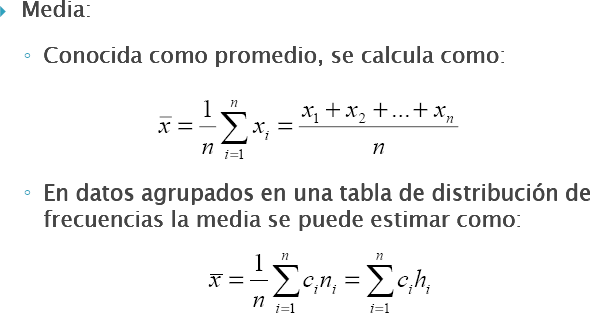
Distribución de Frecuencias



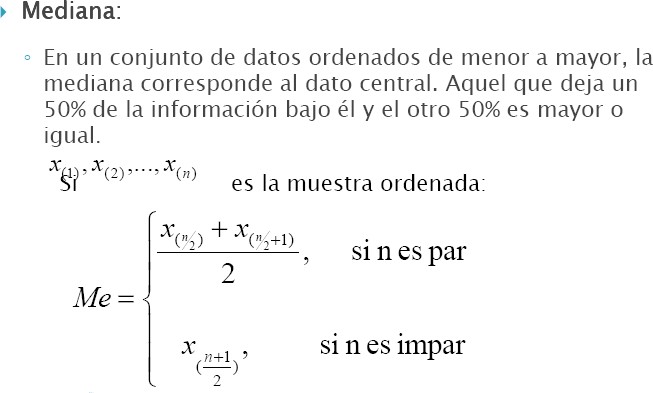
Distribución de Frecuencias cont.

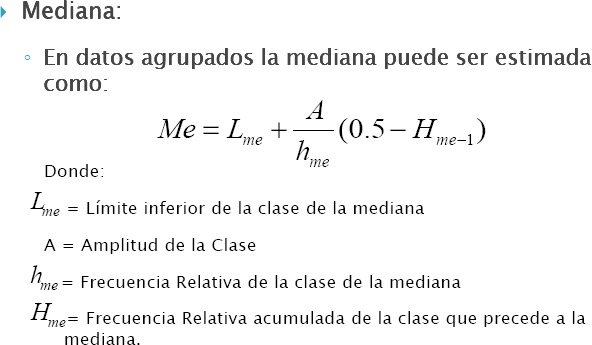


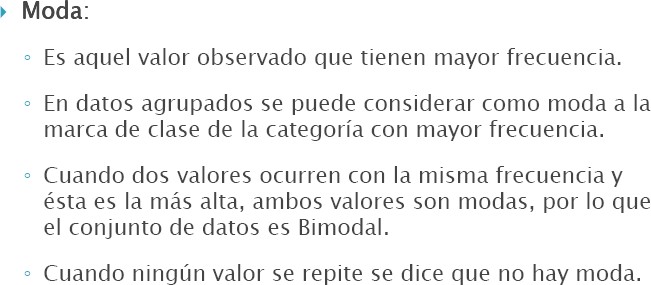
## Medidas de Resumen de Centralización



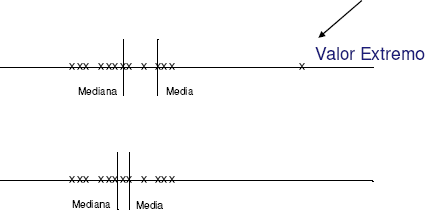
### Medidas de Resumen de Centralización cont.

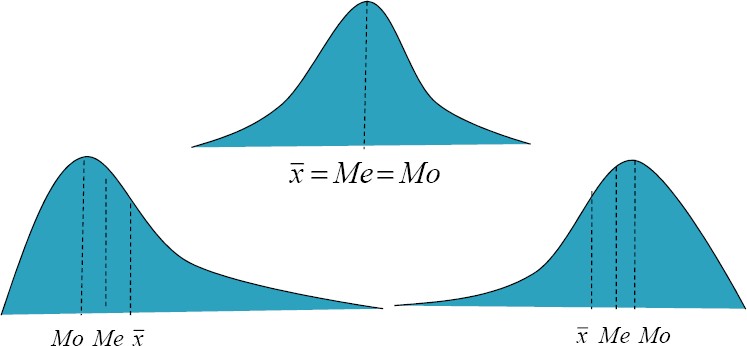




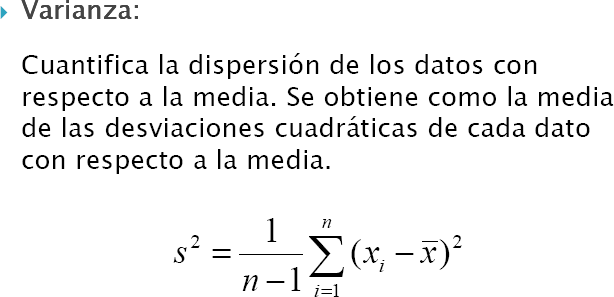


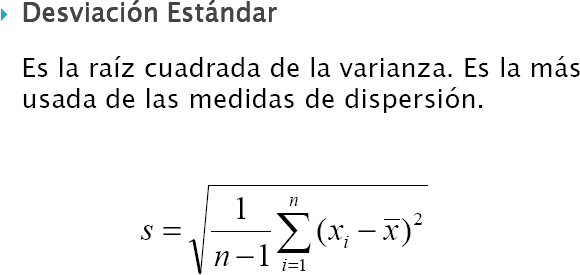
* La media es sensible a la presencia de datos extremos.
* La mediana es muy útil cuando la distribución de la variable es poco simétrica.

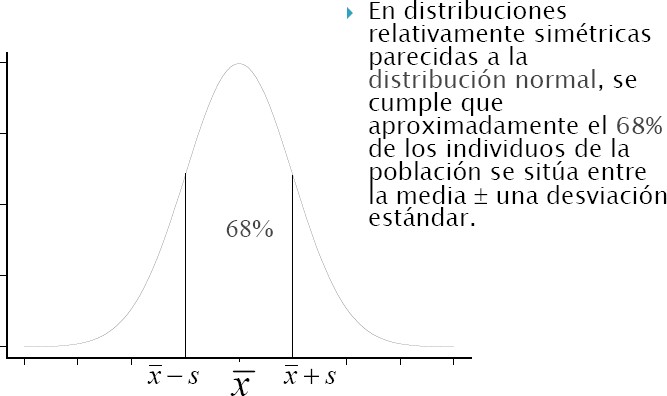


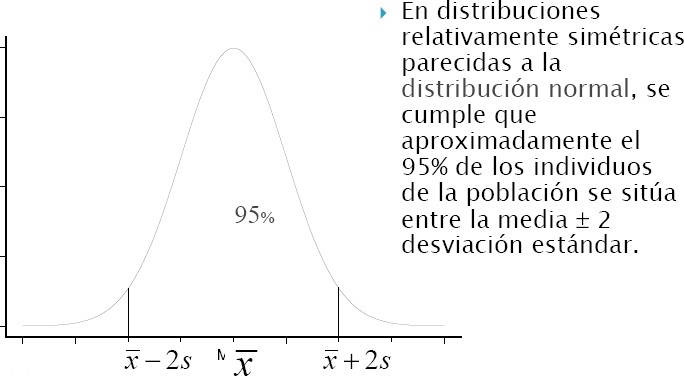


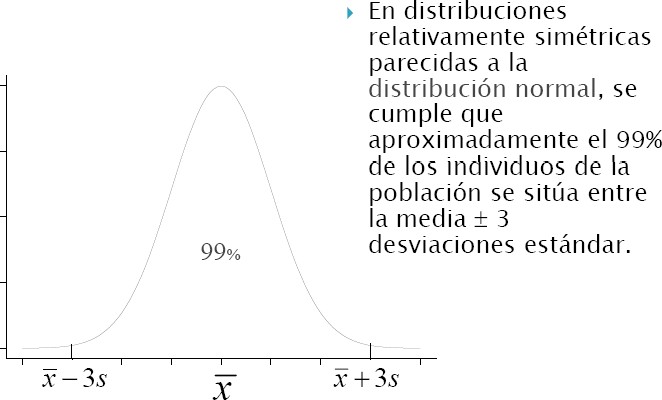
Medidas de Resumen de Dispersión

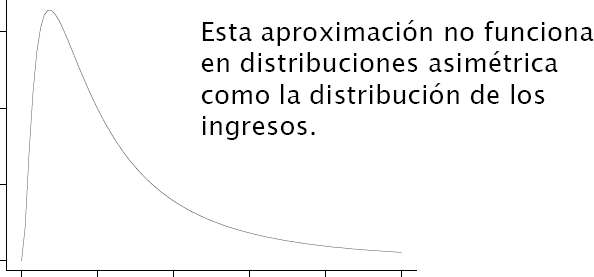


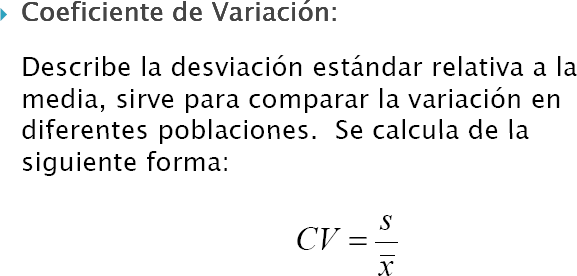


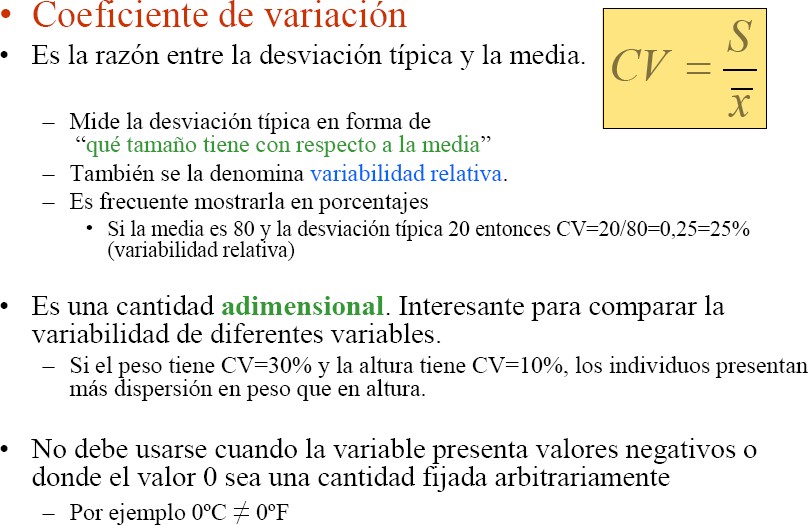


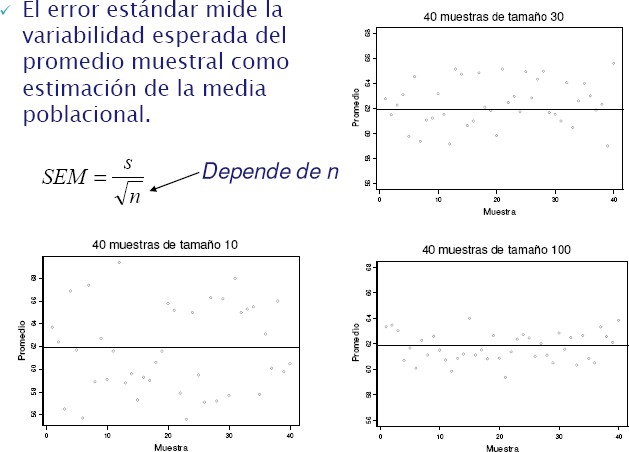


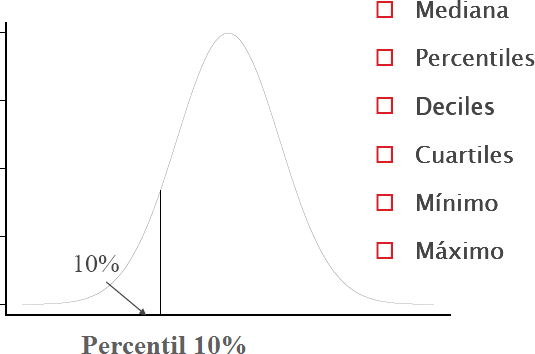


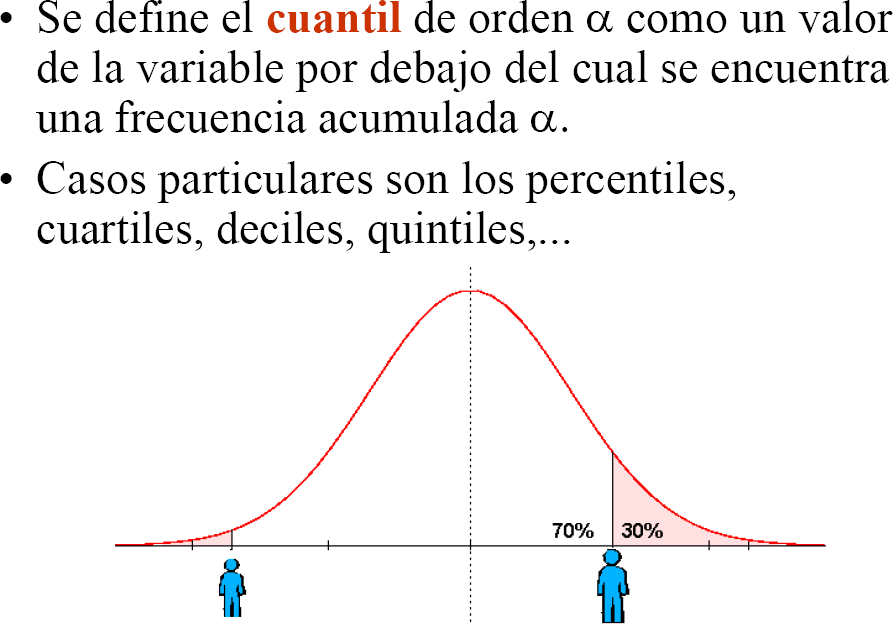


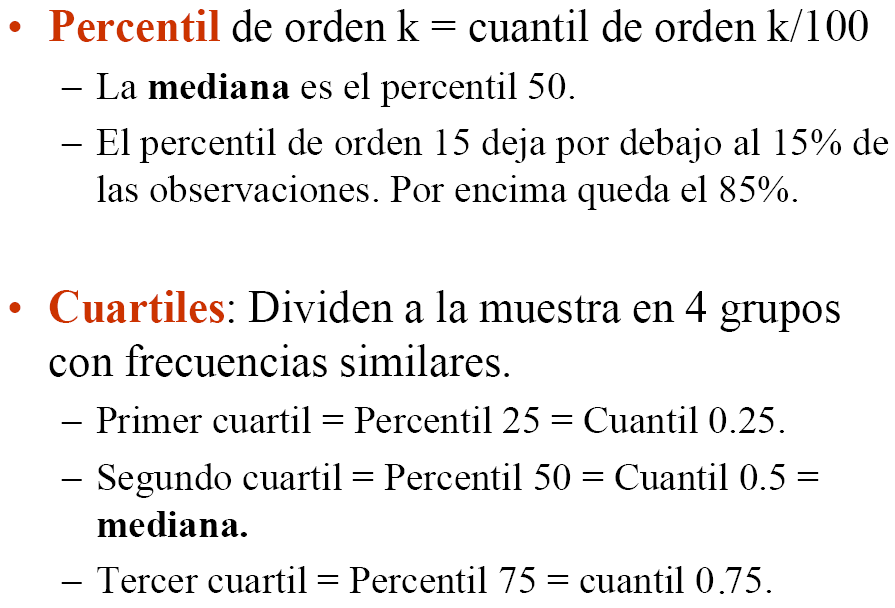








Estadísticos de Posición





Son valores de la variable que dividen a la muestra en partes de igual porcentaje.

Los **percentiles** separan la muestra en grupos de 1% cada uno (son 99).

* **Cuartiles**: agrupan 25% c/u (son 3).
* **Quintiles**: agrupan 20% c/u (son 4).
* **Deciles**: agrupan 10% c/u (son 9).