3. MEDIDAS DE POSICIÓN

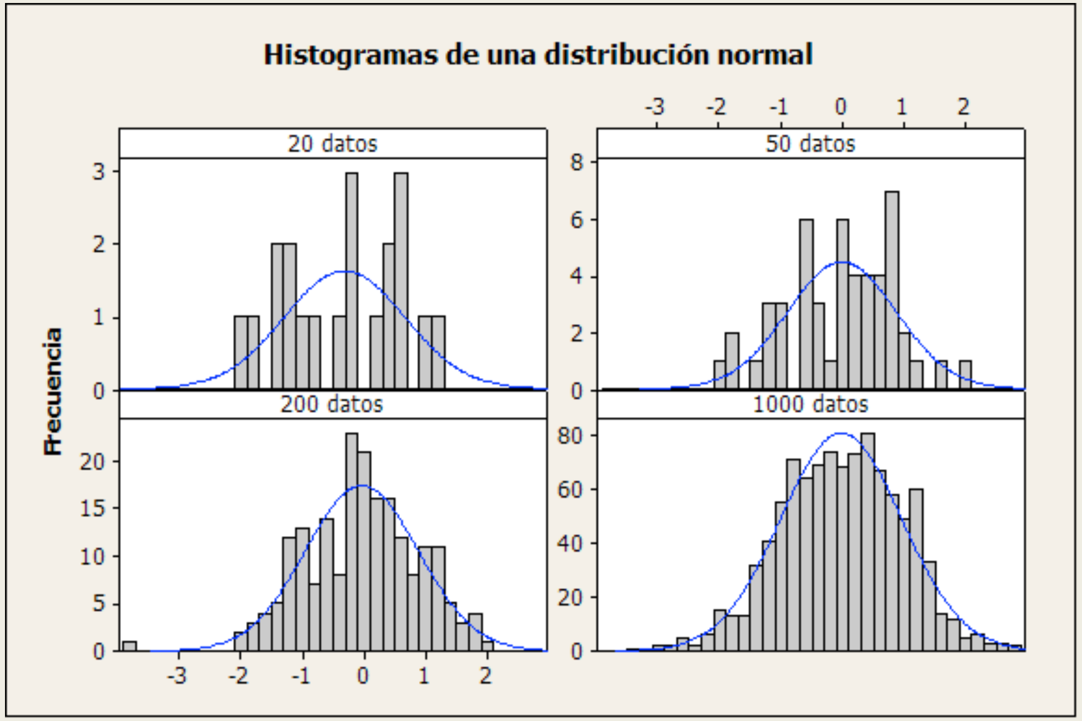
Estadística Descriptiva

Dr. Francisco Rabadán Pérez

Índice

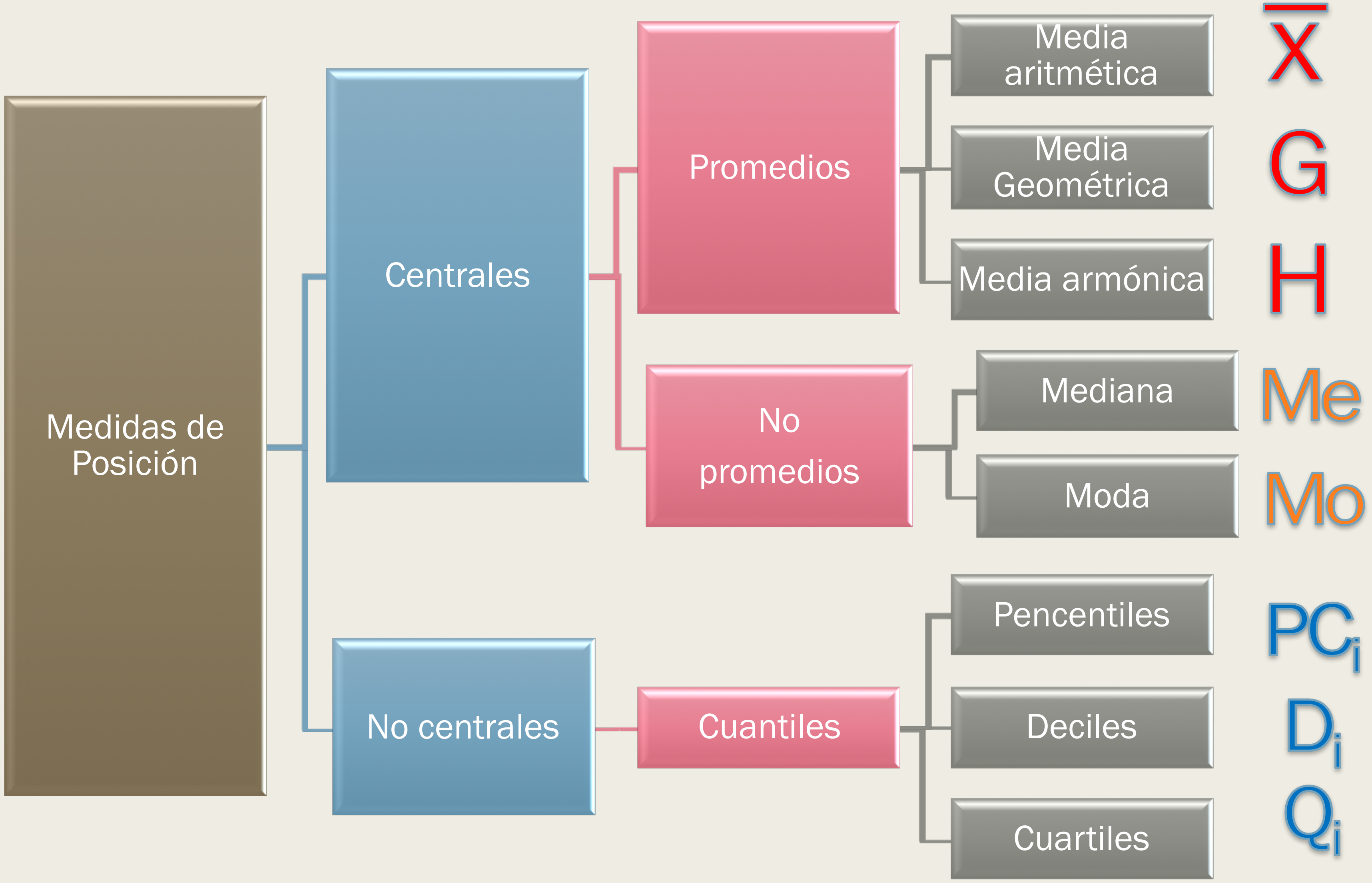
1. Medidas de Posición

- **Objetivo:** resumir o sintetizar la información de los datos.
- La **medida de** posición ideal sería aquella que pudiera sustituir a todos los datos (**valores** y frecuencias) sin que perdiéramos información (capacidad de descripción).
- Nos interesan los valores centrales (porque es donde se suelen acumular mayor cantidad de datos), pero también los extremos (donde suelen aparecer los casos atípicos).



Fuente: http://www.caletec.com/blog/6sigma/histogramas-y-normalidad-de-los-datos/

1. Medidas de Posición



2. Media Aritmética $\bar{x}$

Cálculo

$$\bar{x} = \sum_{i=1}^{n} \frac{x_i}{N}$$

$$\bar{x} = \sum_{i=1}^{n} \frac{x_i n_i}{N}$$

$$\bar{x} = \frac{\sum_{i=1}^{n} x_i w_i}{\sum_{i=1}^{n} w_i}$$

Partiendo de los datos originales

- sumamos n_i veces x_i

Partiendo de la distribución de frecuencias

- En datos agrupados tomamos la marca de clase (x_i)

Media aritmética ponderada

- w_i: ponderación

2. Media Aritmética $\bar{x}$

Ventajas

- Tiene en cuentas todos los datos
- Fácil Cálculo
- Es única
- Centro de gravedad

Inconvenientes

- Afectada por los valores extremos

(Martín Pliego, 2011; pág. 41)

2. Media Aritmética $\bar{X}$

Propiedades

1. Centro de gravedad de la distribución: La suma de las desviaciones de los valores de la variable respecto a la media es cero.

$$\sum_{i=1}^{n}(x_i-\bar{x})\,n_i=\sum_{i=1}^{n}x_in_i-\sum_{i=1}^{n}\bar{x}n_i=\sum_{i=1}^{n}x_in_i-\bar{x}\sum_{i=1}^{n}n_i=N\frac{\sum_{i=1}^{n}x_in_i}{N}-\bar{x}N=N\bar{x}-\bar{x}N=0$$

2. Afectada por los cambios de origen $y = x + k \rightarrow \bar{y}=\bar{x}+k$

3. Afectada por los cambios de escala $y = kx \rightarrow \bar{y}=k\bar{x}$

4. Hace mínima las desviaciones cuadráticas (Prop.2. Martín Pliego, 2011)

Mas en (Martín Pliego, 2011; pag. 39-41) .... N.E. 2016/2017

3. Media geométrica G

- **Definición:** Raíz n-ésima del producto de los N valores de la variable de la distribución

Cálculo

$$G = \sqrt[N]{{x_1}^{n_1} \cdot {x_2}^{n_2} \cdot \cdots \cdot {x_n}^{n_n}} = \sqrt[N]{\prod_{i=1}^{n} {x_i}^{n_i}}$$

Comprendiendo
la geométrica:

Cambiamos la suma por la multiplicación.

- **Ponderar:** en vez de multiplicar por la frecuencia, elevamos a la frecuencia (multiplicar n_i veces x_i)
- **Repartir:** En vez de dividir por N, hacemos la raíz N-ésima.

3. Media geométrica G

Ventajas

- Tiene en cuentas todos los datos
- Menos afectada por los valores extremos

Inconvenientes

- Significado estadístico poco intuitivo
- Computo mas difícil
- Posibilidad de indeterminación
  - Problemas con el cero y números negativos

Recomendado para:

- Porcentajes
- Tasas
- Números Índices
- Variaciones Acumulativas

(Martín Pliego, 2011; pág. 43)

3. Media geométrica G

Incrementos de magnitud

Periodo	incremento	x_i
T1	8%	$X_1=8$
T2	12%	$X_2=12$
T3	18%	$X_3=18$
T4	27%	$X_4=27$
T5	40,5%	$X_5=40,5$
T6	60,75%	$X_6=60,75$

Los valores de la distribución representan incrementos de una magnitud trimestral→ media geométrica

$$G = \sqrt[6]{8 \cdot 12 \cdot 18 \cdot 27 \cdot 40,5 \cdot 60,75}$$
$$= \sqrt[6]{114.791.256} = 22,0454$$

- Significado:
  - *Todos estos aumentos equivalen a un aumento en cada trimestre de 22.0454%*

3. Media geométrica G

Tipos de interés

$$C_1 = C_0(1+i_1)$$
$$C_2 = C_1(1+i_2) = C_0(1+i_1)(1+i_2)$$
$$C_n = C_{n-1}(1+i_n) = C_0(1+i_1)(1+i_2)\dots(1+i_n)$$

- Tipo medio equivalente: $C_n = C_0(1+i)^n$
- Igualamos $\cancel{C_0}(1+i_1)(1+i_2)\dots(1+i_n) = \cancel{C_0}(1+i)^n$
- Despejando

$$\boldsymbol{i = \sqrt[n]{(1+i_1)(1+i_2)\dots(1+i_n)} - 1}$$

IMPORTANTE

(Martín Pliego, 2011; pág. 44)

3. Media geométrica G

Propiedades

1. El logaritmo de la media geométrica es la media aritmética de los logaritmos de los valores de la variable

$$Log\ G = \frac{1}{N}\sum_{i=1}^{n} Log(x_i) \cdot n_i$$

Mas en (Martín Pliego, 2011; pag. 42) .... N.E. 2016/2017

4. Media Armónica H

Cálculo

$$H = \frac{N}{\frac{1}{x_1}n_1 + \frac{1}{x_2}n_2 + \cdots \frac{1}{x_n}n_n} = \frac{N}{\sum_{i=1}^{n}\frac{n_i}{x_i}}$$

- **Nota:** La inversa de H es la Media aritmética de los inversos de los valores de la variable

Más fácil

$$\frac{1}{H} = \frac{\sum_{i=1}^{n}\frac{n_i}{x_i}}{N}$$

Comprendiendo la armónica:

- Trabajamos con la inversa de la variable.
- Ejemplo:
  - 9 Km/h (en un hora recorremos 9 Km)
  - 1/9 h/km (tardamos 1/9 horas en recorrer un kilómetro)

4. Media Armónica H

Ventajas

Contribuyen todos los valores

Mas representativa cuando hablamos de variables cociente de magnitudes simples

Inconvenientes

Queda afectada por valores pequeños

Indeterminada con algún valor cero

Recomendado para:

variables que están relacionadas de forma inversa con el tiempo

- velocidad, rendimiento,...

No recomendada para valores muy pequeños

(Martín Pliego, 2011; pág. 45)

4.2. Relación entre estos promedios

Se verifica que

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

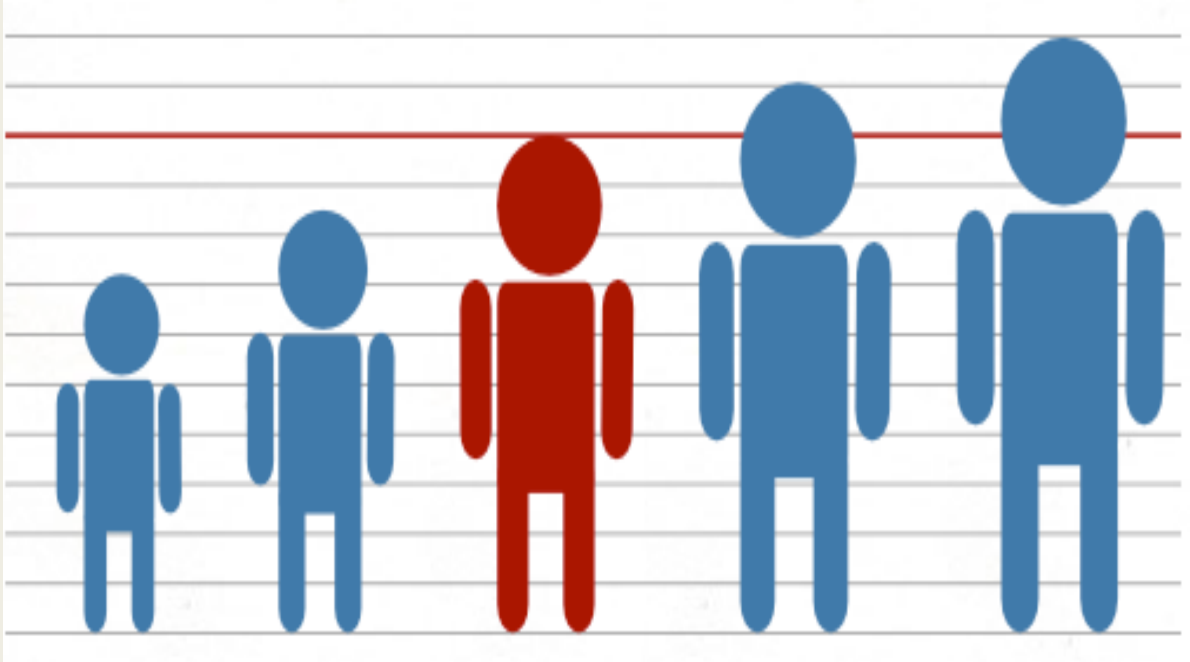
(Martín Pliego, 2011; pág. 46)

5. Mediana Me

Definición: valor de la distribución ordenada de menor a mayor que deja a su izquierda y su derecha el mismo número de frecuencias (individuos).

n impar — $Me = x_{n/2}$

n par — $Me = \dfrac{x_{N/2} + x_{(N+1)/2}}{2}$



Ch 3.15

Mas en (Martín Pliego, 2011; pag. 42) .... N.E. 2016/2017

5. Mediana Me

Ejemplo (Martín-Pliego,2011-pág.48)

x_i	n_i	N_i
1	3	3
2	4	7
5	9	16
7	10	26
10	7	33
13	2	35
	35	

donde

$$\frac{N}{2} = \frac{35}{2} = 17{,}5\,,$$

y el centro será el que ocupe el lugar décimo-octavo, con lo que la mediana será

$$Me = 7$$

Mas en (Martín Pliego, 2011; pag. 42) .... N.E. 2016/2017

5. Mediana Me

Datos agrupados:

- Martín-Pliego estudia la posición teórica de la mediana bajo el presupuesto de uniformidad en la distribución
- Nosotros hablaremos de intervalo mediano: aquel donde se encuentra la frecuencia acumulada N/2.

Mas en (Martín Pliego, 2011; pag. 49) …. N.E. 2016/2017

5. Mediana Me

Ventajas

No le afectan los valores extremos

Inconvenientes

No tiene en cuenta el peso de los valores (xi)

No minimiza las desviaciones absolutas

Recomendado para:

cuando la media no sea representativa

- Lo veremos en DISPERSIÓN

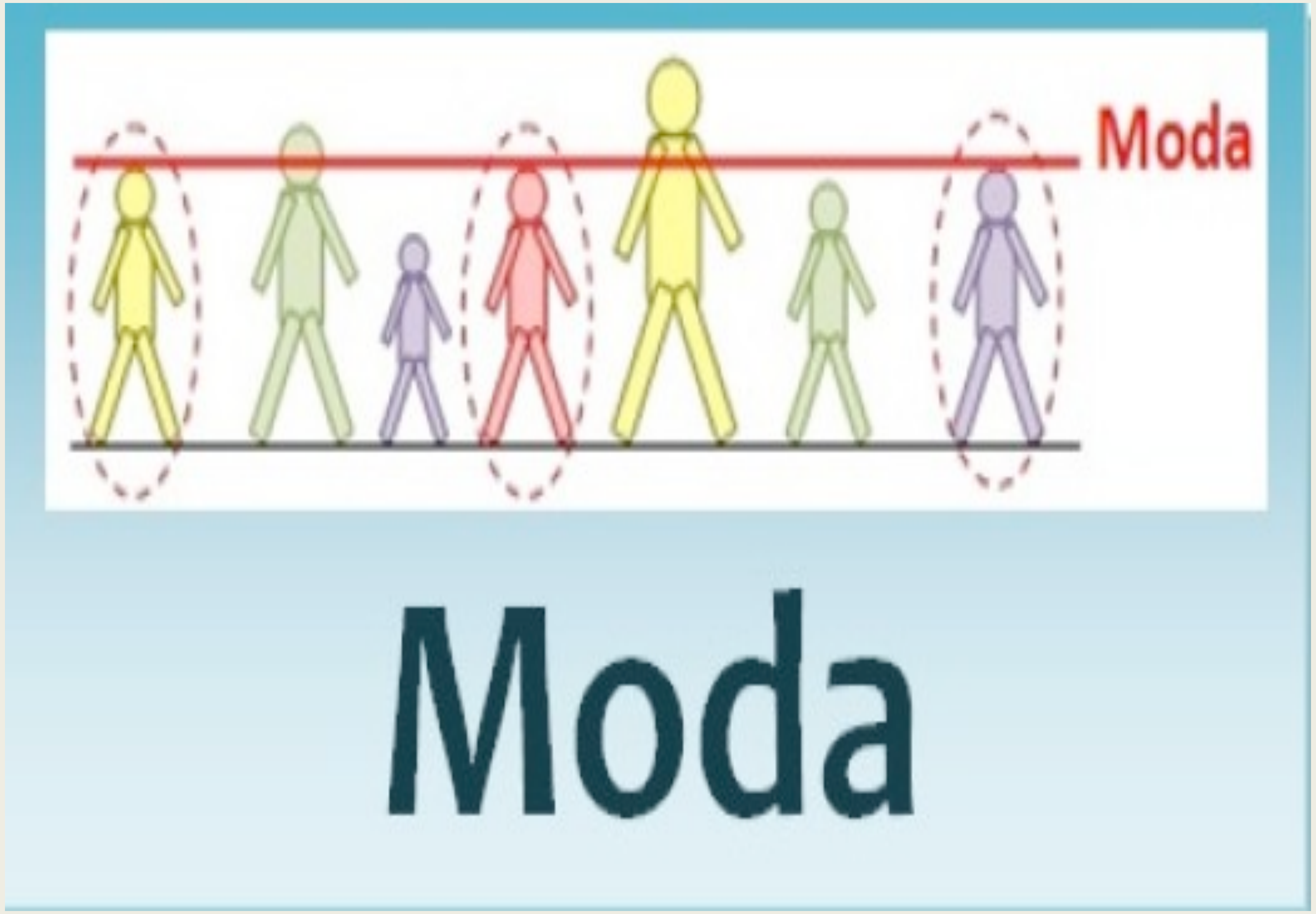
6. Moda

Mo

Definición: valor de la distribución que más se repite.

Cálculo: valor de la variable con frecuencia absoluta máxima.

Intervalos: Hablaremos de intervalo modal (por la misma razón que en la mediana)



6. Moda

Mo

Ventajas

?

Inconvenientes

No tiene porque ser única

No tiene en cuenta el resto de valores ni sus frecuencias

Recomendado para:

Como dato curioso

Cuando su frecuencia absoluta es superior a N/2 y **porque coincide con la mediana**

7. Medidas de Posición no centrales

Cuantiles (Q_i) : n-1 valores (x_i)que dividen las distribución en n partes iguales respecto al número de individuos.

Cuartiles:
- C_1,C_2 y C_3
- *Dividen la distribución en cuatro partes iguales*
- C_2=*Me*

Deciles:
- D_1,D_2,...D_9
- Dividen la distribución en 10 partes iguales
- D5=Me

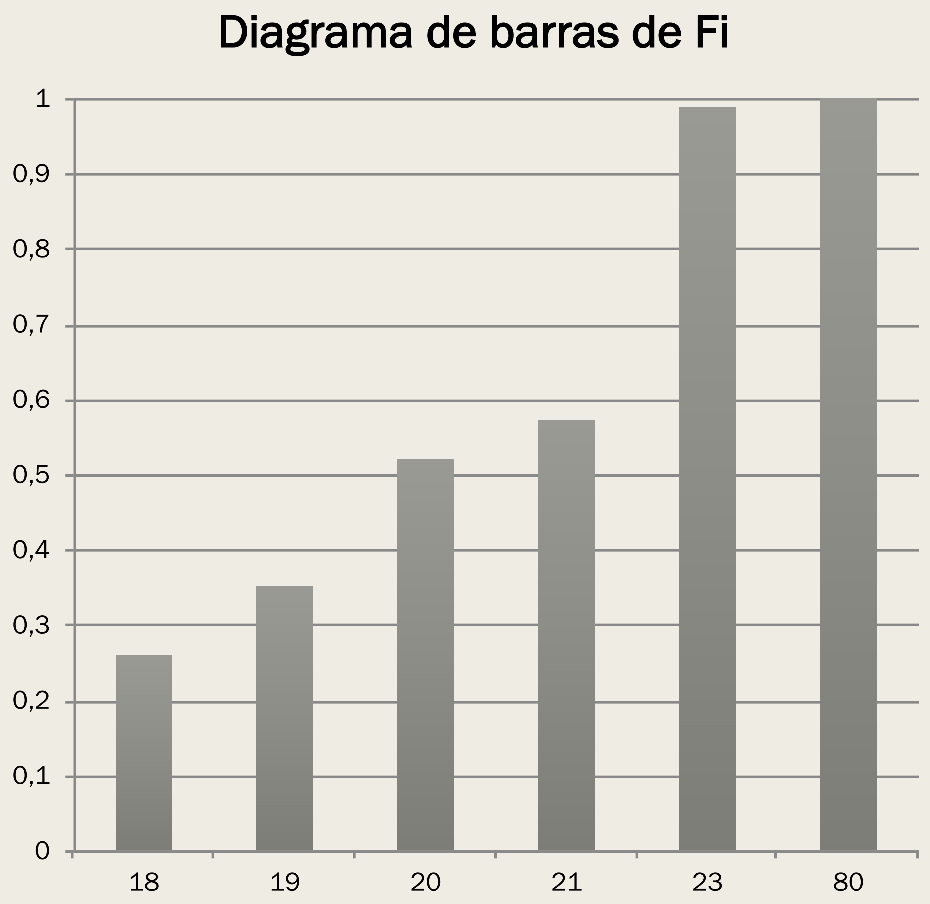
Percentiles:
- PC_1,PC_2,..PC_{99}
- Dividen la distribución en cien partes iguales

7. Medidas de Posición no centrales

En el diagrama de Frecuencias relativas acumuladas podemos identificar la posición (eje vertical) y buscar el valor x_i (eje horizontal)

Posición	Cuantiles	Valor Cuantil
0,5	$Me=C_2=D_5=PC_{50}$	20
0,25	$C_1=PC_{25}$	18
0,75	$C_3=PC_{75}$	23

- En caso de duda consultamos la distribución de frecuencias (xi, Fi)
- Notese que las medidas de posición no centrales pueden ser iguales, incluso, para todo el recorrido de la variable.



Fuente: elaboración propia
X: Edad de los alumnos

7. Medidas de posición no centrales

Ventajas	Inconvenientes	Recomendado para:
No les afectan los valores extremos Conjuntamente dan una elevada capacidad de descripción	No tiene en cuenta los valores (x_i)	Describir el comportamiento de la distribución por tramos

Apéndice: momentos de una distribución de frecuencias

Respecto al origen	$r=_0 \rightarrow a_0=1$
	$r=_1 \rightarrow a_1=\bar{x}$
Respecto a la media	$r=0 \rightarrow m_0=1$
	$r=1 \rightarrow m_1=0$
	$R=2 \rightarrow m_2=\sigma^2$

$$a_r = \sum_{i=1}^{n} \frac{(x_i)^r n_i}{N}$$

$$m_r = \sum_{i=1}^{n} \frac{(x_i - \bar{x})^r n_i}{N}$$

Mas en (Martín Pliego, 2011; pag. 68-72)

Apéndice: momentos de una distribución de frecuencias

- Los momentos respecto a la media y respecto al origen se diferencian en un cambio de escala.
  - *Los a_r tienen como origen el cero, y los m_r en $\bar{x}$*
- Los momentos respecto a la media se pueden calcular partiendo de los momentos respecto al origen aplicando el binomio de Newton.

$$m_0 = a_0$$

$$m_1 = a_1 - a_1$$

$$m_2 = a_2 - {a_1}^2$$

IMPORTANTE

Textos recomendados

- Martín-Pliego, *Introducción a la Estadística Económica y Empresarial*, Editorial AC, 2011, 3ª Edición

Prácticas recomendadas

- Ejercicios: pág 61 -67 (Martín-Pliego,2011)
- Ejercicios resueltos en clase
- Prácticas y recursos web (aula virtual)