TEMA IIB MEDIDAS DE POSICIÓN

IIB.4.- MEDIAS.
- média aritmética $\bar{x}$
- média harmónica H
- média xeométrica G

IIB.5.- MEDIANA
IIB.6.- MODA
IIB.7.- CUANTÍS
IIB.8.- MOMENTOS

IIB.4.- MEDIAS.

Son valores que representan a variábel ou algunha característica da variábel.

-Carácter central
-Carácter non central

Carácter central: un valor en torno ao cal están agrupados os valores desta variábel. Sérve-nos para resumir a información que nos proporciona a variábel. En troques de dar todos os valores, damos un valor meio para representar esa variábel.

Sexa x a variábel en estudo, con modalidades $x_1, x_2, ..., x_m$ e frecuéncias absolutas respectivas $n_1, n_2, ..., n_m$ e sexa m o número de modalidades distintas e n o tamaño mostral:

- **Média aritmética $\bar{x}$**

$$\boxed{\bar{x} = \sum_{i=1}^{m} x_i \cdot n_i / n = \sum_{i=1}^{m} x_i \cdot f_i}$$

4 propriedades da média aritmética

1) $$\sum_{i=1}^{m} (x_i - \bar{x}) \cdot n_i = 0$$

dem.) $$\sum_{i=1}^{m} (x_i - \bar{x}) \cdot n_i = \sum_{i=1}^{m} x_i \cdot n_i - \sum_{i=1}^{m} \bar{x} \cdot n_i = \sum_{i=1}^{m} x_i \cdot (f_i \cdot m) - \bar{x} \cdot \sum_{i=1}^{m} n_i =$$

$$= m \cdot \sum_{i=1}^{m} x_i \cdot f_i - \bar{x} \cdot \sum_{i=1}^{m} n_i = m\,\bar{x} - \bar{x} \cdot m = 0$$

2) $$\min_{b \in R} \sum_{i=1}^{m} (x_1 - b)^2 \cdot n_i = \sum_{i=1}^{m} (x_1 - \bar{x})^2 \cdot n_i$$

A diferéncia entre valores e média resultará nula porque compensamos as diferéncias positivas coas negativas. Obviamente isto non quere dicer que todos os valores sexan iguais á média.

3) $$\bar{x} = \frac{N_1 \cdot \bar{x}_1 + N_2 \cdot \bar{x}_2}{N}$$

Se coñecemos a média de vários sunconxuntos podemos calcular a média da povoación total.

4) $y = a + bx$ ($a,b \in R$; ecuación dunha recta) $\Rightarrow \bar{y} = a + b \cdot \bar{x}$

A média é transparente a cambios lineais (sumas, produtos):

$$\bar{y} = \sum_{i=1}^{m} y_i \cdot n_i/n = \sum_{i=1}^{m} (a+bx_i) \cdot n_i/n = a \cdot \sum_{i=1}^{m} n_i/n + b \cdot \sum_{i=1}^{m} x_i \cdot n_i/n = a + b \cdot \bar{x}$$

- **Média harmónica H**

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

Empréga-se para achar a média de porcentaxes e promedios.
Multiplicando por 1/n en numerador e denomidador temos:

$$\boxed{H = \frac{1}{\sum_{i=1}^{m} \frac{f_i}{x_i}}}$$

Exemplo: queremos calcular a média de velocidade de duas viaxes, ir de A a B (70km/h) e de volta, de B a A (40km/h):

Se calculamos a média aritmética temos 70+40/2 = 55km/h

pero en realidade o axeitado é

$$v = e / t = 2d / (t_1+t_2) = \frac{2d}{\frac{d}{70} + \frac{d}{40}} = \frac{2d}{\frac{4d+7d}{280}} = \frac{560}{11} = 50'9\text{km/h}$$

[onde d = distáncia entre A e B]

- **Média Xeométrica G**

$$G = \sqrt[n]{\prod_{i=1}^{m} x_i^{n_i}}$$

Úsa-se como média de índices e razóns económicas.
Substituíndo n_i por $f_i \cdot n$:

$$G = \prod_{i=1}^{m} x_i^{f_i}$$

Número de índice: cociente do valor dunha variábel nun tempo partido polo valor inicial.

Exemplo: Queremos calcular o rendemento meio deste período:

Ano	Rendemento
78	12%
79	10%
80	7%
81	6%
82	5%

$G = 0'05^{1/5} \cdot 0'06^{1/5} \cdot 0'07^{1/5} \cdot 0'10^{1/5} \cdot 0'12^{1/5} = 0'075906699$

$G = \sqrt[5]{0'05 \cdot 0'06 \cdot 0'07 \cdot 0'10 \cdot 0'12} = 0'75906699$

Exemplo: Durante cinco anos os prezos aos que unha persoa mercou a gasolina para o seu coche foron:

Prezo
70pts/l
80pts/l
75pts/l
90pts/l
95pts/l

Cal foi o prezo promedio da gasolina neste periodo...
a) ... se cada ano mercou a mesma cantidade de gasolina?
b) ... se cada ano gastou a mesma cantidade de diñeiro?

a) média aritmética (igual pts, numerador)

$(70+80+75+90+95)/5 = 82$pts/l

b) média harmónica (igual l, denominador)

$1/(\overline{1/x}) = (\overline{x^{-1}})^{-1}$
$= (1/70 \cdot 1/5 + 1/80 \cdot 1/5 + 1/75 \cdot 1/5 + 1/90 \cdot 1/5 + 1/95 \cdot 1/5)^{-1} =$
$= 80'96$pts/l

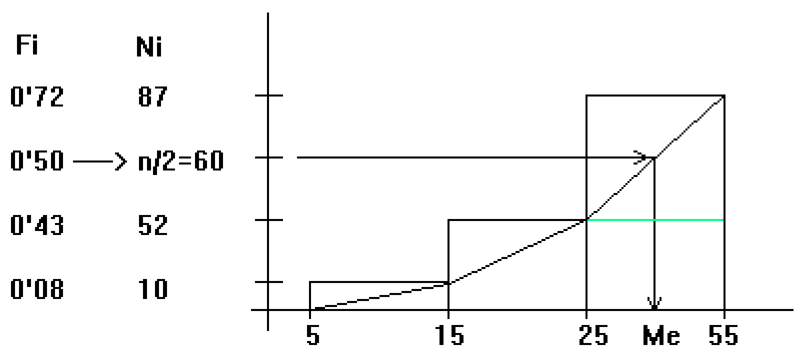
IIB.5.- MEDIANA.

É o valor que deixa, tendo ordenados os valores observados, á esquerda e á dereita o mesmo número de valores. Se o número de observacións (tamaño mostral) é par e non se teñen as observacións agrupadas en clases, a mediana será o valor $[x_{(n \bmod 2 -1)}+x_{(n \bmod 2 +1)}]/2$ —média aritmética dos dous valores centrais—. Sen clases e con n impar, a mediana será o valor $x_{(n \bmod 2 +1)}$ —valor central—. Se as observacións están agrupadas en clases haberá que facer un diagrama acumulativo e interpolar linealmente.

Intervalo Mediano: é aquel intervalo que tén frecuéncia acumulada $\geq n/2$, ousexa aquel que tén frecuéncia acumulada relativa $\geq 0'50$ (*)

Exemplo:

Ventas	Nºfamilias	Ni	Fi
5-15	10	10	0'0833
15-25	42	52	0'4333
25-55	35	87	**0'7250** (*)
55-75	20	97	0'8083
75-95	13	120	1'0000
	n=120		



[trazar sempre as diagonais por dentro dos rectángulos]

$$\frac{87-52}{55-25} = \frac{60-52}{Me-25} \quad // \; 35 \cdot Me-875 = 240 \; // \; Me = (240+875)/35 = 31'857$$

IIB.6.- MODA.

É o valor que mais veces aparece, o valor que tén unha frecuéncia maior.

Exemplo:

x_i	n_i
1	2
2	7 → Mo = 2 (distribución unimodal)
3	3
4	4

x_i	n_i
1	2
2	7 → Mo=2
3	7 → Mo=3 (distribución bimodal)

4 4

Por intervalos: determinar o intervalo modal, é obviamente o que tén maior frecuéncia. Como determinar o ponto que se corresponde coa moda nesa frecuéncia? É a média aritmética dos valores extremos do intervalo modal.

Exemplo:

Ventas	Nºfamilias	
5-15	10	
15-25	**42**	← intervalo modal ⇒ moda = (15+25)/2 = 20
25-55	35	
55-75	20	
75-95	13	

Inconvintes da moda: pode haber muitas á vez.
Avantaxes da moda: emprega todos os valores dos que se dispón. Para a mediana só utilizamos os que están á esquerda e os que están á dereita. Só nos importa o número de veces que se repiten.

A média é mui sensíbel a valores atípicos (mui grandes ou mui pequenos). Polo contrário, a mediana é unha medida ***robusta***, xa que non se ve afectada por estes valores. A solución é calcular tanto mediana como média para ver como se comporta a variábel.

IIB.7.- CUANTÍS.

Son valores que dividen a distribución en partes iguais; dependendo de cantas partes iguais distinguiremos cuartís, quintís, decís e percentís.

CUARTÍS

son tres valores Q_1, Q_2, Q_3 que dividen a distribución en catro partes iguais (contendo fracción o 25% dos valores). Q_1 deixa á sua esquerda o 25% dos dados, as observacións menores, e o 75%, as maiores, á sua dereita. Q_2 é a mediana. Q_3 deixa o 75% dos valores á sua esquerda (os menores), e as observacións maiores en porcentaxe de 25%, están á sua dereita. Todo isto cos dados ordenados crecentemente, claro.

Exemplo:

x_i	n_i
1	2
2	25
7	5
8	8
	n=40

$0.25\cdot n=10 \rightarrow Q_1=[x_{(10)}+x_{(11)}]/2=(2+2)/2=2$
$0.50\cdot n=20 \rightarrow Q_1=[x_{(20)}+x_{(21)}]/2=(2+2)/2=2$
$0.75\cdot n=30 \rightarrow Q_1=[x_{(30)}+x_{(31)}]/2=(7+7)/2=7$

QUINTÍS

catro valores que dividen a distribución en cinco partes iguais:
$x_{(1)}$ — 20% — $\mathbf{Q_1}$ — 20% — $\mathbf{Q_2}$ — 20% — $\mathbf{Q_3}$ — 20% — $\mathbf{Q_4}$ — 20% — $x_{(n)}$

DECÍS

nove valores que dividen a distribución en dez partes iguais.

CENTÍS

son 99 valores que dividen a distribución en 100 partes iguais.
Dá intervalos do 1%.

Exemplo:

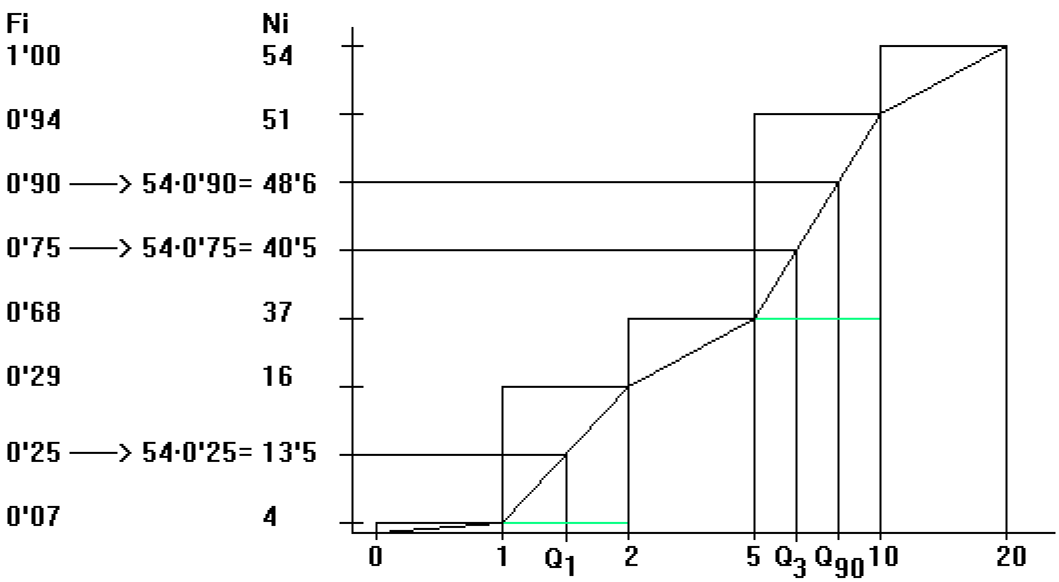
Nunha axéncia de aluger de coches, sobre unha mostra de 54 clientes obtívo-se a seguinte distribución do número de dias que é alugado un automóbil na tempada alta. Calcular o primeiro e terceiro cuartil, o derradeiro decil e interpretá-los.

(L_{i-1},L_i)	n_i	N_i
(0,1]	4	4
(1,2]	12	16
(2,5]	21	37
(5,10]	14	51
(10,20]	3	54
	n=54	

Analogamente ao visto para o cálculo da mediana, como a distribución está agrupada en clases, haberá que aplicar interpolación lineal. Calculamos as frecuéncias acumuladas relativas:

(L_{i-1},L_i)	N_i	F_i
(0,1]	4	0’074
(1,2]	16	0’296
(2,5]	37	0’685
(5,10]	51	0’944
(10,20]	54	1’000

primeiro cuartil: 0’25
terceiro cuartil: 0’75
derradeiro decil: 0’90



$$\frac{51-37}{10-5} = \frac{48'6-37}{Q_{90}-5} \Rightarrow Q_{90} = 9'142857$$

$$\frac{51-37}{10-5} = \frac{40'5-37}{Q_3-5} \Rightarrow Q_3 = 6'25$$

$$\frac{16-4}{2-1} = \frac{13'5-4}{Q_1-1} \Rightarrow Q_1 = 1'79167$$

Interpretación: o 25% dos que alugan coches fan-o durante menos de dous dias. O 75% dos que alugan coches fan-o menos de sete dias. O 90% dos que alugan coches fan-o menos de dez dias. O 10% restante fan-o mais de dez dias.

IIB.8.- MOMENTOS.

Son medidas que caracterizan unha distribución de frecuéncias. Existe unha relación biunívoca entre unha distribución e o seu momento.

MOMENTOS RESPEITO Á ORIXE (respeito ao cero)

$$\boxed{a_r = \sum_{i=1}^{m} x_i^r \cdot f_i}$$

$$a_1 = \bar{x}$$

MOMENTOS RESPEITO Á MÉDIA

$$\boxed{m_r = \sum_{i=1}^{m} (x_i - \bar{x})^r \cdot f_i}$$

$$m_2 = \underline{\text{varianza}} = \sum_{i=1}^{m} (x_i - \bar{x})^2 \cdot f_i$$

A varianza é mais cómodo calculá-la así:

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

obtención desta expresión:

$$m_2 = \sum_{i=1}^{m} (x_i - \bar{x})^2 \cdot f_i = \sum_{i=1}^{m} (x_i^2 - 2 \cdot x_i \cdot \bar{x} + \bar{x}^2) \cdot f_i = \sum_{i=1}^{m} (x_i^2 \cdot f_i - 2 \cdot x_i \cdot \bar{x} \cdot f_i + \bar{x}^2 \cdot f_i) =$$

$$= \sum_{i=1}^{m} x_i^2 \cdot f_i - 2 \cdot \bar{x} \sum_{i=1}^{m} x_i \cdot f_i + \sum_{i=1}^{m} \bar{x}^2 \cdot f_i = a_2 - 2 \cdot \bar{x} \cdot \bar{x} + \bar{x}^2 = a_2 - \bar{x}^2 = a_2 - a_1^2$$

$$\mathbf{m_3 = a_3 - 3a_1a_2 + 2a_1^3}$$

obtención desta expresión:

$$m_3 = \sum_{i=1}^{m} (x_i - \bar{x})^3 \cdot f_i = \sum_{i=1}^{m} (x_i^3 - 3x_i^2\,\bar{x} + 3x_i\,\bar{x}^2 - \bar{x}^3) \cdot f_i =$$

$$= \sum_{i=1}^{m} x_i^3 \cdot f_i - 3\,\bar{x} \sum_{i=1}^{m} x_i^2 \cdot f_i + 3\,\bar{x}^2 \sum_{i=1}^{m} x_i \cdot f_i - \bar{x}^3 \cdot \sum_{i=1}^{m} f_i =$$

$$= a_3 - 3\bar{x}\cdot a_2 + 3\cdot a_1^2\cdot \bar{x} - \bar{x}^3 = a_3 - 3a_1\cdot a_2 + 3\cdot a_1^2\cdot a_1 - a_1^3 =$$

$$= a_3 - 3a_1a_2 + 2a_1^3$$