Departamento de Métodos Cuantitativos para la Economía y la Empresa
GRADO EN ADMINISTRACIÓN Y DIRECCION DE EMPRESAS

2º

- Técnicas Cuantitativas 1
- Técnicas Cuantitativas 2

3º

- Econometría
- Métodos Cuantitativos

1

1

Técnicas Cuantitativas 1

- Estadística Descriptiva
  Temas: 1, 2, 3 y 4
- Probabilidad (tema 5)
  Variables aleatorias (tema 6)
  Discretas (tema 7) y
  Continuas

Técnicas Cuantitativas 2

- Inferencia Estadística

2

2

Estadística Descriptiva

Variables estadísticas
Temas: 1 y 2

Unidimensionales
Tema 1

Bidimensionales
Tema 2

Números índice. Tema 3

Series cronológicas. Tema 4

3

3

Probabilidad

Probabilidad. Temas 5

Variables aleatorias. Tema 6

Distribuciones discretas de probabilidad.
Tema 7

4

4

1.- Variables estadísticas unidimensionales.

5

Estadística (status)
Estadística Descriptiva

Tipos de fenómenos:
- *Fenómenos causales o determinísticos:*
- ***Fenómenos aleatorios o estadísticos:***

Población / muestra.
- *Evitar la destrucción de la población.*
- *Rapidez.*
- *Economía .*

Variable / atributo
- **Variables discretas**
- **Variables continuas**

- ***pocas modalidades*** (discretas,…)
- ***gran número de valores*** (variables continuas,…)

6

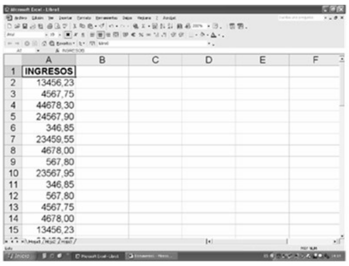
TABLAS ESTADÍSTICAS

TIPO	
Colectiva	56
Unifamiliar	14
total	70

Número de dormitorios	
1	7
2	14
3	21
4	21
5	7
total	70

Salario/hora	
0-10	25
10-20	40
20-40	20
40-50	15
total	100

x_i
x_1
x_2
...
x_n



- **frecuencia absoluta** n_i $\quad \sum_{i=1}^{k} n_i = n$
- **frecuencia relativa** f_i $\quad f_i = \dfrac{n_i}{n}$
- **frecuencia absoluta acumulada** N_i $\quad N_i = \sum_{j=1}^{i} n_j$
- **frecuencia relativa acumulada** F_i

distribución de frecuencias

$L_{i-1} - L_i$	x_i	n_i	f_i	N_i	F_i
$L_0 - L_1$	x_1	n_1	f_1	N_1	F_1
$L_1 - L_2$	x_2	n_2	f_2	N_2	F_2
...	...	...	...	...	...
$L_{k-1} - L_k$	x_k	n_k	f_k	N_k	F_k
		n	1		

x_i	n_i	f_i
Colectiva	56	0,80
Unifamiliar	14	0,20
total	70	1

x_i	n_i	N_i	f_i	F_i
1	7	7	0,10	0,10
2	14	21	0,20	0,30
3	21	42	0,30	0,60
4	21	63	0,30	0,90
5	7	70	0,10	1
total	70		1	

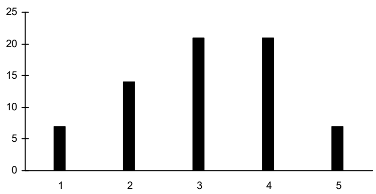
$L_{i-1} - L_i$	x_i	n_i	f_i	N_i	F_i
0-10	5	25	0,25	25	0,25
10-20	15	40	0,40	65	0,65
20-40	30	20	0,20	85	0,85
40-50	45	15	0,15	100	1
total		100	1		

9

REPRESENTACIONES GRÁFICAS

DIAGRAMA DE BARRAS

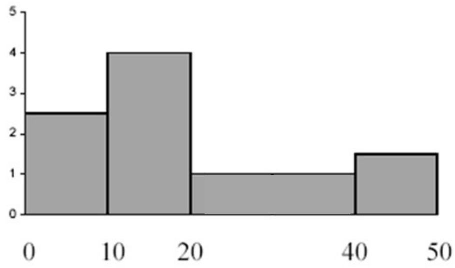
Número de dormitorios	n_i
1	7
2	14
3	21
4	21
5	7



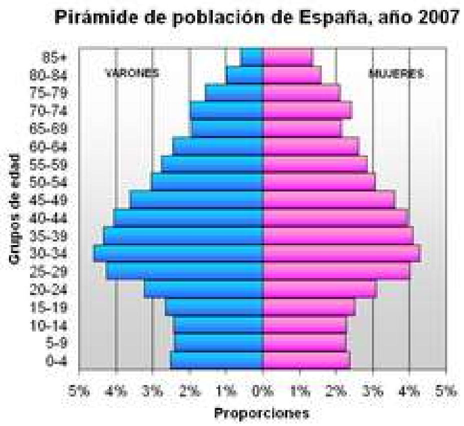
10

HISTOGRAMA

$L_{i-1}-L_i$	n_i	a_i	h_i
0-10	25	10	2,5
10-20	40	10	4
20-40	20	20	1
40-50	15	10	1,5



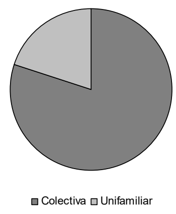
11



12

DIAGRAMA DE SECTORES

x_i	n_i	f_i	sector
Colectiva	56	0,80	0,80x360º=288º
Unifamiliar	14	0,20	0,20x360º=72º
total	70	1	360º



13

x_i
4
5
7
8

$$\sum_{i=1}^{4} x_i = 24 \qquad \bar{x} = \frac{1}{4}\sum_{i=1}^{4} x_i = \frac{24}{4} = 6$$

x_i	n_i	$x_i\, n_i$
4	2	8
5	3	15
7	1	7
8	2	16

$$\sum_{i=1}^{4} x_i n_i = 46 \qquad \bar{x} = \frac{1}{8}\sum_{i=1}^{4} x_i n_i = \frac{46}{8} = 5,75$$

14

Momentos no centrados.

$$a_r = \frac{1}{n}\sum_{i=1}^{k} x_i^r n_i = \sum_{i=1}^{k} x_i^r f_i$$ (para tablas con frecuencias)

$$a_r = \frac{1}{n}\sum_{i=1}^{n} x_i^r$$ (para tablas sin frecuencias)

Nota: $a_1 = \overline{x}$.

Momentos centrados.

$$m_r = \frac{1}{n}\sum_{i=1}^{k} \left(x_i - \overline{x}\right)^r n_i = \sum_{i=1}^{k} \left(x_i - \overline{x}\right)^r f_i$$ (para tablas con frecuencias)

$$m_r = \frac{1}{n}\sum_{i=1}^{n} \left(x_i - \overline{x}\right)^r$$ (para tablas sin frecuencias)

$$S^2 = m_2 = \frac{\sum_{i=1}^{k} \left(x_i - \overline{x}\right)^2 n_i}{n} = \frac{\sum_{i=1}^{k} x_i^2 n_i}{n} + \frac{\sum_{i=1}^{k} \overline{x}^2 n_i}{n} - \frac{\sum_{i=1}^{k} 2\overline{x} x_i n_i}{n} = \frac{\sum_{i=1}^{k} x_i^2 n_i}{n} + \frac{\overline{x}^2 \sum_{i=1}^{k} n_i}{n} - \frac{2\overline{x}\sum_{i=1}^{k} x_i n_i}{n} =$$

$$= a_2 + \overline{x}^2 - 2\overline{x} a_1 = a_2 + a_1^2 - 2a_1 a_1 = a_2 - a_1^2$$

$$m_3 = a_3 - 3a_2 a_1 + 2a_1^3 \qquad m_4 = a_4 - 4a_3 a_1 + 6a_2 a_1^2 - 3a_1^4$$

15

Cálculo de los momentos.

Momentos no centrados

Tabla sin frecuencias.

x_i	x_i^2	x_i^3	x_i^4
3	9	27	81
3	9	27	81
2	4	8	16
2	4	8	16
3	9	27	81
5	25	125	625
18	60	222	900

$$\overline{x} = a_1 = \frac{1}{n}\sum_{i=1}^{n} x_i = \frac{18}{6} = 3 \qquad a_2 = \frac{1}{n}\sum_{i=1}^{n} x_i^2 = \frac{60}{6} = 10$$

$$a_3 = \frac{1}{n}\sum_{i=1}^{n} x_i^3 = \frac{222}{6} = 37 \qquad a_4 = \frac{1}{n}\sum_{i=1}^{n} x_i^4 = \frac{900}{6} = 150$$

16

Tabla con frecuencias. Variable discreta.

x_i	n_i	$x_i n_i$	$x_i^2 n_i$	$x_i^3 n_i$	$x_i^4 n_i$
1	7	7	7	7	7
2	14	28	56	112	224
3	21	63	189	567	1701
4	21	84	336	1344	5376
5	7	35	175	875	4375
total	n=70	217	763	2905	11683

$$\bar{x} = a_1 = \frac{1}{n}\sum_{i=1}^{k} x_i\, n_i = \frac{217}{70} = 3,1 \qquad a_2 = \frac{1}{n}\sum_{i=1}^{k} x_i^2 n_i = \frac{763}{70} = 10,9$$

$$a_3 = \frac{1}{n}\sum_{i=1}^{k} x_i^3 n_i = \frac{2905}{70} = 41,5 \qquad a_4 = \frac{1}{n}\sum_{i=1}^{k} x_i^4 n_i = \frac{11683}{70} = 166,9$$

Tabla con frecuencias. Variable continua.

$L_{i-1} - L_i$	x_i	n_i	$x_i n_i$	$x_i^2 n_i$	$x_i^3 n_i$	$x_i^4 n_i$
0-10	5	25	125	625	3125	15625
10-20	15	40	600	9000	135000	2025000
20-40	30	20	600	18000	540000	16200000
40-45	45	15	675	30375	1366875	61509375
total		n=100	2000	58000	2045000	79750000

$$\bar{x} = a_1 = \frac{1}{n}\sum_{i=1}^{k} x_i\, n_i = \frac{2000}{100} = 20 \qquad a_2 = \frac{1}{n}\sum_{i=1}^{k} x_i^2 n_i = \frac{58000}{100} = 580$$

$$a_3 = \frac{1}{n}\sum_{i=1}^{k} x_i^3 n_i = \frac{2045000}{100} = 20450 \qquad a_4 = \frac{1}{n}\sum_{i=1}^{k} x_i^4 n_i = \frac{79750000}{100} = 797500$$

17

Momentos centrados

Tabla sin frecuencias.

x_i	$x_i - \bar{x}$	$(x_i - \bar{x})^2$	$(x_i - \bar{x})^3$	$(x_i - \bar{x})^4$
3	0	0	0	0
3	0	0	0	0
2	-1	1	-1	1
2	-1	1	-1	1
3	0	0	0	0
5	2	4	8	16
total	0	6	6	18

$$m_1 = \frac{1}{n}\sum_{i=1}^{n} (x_i - \bar{x}) = \frac{0}{6} = 0 \qquad m_2 = \frac{1}{n}\sum_{i=1}^{n} (x_i - \bar{x})^2 = \frac{6}{6} = 1$$

$$m_3 = \frac{1}{n}\sum_{i=1}^{n} (x_i - \bar{x})^3 = \frac{6}{6} = 1 \qquad m_4 = \frac{1}{n}\sum_{i=1}^{n} (x_i - \bar{x})^4 = \frac{18}{6} = 3$$

Tabla con frecuencias. Variable continua.

$L_{i-1} - L_i$	x_i	n_i	$(x_i - \bar{x})$	$(x_i - \bar{x}) n_i$	$(x_i - \bar{x})^2 n_i$	$(x_i - \bar{x})^3 n_i$	$(x_i - \bar{x})^4 n_i$
0-10	5	25	-15	-375	5625	-84375	1265625
10-20	15	40	-5	-200	1000	-5000	25000
20-40	30	20	10	200	2000	20000	200000
40-45	45	15	25	375	9375	234375	5859375
total		n=100		0	18000	165000	7350000

$$m_1 = \frac{1}{n}\sum_{i=1}^{k} (x_i - \bar{x})\, n_i = \frac{0}{100} = 0 \qquad m_2 = \frac{1}{n}\sum_{i=1}^{k} (x_i - \bar{x})^2 n_i = \frac{18000}{100} = 180$$

$$m_3 = \frac{1}{n}\sum_{i=1}^{k} (x_i - \bar{x})^3 n_i = \frac{165000}{100} = 1650 \qquad m_4 = \frac{1}{n}\sum_{i=1}^{k} (x_i - \bar{x})^4 n_i = \frac{7350000}{100} = 73500$$

18

$L_{i-1}-L_i$	x_i	n_i	$x_i n_i$	$x_i^2 n_i$	$x_i^3 n_i$	$x_i^4 n_i$
0-10	5	25	125	625	3125	15625
10-20	15	40	600	9000	135000	2025000
20-40	30	20	600	18000	540000	16200000
40-50	45	15	675	30375	1366875	61509375
total		n=100	2000	58000	2045000	79750000

$$\bar{x} = a_1 = \frac{1}{n}\sum_{i=1}^{k} x_i n_i = \frac{2000}{100} = 20 \qquad a_2 = \frac{1}{n}\sum_{i=1}^{k} x_i^2 n_i = \frac{58000}{100} = 580$$

$$a_3 = \frac{1}{n}\sum_{i=1}^{k} x_i^3 n_i = \frac{2045000}{100} = 20450 \qquad a_4 = \frac{1}{n}\sum_{i=1}^{k} x_i^4 n_i = \frac{79750000}{100} = 797500$$

$$S^2 = m_2 = a_2 - a_1^2 = 580 - 20^2 = 580 - 400 = 180$$

$$m_3 = a_3 - 3a_2a_1 + 2a_1^3 = 20450 - \left(3\times 580\times 20\right) + \left(2\times 20^3\right) = 1650$$

$$m_4 = a_4 - 4a_3a_1 + 6a_2a_1^2 - 3a_1^4 = 797500 - \left(4\times 20450\times 20\right) + \left(6\times 580\times 20^2\right) - \left(3\times 20^4\right) = 73500$$

19

Cálculos como los anteriores, dispuestos en forma de tabla, son fáciles de hacer con la ayuda de una **hoja de cálculo**.

20

MEDIDAS DE POSICIÓN CENTRAL

Media aritmética.

$\bar{x} = \frac{1}{n}\sum_{i=1}^{k} x_i n_i = \sum_{i=1}^{k} x_i f_i$ (para tablas con frecuencias)

$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$ (para tablas sin frecuencias) $a_1 = \bar{x}$

Propiedades de la media aritmética:

$$\bar{x} = \frac{\bar{x}_1 n_1 + \ldots + \bar{x}_s n_s}{n_1 + \ldots + n_s} = \frac{1}{n}\sum_{i=1}^{s} \bar{x}_i n_i$$

$y_i = ex_i + c$ $\bar{y} = e\bar{x} + c$

21

Media geométrica.

$G = \sqrt[n]{\prod_{i=1}^{k} x_i^{n_i}} = \sqrt[n]{x_1^{n_1} \ldots x_k^{n_k}}$ (para tablas con frecuencias)

$G = \sqrt[n]{\prod_{i=1}^{n} x_i} = \sqrt[n]{x_1 \ldots x_n}$ (para tablas sin frecuencias)

22

El valor de la vivienda ha sufrido en los últimos 5 años los siguientes incrementos

incremento
6%
5%
17%
20%
14%

Obtenga el incremento medio del valor de la vivienda en estos cinco años.

$$V_1 = V_0 + \frac{6}{100} V_0 = 1,06 V_0$$

$$V_2 = V_1 + \frac{5}{100} V_1 = 1,05 V_1 = 1,05 \times 1,06 \times V_0$$

$$...$$

$$V_5 = 1,14 \times 1,20 \times 1,17 \times 1,05 \times 1,06 \times V_0$$

$V_5 = 1,7814 V_0$ $\quad \left(\frac{V_5}{V_0} = 1,7814 \right)$

$$1,14 \times 1,20 \times 1,17 \times 1,05 \times 1,06 \times V_0 = (1+r) \times (1+r) \times (1+r) \times (1+r) \times (1+r) \times V_0$$

$$(1+r)^5 = 1,14 \times 1,20 \times 1,17 \times 1,05 \times 1,06 = 1,7814$$

$$1 + r = \sqrt[5]{1,14 \times 1,20 \times 1,17 \times 1,05 \times 1,06} = \sqrt[5]{1,7814} = \sqrt[5]{\frac{V_5}{V_0}} = 1,1224$$

$r = 0,1224 \quad \Rightarrow$ en tanto por ciento $\quad r\% = 12,24\%$

$$\bar{r} = \frac{0,06 + 0,05 + 0,17 + 0,20 + 0,14}{5} = 0,124$$

en tanto por ciento $\Rightarrow$ $\bar{r} = \frac{6\% + 5\% + 17\% + 20\% + 14\%}{5} = 12,4\%$

Una vivienda que en el año 2000 se compró por 125.000€ se ha vendido en el año 2007 por 500.000€. Otra vivienda que se compró en 1995 por 100.000€ se vendió en el 2006 por 700.000€. ¿Cuál de las viviendas incrementó más su valor?

(primera vivienda)

$$\sqrt[7]{\frac{V_{2007}}{V_{2000}}} = \sqrt[7]{\frac{500.000}{125.000}} = \sqrt[7]{4} = 1,219 \quad \Rightarrow \text{ incremento anual medio del 21,9\%}$$

(segunda vivienda)

$$\sqrt[11]{\frac{V_{2006}}{V_{1995}}} = \sqrt[11]{\frac{700.000}{100.000}} = \sqrt[11]{7} = 1,1935 \quad \Rightarrow \text{ incremento anual medio del 19,35\%}$$

25

Media armónica.

$$H = \frac{n}{\sum_{i=1}^{k} \frac{n_i}{x_i}} \qquad \text{(para tablas con frecuencias)}$$

$$H = \frac{n}{\sum_{i=1}^{n} \frac{1}{x_i}} \qquad \text{(para tablas sin frecuencias)}$$

$$H = \frac{D}{\sum_{i=1}^{k} \frac{d_i}{v_i}} = \frac{D}{\frac{d_1}{v_1} + \ldots + \frac{d_k}{v_k}} = \frac{\text{distancia total}}{\text{tiempo total}} \qquad H = \frac{n}{\sum_{i=1}^{k} \frac{n_i}{x_i}} = \frac{\text{Precio total}}{\text{Litros totales}}$$

26

Si se sube en bicicleta a Sierra Nevada a una velocidad de 10km/h y se baja a una velocidad de 50km/h, ¿a qué velocidad media se ha hecho el recorrido completo de ida y vuelta?

La solución no depende de la distancia recorrida, suponemos que el ciclista recorre 45km en la subida y los mismos en la bajada.

La **media aritmética,** $\bar{x} = \frac{10+50}{2} = 30\,km/h$, no es la solución:

A esa velocidad media se tardaría en subir y bajar $3h = \frac{45+45}{30}$.

Realmente ha tardado $4,5\,h = \frac{45}{10}$ (subida) y $0,9\,h = \frac{45}{50}$ (bajada), en total 5,4 horas.

27

La **solución** es la **media armónica:**

$$H = \frac{2}{\sum_{i=1}^{2} \frac{1}{x_i}} = \frac{2}{\frac{1}{10} + \frac{1}{50}} = \frac{2}{0,10 + 0,02} = 16,6667\,km/h$$

(Observe que la distancia, 45km, no interviene en el cálculo)

A esa velocidad media tardaríamos $\frac{45+45}{16,6667} = \frac{90}{16,6667} = 5,4\,h$ en recorrer los 90km.

Utilizando las distancias recorridas, la media armónica se podría haber calculado como

$$H = \frac{n}{\sum_{i=1}^{2} \frac{n_i}{x_i}} = \frac{\sum_{i=1}^{2} n_i}{\sum_{i=1}^{2} \frac{n_i}{x_i}} = \frac{90}{\frac{45}{10} + \frac{45}{50}}$$

$$\frac{\frac{90}{45}}{\frac{\left(\frac{45}{10} + \frac{45}{50}\right)}{45}} = \frac{2}{\frac{1}{10} + \frac{1}{50}}$$

28

$$H=\frac{D}{\sum_{i=1}^{k}\frac{d_i}{v_i}}=\frac{D}{\frac{d_1}{v_1}+...+\frac{d_k}{v_k}}=\frac{\text{distancia total}}{\text{tiempo total como suma de los tiempos parciales en cada trayecto}}$$

$$velocidad\ media=\frac{\text{distancia total}}{\text{tiempo total}}=\frac{90\ km}{5,4\ h}=16,6667\ km/h$$

$$velocidad\ media=\frac{\text{distancia total}}{\text{tiempo total}}=\frac{(2\times 10)+(2\times 50)}{2+2}=\frac{120}{4}=\frac{60}{2}=\bar{x}=\frac{10+50}{2}=30\ km/h$$

$$H=\frac{2}{\frac{1}{10}+\frac{1}{50}}=16,6667<G=\sqrt{10\times 50}=22,3607<\bar{x}=\frac{10+50}{2}=30$$

29

El dueño de una motocicleta ha repostado combustible en 5 ocasiones durante el último mes. En la primera ocasión repostó 26€ a un precio de 1,3€/litro, la segunda vez repostó 30€ a un precio de 1,2€/litro, la tercera 36€ a 1,5€/litro, la cuarta también 36€ pero a 1,2€/litro y por último 35€ a 1,4€/litro. ¿A qué precio medio pagó el litro de combustible durante dicho mes?

Fácilmente calculamos el gasto total en combustible y el total de litros repostados,

$$Precio\ medio\ por\ litro\ (€/l)=\frac{total\ de\ €}{total\ de\ litros}=\frac{163}{124}=1,3145\ €/l$$

Precio/litro (€/l)	*Valor del repostaje (€)*	*Litros (l)*
x_i	n_i	$\frac{n_i}{x_i}$
1,3	26	26/1,3=20
1,2	30	30/1,2=25
1,5	36	36/1,5=24
1,2	36	36/1,2=30
1,4	35	35/1,4=25
	Total de €=n=163	*Total de litros=124*

$$H=\frac{n}{\sum_{i=1}^{5}\frac{n_i}{x_i}}=\frac{26+30+36+36+35}{\frac{26}{1,3}+\frac{30}{1,2}+\frac{36}{1,5}+\frac{36}{1,2}+\frac{35}{1,4}}=\frac{26+30+36+36+35}{20+25+24+30+25}=\frac{163}{124}=1,3145$$

30

Si hubiera repostado en todas las ocasiones la misma cantidad de litros, *L*:

$$\frac{total\ de\ €}{total\ de\ litros} = \frac{1,3L + 1,2L + 1,5L + 1,2L + 1,4L}{L + L + L + L + L} = \frac{(1,3 + 1,2 + 1,5 + 1,2 + 1,4)L}{5L} = \frac{1,3 + 1,2 + 1,5 + 1,2 + 1,4}{5} = \bar{x}$$

$$H = \frac{n}{\sum_{i=1}^{5} \frac{n_i}{x_i}} = \frac{26 + 30 + 36 + 36 + 35}{\frac{26}{1,3} + \frac{30}{1,2} + \frac{36}{1,5} + \frac{36}{1,2} + \frac{35}{1,4}} = \frac{26 + 30 + 36 + 36 + 35}{20 + 25 + 24 + 30 + 25} = \frac{163}{124} = 1,3145$$

Si hubiera repostado en todas las ocasiones por la misma cantidad de euros, *E:*

$$\frac{total\ de\ €}{total\ de\ litros} = \frac{E + E + E + E + E}{\frac{E}{1,3} + \frac{E}{1,2} + \frac{E}{1,5} + \frac{E}{1,2} + \frac{E}{1,4}} = \frac{5E}{\left(\frac{1}{1,3} + \frac{1}{1,2} + \frac{1}{1,5} + \frac{1}{1,2} + \frac{1}{1,4}\right)E} = \frac{5}{\left(\frac{1}{1,3} + \frac{1}{1,2} + \frac{1}{1,5} + \frac{1}{1,2} + \frac{1}{1,4}\right)} = H$$

31

	$\bar{x}$	G	H
1.- Nota media			
2.- Velocidad media			
3.- Precio medio vivienda (€/m^2)			
4.- Precio medio de la vivienda (€, precio total)			
5.- Incremento medio precio vivienda (%)			
6.- Salario medio (€, salario total)			
7.- Salario/hora medio (€/h)			
8.- Incremento medio salarial (%)			
9.- Precio medio del combustible (€/l)			
10.- Cambio medio del dolar (€/$)			
11.- Rentabilidad media de una cuenta corriente (%)			

32

	$\bar{x}$	G	H
1.- Nota media	X		
2.- Velocidad media	X		X
3.- Precio medio vivienda (€/m^2)	X		X
4.- Precio medio de la vivienda (€, precio total)	X		
5.- Incremento medio precio vivienda (%)		X	
6.- Salario medio (€, salario total)	X		
7.- Salario/hora medio (€/h)	X		X
8.- Incremento medio salarial (%)		X	
9.- Precio medio del combustible (€/l)	X		X
10.- Cambio medio del dolar (€/$)	X		X
11.- Rentabilidad media de una cuenta corriente (%)	X	X	

33

MODA

x_i	n_i
1	7
2	14
3	**21**
4	**21**
5	7
	$n = 70$

Mo=3

Mo=4

TIPO	n_i
Colectiva	**56**
Unifamiliar	14
total	70

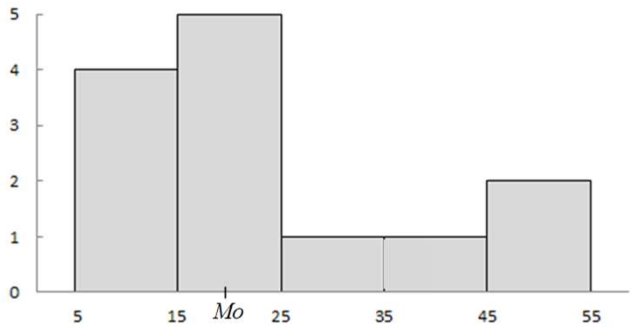
Mo=Colectiva

$L_{i-1} - L_i$	n_i	a_i	h_i
0-10	25	10	2,5
10-20	30	10	**3**
20-40	40	20	2
40-50	15	10	1,5
	$n = 100$		

Intervalo modal: 10-20.

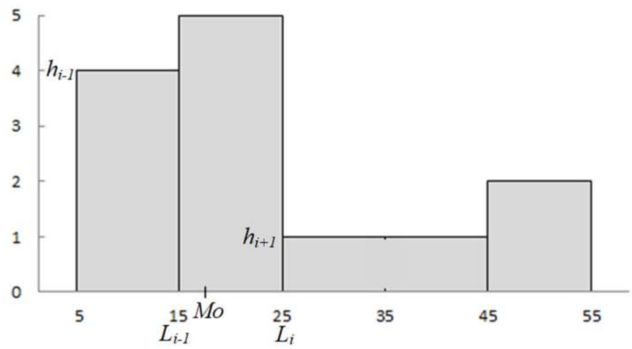
Dentro del intervalo modal hay que seleccionar un punto como moda. Hay diversos criterios,

34



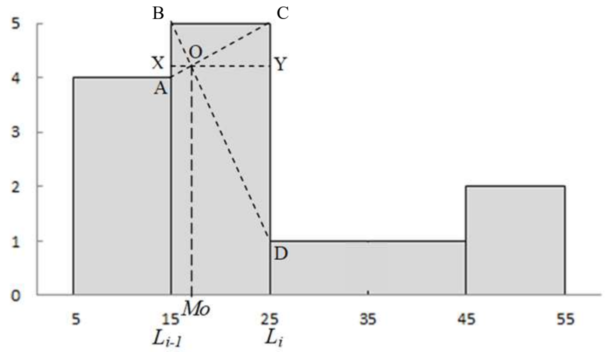
$$Mo(I) = \frac{L_{i-1} + L_i}{2}$$

35



$$\frac{Mo(II) - L_{i-1}}{h_{i+1}} = \frac{L_i - Mo(II)}{h_{i-1}} \quad \Rightarrow \quad Mo(II) = L_{i-1} + \frac{h_{i+1}}{h_{i-1} + h_{i+1}} a_i$$

36



$$\frac{\overline{AB}}{\overline{OX}}=\frac{\overline{CD}}{\overline{OY}} \Leftrightarrow \frac{h_i-h_{i-1}}{Mo(III)-L_{i-1}}=\frac{h_i-h_{i+1}}{L_i-Mo(III)} \qquad \Rightarrow \qquad Mo(III)=L_{i-1}+\frac{h_i-h_{i-1}}{(h_i-h_{i-1})+(h_i-h_{i+1})}a_i$$

37

MODA

$L_{i-1}-L_i$	n_i	a_i	$h_i=\frac{n_i}{a_i}$
5-15	40	10	4
15-25	50	10	**5**
25-45	60	20	3
45-55	20	10	2

Intervalo modal: 15-25.

Como puede verse el intervalo modal no coincide con el de mayor frecuencia ($n_3=60$), es el de mayor altura ($h_2=5$).

$$Mo(I)=\frac{L_{i-1}+L_i}{2}=\frac{15+25}{2}=20$$

$$Mo(II)=L_{i-1}+\frac{h_{i+1}}{h_{i-1}+h_{i+1}}a_i=15+\frac{3}{4+3}10=19,2857$$

$$Mo(III)=L_{i-1}+\frac{h_i-h_{i-1}}{(h_i-h_{i-1})+(h_i-h_{i+1})}a_i=15+\frac{5-4}{(5-4)+(5-3)}10=15+\frac{10}{3}=18,3333$$

38

MEDIANA

5, 10, 30, 45, 50 $\Rightarrow$ Me=30

5, 10, 30, 45 $\Rightarrow$ Me=(10+30)/2=40/2=20

39

Variables discretas: Hay dos posibilidades.

x_i	n_i	N_i
1	2	2
2	3	5
3	5	**10**
4	6	16
5	4	20

$n = 20$

$$N_i = \frac{n}{2}$$

1 1 2 2 2 3 3 3 3 3 ⋮ 4 4 4 4 4 4 5 5 5 5

10 observaciones ← ⋮ → 10 observaciones

x_i	n_i	N_i
1	2	2
2	4	6
3	7	**13**
4	5	18
5	2	20

$n = 20$

$$N_i \neq \frac{n}{2}$$

$$N_3 > \frac{n}{2}$$

1 1 2 2 2 2 ⋮ 3 3 3 3 3 3 3 ⋮ 4 4 4 4 4 5 5

6 obs. ← ⋮ 13 obs. ← ⋮

40

Variables discretas: Hay dos posibilidades.

x_i	n_i	N_i
1	2	2
2	3	5
3	5	**10**
4	6	16
5	4	20

$n = 20$

$$N_i = \frac{n}{2}$$

$$Me = \frac{x_i + x_{i+1}}{2} = \frac{3+4}{2} = 3,5\ .$$

1 1 2 2 2 3 3 3 3 3 ⋮ 4 4 4 4 4 4 5 5 5 5

10 observaciones $\leftarrow$ ⋮ $\rightarrow$ 10 observaciones

x_i	n_i	N_i
1	2	2
2	4	6
3	7	**13**
4	5	18
5	2	20

$n = 20$

$$N_i \neq \frac{n}{2}$$

$$N_3 > \frac{n}{2}$$

1 1 2 2 2 2 ⋮ 3 3 3 3 3 3 3 ⋮ 4 4 4 4 4 5 5

6 obs. $\leftarrow$ ⋮ 13 obs. $\leftarrow$ ⋮

41

Variables discretas: Hay dos posibilidades.

x_i	n_i	N_i
1	2	2
2	3	5
3	5	**10**
4	6	16
5	4	20

$n = 20$

$$N_i = \frac{n}{2}$$

$$Me = \frac{x_i + x_{i+1}}{2} = \frac{3+4}{2} = 3,5\ .$$

1 1 2 2 2 3 3 3 3 3 ⋮ 4 4 4 4 4 4 5 5 5 5

10 observaciones $\leftarrow$ ⋮ $\rightarrow$ 10 observaciones

x_i	n_i	N_i
1	2	2
2	4	6
3	7	**13**
4	5	18
5	2	20

$n = 20$

$$N_i \neq \frac{n}{2}$$

$$N_3 > \frac{n}{2} \quad \Rightarrow \quad Me = x_3 = 3\ .$$

1 1 2 2 2 2 ⋮ 3 3 3 3 3 3 3 ⋮ 4 4 4 4 4 5 5

6 obs. $\leftarrow$ ⋮ 13 obs. $\leftarrow$ ⋮

42

V**ariables continuas.** Se distinguen las dos mismas posibilidades:

$L_{i-1}-L_i$	n_i	N_i
0-10	20	20
10-**20**	30	**50**
20-40	35	85
40-50	15	100

$n=100$

$N_i = \frac{n}{2}$.

La mediana es el extremo superior de dicho intervalo: $Me = L_i = 20$.

43

$L_{i-1}-L_i$	a_i	n_i	N_i
100-110	10	25	25
110-120	**10**	**40**	**65**
120-140	20	20	85
140-150	10	15	100

$n=100$

$N_i \neq n/2$.

La mediana estará, obviamente, en el intervalo donde por primera vez $N_i > n/2$

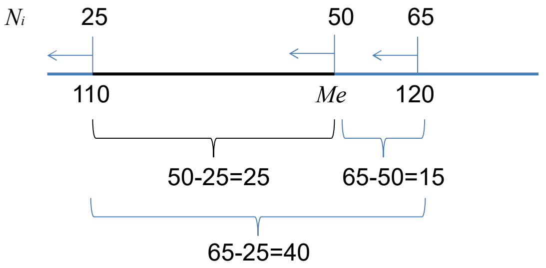
$$\frac{Me-110}{120-110} = \frac{50-25}{40} \quad \Rightarrow \quad Me = 116,25$$

$$\frac{Me-L_{i-1}}{L_i-L_{i-1}} = \frac{\frac{n}{2}-N_{i-1}}{n_i} \quad \Leftrightarrow \quad \frac{Me-L_{i-1}}{a_i} = \frac{\frac{n}{2}-N_{i-1}}{n_i} \quad \Rightarrow \quad Me = L_{i-1} + \frac{\frac{n}{2}-N_{i-1}}{n_i} a_i$$

$$Me = 110 + \frac{50-25}{40}10 = 116,25$$

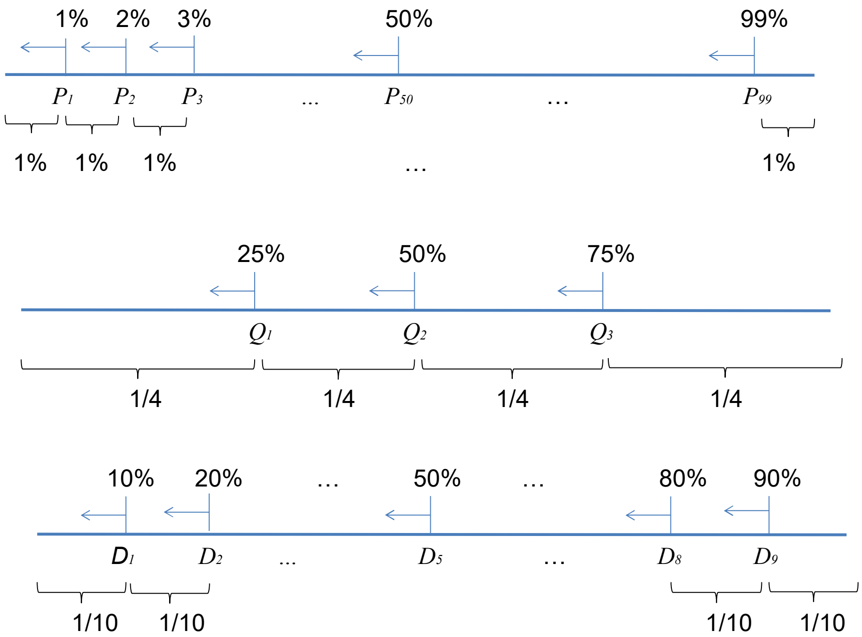
$$Me = L_{i-1} + \frac{\frac{n}{2}-N_{i-1}}{N_i - N_{i-1}} a_i = \left(dividiendo\ numerador\ y\ denominador\ por\ n\right) = L_{i-1} + \frac{\frac{1}{2}-F_{i-1}}{F_i - F_{i-1}} a_i$$

44



$$\frac{Me-110}{120-110}=\frac{50-25}{40} \quad \Rightarrow \quad Me=116,25$$

45



46

MEDIDAS DE POSICIÓN NO CENTRAL

PERCENTILES

►***EJEMPLO***

Calcule sobre las siguientes distribuciones de frecuencias los percentiles 30 y 85

x_i	n_i	N_i
1	7	7
2	14	21
3	14	35
4	28	63
5	7	70

$n = 70$

$$30\frac{n}{100} = 21 \qquad N_2 = 21 \qquad \Rightarrow \qquad P_{30} = \frac{x_2 + x_3}{2} = \frac{2+3}{2} = 2,5$$

$$85\frac{n}{100} = 59,5 \qquad N_3 < 59,5 < N_4 \qquad \Rightarrow \qquad P_{85} = x_4 = 4$$

$L_{i-1} - L_i$	a_i	n_i	N_i
100-110	10	25	25
110-120	10	40	65
120-140	20	20	85
140-150	10	15	100

$n = 100$

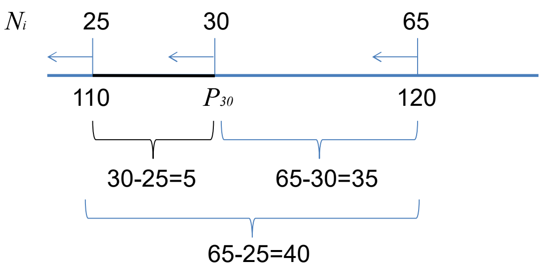
$$85\frac{n}{100} = 85 \qquad N_3 = 85 \qquad \Rightarrow \qquad P_{85} = L_3 = 140$$

$$30\frac{n}{100} = 30, \quad N_i \neq \alpha\frac{n}{100} = 30 \quad \Rightarrow$$

$$\Rightarrow \quad \frac{P_{30} - 110}{120 - 110} = \frac{30 - 25}{40} \quad \Rightarrow \quad P_{30} - 110 = \frac{50}{40} \quad \Rightarrow \quad P_{30} = 111,25$$

$$P_\alpha = L_{i-1} + \frac{\alpha\frac{n}{100} - N_{i-1}}{n_i} a_i \qquad P_{30} = L_{i-1} + \frac{30\frac{n}{100} - N_{i-1}}{n_i} a_i = 110 + \frac{30 - 25}{40} 10 = 111,25$$

47



$$\frac{P_{30} - 110}{120 - 110} = \frac{30 - 25}{65 - 25} \quad \Rightarrow \qquad P_{30} = 111,25$$

48

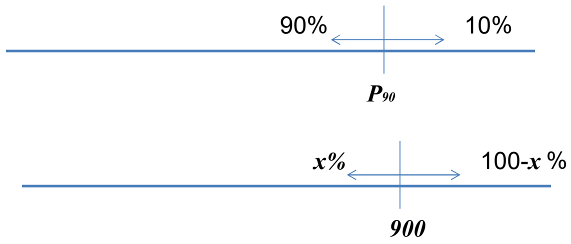
►*EJEMPLO*

Los saldos de las cuentas abiertas por los clientes de una sucursal bancaria se distribuyen de acuerdo a la siguiente tabla

SALDOS	NUMERO DE CLIENTES
0-200	100
200-1.000	400
1.000-5.000	300
5.000-30.000	50

Se consideran clientes preferentes al 10% de los clientes con mayores saldos, ¿cuál ha de ser el saldo para que un cliente sea considerado como tal?

¿Qué porcentaje de clientes tienen un saldo superior a 900€?



49

SALDOS	NUMERO DE CLIENTES	N_i
0-200	100	100
200-1.000	400	500
1.000-5.000	300	800
5.000-30.000	50	850

$$n = 850 \quad \Rightarrow \quad 90\frac{850}{100} = 765$$

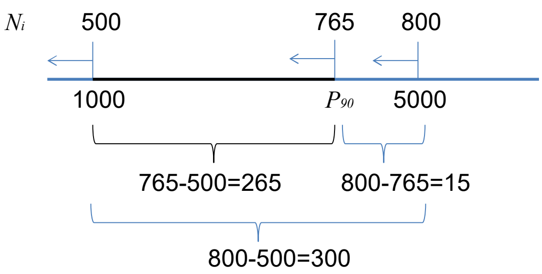
$$P_{90} = L_{i-1} + \frac{90\frac{n}{100} - N_{i-1}}{n_i} a_i = 1000 + \frac{765 - 500}{300} 4000 = 4533{,}33€$$

$$P_\alpha = L_{i-1} + \frac{\alpha\frac{n}{100} - N_{i-1}}{n_i} a_i = 900 = 200 + \frac{\alpha\frac{850}{100} - 100}{400} 800 \quad \Rightarrow \quad \alpha = 52{,}94$$

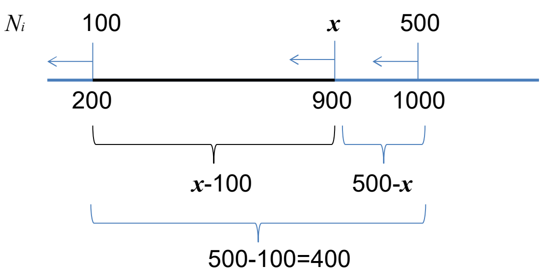
100-52,94=47,06%

50

$$\frac{90}{100}n = \frac{90}{100}850 = 765$$



$$\frac{P_{90} - 1000}{5000 - 1000} = \frac{765 - 500}{800 - 500} \quad \Rightarrow \quad P_{90} = 4533,3333$$



$$\frac{900 - 200}{1000 - 200} = \frac{x - 100}{500 - 100} \quad \Rightarrow \quad x = 450$$

$$\frac{450}{n}100 = \frac{450}{850}100 = 52,9412 \quad \Rightarrow \quad 100 - 52,4912 = 47,0588$$

MEDIDAS DE DISPERSIÓN ABSOLUTA

Recorridos

El recorrido o rango *R:* R=máximo-mínimo

El recorrido intercuartílico R_I : $R_I = Q_3 - Q_1 = P_{75} - P_{25}$

Varianza

$$S^2 = \frac{1}{n}\sum_{i=1}^{k}\left(x_i - \bar{x}\right)^2 n_i$$ (para tablas con frecuencias)

$$S^2 = \frac{1}{n}\sum_{i=1}^{n}\left(x_i - \bar{x}\right)^2$$ (para tablas sin frecuencias)

$$S^2 = a_2 - a_1^2 = \left(\frac{1}{n}\sum_{i=1}^{k} x_i^2 n_i\right) - \bar{x}^2$$ (tablas con frecuencias)

$$S^2 = a_2 - a_1^2 = \left(\frac{1}{n}\sum_{i=1}^{n} x_i^2\right) - \bar{x}^2$$ (tablas sin frecuencias)

$$y_i = ex_i + c \qquad S_y^2 = e^2 S_x^2$$

$$S^2 = m_2 \qquad m_r(y) = e^r m_r(x)$$

53

Desviación típica.

$$S = \sqrt{S^2}$$

$$S = \sigma n \qquad S_y = e\ S_x$$

►***EJEMPLO 1.16.***

Calcule la varianza y la desviación típica para las siguientes distribuciones de frecuencias de los salarios/hora y salarios mensuales, expresados en euros, de los empleados en dos almacenes.

Salario/hora (Almacén A)	n_i
5	1
15	2
25	3
35	4
	10

Salario mensual (Almacén B)	n_i
1005	1
1015	2
1025	3
1035	4
	10

Solución:

x_i	n_i	$x_i n_i$	$x_i^2 n_i$
5	1	5	25
15	2	30	450
25	3	75	1875
35	4	140	4900
	10	250	7250

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{k} x_i n_i = \frac{250}{10} = 25€$$

$$S^2 = a_2 - a_1^2 = \left(\frac{1}{n}\sum_{i=1}^{k} x_i^2 n_i\right) - \bar{x}^2 = \frac{7250}{10} - 25^2 = 100€^2$$

$$S = \sqrt{S^2} = \sqrt{100} = 10€$$

54

¿Qué significa que la desviación típica de un conjunto de datos, con media 70, es igual a 5? ¿Cómo podemos interpretar dicho valor? La desigualdad de Tchebycheff nos ayuda a **entender mejor el significado de la desviación típica** como medida de dispersión en torno a la **media.**

Desigualdad de Tchebycheff

$$p\left[x_i \in \left(\bar{x} - kS, \bar{x} + kS\right)\right] \geq 1 - \frac{1}{k^2} \qquad (k>1)$$

$$k = 2 \qquad \Rightarrow \qquad p\left[x_i \in \left(\bar{x} - 2S\ , \bar{x} + 2S\right)\right] \geq 1 - \frac{1}{k^2} = 1 - \frac{1}{2^2} = 0,75\,.$$

Otra forma:

$$kS = t \Leftrightarrow k = \frac{t}{S} \Leftrightarrow k^2 = \frac{t^2}{S^2} \qquad\qquad p\left[x_i \in \left(\bar{x} - t, \bar{x} + t\right)\right] \geq 1 - \frac{S^2}{t^2}$$

$\left(\bar{x} - 2S\ , \bar{x} + 2S\right) = \left(70 - 2\times 5\ , 70 + 2\times 5\right) = \left(60\ , 80\right)$. En el intervalo (60,80) hay al menos el 75% de las observaciones, $\left(1 - \frac{1}{k^2} = 1 - \frac{1}{2^2} = 0,75\right)$.

k=1,6 $\qquad \left(\bar{x} - 1,6S\ , \bar{x} + 1,6S\right) = \left(70 - 1,6\times 5\ , 70 + 1,6\times 5\right) = \left(62\ , 78\right) \qquad \left(1 - \frac{1}{k^2} = 1 - \frac{1}{1,6^2} = 0,61\right)$

55

Se sabe que el número medio de unidades diarias de un determinado producto que vende un supermercado es $\bar{x} = 100$ y la desviación típica $S = 40$. Si cada día el supermercado repone hasta completar 200 unidades del producto en sus estanterías:

¿Cuántos días al año la demanda será mayor que su oferta?

¿Cuánto habría que reponer para asegurar que no va a faltar producto en las estanterías el 95% de los días?

$\bar{x} = 100 \qquad S = 40$

$$\left(\bar{x} - t, \bar{x} + t\right) = \left(0, 200\right) \qquad t = 100 \qquad 1 - \frac{S^2}{t^2} = 1 - \frac{1600}{10000} = 0,84$$

0,16=1-0,84 . Es decir, en menos del 16% de los días la demanda es superior a 200 unidades.

$$1 - \frac{S^2}{t^2} = 1 - \frac{1600}{t^2} = 0,95 \ \Rightarrow\ \frac{1600}{0,05} = 32000 = t^2 \ \Rightarrow\ t = 178,9 \ \Rightarrow\ \bar{x} + t \approx 279$$

56

Variable tipificada

$$Z = \frac{X - \bar{x}}{S}$$

media cero y desviación típica 1.

Se quiere comparar los precios de dos viviendas con las mismas características, una en Madrid y otra en Granada. El precio medio de las viviendas del tipo considerado es 200.000€ en Madrid y 140.000€ en Granada, las desviaciones típicas son respectivamente 20.000€ y 15.000€. Las dos viviendas a comparar tienen unos precios de 260.000€ (Madrid) y 190.000€ (Granada). ¿cuál de las dos viviendas está alcanzando un mayor valor en su mercado?

Solución:

$$\frac{260.000 - 200.000}{20.000} = \frac{60.000}{20.000} = 3 \qquad \frac{190.000 - 140.000}{15.000} = \frac{50.000}{15.000} = 3,33$$

57

MEDIDAS DE DISPERSIÓN RELATIVA

Coeficiente de Variación.

$$CV = \frac{S}{\bar{x}}$$

Compare utilizando el coeficiente de variación la dispersión en los datos de los almacenes A y B del ejemplo 1.16.

Salario/hora (Almacén A)	n_i
5	1
15	2
25	3
35	4
	10

Salario mensual (Almacén B)	n_i
1005	1
1015	2
1025	3
1035	4
	10

$$CV_A = \frac{S_A}{\bar{x}_A} = \frac{10\,euros}{25\,euros} = 0,4 \qquad CV_B = \frac{S_B}{\bar{x}_B} = \frac{10\,euros}{1025\,euros} = 0,0098$$

58

Ejercicio resuelto 14

La distribución de los salarios mensuales de 10 trabajadores con igual cualificación profesional es

Salarios en cientos de €	n_i
0-10	1
10-20	2
20-30	3
30-40	4
	10

El horario de trabajo no es único para todos, siendo 6 el número medio de horas trabajadas cada día y 1 hora la desviación típica. ¿Es coherente la distribución de los salarios con la de las horas trabajadas?

x=horas trabajadas, y=salario

Si $y = kx$, donde *k=salario/hora,* $\Rightarrow CV_x = CV_y$

$$CV_x \neq CV_y \quad \Rightarrow \quad y \neq kx \,.$$

$$\bar{y} = \frac{1}{n}\sum_{i=1}^{k} y_i n_i = \frac{250}{10} = 25 \qquad S = \sqrt{S^2} = \sqrt{100} = 10 \qquad \Rightarrow \quad CV(salarios) = \frac{10}{25} = 0,4$$

$$CV(horarios) = \frac{1}{6} = 0,167\,.$$

59

MEDIDAS DE FORMA

Coeficiente de asimetría de Fisher.

$$g_1 = \frac{m_3}{S^3}$$

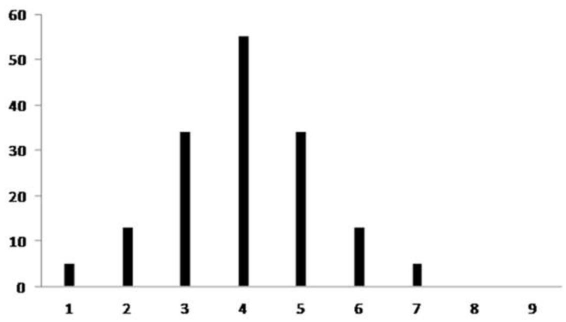
$$\sum_{i=1}^{n}\left(x_i - \bar{x}\right)^3 = 0 \qquad \sum_{i=1}^{k}\left(x_i - \bar{x}\right)^3 n_i = 0$$

- Si la distribución es simétrica $\Rightarrow \quad g_1 = 0$
- Si la distribución es asimétrica a la izquierda $\Rightarrow \quad g_1 < 0$
- Si la distribución es asimétrica a la derecha $\Rightarrow \quad g_1 > 0$

Adimensional e independiente de cambios de origen y escala.

60

x_i	n_i	$x_i n_i$	$x_i^2 n_i$	$x_i^3 n_i$	$x_i^4 n_i$	$x_i - \bar{x}$	$\left(x_i - \bar{x}\right)^3 n_i$
1	5	5	5	5	5	-3	-135
2	13	26	52	104	208	-2	-104
3	34	102	306	918	2754	-1	-34
4	55	220	880	3520	14080	0	0
5	34	170	850	4250	21250	1	34
6	13	78	468	2808	16848	2	104
7	5	35	245	1715	12005	3	135
total	159	636	2806	13320	67150		0



61

MEDIDAS DE FORMA

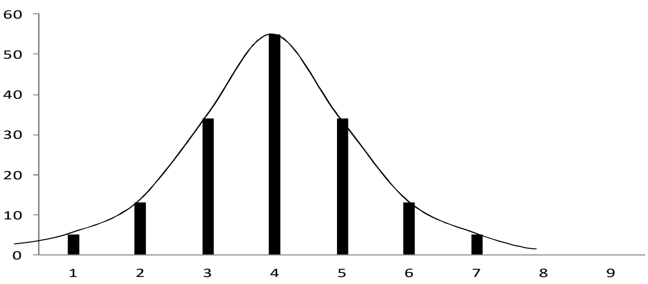
Coeficiente de apuntamiento (o curtosis) de Fisher.

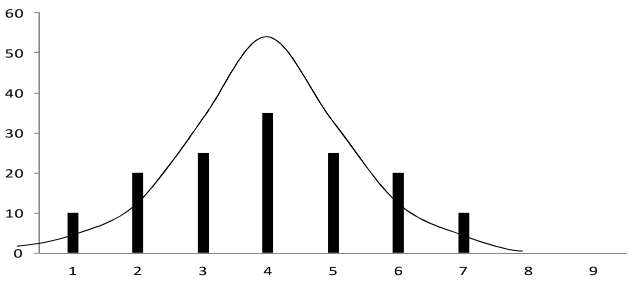
$$g_2 = \frac{m_4}{S^4} - 3$$

- Si la distribución tiene un apuntamiento *normal* (mesocúrtica) $\Rightarrow \quad g_2 = 0$
- Si la distribución es más aplanada que la *normal* (platicúrtica) $\Rightarrow \quad g_2 < 0$
- Si la distribución es más apuntada que la *normal* (leptocúrtica) $\Rightarrow \quad g_2 > 0$

Adimensional e independiente de cambios de origen y escala.

62





63

MEDIDAS DE CONCENTRACIÓN

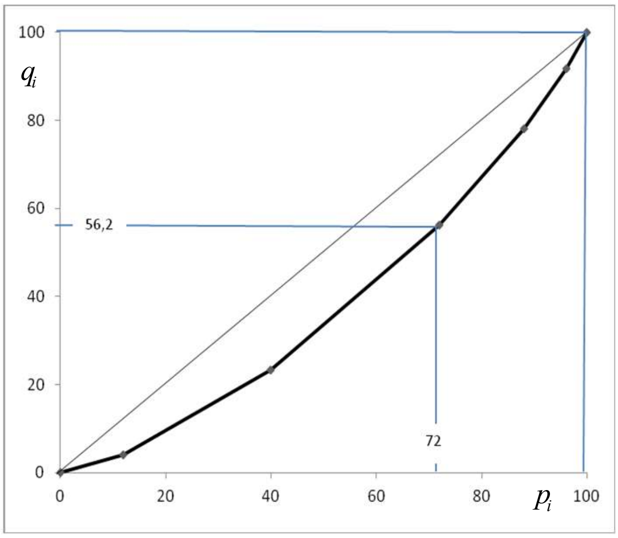
Estudiar la concentración de los salarios de 25 trabajadores recogidos en la siguiente tabla

$L_{i-1} - L_i$	n_i
500-1500	3
1500-2500	7
2500-3500	8
3500-4500	4
4500-5500	2
5500-6500	1

Hacemos los siguientes cálculos

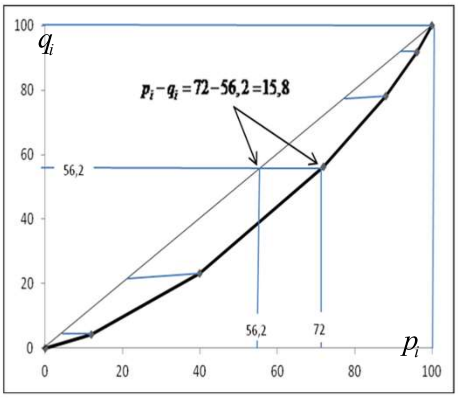
L_{i-1}	L_i	x_i	n_i	$x_i n_i$	N_i	u_i	p_i	q_i
500	1500	1000	3	3000	3	3000	12	4,1
1500	2500	2000	7	14000	10	17000	40	23,3
2500	3500	3000	8	24000	18	41000	72	56,2
3500	4500	4000	4	16000	22	57000	88	78,1
4500	5500	5000	2	10000	24	67000	96	91,8
5500	6500	6000	1	6000	25	73000	100	100,0
total			$n = 25$	$\sum_{j=1}^{k} x_j n_j = 73000$			408	353,5

64

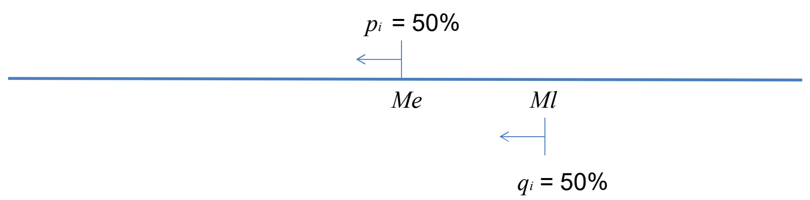


Indice de Gini.

$$I_G = \frac{\sum_{i=1}^{k-1}(p_i - q_i)}{\sum_{i=1}^{k-1} p_i} = 1 - \frac{\sum_{i=1}^{k-1} q_i}{\sum_{i=1}^{k-1} p_i} \qquad 0 \le I_G \le 1$$



$$I_G = 1 - \frac{253,5}{308} = 0,1769$$



67

Mediala.

L_{i-1}	L_i	x_i	n_i	$x_i n_i$	N_i	u_i	p_i	q_i
500	1500	1000	3	3000	3	3000	12	4,1
1500	2500	2000	7	14000	10	17000	40	23,3
2500	3500	3000	8	24000	18	41000	72	56,2
3500	4500	4000	4	16000	22	57000	88	78,1
4500	5500	5000	2	10000	24	67000	96	91,8
5500	6500	6000	1	6000	25	73000	100	100,0
total			$n=25$	$\sum_{j=1}^{k} x_j n_j = 73000$			408	353,5

$$\frac{L_i - L_{i-1}}{p_i - p_{i-1}} = \frac{Me - L_{i-1}}{50 - p_{i-1}} \Rightarrow \frac{3500-2500}{72-40} = \frac{Me-2500}{50-40} \Rightarrow \frac{1000\times 10}{32} + 2500 = Me = 2812,5$$

$$\frac{L_i - L_{i-1}}{q_i - q_{i-1}} = \frac{Ml - L_{i-1}}{50 - q_{i-1}} \Rightarrow \frac{3500-2500}{56,2-23,3} = \frac{Ml-2500}{50-23,3} \Rightarrow \frac{1000\times 26,7}{32,9} + 2500 = Ml = 3311,55$$

68

Dos familias con 4 y 5 hijos respectivamente deciden repartir parte de sus patrimonios entre ellos de la siguiente forma.

Familia A
300000
650000
200000
150000

Familia B
1500000
1000000
2000000
1000000
2000000

¿Cuál de los dos repartos es más equitativo?

69

Familia A:

x_i	n_i	$x_i n_i$	N_i	u_i	p_i	q_i
150000	1	150000	1	150000	25	11,54
200000	1	200000	2	350000	50	26,92
300000	1	300000	3	650000	75	50,00
650000	1	650000	4	1300000	100	100,00
total	4	1300000			250	188,46

$$I_G(A) = 1 - \frac{\sum_{i=1}^{k-1} q_i}{\sum_{i=1}^{k-1} p_i} = 1 - \frac{88,46}{150} = 0,410$$

$$q_3 = 50 \quad \Rightarrow \quad Ml(A) = \frac{x_3 + x_4}{2} = \frac{300000 + 650000}{2} = 475000 .$$

70

Familia B:

x_i	n_i	$x_i n_i$	N_i	u_i	p_i	q_i
1000000	2	2000000	2	2000000	40	26,7
1500000	1	1500000	3	3500000	60	46,7
2000000	2	4000000	5	7500000	100	100,0
total	5	7500000			200	173,4

$$I_G(B) = 1 - \frac{\sum_{i=1}^{k-1} q_i}{\sum_{i=1}^{k-1} p_i} = 1 - \frac{73,4}{100} = 0,266$$

$$q_3 = 100 > 50 \quad \Rightarrow \quad Ml(B) = x_3 = 2000000$$

$I_G(A) = 0,410 > I_G(B) = 0,266$