

FONEMATO 1

Para una muestra de 60 estudiantes, la siguiente tabla recoge información sobre las horas que pelan la pava diariamente:

Horas	1	2	3	4	5	6
Niños	10	12	15	8	6	9

- 1) Determinéense las medias aritmética, geométrica y armónica.
- 2) Determinéense la mediana y la moda.

SOLUCIÓN

x_i	n_i	N_i	$x_i \cdot n_i$	n_i/x_i
1	10	10	10	10
2	12	22	24	6
3	15	37	45	5
4	8	45	32	2
5	6	51	30	1.2
6	9	60	54	1.5
	60		195	25.7

1) Media aritmética: $\bar{x} = \frac{1}{N} \cdot \sum x_i \cdot n_i = \frac{195}{60}$

• Media geométrica: $G = \sqrt[n]{x_1^{n_1} \cdot x_2^{n_2} \cdot \dots \cdot x_k^{n_k}} = \sqrt[60]{1^{10} \cdot 2^{12} \cdot 3^{15} \cdot 4^8 \cdot 5^6 \cdot 6^9}$

• Media armónica: $H = \frac{N}{\sum n_i/x_i} = \frac{60}{25.7}$

2) La mediana es $Me = 3$, pues 3 es el primer valor cuya frecuencia absoluta acumulada es mayor o igual a $2 \cdot N/4 = 2 \cdot 60/4 = 30$.

• La moda es $Mo = 3$, pues 3 es valor observado que tiene mayor frecuencia.

FONEMATO 2

La siguiente tabla recoge información sobre la puntuación obtenida en un test por 60 estudiantes:

Puntuación	≤ 20	(20;30]	(30;50]	(50;70]	(70;80]	(80;100]
Alumnos	10	12	15	8	6	9

- 1) Determinénse las medias aritmética, geométrica y armónica.
- 2) Determinénse la mediana y la moda.

SOLUCIÓN

$(L_{i-1}; L_i]$	x_i	n_i	N_i	$x_i \cdot n_i$	n_i/x_i	$d_i = n_i/(L_i - L_{i-1})$
≤ 20	10	10	10	100	1	0'5
(20;30]	15	12	22	180	0'8	1'2
(30;50]	40	15	37	600	0'375	0'75
(50;70]	60	8	45	480	0'133	0'4
(70;80]	75	6	51	450	0'080	0'6
(80;100]	90	9	60	810	0'1	0'45
		60		2620	2'488	

1) Media aritmética: $\bar{x} = \frac{1}{N} \cdot \sum x_i \cdot n_i = \frac{2620}{60}$

• Media geométrica: $G = \sqrt[N]{x_1^{n_1} \cdot x_2^{n_2} \cdot \dots \cdot x_k^{n_k}} = \sqrt[60]{10^{10} \cdot 15^{12} \cdot 40^{15} \cdot 60^8 \cdot 75^6 \cdot 90^9}$

• Media armónica: $H = \frac{N}{\sum n_i/x_i} = \frac{60}{2'488}$

- 2) El intervalo mediano (el primero con frecuencia absoluta acumulada mayor o igual que $N/2 = 60$) es el (30;50], así:

$$\begin{aligned} \text{Me} &= L_{i-1} + \frac{\frac{N}{2} - N_{i-1}}{n_i} \cdot (L_i - L_{i-1}) = \\ &= 30 + \frac{30 - 22}{15} \cdot (50 - 30) = 46 \end{aligned}$$

- El intervalo de mayor densidad de frecuencia $d_i = n_i/(L_i - L_{i-1})$, es (20;30]:

$$\begin{aligned} \text{Mo} &= L_{i-1} + \frac{d_{i+1}}{d_{i-1} + d_{i+1}} \cdot (L_i - L_{i-1}) = \\ &= 20 + \frac{0'75}{0'56 + 0'75} \cdot (30 - 20) = 25'72 \end{aligned}$$

FONEMATO 3

Para una muestra de 60 niños, la siguiente tabla recoge información sobre las horas de televisión que ven diariamente:

Horas	1	2	3	4	5	6
Niños	10	12	15	8	6	9

Determine los cuartiles, los deciles cuarto y séptimo y los percentiles 62 y 85.

SOLUCIÓN

x_1	n_i	N_i
1	10	10
2	12	22
3	15	37
4	8	45
5	6	51
6	9	60
	60	

- El primer cuartil es $C_1 = 2$, pues 2 es el primer valor cuya frecuencia absoluta acumulada es mayor o igual a $1.N/4 = 60/4 = 15$.
- El segundo cuartil es $C_2 = Me = 3$, pues 3 es el primer valor cuya frecuencia absoluta acumulada es mayor o igual a $2.N/4 = 2.60/4 = 30$.
- El tercer cuartil es $C_3 = (5 + 6)/2 = 5.5$, pues como la frecuencia absoluta acumulada $3.N/4 = 3.60/4 = 45$ está en la tabla, debemos tomar la media aritmética entre la observación que corresponde a $N_i = 45$ y la siguiente.
- El cuarto decil es $D_4 = 3$, pues 3 es el primer valor cuya frecuencia absoluta acumulada es mayor o igual a $4.N/10 = 4.60/10 = 24$.
- El séptimo decil es $D_7 = 4$, pues 4 es el primer valor cuya frecuencia absoluta acumulada es mayor o igual a $7.N/10 = 7.60/10 = 42$.
- El percentil 62 es $P_{62} = 4$, pues 4 es el primer valor cuya frecuencia absoluta acumulada es mayor o igual a $62.N/100 = 62.60/100 = 37.2$.
- El percentil 85 es $P_{85} = (5 + 6)/2 = 5.5$, pues como la frecuencia absoluta acumulada $85.N/100 = 85.60/100 = 51$ está en la tabla, debemos tomar la media aritmética entre la observación que corresponde a $N_i = 51$ y la siguiente.

FONEMATO 4

La siguiente tabla recoge información sobre la puntuación obtenida en un test por 60 estudiantes:

Puntuación	≤ 15	(15;30]	(30;60]	(60;70]	(70;80]	(80;100]
Alumnos	10	12	15	8	6	9

Determine los cuartiles, el tercer decil y el percentil 62.

SOLUCIÓN

$(L_{i-1};L_i]$	n_i	N_i
≤ 15	10	10
(15;30]	12	22
(30;60]	15	37
(60;70]	8	45
(70;80]	6	51
(80;100]	9	60
	60	

- El cuartil C_1 está en el intervalo (15;30], que es el primer intervalo cuya frecuencia absoluta acumulada es mayor o igual a $1.N/4 = 60/4 = 15$:

$$C_1 = L_{i-1} + \frac{\frac{1.N}{4} - N_{i-1}}{n_i} \cdot (L_i - L_{i-1}) = 15 + \frac{15 - 10}{12} \cdot (30 - 15) = 21'25$$

- El cuartil $C_2 \equiv Me$ está en el intervalo (30;60], que es el primer intervalo cuya frecuencia absoluta acumulada es mayor o igual a $2.N/4 = 2.60/4 = 30$:

$$C_2 = L_{i-1} + \frac{\frac{2.N}{4} - N_{i-1}}{n_i} \cdot (L_i - L_{i-1}) = 30 + \frac{30 - 22}{15} \cdot (60 - 30) = 46$$

- El cuartil C_3 está en el intervalo (60;70], que es el primer intervalo cuya frecuencia absoluta acumulada es mayor o igual a $3.N/4 = 3.60/4 = 45$:

$$C_3 = L_{i-1} + \frac{\frac{3.N}{4} - N_{i-1}}{n_i} \cdot (L_i - L_{i-1}) = 60 + \frac{45 - 37}{8} \cdot (70 - 60) = 70$$

- El tercer decil D_3 está en el intervalo (60;70], que es el primer intervalo cuya frecuencia absoluta acumulada es mayor o igual a $3.N/10 = 3.60/10 = 18$:

$$D_3 = L_{i-1} + \frac{\frac{3.N}{10} - N_{i-1}}{n_i} \cdot (L_i - L_{i-1}) = 15 + \frac{18 - 10}{12} \cdot (30 - 15) = 25$$

- El percentil P_{62} está en el intervalo (60;70], que es el primer intervalo cuya frecuencia absoluta acumulada es mayor o igual a $62.N/100 = 62.60/100 = 37'2$:

$$P_{62} = L_{i-1} + \frac{\frac{62.N}{100} - N_{i-1}}{n_i} \cdot (L_i - L_{i-1}) = 60 + \frac{37'2 - 37}{15} \cdot (70 - 60) = 60'13$$

FONEMATO 5

Al tomar una muestra de 40 personas y observar el número de caries que presenta, se han registrado los siguientes datos:

Número de caries	1	2	3	4	5	6	7	8
Número de personas	2	6	10	5	10	3	2	2

Se pide:

- 1) La media aritmética, la geométrica y la armónica.
- 2) La mediana, la moda y los cuartiles.
- 3) El recorrido semiintercuartílico.
- 4) Los tres primeros momentos respecto al origen.
- 5) Los tres primeros momentos centrales.
- 6) La desviación típica y el coeficiente de variación.
- 7) El coeficiente de asimetría de Fisher.

SOLUCIÓN

x_i	n_i	N_i	$x_i \cdot n_i$	n_i/x_i	$x_i^2 \cdot n_i$	$x_i^3 \cdot n_i$
1	2	2	2	2	2	2
2	6	8	12	3	24	48
3	10	18	30	3'3	90	270
4	5	23	20	1'25	80	320
5	10	33	50	2	250	1250
6	3	36	18	0'5	108	648
7	2	38	14	0'28	98	686
8	2	40	16	0'25	128	1024
			162	12'58	780	4248

1) Media aritmética: $\bar{x} = \frac{1}{N} \cdot \sum x_i \cdot n_i = \frac{162}{40} = 4'05$

• Media geométrica: $G = \sqrt[40]{x_1^{n_1} \cdot x_2^{n_2} \cdot \dots \cdot x_k^{n_k}} = \sqrt[40]{1^2 \cdot 2^6 \cdot 3^{18} \cdot 4^5} = 3'69$

• Media armónica: $H = \frac{N}{\sum n_i/x_i} = \frac{40}{12'58} = 3'17$

2) La mediana es 4, pues 4 es el primer valor cuya frecuencia absoluta acumulada es mayor o igual a $N/2 = 40/2 = 20$.

- La distribución es **bimodal**: la mayor frecuencia observada corresponde a 3 caries y 5 caries (10 observaciones en ambos casos).
- El primer cuartil es $C_1 = 3$, pues 3 es el primer valor cuya frecuencia absoluta acumulada es mayor o igual a $N/4 = 40/4 = 10$.
- El segundo cuartil es $C_2 = Me = 4$, pues 4 es el primer valor cuya frecuencia absoluta acumulada es mayor o igual a $2 \cdot N/4 = 2 \cdot 40/4 = 20$.

- El tercer cuartil es $C_3 = 5$, pues 5 es el primer valor cuya frecuencia absoluta acumulada es mayor o igual a $3.N/4 = 3.40/4 = 30$.

3) Recorrido semiintercuartílico es $R_s = \frac{C_3 - C_1}{C_3 + C_1} = \frac{5 - 3}{5 + 3}$.

4) Es $a_1 = \bar{x} = \frac{1}{N} \cdot \sum x_i \cdot n_i = \frac{162}{40} = 4'05$.

- Es $a_2 = \frac{1}{N} \cdot \sum x_i^2 \cdot n_i = \frac{780}{40} = 19'5$.

- Es $a_3 = \frac{1}{N} \cdot \sum x_i^3 \cdot n_i = \frac{4248}{40} = 106'2$.

4) Es $m_1 = \frac{1}{N} \cdot \sum (x_i - \bar{x}) \cdot n_i = 0$.

- Es $m_2 = \frac{1}{N} \cdot \sum (x_i - \bar{x})^2 \cdot n_i = a_2 - a_1^2 = 19'5 - 4'05^2$.

- Es $m_3 = \frac{1}{N} \cdot \sum (x_i - \bar{x})^3 \cdot n_i = a_3 - 3 \cdot a_2 \cdot a_1 + 3 \cdot a_1^3 = 2'13$.

5) La varianza es $S^2 = a_2 - a_1^2 = 19'5 - 4'05^2 = 3'0975$, y la desviación típica es su raíz cuadrada; o sea: $S = \sqrt{3'0975} = 1'759$

- Coeficiente de variación es $C_V = S/\bar{x} = 1'759/4'05 = 0'434$

6) Coeficiente de asimetría de Fisher.

$$g_1 = \frac{m_3}{S^3} = \frac{2'13}{1'759^3} = 0'391 > 0$$

Como $g_1 > 0$, la distribución es asimétrica positiva o hacia la derecha y si, como en nuestro caso, la distribución no es unimodal, queda en suspenso eso de que, cuando $g_1 > 0$, sucede que $Mo < Me < \bar{x}$.

FONEMATO 6

La tabla recoge el precio (euros/kg) del cobre en las 5 subastas de una semana:

Día	1	2	3	4	5
Precio	12	15	17	20	22

- 1) Determine el precio medio si cada día se subasta igual cantidad de cobre.
- 2) Determine el precio medio si el valor de lo subastado cada día es el mismo.

SOLUCIÓN

- 1) Si "c" es la cantidad subastada diariamente, se tiene que:

$$\begin{aligned}\text{Precio Medio} &= \frac{\text{Valor total de lo subastado}}{\text{Cantidad total subastada}} = \\ &= \frac{12 \cdot c + 15 \cdot c + 17 \cdot c + 20 \cdot c + 22 \cdot c}{5 \cdot c} = 17'2 \text{ €/kg}\end{aligned}$$

- 2) Siendo x_i es la cantidad de cobre subastada el i -ésimo día ($i = 1, 2, \dots, 5$), es:

$$\begin{aligned}\text{Precio Medio} &= \frac{\text{Valor total de lo subastado}}{\text{Cantidad total subastada}} = \\ &= \frac{12 \cdot x_1 + 15 \cdot x_2 + 17 \cdot x_3 + 20 \cdot x_4 + 22 \cdot x_5}{x_1 + x_2 + x_3 + x_4 + x_5}\end{aligned}$$

Si el valor de lo subastado cada día es el mismo, sucede que:

$$12 \cdot x_1 = 15 \cdot x_2 = 17 \cdot x_3 = 20 \cdot x_4 = 22 \cdot x_5$$

Por tanto:

$$x_2 = \frac{12}{15} \cdot x_1 ; x_3 = \frac{12}{17} \cdot x_1 ; x_4 = \frac{12}{20} \cdot x_1 ; x_5 = \frac{12}{22} \cdot x_1$$

$$= \frac{(12 \cdot x_1) \cdot 5}{x_1 + \frac{12}{15} \cdot x_1 + \frac{12}{17} \cdot x_1 + \frac{12}{20} \cdot x_1 + \frac{12}{22} \cdot x_1} = 16'43 \text{ €/kg}$$

FONEMATO 7

Una empresa tiene tres factorías. La tabla recoge la producción de cada una de ellas y la productividad media por empleado.

Factoría	Producción	Productividad Media
1	4000	300
2	7000	280
3	8000	310

Determine la productividad media por empleado de la empresa.

SOLUCIÓN

$$\begin{aligned} \text{Productividad Media} &= \frac{\text{Producción total}}{\text{Número total de empleados}} = \\ &= \frac{4000 + 7000 + 8000}{\frac{4000}{300} + \frac{7000}{280} + \frac{8000}{310}} \end{aligned}$$

FONEMATO 8

Determinese la velocidad media de una hormiga que va Pinto a Valdemoro a velocidad constante de 20 metros/hora y vuelve de Valdemoro a Pinto a una velocidad constante de 30 metros/hora.

SOLUCIÓN

$$\text{Velocidad Media} = \frac{\text{Distancia total}}{\text{Tiempo total}} =$$

siendo "d" (en metros) la distancia entre Pinto y Valdemoro

$$= \frac{2 \cdot d}{\frac{d}{20} + \frac{d}{30}} = \frac{2}{\frac{1}{20} + \frac{1}{30}} = \frac{1200}{50} = 24 \frac{\text{metros}}{\text{hora}}$$

Si el viaje de ida y vuelta lo hiciera tres veces a velocidades constantes (en metros/hora) de 23, 22, 26, 28, 21 y 29:

$$\begin{aligned} \text{Velocidad Media} &= \frac{\text{Distancia total}}{\text{Tiempo total}} = \\ &= \frac{6 \cdot d}{\frac{d}{23} + \frac{d}{22} + \frac{d}{26} + \frac{d}{28} + \frac{d}{21} + \frac{d}{29}} = \frac{6}{\frac{1}{23} + \frac{1}{22} + \frac{1}{26} + \frac{1}{28} + \frac{1}{21} + \frac{1}{29}} \frac{\text{metros}}{\text{hora}} \end{aligned}$$

FONEMATO 9

Si entre Málaga y Malagón hay 15 km, determinese la velocidad media de un coche que tarda 20 minutos en el viaje de ida y 30 minutos en el de vuelta.

SOLUCIÓN

$$\text{Velocidad Media} = \frac{\text{Distancia total}}{\text{Tiempo total}} = \frac{15 + 15}{\frac{1}{3} + \frac{1}{2}} = \frac{180}{5} = 36 \frac{\text{Km}}{\text{hora}}$$

FONEMATO 10

La tabla recoge la información relativa al consumo anual de kamimocho (en millones de metros cúbicos) durante 6 años consecutivos..

Año	1	2	3	4	5	6
Consumo	60	63	65	66	68	70

- 1) Calcule la tasa de variación interanual para el periodo 1-6.
- 2) Calcule la tasa media de variación anual para el periodo 1-6.
- 3) Estime la producción del año "7".
- 4) Si el crecimiento anual fuese constante en términos absolutos, ¿cuál sería el consumo en el año 7?

SOLUCIÓN

- 1) Si x_{i-1} es el consumo en el año " $i-1$ " y x_i es el consumo en el año " i ", la tasa t_i de variación anual (en tanto por uno) es

$$t_i = \frac{x_i - x_{i-1}}{x_{i-1}} \Rightarrow x_i = (1 + t_i) \cdot x_{i-1}$$

Multiplicando t_i por 100 obtenemos la tasa de variación porcentual anual.

	Tasa interanual Tanto por uno	Tasa interanual Tanto por ciento
01 - 02	$(63 - 60)/60 = 0'050$	5 %
02 - 03	$(65 - 63)/63 = 0'031$	3'1 %
03 - 04	$(66 - 65)/65 = 0'0153$	1'53 %
04 - 05	$(68 - 66)/66 = 0'0303$	3'03 %
05 - 06	$(70 - 68)/68 = 0'0285$	2'85 %

- 2) La tasa media t_m de variación anual para el periodo 1-6 es tal que

$$x_6 = (1 + t_m)^{6-1} \cdot x_1$$

O sea:

$$70 = (1 + t_m)^5 \cdot 60 \Rightarrow t_m = 0'031$$

El que $t_m = 0'031$ significa que partiendo de un consumo 60 en el año "1", si la tasa de variación interanual fuese 0'031 durante todo los 6 años, el consumo en el año "6" también sería 70.

- 3) Es $x_7 = (1 + t_m)^{7-1} \cdot x_1 = 1'031^6 \cdot 60 = 72'06$

- 4) Siendo "d" el crecimiento anual constante en el consumo, sería:

$$x_2 = x_1 + d = 60 + d$$

$$x_3 = x_2 + d = 60 + 2 \cdot d$$

$$x_4 = x_3 + d = 60 + 3 \cdot d$$

$$x_5 = x_4 + d = 60 + 4 \cdot d$$

$$x_6 = x_5 + d = 60 + 5 \cdot d = 70 \Rightarrow d = 2$$

$$x_7 = x_6 + d = 72$$

FONEMATO 11

La tabla recoge información sobre una cadena de tiendas de ropa.

Superficie (m ²)	40 - 50	50 - 60	60 - 80	80 - 100	100 - 120
Número de parcelas	30	40	60	40	30

- 1) Mediante un cambio de origen y de escala, calcule la media y la varianza.
- 2) ¿Qué tamaño de tienda es más frecuente?
- 3) Repita 1) y 2) si la superficie de cada tienda aumenta 5 m².

SOLUCIÓN

	x_i	n_i	$d_i = \frac{n_i}{L_{i-1} - L_i}$	$z_i = \frac{x_i - 70}{5}$	$z_i \cdot n_i$	$z_i^2 \cdot n_i$
40 - 50	45	30	3	-5	-150	750
50 - 60	55	40	4	-3	-120	360
60 - 80	70	60	3	0	0	0
80 - 100	90	40	2	4	160	640
100 - 120	110	30	1'5	8	240	1920
		200			130	3670

1) Es

$$\bar{z} = \frac{1}{N} \cdot \sum z_i \cdot n_i = \frac{130}{200} = 0'65 \Rightarrow \bar{x} = 5 \cdot \bar{z} + 70 = 73'25$$

$$z_i = \frac{x_i - 70}{5} \Rightarrow x_i = 5 \cdot z_i + 70$$

$$S_Z^2 = \left(\frac{1}{N} \cdot \sum z_i^2 \cdot n_i \right) - (\bar{z})^2 = \frac{3670}{200} - 0'65^2 \Rightarrow S_X^2 = 5^2 \cdot S_Z^2$$

2) El intervalo modal, el de mayor densidad de frecuencia $d_i = n_i / (L_{i-1} - L_i)$, es el 50 - 60; así:

$$Mo = L_{i-1} + \frac{d_{i+1}}{d_{i-1} + d_{i+1}} \cdot (L_i - L_{i-1}) = 50 + \frac{3}{3+3} \cdot (60 - 50) = 55$$

3) Siendo $w_i = x_i + 5$, es:

$$\bar{w} = \bar{x} + 5 = 78'25 ; S_W^2 = S_X^2 ; Mo(W) = Mo(X) + 5$$

FONEMATO 12

La tabla recoge información sobre los salarios en una empresa.

Salarios (euros)	25 - 30	30 - 45	45 - 55	55 - 70
Número de empleados	7	19	16	8

- Determine la moda.
- Determine el salario que cobran al menos la mitad de los empleados.
- Determine la varianza.
- ¿Qué salarios definen un intervalo que agrupa el 80% central de la distribución?
- Halle la varianza de los nuevos salarios si éstos suben un 12 % y luego 2 euros.

SOLUCIÓN

	x_i	n_i	N_i	$d_i = \frac{n_i}{L_{i-1} - L_i}$	$z_i = \frac{x_i}{2'5}$	$z_i \cdot n_i$	$z_i^2 \cdot n_i$
25 - 30	27'5	7	7	1'4	11	77	847
30 - 45	37'5	19	26	1'26	15	285	4275
45 - 55	50	16	42	1'6	20	320	6400
55 - 70	62'5	8	50	0'53	25	200	5000
		50				882	16522

- 1) El intervalo de mayor densidad de frecuencia $d_i = n_i / (L_{i-1} - L_i)$, es 45 - 55:

$$Mo = L_{i-1} + \frac{d_{i+1}}{d_{i-1} + d_{i+1}} \cdot (L_i - L_{i-1}) = 65 + \frac{0'53}{1'26 + 0'53} \cdot (55 - 45) = 47'96$$

- 2) El intervalo mediano (el primero con frecuencia absoluta acumulada mayor o igual que $N/2 = 25$) es el 30 - 45; así:

$$Me = L_{i-1} + \frac{\frac{N}{2} - N_{i-1}}{n_i} \cdot (L_i - L_{i-1}) = 30 + \frac{25 - 7}{19} \cdot (45 - 30) = 44'2$$

- 3) Es: $S_Z^2 = \left(\frac{1}{N} \cdot \sum z_i^2 \cdot n_i \right) - (\bar{z})^2 = \frac{16522}{30} - \left(\frac{882}{50} \right)^2 \Rightarrow S_X^2 = 2'5^2 \cdot S_Z^2$

$$\boxed{z_i = x_i / 2'5 \Rightarrow x_i = 2'5 \cdot z_i} \uparrow$$

- 4) D_1 y D_9 definen un intervalo que agrupa el 80% central de la distribución.

- El decil D_1 está en el primer intervalo con frecuencia absoluta acumulada mayor o igual que $1.50/10 = 5$; así, $D_1 \in (55; 30]$, siendo:

$$D_1 = L_{i-1} + \frac{\frac{N}{10} - N_{i-1}}{n_i} \cdot (L_i - L_{i-1}) = 25 + \frac{5 - 0}{7} \cdot (30 - 25) = 28'357$$

$$D_9 = L_{i-1} + \frac{\frac{9 \cdot N}{10} - N_{i-1}}{n_i} \cdot (L_i - L_{i-1}) = 55 + \frac{45 - 42}{8} \cdot (70 - 55) = 60'625$$

- 5) Siendo $w_i = 1'12 \cdot x_i + 2$, es $S_W^2 = 1'12^2 \cdot S_X^2$.

FONEMATO 13

La tabla recoge información sobre los 230 trabajadores de empresa.

Edad	18 - 25	25 - 40	40 - 50	50 - 60
Mujeres	15	45	21	14
Hombres	18	52	40	25

Determine la edad media. Empleando la media y la moda de cada distribución, analice la posible asimetría de cada una.

SOLUCIÓN

Siendo "X" la variable estadística que expresa la edad de los trabajadores, es:

$$\bar{x} = \frac{1}{N} \cdot \sum x_i \cdot n_i = \frac{8752}{230} = 38'05$$

	x_i	Mujeres	Hombres	n_i	$x_i \cdot n_i$
18 - 25	21'5	15	18	33	709'5
25 - 40	32'5	45	52	97	3152'5
40 - 50	45	21	40	61	2745
50 - 60	55	14	25	39	2145
		95	135	230	8752

Las respectivas edades medias de mujeres y hombres, \bar{x}_M y \bar{x}_H son:

$$\bar{x}_M = \frac{21'5 \cdot 15 + 32'5 \cdot 45 + 45 \cdot 21 + 55 \cdot 14}{95} = 36'48$$

$$\bar{x}_H = \frac{21'5 \cdot 18 + 32'5 \cdot 52 + 45 \cdot 40 + 55 \cdot 25}{135} = 38'90$$

En cada distribución, el intervalo modal es el de mayor densidad de frecuencia $d_i = n_i / (L_{i-1} - L_i)$; o sea, el 25 - 40 en mujeres y el 40 - 50 en hombres:

	$L_{i-1} - L_i$	n_{iM}	d_{iM}	n_{iH}	d_{iH}
18 - 25	7	15	2'1	18	2'5
25 - 40	15	45	3	52	3'4
40 - 50	10	21	2'1	40	4
50 - 60	10	14	1'4	25	2'5

Así, siendo $Mo = L_{i-1} + \frac{d_{i+1}}{d_{i-1} + d_{i+1}} \cdot (L_i - L_{i-1})$, es::

$$(Mo)_M = 25 + \frac{2'1}{2'1 + 2'1} \cdot (40 - 25) = 32'5$$

$$(Mo)_H = 40 + \frac{2'5}{3'4 + 2'5} \cdot (50 - 40) = 44'23$$

Como $\left\{ \begin{array}{l} (Mo)_M = 32'5 < \bar{x}_M = 36'48 \\ (Mo)_H = 44'23 > \bar{x}_H = 38'90 \end{array} \right\}$, la distribución entre las $\left\{ \begin{array}{l} \text{mujeres} \\ \text{hom bres} \end{array} \right\}$ es asimétrica a la $\left\{ \begin{array}{l} \text{derecha} \\ \text{izquierda} \end{array} \right\}$.

FONEMATO 14

La tabla recoge información sobre los salarios en una empresa.

Salario	2'5 - 5'5	5'5 - 10'5	10'5 - 14'5	14'5 - 20'5	20'5 - 29'5
Trabajadores	21	20	32	18	9

Determinése el salario medio, la mediana, el tercer cuartil y el número de trabajadores con salario inferior a 12.

SOLUCIÓN

	x_i	n_i	N_i	$x_i \cdot n_i$
2'5 - 5'5	4	21	21	84
5'5 - 10'5	8	20	41	160
10'5 - 14'5	12'5	32	73	400
14'5 - 20'5	17'5	18	91	315
20'5 - 29'5	25	9	100	225
		100		1184

- Salario medio: $\bar{x} = \frac{1}{N} \cdot \sum x_i \cdot n_i = \frac{1184}{100} = 11'84$
- El intervalo mediano (el primero con frecuencia absoluta acumulada mayor o igual que $N/2 = 50$) es el 10'5 - 14'5; así:

$$Me = L_{i-1} + \frac{\frac{N}{2} - N_{i-1}}{n_i} \cdot (L_i - L_{i-1}) = 10'5 + \frac{50 - 41}{32} \cdot (14'5 - 10'5) = 11'625$$

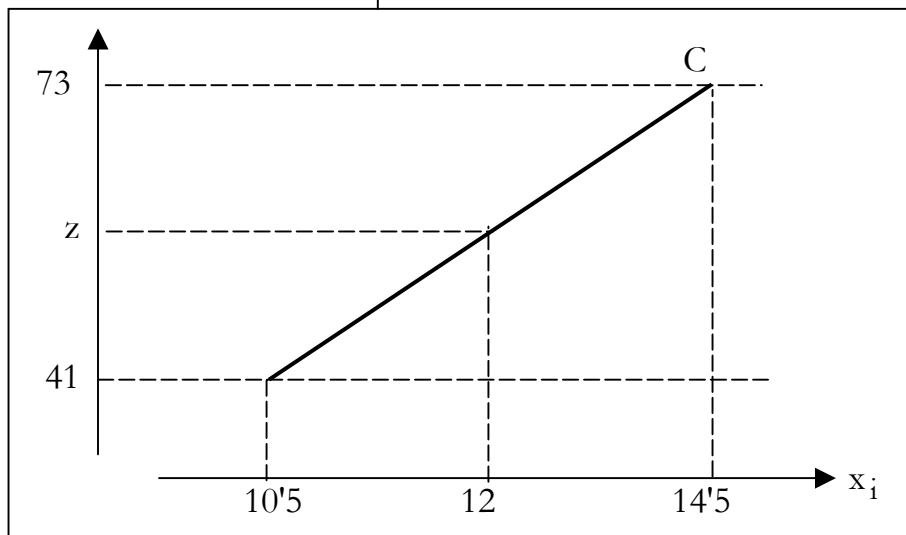
- Como $3 \cdot N/4 = 75$, es $C_3 \in (14'5; 20'5]$, siendo:

$$C_3 = L_{i-1} + \frac{\frac{3 \cdot N}{4} - N_{i-1}}{n_i} \cdot (L_i - L_{i-1}) = 14'5 + \frac{75 - 73}{18} \cdot (20'5 - 14'5) = 15'16$$

Así, el 75% de trabajadores tienen salario inferior a 15'16.

- El número "z" de trabajadores con salario inferior a 12 $\in (10'5; 14'5]$, es 53:

$$\frac{z - 41}{12'5 - 10'5} = \frac{73 - 41}{14'5 - 10'5} \Rightarrow z = 53$$



FONEMATO 15

Una empresa decide reajustar la categoría de sus empleados según los resultados de un test al que se les somete y que da los siguientes resultados:

Puntuación	0 – 30	30 – 50	50 – 70	70 – 90	90 – 100
Trabajadores	94	140	160	98	8

Si la planificación óptima determina que el 65% sean administrativos, el 20% jefes de sección, el 10% jefes de departamento y el 5% inspectores, determinense las puntuaciones de corte.

SOLUCIÓN

	n_i	N_i
0 – 30	94	94
30 – 50	140	234
50 – 70	160	394
70 – 90	98	492
90 – 100	8	500
	500	100

Debemos determinar los percentiles P_{65} , P_{65+20} y $P_{65+20+10}$ de la distribución de puntuaciones.

- El percentil P_{65} está en el primer intervalo con frecuencia absoluta acumulada mayor o igual que $65.N/100 = 65.500/100 = 350$; así, $P_{65} \in (50; 70]$, siendo:

$$P_{65} = L_{i-1} + \frac{350 - N_{i-1}}{n_i} \cdot (L_i - L_{i-1}) = 50 + \frac{350 - 234}{160} \cdot (70 - 50) = 61'375$$

- El percentil P_{85} está en el primer intervalo con frecuencia absoluta acumulada mayor o igual que $85.N/100 = 85.500/100 = 425$; así, $P_{85} \in (70; 90]$, siendo:

$$P_{85} = L_{i-1} + \frac{425 - N_{i-1}}{n_i} \cdot (L_i - L_{i-1}) = 70 + \frac{425 - 394}{98} \cdot (90 - 70) = 76'326$$

- El percentil P_{95} está en el primer intervalo con frecuencia absoluta acumulada mayor o igual que $95.N/100 = 95.500/100 = 475$; así, $P_{95} \in (70; 90]$, siendo:

$$P_{95} = L_{i-1} + \frac{475 - N_{i-1}}{n_i} \cdot (L_i - L_{i-1}) = 70 + \frac{475 - 394}{98} \cdot (90 - 70) = 86'53$$

FONEMATO 16

Encuesta a 1000 bares sobre el número de horas de apertura semanal:

Horas	Bares
(30;35]	50
(35;37]	100
(37;39]	200
(39;40]	150
(40;41]	120
(41;43]	n_6
(43;45]	130
(45;50]	n_8

- Determine n_6 y n_8 si la media de la distribución es 40'38
- Determine la moda y la mediana.
- ¿Cuántas horas abren, como mínimo, los 230 bares que más horas abren?

SOLUCIÓN

Intervalos	n_i	x_i	$x_i \cdot n_i$	N_i	d_i	$N_i \downarrow$
(30;35]	50	32'5	1625	50	10	
(35;37]	100	36	3600	150	50	
(37;39]	200	38	7600	350	100	
(39;40]	150	39'5	5925	500	150	
(40;41]	120	40'5	4860	620	120	
(41;43]	n_6	42	$42 \cdot n_6$	770	75	
(43;45]	130	44	5720	900	65	230
(45;50]	n_8	47'5	$47'5 \cdot n_8$	1000	20	100

$$\sum x_i \cdot n_i = 29330 + 42 \cdot n_6 + 47'5 \cdot n_8$$

$$\left\{ \begin{array}{l} \frac{29330 + 42 \cdot n_6 + 47'5 \cdot n_8}{1000} = 40'38 \\ 750 + n_6 + n_8 = 1000 \end{array} \right\} \Rightarrow \left\{ \begin{array}{l} n_6 = 150 \\ n_8 = 100 \end{array} \right.$$

- El intervalo de mayor densidad de frecuencia $d_i = n_i / (L_{i-1} - L_i)$, es (39;40]:

$$Mo = L_{i-1} + \frac{d_{i+1}}{d_{i-1} + d_{i+1}} \cdot (L_i - L_{i-1}) = 39 + \frac{120}{100 + 120} \cdot (40 - 39) = 39'54$$

- El intervalo mediano (el primero con frecuencia absoluta acumulada mayor o igual que $N/2 = 500$) es el (39;40]; así:

$$Me = L_{i-1} + \frac{\frac{N}{2} - N_{i-1}}{n_i} \cdot (L_i - L_{i-1}) = 39 + \frac{500 - 350}{150} \cdot (40 - 39) = 40$$

- La columna $N_i \downarrow$ de la frecuencia absoluta acumulada descendente indica que los 230 bares que más horas abren, como mínimo abren 43 horas.

FONEMATO 17

Si $a_1 = 6, a_2 = 12, a_3 = 12, a_4 = 270, a_5 = 4, a_6 = 567$ son los 6 primeros momentos respecto al origen de la distribución de frecuencias $(x_i; n_i)$, estúdiense la simetría de la distribución $(z_i; n_i)$ si $z_i = x_i^2/3$.

SOLUCIÓN

La distribución de frecuencias $(z_i; n_i)$ es asimétrica negativa, pues su **coeficiente de asimetría de Fisher** es negativo:

$$g_1(Z) = \frac{m_3(Z)}{S_Z^3} = -\frac{147}{15^{3/2}} < 0$$

Sea $(x_i; n_i)$ la distribución de frecuencias correspondiente a una muestra de tamaño "N" de una variable estadística "X"; así, siendo constantes "C" y "r", si "Z" es la variable estadística tal que $Z = C \cdot X^r$, para la distribución de frecuencias $(z_i; n_i)$ sucede que:

$$\begin{aligned} a_s(Z) &= \frac{1}{N} \cdot \sum z_i^s \cdot n_i = \frac{1}{N} \cdot \sum (C \cdot x_i^r)^s \cdot n_i = \\ &= C^s \cdot \underbrace{\left(\frac{1}{N} \cdot \sum (x_i^{r \cdot s}) \cdot n_i \right)}_{a_{r \cdot s}(X)} = C^s \cdot a_{r \cdot s}(X) \end{aligned}$$

Si $Z = \frac{1}{3} \cdot X^2$ (o sea, $C = \frac{1}{3}$ y $r = 2$), es:

$$a_1(Z) = \frac{1}{3} \cdot a_2(X) = \frac{1}{3} \cdot 12 = 4$$

$$a_2(Z) = \left(\frac{1}{3}\right)^2 \cdot a_4(X) = \left(\frac{1}{3}\right)^2 \cdot 270 = 30$$

$$a_3(Z) = \left(\frac{1}{3}\right)^3 \cdot a_6(X) = \left(\frac{1}{3}\right)^3 \cdot 567 = 21$$

Por tanto:

$$S_Z = +\sqrt{a_2(Z) - (a_1(Z))^2} = \sqrt{30 - 4^2} = \sqrt{15}$$

$$m_3(Z) = a_3(Z) - 3 \cdot a_2(Z) \cdot a_1(Z) + 3 \cdot a_1^3(Z) = 21 - 3 \cdot 30 \cdot 4 + 3 \cdot 4^3 = -147$$

FONEMATO 18

Se toman dos muestras de una población y se obtienen los siguientes resultados relativos al consumo de cerveza:

	Tamaño	Media	Varianza
Muestra 1	1000	10	80
Muestra 2	1500	12	120

Determinése la media y la varianza para el conjunto de todas las observaciones.

SOLUCIÓN

Si la distribución

x_1	x_2	x_h	x_{h+1}	x_{h+2}	x_k
n_1	n_2	n_h	n_{h+1}	n_{h+1}	n_k

se particiona en las siguientes dos distribuciones

x_1	x_2	x_h
n_1	n_2	n_h

,

x_{h+1}	x_{h+2}	x_k
n_{h+1}	n_{h+2}	n_k

siendo $N_1 = \sum_{i=1}^h n_i$ y $N_2 = \sum_{i=h+1}^k n_i$, sabemos que si $(\bar{x})_r$ es la media aritmética del r-ésimo estrato ($i=1,2$), la media aritmética \bar{x} de la distribución dada es $\bar{x} = \frac{N_1 \cdot (\bar{x})_1 + N_2 \cdot (\bar{x})_2}{N}$ y si $(a_2)_r$ es el momento de orden 2 respecto al origen del r-ésimo estrato ($i=1,2$), el momento de orden 2 respecto al origen de la distribución dada es $a_2 = \frac{N_1 \cdot (a_2)_1 + N_2 \cdot (a_2)_2}{N}$. Así, la varianza de la distribución dada, es:

$$S^2 = a_2 - (\bar{x})^2 = \frac{N_1 \cdot (a_2)_1 + N_2 \cdot (a_2)_2}{N} - \left(\frac{N_1 \cdot (\bar{x})_1 + N_2 \cdot (\bar{x})_2}{N} \right)^2$$

En nuestro caso:

$$\bar{x} = \frac{N_1 \cdot (\bar{x})_1 + N_2 \cdot (\bar{x})_2}{N} = \frac{1000 \cdot 10 + 1500 \cdot 12}{1000 + 1500} = 11'2$$

$$a_2 = \frac{N_1 \cdot (a_2)_1 + N_2 \cdot (a_2)_2}{N} = \frac{1000 \cdot 180 + 1500 \cdot 264}{1000 + 1500} = 230'4$$

$(a_2)_1 = S_1^2 + (\bar{x})_1^2 = 80 + 10^2 = 180$
$(a_2)_2 = S_2^2 + (\bar{x})_2^2 = 120 + 12^2 = 264$

Por tanto:

$$S^2 = a_2 - (\bar{x})^2 = 230'4 - 11'2^2 = 104'96$$

FONEMATO 19

De una mancomunidad de cuatro pueblos se conocen los siguientes datos:

Pueblo	Habitantes	Renta per cápita	Varianza de la renta per cápita
A	50	150	56200
B	40	140	36000
C	12	135	64000
D	11	120	25000

Determinése la renta per cápita de la mancomunidad y su varianza.

Determinése el coeficiente de variación de la renta per cápita.

SOLUCIÓN

La renta per cápita de la mancomunidad

$$\bar{x} = \frac{150 \cdot 50 + 140 \cdot 40 + 135 \cdot 12 + 120 \cdot 11}{50 + 40 + 12 + 11}$$

Siendo

$$(a_2)_1 = S_1^2 + (\bar{x})_1^2 = 56200 + 150^2$$

$$(a_2)_2 = S_2^2 + (\bar{x})_2^2 = 36000 + 140^2$$

$$(a_2)_3 = S_3^2 + (\bar{x})_3^2 = 64000 + 135^2$$

$$(a_2)_4 = S_4^2 + (\bar{x})_4^2 = 25000 + 120^2$$

es:

$$a_2 = \frac{50 \cdot (a_2)_1 + 40 \cdot (a_2)_2 + 12 \cdot (a_2)_3 + 11 \cdot (a_2)_4}{50 + 40 + 12 + 11}$$

por tanto:

$$S^2 = a_2 - (\bar{x})^2$$

El coeficiente de variación de la renta per cápita es $C_V = \frac{S}{\bar{x}}$.

FONEMATO 20

En el año actual, de una empresa con tres categorías de empleados, se sabe que:

Categoría	Empleados	Salario medio	Desviación típica
A	130	145	22'5
B	50	200	42
C	20	300	70

Halle el salario medio en la empresa, dando un indicador de su representatividad.

Para fijar los salarios del próximo año se proponen tres alternativas:

- 1) Elevación de todos los salarios un 5 %.
 - 2) Incremento lineal de 10 u.m.
 - 3) Aumento de un 10 % en la categoría "A", un 8 % en la "B" y un 4 % en la "C"
- Determinése el salario medio en la empresa con cada alternativa, analizando cuál de ellas reduce más la dispersión actual.

SOLUCIÓN

$$\text{Salario medio en la empresa: } \bar{x} = \frac{130 \cdot 145 + 50 \cdot 200 + 20 \cdot 300}{130 + 50 + 20} = 174'25$$

$$\text{El coeficiente de variación es } C_V = S/\bar{x} = \sqrt{3563'2}/174'25 = 0'34.$$

$$S^2 = a_2 - (\bar{x})^2 = 33926'3 - 174'25^2 = 3563'2$$
$$a_2 = \frac{130 \cdot (a_2)_1 + 50 \cdot (a_2)_2 + 20 \cdot (a_2)_3}{130 + 50 + 20} = 33926'3$$
$$(a_2)_1 = 22'5^2 + 145^2 ; (a_2)_2 = 42^2 + 200^2 ; (a_2)_3 = 70^2 + 300^2$$

- 1) Si los salarios se multiplican por 1'05, el salario medio en la empresa también se multiplica por 1'05; así el nuevo salario medio es $174'25 \cdot 1'05 = 182'96 \dots$ y el coeficiente de variación no cambia, pues es insensible a los cambios de escala.
- 2) Si los salarios suben 10 u.m, el salario medio en la empresa sube 10 u.m, pasando a ser $174'25 + 10 = 184'25 \dots$ y como la desviación típica es insensible al cambio de origen, es $C_V = \sqrt{3563'2}/184'25 = 0'324$.

- 3) Salario medio en la empresa:

$$\bar{x} = \frac{130 \cdot 145 \cdot 1'1 + 50 \cdot 200 \cdot 1'08 + 20 \cdot 300 \cdot 1'04}{130 + 50 + 20} = 188'87$$

$$\text{El coeficiente de variación es } C_V = S/\bar{x} = \sqrt{3703'3}/188'87 = 0'322.$$

$$S^2 = a_2 - (\bar{x})^2 = 3526'8 - 188'87^2 = 3703'3$$
$$a_2 = \frac{130 \cdot (a_2)_1 + 50 \cdot (a_2)_2 + 20 \cdot (a_2)_3}{130 + 50 + 20} = 3526'8$$
$$(a_2)_1 = (22'5^2 + 145^2) \cdot 1'1^2 ; (a_2)_2 = (42^2 + 200^2) \cdot 1'08^2$$
$$(a_2)_3 = (70^2 + 300^2) \cdot 1'04^2$$

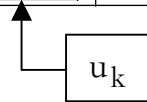
FONEMATO 21

Determine el índice de Gini en la siguiente distribución de salarios.

Salario	Empleados
355	10
457	17
600	12
750	11

SOLUCIÓN

x_i	n_i	$x_i \cdot n_i$	N_i	u_i	$p_i = \frac{N_i}{N} \cdot 100$	$q_i = \frac{u_i}{u_k} \cdot 100$	$p_i - q_i$
355	10	3550	10	3550	20	13'26	6'74
457	17	7769	27	11319	54	42'28	11'72
600	12	7200	39	18519	78	69'18	8'81
750	11	8250	50	26769	×	×	×
		26769			152		27'27



$$I_G = \frac{\sum_{i=1}^{k-1} (p_i - q_i)}{\sum_{i=1}^{k-1} p_i} = \frac{27'27}{152} = 0'179$$

FONEMATO 22

En 20 empresas europeas se obtienen los siguientes resultados:

Beneficio (Euros)	Número de Empresas
10	4
11	6
12	6
13	2
14	2

En 20 empresas americanas, resulta un beneficio medio de 50000 \$ con desviación típica de 6000 \$.

Determinar la concentración del beneficio de las empresas europeas mediante el índice de Gini. ¿A qué lado del Atlántico hay mayor dispersión del beneficio?

SOLUCIÓN

x_i	n_i	$x_i \cdot n_i$	N_i	u_i	$p_i = \frac{N_i}{N} \cdot 100$	$q_i = \frac{u_i}{u_k} \cdot 100$	$p_i - q_i$	$x_i^2 \cdot n_i$
10	4	40	4	40	20	17'24	2'76	400
11	6	66	10	106	50	45'68	4'32	726
12	6	66	16	178	80	76'72	3'28	864
13	2	26	18	204	90	87'93	2'07	338
14	2	28	20	232	×	×	×	392
		232			240		12'43	2720

$$I_G = \frac{\sum_{i=1}^{k-1} (p_i - q_i)}{\sum_{i=1}^{k-1} p_i} = \frac{12'43}{240} = 0'051$$

- Como I_G es muy próximo a cero, la equidistribución es muy elevada.
- Como el beneficio se mide en unidades distintas a uno y otro lado del charco, estudiamos la dispersión mediante el coeficiente de variación.

En Europa:

$$C_V = \frac{S}{\bar{x}} = \frac{\sqrt{1'44}}{11'6} = 0'1034$$

$$\bar{x} = \frac{\sum x_i \cdot n_i}{20} = \frac{232}{20} = 11'6$$

$$S^2 = a_2 - \bar{x}^2 = \frac{\sum x_i^2 \cdot n_i}{20} - \bar{x}^2 = \frac{2720}{20} - 11'6^2 = 1'44$$

En América: $C_V = \frac{S}{\bar{x}} = \frac{6000}{50000} = 0'12$

Por tanto, en Europa hay menor dispersión.

FONEMATO 23

La tabla recoge información sobre los salarios en una empresa.

Salario	≤ 6	(6;10]	(10;15]	(15;20]	(20;30]	(30;60]	> 60
Frecuencia relativa	0'054	0'157	0'323	0'222	0'150	0'075	0'019

Determine el decil 7, el percentil 39, y la mediana. Calcule la varianza. Determine el índice de Gini.

SOLUCIÓN

Para trabajar, suponemos que el extremo superior del último intervalo es 90.

	f_i	F_i	x_i	$x_i \cdot f_i$	$x_i^2 \cdot f_i$
(0;6]	0'054	0'054	3	0'162	0'486
(6;10]	0'157	0'211	8	1'256	10'048
(10;15]	0'323	0'534	12'5	4'037	50'468
(15;20]	0'222	0'756	17'5	3'885	67'987
(20;30]	0'150	0'906	25	3'750	93'750
(30;60]	0'075	0'981	45	3'375	151'875
(60;90]	0'019	1	80	1'520	93'100
				17'985	496'214

- El decil D_7 está en el primer intervalo con frecuencia relativa acumulada mayor o igual que 0'7; así, $D_7 \in (15;20]$, siendo:

$$D_7 = L_{i-1} + \frac{0'7 - F_{i-1}}{f_i} \cdot (L_i - L_{i-1}) = 20 + \frac{0'7 - 0'534}{0'222} \cdot (20 - 15) = 23'73$$

Así, el 70% de trabajadores tienen salario inferior a 23'73.

- El percentil P_{39} está en el primer intervalo con frecuencia relativa acumulada mayor o igual que 0'39; así, $P_{39} \in (10;15]$, siendo:

$$P_{39} = L_{i-1} + \frac{0'39 - F_{i-1}}{f_i} \cdot (L_i - L_{i-1}) = 10 + \frac{0'39 - 0'211}{0'323} \cdot (15 - 10) = 12'77$$

Así, el 39% de trabajadores tienen salario inferior a 12'77.

- La mediana $Me \equiv C_2 \equiv D_5 \equiv P_{50}$ está en el primer intervalo con frecuencia relativa acumulada mayor o igual que 0'5; así, $Me \in (10;15]$, siendo:

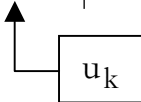
$$Me = L_{i-1} + \frac{0'5 - F_{i-1}}{f_i} \cdot (L_i - L_{i-1}) = 10 + \frac{0'5 - 0'211}{0'323} \cdot (15 - 10) = 14'47$$

- Es:

$$\begin{aligned} S^2 &= a_2 - (\bar{x})^2 = \\ &= \left(\frac{1}{N} \cdot \sum x_i^2 \cdot n_i \right) - \left(\frac{1}{N} \cdot \sum x_i \cdot n_i \right)^2 = \left(\sum x_i^2 \cdot f_i \right) - \left(\sum x_i \cdot f_i \right)^2 = \\ &= 496'214 - 17'985^2 \end{aligned}$$

- Índice de Gini

$x_i \cdot f_i$	$u_i = x_1 \cdot f_1 + \dots + x_i \cdot f_i$	$p_i = F_i \cdot 100$	$q_i = \frac{u_i}{u_k} \cdot 100$	$p_i - q_i$
0'162	0'162	5'4	0'9	4'5
1'256	1'418	21'1	7'88	13'22
4'037	5'455	53'4	30'33	23'07
3'885	9'340	75'6	51'93	23'67
3'750	13'09	90'6	72'78	17'82
3'375	16'465	98'1	91'54	6'56
1'520	17'985	×	×	×
17'985		334'2		88'84



$$I_G = \frac{\sum_{i=1}^{k-1} (p_i - q_i)}{\sum_{i=1}^{k-1} p_i} = \frac{88'84}{334'2} = 0'258$$

Si no te gusta trabajar con frecuencias relativas, basta llevar el agua al molino de las frecuencias absolutas y para ello basta considerar que la tabla es así:

Salario	≤ 6	(6;10]	(10;15]	(15;20]	(20;30]	(30;60]	> 60
Número de empleados	54	157	323	222	150	75	19

FONEMATO 24

El ejército de Pancho Villa lo forman 25 cabos y 30 generales, y la distribución de salarios entre los cabos es la siguiente

Salario	55	64	70	78	84	88
n_i	4	2	3	5	6	5

La distribución de salarios entre los generales es la siguiente

Salario	100 - 120	120 - 130	130 - 150	150 - 170	170 - 180
n_i	3	6	7	10	4

- 1) Determine el salario medio del ejército. ¿En cuál de los dos colectivos hay menor dispersión relativa de los salarios?
- 2) Si el salario de los soldados disminuye primero un 10% y después un aumento lineal de 12 u.m, y el salario de los generales aumenta el 12%, ¿en cuál de los dos colectivos hay ahora menor dispersión relativa de los salarios?
- 3) Estudie la simetría de la distribución inicial de salarios de los soldados. ¿Qué influencia tiene la revisión de salarios en la simetría?
- 4) Determine el salario más frecuente en los generales, y el salario mínimo del 30% de los generales de mayores ingresos.

SOLUCIÓN

x_i	n_i	$x_i \cdot n_i$	$x_i^2 \cdot n_i$	$x_i^3 \cdot n_i$
55	4	220	12100	665500
64	2	128	8192	524288
70	3	210	14700	1029000
78	5	390	30420	2772760
84	6	504	42336	3556224
88	5	440	38720	3407360
	25	1892	146468	11555132

$$\Rightarrow \left\{ \begin{array}{l} \bar{x}_C = \frac{1892}{25} = 75'68 \\ S_C^2 = \frac{146468}{25} - 75'68^2 = 131'25 \\ C_{VC} = \frac{S_C}{\bar{x}_C} = \frac{\sqrt{131'25}}{75'68} = 0'15 \end{array} \right.$$

$(L_{i-1}; L_i]$	z_i	n_i	$z_i \cdot n_i$	$z_i^2 \cdot n_i$
100 - 120	110	3	330	36300
120 - 130	125	6	750	93750
130 - 150	140	7	980	137200
150 - 170	160	10	1600	256000
170 - 180	175	4	700	122500
		30	4360	645750

$$\Rightarrow \left\{ \begin{array}{l} \bar{x}_G = \frac{4360}{30} = 145'33 \\ S_G^2 = \frac{645750}{30} - 145'33^2 = 403'3 \\ C_{VG} = \frac{S_G}{\bar{x}_G} = \frac{\sqrt{403'3}}{145'33} = 0'138 \end{array} \right.$$

- 1) El salario medio del ejército es:

$$\bar{x} = \frac{25 \cdot \bar{x}_C + 30 \cdot \bar{x}_G}{25 + 30} = \frac{1892 + 4360}{25 + 30} = 113'67$$

Si las medias de ambas poblaciones fueran parecidas, para comparar la dispersión bastaría comparar las respectivas varianzas; pero como dichas medias son muy diferentes, compararemos la dispersión mediante los respectivos coeficientes de variación ... y como $C_{VC} = 0'15$ y $C_{VG} = 0'138$, la dispersión es menor entre los generales.

2) Si $x_i^* = 0'9 \cdot x_i + 12$, es $C_{VC}^* = \frac{S_C^*}{\bar{x}_C^*} = \frac{0'9 \cdot S_C}{0'9 \cdot \bar{x}_C + 12} = 0'128$

Si $z_i^* = 1'12 \cdot z_i$, es $C_{VG}^* = \frac{S_G^*}{\bar{z}_G^*} = \frac{1'12 \cdot S_C}{1'12 \cdot \bar{z}_C} = \frac{S_C}{\bar{z}_C} = 0'138$

Ahora la dispersión es menor entre los soldados.

- 3) El coeficiente de asimetría de Fisher de la distribución inicial de salarios de los soldados:

$$g_1 = \frac{m_3}{S^3} = \frac{-1049'82}{(\sqrt{131'25})^3} = -0'69$$

$m_3 = a_3 - 3 \cdot a_2 \cdot a_1 + 2 \cdot a_1^3 = -1049'82$
$a_1 = \frac{1}{N} \cdot \sum x_i \cdot n_i = \frac{1892}{25} = 75'68 ; a_2 = \frac{1}{N} \cdot \sum x_i^2 \cdot n_i = \frac{146468}{25}$
$a_3 = \frac{1}{N} \cdot \sum x_i^3 \cdot n_i = \frac{11555132}{25}$

Como todo el mundo sabe, coeficiente de asimetría de Fisher g_1 es insensible a cambios de origen y de escala; por tanto, no cambia tras la revisión del salario de los soldados.

- 4) El intervalo de mayor densidad de frecuencia $d_i = n_i / (L_{i-1} - L_i)$, es el 2º:

$$\begin{aligned} Mo &= L_{i-1} + \frac{d_{i+1}}{d_{i-1} + d_{i+1}} \cdot (L_i - L_{i-1}) = \\ &= 120 + \frac{0'35}{0'15 + 0'35} \cdot (130 - 120) = 127 \end{aligned}$$

$(L_{i-1}; L_i]$	n_i	$d_i = n_i / (L_{i-1} - L_i)$	N_i
100 - 120	3	0'15	3
120 - 130	6	0'6	9
130 - 150	7	0'35	16
150 - 170	10	0'5	26
170 - 180	4	0'4	30
	30	30	

El salario mínimo del 30% de los generales de mayores ingresos es el séptimo decil D_7 , que está en el intervalo 150 - 170, por ser el primero con frecuencia absoluta acumulada mayor o igual a $7 \cdot N / 10 = 7 \cdot 30 / 10 = 21$