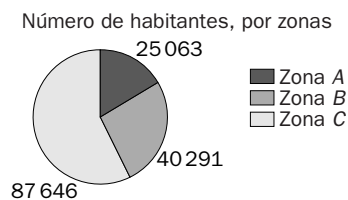


# 16 Estadística descriptiva

1. Cierta ayuntamiento va a realizar una encuesta para saber la opinión de los ciudadanos sobre las medidas tomadas para mejorar la circulación urbana. El ayuntamiento pretende que la muestra contenga información de distintas zonas de la ciudad. El gráfico adjunto proporciona la distribución de habitantes por zonas.
- Si el tamaño de la muestra es de 200 personas, determina el tamaño muestral de cada zona, redondeando si es preciso.

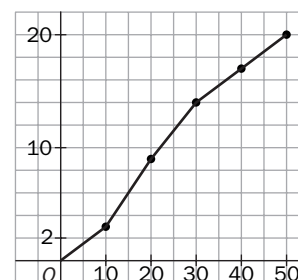


2. Se ha medido la longitud, en centímetros, de las barras de acero producidas por una máquina, obteniéndose los siguientes resultados:
- 10, 29, 27, 19, 24, 21, 13, 16, 17, 16, 16, 10, 15, 16, 15, 20, 15, 14, 21, 11, 12, 14, 12, 17, 14, 18, 17, 15, 25, 13, 19, 17, 18, 16, 16, 17, 16, 18, 20, 11, 26, 11, 13, 12, 15, 12, 14, 28, 17, 17.
- Agrupar los datos en clases de amplitud 4 cm y elaborar la tabla de frecuencias absolutas, relativas, absolutas acumuladas, relativas acumuladas, porcentajes y porcentajes acumulados.

3. Completa la siguiente tabla de frecuencias:

$x_i$	$f_i$	$h_i$	$p_i$	$P_i$
1	10	0,2		
2		0,12	12 %	
3			18 %	
4	12	0,24	24 %	
5	5			
6				
	50			

4. Se considera una distribución de datos agrupados en clases cuyo polígono de frecuencias acumuladas es el de la figura.
- a) Elabora la tabla de frecuencias absolutas, frecuencias relativas y porcentajes.  
b) Representa el histograma de frecuencias absolutas correspondiente.



5. Se ha realizado una prueba de 100 preguntas a un grupo de alumnos, obteniéndose los siguientes resultados:

N.º de respuestas acertadas	[0, 20)	[20, 40)	[40, 60)	[60, 80)	[80, 100)
Número de alumnos	6	9	21	15	9

- a) Elabora la tabla de frecuencias de forma que contenga las marcas de clase, frecuencias relativas, porcentajes y porcentajes acumulados.  
b) Representa el polígono de porcentajes acumulados.  
c) ¿A partir de qué número de respuestas acertadas se encuentra el 40 % de los alumnos que han obtenido mejor resultado?

# SOLUCIONES

1. La población consta de:

$$25\,063 + 40\,291 + 87\,646 = 153\,000 \text{ personas.}$$

El número de integrantes de la muestra de cada zona es:

• Zona A:  $\frac{25\,063}{153\,000} \cdot 200 = 32,76 \approx 33$  personas.

• Zona B:  $\frac{40\,291}{153\,000} \cdot 200 = 52,66 \approx 53$  personas.

• Zona C:  $\frac{87\,646}{153\,000} \cdot 200 = 114,56 \approx 114$  personas.

2.

Clases	$x_i$	$f_i$	$h_i$	$F_i$	$H_i$	$p_i$	$P_i$
[10, 14)	12	12	0,24	12	0,24	24 %	24 %
[14, 18)	16	23	0,46	35	0,7	46 %	70 %
[18, 22)	20	9	0,18	44	0,88	18 %	88 %
[22, 26)	24	2	0,04	46	0,92	4 %	92 %
[26, 30)	28	4	0,08	50	1	8 %	100 %
		50	1			100 %	

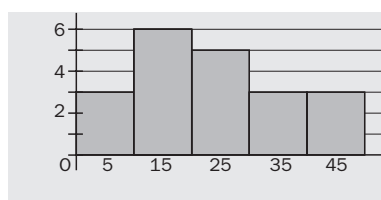
3. Los datos en negrita son los que completan la tabla del enunciado.

$x_i$	$f_i$	$h_i$	$p_i$	$P_i$
1	10	0,2	$0,2 \cdot 100 = 20$	20
2	$0,12 \cdot 50 = 6$	0,12	12	32
3	$0,18 \cdot 50 = 9$	$\frac{18}{100} = 0,18$	18	50
4	12	0,24	24	74
5	5	$\frac{5}{50} = 0,10$	$0,10 \cdot 100 = 10$	84
6	$50 - 42 = 8$	$\frac{8}{50} = 0,16$	$0,16 \cdot 100 = 16$	100
	50			

4.

Intervalos	$x_i$	$f_i$	$h_i$	$p_i$
[0, 10)	5	3	0,15	15 %
[10, 20)	15	$9 - 3 = 6$	0,30	30 %
[20, 30)	25	$14 - 9 = 5$	0,25	25 %
[30, 40)	35	$17 - 14 = 3$	0,15	15 %
[40, 50)	45	$20 - 17 = 3$	0,15	15 %
		20	1	

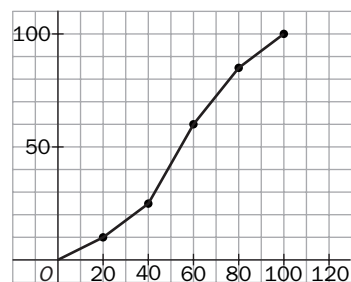
Histograma de frecuencias absolutas:



5.

Clases	$x_i$	$f_i$	$h_i$	$p_i$	$P_i$
[0, 20)	10	6	0,1	10 %	10 %
[20, 40)	30	9	0,15	15 %	25 %
[40, 60)	50	21	0,35	35 %	60 %
[60, 80)	70	15	0,25	25 %	85 %
[80, 100)	90	9	0,15	15 %	100 %
		60	1		

Polígono de porcentajes acumulados:



Como se observa en el polígono de porcentajes acumulados, el número de respuestas acertadas es 60.

# 16 Estadística descriptiva

## CRITERIOS

A. Saber utilizar correctamente el lenguaje estadístico, distinguir y clasificar caracteres y determinar las variables estadísticas que se generan en cada caso.

B. Saber agrupar datos en intervalos o clases eligiendo razonadamente el número y amplitud de los mismos, y elaborar tablas de frecuencias y porcentajes.

C. Elaborar de forma razonada los gráficos estadísticos apropiados a cada tipo de tabla.

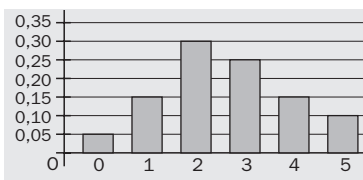
## ACTIVIDADES

1. A un grupo de personas se le somete a una serie de pruebas con objeto de rellenar una ficha que contempla los siguientes aspectos: grupo sanguíneo, sexo, peso, temperatura corporal, número de integrantes de la unidad familiar, resultado de un test sobre diabetes.

- Indica si son caracteres cualitativos o cuantitativos.
- Expresa modalidades posibles del carácter cualitativo o valores de la variable discreta o continua, si son cuantitativos.

2. Se ha preguntado a un grupo de 80 alumnos de un instituto sobre el número de horas que semanalmente dedican a actividades deportivas.

Los resultados vienen dados en el siguiente diagrama de barras de frecuencias relativas:



- Calcula la frecuencia absoluta del dato más frecuente.
- ¿Cuál es el porcentaje del valor 5?

3. Las alturas en centímetros de veinte personas son:

165, 171, 154, 165, 149, 159, 151, 171, 191, 163  
173, 193, 176, 152, 188, 169, 171, 184, 152, 183

Construye una tabla agrupando los datos en intervalos de amplitud 10 desde 140 hasta 200, en la que figuren las marcas de clase, frecuencias absolutas, relativas, absolutas y relativas acumuladas, y porcentajes.

4. Los alumnos de un grupo de primero de universidad se clasifican por edad según la siguiente tabla:

Edad	[17, 19)	[19, 21)	[21, 23)	[23, 25)
Número	52	33	12	3

Elabora una tabla de frecuencias que incluya frecuencias absolutas, relativas, absolutas y relativas acumuladas, porcentajes y porcentajes acumulados.

5. Se ha preguntado a 100 alumnos de un instituto por los gastos semanales, en euros, y se han obtenido los resultados siguientes:

Gastos	[0, 10)	[10, 20)	[20, 30)	[30, 40)
N.º de alumnos	15	42	34	9

Elabora la tabla de frecuencias necesaria para representar el histograma de frecuencias absolutas acumuladas y el polígono de porcentajes.

# SOLUCIONES

1. a) Grupo sanguíneo: cualitativo

Sexo: cualitativo

Peso: cuantitativo

Temperatura corporal: cuantitativo

Número de integrantes de la unidad familiar: cuantitativo

Resultado de un test sobre diabetes: cualitativo

b) Grupo sanguíneo:

$A, B, O, AB$

Sexo: femenino, masculino

Peso:

56 kg, 75 kg

Temperatura corporal:

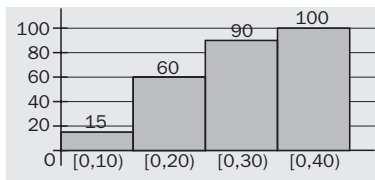
$36^\circ; 37,5^\circ; 39^\circ$

Número de integrantes de la unidad familiar:

1, 2, 3, 4, 5

Resultado de un test sobre diabetes:

positivo, negativo

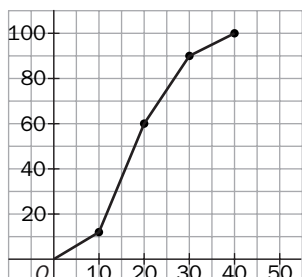


2. a) El dato más frecuente es 2, ya que en el diagrama de barras se observa que le corresponde la barra de mayor longitud.

Su frecuencia absoluta es:

$$0,30 \cdot 80 = 24$$

b) El porcentaje correspondiente a 5 es: 15 %



3.

Clases	$x_i$	$f_i$	$h_i$	$F_i$	$H_i$	$p_i$
[140, 150)	145	1	0,05	1	0,05	5 %
[150, 160)	155	5	0,25	6	0,3	25 %
[160, 170)	165	4	0,2	10	0,5	20 %
[170, 180)	175	5	0,25	15	0,75	25 %
[180, 190)	185	3	0,15	18	0,9	15 %
[190, 200)	195	2	0,1	20	1	10 %
		20	1			100 %

4.

Clases	$x_i$	$f_i$	$h_i$	$F_i$	$H_i$	$p_i$	$P_i$
[17, 19)	18	52	0,52	52	0,52	52 %	52 %
[19, 21)	20	33	0,33	85	0,85	33 %	85 %
[21, 23)	22	12	0,12	97	0,97	12 %	97 %
[23, 25)	24	3	0,03	100	1	3 %	100 %
		100	1			100 %	

5.

Clases	$x_i$	$f_i$	$h_i$	$F_i$	$H_i$	$p_i$	$P_i$
[0, 10)	5	15	0,15	15	0,15	15 %	15 %
[10, 20)	15	45	0,45	60	0,6	45 %	60 %
[20, 30)	25	30	0,3	90	0,9	30 %	90 %
[30, 40)	35	10	0,1	100	1	10 %	100 %
		100	1			100 %	

# 17 | Parámetros estadísticos

1. La siguiente serie da el número de espectadores, por intervalos, y de días de un cine a lo largo de un año:

Número de espectadores	[150, 170)	[170, 190)	[190, 210)	[210, 230)	[230, 250)	[250, 270)	[270, 290)
Número de días	12	30	88	135	63	17	20

- a) Halla la tabla de distribución de frecuencias y de frecuencias acumuladas.  
 b) Calcula la media, la varianza, la desviación típica y el coeficiente de variación.

2. A partir de la siguiente tabla de frecuencias:

Intervalos	[5, 15)	[15, 25)	[25, 35)	[35, 45)
Frecuencias	15	$2p$	$p$	5

- a) Calcula el valor de  $p$  si se sabe que la media es 21.  
 b) Calcula la moda y la mediana.

3. Un estudio revela que en una determinada región,  $A$ , con 250 000 habitantes, el ingreso anual medio por persona es de 12 000 euros, mientras que en otra región,  $B$ , con 500 000 habitantes, el ingreso anual medio es de 15 000 euros.

- a) ¿Cuál es el ingreso anual total de cada región?  
 b) Si consideramos la unión de las dos regiones, ¿cuál es el ingreso anual medio de dicha unión?

4. a) Considera la variable estadística  $X$ , que toma los valores: 5, 8, 9, 12, 15. Calcula su media y su desviación típica.  
 b) Considera la variable  $X + 3$  y calcula, para los nuevos datos, la media y la desviación típica. ¿Qué relación hay entre las medias y las desviaciones típicas de ambas variables,  $X$  y  $X + 3$ ?  
 c) Genera una nueva variable  $2X$  y calcula, para los nuevos números, la media y la desviación típica. ¿Qué relación hay entre las medias y las desviaciones típicas de ambas variables,  $X$  y  $2X$ ?

5. Al estudiar la distribución de la duración de un lote de bombillas, en horas, se obtuvieron los siguientes resultados:

Duración	[0, 20)	[20, 40)	[40, 60)	[60, 80)
N.º de bombillas	15		15	16

- a) ¿Cuál será el dato que falta si la duración media es de 35 horas?  
 b) ¿Cuál sería la desviación típica si ese dato fuera 16?

6. De la distribución de una variable, sabemos que el recorrido es 20 y la distribución está dividida en 4 intervalos de amplitud constante. Las frecuencias de cada intervalo son, por orden, 6, 14, 20 y 10. La media aritmética es 10,9. Una vez hallada la distribución, calcula la desviación típica.

# SOLUCIONES

1. a)

Intervalos	$x_i$	$f_i$	$F_i$	$h_i$	$H_i$
[150, 170)	160	12	12	0,0328	0,0328
[170, 190)	180	30	42	0,0821	0,1149
[190, 210)	200	88	130	0,2410	0,3559
[210, 230)	220	135	265	0,3698	0,7257
[230, 250)	240	63	328	0,1726	0,8983
[250, 270)	260	17	345	0,0465	0,9448
[270, 290)	280	20	365	0,0547	1
		365		1	

$$b) \bar{x} = \frac{\sum_{i=1}^7 x_i \cdot f_i}{\sum_{i=1}^7 f_i} = 218,52$$

$$s^2 = \frac{\sum_{i=1}^7 x_i^2 \cdot f_i}{\sum_{i=1}^7 f_i} - \bar{x}^2 = 684,93$$

$$s = \sqrt{\frac{\sum_{i=1}^7 x_i^2 \cdot f_i}{\sum_{i=1}^7 f_i} - \bar{x}^2} = 26,17$$

$$CV = \frac{s}{\bar{x}} = \frac{26,17}{218,52} = 0,12$$

2. a)  $\bar{x} = 21 \frac{10 \cdot 15 + 20 \cdot 2p + 30 \cdot p + 40 \cdot 5}{15 + 2p + p + 5} =$   
 $= \frac{350 + 70p}{20 + 3p}; 420 + 63p = 350 + 70p;$   
 $7p = 70; p = 10$

b) La moda  $M_o = 20$ , marca de la clase modal, es decir, de la que tiene mayor frecuencia.

c) La mediana  $M = 20$ , marca de clase del primer intervalo cuya frecuencia acumulada, 35, supera el valor  $\frac{n}{2} = 25$ .

3. a) El ingreso anual de la región A es:  
 $12\,000 \cdot 250\,000 = 3\,000\,000\,000$  de euros  
 El ingreso anual de la región B es:  
 $15\,000 \cdot 500\,000 = 7\,500\,000\,000$  de euros

b) El ingreso medio anual de la unión de ambas regiones es:  

$$\frac{12\,000 \cdot 250\,000 + 15\,000 \cdot 500\,000}{250\,000 + 500\,000} =$$
  

$$= \frac{10\,750\,000\,000}{750\,000} = \frac{10\,750\,000\,000}{750\,000} =$$
  

$$= 14\,333,33 \text{ euros}$$

4. Resumimos los resultados en la siguiente tabla.

		$\bar{x}$	$s$
$X$	5, 8, 9, 12, 15	9,80	3,43
$X + 3$	8, 11, 12, 15, 18	$\frac{12,80}{= 9,80 + 3}$	3,43
$2X$	10, 16, 18, 24, 30	$\frac{19,60}{= 2 \cdot 9,80}$	$\frac{6,86}{= 2 \cdot 3,43}$

Se observa que la media aritmética sufre la misma transformación que los datos de la variable.

La desviación típica es invariante por transformaciones del tipo  $x + a$ , pero se ve afectada por transformaciones del tipo  $x \cdot a$ , en la misma medida que los datos de la variable.

5. a)  $\bar{x} = 35 =$   
 $= \frac{10 \cdot 15 + 30 \cdot x + 50 \cdot 15 + 70 \cdot 16}{15 + x + 15 + 16} =$   
 $= \frac{2\,020 + 30x}{46 + x}; 1\,610 + 35x = 2\,020 + 30x$   
 $5x = 410; x = 82$

b)  $S =$   
 $= \sqrt{\frac{10^2 \cdot 15 + 30^2 \cdot 16 + 50^2 \cdot 15 + 70^2 \cdot 16}{15 + 16 + 15 + 16} - \left(\frac{2\,500}{62}\right)^2} =$   
 $= \sqrt{\frac{131\,800}{62} - \left(\frac{2\,500}{62}\right)^2} = 22,36$

6. Como el recorrido de la variable es 20, el número de intervalos es 4, y la amplitud es constante, cada intervalo tiene una amplitud 5. Sea  $x$  la marca de clase del primer intervalo;  $x + 5$ ,  $x + 10$  y  $x + 15$  serán las restantes marcas de los sucesivos intervalos, así:  
 $10,9 =$

$$= \frac{x \cdot 6 + (x + 5) \cdot 14 + (x + 10) \cdot 20 + (x + 15) \cdot 10}{50}$$

$$545 = 50x + 420; 50x = 125; x = 2,5$$

Por tanto, la distribución es:

Intervalos	Marcas	Frecuencia
[0, 5)	2,5	6
[5, 10)	7,5	14
[10, 15)	12,5	20
[15, 20)	17,5	10
		50

La desviación típica:  $s =$   
 $= \sqrt{\frac{2,5^2 \cdot 6 + 7,5^2 \cdot 14 + 12,5^2 \cdot 20 + 17,5^2 \cdot 10}{50} - 10,9^2} =$   
 $= 4,63$

# 17 | Parámetros estadísticos

## CRITERIOS

A. Calcular e interpretar parámetros de centralización con datos agrupados.

B. Calcular e interpretar parámetros de dispersión con datos agrupados.

C. Utilizar conjuntamente la media aritmética y la desviación típica en la especificación del porcentaje de datos en intervalos del tipo  $(\bar{x} - ks, \bar{x} + ks)$ , en distribuciones unimodales y simétricas.

D. Utilizar el coeficiente de variación en la comparación de distribuciones no homogéneas.

## ACTIVIDADES

1. Se ha realizado una encuesta compuesta por diez preguntas a 40 personas. El número de respuestas contestadas figura en la siguiente tabla:

N.º de respuestas contestadas	[0, 2)	[2, 4)	[4, 6)	[6, 8)	[8, 10)
Número de personas	4	9	15	7	5

Halla:

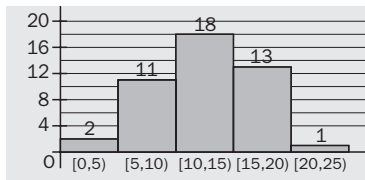
- a) La media.    b) La mediana.    c) La moda.

2. Las ventas de gasolina, en litros, registradas en una estación de servicio durante un día están recogidas en la siguiente tabla:

Ventas	[0, 10)	[10, 20)	[20, 30)	[30, 40)	[40, 50)	[50, 60)	[60, 70)
Vehículos	9	14	16	21	10	12	8

Calcula la varianza, la desviación típica y el coeficiente de variación.

3. El histograma de frecuencias absolutas para ciertos datos es:



- a) Construye la tabla de distribución de frecuencias y la de frecuencias acumuladas.  
 b) Calcula la media aritmética y la desviación típica.  
 c) Calcula el porcentaje de datos que se espera en los siguientes intervalos:

$$(\bar{x} - s, \bar{x} + s) \quad (\bar{x} - 2s, \bar{x} + 2s) \quad (\bar{x} - 3s, \bar{x} + 3s)$$

4. Una distribución tiene las siguientes características: es unimodal, bastante simétrica, el número total de datos es 1 200, la media aritmética es 30 y la desviación típica es 6. Indica cuántos datos de la distribución se espera que se distribuyan en los siguientes intervalos:

- a) (24, 36)  
 b) (18, 42)  
 c) (12, 48)

5. Los ingresos mensuales de un negocio A son de una media de 60 000 € y una desviación típica de 7 500 €. Otro negocio B, del mismo ramo, tiene unos ingresos medios mensuales de 9 000 € con una desviación típica de 1 500 €. ¿Cuál de los dos negocios tiene mayor variación en sus ingresos?

# SOLUCIONES

1. Disponemos los datos en la siguiente tabla:

Intervalos	Marcas de clase $x_i$	$f_i$	$F_i$	$x_i f_i$
[0, 2)	1	4	4	4
[2, 4)	3	9	13	27
[4, 6)	5	15	28	75
[6, 8)	7	7	35	49
[8, 10)	9	5	40	45
		40		200

- a) La media aritmética es:  $\bar{x} = \frac{200}{40} = 5$
- b) La mediana:  $M = 5$ , ya que  $\frac{n}{2} = 20$ , siendo la primera frecuencia absoluta acumulada que supera dicho valor, 28, que corresponde a la clase mediana [4, 6), cuya marca de clase es 5.
- c) La moda:  $M_o = 5$ , marca de la clase modal, que es la clase más frecuente.

2. Elaboramos la tabla:

Intervalos	Marcas	$f_i$	$x_i f_i$	$x_i^2 f_i$
[0, 10)	5	9	45	225
[10, 20)	15	14	210	3 150
[20, 30)	25	16	400	10 000
[30, 40)	35	21	735	25 725
[40, 50)	45	10	450	20 250
[50, 60)	55	12	660	36 300
[60, 70)	65	8	520	33 800
		90	3 020	129 450

Media aritmética:  $\bar{x} = \frac{3 020}{90} = 33,56$

Varianza:  $s^2 = \frac{129 450}{90} - 33,56^2 = 312,36$

Desviación típica:  $s = \sqrt{\frac{129 450}{90} - 33,56^2} = 17,67$

Coefficiente de variación:  $CV = \frac{17,67}{33,56} = 0,53$

3. a) A partir del histograma se obtiene la tabla de distribución de frecuencias. La completamos con las columnas  $x_i f_i$  y  $x_i^2 f_i$ , que posteriormente utilizaremos para calcular diversos parámetros:

Intervalos	Marcas	$f_i$	$F_i$	$x_i f_i$	$x_i^2 f_i$
[0, 5)	2,5	2	2	5	12,5
[5, 10)	7,5	11	13	82,5	618,75
[10, 15)	12,5	18	31	225	2 812,5
[15, 20)	17,5	13	44	227,5	3 981,25
[20, 25)	22,5	1	45	22,5	506,25
		45		562,5	7 931,25

b) Media aritmética:  $\bar{x} = \frac{562,5}{45} = 12,5$

Desviación típica:  $s = \sqrt{\frac{7 931,25}{45} - 12,5^2} = 4,47$

- c) Del histograma de frecuencias absolutas se deduce que la distribución es unimodal y bastante simétrica; por tanto, es de esperar que en:

$(\bar{x} - s, \bar{x} + 2)$  se distribuye el 68 %

$(\bar{x} - 2s, \bar{x} + 2s)$  se distribuye el 94 %

$(\bar{x} - 3s, \bar{x} + 3s)$  se distribuye el 100 %

4. a) En el intervalo  $(\bar{x} - s, \bar{x} + s) = (24, 36)$  se distribuye el 68 %, esto es, 816 datos.
- b) En el intervalo  $(\bar{x} - 2s, \bar{x} + 2s) = (18, 42)$  se distribuye el 94 %, esto es, 1 128 datos.
- c) En el intervalo  $(\bar{x} - 3s, \bar{x} + 3s) = (12, 48)$  se distribuye el 100 %, esto es, 1 200 datos.

5.  $CV_A = \frac{7 500}{60 000} = 0,125$

$CV_B = \frac{1 500}{9 000} = 0,167$

El negocio B tiene mayor variación en sus ingresos.



# 17 | Parámetros estadísticos

1. Desarrolla los siguientes sumatorios:

a)  $\sum_{i=1}^4 x_i \cdot x_{i+1}$

b)  $\sum_{i=1}^5 x_i (y_i - 3)$

2. Expresa, mediante el símbolo sumatorio, los siguientes desarrollos:

a)  $x_1^2 + x_2^2 + x_3^2 + x_4^2$

b)  $(x_1 - 4)f_2 + (x_2 - 4)f_3 + (x_3 - 4)f_4$

3. Halla la media aritmética, la mediana y la moda del siguiente conjunto de datos:

5, 3, 6, 5, 4, 5, 2, 8, 6, 5, 4, 8, 3, 4, 5, 4, 8, 2, 5, 4

4. Se han medido las longitudes de una serie de varillas metálicas, los resultados están reflejados en la siguiente tabla. Calcula la media, la mediana y la moda.

Longitudes	Número de varillas
[60, 63)	5
[63, 66)	18
[66, 69)	43
[69, 72)	27
[72, 75)	7

5. Una distribución estadística viene dada por la siguiente tabla de frecuencias:

Clases	$f_i$
[0, 10)	8
[10, 20)	12
[20, 30)	10
[30, 40)	6
	36

Halla:

- a) La media aritmética.
- b) El rango.
- c) La varianza y la desviación típica.

6. Una distribución viene dada por la siguiente tabla de frecuencias:

Datos	2	5	6	9	11	13
Frecuencia	3	7	15	17	6	2

- a) Halla el coeficiente de variación.
- b) Calcula el porcentaje de datos que se encuentran en el intervalo  $(\bar{x} - 2s, \bar{x} + 2s)$ .

# SOLUCIONES

1. a)  $\sum_{i=1}^4 x_i \cdot x_{i+1} = x_1 \cdot x_2 + x_2 \cdot x_3 + x_3 \cdot x_4 + x_4 \cdot x_5$

b)  $\sum_{i=1}^5 x_i(y_i - 3) = x_1(y_1 - 3) + x_2(y_2 - 3) + x_3(y_3 - 3) + x_4(y_4 - 3) + x_5(y_5 - 3)$

2. a)  $x_1^2 + x_2^2 + x_3^2 + x_4^2 = \sum_{i=1}^4 x_i^2$

b)  $(x_1 - 4)f_2 + (x_2 - 4)f_3 + (x_3 - 4)f_4 = \sum_{i=1}^3 (x_i - 4)f_{i+1}$

3. A partir de los datos del enunciado elaboramos la siguiente tabla:

$x_i$	$f_i$	$F_i$	$x_i \cdot f_i$
2	2	2	4
3	2	4	6
4	5	9	20
5	6	15	30
6	2	17	12
8	3	20	24
	20		96

$$\bar{x} = \frac{\sum_{i=1}^n x_i \cdot f_i}{\sum_{i=1}^n f_i} = \frac{96}{20} = 4,8$$

La primera frecuencia acumulada que supera a 10 es 15, que corresponde al 5. La mediana es 5.

La moda  $M_o = 5$ , es el valor más frecuente.

4. Elaboramos la siguiente tabla:

Longitudes	Marcas de clase $x_i$	Número de varillas $f_i$	$F_i$	$x_i \cdot f_i$
[60, 63)	61,5	5	5	307,5
[63, 66)	64,5	18	23	1 161
[66, 69)	67,5	43	66	2 902,5
[69, 72)	70,5	27	93	1 903,5
[72, 75)	73,5	7	100	514,5
		100		6 789

$$\bar{x} = \frac{6 789}{100} = 67,89$$

La primera frecuencia acumulada que supera a 50 es 66, que corresponde a la clase mediana [66, 69), cuya marca de clase es 67,5. La mediana es 67,5.

La moda  $M_o = 67,5$ , marca de la clase modal.

5. Elaboramos la siguiente tabla:

Clases	Marcas de clase $x_i$	$f_i$	$x_i \cdot f_i$	$x_i^2 \cdot f_i$
[0, 10)	5	8	40	200
[10, 20)	15	15	225	3 375
[20, 30)	25	11	275	6 875
[30, 40)	35	6	210	7 350
		40	750	17 800

a)  $\bar{x} = \frac{750}{40} = 18,75$ .      b) Rango: 40

c)  $V = \frac{\sum_{i=1}^n f_i \cdot x_i^2}{\sum_{i=1}^n f_i} - \bar{x}^2 = 93,44$

Desviación típica:  $s = \sqrt{93,44} = 9,67$

6. Calculamos la media aritmética y la desviación típica, para poder obtener el coeficiente de variación:

$x_i$	$f_i$	$x_i \cdot f_i$	$x_i^2 \cdot f_i$
2	3	6	12
5	7	35	175
6	15	90	540
9	17	153	1 377
11	6	66	726
13	2	26	338
	50	376	3 168

$$\bar{x} = \frac{376}{50} = 7,52$$

$$s = \sqrt{\frac{\sum_{i=1}^n f_i \cdot x_i^2}{\sum_{i=1}^n f_i} - \bar{x}^2} = \sqrt{\frac{3 168}{50} - 7,52^2} = 2,61$$

a) Coeficiente de variación:  $CV = \frac{s}{\bar{x}} = \frac{2,61}{7,52} = 0,35$

b) El intervalo  $(\bar{x} - 2s, \bar{x} + 2s) = (2,3; 12,74)$

En este intervalo hay 45 datos, que suponen el 90 % del total.