9 Soluciones a las actividades de cada epígrafe

PÁGINA 188

En este torneo de ajedrez participan 80 jugadores. La finalidad del torneo es clasificar al 10% de los participantes para la fase final. La tabla resume los resultados finales:

PUNTUACIÓN	6	5,5	5	4,5	4	…
N.° DE JUG.	4	3	3	8	7	…

1 **Teniendo en cuenta la tabla anterior, ¿cuál es la menor puntuación con la que se ha clasificado un jugador?**

El 10% de 80 es 8. Este es el número de jugadores que se clasifican para la fase final. Por tanto, se clasificarán los cuatro que han obtenido 6 puntos, los tres de 5,5 puntos y uno de los que han obtenido 5 puntos.

Por tanto, la menor puntuación con la que se ha clasificado un jugador es 5 puntos.

2 **De los 80 jugadores, hay 16 con el título de Maestro Provincial (M.P.). Sus puntuaciones en el torneo se dan en la tabla siguiente:**

PUNTOS	6	5,5	5	4,5	4
N.° DE M.P.	4	2	2	6	2

¿Cuál ha sido la puntuación media de estos jugadores (M.P.)?

$$\text{Puntuación media} = \frac{4 \cdot 6 + 2 \cdot 5{,}5 + 2 \cdot 5 + 6 \cdot 4{,}5 + 2 \cdot 4}{4 + 2 + 2 + 6 + 2} = \frac{80}{16} = 5$$

La puntuación media de los M.P. ha sido de 5 puntos.

PÁGINA 189

ANTES DE COMENZAR, RECUERDA

1 **Deseamos hacer un estudio comparativo de algunos aspectos de los distintos países del mundo (número de habitantes, renta per cápita, religión predominante y número de ciudades con más de 500 000 habitantes). En este estudio estadístico, ¿cuál es la población? ¿Cuáles son los individuos? Di cuáles son las variables y de qué tipo son.**

- Población → los países del mundo.
- Individuos → cada uno de los países.
- Variables:
  - — N.° de habitantes → cuantitativa discreta.
  - — Renta per cápita → cuantitativa continua.
  - — Religión predominante → cualitativa.
  - — N.° de ciudades con más de 500 000 habitantes → cuantitativa discreta.

PÁGINA 191

1 **Reparte los cuarenta datos del ejercicio resuelto anterior en 10 intervalos con el mismo recorrido total.**

Tomando $r' = 30$ y siendo 10 el número de intervalos, la longitud de cada intervalo será de $\frac{30}{10} = 3$.

INTERVALO	MARCAS DE CLASE	FRECUENCIAS
148,5 - 151,5	150	2
151,5 - 154,5	153	1
154,5 - 157,5	156	1
157,5 - 160,5	159	6
160,5 - 163,5	162	7
163,5 - 166,5	165	9
166,5 - 169,5	168	6
169,5 - 172,5	171	3
172,5 - 175,5	174	4
175,5 - 178,5	177	1

2 **Reparte los cuarenta datos del ejercicio resuelto anterior en 8 intervalos. Para ello, toma $r' = 32$.**

Tomando $r' = 32$ y siendo 8 el número de intervalos, la longitud de cada uno de ellos será $\frac{32}{8} = 4$.

INTERVALO	MARCAS DE CLASE	FRECUENCIAS
147,5 - 151,5	149,5	2
151,5 - 155,5	153,5	1
155,5 - 159,5	157,5	4
159,5 - 163,5	161,5	10
163,5 - 167,5	165,5	12
167,5 - 171,5	169,5	6
171,5 - 175,5	173,5	4
175,5 - 179,5	177,5	1

PÁGINA 193

1 **Halla, manualmente y con calculadora, $\bar{x}$, σ y C.V. en la tabla obtenida en el ejerecicio resuelto de la página 191:**

x_i	151	156	161	166	171	176
f_i	2	4	11	14	5	4

x_i	f_i	$f_i x_i$	$f_i x_i^2$
151	2	302	45 602
156	4	624	97 344
161	11	1 771	285 131
166	14	2 324	385 784
171	5	855	146 205
176	4	704	123 904
	40	6 580	1 083 970

$$\bar{x} = \frac{6\,580}{40} = 164{,}5$$

$$\sigma = \sqrt{\frac{1\,083\,970}{40} - 164{,}5^2} = 6{,}24$$

$$\text{C.V.} = \frac{6{,}24}{164{,}5} = 0{,}038 \rightarrow 3{,}8\%$$

2 **Halla, manualmente y con calculadora, $\bar{x}$, σ y C.V. en la distribución de los ejercicios 1 y 2 de la página 191.**

Compara los resultados entre sí y con los del ejercicio 1 de esta página.

1.ª distribución

INTERVALOS	x_i	f_i	$f_i x_i$	$f_i x_i^2$
148,5 - 151,5	150	2	300	45 000
151,5 - 154,5	153	1	153	23 409
154,5 - 157,5	156	1	156	24 336
157,5 - 160,5	159	6	954	151 686
160,5 - 163,5	162	7	1 134	183 708
163,5 - 166,5	165	9	1 485	245 025
166,5 - 169,5	168	6	1 008	169 344
169,5 - 172,5	171	3	513	87 723
172,5 - 175,5	174	4	696	121 104
175,5 - 178,5	177	1	177	31 329
		40	6 576	1 082 664

MEDIA: $\bar{x} = \dfrac{\Sigma f_i x_i}{\Sigma f_i} = \dfrac{6\,576}{40} = 164{,}4$ cm

VAR.: $\dfrac{\Sigma f_i x_i^2}{\Sigma f_i} - \bar{x}^2 = \dfrac{1\,082664}{40} - 164{,}4^2 = 39{,}24$

DESVIACIÓN TÍPICA: $\sigma = \sqrt{39{,}24} = 6{,}26$ cm

$$\text{C.V.} = \frac{\sigma}{\bar{x}} = \frac{6{,}26}{164{,}4} = 0{,}038 \rightarrow 3{,}8\%$$

2.ª distribución

INTERVALOS	x_i	f_i	$f_i x_i$	$f_i x_i^2$
147,5 - 151,5	149,5	2	299	44 700,5
151,5 - 155,5	153,5	1	153,5	23 562,25
155,5 - 159,5	157,5	4	630	99 225
159,5 - 163,5	161,5	10	1 615	260 822,5
163,5 - 167,5	165,5	12	1 986	328 683
167,5 - 171,5	169,5	6	1 017	172 381,5
171,5 - 175,5	173,5	4	694	120 409
175,5 - 179,5	175,5	1	177,5	31 506,25
		40	6 572	1 081 290

MEDIA: $\bar{x} = \dfrac{\Sigma f_i x_i}{\Sigma f_i} = \dfrac{6\,572}{40} = 164{,}3$ cm

VAR.: $\dfrac{\Sigma f_i x_i^2}{\Sigma f_i} - \bar{x}^2 = \dfrac{1\,081\,290}{40} - 164{,}3^2 = 37{,}76$

DESVIACIÓN TÍPICA: $\sigma = \sqrt{37{,}76} = 6{,}14$ cm

C.V. $= \dfrac{\sigma}{\bar{x}} = \dfrac{6{,}14}{164{,}3} = 0{,}037 \rightarrow 3{,}7\%$

Como se puede ver, las diferencias entre unas y otras son inapreciables.

PÁGINA 195

1 **En la siguiente distribución de notas, halla** Me, Q_1, Q_3, p_{80}, p_{90} **y** p_{99}.

NOTAS	1	2	3	4	5	6	7	8	9	10
N.° DE ALUMNOS	7	15	41	52	104	69	26	13	19	14

x_i	f_i	F_i	EN %
1	7	7	1,94
2	15	22	6,11
3	41	63	17,5
4	52	115	31,94
5	104	219	60,83
6	69	288	80
7	26	314	87,22
8	13	327	90,83
9	19	346	96,11
10	14	360	100

$Me = p_{50} = 5$, porque F_i supera el 50% para $x_i = 5$.

$Q_1 = p_{25} = 4$, porque F_i supera el 25% para $x_i = 4$.

$Q_3 = p_{75} = 6$, porque F_i supera el 75% para $x_i = 6$.

$p_{80} = 6{,}5$, porque para $x_i = 6$, la F_i iguala al 80%. En ese caso, se toma como p_{80} el valor intermedio entre 6 y 7.

$p_{90} = 8$, porque para $x_i = 8$, F_i supera el 90%.

$p_{99} = 10$, porque para $x_i = 10$, F_i supera el 99%.

PÁGINA 197

1 Haz el diagrama de caja correspondiente a esta distribución de notas.

x_i	f_i
1	6
2	15
3	22
4	24
5	33
6	53
7	22
8	16
9	8
10	1

x_i	f_i	F_i
1	6	6
2	15	21
3	22	43
4	24	67
5	33	100
6	53	153
7	22	175
8	16	191
9	8	199
10	1	200

Comenzamos hallando Me, Q_1 y Q_3:

$n = 200$

$\frac{n}{2} = 100 \rightarrow Me = 5{,}5$

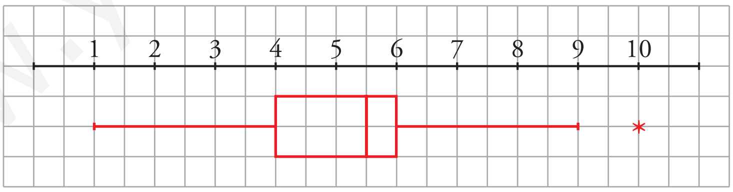
$\frac{n}{4} = 50 \rightarrow Q_1 = 4$

$\frac{3}{4} \cdot n = 150 \rightarrow Q_3 = 6$

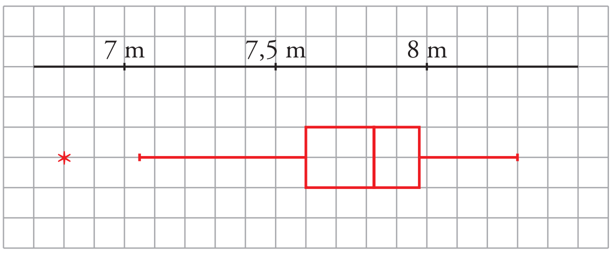
La longitud de la caja será $Q_3 - Q_1 = 6 - 4 = 2$.

$1{,}5 \cdot 2 = 3 \rightarrow$ Los bigotes llegarán hasta $4 - 3 = 1$ y hasta $6 + 3 = 9$.

Por tanto, el diagrama de caja y bigotes será:



2 Interpreta el siguiente diagrama de caja relativo a marcas de saltadores de longitud.



$Me = 7{,}825$ m; $Q_1 = 7{,}6$ m; $Q_3 = 7{,}975$ m

Todos saltaron entre 7,05 m y 8,3 m, excepto uno que saltó 6,8 m.

Un 25% de los saltadores saltó menos de 7,6 m.

Un 25% saltó entre 7,6 m y 7,825 m.

Un 25% saltó entre 7,825 m y 7,975 m.

Un 25% saltó más de 7,975 m.

PÁGINA 198

1 **Un fabricante de tornillos desea hacer un control de calidad. Recoge uno de cada 100 tornillos fabricados y lo analiza. El conjunto de tornillos analizados, ¿es población o muestra? ¿Por qué?**

Los tornillos analizados constituyen una muestra, pues solo se analiza uno de cada cien tornillos fabricados.

2 **Un fabricante de vasos de vidrio quiere estudiar la resistencia que presentan a la rotura. El procedimiento consiste en someterlos a presiones paulatinamente crecientes hasta que se parten. ¿Puede hacer el estudio sobre la población o debe recurrir a una muestra? ¿Por qué?**

Sí, debe recurrir a una muestra porque el proceso de estudio es destructivo.

3 **Un campesino posee 127 gallinas. Para probar la eficacia de un nuevo tipo de alimentación, las pesa a todas antes y después de los veinte días que dura el tratamiento. El conjunto de esas 127 gallinas, ¿es población o muestra? ¿Por qué?**

Es una población, pues el análisis se hace sobre las 127 gallinas del campesino. Ahora bien, si de ese estudio se pretende sacar conclusiones para el conjunto de todas las gallinas de España, o de Europa, o del mundo, estaríamos hablando de una muestra.

PÁGINA 200

4 **Se les ha pasado un test a los 64 individuos de una muestra seleccionada aleatoriamente. Con los resultados obtenidos se ha llegado a la siguiente conclusión: "La puntuación media que alcanzarían los individuos de toda la población si se les pasara este test estaría entre 42,7 puntos y 44,1 puntos. Y esto lo podemos afirmar con un nivel de confianza del 95%".**

a) Si el intervalo que se diera fuera 42-44,8, el nivel de confianza sería del...

• 90% • 95% • 98%

b) Si quisiéramos un nivel de confianza del 99% y un intervalo de la misma amplitud, ¿cómo tendría que ser la muestra?

• De menos de 64 • De 64 • De más de 64

a) 98%, pues al ampliar la longitud del intervalo, también aumenta el nivel de confianza.

b) De más de 64, pues cuanto mayor es la muestra, mayor es el nivel de confianza.

PÁGINA 201

PRACTICA

Tablas de frecuencias

1 **El número de faltas de ortografía que cometieron un grupo de estudiantes en un dictado fue:**

0	**3**	**1**	**2**	**0**	**2**	**1**	**3**	**0**	**4**
0	**1**	**1**	**4**	**3**	**5**	**3**	**2**	**4**	**1**
5	**0**	**2**	**1**	**0**	**0**	**0**	**0**	**2**	**1**
2	**1**	**0**	**0**	**3**	**0**	**5**	**3**	**2**	**1**

a) Di cuál es la variable y de qué tipo es.

b) Haz una tabla de frecuencias y representa los datos en un diagrama adecuado.

a) Variable: "Número de faltas de ortografía"

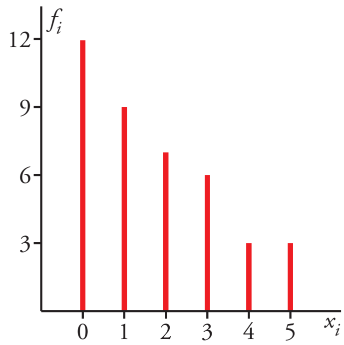
Es una variable cuantitativa discreta.

Llamamos x_i a dicha variable y sus valores son 0, 1, 2, 3, 4 y 5.

b) Tabla de frecuencias:

x_i	f_i
0	12
1	9
2	7
3	6
4	3
5	3
	40

Diagrama de barras:



2 **Las urgencias atendidas durante un mes en un centro de salud fueron:**

1	**5**	**3**	**2**	**1**	**6**	**4**	**2**	**2**	**3**
4	**3**	**5**	**1**	**0**	**1**	**5**	**3**	**3**	**6**
2	**4**	**6**	**3**	**2**	**4**	**3**	**2**	**1**	**5**

a) ¿Cuál es la variable y de qué tipo es?

b) Haz una tabla de frecuencias y representa los datos.

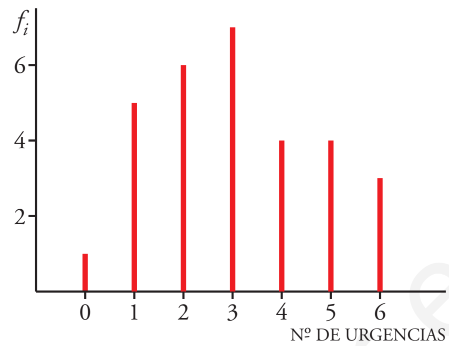
a) Variable: n.º de urgencias atendidas.

Tipo: cuantitativa discreta.

b)

x_i = URGENCIAS ATENDIDAS	f_i
0	1
1	5
2	6
3	7
4	4
5	4
6	3

Representamos los datos en un diagrama de barras:



3 **En una maternidad se han tomado los pesos (en kilogramos) de 50 recién nacidos:**

2,8	**3,2**	**3,8**	**2,5**	**2,7**	**3,7**	**1,9**	**2,6**	**3,5**	**2,3**
3,0	**2,6**	**1,8**	**3,3**	**2,9**	**2,1**	**3,4**	**2,8**	**3,1**	**3,9**
2,9	**3,5**	**3,0**	**3,1**	**2,2**	**3,4**	**2,5**	**1,9**	**3,0**	**2,9**
2,4	**3,4**	**2,0**	**2,6**	**3,1**	**2,3**	**3,5**	**2,9**	**3,0**	**2,7**
2,9	**2,8**	**2,7**	**3,1**	**3,0**	**3,1**	**2,8**	**2,6**	**2,9**	**3,3**

a) ¿Cuál es la variable y de qué tipo es?

b) Construye una tabla con los datos agrupados en 6 intervalos de 1,65 a 4,05.

c) Representa gráficamente esta distribución.

Localizamos los valores extremos: 1,8 y 3,9.

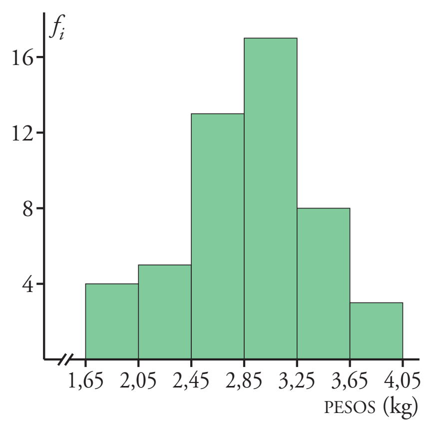
Recorrido = 3,9 – 1,8 = 2,1

a) Variable: peso de los recién nacidos.
Tipo: cuantitativa continua.

b)

INTERVALOS	MARCA DE CLASE (x_i)	f_i
1,65 - 2,05	1,85	4
2,05 - 2,45	2,25	5
2,45 - 2,85	2,65	13
2,85 - 3,25	3,05	17
3,25 - 3,65	3,45	8
3,65 - 4,05	3,85	3
		50

c) Representamos los datos en un histograma; al ser los intervalos de la misma amplitud, la altura de cada barra corresponde a la frecuencia (f_i) de cada intervalo.



4 □□□ **A un grupo de 30 personas se les ha tomado el número de pulsaciones por minuto (ritmo cardíaco) obteniéndose los siguientes resultados:**

87	**85**	**61**	**51**	**64**	**75**	**80**	**70**	**69**	**82**
80	**79**	**82**	**74**	**92**	**76**	**72**	**73**	**63**	**65**
67	**71**	**88**	**76**	**68**	**73**	**70**	**76**	**71**	**86**

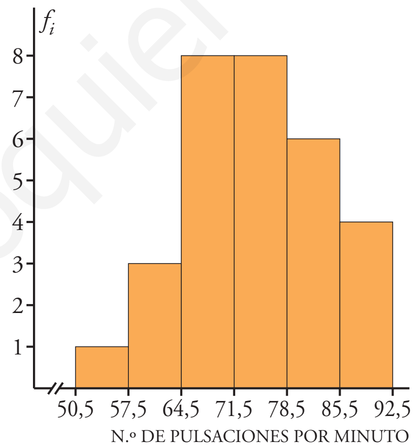
Representa gráficamente esta distribución agrupando los datos en 6 intervalos (desde 50,5 a 92,5).

Cada intervalo medirá $\dfrac{92{,}5 - 50{,}5}{6} = 7$.

Tabla de frecuencias:

INTERVALO	MARCAS DE CLASE	FRECUENCIA
50,5 - 57,5	54	1
57,5 - 64,5	61	3
64,5 - 71,5	68	8
71,5 - 78,5	75	8
78,5 - 85,5	82	6
85,5 - 92,5	89	4

Histograma:



Puesto que los intervalos son de la misma longitud, la altura de cada barra en este histograma coincide con la frecuencia.

Media, desviación típica y C.V.

Halla la media, la desviación típica y el coeficiente de variación en las siguientes distribuciones:

5

x_i	f_i
0	12
1	9
2	7
3	6
4	3
5	3

x_i	f_i	$f_i x_i$	$f_i x_i^2$
0	12	0	0
1	9	9	9
2	7	14	28
3	6	18	54
4	3	12	48
5	3	15	75
	40	68	214

MEDIA: $\bar{x} = \dfrac{\Sigma f_i x_i}{\Sigma f_i} = \dfrac{68}{40} = 1{,}7$

VAR.: $\dfrac{\Sigma f_i x_i^2}{\Sigma f_i} - \bar{x}^2 = \dfrac{214}{40} - 1{,}7^2 = 2{,}46$

DESVIACIÓN TÍPICA: $\sigma = \sqrt{2{,}46} = 1{,}57$

COEFICIENTE DE VARIACIÓN: $\text{C.V.} = \dfrac{\sigma}{\bar{x}} = 0{,}9235 \to 92{,}35\%$

6

x_i	f_i
0	1
1	5
2	6
3	7
4	4
5	4
6	3

x_i	f_i	$f_i x_i$	$f_i x_i^2$
0	1	1	0
1	5	5	5
2	6	12	24
3	7	21	63
4	4	16	64
5	4	20	100
6	3	18	108
	30	93	364

MEDIA: $\bar{x} = \dfrac{\Sigma f_i x_i}{\Sigma f_i} = \dfrac{93}{30} = 3{,}1$

VAR.: $\dfrac{\Sigma f_i x_i^2}{\Sigma f_i} - \bar{x}^2 = \dfrac{364}{30} - 3{,}1^2 = 2{,}52$

DESVIACIÓN TÍPICA: $\sigma = \sqrt{2{,}52} = 1{,}59$

COEFICIENTE DE VARIACIÓN: $\text{C.V.} = \dfrac{\sigma}{\bar{x}} = 0{,}5129 \to 51{,}29\%$

7

INTERVALO	f_i
1,65-2,05	4
2,05-2,45	5
2,45-2,85	13
2,85-3,25	17
3,25-3,65	8
3,65-4,05	3

INTERVALOS	x_i	f_i	$f_i x_i$	$f_i x_i^2$
1,65 - 2,05	1,85	4	7,4	13,69
2,05 - 2,45	2,25	5	11,25	25,31
2,45 - 2,85	2,65	13	34,45	91,29
2,85 - 3,25	3,05	17	51,85	158,14
3,25 - 3,65	3,45	8	27,6	95,22
3,65 - 4,05	3,85	3	11,55	44,47
		50	144,1	428,12

$\bar{x} = \dfrac{144,1}{50} = 2,9$

$\text{VAR.} = \dfrac{428,12}{50} - 2,8^2 = 0,1524$

$\sigma = \sqrt{0,1524} = 0,39$

$\text{C.V.} = \dfrac{0,39}{2,9} = 0,1345 \rightarrow 13,45\%$

8

INTERVALO	f_i
50,5-57,5	1
57,5-64,5	3
64,5-71,5	8
71,5-78,5	8
78,5-85,5	6
85,5-92,5	4

INTERVALO	x_i	f_i	$f_i x_i$	$f_i x_i^2$
50,5-57,5	54	1	54	2 916
57,5-64,5	61	3	183	11 163
64,5-71,5	68	8	544	36 992
71,5-78,5	75	8	600	45 000
78,5-85,5	82	6	492	40 344
85,5-92,5	89	4	356	31 684
		30	2 229	168 099

$\bar{x} = \dfrac{2\,229}{30} = 74,3$

$\sigma = \sqrt{\dfrac{168\,099}{30} - 74,3^2} = 9,1$

$\text{C.V.} = \dfrac{9,1}{74,3} = 0,1225 \rightarrow 12,25\%$

9 **Los gastos mensuales de una empresa *A* tienen una media de 100 000 euros y una desviación típica de 12 500 euros. En otra empresa *B* la media es 15 000 euros, y la desviación típica, 2 500 euros. Calcula el coeficiente de variación y di cuál de las dos tiene más variación relativa.**

Empresa *A*: $\left.\begin{array}{l}\bar{x} = 100\,000 \text{ €} \\ \sigma = \ \ 12\,500 \text{ €}\end{array}\right\}$ $\text{C.V.} = \dfrac{\sigma}{\bar{x}} = \dfrac{12\,500}{100\,000} = 0,125$ o bien 12,5%

Empresa *B*: $\left.\begin{array}{l}\bar{x} = 15\,000 \text{ €} \\ \sigma = \ \ 2\,500 \text{ €}\end{array}\right\}$ $\text{C.V.} = \dfrac{2\,500}{15\,000} = 0,1\widehat{6}$ o bien 16,67%

Tiene mayor variación relativa la empresa *B*.

10 □□□ **El peso medio de los alumnos de una clase es de 58,2 kg, y su desviación típica, 3,1 kg. El de las alumnas de esa clase es 52,4 kg y su desviación típica es 5,2 kg. Calcula el coeficiente de variación y compara la dispersión de ambos grupos.**

Alumnos $\left\{\begin{array}{l}\bar{x} = 58{,}2 \text{ kg} \\ \sigma = \ \ 3{,}1 \text{ kg}\end{array}\right\}$ C.V. $= \dfrac{3{,}1}{58{,}2} = 0{,}053 \ \rightarrow\ 5{,}3\%$

Alumnas $\left\{\begin{array}{l}\bar{x} = 52{,}4 \text{ kg} \\ \sigma = \ \ 5{,}2 \text{ kg}\end{array}\right\}$ C.V. $= \dfrac{5{,}2}{52{,}4} = 0{,}099 \ \rightarrow\ 9{,}9\%$

El peso medio de las alumnas es más variable que el peso de los alumnos.

11 □□□ **Se han pedidos los pesos y las alturas de 6 personas, obteniéndose los siguientes datos:**

PESO (kg)	ALTURA (m)
65	1,70
60	1,50
63	1,70
63	1,70
68	1,75
68	1,80

Calcula el coeficiente de variación y di si están más dispersos los pesos o las alturas.

PESOS (x_i)	f_i	$f_i x_i$	$f_i x_i^2$
60	1	60	3 600
63	2	126	7 938
65	1	65	4 225
68	2	136	9 248
	6	387	25 011

$$\bar{x} = \frac{\Sigma f_i x_i}{\Sigma f_i} = \frac{387}{6} = 64{,}5 \text{ kg}$$

$$\text{VAR.} = \frac{\Sigma f_i x_i^2}{\Sigma f_i} - \bar{x}^2 = \frac{25\,011}{6} - 64{,}5^2 = 8{,}25 \ \rightarrow\ \sigma = \sqrt{8{,}25} = 2{,}87 \text{ kg}$$

$$\text{C.V.} = \frac{\sigma}{\bar{x}} = \frac{2{,}87}{64{,}5} = 0{,}044 \text{ o bien } 4{,}4\%$$

ALTURAS (y_i)	f_i	$f_i y_i$	$f_i y_i^2$
1,5	1	1,5	2,25
1,7	3	5,1	8,67
1,75	1	1,75	3,06
1,8	1	1,8	3,24
	6	10,15	17,22

$$\bar{y} = \frac{\Sigma f_i y_i}{\Sigma f_i} = \frac{10{,}15}{6} = 1{,}69 \text{ m}$$

$$\text{VAR.} = \frac{\Sigma f_i y_i^2}{\Sigma f_i} - \bar{y}^2 = \frac{17{,}22}{6} - 1{,}69^2 = 0{,}0139 \ \rightarrow\ \sigma = \sqrt{20{,}0139} = 0{,}12 \text{ m}$$

$$\text{C.V.} = \frac{\sigma}{\bar{y}} = \frac{0{,}12}{1{,}69} = 0{,}071 \text{ o bien } 7{,}1\%$$

Están más dispersas las alturas que los pesos.

PÁGINA 202

Medidas de posición

12 **La mediana y los cuartiles de la distribución de "Aptitud para la música" (escala 1-100) en un colectivo de personas son $Q_1 = 31$, $Me = 46$ y $Q_3 = 67$.**

Completa las siguientes afirmaciones:

a) El 75% tiene una aptitud superior o igual a ____.

b) El 25% tiene una aptitud superior o igual a ____.

c) El ____% tiene una aptitud igual o menor a 46 puntos.

d) El ____% tiene una aptitud superior o igual a 46 e inferior o igual a 67.

e) El ____% tiene una aptitud superior o igual a 31 e inferior o igual a 67.

a) $Q_1 = 31$ b) $Q_3 = 67$ c) 50%

d) 25% e) 50%

13 **La altura, en centímetros, de un grupo de alumnos y alumnas de una misma clase es:**

150	**169**	**171**	**172**	**172**	**175**	**181**
182	**183**	**177**	**179**	**176**	**184**	**158**

Calcula la mediana y los cuartiles y explica el significado de estos parámetros.

Colocamos los datos en orden creciente:

150 - 158 - 169 - 171 - 172 - 172 - 175 - 176 - 177 - 179 - 181 - 182 - 183 - 184

Hay 14 datos:

$\frac{14}{2} = 7 \rightarrow$ Mediana: valor intermedio de los dos centrales situados en séptima y octava posición:

$$Me = \frac{175 + 176}{2} = 175{,}5 \text{ cm}$$

Significa que la mitad de los alumnos tiene una estatura inferior a 175,5 cm.

$\frac{14}{4} = 3{,}5 \rightarrow$ $Q_1 = 171$ cm (4.° lugar)

El 25% de los alumnos mide menos de 171 cm de altura.

$14 \cdot \frac{3}{4} = 10{,}5 \rightarrow$ $Q_3 = 181$ cm (posición 11)

El 75% de los alumnos tiene una estatura inferior a 181 cm.

14 □□□ **Calcula la mediana y los cuartiles de la siguiente distribución:**

x_i	0	1	2	3	4	5
f_i	12	9	7	6	3	3

x_i	f_i	F_i	EN %
0	12	12	30
1	9	21	52,5
2	7	28	70
3	6	34	85
4	3	37	92,5
5	3	40	100

- $Me = 1$, porque para $x_i = 1$ la F_i supera el 50%
- $Q_1 = 0$, porque F_i supera el 25% para $x_i = 0$
- $Q_3 = 3$, porque F_i supera el 75% para $x_i = 3$

15 □□□ **Halla la mediana, los cuartiles y el percentil 60 en cada una de las siguientes distribuciones, correspondientes a las notas obtenidas en un test que han hecho dos grupos de estudiantes:**

A: 25 – 22 – 27 – 30 – 23 – 22 – 31 – 18
24 – 25 – 32 – 35 – 20 – 28 – 30

B: 27 – 32 – 19 – 22 – 25 – 30 – 21
29 – 23 – 31 – 21 – 20 – 18 – 27

Colocamos en orden creciente los datos:

A 18 - 20 - 22 - 22 - 23 - 24 - 25 - 25 - 27 - 28 - 30 - 30 - 31 - 32 - 35

Hay 15 datos:

- La mediana es el valor central (posición 8) $\rightarrow$ $Me = 25$
- $\frac{15}{4} = 3,75 \rightarrow Q_1 = 22$ (4.ª posición)
- $15 \cdot \frac{3}{4} = 11,25 \rightarrow Q_3 = 30$ (12.ª posición)
- $15 - \frac{60}{100} = 9 \rightarrow$ p_{60} será el valor intermedio de los datos situados en 9.ª y 10.ª posición, es decir:

$$p_{60} = \frac{27 + 28}{2} \rightarrow p_{60} = 27,5$$

B 18 - 19 - 20 - 21 - 21 - 22 - 23 - 25 - 27 - 27 - 29 - 30 - 31 - 32

Hay 14 datos:

- Los dos valores centrales son 23 y 25 $\rightarrow Me = \frac{23 + 25}{2} = 24$
- $\frac{14}{4} = 3,5 \rightarrow Q_1 = 21$ (4.ª posición)
- $14 \cdot \frac{3}{4} = 10,5 \rightarrow Q_3 = 29$ (11.ª posición)
- $14 \cdot \frac{60}{100} = 8,4 \rightarrow p_{60} = 27$ (9.ª posición)

16 **En la fabricación de cierto tipo de bombillas se han detectado algunas defectuosas. Se han estudiado 200 cajas de 100 bombillas cada una, obteniéndose la siguiente tabla:**

DEFECTUOSAS	1	2	3	4	5	6	7	8
N.° DE CAJAS	5	15	38	42	49	31	18	2

Calcula la mediana, los cuartiles y los percentiles p_{10}, p_{90} y p_{95}.

Hacemos la tabla de frecuencias acumuladas.

x_i	f_i	F_i	EN %
1	5	5	2,5
2	15	20	10
3	38	58	29
4	42	100	50
5	49	149	74,5
6	31	180	90
7	18	198	99
8	2	200	100

Para $x_i = 4$, F_i iguala el 50%, luego la mediana será el valor intermedio entre 4 y el siguiente, 5, esto es, $Me = 4,5$.

$Q_1 = p_{25} = 3$

$Q_3 = p_{75} = 6$

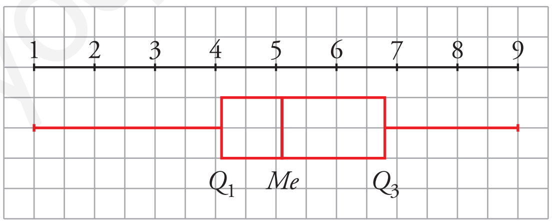
$p_{10} = 2,5$

$p_{90} = 6,5$

$p_{95} = 7$

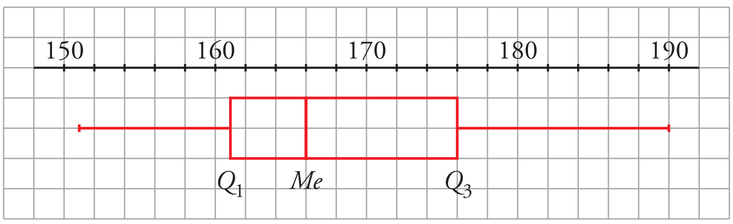
Diagramas de caja

17 **Las puntuaciones obtenidas por 87 personas tienen los siguientes parámetros de posición: $Q_1 = 4,1$; $Me = 5,1$ y $Q_3 = 6,8$. Todas las puntuaciones están en el intervalo 1 a 9. Haz el diagrama de caja.**



18 **Las estaturas de 35 alumnos de una clase están comprendidas entre 153 y 188. Los tres restantes miden 151, 152 y 190. Conocemos los siguientes parámetros: $Q_1 = 161$; $Me = 166$ y $Q_3 = 176$.**

Haz un diagrama de caja para esta distribución.

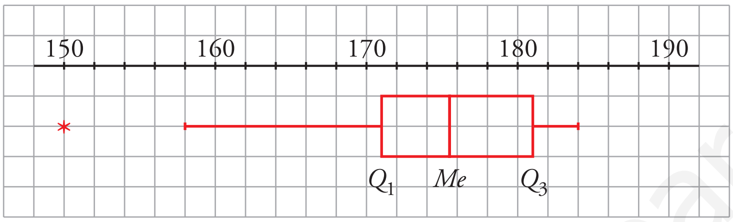


Haz el diagrama de caja correspondiente a las siguientes distribuciones.

19 **La misma que la del ejercicio 13 anterior.**

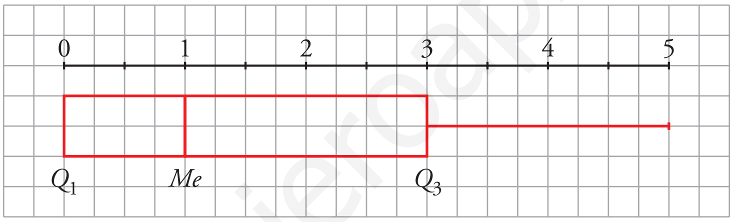
$Q_1 = 171$; $Me = 175{,}5$; $Q_3 = 181$

$(Q_3 - Q_1) \cdot 1{,}5 = (181 - 171) \cdot 1{,}5 = 10 \cdot 1{,}5 = 15 \begin{cases} 171 - 15 = 156 \\ 181 + 15 = 196 \end{cases}$



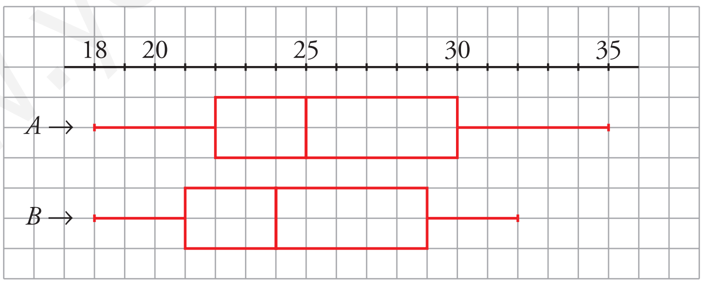
20 **La misma que la del ejercicio 14 anterior.**

$Q_1 = 0$; $Me = 1$; $Q_3 = 3$



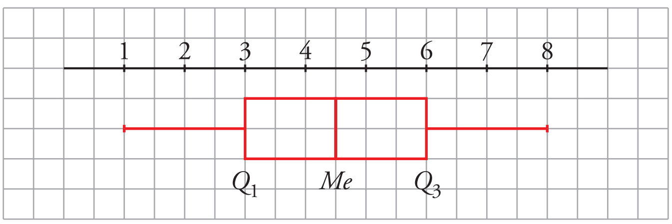
21 **La A y la B que se propusieron en el ejercicio 15 anterior.**

A: $Q_1 = 22$; $Me = 25$; $Q_3 = 30$ B: $Q_1 = 21$; $Me = 24$; $Q_3 = 29$



22 **La misma que la del ejercicio 16 anterior.**

$Q_1 = 3$; $Me = 4{,}5$; $Q_3 = 6$



Muestreo

23 □□□ Se quiere realizar los siguientes estudios:

I. Tipo de transporte que utilizan los vecinos de un barrio para acudir a su trabajo.

II. Estudios que piensan seguir los alumnos y alumnas de un centro escolar al terminar la ESO.

III. Edad de las personas que han visto una obra de teatro en una ciudad.

IV. Número de horas diarias que ven la televisión los niños y niñas de tu comunidad autónoma con edades comprendidas entre 5 y 10 años.

a) Di en cada uno de estos casos cuál es la población.

b) ¿En cuáles de ellos es necesario recurrir a una muestra? ¿Por qué?

a) I → Los vecinos del barrio.

II → Alumnos y alumnas de la ESO de un centro.

III → Personas que han visto la obra.

IV → Niños y niñas de mi comunidad autónoma de entre 5 y 10 años.

b) I → Dependiendo del número de vecinos del barrio: si son pocos, población; si son muchos, una muestra. Aunque teniendo en cuenta que es difícil cogerlos a todos y que todos contesten a la encuesta, quizás sería mejor una muestra.

II → Población. Con encuestas en clase en las que participan todos (obviamente, siempre falta alguno).

III → Muestra. Son muchas personas y sería inoportuno molestar a tanta gente, se formarían colas…

IV → Muestra. Son demasiadas personas.

PÁGINA 203

24 □□□ ¿Cómo se puede contar el número aproximado de palabras que tiene un cierto libro?

— Se seleccionan, abriendo al azar, unas cuantas páginas y se cuentan las palabras en cada una.

— Se calcula el número medio de palabras por página.

— Se da un intervalo en el que pueda estar comprendido el número total de palabras.

Hazlo con algún libro. O si no, imagina que lo has hecho e inventa los resultados.

- En un libro de 200 páginas, seleccionamos al azar 5 páginas. Contamos el número de palabras de estas páginas: 537, 562, 548, 324, 600.

- Calculamos el número medio de palabras:

$$\frac{537 + 562 + 548 + 324 + 600}{5} = 514{,}2$$

En 200 páginas, habrá 102 840 palabras.

- El número de palabras del libro estará entre 100 000 y 105 000.

25 **Para hacer un sondeo electoral en un pueblo de 400 electores, aproximadamente, se va a elegir una muestra de 200 individuos. Di si te parece válido cada uno de los siguientes modos de seleccionarlos y explica por qué.**

a) Se le pregunta al alcalde, que conoce a todo el pueblo, qué individuos le parecen más representativos.

b) Se eligen 200 personas al azar entre las que acuden a la verbena el día del patrón.

c) Se seleccionan al azar en la guía telefónica y se les encuesta por teléfono.

d) Se acude a las listas electorales y se seleccionan al azar 200 de ellos.

a) No es válido. Se trata de una elección subjetiva.

b) No es válido. Probablemente haya grupos de edades mucho más representados que otros.

c) Sí es válido.

d) Sí es válido.

PIENSA Y RESUELVE

26 **Deseamos hacer una tabla con datos agrupados a partir de 384 datos, cuyos valores extremos son 19 y 187.**

a) Si queremos que sean 10 intervalos de amplitud 17, ¿cuáles serán esos intervalos?

b) Haz otra distribución en 12 intervalos de la amplitud que creas conveniente.

Recorrido: $r = 187 - 19 = 168$

a) Buscamos un número algo mayor que el recorrido y que sea múltiplo de 10. Por ejemplo, $r' = 170$. De este modo, cada intervalo tendrá una longitud de 17.

Los intervalos son:

[18, 35); [35, 52); [52, 69); [69, 86); [86, 103); [103, 120)

[120, 137); [137, 154); [154, 171); [171, 188)

b) Buscamos ahora un número que sea múltiplo de 12, que es el número de intervalos en este caso.

$168 = 12 \cdot 14 \quad \rightarrow$ la amplitud de cada intervalo será 14.

Los intervalos son:

[19, 33); [33, 47); [47, 61); [61, 75); [75, 89); [89, 103)

[103, 117); [117, 131); [131, 145); [145, 159); [159, 173); [173, 187)

27 □□□ **En una urbanización de 25 familias se ha observado la variable "número de coches que tiene la familia" y se han obtenido los siguientes datos:**

0	**1**	**2**	**3**	**1**	**0**	**1**	**2**	**3**	**1**
0	**1**	**1**	**1**	**4**	**0**	**1**	**1**	**1**	**4**
3	**2**	**2**	**1**	**1**					

a) Construye la tabla de frecuencias de la distribución.

b) Haz el diagrama de barras.

c) Calcula la media y la desviación típica.

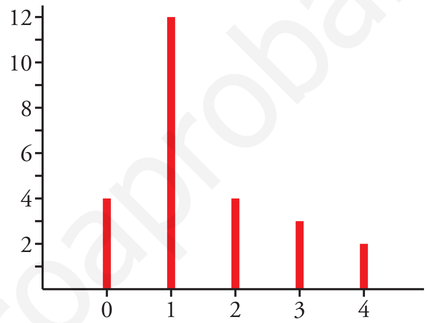
d) Halla la mediana y los cuartiles.

e) Haz el diagrama de caja.

a)

x_i	f_i
0	4
1	12
2	4
3	3
4	2

b)



c)

x_i	f_i	$f_i x_i$	$f_i x_i^2$
0	4	0	0
1	12	12	12
2	4	8	16
3	3	9	27
4	2	8	32
	25	37	87

$$\bar{x} = \frac{37}{25} = 1,48$$

$$\sigma = \sqrt{\frac{87}{25} - 1,48^2} = 1,14$$

d)

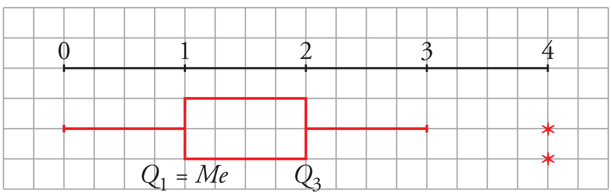
x_i	f_i	F_i	EN %
0	4	4	16
1	12	16	64
2	4	20	80
3	3	23	92
4	2	25	100

$Me = 1$

$Q_1 = 1$

$Q_3 = 2$

e)



28 ■■□ **El número de personas que acudieron cada día a las clases de natación de una piscina municipal fueron:**

38	**31**	**54**	**47**	**50**	**56**	**52**	**48**	**55**	**60**
58	**46**	**47**	**55**	**60**	**53**	**43**	**52**	**46**	**55**
43	**60**	**45**	**48**	**40**	**56**	**54**	**48**	**39**	**50**
53	**59**	**48**	**39**	**48**					

a) Haz una tabla de frecuencias agrupando los datos en intervalos.

b) Representa gráficamente la distribución.

c) Halla $\bar{x}$ y σ.

Localizamos los valores extremos: 31 y 60.

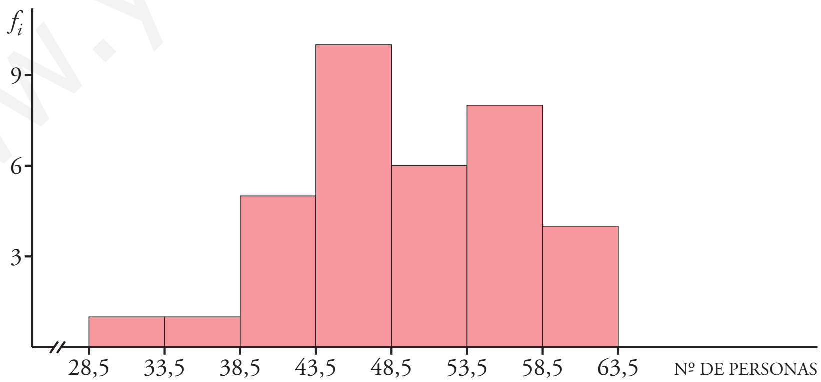
Recorrido = 60 – 31 = 29

Agrupamos los datos en 7 intervalos de longitud 5.

a)

INTERVALOS	x_i	f_i	$f_i x_i$	$f_i x_i^2$
28,5 - 33,5	31	1	31	961
33,5 - 38,5	36	1	36	1 296
38,5 - 43,5	41	5	205	8 405
43,5 - 48,5	46	10	460	21 160
48,5 - 53,5	51	6	306	15 606
53,5 - 58,5	56	8	448	25 088
58,5 - 63,5	61	4	244	14 884
		35	1 730	87 400

b) Representamos los datos en un histograma. La altura de cada rectángulo coincidirá con la frecuencia absoluta, por ser los intervalos de igual amplitud.



c) MEDIA: $\bar{x} = \dfrac{1\,730}{35} \approx 49{,}43$

VAR. $= \dfrac{87\,400}{35} - 49{,}43^2 = 53{,}82$

DESVIACIÓN TÍPICA: $\sigma = \sqrt{53{,}82} \approx 7{,}34$

29 **Un dentista observa el número de caries en cada uno de los 100 niños de un colegio y obtiene los resultados resumidos en esta tabla:**

N.° DE CARIES	F. ABSOLUTA	F. RELATIVA
0	25	0,25
1	20	0,2
2	y	z
3	15	0,15
4	x	0,05

a) Completa la tabla obteniendo x, y, z.

b) Calcula el número medio de caries.

a) La frecuencia relativa es el cociente entre la frecuencia absoluta y el número total de individuos (100, en nuestro caso).

$$0{,}05 = \frac{x}{100} \rightarrow x = 5$$

$$25 + 20 + y + 15 + 5 = 100 \rightarrow y = 35$$

$$z = \frac{y}{100} = \frac{35}{100} \rightarrow z = 0{,}35$$

b)

N.° DE CARIES (x_i)	f_i	$f_i x_i$
0	25	0
1	20	20
2	35	70
3	15	45
4	5	20
	100	155

$$\bar{x} = \frac{155}{100} = 1{,}55$$

El número medio de caries es de 1,55.

30 **El número de errores cometidos en un test por un grupo de personas viene reflejado en la siguiente tabla:**

NÚMERO DE ERRORES	0	1	2	3	4	5	6
NÚMERO DE PERSONAS	10	12	8	7	5	4	3

a) Halla la mediana y los cuartiles inferior y superior, y explica su significado.

b) ¿Cuál es el número medio de errores por persona?

Construimos la tabla de frecuencias acumuladas:

N.° DE ERRORES (x_i)	N.° DE PERSONAS (f_i)	$x_i f_i$	F_i	EN %
0	12	0	12	23,53
1	12	12	24	47,06
2	8	16	32	62,75
3	7	21	39	76,47
4	5	20	44	86,27
5	4	20	48	94,12
6	3	18	51	100
	51	107		

a) $Me = 2$. Significa que el 50% de las personas cometen 0, 1 ó 2 errores.

$Q_1 = 1$. El 25% de las personas comete 1 error o ninguno.

$Q_3 = 3$. El 75% de las personas comente 3 errores o menos de 3 errores.

b) $\bar{x} = \dfrac{\Sigma f_i x_i}{\Sigma f_i} = \dfrac{107}{51} \approx 2,1$

El número medio de errores por persona es ligeramente superior a 2.

31 **Al preguntar a un grupo de personas cuánto tiempo dedicaron a ver televisión durante un fin de semana, se obtuvieron estos resultados:**

TIEMPO EN HORAS	N.° DE PERSONAS
[0; 0,5)	**10**
[0,5; 1,5)	**10**
[1,5; 2,5)	**18**
[2,5; 4)	**12**
[4; 8)	**12**

Dibuja el histograma correspondiente y halla la media y la desviación típica.

Como los intervalos no son de la misma longitud, para representar la distribución mediante un histograma pondremos en cada barra una altura tal que el área sea proporcional a la frecuencia:

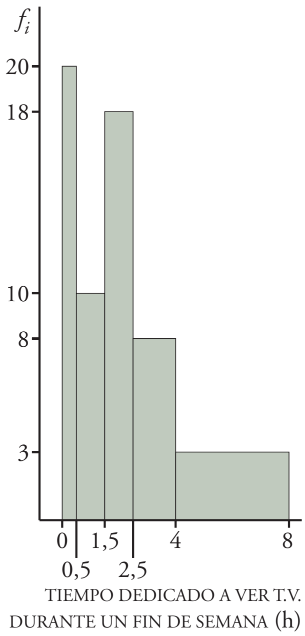
$[0; 0,5) \quad \to \quad a_1 = 0,5 \quad f_1 = 10 \quad \to \quad h_1 = \dfrac{10}{0,5} = 20$

$[0,5; 1,5) \quad \to \quad a_2 = 1 \quad f_2 = 10 \quad \to \quad h_2 = 10$

$[1,5; 2,5) \quad \to \quad a_3 = 1 \quad f_3 = 18 \quad \to \quad h_3 = 18$

$[2,5; 4) \quad \to \quad a_4 = 1,5 \quad f_4 = 12 \quad \to \quad h_4 = \dfrac{12}{1,5} = 8$

$[4; 8) \quad \to \quad a_5 = 4 \quad f_5 = 12 \quad \to \quad h_5 = \dfrac{12}{4} = 3$



TIEMPO	MARCA (x_i)	f_i	$x_i f_i$	$x_i f_i^2$
[0; 0,5)	0,25	10	2,5	0,625
[0,5; 1,5)	1	10	10	10
[1,5; 2,5)	2	18	36	72
[2,5; 4)	3,25	12	39	126,75
[4; 8)	6	12	72	432
		62	159,5	641,375

$$\bar{x} = \frac{159,5}{62} = 2,57$$

$$\sigma^2 = \frac{641,375}{62} - 2,57^2 = 3,74 \rightarrow \sigma = \sqrt{3,74} = 1,93$$

PÁGINA 204

32 **Estas tablas recogen la frecuencia de cada signo en las quinielas durante las 20 primeras jornadas:**

JORNADA	1	X	2
1.ª	4	4	6
2.ª	9	3	2
3.ª	11	2	1
4.ª	10	2	2
5.ª	8	4	2
6.ª	9	4	1
7.ª	10	4	0
8.ª	8	4	2
9.ª	9	5	0
10.ª	5	6	3

JORNADA	1	X	2
11.ª	9	3	2
12.ª	5	6	3
13.ª	7	5	2
14.ª	4	9	1
15.ª	7	3	4
16.ª	6	4	4
17.ª	8	2	4
18.ª	6	7	1
19.ª	7	4	3
20.ª	7	5	2

a) Haz una tabla de frecuencias para el número de veces que sale el "1" en cada una de las 20 jornadas:

x_i	4	5	6	7	8	9	10	11
f_i								

Halla su media y su desviación típica.

b) Haz lo mismo para la "X" y para el "2".

c) Halla el C.V. en los tres casos y compáralos.

a) TABLA DE FRECUENCIAS PARA LOS UNOS

x_i	f_i	$f_i x_i$	$f_i x_i^2$
4	2	8	32
5	2	10	50
6	2	12	72
7	4	28	196
8	3	24	192
9	4	36	324
10	2	20	200
11	1	11	121
	20	149	1 187

$$\bar{x} = \frac{149}{20} = 7,45$$

$$\sigma = \sqrt{\frac{1\,187}{20} - 7,45^2} = \sqrt{3,8475} = 1,96$$

b) TABLA DE FRECUENCIAS PARA LAS EQUIS

x_i	f_i	$f_i x_i$	$f_i x_i^2$
2	3	6	12
3	3	9	27
4	7	28	112
5	3	15	75
6	2	12	72
7	1	7	49
9	1	9	81
	20	86	428

$$\bar{x} = \frac{86}{20} = 4,3$$

$$\sigma = \sqrt{\frac{428}{20} - 4,3^2} = \sqrt{2,91} = 1,71$$

TABLA DE FRECUENCIAS PARA LOS DOSES

x_i	f_i	$f_i x_i$	$f_i x_i^2$
0	2	0	0
1	4	4	4
2	7	14	28
3	3	9	27
4	3	12	48
6	1	6	36
	20	45	143

$$\bar{x} = \frac{45}{20} = 2,25$$

$$\sigma = \sqrt{\frac{143}{20} - 2,25^2} = \sqrt{2,0875} = 1,44$$

c) UNOS $\rightarrow$ $\text{C.V.} = \dfrac{1,96}{7,45} = 0,2631 \rightarrow 26,31\%$

EQUIS $\rightarrow$ $\text{C.V.} = \dfrac{1,71}{4,3} = 0,3977 \rightarrow 39,77\%$

DOSES $\rightarrow$ $\text{C.V.} = \dfrac{1,44}{2,25} = 0,64 \rightarrow 64\%$

33 ■■■ **Cada alumno de un grupo cuenta el número de personas y el número de perros que viven en su portal.**

Suman sus resultados y obtienen una muestra con la que se puede estimar el número de perros que hay en su ciudad.

Por ejemplo, supongamos que en su observación obtienen un total de 747 personas y 93 perros. Y saben que en su ciudad viven 75 000 personas.

a) ¿Cuántos perros estiman que habrá en la ciudad?

b) ¿Cómo es de fiable esta estimación?

c) ¿Es aleatoria la muestra que han utilizado?

a) $\dfrac{93}{747} = \dfrac{x}{75\,000} \rightarrow x = \dfrac{93 \cdot 75\,000}{747} = 9\,337{,}3$

Estiman que habrá unos 9 337 perros, aproximadamente.

b) No es fiable. La muestra estudiada no es representativa de la ciudad.

c) No es aleatoria.

34 ■■■ **Para hacer un estudio sobre los hábitos ecológicos de las familias de una ciudad, se han seleccionado por sorteo las direcciones, calle y número, que serán visitadas.**

Si en un portal vive más de una familia, se sorteará entre ellas la que será seleccionada.

¿Obtendremos con este procedimiento una muestra aleatoria?

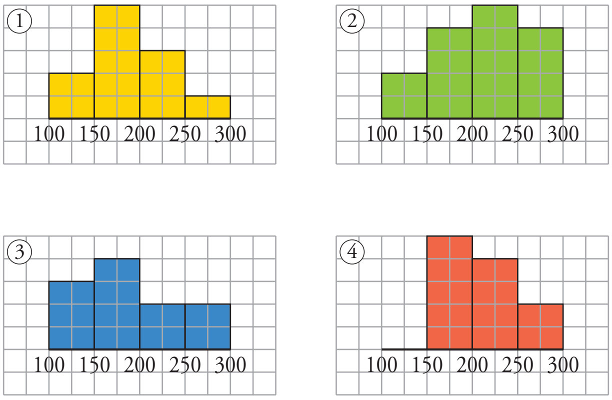
☞ ***Piensa si tiene la misma probabilidad de ser incluida en la muestra una familia que vive en una vivienda unifamiliar que otra que vive, por ejemplo, en un bloque de 32 viviendas.***

No se obtiene una muestra aleatoria, porque una familia que vive en una vivienda unifamiliar tiene más probabilidades de ser elegida que una familia que vive en un bloque de viviendas.

35 ■■■ **Se ha medido el nivel de colesterol en cuatro grupos de personas sometidas a diferentes dietas. Las medias y las desviaciones típicas son las que figuran en esta tabla:**

DIETA	A	B	C	D
$\bar{x}$	211,4	188,6	211,7	188,6
σ	37,5	52,6	49,9	43,1

Las gráficas son, no respectivamente:



Asocia a cada dieta la gráfica que le corresponde.

Fijándonos en las gráficas, se observa que los grupos 1 y 3 tienen una media inferior a 200, mientras que las medias de 2 y 4 son superiores a ese número. Luego podemos asociar:

$$\text{A y C} \;\rightarrow\; \text{2 y 4}$$
$$\text{B y D} \;\rightarrow\; \text{1 y 3}$$

Por otra parte, las personas de 2 tienen el nivel de colesterol más disperso que las de 4. Según esto, su desviación típica será mayor, por lo que C $\leftrightarrow$ 2 y A $\leftrightarrow$ 4. Análogamente, B $\leftrightarrow$ 3 y D $\leftrightarrow$ 1.

PÁGINA 205

REFLEXIONA SOBRE LA TEORÍA

36 **Justifica que la suma de las frecuencias relativas es siempre igual a 1.**

Supongamos que tenemos n datos:

$$fr_1 + fr_2 + \ldots + fr_n = \frac{f_1}{n} + \frac{f_2}{n} + \ldots + \frac{f_n}{n} = \frac{f_1 + f_2 + \ldots + f_n}{n} = \frac{n}{n} = 1$$

Siendo f_i la frecuencia absoluta del dato x_i.

37 **Completa la tabla de esta distribución en la que sabemos que su media es 2,7.**

x_i	1	2	3	4
f_i	3	…	7	5

Llamamos z a la frecuencia absoluta del dato $x_i = 2$.

Aplicamos la definición de la media:

$$\bar{x} = \frac{\Sigma f_i x_i}{\Sigma f_i} \;\rightarrow\; 2{,}7 = \frac{3 + 2z + 21 + 20}{15 + z}$$

$$2{,}7 \cdot (15 + z) = 44 + 2z$$

$$40{,}5 + 2{,}7z = 44 + 2z \;\rightarrow\; 0{,}7z = 3{,}5 \;\rightarrow\; z = 5$$

38 **Si a todos los datos de una distribución le sumamos un mismo número, ¿qué le ocurre a la media? ¿Y a la desviación típica? ¿Y si multiplicamos todos los datos por un mismo número?**

Llamamos a al valor sumado a cada dato de la distribución:

• MEDIA

$$\frac{(x_1 + a)f_1 + (x_2 + a)f_2 + \ldots + (x_n + a)f_n}{n} =$$

$$= \frac{x_1 f_1 + x_2 f_2 + \ldots + x_n f_n + a(f_1 + f_2 + \ldots + f_n)}{n} =$$

$$= \frac{\Sigma f_i x_i}{n} + a\frac{\Sigma f_i}{n} = \bar{x} + a, \text{ puesto que } \frac{\Sigma f_i}{n} = \frac{n}{n} = 1$$

La nueva media es el valor de la media original más el valor que hemos sumado a cada dato.

• DESVIACIÓN TÍPICA:

$$\frac{\Sigma f_i (x_i + a)^2}{\Sigma f_i} - (\bar{x} + a)^2 = \frac{\Sigma f_i x_i^2 + \Sigma f_i a^2 + \Sigma f_i 2x_i a}{\Sigma f_i} - \bar{x}^2 - a^2 - 2\bar{x}a =$$

$$= \frac{\Sigma f_i x_i^2}{\Sigma f_i} + a^2 + 2a\bar{x} - \bar{x}^2 - a^2 - 2\bar{x}a =$$

$$= \frac{\Sigma f_i x_i^2}{\Sigma f_i} - \bar{x}^2$$

La desviación típica no se ve alterada al sumar a todos los datos de la distribución un mismo número.

Supongamos ahora que todos los datos se multiplican por un mismo valor a:

• MEDIA $= \dfrac{ax_1 f_1 + ax_2 f_2 + \ldots + ax_n f_n}{n} = a\bar{x} \;\rightarrow$ la media queda multiplicada por dicho valor.

• DESVIACIÓN TÍPICA:

$$\frac{\Sigma f_i (x_i a)^2}{\Sigma f_i} - (\bar{x}a)^2 = \frac{a^2 \Sigma f_i x_i^2}{\Sigma f_i} - a^2\bar{x}^2 = a^2\left(\frac{\Sigma f_i x_i^2}{\Sigma f_i} - \bar{x}^2\right)$$

La varianza quedaría multiplicada por a^2, luego la desviación típica queda multiplicada por a.

39 **Dos distribuciones estadísticas, A y B, tienen la misma desviación típica.**

a) Si la media de A es mayor que la de B, ¿cuál tiene mayor coeficiente de variación?

b) Si la media de A es el doble que la de B, ¿cómo serán sus coeficientes de variación?

$\text{C.V.} = \dfrac{\sigma}{\bar{x}}$

a) Si $\bar{x}_A > \bar{x}_B \rightarrow \dfrac{1}{\bar{x}_A} < \dfrac{1}{\bar{x}_B} \xrightarrow{\sigma > 0} \dfrac{\sigma}{\bar{x}_A} < \dfrac{\sigma}{\bar{x}_B} \rightarrow$ B tiene mayor coeficiente de variación.

b) Si $\bar{x}_A = 2\bar{x}_B$

$\left.\begin{array}{l} \text{C.V. de } A \rightarrow \dfrac{\sigma}{\bar{x}_A} = \dfrac{\sigma}{2\bar{x}_B} \\ \text{C.V. de } B \rightarrow \dfrac{\sigma}{\bar{x}_B} \end{array}\right\}$ El coeficiente de variación de A es la mitad que el de B.

40 **La validez de la información que nos proporciona una encuesta depende, en gran medida, de la cuidadosa elaboración del cuestionario.**

Algunas características que deben tener las preguntas son:

— Ser cortas y con un lenguaje sencillo.

— Sus esquemas deben presentar opciones no ambiguas y equilibradas.

— Que no requieran esfuerzo de memoria.

— Que no levanten prejuicios en los encuestados.

Estudia si las siguientes preguntas son adecuadas para formar parte de una encuesta y corrige los errorres que observes:

a) ¿Cuánto tiempo sueles estudiar cada día?

☐Mucho ☐Poco ☐Según el día

b) ¿Cuántas veces has ido al cine este año?

c) ¿Qué opinión tienes sobre la gestión del director?

☐Muy buena ☐Buena ☐Indiferente

d) ¿Pierden sus hijos el tiempo viendo la televisión?

☐Sí ☐No

e) ¿En qué grado cree usted que la instalación de la planta de reciclado afectaría al empleo y a las condiciones de salud de nuestra ciudad?

a) La pregunta es muy subjetiva: lo que para un alumno puede ser mucho tiempo de estudio, para otro puede ser poco.

Sería mejor dar opciones con horas determinadas:

☐Menos de 1 hora ☐Entre 1 y 2 horas ☐Más de 2 horas

b) La pregunta requiere un gran esfuerzo de memoria.

Sería mejor preguntar cuántas veces se ha ido al cine en el último mes.

c) Falta la opción de manifestar que la opinión sobre la gestión del director es mala.

d) La pregunta es subjetiva.

Sería mejor preguntar cuántas horas al día ven televisión los hijos de los encuestados:

☐Menos de 1 hora ☐Entre 1 y 2 horas ☐Más de 2 horas

e) La pregunta es demasiado larga y no ofrece opciones claras de respuesta.

Además, mezcla el "empleo" con las "condiciones de salud". Sería mejor hacer dos preguntas separadas dando una graduación adecuada en cada caso.

PROFUNDIZA

41 ■■■ **En una fábrica se ha medido la longitud de 1 000 piezas de las mismas características y se han obtenido los datos que puedes ver en esta tabla.**

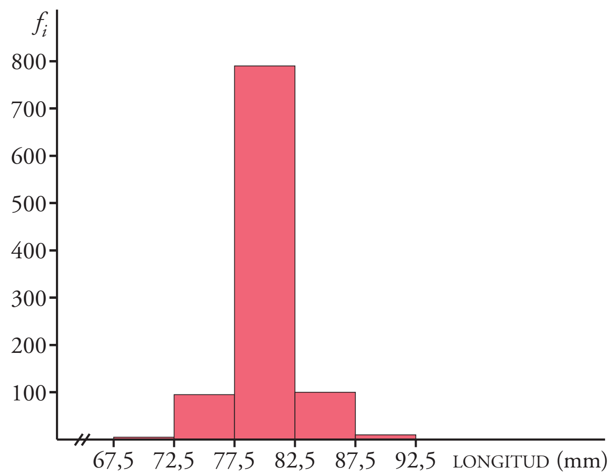
LONGITUD (EN mm)	N.° DE PIEZAS
67,5-72,5	5
72,5-77,5	95
77,5-82,5	790
82,5-87,5	100
87,5-92,5	10

a) Representa el histograma correspondiente.

b) Se consideran aceptables las piezas cuya longitud está en el intervalo [75, 86]. ¿Cuál es el porcentaje de piezas defectuosas?

☞ ***Del 2.° intervalo habrá que rechazar las que midan entre 72,5 y 75. Calcula qué tanto por ciento de la amplitud representa la diferencia 75 – 72,5 y halla el porcentaje de la frecuencia correspondiente. Procede análogamente en el 4.° intervalo.***

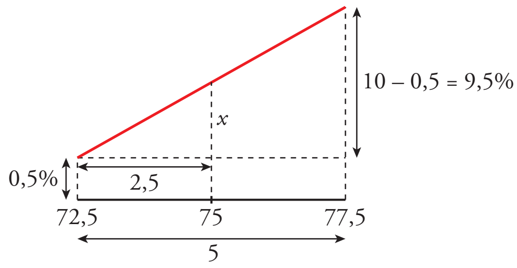
a) Por tener todos los intervalos la misma longitud, la altura de cada una de las barras coincidirá con la frecuencia de cada intervalo.



b) Construimos la tabla de frecuencias absolutas acumuladas:

INTERVALO	f_i	F_i	EN %
67,5 - 72,5	5	5	0,5
72,5 - 77,5	95	100	10
77,5 - 82,5	790	890	89
82,5 - 87,5	100	990	99
87,5 - 92,5	10	1 000	100

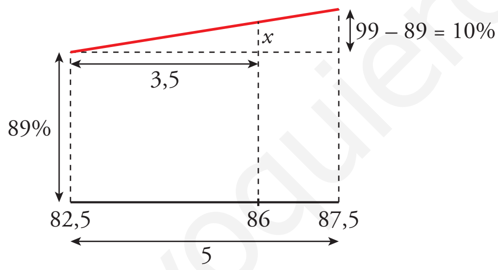
- Calculamos el porcentaje de piezas que hay por debajo de 75 mm:



$$\frac{9,5}{5} = \frac{x}{2,5} \to x = 4,75$$

Por debajo de 75 mm están el $4,75 + 0,5 = 5,25\%$ de las piezas.

- Calculamos el porcentaje de piezas que están por debajo de 86 mm:



$$\frac{10}{5} = \frac{x}{3,5} \to x = 7$$

Por debajo de 86 mm están el $89 + 7 = 96\%$ de las piezas.

El porcentaje de piezas que hay en el intervalo [75, 86] es:

$$96 - 5,25 = 90,75\%$$

Por tanto, el $100 - 90,75 = 9,25\%$ de las piezas serán defectuosas.

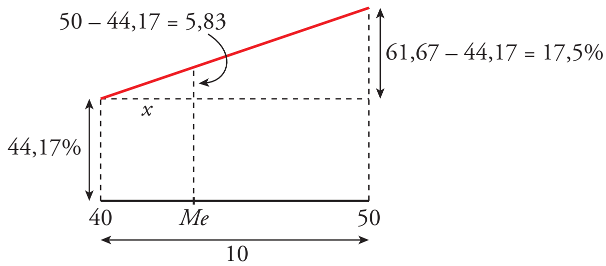
42 **Se ha pasado un test de 80 preguntas a 600 personas. Este es el número de respuestas correctas:**

RESPUESTAS CORRECTAS	[0-10)	[10-20)	[20-30)	[30-40)	[40-50)	[50-60)	[60-70)	[70-80)
NÚMERO DE PERSONAS	40	60	75	90	105	85	80	65

a) Comprueba que la mediana está en el intervalo [40-50). Asígnale un valor repartiendo homogéneamente los 105 individuos que hay en el intervalo.

b) Haz lo mismo para los cuartiles.

a) MEDIANA $\rightarrow$ El 50% se alcanza en el intervalo 40 - 50

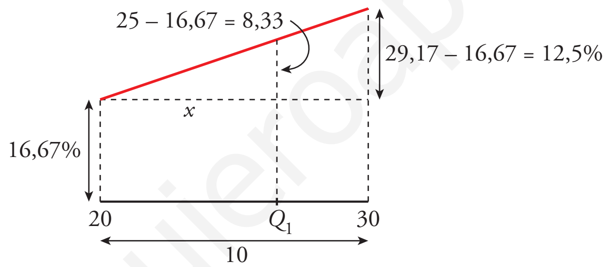


INTERVALO	f_i	F_i	EN %
[0, 10)	40	40	6,67
[10, 20)	60	100	16,67
[20, 30)	75	175	29,17
[30, 40)	90	265	44,17
[40, 50)	105	370	61,67
[50, 60)	85	455	75,83
[60, 70)	80	535	89,17
[70, 80)	65	600	100

$$\frac{10}{17{,}5} = \frac{x}{5{,}83} \rightarrow x = 3{,}33 \rightarrow \text{Luego } Me = 40 + 3{,}33 = 43{,}33$$

b) CUARTILES

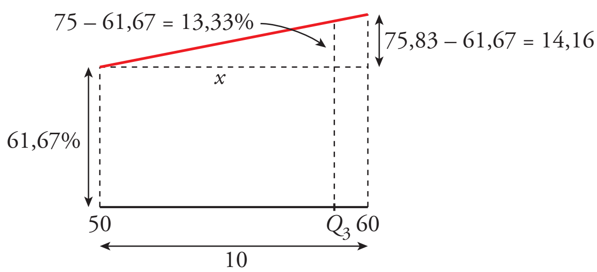
$Q_1 \rightarrow$ El 25% se alcanza en el intervalo 20 - 30.



$$\frac{10}{12{,}5} = \frac{x}{8{,}33} \rightarrow x = 6{,}66$$

$Q_1 = 20 + 6{,}66 = 26{,}66$

$Q_3 \rightarrow$ El 75% se alcanza en el intervalo 50 - 60.



$$\frac{10}{14{,}16} = \frac{10}{13{,}33} \rightarrow x = 9{,}41$$

$Q_3 = 50 + 9{,}41 = 59{,}41$

43 **a) Para estimar la estatura media de los 934 soldados de un regimiento, extraemos una muestra de 53 de ellos. La media de la muestra es 172,6 cm. Expresa este resultado sabiendo que en la ficha técnica se dice que el error máximo es de ±1,8 cm, con una probabilidad de 0,90.**

b) Si con el mismo estudio anterior admitimos que se cometa un error de ±2,6 cm, el nivel de confianza ¿será superior o inferior al 90%?

c) ¿Cómo podríamos aumentar el nivel de confianza manteniendo la cota de error en ±1,8 cm?

a) La altura media de los soldados está en el intervalo (170,8; 174,4) con un nivel de confianza del 90%.

b) El nivel de confianza, al aumentar la longitud del intervalo, también aumenta. Por tanto, será superior al 90%.

c) Tendríamos que aumentar el tamaño de la muestra. Es decir, habría que estudiar a más de 53 soldados.