2 Tablas de frecuencias

Página 173

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 $\frac{30}{10} = 3$.

INTERVALOS	MARCA 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 $\frac{32}{8} = 4$.

INTERVALOS	MARCA 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

3 Parámetros estadísticos: $\bar{x}$ y σ

Página 175

1. Halla, manualmente y con calculadora, $\bar{x}$, σ y C.V. en la tabla obtenida en el ejercicio resuelto de la página 173:

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

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

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

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

2. Halla, manualmente y con calculadora, $\bar{x}$, σ y C.V. en la distribución de los ejercicios 1 y 2 de la página 173:

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.: $\dfrac{\Sigma f_i x_i^2}{\Sigma f_i} - \bar{x}^2 = \dfrac{1\,082\,664}{40} - 164{,}4^2 = 39{,}24$

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

C.V. $= \dfrac{\sigma}{\bar{x}} = \dfrac{6,26}{164,4} = 0{,}038 \;\to\; 3{,}8\%$

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	177,5	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. $= \dfrac{\sigma}{\bar{x}} = \dfrac{6,14}{164,3} = 0{,}037 \;\to\; 3{,}7\%$

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

4 Parámetros de posición

Página 176

1. **Halla Q_1, Me, Q_3 y p_{40} en esta distribución:**

0	**1**	**1**	**2**	**2**	**2**	**3**	**3**	**4**	**4**	**4**	**4**	**5**	**5**	**5**
6	**6**	**7**	**7**	**7**	**7**	**8**	**8**	**8**	**8**	**9**	**9**	**9**	**10**	**10**

Hay 30 individuos en la distribución.

30 : 4 = 7,5 individuos en cada grupo

7,5 $\to$ individuo 8.° $\to$ $Q_1 = 3$

$7{,}5 \cdot 2 = 15$ $\to$ individuo entre 15.° y 16.° $\to$ $Me = 5{,}5$

$7{,}5 \cdot 3 = 22{,}5$ $\to$ individuo 23.° $\to$ $Q_3 = 8$

Para calcular p_{40}:

$30 \cdot \dfrac{40}{100} = 12$ $\to$ individuo entre 12.° y 13° $\to$ $p_{40} = 4{,}5$

Página 177

2. 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	6	19	37	45	109	81	39	22	30	12

NOTAS	f_i	F_i	% ACUM.
1	6	6	1,5
2	19	25	6,25
3	37	62	15,5
4	45	107	26,75
5	109	216	54
6	81	297	74,25
7	39	336	84
8	22	358	89,5
9	30	388	97
10	12	400	100

$Me = p_{50} = 5$

$Q_1 = p_{25} = 4$

$Q_3 = p_{75} = 7$

$p_{80} = 7$

$p_{90} = 9$

$p_{99} = 10$

5 Diagramas de caja

Página 179

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:

$n = 200$

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

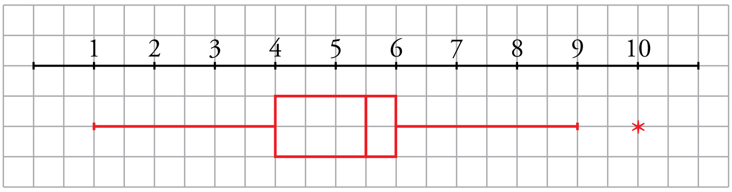
$\frac{n}{4} = 50 \rightarrow Q_1 = 4$

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

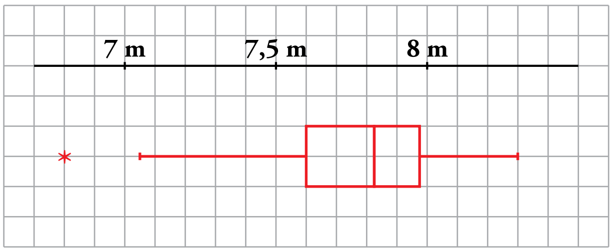
La longitud de la caja será $Q_3 - Q_1 = 6 - 4 = 2$.

$1{,}5 \cdot 2 = 3 \rightarrow$ 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 y bigotes relativo a las marcas de algunos 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.

6 Estadística inferencial

Página 180

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. El responsable de calidad de una empresa que fabrica pilas quiere estudiar la energía suministrada por cada pila hasta que se gasta.

¿Puede hacer el estudio sobre la población o debe recurrir a una muestra? ¿Por qué?

Debe recurrir a una muestra porque el estudio requiere el consumo de las pilas.

3. El dueño de un vivero tiene 285 plantas de interior. Para probar la eficacia de un nuevo fertilizante, las mide todas antes y después del semestre que dura el tratamiento.

El conjunto de esas 285 plantas, ¿es población o muestra? ¿Por qué?

Las 285 plantas sería la población. En este caso, es posible estudiar toda la población, no hace falta trabajar con una muestra.

Ejercicios y problemas

Página 181

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**

Di cuál es la variable y de qué tipo es.

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

• Variable: "Número de faltas de ortografía"

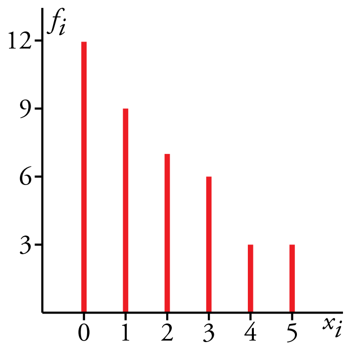
Es una variable cuantitativa discreta.

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

• Tabla de frecuencias:

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

Diagrama de barras:



2. **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**
3,3	**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 desde 1,65 hasta 4,05 y haz una representación adecuada.

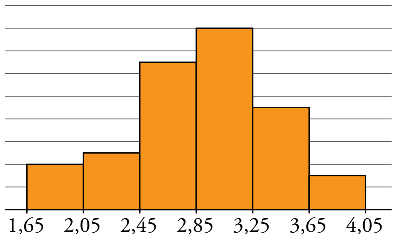
Localizamos los valores extremos: 1,9 y 3,9. Recorrido = 3,9 – 1,8 = 2,1

a) Variable: peso de los recién nacidos.

Tipo: cuantitativa continua.

b) La mejor representación es un histograma:

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	16
3,25-3,65	3,45	9
3,65-4,05	3,85	3
		50



Media, desviación típica y C.V.

3. **Halla la media, la desviación típica y el coeficiente de variación en estas distribuciones:**

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

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**

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

$$\bar{x} = \frac{\Sigma f_i x_i}{\Sigma f_i} = \frac{68}{40} = 1,7$$

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

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

$$\text{C.V.} = \frac{\sigma}{\bar{x}} = 0,9235 \rightarrow 92,35\,\%$$

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

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

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

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

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

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

Parámetros de posición

5. **La altura, en centímetros, de un grupo de estudiantes de una misma clase es:**

150 169 171 172 172 175 181

182 183 177 179 176 184 158

Halla 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:

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

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

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

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

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

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

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

6. **Halla la mediana, los cuartiles y el percentil 60 en cada una de las siguientes distribuciones correspondientes al número de respuestas correctas en un test realizado por 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)
- $15 \cdot \frac{60}{100} = 9 \rightarrow p_{60}$ 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)

7. Rellena la columna de los porcentajes acumulados en la siguiente tabla. Calcula, a partir de la tabla, la mediana, los cuartiles y los percentiles p_{70} y p_{90}.

x_i	f_i	F_i	% ACUM.
0	**12**	**12**	32,4
1	**9**	**21**	56,8
2	**7**	**28**	75,7
3	**6**	**34**	91,9
4	**3**	**37**	100

$Q_1 = 0$

$Me = 1$

$Q_3 = 2$

$p_{70} = 2$

$p_{90} = 3$

8. En la fabricación de cierto tipo de bombillas se han detectado algunas defectuosas. Se analiza el contenido de 200 cajas de 100 bombillas cada una y se obtienen los siguientes resultados:

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	% ACUM.
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, $Me = 4{,}5$.

$Q_1 = p_{25} = 3$

$Q_3 = p_{75} = 6$

$p_{10} = 2{,}5$

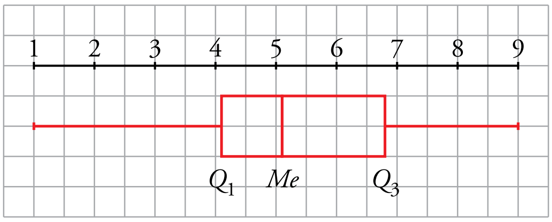
$p_{90} = 6{,}5$

$p_{95} = 7$

Página 182

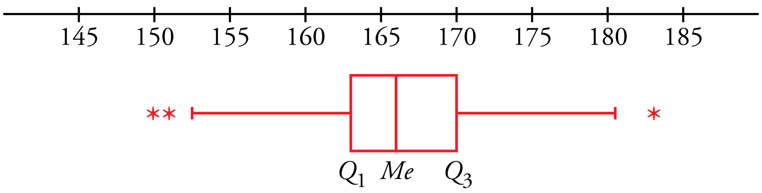
Diagramas de caja

9. **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.**



10. **En una clase de 38 estudiantes de Primaria, las estaturas de 35 de ellos están comprendidas entre 153 cm y 179 cm. Los tres restantes miden 150 cm, 151 cm y 183 cm. Sabemos que $Q_1 = 163$; $Me = 166$ y $Q_3 = 170$.**

Representa los datos en un diagrama de caja.



11. **Haz el diagrama de caja correspondiente a las siguientes distribuciones.**

a) La del ejercicio 5.

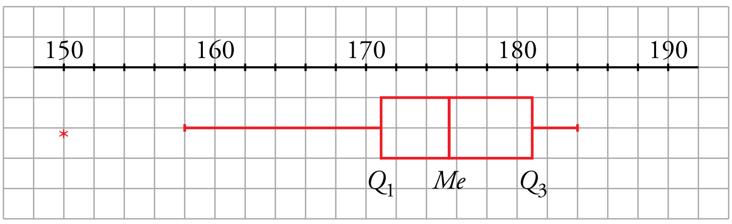
b) La A y la B del ejercicio 6.

c) La del ejercicio 7.

d) La del ejercicio 8.

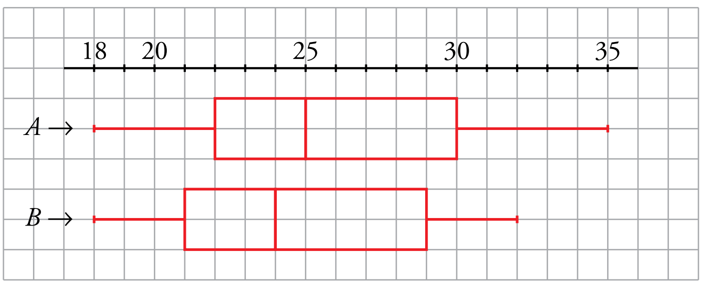
a) $Q_1 = 171$; $Me = 175{,}5$; $Q_3 = 181$

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

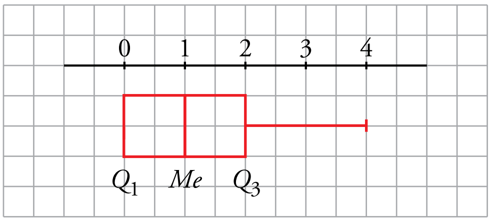


b) A: $Q_1 = 22$; $Me = 25$; $Q_3 = 30$

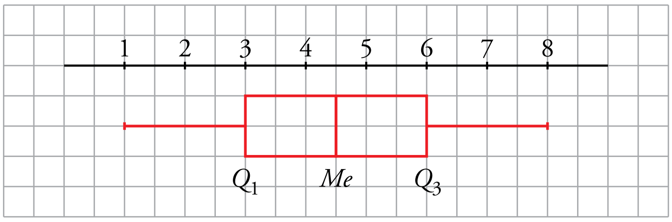
B: $Q_1 = 21$; $Me = 24$; $Q_3 = 29$



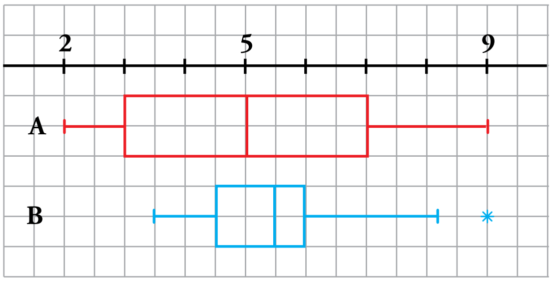
c) $Q_1 = 0$; $Me = 1$; $Q_3 = 2$



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



12. **A los estudiantes de dos clases numerosas de un mismo centro les han puesto un test. Las notas vienen reflejadas en los siguientes diagramas de caja:**



a) ¿Cuál de las clases es más homogénea?

b) ¿En cuál ha aprobado la mitad de la clase?

c) En una de las clases, la tercera nota más alta ha sido un 6,5. ¿De qué clase se trata?

d) ¿En qué clase las notas del 25 % de los estudiantes difieren en medio punto o menos?

e) ¿Cuál es el rango de las notas de cada clase?

a) Es más homogénea la clase B.

b) En la clase A ha aprobado exactamente la mitad de la clase.

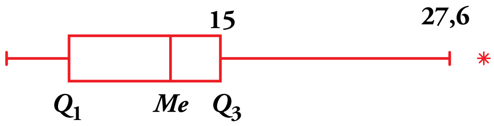
c) En la clase B, puesto que en la clase A el 25 % tiene notas entre 5 y 7.

d) En la clase B.

e) En la clase A el rango es 9 – 2 = 7.

En la clase B el rango es 9 – 3,5 = 5,5.

13. **Calcula el valor del primer cuartil correspondiente al siguiente diagrama de caja:**



27,6 – 15 = 12,6

12,6 : 1,5 = 8,4, por tanto, $Q_3 - Q_1 = 8{,}4 \rightarrow 15 - Q_1 = 8{,}4$.

Luego, $Q_1 = 6{,}6$.

Muestreo

14. **Se quieren realizar estos estudios estadísticos:**

I. Tipo de transporte que utilizan los vecinos de un barrio para acudir a sus trabajos.

II. Estudios que piensan seguir los estudiantes 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 las niñas de tu comunidad autónoma con edades comprendidas entre 5 y 10 años.

V. Tiempo de conversación que aguantan las baterías de los móviles que fabrican en una empresa.

VI. Preferencia de emisora de radio musical de los asistentes a un concierto.

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.

V → Los móviles que fabrica la empresa.

VI → Los asistentes a un concierto.

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.

V → Es necesario recurrir a una muestra para el estudio porque llevarlo a cabo requiere el desgaste de las baterías.

VI → Será necesario recurrir a una muestra.

15. **¿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 alguna novela que encuentres en casa. Cuanto más homogéneas sean sus páginas, más precisión tendrás en el resultado.

- 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:

$$\frac{538 + 562 + 548 + 324 + 600}{5} = 514{,}2$$

En 200 páginas, habrá 102 840 palabras.

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

16. **Para hacer un sondeo electoral en un pueblo de 2 000 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.

Página 183

Aplica lo aprendido

17. **El número de errores cometidos en un test por un grupo de personas viene reflejado en esta tabla:**

N.° DE ERRORES	0	1	2	3	4	5	6
N.° DE PERSONAS	10	12	8	7	5	4	3

a) Halla la mediana, los cuartiles inferior y superior y los percentiles p_{20}, p_{40} y p_{90}. Explica su significado.

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

Completamos la siguiente tabla:

N.° DE ERRORES	0	1	2	3	4	5	6
N.° DE PERSONAS	10	12	8	7	5	4	3
F_i	10	22	30	37	42	46	49
% ACUMULADO	20,4	44,9	61,2	75,5	85,7	93,9	100

a) $p_{20} = 0$ $\quad Q_1 = p_{25} = 1$ $\quad p_{40} = 1$ $\quad Me = p_{50} = 2$ $\quad p_{90} = 5$ $\quad Q_3 = p_{75} = 3$

$p_m = n$ significa que el $m\%$ de las personas comete un máximo de n errores.

b) $\bar{x} = \dfrac{[0 \cdot 10 + 1 \cdot 12 + 2 \cdot 8 + 3 \cdot 7 + 4 \cdot 5 + 5 \cdot 4 + 6 \cdot 3]}{49} = 2,18$ errores por persona.

18. **Deseamos hacer una tabla de datos agrupados a partir de 384 datos, cuyos valores extremos son 19 y 188.**

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 $r = 188 - 19 = 169$

a) Buscamos un número mayor que r que sea múltiplo de 10 $\rightarrow r' = 170$.

Cada intervalo tendrá longitud 17.

Como $r' - r = 1$, comenzamos 0,5 antes del primer dato y finalizamos 0,5 después del último dato.

Los intervalos son:

[18,5; 35,5); [35,5; 52,5); [52,5; 69,5); [69,5; 86,5); [86,5; 103,5);

[103,5; 120,5); [120,5; 137,5); [137,5; 154,5); [154,5; 171,5); [171,5; 188,5)

b) Ahora buscamos un múltiplo de 12 mayor que 169 $\rightarrow r' = 180$.

Como $r' - r = 180 - 169 = 11$, comenzamos 5,5 antes del primer dato y finalizamos 5,5 después del último dato y cada intervalo tendrá amplitud 180 : 12 = 15.

Los intervalos son:

[13,5; 28,5); [28,5; 43,5); [43,5; 58,5); [58,5; 73,5);

[73,5; 88,5); [88,5; 103,5); [103,5; 118,5); [118,5; 133,5);

[133,5; 148,5); [148,5; 163,5); [163,5; 178,5); [178,5; 193,5)

19. **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**	**1**	**1**	**3**	**1**	**2**
0	**1**	**1**	**1**	**4**	**1**	**0**	**1**	**3**	**4**
3	**2**	**2**	**1**	**1**					

a) Construye la tabla de frecuencias.

b) Haz el diagrama de barras.

c) Calcula la media y la desviación típica.

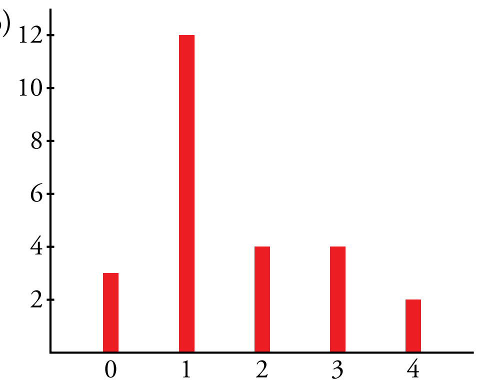
d) Halla la mediana, los cuartiles y los percentiles p_{40} y p_{90}.

e) Dibuja el diagrama de caja.

a)

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

b)



c)

x_i	f_i	$f_i x_i$	$f_i x_i^2$
0	3	0	0
1	12	12	12
2	4	8	16
3	4	12	36
4	2	8	32
	25	40	96

$$\bar{x} = \frac{40}{25} = 1{,}6$$

$$\sigma = \sqrt{\frac{96}{25} - 1{,}6^2} = 1{,}13$$

d)

x_i	f_i	F_i	% ACUM.
0	3	3	12
1	12	15	60
2	4	19	76
3	4	23	92
4	2	25	100

$Q_1 = 1$

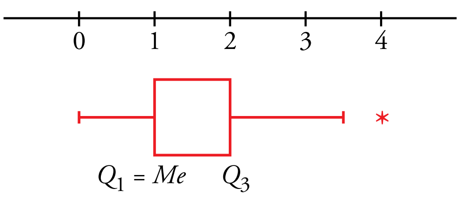
$Me = 1$

$Q_3 = 2$

$p_{40} = 1$

$p_{90} = 3$

e)

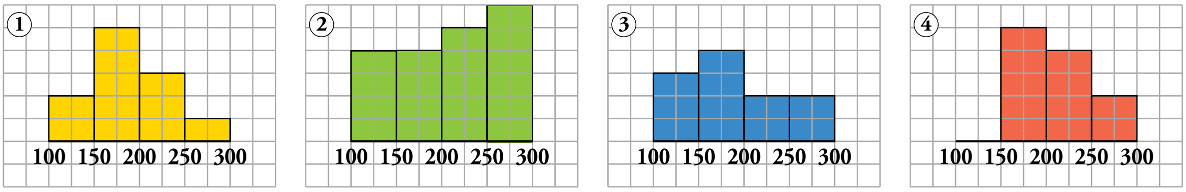


Resuelve problemas

20. **Se ha medido el nivel de colesterol en cuatro grupos de personas sometidas a diferentes dietas. Las medias y las desviaciones típicas son las de la tabla:**

DIETA	A	B	C	D
$\bar{x}$	211,4	188,6	209,2	188,6
σ	37,5	52,6	56,3	43,1

Asocia a cada dieta la gráfica que le corresponde.



Observamos que las gráficas 1 y 3 muestran distribuciones con una media inferior a 200, mientras que las gráficas 2 y 4 muestran distribuciones con una media superior a 200.

Por tanto: A y C $\rightarrow$ 2 y 4; B y D $\rightarrow$ 1 y 3

Por otro lado, los datos en la gráfica 2 están más dispersos que en la gráfica 4, por tanto, la desviación típica es mayor. Así: C $\rightarrow$ 2; A $\rightarrow$ 4

De igual forma, los datos están más dispersos en la gráfica 3 que en la gráfica 1 y, por tanto: B $\rightarrow$ 3; D $\rightarrow$ 1.

21. **En la clase de educación física se ha pedido a cada estudiante que lance 10 veces la pelota de baloncesto desde la línea de personal. Estos resultados son las canastas conseguidas por cada estudiante:**

4 5 7 3 5   2 6 5 4 4   5 8 6 5 7
4 3 5 7 1   2 4 3 6 3   3 5 4 4 2

a) Construye y representa una tabla de frecuencias. Amplía la tabla con las columnas necesarias para hallar la media y la desviación típica. Calcula también el coeficiente de variación.

b) Construye la tabla de frecuencias acumuladas y de porcentajes acumulados y, a partir de ella, halla Q_1, Me, Q_3, p_{30}, p_{90} y p_{99}.

c) Representa los datos en un diagrama de caja.

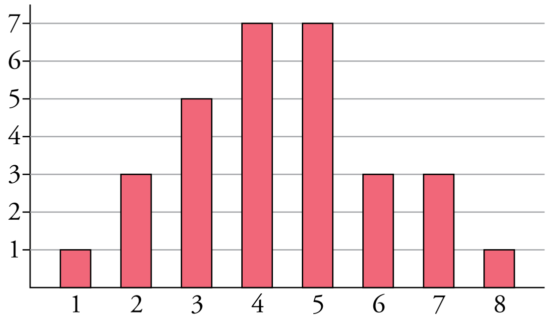
a) En la clase hay 30 estudiantes:

x_i	f_i	$f_i x_i$	$f_i x_i^2$
1	1	1	1
2	3	6	12
3	5	15	45
4	7	28	112
5	7	35	175
6	3	18	108
7	3	21	147
8	1	8	64
		132	664

$$\bar{x} = \frac{\Sigma f_i x_i}{\Sigma f_i} = \frac{132}{30} = 4,4$$

$$\sigma = \sqrt{\frac{\Sigma f_i x_i^2}{\Sigma f_i} - \bar{x}^2} = \sqrt{\frac{664}{30} - 4,4^2} = \sqrt{2,77} = 1,66$$

$$\text{C.V.} = \frac{\sigma}{\bar{x}} = \frac{1,66}{4,4} = 0,38$$



b)

x_i	f_i	F_i	% ACUM.
1	1	1	3,33
2	3	4	13,33
3	5	9	30
4	7	16	53,33
5	7	23	76,66
6	3	26	86,66
7	3	29	96,66
8	1	30	100

$Q_1 = 3$

$Me = 4$

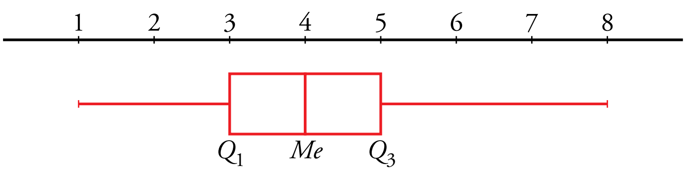
$Q_3 = 5$

$p_{30} = 3,5$

$p_{90} = 7$

$p_{99} = 7$

c)



Curiosidades matemáticas

¿Sabías que…?

Los teclados de los ordenadores tienen, todos, la misma distribución de los caracteres; cada letra, número o signo tiene su lugar, fijo. Esa distribución, heredada de las antiguas máquinas de escribir, fue ideada por Christopher Sholes (Inglaterra, 1867), basándose en un estudio estadístico sobre la frecuencia de aparición de cada letra en la lengua inglesa. Puso las más frecuentes "más a mano".

- **Si metieras en un bombo todas las letras de las dos líneas que estás leyendo y sacaras una al azar, ¿cuál de ellas tendría mayor probabilidad de ser elegida?**
- **¿Qué letra es la más usada en castellano? Diseña un proyecto para averiguarlo.**

- A simple vista vemos que las letras que más aparecen son la "a" y la "e". Respectivamente aparecen 21 y 17 veces.
- Respuesta abierta.