

EJERCICIOS PROPUESTOS

14.1 Clasifica los siguientes caracteres estadísticos.

- a) Número de canastas encestadas en un partido de baloncesto.
- b) Canal de televisión preferido por los vecinos de una casa.
- c) Medida, en metros, del salto de longitud en unos juegos olímpicos.

- a) Cuantitativo discreto
- b) Cualitativo
- c) Cuantitativo continuo

14.2 En una ciudad hay tres millones de personas con derecho a voto, de las que el 53 % son mujeres. Se quiere elegir una muestra constituida por 3000 personas.

¿Cuántos hombres y mujeres deberán formar parte de la muestra para que sea representativa de la población?

Para que la muestra guarde la misma proporción de hombres y mujeres que en la población, se deberá elegir:

$$\frac{53 \cdot 3000}{100} = 1590 \text{ mujeres}$$

$$\frac{47 \cdot 3000}{100} = 1410 \text{ hombres}$$

14.3 La tabla adjunta muestra el número de faltas de asistencia en una clase a lo largo de un mes.

N.º de faltas	0	1	2	3	4	5
N.º de alumnos	10	7	6	2	1	4

Calcula la media aritmética y la moda.

$$\text{Media aritmética: } \bar{x} = \frac{0 \cdot 10 + 1 \cdot 7 + 2 \cdot 6 + 3 \cdot 2 + 4 \cdot 1 + 5 \cdot 4}{10 + 7 + 6 + 2 + 1 + 4} = \frac{49}{30} = 1,6 \text{ faltas}$$

$$\text{Moda} = M_o = 0 \text{ faltas}$$

14.4 La siguiente tabla muestra los resultados de unos alumnos en la prueba de salto de longitud.

Salto (m)	[2; 2,5)	[2,5; 3)	[3; 3,5)	[3,5; 4)
N.º de alumnos	6	12	15	4

Halla la media aritmética y la moda.

$$\text{Media aritmética: } \bar{x} = \frac{2,25 \cdot 6 + 2,75 \cdot 12 + 3,25 \cdot 15 + 3,75 \cdot 4}{6 + 12 + 15 + 4} = \frac{110,25}{37} = 2,98 \text{ m}$$

Moda: La clase modal es [3; 3,5), y tomaremos como valor aproximado de la moda la marca de la clase modal.

$$\text{Así, } M_o = 3,25 \text{ m.}$$

14.5 Calcula la mediana y los cuartiles de la distribución estadística dada por esta tabla.

x_i	2	3	4	5	6
f_i	11	17	23	24	15

Primer cuartil. El número de datos es 90. La cuarta parte es 22,5. El primer cuartil es 3, ya que es el valor de la variable cuya frecuencia absoluta acumulada excede por primera vez la cuarta parte del número de datos. Por tanto, $Q_1 = 3$.

x_i	f_i	F_i
2	11	$11 < 22,5$
3	17	$28 > 22,5$
4	23	51
5	24	75
6	15	90

Mediana. El número de datos es 90. La mitad es 45. La mediana es 4, ya que es el valor de la variable cuya frecuencia absoluta acumulada excede por primera vez la mitad del número de datos. Por tanto, $M = 4$.

x_i	f_i	F_i
2	11	11
3	17	$28 < 45$
4	23	$51 > 45$
5	24	75
6	15	90

Tercer cuartil. Las tres cuartas partes del número de datos son 67,5. El tercer cuartil es 5, ya que es el valor de la variable cuya frecuencia absoluta acumulada excede por primera vez las tres cuartas partes del número de datos. Por tanto, $Q_3 = 5$.

x_i	f_i	F_i
2	11	11
3	17	28
4	23	$51 < 67,5$
5	24	$75 > 67,5$
6	15	90

14.6 Calcula la mediana, Q_1 y Q_3 de la siguiente distribución.

x_i	[0, 10)	[10, 20)	[20, 30)	[30, 40)
f_i	12	16	17	11

Primer cuartil. El número de datos es 56. La cuarta parte del número de datos es 14. La clase que contiene al primer cuartil es [0, 10), ya que su frecuencia absoluta acumulada excede por primera vez la cuarta parte del número de datos. Tomaremos como aproximación al primer cuartil la marca de la clase [0, 10). Por tanto, $Q_1 = 5$.

x_i	f_i	F_i
[0, 10)	12	$12 < 14$
[10, 20)	16	$28 > 14$
[20, 30)	17	45
[30, 40)	11	56

Mediana. El número de datos es 56. La mitad del número de datos es 28. En este caso, la mediana será la media entre las marcas de la clase [10, 20) y [20, 30). Por tanto, $M = 20$.

x_i	f_i	F_i
[0, 10)	12	12
[10, 20)	16	28
[20, 30)	17	45
[30, 40)	11	56

Tercer cuartil. Las tres cuartas partes del número de datos son 42. La clase que contiene al tercer cuartil es [20, 30), ya que su frecuencia absoluta acumulada excede por primera vez las tres cuartas partes del número de datos. Tomaremos como aproximación al tercer cuartil la marca de la clase [20, 30). Por tanto, $Q_3 = 25$.

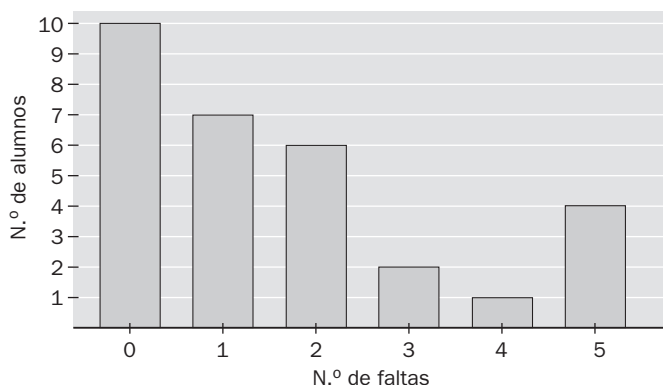
x_i	f_i	F_i
[0, 10)	12	12
[10, 20)	16	$28 < 42$
[20, 30)	17	$45 > 42$
[30, 40)	11	56

14.7 La tabla adjunta muestra el número de faltas en una clase a lo largo de un mes.

N.º de faltas	0	1	2	3	4	5
N.º de alumnos	10	7	6	2	1	4

Representa los datos gráficamente.

Representamos los datos mediante el siguiente diagrama de barras.



14.8 La siguiente tabla muestra los resultados de unos alumnos en la prueba de salto de longitud.

Salto (m)	[2; 2,5)	[2,5; 3)	[3; 3,5)	[3,5; 4)
N.º de alumnos	10	9	11	7

Dibuja el diagrama de cajas y bigotes.

L_i : 2,25

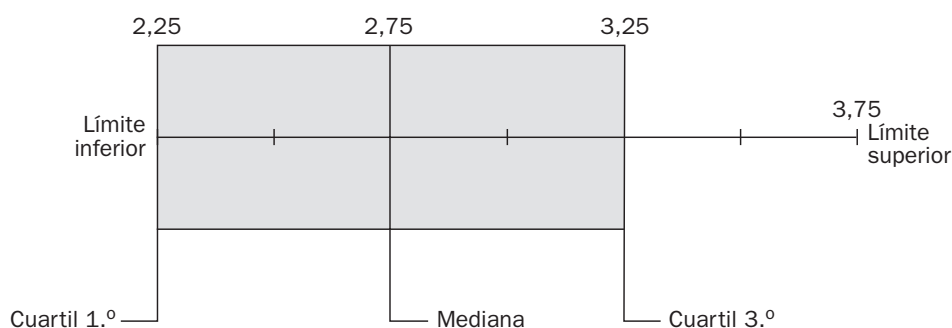
Q_1 : La clase que contiene el primer cuartil es [2; 2,5), ya que su frecuencia absoluta acumulada excede por primera vez la cuarta parte de los datos, $\frac{37}{4} = 9,25$. El primer cuartil será la marca de clase, es decir, $Q_1 = 2,25$.

M : La clase que contiene la mediana es [2,5; 3), ya que su frecuencia absoluta acumulada excede por primera vez la mitad de los datos, $\frac{37}{2} = 18,5$. La mediana será la marca de clase, es decir, $M = 2,75$.

Q_3 : La clase que contiene el tercer cuartil es [3; 3,5), ya que su frecuencia absoluta acumulada excede por primera vez las tres cuartas partes de los datos, $\frac{111}{4} = 27,75$. El tercer cuartil será la marca de clase, es decir, $Q_3 = 3,25$.

L_s : 3,75

Con estos datos, el diagrama de cajas y bigotes es el siguiente:



14.9 La siguiente tabla muestra el número de faltas en una clase a lo largo de un mes.

N.º de faltas	0	1	2	3	4	5
N.º de alumnos	10	7	6	2	1	4

a) Calcula el rango.

b) Halla la varianza y la desviación típica.

Formamos la siguiente tabla.

x_i	f_i	x_i^2	$f_i x_i^2$
0	10	0	0
1	7	1	7
2	6	4	24
3	2	9	18
4	1	16	16
5	4	25	100
	30		165

a) Rango = $5 - 0 = 5$

b) Varianza: $s^2 = \frac{165}{30} - \left(\frac{49}{39}\right)^2 = 2,83 \text{ faltas}^2$

Desviación típica: $s = \sqrt{2,83} = 1,68 \text{ faltas}$

14.10 La tabla adjunta representa los resultados de unos alumnos en la prueba de salto de longitud.

Salto (m)	[2; 2,5)	[2,5; 3)	[3; 3,5)	[3,5; 4)
N.º de alumnos	6	12	15	4

a) Halla el rango.

b) Calcula la varianza y la desviación típica.

Formamos la siguiente tabla.

x_i	f_i	x_i^2	$f_i x_i^2$
2,25	6	5,0625	30,375
2,75	12	7,5625	90,75
3,25	15	10,5625	158,4375
3,75	4	14,0625	56,25
			335,8125

a) Rango = $3,75 - 2,25 = 1,5$

b) Varianza: $s^2 = \frac{335,8125}{37} - 2,9797^2 = 0,197 \text{ m}^2$

Desviación típica: $s = \sqrt{0,197} = 0,444 \text{ m}$

14.11 Una distribución viene dada por la siguiente tabla.

x_i	[10, 20)	[20, 30)	[30, 40)	[40, 50)	[50, 60)
f_i	5	12	20	11	6

Halla el porcentaje de datos incluidos en los intervalos $(\bar{x} - s, \bar{x} + s)$, $(\bar{x} - 2s, \bar{x} + 2s)$ y $(\bar{x} - 3s, \bar{x} + 3s)$.

$$\bar{x} = \frac{1900}{54} = 35,185$$

$$s = \sqrt{\frac{73550}{54} - 35,185^2} = 11,137$$

$$(\bar{x} - s, \bar{x} + s) = (24,048; 46,322).$$

Si suponemos que los datos se distribuyen de forma uniforme, tenemos, por una simple proporcionalidad, que si de 20 a 30 se distribuyen 12 datos, de 24 a 30 se distribuirán: $\frac{6 \cdot 12}{10} \approx 7$ datos.

Del mismo modo, si de 40 a 50 se distribuyen 11 datos, de 40 a 46 se distribuirán: $\frac{6 \cdot 11}{10} \approx 7$ datos.

Luego en el intervalo $(\bar{x} - s, \bar{x} + s)$ se distribuyen: $7 + 20 + 7 = 34$ datos. 34 de 54 datos totales representan el 63 %.

Razonando de forma análoga se tiene $(\bar{x} - 2s, \bar{x} + 2s) = (12,911; 57,459)$.

En este intervalo se distribuyen $4 + 12 + 20 + 11 + 4 = 51$ datos. 51 de 54 datos totales representan el 94,4 %.

$$(\bar{x} - 3s, \bar{x} + 3s) = (1,774; 68,596)$$

En este intervalo se distribuye el 100 %.

14.12 Los porcentajes de uso del cinturón de seguridad en dos ciudades durante 4 días son:

A	87	78	67	82
B	60	95	92	47

Calcula el coeficiente de variación en cada ciudad e interpreta el resultado.

$$\text{Ciudad A: } \bar{x} = \frac{87 + 78 + 67 + 82}{4} = 78,5$$

$$s = \sqrt{\frac{24\,866}{4} - (78,5)^2} = 7,36$$

$$\text{Luego } CV_A = \frac{s}{\bar{x}} = \frac{7,36}{78,5} = 0,094$$

$$\text{Ciudad B: } \bar{x} = \frac{60 + 95 + 92 + 47}{4} = 73,5$$

$$s = \sqrt{\frac{23\,298}{4} - (73,5)^2} = 20,55$$

$$\text{Luego } CV_B = \frac{s}{\bar{x}} = \frac{20,55}{73,5} = 0,28$$

Como se observa, el coeficiente de variación de la ciudad A es menor que el de la ciudad B. Esto quiere decir que el comportamiento de los ciudadanos de A con respecto al uso del cinturón de seguridad es más homogéneo que el de los ciudadanos de B.

- 14.13** La media y la mediana de un conjunto de cinco números naturales distintos es 7, y el rango es 6. Halla dichos números.

Sean a, b, c, d y e , ordenados de menor a mayor, los cinco números naturales distintos:

Como la media es 7 $\Rightarrow a + b + c + d + e = 35$.

Como la mediana es 7 $\Rightarrow c = 7$.

Como el rango es 6 $\Rightarrow e - a = 6$.

Mediante la estrategia de ensayo – error se tiene que las posibles soluciones son:

Solución 1: 4, 5, 7, 9 y 10

Solución 2: 4, 6, 7, 8 y 10

- 14.14** Siete números naturales, no necesariamente distintos, tienen de media 8 y de mediana 11. ¿Cuál puede ser el mayor valor del rango? ¿Y el menor?

Sean los números $n_1 \leq n_2 \leq n_3 \leq n_4 \leq n_5 \leq n_6 \leq n_7$.

Como la mediana es 11, se deduce que $n_4 = 11$.

Como la media es 8, la suma de los 7 números es $8 \cdot 7 = 56$.

Para corregir el mayor rango hacemos:

$n_1 = n_2 = n_3 = 1$ y $n_5 = n_6 = 11$; entonces, $n_7 = 56 - 3 - 33 = 20$. En este caso, el rango es $20 - 1 = 19$.

Para corregir el menor rango hacemos:

$n_4 = n_5 = n_6 = n_7 = 11$; entonces, $n_1 + n_2 + n_3 = 56 - 4 \cdot 11 = 12$, y como queremos el rango menor, $n_1 = n_2 = n_3 = 4$. En este caso, el rango es $11 - 4 = 7$.

ACTIVIDADES

EJERCICIOS PARA ENTRENARSE

Caracteres y parámetros estadísticos. Gráficos

- 14.15** Indica si los siguientes caracteres son cualitativos o cuantitativos, y, en su caso, expresa si la variable estadística es discreta o continua.

a) Número de faltas de asistencia de los alumnos de una clase de 4.º de ESO en un mes.

b) Número de horas de productividad entre los trabajadores de una oficina.

c) Número de móviles que poseen los miembros de las familias de un edificio.

a) Cuantitativa; variable discreta

b) Cuantitativa; variable continua

c) Cuantitativa; variable discreta

- 14.16** Completa los valores de la tabla.

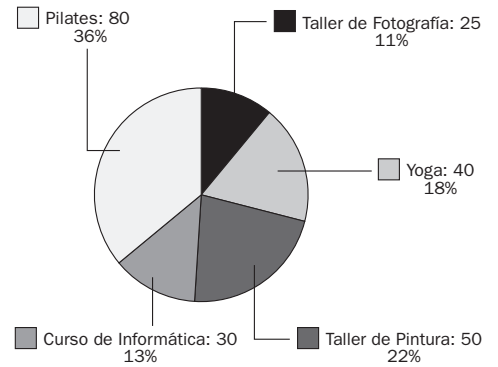
x_i	f_i	F_i	$x_i f_i$
1	5		
3			21
4		14	
	8		56
8		23	
	4		
11			33
Total			171

\Rightarrow

x_i	f_i	F_i	$x_i f_i$
1	5	5	5
3	7	12	21
4	2	14	8
7	8	22	56
8	1	23	8
10	4	27	40
11	3	30	33
Total	33		171

14.17 La siguiente tabla muestra las actividades ofertadas por un centro cultural y el número de vecinos del barrio que cursan dichas actividades.

Actividades	Número de vecinos
Yoga	40
Taller de pintura	50
Curso de informática	30
Pilates	80
Taller de fotografía	25



Representa el diagrama de sectores asociado a la tabla anterior.

14.18 Se realiza una encuesta a un grupo de 20 personas acerca del número de veces que acuden al cine a lo largo de un año, obteniéndose los siguientes resultados.

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

- Agrupar los datos en una tabla.
- Halla la media, la moda, la mediana y el primer cuartil.
- Calcula el rango, la varianza y la desviación típica.
- Representa el diagrama de barras y el polígono de frecuencias de los datos.

a)

N.º películas x_i	N.º personas f_i	$x_i \cdot f_i$	$x_i^2 \cdot f_i$
1	2	2	2
2	3	6	12
3	4	12	36
4	4	16	64
5	3	15	75
6	1	6	36
7	2	14	98
8	1	8	64
	20	79	387

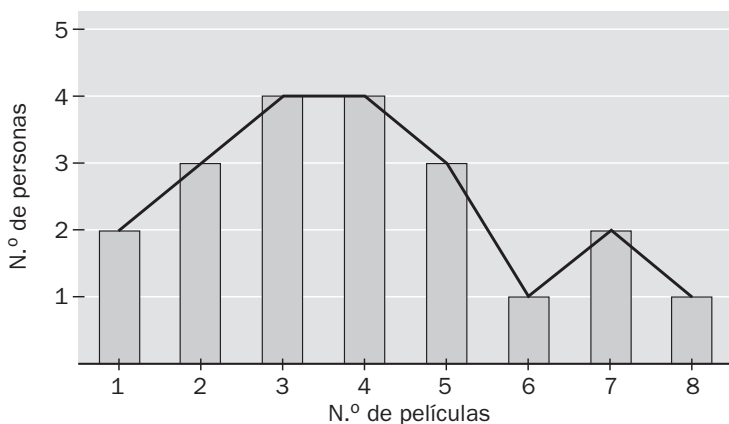
b) $\bar{x} = \frac{79}{20} = 3,95$

Es una distribución bimodal: $M_o = 3$; $M_o = 4$

$M = 4$ y $Q_1 = 3$

c) Rango = $8 - 1 = 7$; $s^2 = \frac{387}{20} - 3,95^2 = 3,7475$; $s = \sqrt{3,7475} \approx 1,93$

d)



14.19 La siguiente tabla presenta el número de horas semanales que dedican al estudio los 30 alumnos de una clase de 4.º de ESO.

N.º de horas	N.º de alumnos
[0, 4)	8
[4, 8)	10
[8, 12)	8
[12, 16)	4

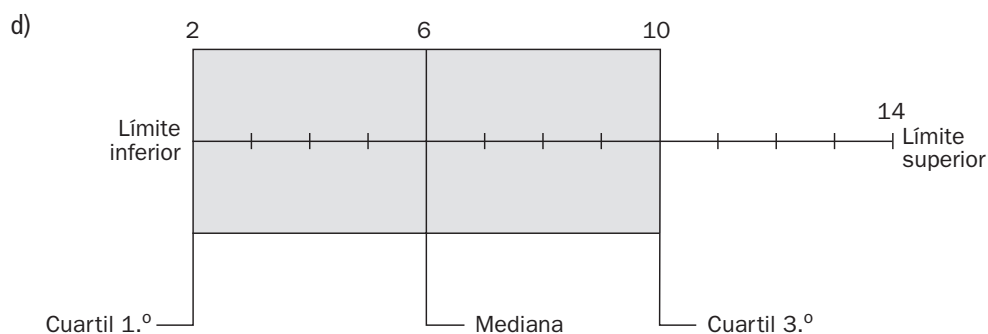
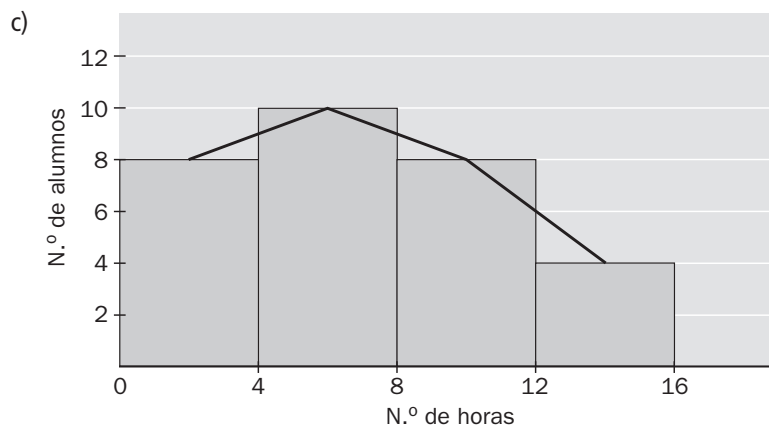
- Halla la media, la moda, la mediana y los otros dos cuartiles.
- Calcula el rango, la varianza y la desviación típica.
- Representa el histograma y el polígono de frecuencias.
- Dibuja el diagrama de cajas y bigotes.

a)

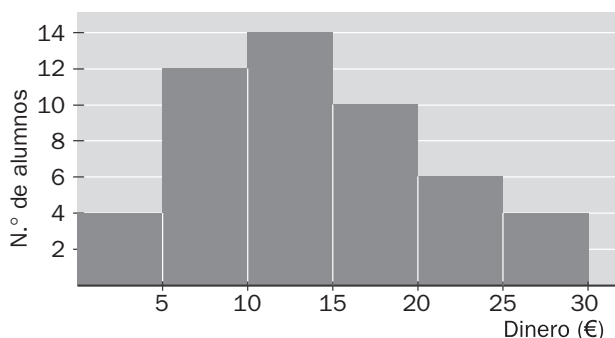
Marcas (x_i)	f_i	F_i	$x_i \cdot f_i$	$x_i^2 \cdot f_i$
2	8	8	16	32
6	10	18	60	360
10	8	26	80	800
14	4	30	56	784
	30		212	1976

$$\bar{x} = \frac{212}{30} = 7,07; M_o = 6; M = 6; Q_1 = 2; Q_3 = 10$$

b) Rango = $14 - 2 = 12$; $s^2 = \frac{1976}{30} - 7,07^2 \approx 15,88$; $s = \sqrt{15,88} \approx 3,98$



14.20 El siguiente histograma representa el dinero gastado en telefonía móvil en un mes por un grupo de 50 estudiantes de Bachillerato.



Calcula el gasto medio en móvil del grupo de estudiantes y la desviación típica.

La tabla asociada al histograma es la siguiente.

Dinero (€)	N.º alumnos f_i	Marcas x_i	$x_i \cdot f_i$	$x_i^2 \cdot f_i$
[0, 5)	4	2,5	10	25
[5, 10)	12	7,5	90	675
[10, 15)	14	12,5	175	2187,5
[15, 20)	10	17,5	175	3062,5
[20, 25)	6	22,5	135	3037,5
[25, 30)	4	27,5	110	3025
	50		695	12 012,5

El gasto medio es: $\bar{x} = \frac{695}{50} = 13,9 \text{ €}$

Desviación típica: $s = \sqrt{\frac{12\,012,5}{50} - 13,9^2} \approx 6,86$

14.21 De una distribución estadística compuesta por cinco datos, se sabe que su moda es 4, su media es 6 y su mediana es 5. Halla una distribución de cinco elementos con estas características.

Una distribución con esas características es: 4, 4, 5, 8, 9.

Utilización conjunta de \bar{x} y s

14.22 Se ha realizado un estudio con el fin de averiguar la cantidad de papel reciclado en toneladas de los distintos distritos y se han obtenido los siguientes resultados. 64, 65, 68, 67, 68, 67, 72, 74, 80, 74, 68, 74, 68, 72, 68, 65, 72, 67, 68, 85.

a) Halla la media y la desviación típica.

b) Calcula el porcentaje de distritos cuyas cantidades recicladas se encuentran dentro del intervalo $(\bar{x} - 2s, \bar{x} + 2s)$

a)

x_i	f_i	$x_i \cdot f_i$	$x_i^2 \cdot f_i$
64	1	64	4096
65	2	130	8450
67	3	201	13 467
68	6	408	27 744
72	3	216	15 552
74	3	222	16 428
80	1	80	6400
85	1	85	7225
	20	1406	99 362

$$\bar{x} = \frac{1406}{20} = 70,3; s = \sqrt{\frac{99\,362}{20} - 70,3^2} \approx 5,1$$

b) $(\bar{x} - 2s, \bar{x} + 2s) = (60,1; 80,5)$. Luego los distritos que se encuentran en ese intervalo son 19, lo que supone un porcentaje del 95 %.

14.23 Se han medido las temperaturas máximas alcanzadas en dos ciudades durante 10 días consecutivos del mes de agosto, obteniéndose los siguientes resultados.

A	32	33	24	22	35	30	29	31	20	19
B	27	28	25	31	24	25	24	26	22	28

a) ¿Los habitantes de qué ciudad han tenido una temperatura media más alta a lo largo de esos 10 días?

b) ¿Qué ciudad ha sufrido una variabilidad de temperatura mayor?

c) ¿Qué parámetro has empleado para contestar el anterior apartado? ¿Por qué?

a) Ya que $\bar{x}_A = \frac{275}{10} = 27,5^\circ$ y $\bar{x}_B = \frac{260}{10} = 26^\circ$, la ciudad A ha tenido una temperatura media más alta.

b) Para ver cuál ha sufrido una variabilidad mayor, debemos calcular el coeficiente de variación de ambas ciudades:

$$s_A^2 = \frac{7861}{10} - 27,5^2 = 29,85 \Rightarrow s_A \approx 5,46 \Rightarrow CV_A = \frac{5,46}{27,5} \approx 0,198$$

$$s_B^2 = \frac{6820}{10} - 26^2 = 6 \Rightarrow s_B \approx 2,45 \Rightarrow CV_B = \frac{2,45}{26} \approx 0,09$$

Por tanto, la ciudad A ha tenido una variabilidad de temperatura mayor.

c) El parámetro empleado es el CV, y se utiliza porque las medias y las desviaciones típicas de las dos distribuciones son distintas, y es la única manera de comparar sus dispersiones.

14.24 Las notas obtenidas en la asignatura de Matemáticas por los alumnos de dos clases de 4.º de ESO son las siguientes.

Notas	0	1	2	3	4	5	6	7	8	9	10
4.º A	5	4	1	0	0	0	0	0	1	4	5
4.º B	0	0	2	2	3	6	3	2	2	0	0

- a) ¿Cuál es la calificación media de cada una de las dos clases?
b) ¿Cuál de ellas tiene las notas menos dispersas?
c) ¿Es necesario calcular el coeficiente de variación para poder determinarlo? ¿Por qué?

a)

$x_i \cdot f_i^A$	$x_i \cdot f_i^B$	$x_i^2 \cdot f_i^A$	$x_i^2 \cdot f_i^B$
0	0	0	0
4	0	4	0
2	4	4	8
0	6	0	18
0	12	0	48
0	30	0	150
0	18	0	108
0	14	0	98
8	16	64	128
36	0	324	0
50	0	500	0
100	100	896	558

La nota media de 4.º A es: $\bar{x}_A = \frac{100}{20} = 5$.

La nota media de 4.º B es: $\bar{x}_B = \frac{100}{20} = 5$.

- b) Para observar la dispersión calculamos las desviaciones típicas de ambas clases:

$$s_A = \sqrt{\frac{896}{20} - 5^2} = 4,45; s_B = \sqrt{\frac{558}{20} - 5^2} \approx 1,7$$

Por tanto, 4.º B tiene las notas menos dispersas.

- c) No ha sido necesario calcular CV, ya que cuando las medias son iguales tiene menor dispersión la distribución que tenga menor desviación típica.

14.25 La profesora de Educación Física realiza un estudio referente a la altura y el peso de los alumnos de una clase, obteniendo los siguientes resultados: la altura, en metros, del 95 % de los alumnos se encuentra dentro del intervalo (1,52; 1,92), y el peso, en kilogramos, del mismo porcentaje de alumnos se incluye en el intervalo (56,9; 66,1).

¿Cuál de las distribuciones tiene una dispersión relativa mayor?

El intervalo que contiene el 95 % de los datos de una distribución es $(\bar{x} - 2s, \bar{x} + 2s)$. Sustituyendo los datos del enunciado obtenemos el siguiente sistema de dos ecuaciones con dos incógnitas para la altura:

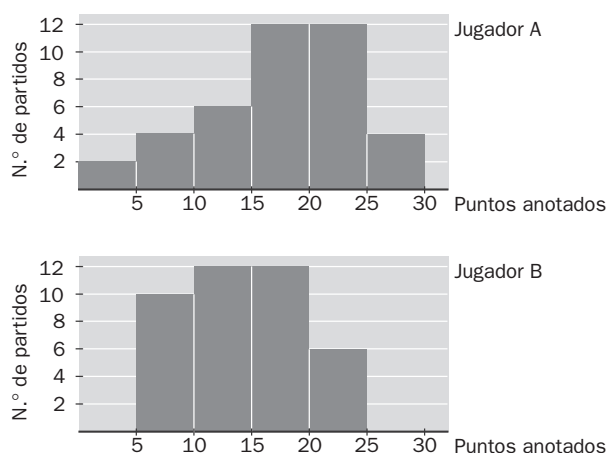
$$\begin{cases} \bar{x} - 2s = 1,52 \\ \bar{x} + 2s = 1,92 \end{cases} \Rightarrow \bar{x} = 1,72; s = 0,1. \text{ Por tanto: } CV_a = \frac{0,1}{1,72} \approx 0,06$$

De igual manera, para el peso obtenemos el siguiente sistema de dos ecuaciones con dos incógnitas:

$$\begin{cases} \bar{x} - 2s = 56,9 \\ \bar{x} + 2s = 66,1 \end{cases} \Rightarrow \bar{x} = 61,5; s = 2,3. \text{ Por tanto: } CV_p = \frac{2,3}{61,5} \approx 0,04$$

En definitiva, la distribución de las alturas tiene una dispersión relativa mayor, ya que su CV es mayor.

14.26 Los dos siguientes histogramas muestran los puntos anotados por dos jugadores de baloncesto a lo largo de toda la liga.



a) ¿Cuál de ellos alcanza mejor media anotadora?

b) ¿Quién es más regular en su anotación?

a) Nos creamos la tabla asociada a los histogramas anteriores para poder responder a las cuestiones.

Marcas x_i	f_i^A	f_i^B	$x_i \cdot f_i^A$	$x_i \cdot f_i^B$	$x_i^2 \cdot f_i^A$	$x_i^2 \cdot f_i^B$
2,5	2	0	5	0	12,5	0
7,5	4	10	30	75	225	562,5
12,5	6	12	75	150	937,5	1875
17,5	12	12	210	210	3675	3675
22,5	12	6	270	135	6075	3037,5
27,5	4	0	110	0	3025	0
	40	40	700	570	13 950	9150

La media anotadora del jugador A es: $\bar{x}_A = \frac{700}{40} = 17,5$.

La media anotadora del jugador B es: $\bar{x}_B = \frac{570}{40} = 14,25$.

Por tanto, alcanza mejor media anotadora el jugador A.

b) Para determinar quién es más regular debemos calcular los CV de ambos:

$$s_A = \sqrt{\frac{13\,950}{40} - 17,5^2} \approx 6,52 \Rightarrow CV_A = \frac{6,52}{17,5} \approx 0,37$$

$$s_B = \sqrt{\frac{9150}{40} - 14,25^2} \approx 5,068 \Rightarrow CV_B = \frac{5,068}{14,25} \approx 0,35$$

Por tanto, es más regular en su anotación el jugador B.

CUESTIONES PARA ACLARARSE

14.27 ¿Puede ser que la media no coincida con ningún valor de la variable? ¿Y la moda?

La media puede no coincidir con ningún valor de la variable; sin embargo, la moda siempre estará asociada a un valor concreto de la distribución.

14.28 Averigua el dato que falta en la siguiente distribución para que la media sea 18.

7, 12, 15, 22, 23, 28, 32

$$\frac{7 + 12 + 15 + 22 + 23 + 28 + 32 + x}{8} = 18 \Rightarrow x = 5$$

14.29 ¿Por qué en la siguiente tabla la mediana resulta poco significativa?

x_i	3	12	2000
f_i	50	1	50

Claramente, la mediana de la distribución es 12, al ser el valor central de la misma; sin embargo, no informa en absoluto de la realidad de la distribución, que está repleta de los valores 3 y 2000.

14.30 Al preguntar a 30 alumnos cuántas asignaturas han suspendido, 13 de ellos contestan que 5 materias. ¿Qué sector circular le corresponde al valor 5 de la variable *Número de suspensos*?



Le corresponde el sector circular de 156° , ya que si 360° corresponde al total de los alumnos, entonces a 13 alumnos les corresponden 156° .

14.31 Si se suma a todos los valores de la variable una constante, ¿cómo quedan afectadas la media y la varianza? ¿Y si se multiplican por una constante?

Si se suma a todos los valores de la variable una constante, la media queda aumentada en esa constante, mientras que la varianza no varía.

Si se multiplican todos los valores de la variable por una constante, la media queda multiplicada por esa constante, mientras que la varianza queda multiplicada por el cuadrado de dicha constante.

14.32 Pon un ejemplo de una distribución donde la media, la moda y la mediana coincidan.

Por ejemplo, la siguiente distribución: 1, 2, 2, 3, 3, 3, 4, 4, 5

En este caso, Media = Moda = Mediana = 3

Puede considerarse también cualquier otro ejemplo que conserve una simetría.

14.33 El rango es un valor que mide la dispersión de valores de la variable. Calcula el rango de estas dos distribuciones e indica en cuál de ellas el rango es más significativo de la realidad de los datos y por qué.

x_i	5	15	40
f_i	2	25	3

y_i	5	15	40
f_i	12	2	16

El rango de las dos distribuciones es $40 - 5 = 35$, es decir, indica una dispersión grande de los datos que es mucho más real en y_i que en x_i , ya que, en general, en esta última están muy poco dispersos.

14.34 Relaciona cada variable con su desviación típica. Razónalo.

x_i	f_i
2	4
4	4
6	3
8	4

$$s = 1,32$$

y_i	f_i
2	1
4	1
6	13
8	1

$$s = 2,7$$

z_i	f_i
2	6
4	2
6	1
8	6

$$s = 2,29$$

x_i	f_i
2	4
4	4
6	3
8	4

$$s = 2,29$$

y_i	f_i
2	1
4	1
6	13
8	1

$$s = 1,32$$

z_i	f_i
2	6
4	2
6	1
8	6

$$s = 2,7$$

Ya que la mayor dispersión se produce en z_i , después en x_i , y, sin embargo, en y_i están muy concentrados los datos.

PROBLEMAS PARA APLICAR

- 14.35 En una ciudad de Castilla y León, el 35 % de sus habitantes son hombres y el 65 %, mujeres. Entre las mujeres, el 20 % son niñas; el 25 %, adultas, y el 55 %, mayores. Entre los hombres, el 15 % son niños; el 25 %, adultos, y el 60 %, mayores.

Para elaborar un estudio exhaustivo sobre los hábitos cotidianos de la población de esa ciudad, se elige una muestra de 1200 habitantes. ¿Cuál es la muestra más representativa de la población?

	Muestra 1			Muestra 2			Muestra 2	
	H	M		H	M		H	M
Niñ.	63	156		58	160		70	145
Adul.	105	195		100	182		95	170
May.	252	429		262	438		255	465

La muestra más representativa es la 1:

	Muestra 1	
	H	M
Niñ.	63	156
Adul.	105	195
May.	252	429

- 14.36 La moda y la media aritmética tratan de resumir en un número los valores que la variable ha ido tomando. En ocasiones, la media se ajusta más que la moda a la distribución, y a veces lo contrario. En cada una de las siguientes tablas, ¿qué parámetro (\bar{x} o M_o) es más significativo y por qué?

x_i	f_i
138	2
254	1
351	1
2	30

y_i	f_i
3	8
7	6
12	7
21	5

En la variable A: $\bar{x} \approx 27,7$; $M_o = 2$. La moda es mucho más representativa, ya que de 34 datos, 30 toman el valor 2, que es la moda.

En la variable B: $\bar{x} \approx 9,8$; $M_o = 3$, y ya que los datos están mucho más repartidos, es más representativa la media.

- 14.37 Los resultados obtenidos al tirar un dado han sido los siguientes.

3, 2, 4, 3, 1, 2, 6, 3, 5, 5, 1, 3, 2

Ordena los datos y averigua los tres cuartiles y la mediana.

Ordenación: 1, 1, 2, 2, 2, 3, 3, 3, 3, 4, 5, 5, 6

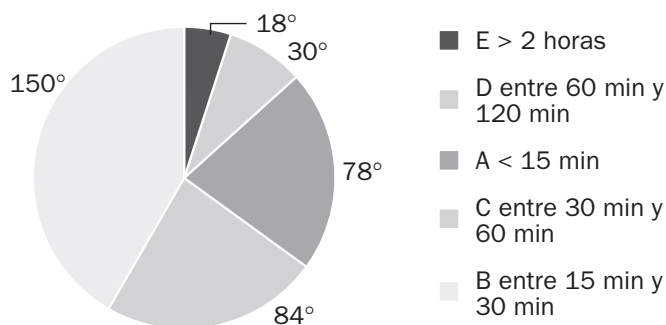
Los cuartiles son:

$$Q_1 = 2$$

$$Q_2 = 3 = M$$

$$Q_3 = 4$$

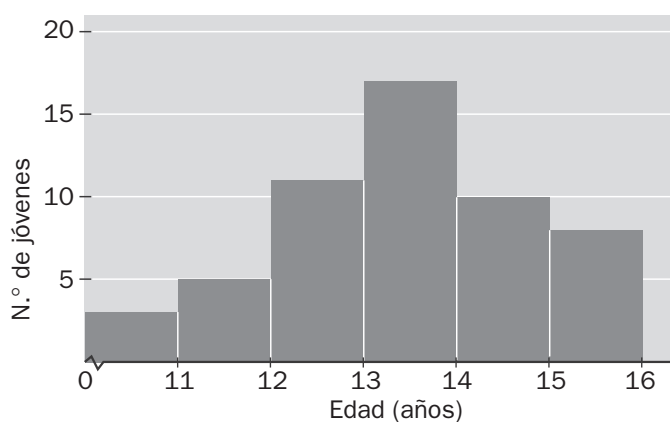
- 14.38 A 60 alumnos de Secundaria se les pregunta sobre el tiempo que dedican al ordenador diariamente. Las respuestas obtenidas las refleja este diagrama de sectores.



¿Cuántos alumnos respondieron a las distintas categorías del estudio?

Realizando sendas reglas de tres obtenemos los siguientes resultados: A: 13 alumnos; B: 25; C: 14; D: 5, y E: 3.

14.39 Según datos del Plan Nacional de Drogas, la edad media de inicio en el consumo de alcohol es de 13,6 años. En un estudio que se ha realizado en un centro de Secundaria a 54 jóvenes sobre la edad en la que comenzaron a ingerir alcohol, se han obtenido los siguientes resultados.



a) Elabora una tabla con los datos del gráfico y calcula su media y su mediana.

b) Compara los resultados con la media nacional.

a)

Edad (años)	N.º de alumnos f_i	Marcas x_i	F_i	$x_i \cdot f_i$
[10, 11)	3	10,5	3	31,5
[11, 12)	5	11,5	8	57,5
[12, 13)	11	12,5	19	137,5
[13, 14)	17	13,5	36	229,5
[14, 15)	10	14,5	46	145
[15, 16)	8	15,5	54	124
	54			725

$$\text{Media: } \bar{x} = \frac{725}{54} \approx 13,4$$

$$\text{Mediana: } M = 13,5$$

b) La media nacional es de 13,6, así que el estudio ratifica la bajada de edad en el consumo de alcohol e incluso con una media menor.

14.40 El número de asignaturas suspensas en 4.º de ESO en un centro de Secundaria en septiembre viene reflejado en la siguiente tabla.

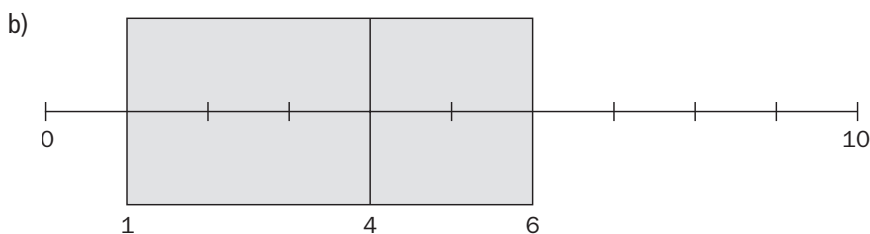
x_i	0	1	2	3	4	5	6	7	8	9	10
f_i	20	12	8	7	17	13	8	7	5	3	5

- Calcula la mediana y los cuartiles.
- Haz un diagrama de cajas y bigotes, y analiza la simetría de la distribución.
- ¿Cuál es la media de suspensos?
- En este centro han definido los siguientes parámetros: el *índice de éxito escolar*, que mide el porcentaje de alumnos con posibilidad de titular (dos o menos suspensos); el *índice de repetición*, que establece los alumnos entre tres y seis suspensos, y el *índice de fracaso escolar*, que determina los alumnos con siete o más suspensos.

Calcula los tres índices relativos a la tabla expresados en tantos por ciento.

a)	x_i	0	1	2	3	4	5	6	7	8	9	10
	f_i	20	12	8	7	17	13	8	7	5	3	5
	F_i	20	32	40	47	64	77	85	92	97	100	105

$Q_1 = 1$
 $Q_2 = M = 4$
 $Q_3 = 6$



Analizando el diagrama de cajas y bigotes, se observa que hay más concentración de datos en la primera cuarta parte al tener el bigote más corto. Se produce mayor dispersión entre el 25 % y el 50 % que entre el 50 % y el 75 %. Por último, la vertical que corta la caja no está en el centro, lo que indica asimetría.

- $\bar{x} = \frac{396}{105} \approx 3,77$
- Índice de éxito escolar $\approx 38,09 \%$
Índice de repetición $\approx 42,86 \%$
Índice de fracaso escolar $\approx 19,05 \%$

14.41 En una población nórdica con 2500 habitantes adultos se ha realizado un estudio sobre su altura. La distribución de alturas es normal (unimodal y simétrica).

Sabiendo que en el intervalo (172, 196) se encuentran 2375 habitantes y que la altura media es de 184 centímetros, calcula:

- La desviación típica de la distribución.
 - El número de habitantes que miden entre 178 y 190 centímetros.
- En el intervalo (172, 196) se encuentran 2375 habitantes, lo que supone el 95 % de la población total. Como el intervalo que contiene el 95 % de los datos de una distribución es $(\bar{x} - 2s, \bar{x} + 2s)$, se tiene que $\bar{x} - 2s = 172$. Sustituyendo los datos del enunciado obtenemos la siguiente ecuación:

$$184 - 2s = 172 \Rightarrow 2s = 184 - 172 \Rightarrow s = \frac{184 - 172}{2} = 6$$

- El número de habitantes que miden entre 178 y 190 cm será igual al número de habitantes que se encuentren en el intervalo $(178, 190) = (\bar{x} - s, \bar{x} + s)$. Este último intervalo contiene el 68 % de los datos, lo que supone que hay 1700 habitantes que miden entre 178 y 190 cm.

Caracteres y parámetros estadísticos. Gráficos

14.42 La siguiente tabla muestra las edades de las personas que acuden a un bibliobús de barrio solicitando préstamos de libros en un día cualquiera.

Edad	N.º de personas
[6, 8)	5
[8, 10)	12
[10, 12)	14
[12, 14)	13
[14, 16)	4
[16, 18)	2

- Halla la media, la moda y el tercer cuartil.
- Calcula la desviación típica.
- Representa el histograma y el polígono de frecuencias.

a)

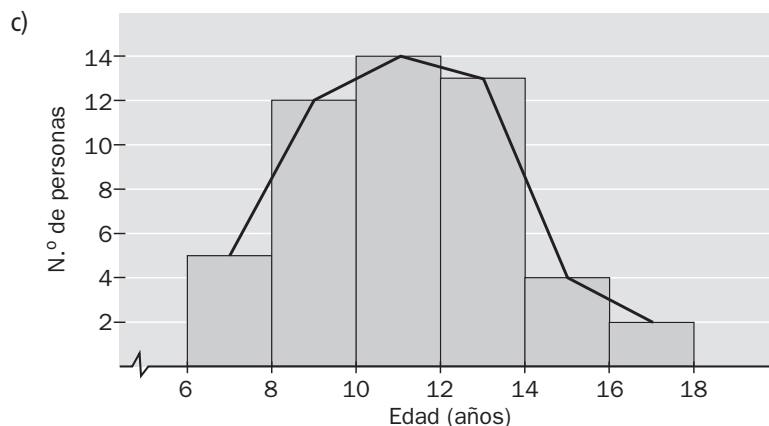
Marcas (x_i)	N.º personas (f_i)	F_i	$x_i \cdot f_i$	$x_i^2 \cdot f_i$
7	5	5	35	245
9	12	17	108	972
11	14	31	154	1694
13	13	44	169	2197
15	4	48	60	900
17	2	50	34	578
	50		560	6586

La media es: $\bar{x} = \frac{560}{50} = 11,2$ años.

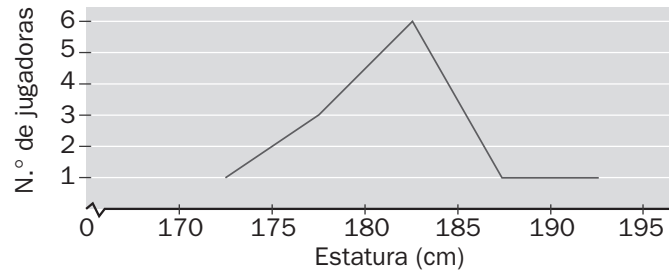
La moda es: $M_o = 11$ años.

El tercer cuartil es: $Q_3 = 13$ años.

b) $s = \sqrt{\frac{6586}{50} - 11,2^2} \approx 2,5$



- 14.43 Observa el siguiente polígono de frecuencias, que muestra las estaturas de las 12 jugadoras de un equipo de voleibol femenino.



Calcula la media aritmética y la desviación típica.

Nos creamos la tabla asociada al polígono de frecuencias anterior para poder responder a las cuestiones:

Marcas (cm) (x_i)	N.º jugadoras (f_i)	$x_i \cdot f_i$	$x_i^2 \cdot f_i$
172,5	1	172,5	29 756,25
177,5	3	532,5	94 518,75
182,5	6	1095	199 837,5
187,5	1	187,5	35 156,25
192,5	1	192,5	37 056,25
	12	2180	396 325

La media aritmética es: $\bar{x} = \frac{2180}{12} = 181,67$ cm.

La desviación típica es: $s = \sqrt{\frac{396\,325}{12} - 181,67^2} = 4,93$.

- 14.44 Susana, profesora de Educación Plástica y Visual, evalúa a sus alumnos cada trimestre con la media de 10 calificaciones sobre distintas pruebas y trabajos. Bilal ha obtenido, de momento, las siguientes notas:

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

Le queda tan solo la nota del último examen que debe realizar. ¿Qué calificación debe obtener para aprobar la asignatura con un 5?

Lo que necesita Bilal es que su media sea 5, es decir:

$$\bar{x} = \frac{2 + 4 + 4 + 5 + 8 + 3 + 6 + 3 + 5 + a}{10} = 5 \Rightarrow \frac{40 + a}{10} = 5 \Rightarrow a = 10$$

Por tanto, para aprobar la asignatura con un 5, Bilal debe obtener un 10.

Utilización conjunta de \bar{x} y s

14.45 Se ha registrado en la siguiente tabla el número de goles marcados por dos equipos de balonmano en 8 partidos del campeonato nacional de liga.

EQ. 1	25	24	27	24	26	25	27	24
EQ. 2	28	30	21	22	27	20	28	30

a) Calcula el número medio de goles de cada uno de los equipos.

b) ¿Cuál de ellos es más regular en su tanteo?

a)

Goles (x_i)	f_{i1}	f_{i2}	$x_i \cdot f_{i1}$	$x_i \cdot f_{i2}$	$x_i^2 \cdot f_{i1}$	$x_i^2 \cdot f_{i2}$
20	0	1	0	20	0	400
21	0	1	0	21	0	441
22	0	1	0	22	0	484
23	0	0	0	0	0	0
24	3	0	72	0	1728	0
25	2	0	50	0	1250	0
26	1	0	26	0	676	0
27	2	1	54	27	1458	729
28	0	2	0	56	0	1568
29	0	0	0	0	0	0
30	0	2	0	60	0	1800
	8	8	202	206	5112	5422

La media del equipo 1 es: $\bar{x}_1 = \frac{202}{8} = 25,25$ goles.

La media del equipo 2 es: $\bar{x}_2 = \frac{206}{8} = 25,75$ goles.

b) Ya que las medias son distintas, para comprobar la regularidad de cada uno debemos calcular el CV:

$$s_1 = \sqrt{\frac{5112}{8} - 25,25^2} \approx 1,2 \Rightarrow CV_1 = \frac{1,2}{25,25} \approx 0,047$$

$$s_2 = \sqrt{\frac{5422}{8} - 25,75^2} \approx 3,83 \Rightarrow CV_2 = \frac{3,83}{25,75} \approx 0,15$$

Por tanto, es más regular en su tanteo el equipo 1, ya que su CV es menor.

14.46 Se ha realizado una encuesta entre los alumnos de un colegio de Enseñanza Primaria con el objeto de conocer el número de horas semanales que ven la televisión. El estudio arroja la siguiente información: el número de horas del 95 % de los alumnos se encuentra en el intervalo (3,18; 17,1). Calcula la media aritmética y la desviación típica.

El intervalo que contiene el 95 % de los datos de una distribución es $(\bar{x} - 2s, \bar{x} + 2s)$. Sustituyendo los datos del enunciado obtenemos el siguiente sistema de dos ecuaciones con dos incógnitas:

$$\begin{cases} \bar{x} - 2s = 3,18 \\ \bar{x} + 2s = 17,1 \end{cases} \Rightarrow \bar{x} = 10,14; s = 3,48$$

14.47 Completa la siguiente distribución con dos datos más de forma que:

2, 5, 6, 7, 7, 9

a) Se conserve la media, pero la desviación típica aumente.

b) Se conserve la media, pero la desviación típica disminuya.

a) Para que se conserve la media y aumente la desviación típica, podemos añadir los datos 3 y 9.

b) Para que se conserve la media y disminuya la desviación típica, podemos añadir los datos 6 y 6.

14.48 Dadas las distribuciones:

x_i	f_i
1	3
3	7
5	11
7	18
9	21

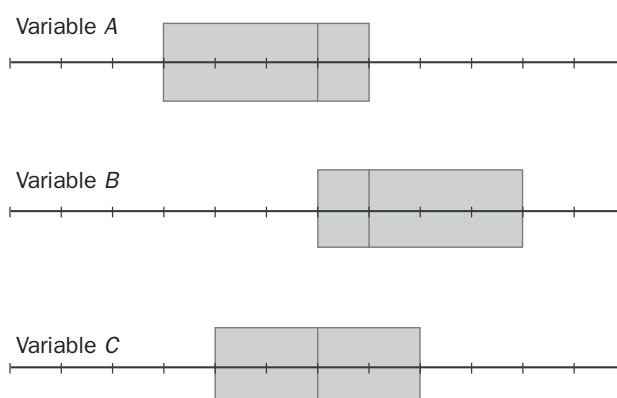
x_i	f_i
1	15
3	7
5	15
7	8
9	15

x_i	f_i
1	5
3	12
5	25
7	14
9	4

¿En cuál de ellas se puede decir que en el intervalo $(\bar{x} - s, \bar{x} + s)$ se encuentra el 68 % de los datos?
¿Por qué?

La tercera variable es la única donde se puede asegurar dicha afirmación, ya que es simétrica y unimodal. En la primera variable no se puede garantizar por ser asimétrica, y en la segunda tampoco por ser trimodal.

14.49 Estos son los diagramas de cajas y bigotes de tres variables.



Determina qué se puede decir de cada uno de ellos en cuanto a los siguientes aspectos.

a) Concentración de los datos.

b) Dispersión de los datos.

c) Simetría.

Variable A: Más concentración de datos en la primera cuarta parte al tener el bigote más corto. Más dispersión entre el 25 % y el 50 % que entre el 50 % y el 75 %. La vertical que corta la caja no está en el centro, lo que indica asimetría.

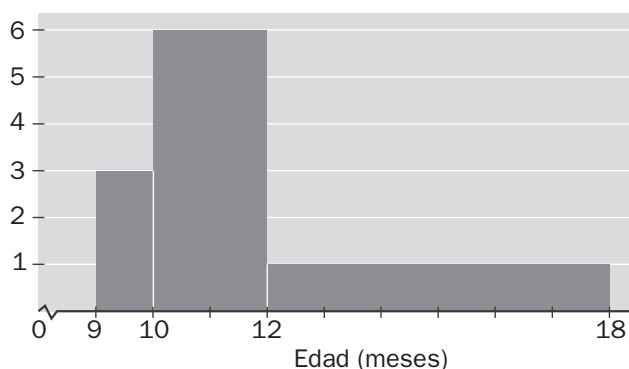
Variable B: Más concentración de datos en la cuarta parte más alta al tener el bigote más corto. Más concentración entre el 25 % y el 50 % que entre el 50 % y el 75 %. La vertical que corta la caja no está en el centro, lo que indica asimetría.

Variable C: Concentración simétrica a lo largo de los cuatro cuartos.

14.50 Echar a andar

Se ha realizado un estudio sobre los meses de edad que tenían un grupo de bebés en el momento en que comenzaron a caminar.

Los resultados vienen expresados mediante el siguiente histograma, donde las áreas de los rectángulos son proporcionales a las frecuencias.



Se sabe que hay 21 bebés que echaron a andar entre los 9 y 10 meses.

- a) ¿Cuántos bebés se observaron para realizar el estudio? ¿Cuántos comenzaron entre los 10 y 12 meses? ¿Y entre los 12 y 18 meses?
- b) Se consideran como casos “dentro de lo habitual” todos aquellos cuya edad de echar a andar no se aleje de la media más de tres desviaciones típicas. En estas condiciones, ¿cuántos casos no se pueden considerar como “dentro de lo habitual”? ¿Qué porcentaje representan?

a) Construimos la siguiente tabla:

Intervalos	Marcas	f_i	$f_i \cdot x_i$	$f_i \cdot x_i^2$
[9, 10)	9,5	21	199,5	1895,25
[10, 12)	11	84	924	10 164
[12, 18)	15	42	630	9450
Totales		147	1753,5	21 509,25

Analizando dicha tabla se deduce que se observaron 147 bebés para realizar el estudio. 84 de ellos echaron a andar entre los 10 y 12 meses, y 42, entre los 12 y 18 meses.

$$b) \bar{x} = \frac{\sum f_i x_i}{\sum f_i} = \frac{1753,5}{147} \approx 11,93 \quad s = \sqrt{\frac{\sum f_i \cdot x_i^2}{\sum f_i} - \bar{x}^2} = \sqrt{\frac{21 509,25}{147} - 11,93^2} \approx 2$$

$$(\bar{x} - 3s, \bar{x} + 3s) = (5,93; 17,93)$$

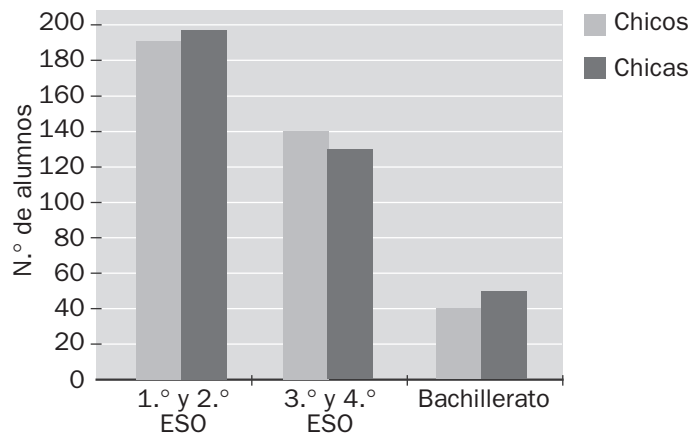
Solo se consideran como “fuera de lo habitual” aquellos casos cuya edad de echar a andar haya sido superior a 17,93 meses.

$$\text{Por tanto: } \frac{18 - 12}{42} = \frac{18 - 17,93}{x} \Rightarrow x = 0,49, \text{ que representan el } \frac{0,49}{147} \cdot 100 = 0,33 \%$$

14.51 La encuesta

Se quiere realizar una encuesta entre los alumnos de un centro escolar sobre sus preferencias en materia de actividades extraescolares.

El siguiente gráfico muestra la distribución de los alumnos según el sexo y el nivel que estudian.



Para el estudio se ha elegido una muestra de 50 alumnos distribuidos de la siguiente forma.

	1.º - 2.º ESO	3.º - 4.º ESO	Bachillerato
Chicos	12	10	3
Chicas	13	9	3

¿Crees que la composición de la muestra es la adecuada o, por el contrario, habría que variar la misma? Justifica tu respuesta.

La distribución de los alumnos, en porcentajes, es:

	1.º - 2.º ESO	3.º - 4.º ESO	Bachillerato
Chicos	24,14 %	19,31 %	5,52 %
Chicas	26,21 %	17,93 %	6,90 %

Repartiendo los 50 alumnos de la muestra de acuerdo con los porcentajes quedaría:

	1.º - 2.º ESO	3.º - 4.º ESO	Bachillerato
Chicos	12,07	9,655	2,76
Chicas	13,105	8,965	3,45

Considerando que los números deben ser naturales y sumar 50, la distribución más ajustada de la muestra es precisamente la aportada en el enunciado.

14.A1 El entrenador de un saltador de longitud decide estudiar las últimas diez pruebas en las que participó su pupilo, obteniendo la siguiente tabla.

Longitud del salto (m)	N.º de saltos
[7,8; 7,9)	1
[7,9; 8,0)	2
[8,0; 8,1)	3
[8,1; 8,2)	2
[8,2; 8,3)	2

a) Halla la media, la moda, la mediana y el primer cuartil.

b) Calcula el rango y la desviación típica.

c) Representa un histograma.

a)

Marcas (x_i)	N.º saltos (f_i)	$x_i \cdot f_i$	$x_i^2 \cdot f_i$
7,85	1	7,85	61,6225
7,95	2	15,9	126,405
8,05	3	24,15	194,4075
8,15	2	16,3	132,845
8,25	2	16,5	136,125
	10	80,7	651,405

La media es: $\bar{x} = \frac{80,7}{10} = 8,07$ m.

La moda es: $M_o = 8,05$ m.

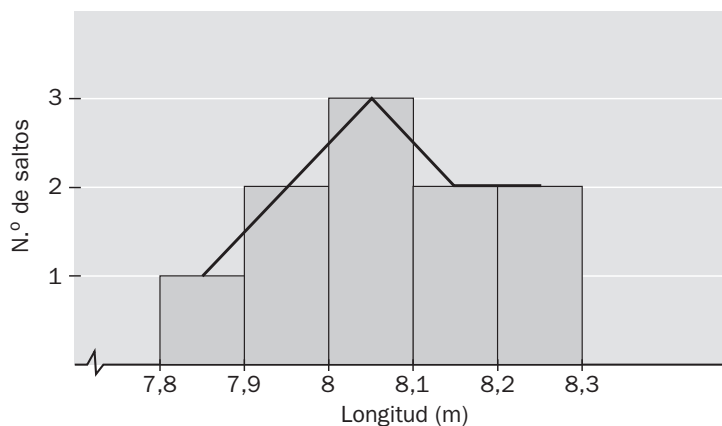
La mediana es: $M = 8,05$ m.

El primer cuartil es: $Q_1 = 7,95$ m.

b) Rango = $8,25 - 7,85 = 0,4$ m.

La desviación típica es: $s = \sqrt{\frac{651,405}{10} - 8,07^2} = \sqrt{0,0156} \approx 0,125$.

c) El histograma que se obtiene es el siguiente:

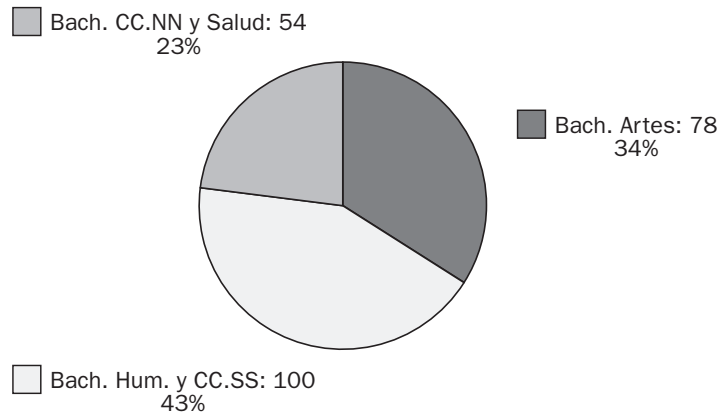


14.A2 Si la media de una distribución es $\bar{x} = 1,97$ y la desviación típica es $s = 0,08$, halla el intervalo en el cual se encuentran el 99 % de los datos de la distribución.

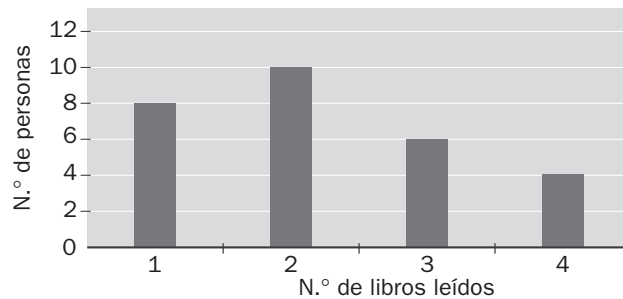
El intervalo que contiene el 99 % de los datos de una distribución es $(\bar{x} - 3s, \bar{x} + 3s)$. Por tanto, al sustituir los datos del enunciado en el intervalo obtenemos $(1,73; 2,21)$, que es el intervalo pedido.

14.A3 En las aulas de Bachillerato de un centro escolar hay 100 alumnos en la modalidad de Humanidades y CC. SS., 54 en la de CC. NN. y de la Salud, y 78 en la de Artes.

Representa estos datos en un diagrama de sectores.



14.A4 El siguiente diagrama de barras nos muestra el número de libros leídos en un año por las treinta personas que trabajan en una oficina.



a) Calcula la media aritmética y la desviación típica.

b) ¿Cuál es el coeficiente de variación?

a) Nos creamos la tabla asociada al diagrama de barras anterior para poder responder a las cuestiones:

Libros leídos (x_i)	N.º personas (f_i)	$x_i \cdot f_i$	$x_i^2 \cdot f_i$
1	8	8	8
2	10	20	40
3	7	21	63
4	5	20	80
	30	69	191

La media es: $\bar{x} = \frac{69}{30} = 2,3$

La desviación típica es: $s = \sqrt{\frac{191}{30} - 2,3^2} \approx 1,04$

b) $CV = \frac{s}{\bar{x}} = \frac{1,04}{2,3} \approx 0,45$

MATE TIEMPOS

El precio de la gasolina

La capacidad media de un barril de petróleo es de 158,98 litros. El coste medio del barril en el mercado de Londres durante un determinado mes es de 70 dólares. Si el precio de la gasolina, en ese mismo mes, fue de 1 euro por litro y un dólar se cotizó a 0,7 euros, ¿cuál fue la diferencia entre el precio de coste y el de venta?

Precio de un litro de petróleo en dólares: $\frac{70}{158,98} = 0,4403$

Precio de un litro de petróleo en euros: $0,4403 \cdot 0,7 = 0,3082$

Diferencia, en euros, entre el precio de venta y el de coste: $1 - 0,3082 = 0,6918 \text{ €}$