MATEMÁTICAS ORIENTADAS A LAS ENSEÑANZAS ACADÉMICAS
3.º ESO

somoslink

SOLUCIONES AL LIBRO DEL ALUMNO

UNIDAD 8. ESTADÍSTICA

Unidad 8. Estadística

SOLUCIONES PÁG. 165

1 **Indica de estos caracteres estadísticos cuáles son cualitativos y cuáles cuantitativos:**

a. La renta per cápita de un país.
Cuantitativa continua, porque se puede expresar numéricamente y toma infinitos valores

b. La variación del nivel de agua de un embalse.
Cuantitativa continua, porque se puede expresar numéricamente y toma infinitos valores.

c. Tu marca de ropa deportiva preferida.
Cualitativa, porque no se puede expresar o medir numéricamente.

d. El estado civil de las personas.
Cualitativa, porque no se puede expresar o medir numéricamente.

2 **Escribe tres caracteres cuantitativos y otros tres cualitativos.**
Respuesta abierta.

3 **Escribe tres variables estadísticas discretas y otras tres continuas.**
Respuesta abierta.

4 **Clasifica los siguientes caracteres estadísticos en cualitativos o cuantitativos. De los cualitativos indica alguna posible modalidad, y de los cuantitativos señala cuáles son discretos y cuáles continuos.**

a. El cociente intelectual de las personas.
Cuantitativa continua, porque se puede expresar numéricamente y toma infinitos valores.

b. Las materias optativas que eliges en 3.º de ESO.
Cualitativa. Respuesta abierta, por ejemplo, Matemáticas orientadas a las enseñanzas sociales o a las enseñanzas académicas.

c. La densidad de la población en las localidades de un país.
Cuantitativa continua, porque se puede expresar numéricamente y toma infinitos valores.

d. Los monumentos patrimonios de la humanidad.
Cualitativa. Respuesta abierta, por ejemplo, monumentos situados en Asia, en Europa o en América.

5 **Se desea conocer la oferta hotelera que ofrece una capital para el desarrollo de un gran evento deportivo.**

a. ¿Cuál sería la población y los individuos del estudio?
La población sería todos los hoteles de la ciudad, y los individuos cada uno de los hoteles.

b. ¿Sería conveniente hacer el estudio sobre una muestra? ¿Por qué? ¿Cómo harías la muestra?
Sí, porque la población es grande y llevaría demasiado tiempo realizarla. Respuesta abierta.

6 **Completa en tu cuaderno las siguientes frases con el término que consideres más adecuado: continua – discreta – modalidad – muestra – tamaño – variable cuantitativa**

a. **La variable estadística que toma un número infinito de valores se llama.** continua, **y la que toma un número finito,** discreta.

b. **El carácter estadístico que se expresa con números se denomina** variable cuantitativa.

c. **Una** muestra **es una parte de la población.**

d. **El** tamaño **de una muestra es el número de individuos que tiene.**

e. **Cada uno de los valores que toma un carácter estadístico cualitativo se llama** modalidad.

7 **Indica cuáles de estas afirmaciones son ciertas y corrige las erróneas:**

a. **Una muestra contiene siempre a todos los individuos del estudio estadístico.**
Falso, es la población.

b. **Una muestra se utiliza para economizar o agilizar el estudio estadístico.**
Verdadero.

c. **Cualquier muestra es representativa de su población.**
Falso, debe tomarse aleatoriamente.

d. **La población es un subconjunto de la muestra.**
Falso, sucede al contrario.

8 **Confecciona una encuesta para obtener información sobre las siguientes variables estadísticas:**

a. **Los hábitos de lectura.**
Respuesta abierta.

b. **Una dieta equilibrada.**
Respuesta abierta.

c. **Un correcto horario de estudio.**
Respuesta abierta

d. **El conocimiento de los diferentes itinerarios en 4.º de ESO.**
Respuesta abierta.

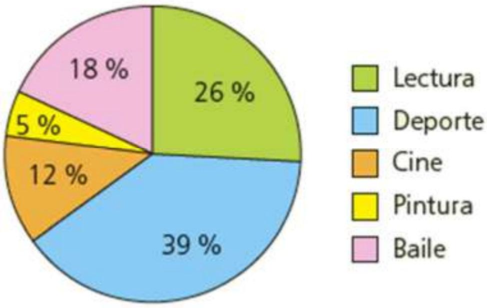
9 **En el municipio de Petra se quiere potenciar el reciclaje de materiales. Para ello, el ayuntamiento ha colgado en su web una encuesta a fin de mejorar los servicios prestados y fomentar el reciclaje entre los vecinos.**

a. **Indica una variable estadística que pertenezca a este estudio.**
Respuesta abierta.

b. **¿Cuál es la población?**
Los habitantes del municipio.

c. **¿Harías el estudio sobre la población o sobre una muestra? En el caso de optar por la muestra, ¿cómo la elaborarías?**
Sobre una muestra. Respuesta abierta.

10 Un estudio referente a las actividades preferidas por los adolescentes mayores de edad en su tiempo libre refleja estos resultados:



Indica la población, el carácter estadístico y las diferentes modalidades que aparecen en este estudio.

La población son todos los adolescentes mayores de edad. El carácter estadístico es la actividad preferida para el tiempo libre. Las modalidades son: lectura, deporte, cine, pintura y baile.

11 Visita esta dirección de Internet para repasar los distintos tipos de variables: http://conteni2.educarex.es/mats/12028/contenido/
Respuesta abierta.

SOLUCIONES PÁG. 167

12 Se ha preguntado a 38 personas sobre sus preferencias relativas al tipo de libros que leen y se han obtenido los siguientes resultados: [novela policíaca (P), novela histórica (H), ensayo (E), libro de temas de actualidad (A)].

P H P P A H H A P H
H P H E H H P P A H
A H A A A H H E A A
H A H H H A A A

Representa los datos en una tabla estadística que incluya todas las frecuencias y el porcentaje.

Para reflejar los datos en una tabla ha de tenerse en cuenta:

- Se sitúan los datos, que son los valores que toma la variable, x_i, en este caso se trata de las claves P, H, etc, es decir, los tipos de lectura.
- Se indica la frecuencia absoluta, n_i, de cada uno de los valores de P, H, etc., es decir, el número de veces que se repite cada valor.
- Se calcula la frecuencia absoluta acumulada, N_i, es decir, si P es el valor menos frecuente, con 7 veces, y H es el siguiente menos repetido con 16 veces, la frecuencia absoluta de H es 7 + 16 = 23.
- Se calcula la frecuencia relativa, f_i, como el cociente entre la frecuencia absoluta de un valor y el número total de datos, en el caso de P, por ejemplo, es $f_i(P) = \frac{n_i}{N} = \frac{7}{38} = 0{,}18$.
- Se calcula la frecuencia relativa acumulada, F_i, como la suma de todas las frecuencias relativas de los valores menores o iguales a dicho valor. Por ejemplo, si la frecuencia relativa del valor P es 0,18 y la de H es 0,42, la frecuencia acumulada de H es 0,18 + 0,42 = 0,60.

- Se calculan los porcentajes, p_i, multiplicando por 100 la frecuencia relativa de cada valor.

Datos, x_i	n_i	N_i	f_i	F_i	p_i
P	7	7	0,18	0,18	18 %
H	16	23	0,42	0,60	42 %
A	13	36	0,34	0,94	34 %
E	2	38	0,05	0,99	5 %
Total	38		0,99		99 %

13 En la ficha personal de los componentes de un equipo de ciclismo se detalla la masa, en kilogramos, de cada ciclista. La relación de las masas es la siguiente:

58,4 – 63,7 – 68,4 – 72,1 – 69,6
79,0 – 71,3 – 72,5 – 80,0 – 59,7
64,6 – 77,9 – 74,2 – 65,0 – 70,3
74,9 – 73,6 – 64,8 –72,7 – 68,8

Realiza un recuento y completa una tabla estadística, tomando como amplitud de clase 5.

Para reflejar los datos en una tabla ha de tenerse en cuenta:

- Se sitúan los datos, que son los valores de x_i, la masa de los ciclistas, agrupados en intervalos, cerrados por la izquierda y abiertos por la derecha, con excepción del último intervalo, cerrado por ambos lados.
- Se sitúan las marcas de clase, c_i, que coincide con el valor medio de los extremos de cada intervalo, por ejemplo, en el caso del primer intervalo será $c_i = \frac{55+60}{2} = 57{,}5$
- El resto de parámetros de la tabla se calculan de forma similar a la de la actividad 12.

Datos	c_i	n_i	N_i	f_i	F_i	p_i
[55 , 60)	57,50	2	2	0,10	0,10	10 %
[60 , 65)	62,50	3	5	0,15	0,25	15 %
[65 , 70)	67,50	4	9	0,20	0,45	20 %
[70 , 75)	72,50	8	17	0,40	0,85	40 %
[75 , 80]	77,50	3	20	0,15	1	15 %
Total		20		1		100 %

SOLUCIONES PÁG. 169

14 Representa un diagrama de barras y un polígono de frecuencias con los siguientes datos:

6 8 10 6 7 10 6 7 10 6 6 6 10 7 8
6 9 10 10 7 6 6 8 7 6 10 6 6 7 10

Se sitúan en el eje de abscisas los valores de la variable, x_i, y en el eje de ordenadas las frecuencias de los valores, n_i.

x_i	6	7	8	9	10
n_i	12	6	3	1	8

15

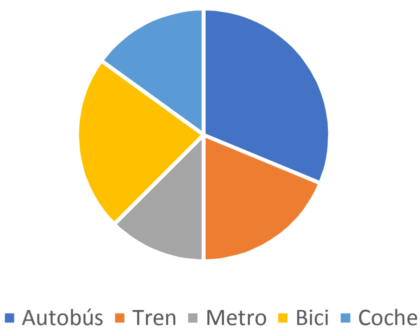
Transporte	Autobús	Tren	Metro	Bici	Coche
Frecuencia absoluta	25	15	10	18	12

Se ha tomado una muestra para averiguar el medio de transporte más utilizado por los trabajadores de una empresa:

Haz una representación de los datos en un diagrama de sectores. Recuerda que hay que calcular previamente la amplitud de cada sector.

Para calcular las amplitudes, se multiplican las frecuencias relativas, f_i por 360º, por ejemplo, en el caso del autobús, la amplitud es 0,3 · 360º = 108º

Transporte	Autobús	Tren	Metro	Bici	Coche	Total
Frecuencia absoluta, n_i	25	15	10	18	12	80
Frecuencia relativa, f_i	0,3	0,18	0,12	0,22	0,15	
Amplitud sector	108	64,8	43,2	79,2	54	

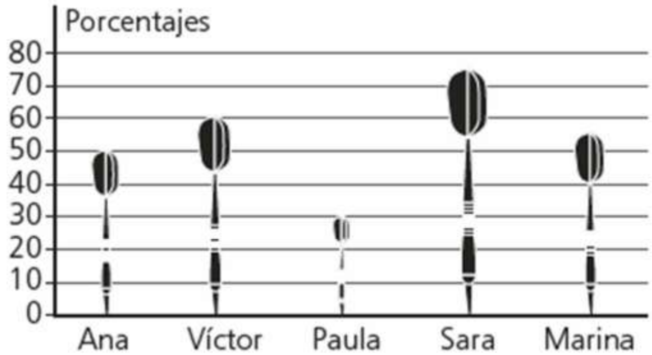


16 Unos amigos que juegan a los dardos tienen estos porcentajes de acierto al número 20 en la diana:

Jugadores	Ana	Víctor	Paula	Sara	Marina
Porcentajes	50	60	30	75	55

Representa los datos en un pictograma.

En un pictograma las barras del diagrama de barras se sustituyen por un dibujo relacionado con el carácter estadístico estudiado, en este caso, un dardo.



17 Las alturas, en metros, de 20 personas son las siguientes:

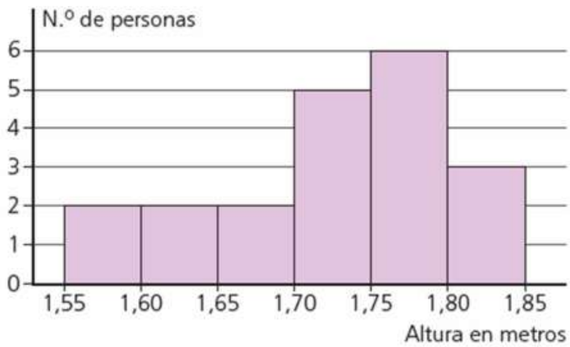
1,59 – 1,73 – 1,80 – 1,67 – 1,75
1,62 – 1,76 – 1,83 – 1,77 – 1,81
1,78 – 1,70 – 1,64 – 1,72 – 1,73
1,68 – 1,71 – 1,79 – 1,76 – 1,57

Realiza una tabla de frecuencias agrupando los datos en intervalos de 0,05 de amplitud. Toma como primer valor del primer intervalo 1,55. Efectúa luego una representación gráfica utilizando un histograma.

Para reflejar los datos en una tabla han de tenerse en cuenta las mismas recomendaciones que en la actividad 13.

Se sitúan en el eje de abscisas los valores de la variable, x_i, y en el eje de ordenadas las frecuencias de los valores, n_i.

Datos	c_i	n_i	N_i	f_i	F_i	p_i
[1,55 , 160)	1,575	2	2	0,10	0,10	10 %
[1,60 , 1,65)	1,625	2	4	0,10	0,20	10 %
[1,65 , 1,70)	1,675	2	6	0,10	0,30	10 %
[1,70 , 1,75)	1,725	5	11	0,25	0,55	25 %
[1,75 , 1,80)	1,775	6	17	0,30	0,85	30 %
[1,80 , 1,85]	1,825	3	20	0,15	1	15 %
Total		20		1		100 %



18 Formad grupos de cuatro o cinco alumnos. Cada grupo debe elegir un tema de investigación: dieta saludable, provincias de nacimiento, serie de televisión favorita, música preferida, profesión de los padres, número de hijos, etcétera. Para cada tema, recabad información, resumidla en una tabla estadística y representadla en un gráfico. Después, cada grupo expondrá a la clase sus objetivos iniciales y extraerá sus conclusiones.

Respuesta abierta.

19 En esta dirección de Internet podrás repasar los gráficos estadísticos: http://conteni2.educarex.es/mats/12031/contenido/

Respuesta abierta.

SOLUCIONES PÁG. 171

20 Actividad resuelta.

21 Halla la moda, la media, la mediana y los cuartiles de los siguientes valores:

Datos	[18 , 23)	[23 , 28)	[28 , 33)	[33 , 38]
Fr. absoluta, n_i	7	12	6	15

Se recogen los datos en una tabla:

Datos	Fr. absoluta, n_i	c_i	$n_i \cdot c_i$	Fr. absoluta acumulada
[18 , 23)	7	20,5	143,5	7
[23 , 28)	12	25,5	306	19
[28 , 33)	6	30,5	183	25
[33 , 38]	15	35,5	532,5	40
	40		1165	

La moda es el valor de la variable que más veces se repite, es decir, el de mayor frecuencia, en este caso el intervalo modal es, entonces [33 , 38]. La moda es la marca de clase de dicho intervalo, es decir, $M_o = \frac{33+38}{2} = 35{,}5$.

La media es $\bar{x} = \frac{\sum c_i \cdot n_i}{N} = \frac{143{,}5+306+183+532{,}5}{40} = \frac{1165}{40} = 29{,}13$

La mediana corresponde al intervalo [28 , 33) porque $\frac{N}{2} = \frac{40}{2} = 20$. Fijándonos en la columna de frecuencias absolutas cumuladas, el valor que sobrepasa la mitad del tamaño de la población es el correspondiente a ese intervalo (25 pertenece al intervalo [28 , 33)).

En cuanto a los cuartiles, dividen los datos en cuatro partes iguales, así que:

El primer cuartil, Q_1, corresponde al intervalo [23 , 28) porque $\frac{N}{4} = \frac{40}{4} = 10$ y fijándonos en la columna de frecuencias absolutas cumuladas, el valor que sobrepasa el 10, es ese intervalo.

El segundo cuartil, Q_2, al intervalo [28 , 33) porque $\frac{N}{4} \cdot 2 = 20$ y el tercer cuartil, Q_3, al intervalo [33, 38], pues $\frac{N}{4} \cdot 3 = 30$.

22 **Según la oficina comunitaria de estadística Eurostat, España es uno de los países en los que los jóvenes se independizan con una media de edad más avanzada.**

Estos son los datos de una encuesta realizada al respecto:

33 28 32 29 30 31 28 30 32 29
30 33 31 26 29 30 28 30 31 30
33 28 32 32 26 31 29 31 32 26

Halla la moda, la media y la mediana de esta distribución.

Se recogen los datos en una tabla:

Datos	Fr. absoluta, n_i	$n_i \cdot x_i$	Fr. Absoluta acumulada
26	3	78	3
28	4	112	7
29	4	116	11
30	6	180	17
31	5	155	22
32	5	160	27
33	3	99	30
	30	900	

El valor de la variable que más se repite es 30 de manera que la moda es $M_o = 30$

La media es: $\bar{x} = \frac{x_1 \cdot n_1 + \ldots + x_n \cdot n_n}{N} = \frac{72 + 112 + 116 + 180 + 155 + 160 + 99}{30} = \frac{900}{30} = 30$

La mediana es el valor central de los datos en orden creciente. Como el número de datos es 30 (par), hay que hacer la media aritmética de los dos datos centrales del listado ordenado $\frac{N}{2} = \frac{30}{2} = 15$, que en este caso son 30 y 30.

26 26 26 28 28 28 28 29 29 29 29 30 30 30 30 30 30 31 31 31 31 31 32 32 32 32
32 33 33 33

Otra manera de hallar la mediana es considerar que, el valor 15 está en el valor de la frecuencia absoluta acumulada 17, por lo que corresponde al dato de 30, luego la mediana es $M_e = 30$.

23 Daniel ha jugado 31 partidos con su equipo de baloncesto, estos son los puntos que ha conseguido:

5 6 7 8 9 10 6 9 7 8 9 5 10 11 8 9
6 9 10 9 5 6 7 7 11 10 8 10 7 11 20

a. Halla la media, la mediana y los cuartiles.

Se recogen los datos en una tabla:

Datos	Fr. absoluta, n_i	ni · x_i	Fr. Absoluta acumulada
5	3	15	3
6	4	24	7
7	5	35	12
8	4	32	16
9	6	54	22
10	5	50	27
11	3	33	30
20	1	20	31
	30	263	

La media es $\overline{x} = \frac{x_1 + \ldots + x_n}{N} = \frac{15+24+35+32+54+50+33+20}{31} = \frac{263}{31} = 8{,}48$

La mediana es el valor central de los datos en orden creciente. Como el número
de datos es 31 (impar), el que corresponde al valor de la mediana es el dato número 16, que en este caso es 8.

Los cuartiles dividen los datos en cuatro partes iguales:

Q_1 = 7 porque $\frac{N}{4} = \frac{31}{4} = 7{,}75$. Fijándonos en la columna de frecuencias absolutas
acumuladas, el valor que sobrepasa el 7,75, es 7.

Q_2 = 8 porque $\frac{N}{4} \cdot 2 = \frac{31 \cdot 2}{4} = 15{,}5$. Según la columna de frecuencias absolutas
acumuladas, el valor que sobrepasa 15,5 es 8.

Q_3 = 10 porque $\frac{N}{4} \cdot 3 = \frac{31 \cdot 3}{4} = 23{,}25$. Según la columna de frecuencias absolutas
acumuladas, el valor que sobrepasa 23,25 es 10.

b. ¿Se ven muy alterados los parámetros calculados si se elimina el dato más alejado?

El dato más alejado es 20, de manera que, si se elimina, la media queda:

$$\overline{x} = \frac{x_1 \cdot n_1 + \ldots + x_n \cdot n_n}{N} = \frac{15+24+35+32+54+50+33}{30} = \frac{243}{30} = 8{,}1$$

La mediana se halla haciendo la media de las posiciones 15 y 16, es decir, M_e = 8, pues según la columna de las frecuencias absolutas acumuladas esas dos posiciones corresponden al valor 8.

Q_1 = 7, porque $\frac{N}{4} = \frac{30}{4} = 7{,}5$. Fijándonos en la columna de frecuencias absolutas acumuladas, el valor que sobrepasa el 7,5, es 7.

Q_2 = 8, porque $\frac{N}{4} = \frac{30}{4} \cdot 2 = 15$. Fijándonos en la columna de frecuencias absolutas acumuladas, el valor que sobrepasa el 15, es 8.

Q_3 = 10, porque $\frac{N}{4} = \frac{30}{4} \cdot 3 = 22{,}5$. Fijándonos en la columna de frecuencias absolutas acumuladas, el valor que sobrepasa el 22,5, es 10.

En conclusión, no se ven muy afectados los parámetros calculados.

c. ¿Cuál resulta más alterado?

Solamente se ve afectada la media, el resto no varía.

24 Sandra ha obtenido las siguientes notas en Matemáticas a lo largo de todo el trimestre: 4 6 7 9
Halla la media aritmética de estas notas teniendo en cuenta las condiciones de cada apartado:

a. Calcula la media aritmética de sus notas.

$$\overline{x} = \frac{x_1 + \ldots + x_n}{N} = \frac{4+6+7+9}{4} = 6{,}5$$

b. El último examen es el final, se evalúa todo el trimestre y puntúa el doble que los demás.

$$\overline{x} = \frac{x_1 + \ldots + x_n}{N} = \frac{4+6+7+(9 \cdot 2)}{5} = 7$$

c. En cada examen vuelven a evaluarse los contenidos anteriores, y la puntuación de cada examen en el cómputo global corresponde al 10 %, 20 %, 30 % y 40 %, respectivamente.

Como Sandra ha obtenido un 4 en el primer examen, en el cómputo global ese examen contará como un 10 %, el segundo sacó un 6, que computará como un 20 % de la nota media, etc.

$$\overline{x} = (4 \cdot 0{,}1) + (6 \cdot 0{,}2) + (7 \cdot 0{,}3) + (9 \cdot 0{,}4) = 7{,}3$$

25 Con las calificaciones de siete alumnos de una clase se elabora una distribución que tiene como media un 5 y como mediana un 2.

a. ¿Es posible que ocurra esto? Razona tu respuesta.

Sí es posible. La media puede ser un 5 porque pueden repetirse algunas calificaciones de los siete alumnos. La mediana puede ser 2 si el valor central de las siete calificaciones es ese, lo que significa que al menos tres alumnos tienen una calificación menor o igual a 2.

b. En caso afirmativo, halla siete notas que lo verifiquen.

Si la mediana es 2 las tres notas más bajas deben ser inferiores a 2 y las tres más altas superiores a 2, por ejemplo, 0 1 2 2 10 10 10.

c. Con la misma media, ¿la mediana podría ser un 1?

No es posible porque si la mediana es 1 las cuatro notas más bajas deben ser al menos 1, por ejemplo, 1 1 1 1 10 10 10, ya media daría como máximo 4,85.

26 La nota media conseguida en una clase de 25 alumnos ha sido de 7. Dos alumnos han suspendido con un 1, cuatro han sacado un 3 y dos obtuvieron un 4. Si el resto de alumnos aprobó, ¿cuál es la nota media de los alumnos aprobados?

El número de dato es N = 25

La media de calificaciones es $\overline{x} = 7$

Como $\overline{x} = \dfrac{x_1 \cdot \mathrm{n}_1 + \ldots + x_n \cdot \mathrm{n}_n}{\mathrm{N}} \Rightarrow x_1 \cdot \mathrm{n}_1 + \ldots + x_n \cdot \mathrm{n}_n = \overline{x} \cdot \mathrm{N} = 25 \cdot 7 = 175$

Se puede calcular la suma de las notas de los 8 alumnos que han suspendido: (2 · 1 + 4 · 3 + 2 · 4 = 22), y restarlas al sumatorio total con el fin de obtener la suma de las notas de los aprobados: 175 – 22 = 153

Para calcular la media de los alumnos aprobados, se divide dicho sumatorio entre el número de alumnos aprobados (25 – 8 = 17):

$$\overline{x} = \frac{x_1 + \ldots + x_n}{\mathrm{N}} = \frac{153}{17} = 9$$

La nota media de los aprobados es 9.

SOLUCIONES PÁG. 173

27 Halla los parámetros de dispersión de estos datos:
7 7 8 7 8 7 8 10 7 9 8 8 7 9 8 7 10 8 7 7 8 10 7 8

El valor máximo es 10 y el mínimo es 7, luego el rango es R = 10 – 7 = 3

Datos	Fr. absoluta, n_i	$n_i \cdot x_i$	Fr. Absoluta acumulada	$x_i - \bar{x}$	$\lvert x_i - \bar{x}\rvert$	$\lvert x_i - \bar{x}\rvert \cdot n_i$
7	10	70	10	–0,91	0,91	9,1
8	9	72	19	0,09	0,09	0,81
9	2	18	21	1,09	1,09	2,18
10	3	30	24	2,09	2,09	6,27
	24	190				18,36

La media es:

$$\bar{x} = \frac{x_1 \cdot n_1 + \ldots + x_n \cdot n_n}{N} = \frac{70+72+18+30}{24} = \frac{190}{24} = 7{,}91$$

La desviación media es:

$$DM = \frac{\sum_{i=1}^{n} \lvert x_i - \bar{x}\rvert \cdot n_i}{N} = \frac{9{,}1 + 0{,}81 + 2{,}18 + 6{,}27}{24} = \frac{18{,}36}{24} = 0{,}76$$

Para calcular la varianza se debe conocer el valor de $x_i^2 \cdot n_i$:

x_i	7	8	9	10
x_i^2	49	64	81	100
n_i	10	9	2	3
$x_i^2 \cdot n_i$	490	576	162	300

$$V(x) = \frac{\sum_{i=1}^{n} x_i^2 \cdot n_i}{N} - \bar{x}^2 = \frac{490+576+162+300}{24} - (7{,}91)^2 = \frac{1528}{24} - (7{,}91)^2 = 1{,}09$$

La desviación típica es:

$$S(x) = +\sqrt{V(x)} = +\sqrt{1{,}09} = 1{,}04$$

El coeficiente de variación de Pearson es:

$$CV = \frac{S(x)}{\bar{x}} = \frac{1{,}04}{7{,}91} = 0{,}13$$

28 Transcurridos dos minutos de la realización de una prueba física, se toman las pulsaciones a dos grupos diferentes de estudiantes, con estos

	$\bar{x}$	σ
A	90	1,5
B	80	4,2

resultados:

En uno de los grupos, varios alumnos han sobrepasado las 110 pulsaciones, mientras que en el otro el alumno con menos pulsaciones tenía 70. ¿A qué grupo pertenece cada caso?

El grupo con integrantes por encima de 110 pulsaciones es el B, porque la desviación típica es alta, es decir, existe gran diferencia entre los cuadrados de las desviaciones y la media al cuadrado. Y el otro es el A porque la desviación típica es baja, es decir, la media al cuadrado se acerca más a los valores del cuadrado de las desviaciones.

29 Se ha realizado un estudio sobre el tiempo que los alumnos de un instituto pasan en el aula virtual repasando y realizando ejercicios. Los datos obtenidos entre los 30 alumnos encuestados son los siguientes:

- **Hay 8 alumnos que la utilizan un máximo de 15 minutos.**
- **Entre un cuarto de hora y media hora la usan 14 alumnos.**
- **Un total de 7 alumnos dedican entre media hora y cuarenta y cinco minutos.**
- **Un alumno permanece en al aula entre tres cuartos de hora y una hora.**

Calcula la media de tiempo empleado en el aula virtual, la desviación media de los datos y su desviación típica.

Se expresan los datos en forma de tabla:

Datos	Fr. absoluta, n_i	c_i	$n_i \cdot c_i$
(0 , 15]	8	7,5	60
(15 , 30]	14	22,5	315
(30 , 45]	7	37,5	262,5
(45 , 60]	1	52,5	52,5
	30		690

La media es: $\bar{x} = \frac{\sum c_i \cdot n_i}{N} = \frac{60 + 315 + 262,5 + 52,5}{30} = \frac{690}{30} = 23$

Para calcular los demás parámetros se deben expresar antes estas operaciones de la tabla:

$c_i - \overline{x}$	$\lvert c_i - \overline{x}\rvert$	$\lvert c_i - \overline{x}\rvert \cdot n_i$	$c_i^2 \cdot n_i$
–15,50	15,5	124	450
–0,50	0,5	7	7087,5
14,50	14,5	101,5	9843,75
29,50	29,5	29,5	2756,25
		262	20137,5

La desviación media es: $DM = \frac{\sum_{i=1}^{n} \lvert c_i - \overline{x}\rvert \cdot n_i}{N} = \frac{124+7+101{,}5+29{,}5}{30} = \frac{262}{30} = 8{,}7$

La varianza es:

$$V(x)\frac{\sum_{i=1}^{n} c_i^2 \cdot n_i}{N} - \overline{x}^2 = \frac{20137{,}5}{30} - (23)^2 = 142{,}25$$

La desviación típica es: $S(x) = +\sqrt{V(x)} = \sqrt{142{,}25} = 11{,}926$

El tiempo medio empleado es de 23 minutos, la desviación media de los datos es 8,73 minutos y la desviación típica es 11,93 minutos.

30 Dos productos financieros han dado los siguientes intereses de beneficios para el capital invertido en los últimos meses:

Producto A	Producto B
2 %	2,2 %
0,2 %	1 %
0,9 %	1,5 %
1,5 %	2,4 %
3 %	2,5 %
3,8 %	2,5 %
5,2 %	2,8 %
2,1 %	2,1 %

a. Halla la media aritmética, el rango y la desviación típica de cada distribución.

Se genera la tabla de datos para calcular los parámetros estadísticos:

x_i (A)	n_i	$n_i \cdot x_i$	$x_i^2 \cdot n_i$		x_i (B)	n_i	$n_i \cdot x_i$	$x_i^2 \cdot n_i$
2	1	2	4		2,2	1	2,2	4,84
0,2	1	0,2	0,04		1	1	1	1
0,9	1	0,9	0,81		1,5	1	1,5	2,25
1,5	1	1,5	2,25		2,4	1	2,4	5,76
3	1	3	9		2,5	1	2,5	6,25
3,8	1	3,8	14,44		2,5	1	2,5	6,25
5,2	1	5,2	27,04		2,8	1	2,8	7,84
2,1	1	2,1	4,41		2,1	1	2,1	4,41

La media del producto A es:

$$\bar{x}_A = \frac{x_1 \cdot n_1 + \ldots + x_n \cdot n_n}{N} = \frac{2+0,2+0,9+1,5+3+3,8+5,2+2,1}{8} = \frac{18,7}{8} = 2,34$$

La media del producto B es:

$$\bar{x}_B = \frac{x_1 \cdot n_1 + \ldots + x_n \cdot n_n}{N} = \frac{2,2+1+1,5+2,4+2,5+2,5+2,8+2,1}{8} = \frac{17}{8} = 2,12$$

El rango de cada producto es el valor mayor menos el menor:

$R_{A} =$ 5,2 – 0,2 = 5

$R_{B} =$ 2,8 – 1 = 1,8

La varianza para A es:

$$V(x) = \frac{\sum_{i=1}^{n} x_i^2 \cdot n_i}{N} - \bar{x}^2 =$$

$$= \frac{4+0,04+0,81+2,25+9+14,44+27,04+4,41}{8} - (2,34)^2 = 2,27$$

La desviación típica para A es:

$$S = S(x) = +\sqrt{V(x)} = \sqrt{2,27} = 1,51$$

La varianza para B es:

$$V(x) = \frac{\sum_{i=1}^{n} x_i^2 \cdot n_i}{N} - \bar{x}^2 =$$

$$= \frac{4,84+1+2,25+5,76+6,25+6,25+7,84+4,41}{8} - (2,12)^2 = 0,33$$

La desviación típica para B es:

$$S = S(x) = +\sqrt{V(x)} = \sqrt{0,33} = 0,57$$

En el producto A, la media es 2,34, el rango es 5 y la desviación típica es 1,51, y en el producto B, la media es 2,12, el rango es 1,8 y la desviación típica es 0,57.

b. Razona en cuál de los dos productos invertirías.

Aunque el producto A ofrece una mayor rentabilidad media, es mucho más estable el producto B, porque su rango es menor y su desviación típica también. No es mucha la diferencia de rentabilidad entre ambos, por tanto, es más aconsejable invertir en el producto B.

31 Actividad resuelta.

32 **Los diez primeros exámenes que se han corregido en dos grupos de 3.º ESO han registrado las siguientes calificaciones:**

A: 4, 7, 7, 5, 6, 8, 8, 6, 9, 10
B: 5, 7, 8, 6, 10, 5, 4, 9, 9, 7

Calcula la desviación media y la desviación típica en estos dos grupos de datos.

Se recogen los datos en una tabla:

Notas, x_i	Frecuencia absoluta, n_i	$n_i \cdot x_i$	Frecuencia absoluta acumulada	$x_i - \bar{x}$	$\lvert x_i - \bar{x}\rvert$	$\lvert x_i - \bar{x}\rvert \cdot n_i$	$x_i^2 \cdot n_i$
4	1	4	1	–3	3	3	16
5	1	5	2	–2	2	2	25
6	2	12	4	–1	1	2	72
7	2	14	6	0	0	0	98
8	2	16	8	1	1	2	128
9	1	9	9	2	2	2	81
10	1	10	10	3	3	3	100
	10	70				14	520

Los parámetros para el grupo A son:

$$\bar{x} = \frac{x_1 \cdot n_1 + \ldots + x_n \cdot n_n}{N} = \frac{4+5+12+14+16+9+10}{10} = \frac{70}{10} = 7$$

$$DM = \frac{\sum_{i=1}^{n} \lvert x_i - \bar{x}\rvert \cdot n_i}{N} = \frac{3+2+2+0+2+2+3}{10} = \frac{14}{10} = 1{,}4$$

$$V(x) = \frac{\sum_{i=1}^{n} x_i^2 \cdot n_i}{N} - \bar{x}^2 =$$

$$= \frac{16+25+72+98+128+81+100}{10} - 7^2 = 3$$

$$S(x) = +\sqrt{V(x)} = \sqrt{3} = 1{,}73$$

Notas, x_i	Frecuencia absoluta, n_i	ni · x_i	Frecuencia absoluta acumulada	$x_i - \bar{x}$	$\lvert x_i - \bar{x}\rvert$	$\lvert x_i - \bar{x}\rvert$ · ni	x_i^2 · ni
4	1	4	1	–3	3	3	16
5	2	10	3	–2	2	4	50
6	1	6	4	–1	1	1	36
7	2	14	6	0	0	0	98
8	1	8	7	1	1	1	64
9	2	18	9	2	2	4	162
10	1	10	10	3	3	3	100
	10	70				16	526

Los parámetros para el grupo B son:

$$\overline{x} = \frac{x_1 \cdot \text{n}_1 + \ldots + x_n \cdot \text{n}_n}{\text{N}} = \frac{4+10+6+14+8+18+10}{10} = \frac{70}{10} = 7$$

$$\text{DM} = \frac{\sum_{i=1}^{n} |x_i - \overline{x}| \cdot n_i}{N} = \frac{3+4+1+0+1+4+3}{10} = \frac{16}{10} = 1{,}6$$

$$\text{V}(x) = \frac{\sum_{i=1}^{n} x_i^2 \cdot \text{n}_i}{\text{N}} - \overline{x}^2 =$$

$$= \frac{16+50+36+98+64+162+100}{10} - 7^2 = 3{,}6$$

$$\text{S}(x) = +\sqrt{\text{V}(x)} = \sqrt{3{,}6} = 1{,}89$$

La media en ambos casos coincide y es un 7. La desviación media del grupo A es 1,4 y del grupo B es 2,1. La desviación típica del grupo A es 1,73 y del B, 1,89.

33 Formad grupos de cuatro o cinco alumnos en la clase. Cada grupo debe elegir un tema sobre el que se preguntará al resto de alumnos de la clase a fin de obtener información. A continuación, someterá los datos así recopilados a un tratamiento estadístico completo: tabla estadística, representación gráfica y cálculo de parámetros estadísticos. Por último, cada grupo expondrá a la clase los resultados del estudio y extraerá sus conclusiones.

Respuesta abierta.

34 En esta dirección de Internet puedes repasar las de medidas de dispersión:
http://conteni2.educarex.es/mats/12037/contenido/

Respuesta abierta.

SOLUCIONES PÁG. 175

35 Construye un diagrama de caja y bigotes con estos datos:
3 5 3 1 8 4 4 3 5 2 5 4 1 6 3 5 4 4 1 3

Se genera la tabla de datos:

x_i	Frecuencia absoluta, n_i	Frecuencia absoluta acumulada
1	3	3
2	1	4
3	5	9
4	5	14
5	4	18
6	1	19
8	1	20
	20	

Se calculan los cuartiles y el rango intercuartílico:

$\frac{N}{4}=\frac{20}{4}=5$, según la columna de frecuencias absolutas acumuladas, el valor que sobrepasa el 5 es 3, con lo que $Q_1 = 3$.

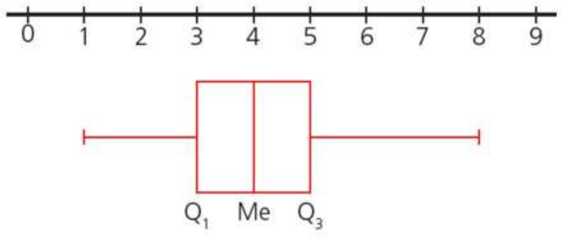
$\frac{N}{4}\cdot 2=\frac{20}{4}\cdot 2=10$, según la columna de frecuencias absolutas acumuladas, el valor que sobrepasa el 10 es 4, con lo que $M_e = Q_2 = 4$.

$\frac{N}{4}\cdot 3=\frac{20}{4}\cdot 3=15$, según la columna de frecuencias absolutas acumuladas, el valor que sobrepasa el 15 es 5, con lo que $Q_3 = 5$.

El rango intercuartílico es $R_i = Q_3 – Q_1 = 5 – 3 = 2$.

Para representar el diagrama de cajas y bigotes se tiene en cuenta que:

- El rectángulo se construye con el valor de $Q_1 = 3$ en el extremo izquierdo del rectángulo y el valor de $Q_3 = 5$ en el derecho.
- La frontera inferior es $f_i = Q_1 – 1,5 \cdot R = 3 – 1,5 \cdot 2 = 0$ y se representa en el extremo izquierdo. La frontera superior es $f_s = Q_3 + 1,5 \cdot R_i = 5 + 1,5 \cdot 2 = 8$ y se representa en el extremo derecho.
- Como los valores de los datos mínimo y máximo son, respectivamente $x_i = 1$ y $x_i = 8$ y están dentro del intervalo de las fronteras, que es (0 , 8), la longitud de los bigotes llega hasta los valores mínimo y máximo, y no hace falta representar el valor de las fronteras.
- El rango intercuartílico, $R_i = 2$ coincide con la longitud de la caja.
- La mediana, $M_e = Q_2 = 4$ divide al rectángulo en dos partes.



36 Elabora un diagrama de caja y bigotes para este conjunto de datos, que representan las notas de un examen de Matemáticas de una clase de 3.º de ESO:

7 5 3 9 6 4 8 7 5 2 5 8 9 6 7 5 4 8 9 6 3 7 9 2 8 6 7 5

Se genera la tabla de datos:

x_i	Frecuencia absoluta, n_i	Frecuencia absoluta acumulada
2	2	2
3	2	4
4	2	6
5	5	11
6	4	15
7	5	20
8	4	24
9	4	28
	28	

Se calculan los cuartiles y el rango intercuartílico:

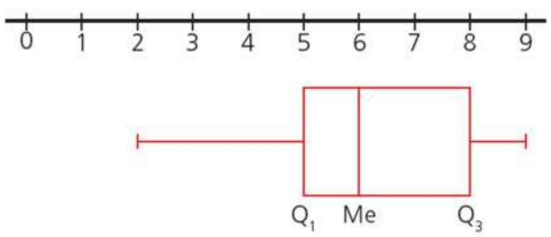
$\frac{N}{4}=\frac{28}{4}=7$, según la columna de frecuencias absolutas acumuladas, el valor que sobrepasa el 7 es 5, con lo que $Q_1 = 5$.

$\frac{N}{4}\cdot 2=\frac{28}{4}\cdot 2=14$, según la columna de frecuencias absolutas acumuladas, el valor que sobrepasa el 14 es 6, con lo que $M_e = Q_2 = 6$.

$\frac{N}{4}\cdot 3=\frac{28}{4}\cdot 3=21$, según la columna de frecuencias absolutas acumuladas, el valor que sobrepasa el 21 es 8, con lo que $Q_3 = 8$.
El rango intercuartílico es $R_i = Q_3 – Q_1 = 8 – 5 = 3$.

Para representar el diagrama de cajas y bigotes se tiene en cuenta que:

- El rectángulo se construye con el valor de $Q_1 = 5$ en el extremo izquierdo del rectángulo y el valor de $Q_3 = 8$ en el derecho.
- La frontera inferior es $f_i = Q_1 – 1,5 \cdot R = 5 – 1,5 \cdot 3 = 0,5$ y se representa en el extremo izquierdo. La frontera superior es $f_s = Q_3 + 1,5 \cdot R_i = 8 + 1,5 \cdot 3 = 12,5$ y se representa en el extremo derecho.
- Como los valores de los datos mínimo y máximo son, respectivamente $x_i = 2$ y $x_i = 9$ y están dentro del intervalo de las fronteras, que es (0,5 , 12,5), la longitud de los bigotes llega hasta los valores mínimo y máximo, y no hace falta representar el valor de las fronteras.
- El rango intercuartílico, $R_i = 3$ coincide con la longitud de la caja.
- La mediana, $M_e = Q_2 = 4$ divide al rectángulo en dos partes.



37 Actividad resuelta.

38 Elabora un diagrama de caja y bigotes para los datos expresados en la siguiente tabla:

Datos, x_i	10	11	12	13	14	15	16
Fr. absoluta, n_i	3	7	39	53	56	21	25

Se genera la tabla de datos:

x_i	Frecuencia absoluta, n_i	Frecuencia absoluta acumulada
10	3	3
11	7	10
12	39	49
13	53	102
14	56	158
15	21	179
16	25	204
	204	

Se calculan los cuartiles y el rango intercuartílico:

$\frac{N}{4} = \frac{204}{4} = 51$, según la columna de frecuencias absolutas acumuladas, el valor que sobrepasa el 51 es 13, con lo que $Q_1 = 13$.

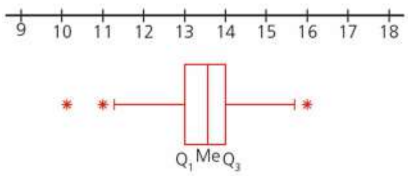
$\frac{N}{4} \cdot 2 = \frac{204}{4} \cdot 2 = 102$, según la columna de frecuencias absolutas acumuladas, el valor que sobrepasa el 102 es 13,5, porque está entre el valor 13 y el 14.Así que el valor de la mediana es $M_e = Q_2 = 13{,}5$.

$\frac{N}{4} \cdot 3 = \frac{204}{4} \cdot 3 = 153$, según la columna de frecuencias absolutas acumuladas, el valor que sobrepasa el 153 es 14, con lo que $Q_3 = 14$.

El rango intercuartílico es $R_i = Q_3 - Q_1 = 14 - 13 = 1$.

Para representar el diagrama de cajas y bigotes se tiene en cuenta que:

- El rectángulo se construye con el valor de $Q_1 = 13$ en el extremo izquierdo del rectángulo y el valor de $Q_3 = 14$ en el derecho.
- La frontera inferior es $f_i = Q_1 - 1{,}5 \cdot R = 13 - 1{,}5 \cdot 3 = 11{,}5$ y se representa en el extremo izquierdo. La frontera superior es $f_s = Q_3 + 1{,}5 \cdot R_i = 14 + 1{,}5 \cdot 3 = 15{,}5$ y se representa en el extremo derecho.
- Como los valores de los datos mínimo y máximo son, respectivamente $x_i = 10$ y $x_i = 16$ y no están dentro del intervalo de las fronteras, que es (11,5 , 15,5), la longitud de los bigotes llega hasta los valores de la frontera y se representan los valores atípicos (con asterisco) que quedan fuera de dicho intervalo frontera.
- El rango intercuartílico, $R_i = 1$ coincide con la longitud de la caja.
- La mediana, $M_e = Q_2 = 13{,}5$ divide al rectángulo en dos partes.



39 **Se han registrado las temperaturas máximas de cada uno de los días de la primera quincena del mes de marzo:**

24 ºC, 26 ºC, 29 ºC, 20 ºC, 21 ºC, 18 ºC, 16 ºC, 17 ºC, 22 ºC, 23 ºC, 19 ºC, 25 ºC, 15 ºC, 20 ºC, 17 ºC

Elabora un diagrama de caja y bigotes para estos datos.

Se genera la tabla de datos:

x_i	Frecuencia absoluta, n_i	Frecuencia absoluta acumulada
15	1	1
16	1	2
17	2	4
18	1	5
19	1	6
20	2	8
21	1	9
22	1	10
23	1	11
24	1	12
25	1	13
26	1	14
29	1	15
	15	

Se calculan los cuartiles y el rango intercuartílico:

$\frac{N}{4} = \frac{15}{4} = 3{,}75$, según la columna de frecuencias absolutas acumuladas, el valor que sobrepasa el 3,75 es 17, con lo que $Q_1 = 17$.

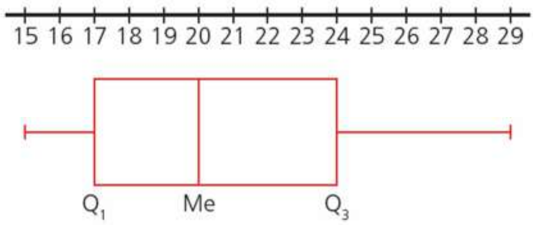
$\frac{N}{4} \cdot 2 = \frac{15}{4} \cdot 2 = 7{,}5$, según la columna de frecuencias absolutas acumuladas, el valor que sobrepasa el 7,5 es 20, con lo que $M_e = Q_2 = 20$.

$\frac{N}{4} \cdot 3 = \frac{15}{4} \cdot 2 = 11{,}25$, según la columna de frecuencias absolutas acumuladas, el valor que sobrepasa el 11,25 es 24, con lo que $Q_3 = 24$.

El rango intercuartílico es $R_i = Q_3 - Q_1 = 24 - 17 = 7$.

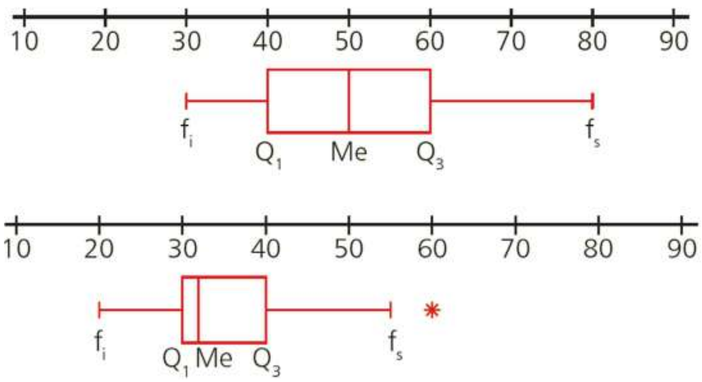
Para representar el diagrama de cajas y bigotes se tiene en cuenta que:

- El rectángulo se construye con el valor de $Q_1 = 17$ en el extremo izquierdo del rectángulo y el valor de $Q_3 = 24$ en el derecho.
- La frontera inferior es $f_i = Q_1 - 1{,}5 \cdot R = 17 - 1{,}5 \cdot 7 = 6{,}5$ y se representa en el extremo izquierdo. La frontera superior es $f_s = Q_3 + 1{,}5 \cdot R_i = 24 + 1{,}5 \cdot 3 = 34{,}5$ y se representa en el extremo derecho.
- Como los valores de los datos mínimo y máximo son, respectivamente $x_i = 15$ y $x_i = 29$ y están dentro del intervalo de las fronteras, que es (6,5 , 34,5), la longitud de los bigotes llega hasta los valores mínimo y máximo, y no hace falta representar el valor de las fronteras.
- El rango intercuartílico, $R_i = 7$ coincide con la longitud de la caja.
- La mediana, $M_e = Q_2 = 20$ divide al rectángulo en dos partes.



40 Actividad resuelta.

41 Las edades de dos grupos de personas se reflejan en estos dos diagramas:



Compáralos para responder a las siguientes preguntas:

a. ¿En cuál de los dos grupos hay mayor diferencia de edad?

En el primer grupo, porque el rango intercuartílico es mayor que en el segundo grupo (20 frente a 10)

b. ¿En qué grupo está la persona más joven? ¿Y la de más edad?

En el segundo grupo, en el que hay, al menos, un valor de edad de 20 años, pues el valor mínimo que marca el bigote es ese frente al valor mínimo del ptro grupo, que es de 30.

La persona de más edad está en el primer grupo, pues el valor máximo que marca el bigote es de 80 años, frente al valor máximo del segundo grupo, que es de 60.

c. ¿Hay algún punto atípico? ¿Qué edad representa?

Sí, en el segundo grupo hay un valor atípico marcado con un asterisco, es decir, fuera del intervalo de fronteras, que es de 60 años.

d. ¿En qué grupo están más concentradas las edades?

En el segundo grupo están más concentradas, porque el diagrama tiene la caja más corta, lo que indica valores más concentrados.

e. ¿Cuáles son los cuartiles de cada distribución?

En la primera distribución $Q_1 = 40$, $Q_2 = 50$ y $Q_3 = 60$, y en la segunda distribución $Q_1 = 30$, $Q_2 = 33$ y $Q_3 = 40$.

SOLUCIONES PÁG. 177

1 **Escribe un ejemplo de cada una de las variables estadísticas indicadas.**
Respuesta abierta.

2 **Los datos que se agrupan en intervalos ¿a qué tipo de variables estadísticas pertenecen?**
A cuantitativas continuas, o en el caso excepcional de cuantitativa discreta con muchos valores.

3 **¿Qué tipo de variable se puede representar con un histograma?**
Cuantitativa continua.

4 **En un diagrama de sectores, ¿se puede representar una variable cuantitativa o cualitativa? Razona tu respuesta.**
En un diagrama de sectores se puede representar cualquier tipo de variable.

5 **¿Qué indican los parámetros de centralización?**
Los valores centrales en torno a los cuales se distribuyen los datos y que representan, de forma global, los datos estadísticos de esa población.

6 **¿Qué información aportan los parámetros de dispersión?**
El índice de agrupamiento de los datos.

7 **¿Qué elementos se representan en un diagrama de caja y bigotes?**
Los cuartiles Q_1, Q_2 y Q_3 y el rango intercuartílico, y los valores extremos, si hiciera falta.

8 **Si, en un diagrama de caja y bigotes, los bigotes son largos, ¿qué significa?**
El rango de los valores es grande, y por lo tanto los datos están dispersos.

9 **Prepara una presentación para tus compañeros. Puedes hacer un documento PowerPoint, utilizar Glogster…**
Respuesta abierta.

SOLUCIONES PÁG. 178

REPASO FINAL
ESTADÍSTICA. POBLACIÓN Y MUESTRA

1 **Indica cuáles de estos caracteres estadísticos son cualitativos y cuáles cuantitativos:**

a. La capacidad del disco duro de los ordenadores.
Cuantitativo discreto. Un ejemplo de modalidad es un disco duro de 500 GB, de 1TB, etc.

b. El tiempo en una carrera de 100 m lisos.
Cuantitativo continuo (se pueden obtener marcas como, 9,85 s o 10,2 s).

c. Los tipos de plantas de un jardín botánico.
Cualitativo. Ejemplos de modalidad pueden ser plantas de zonas tropicales, plantas de bosque mediterráneo, etc.

d. Los dígitos de nuestro sistema numérico decimal.

Cuantitativo discreto. La relación de los dígitos del sistema decimal es 1, 2, 3, 4, 5, etc., por tanto, es discreto.

Con respecto a los cualitativos, señala alguna posible modalidad, y, en relación con los cuantitativos, indica cuáles son discretos y cuáles continuos.

FRECUENCIAS Y TABLAS ESTADÍSTICAS

2 Al contabilizar el número de actividades extraescolares que ha realizado un grupo de alumnos, se obtienen estos datos:

3 3 4 3 4 3 4 5 3 2 4 4 3 2 4
3 5 4 3 3 4 5 3 4 2 4 2 3 3 4

a. Realiza un recuento de los datos y elabora una tabla estadística.

x_i	Frecuencia absoluta, n_i	Frecuencia absoluta acumulada, N_i	Frecuencia relativa, f_i	Frecuencia relativa acumulada, F_i	Porcentaje, p_i
2	4	4	0,13	0,13	13,33
3	12	16	0,40	0,53	40,00
4	11	27	0,37	0,90	36,67
5	3	30	0,10	1,00	10,00
	30		1		100,00

b. ¿Cuál es el tamaño de la muestra?

El número total de datos o tamaño de la muestra, N, es la suma de todas las frecuencias absolutas:

N = 4 + 12 +11 +3 = 30.

c. ¿Cuánto suman las frecuencias relativas de todos los datos?

La frecuencia relativa $f_i = \frac{n_i}{N_i}$ siempre suma 1.

d. ¿Qué porcentaje de alumnos ha tenido cuatro actividades extraescolares?

El porcentaje se obtiene multiplicando por 100 el valor de la frecuencia relativa, en este caso, el porcentaje del valor 4 (el que marca 4 actividades extraescolares) es 36,67 %.

e. ¿Cuántos alumnos han tenido hasta tres actividades extraescolares?

La frecuencia absoluta es el número de veces que se repite dicho valor, es decir, han tenido 3 extraescolares 12 alumnos. No obstante, han sido más los que han tenido hasta 3 actividades, porque también cuentan los que han tenido 2 actividades extraescolares. En ese caso habrá que consultar la frecuencia absoluta acumulada para 3l valor de 3 actividades, que es 18 alumnos.

3 **Un radar fijo situado en un tramo de una carretera nacional en el que la velocidad máxima permitida es de 70 km/h ha registrado estas velocidades de vehículos:**

58 – 76 – 90 – 71 – 100 – 89 – 107 – 52 – 65 – 108 – 94 – 73 – 62 – 65 – 79 – 110 – 88 – 64 – 93 – 68

a. **Realiza un recuento y completa una tabla estadística, tomando 10 como amplitud de clase.**
Se agrupan los datos en intervalos de 10, teniendo en cuenta que la marca de clase es la media de los valores extremos del intervalo.

Datos	c_i	n_i	N_i	f_i	F_i	p_i
[50 , 60)	55	2	2	0,10	0,10	10 %
[60 , 70)	65	5	7	0,25	0,35	25 %
[70 , 80)	75	4	11	0,2	0,55	20 %
[80 , 90)	85	2	13	0,10	0,65	10 %
[90 , 100)	95	3	16	0,15	0,80	15 %
[100 , 110]	105	4	20	0,20	1	20 %
Total		20		1		100 %

b. **¿Cuál es el tamaño de la muestra?**
El tamaño es 20, pues se cuenta con 20 datos de velocidad.

c. **¿Cuántos vehículos circulan entre 60 km/h y 70 km/h?**
El valor de n_i en el intervalo [60, 70) es 5, luego circulan a esa velocidad 5 vehículos.

d. **¿Qué porcentaje de vehículos circulan entre 90 km/h y 100 km/h?**
El valor de p_i en el intervalo [90, 100) es 15 %.

e. **¿Cuántos vehículos exceden los 80 km/h?**
El valor de n_i en el intervalo [80 , 90) es 2, el valor de n_i en el intervalo [90 , 100) es 3 y el del intervalo [100 , 110] es 4, con lo que la suma es 2 + 3 + 4 = 9 vehículos.

4 **Copia en tu cuaderno y completa los huecos de esta tabla estadística:**

Datos	n_i	N_i	f_i	F_i	p_i
2	20				
4			0,15		
6		40			
8					
Total	80		1		100

Para completar la tabla es se debe calcular el valor preciso en cada celda.
En primer lugar:

La **frecuencia acumulada** del valor 2 es la suma del valor de ni de 2, es decir, 20.
La **frecuencia relativa** del valor 2 es 20/80=**0,25**

Datos	n_i	N_i	f_i	F_i	pi
2	**20**	20	0,25		
4			**0,15**		
6		40			
8					
Total	**80**		**1**		**100**

En segundo lugar:

fi = ni/N; Si fi es 0,15 y N = 80, entonces ni = 12

Datos	n_i	N_i	f_i	F_i	pi
2	**20**	20	0,25		
4	12		**0,15**		
6		40			
8					
Total	**80**		**1**		**100**

En tercer lugar:

Datos	n_i	N_i	f_i	F_i	pi
2	**20**	20	0,25		
4	12	20 +12 = 32	**0,15**		
6	8	20 + 12 + 8 = **40**			
8	40	80			
Total	**80**		**1**		**100**

Las frecuencias acumuladas se van hallando sumando las frecuencias absolutas de cada valor y sus anteriores.

En cuarto lugar:

Datos	n_i	N_i	f_i	F_i	pi
2	**20**	20	0,25	0,25	0,25 · 100 = 25 %
4	12	20 +12 = 32	**0,15**	0,25+0,15=0,40	0,15 · 100 = 15 %
6	8	20 + 12 + 8 = **40**	0,1	0,40 + 0,1 = 0,5	0,1 · 100 = 10 %
8	40	80	0,5	0,5 + 0, 5 = 1	0,5 · 100 = 50 %
Total	**80**		1		**100**

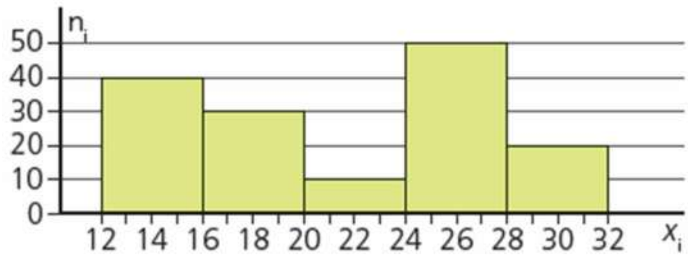
Una vez obtenidos los valores de ni, se hallan las frecuencias relativas, las frecuencias acumuladas y los porcentajes de todos los valores.

Al final, la tabla queda así:

Datos	n_i	N_i	f_i	F_i	p_i
2	20	20	0,25	0,25	25 %
4	12	32	0,15	0,40	15 %
6	8	40	0,1	0,50	10 %
8	40	80	0,5	1	50 %
Total	80		1		100 %

GRÁFICOS ESTADÍSTICOS

5 Las edades de los alumnos de una escuela de música están representadas en el siguiente histograma:



Indica, razonando tu respuesta, si estas afirmaciones son verdaderas o falsas:

a. Más de la mitad de los alumnos son mayores de 23 años.

Cierto, porque si expresamos en una tabla las edades y frecuencias de la edad de los alumnos obtenemos:

x_i	[12 , 16)	[16 , 20)	[20 , 24)	[24 , 28)	[28 , 32)
n_i	40	30	10	50	20

Se observa que el acumulado de alumnos de edad menor o igual a 23 es:
40 + 30 + 10 = 80, mientras que el acumulado de alumnos mayores de 24 es:
50 + 20 = 70

b. Los alumnos de entre 20 y 26 años no superan el 40 % del total.

El número de alumnos entre 20 y 26 es la suma de las frecuencias de los alumnos del intervalo [20 , 24), que es 10 y una parte que no se conoce del intervalo [24 , 28), que es 50, de modo que no se sabe con certeza.

c. Si no ingresan nuevos alumnos en la escuela, dentro de 4 años los alumnos menores de 16 años serán un 1 %.

Falso. Si se desplaza el histograma 4 años hacia la derecha, el grupo de alumnos que ahora está en el intervalo [12 , 16) estará en el intervalo [16, 20) y en el intervalo de edad anterior no habrá alumnos.

d. Dentro de 4 años casi todos los alumnos serán mayores de edad.

No es seguro, depende de cómo estén distribuidos en el intervalo [12 , 16), que dentro de 4 años serán el intervalo [16 , 20).

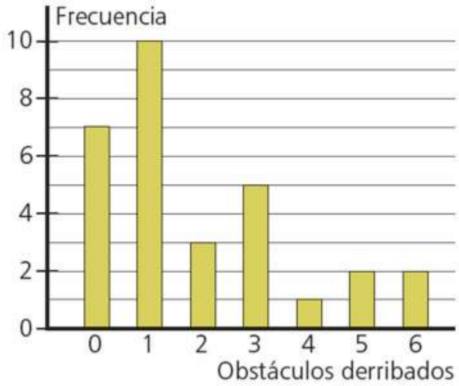
PARÁMETROS DE CENTRALIZACIÓN

6 En una prueba equina de salto de obstáculos se contabiliza el número de obstáculos derribados por los caballos participantes, con los siguientes resultados:

Obstáculos derribados	0	1	2	3	4	5	6
Frecuencia absoluta, n_i	7	10	3	5	1	2	2

a. Representa los datos en un diagrama de barras.

En el eje de abscisas se representan los obstáculos derribados y en el de ordenadas la frecuencia.



b. Calcula los parámetros de centralización.

La moda, $M_o = 1$ porque es el número de obstáculos que más veces se repite.

La media es:

$$\bar{x} = \frac{(7 \cdot 0) + (10 \cdot 1) + (3 \cdot 2) + (5 \cdot 3) + (1 \cdot 4) + (2 \cdot 5) + (2 \cdot 6)}{30} = \frac{57}{30} = 1{,}9$$

x_i	Frecuencia absoluta, n_i	Frecuencia absoluta acumulada
0	7	7
1	10	17
2	3	20
3	5	25
4	1	26
5	2	28
6	2	30
	30	

Los cuartiles y la mediana son:

Q_1 = 1 porque $\frac{N}{4}=\frac{30}{4}=7{,}5$ y el primer valor que supera el 7,5 en la frecuencia absoluta acumulada es el 17, que se corresponde con el valor 1 de obstáculos derribados.

M_e = Q_2 = 1 porque $\frac{N}{4}\cdot 2=\frac{30}{4}\cdot 2=15$ y el primer valor que supera el 15 en la frecuencia absoluta acumulada es el 17, que se corresponde con el valor 1 de obstáculos derribados.

Q_3 = 3 porque $\frac{N}{4}\cdot 3=\frac{30}{4}\cdot 3=22{,}5$ y el primer valor que supera el 22,5 en la frecuencia absoluta acumulada es el 25, que se corresponde con el valor 3 de obstáculos derribados.

SOLUCIONES PÁG. 179

7 El número de faltas personales que han cometido los jugadores en un partido de baloncesto son las siguientes:

Faltas	0	1	2	3	4	5
Frecuencia absoluta, n_i	1	2	1	4	5	3

Calcula la moda, la media aritmética y la mediana de estos valores.

x_i	Frecuencia absoluta, n_i	Frecuencia absoluta acumulada
0	1	1
1	2	3
2	1	4
3	4	8
4	5	13
5	3	16
	16	

La moda es M_o = 4 porque es el número de faltas que más veces se cometen.

La media es:

$$\overline{x}=\frac{(1\cdot 0)+(2\cdot 1)+(1\cdot 2)+(4\cdot 3)+(5\cdot 4)+(3\cdot 5)}{16}=\frac{51}{16}=3{,}18$$

La mediana se calcula con la media de los dos datos centrales:

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

$$M_e=\frac{3+4}{2}=3{,}5$$

PARÁMETROS DE DISPERSIÓN

8 Las notas obtenidas por 30 alumnos son:

Notas	2	3	4	5	6	7	8	9	10
N.º alumnos	2	1	7	4	1	6	5	3	1

a. Calcula los parámetros de centralización.

Se recogen los datos en una tabla:

Notas, x_i	Frecuencia absoluta, n_i	$n_i \cdot x_i$	Frecuencia absoluta acumulada	$x_i - \bar{x}$	$\lvert x_i - \bar{x} \rvert$	$\lvert x_i - \bar{x} \rvert \cdot n_i$	$x_i^2 \cdot n_i$
2	2	4	2	–4	4	8	8
3	1	3	3	–3	3	3	9
4	7	28	10	–2	2	14	112
5	4	20	14	–1	1	4	100
6	1	6	15	0	0	0	36
7	6	42	21	1	1	6	294
8	5	40	26	2	2	10	320
9	3	27	29	3	3	9	243
10	1	10	30	4	4	4	100
	30	180			20	58	1222

La moda es $M_o = 4$ porque es la nota que más veces se repite.

La media es:

$$\bar{x} = \frac{x_1 \cdot n_1 + \ldots + x_n \cdot n_n}{N} = \frac{180}{30} = 6$$

La mediana se calcula con la media de los dos datos centrales:

2 2 3 4 4 4 4 4 4 4 5 5 5 5 6 7 7 7 7 7 7 8 8 8 8 8 9 9 9 10

$$M_e = \frac{7+6}{2} = 6{,}5$$

$Q_1 = 4$ porque $\frac{N}{4} = \frac{30}{4} = 7{,}5$ y el primer valor que supera el 7,5 en la frecuencia absoluta acumulada es el 10, que se corresponde con el valor 4 de calificación.

$Q_3 = 8$ porque $\frac{N}{4} \cdot 3 = \frac{30}{4} \cdot 3 = 22{,}5$, y el primer valor que supera el 22,5 en la frecuencia absoluta acumulada es el 26, que se corresponde con el valor 8 de calificación.

b. Halla los parámetros de dispersión.

El rango es la diferencia entre el valor máximo y el mínimo, es decir:

R = 10 – 2 = 8

La desviación media es:

$$\text{DM} = \frac{\sum_{i=1}^{n} |x_i - \overline{x}| \cdot \text{n}_i}{\text{N}} = \frac{8+3+14+4+0+6+10+9+4}{30} = \frac{58}{30} = 1{,}93$$

La varianza es:

$$\text{V}(x) = \frac{\sum_{i=1}^{n} x_i^2 \cdot \text{n}_i}{\text{N}} - \overline{x}^2 =$$

$$= \frac{8+9+112+100+36+294+320+243+100}{30} - 6^2 = 4{,}7$$

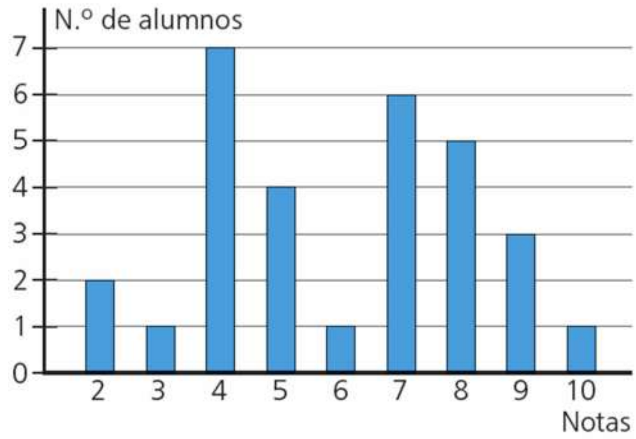
La desviación típica es:

$$\text{S}(x) = +\sqrt{\text{V}(x)} = \sqrt{4{,}7} = 2{,}17$$

El coeficiente de Pearson es;

$$\text{CV} = \frac{\text{S}(x)}{\overline{x}} = \frac{2{,}17}{6} = 0{,}36$$

c. Representa la información en un diagrama de barras.



d. ¿Es representativo el valor de la nota media obtenida?

No, porque el coeficiente de Pearson tiene un valor superior a 30%, concretamente del 36%, y eso significa que la media no es representativa.

9 Se ha realizado un control de masa, en kilogramos, a toda una manada de perros de trineo. Se han anotado los siguientes datos:

22,6 26,1 24,5 25,3 24,8 27,0 25,8 21,8 24,6 26,5
22,6 26,3 25,8 23,9 26,2 24,7 23,8 26,2 25,8 23,9

a. Ordena los datos en una tabla.

Se agrupan los datos por intervalos y se calculan los diferentes parámetros estadísticos:

Datos	c_i	n_i	$n_i \cdot c_i$	$c_i - \bar{x}$	$\lvert c_i - \bar{x}\rvert$	$\lvert c_i - \bar{x}\rvert \cdot n_i$	$c_i^2 \cdot n_i$	N_i
[21 ; 22)	21,5	1	21,50	–3	3	3	462,25	1
[22 ; 23)	22,5	2	45	–2	2	4	1012,5	3
[23 ; 24)	23,5	3	70,50	–1	1	3	1656,75	6
[24 ; 25)	24,5	4	98	0	0	0	2401	10
[25 ; 26)	25,5	4	102	1	1	4	2601	14
[26 ; 27)	26,5	6	159	2	2	12	4213,5	20
		20	496			26	12347	

b. Calcula los parámetros de centralización.

El intervalo modal, es decir, el intervalo que con más frecuencia se repite es [26 ; 27), con una frecuencia de 6.

La media es:

$$\bar{x} = \frac{\sum c_i \cdot n_i}{N} = \frac{70{,}50 + 98 + 102 + 159}{20} = \frac{496}{20} = 24{,}8$$

El intervalo mediano deja su izquierda el 50 % de los datos, es decir: $\frac{N}{4} \cdot 2 = \frac{20}{4} \cdot 2 = 10$ Si se consulta en la columna de los intervalos de frecuencia acumulados, N_i, se ve que los valores superiores a 10 corresponden al acumulado de 14, es decir, al intervalo de datos de [25 ; 26). Es decir, Q_2 está en [25 ; 26).

El primer cuartil, Q_1, se localiza sabiendo que $\frac{N}{4} = \frac{20}{4} = 5$.

Si se consulta en la columna de los intervalos de frecuencia acumulados, N_i, se ve que los valores superiores a 5 corresponden al acumulado de 6, es decir, al intervalo de datos de [23 ; 24). Es decir, Q_1 está en [23 ; 24).

El tercer cuartil, Q_3, se localiza sabiendo que $\frac{N}{4} \cdot 3 = \frac{20}{4} \cdot 3 = 15$.

Si se consulta en la columna de los intervalos de frecuencia acumulados, N_i, se ve que los valores superiores a 15 corresponden al acumulado de 20, es decir, al intervalo de datos de [26 ; 27). Es decir, Q_3 está en [26 ; 27).

c. Calcula los parámetros de dispersión.

El rango es la diferencia entre el valor máximo y el mínimo,
R = 27 – 21,8 = 5,2
La desviación media es:

$$DM = \frac{\sum_{i=1}^{n} \lvert c_i - \bar{x} \rvert \cdot n_i}{N} = \frac{3 + 4 + 3 + 0 + 4 + 12}{20} = \frac{26}{20} = 1{,}3$$

La varianza es:

$$V(x) = \frac{\sum_{i=1}^{n} c_i^2 \cdot n_i}{N} - \bar{x}^2 =$$

$$= \frac{462{,}25 + 1012{,}5 + 1656{,}75 + 2401 + 2601 + 4213{,}5}{20} - 24{,}8^2 = 2{,}31$$

La desviación típica es:

$S(x) = S(x) = +\sqrt{V(x)} = \sqrt{2{,}31} = 1{,}52$

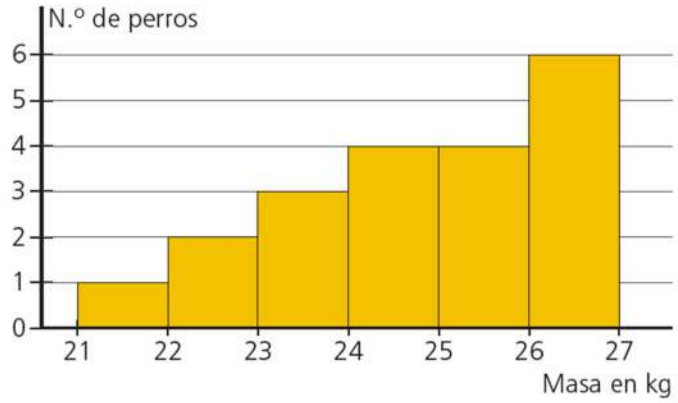
El coeficiente de Pearson es:

$$CV = \frac{S(x)}{\overline{x}} = \frac{1{,}52}{24{,}8} = 0{,}06$$

El CV es del 6 %, lo que significa que la masa media es bastante representativa de esta población.

d. Representa los datos en un histograma.

En el eje de abscisas se representan los datos de masa en kg, por intervalos (primera columna de la tabla anterior) y en el de ordenadas se representa la frecuencia de cada intervalo (columna n_i de la tabla).



Comprueba las soluciones usando una hoja de cálculo Excel.

Respuesta abierta.

10 Calcula la media aritmética, la desviación típica y el coeficiente de variación de estas dos distribuciones:

a. 10, 8, 6, 5, 7, 8

Notas, x_i	Frecuencia absoluta, n_i	$n_i \cdot x_i$	Frecuencia absoluta acumulada	$x_i^2 \cdot n_i$
5	1	5	1	25
6	1	6	2	36
7	1	7	3	49
8	2	16	5	128
10	1	10	6	100
	6	44		338

La media es:

$$\overline{x} = \frac{x_1 \cdot n_1 + \ldots + x_n \cdot n_n}{N} = \frac{5+6+7+16+10}{6} = \frac{44}{6} = 7{,}33$$

Para calcular la desviación típica se tiene que conocer primero la varianza:

$$V(x) = \frac{\sum_{i=1}^{n} x_i^2 \cdot n_i}{N} - \overline{x}^2 = \frac{25+36+49+128+100}{6} - (7{,}33)^2 = 2{,}6$$

La desviación típica es:

$S(x) = S(x) = +\sqrt{V(x)} = \sqrt{2,6} = 1,61$

El coeficiente de variación o coeficiente de Pearson es:

$$CV = \frac{S(x)}{\overline{x}} = \frac{1,61}{7,33} = 0,22$$

b. 16, 21, 16, 18, 17

Notas, x_i	Frecuencia absoluta, n_i	$n_i \cdot x_i$	Frecuencia absoluta acumulada	$x_i^2 \cdot n_i$
16	2	32	2	512
17	1	17	3	289
18	1	18	4	324
21	1	21	5	441
	5	88		1566

$$\overline{x} = \frac{x_1 \cdot n_1 + \ldots + x_n \cdot n_n}{N} = \frac{32+17+18+21}{5} = \frac{88}{5} = 17,6$$

Para calcular la desviación típica se tiene que conocer primero la varianza:

$$V(x) = \frac{\sum_{i=1}^{n} x_i^2 \cdot n_i}{N} - \overline{x}^2 = \frac{512+289+324+441}{5} - (17,6)^2 = 3,44$$

La desviación típica es:

$S(x) = S(x) = +\sqrt{V(x)} = \sqrt{3,44} = 1,85$

El coeficiente de variación o coeficiente de Pearson es:

$$CV = \frac{S(x)}{\overline{x}} = 0,10$$

¿Cuál de ellas presenta mayor dispersión?

Están más dispersos los datos de la distribución a., porque el coeficiente de Pearson es superior, un 22 % frente a un 10 % de la distribución b.

DIAGRAMA DE CAJA Y BIGOTES

11 Elabora un diagrama de caja y bigotes con estos datos referentes a la nota de Música de 28 alumnos e interprétalo:

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

Notas, x_i	Frecuencia absoluta, n_i	Frecuencia absoluta acumulada
1	1	1
2	5	6
3	8	14
4	4	18
5	2	20
6	2	22
7	1	23
8	2	25
9	2	27
10	1	28
	28	

Se calculan los cuartiles y el rango intercuartílico:

$\frac{N}{4} = \frac{28}{4} = 7$, según la columna de frecuencias absolutas acumuladas, el valor que sobrepasa el 7 es 14, con lo que $Q_1 = 3$.

$\frac{N}{4} \cdot 2 = \frac{28}{4} \cdot 2 = 14$, según la columna de frecuencias absolutas acumuladas, el valor 14 queda entre los valores de los datos 3 y 4, por lo que el valor de la mediana se calcula haciendo la media de estos dos valores:

$$M_e = Q_2 = \frac{3+4}{2} = 3{,}5$$

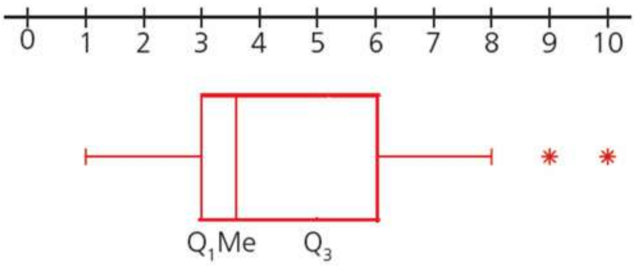
$\frac{N}{4} \cdot 3 = \frac{28}{4} \cdot 3 = 21$, según la columna de frecuencias absolutas acumuladas, el valor que sobrepasa el 21 es 22, con lo que $Q_3 = 6$.

El rango intercuartílico es $R_i = Q_3 - Q_1 = 6 - 3 = 3$.

Para representar el diagrama de cajas y bigotes se tiene en cuenta que:

- El rectángulo se construye con el valor de $Q_1 = 3$ en el extremo izquierdo del rectángulo y el valor de $Q_3 = 6$ en el derecho.
- La frontera inferior es $f_i = Q_1 - 1{,}5 \cdot R = 3 - 1{,}5 \cdot 3 = -1{,}5$. La frontera superior es $f_s = Q_3 + 1{,}5 \cdot R_i = 6 + 1{,}5 \cdot 3 = 10{,}5$.
- Los valores de los datos mínimo y máximo son, respectivamente $x_i = 1$ y $x_i = 10$ y están dentro del intervalo de las fronteras, que es (–1,5 , 10,5), la longitud de los bigotes llega hasta los valores mínimo y máximo, y no hace falta representar el valor de las fronteras.
- El rango intercuartílico, $R_i = 3$ coincide con la longitud de la caja.
- La mediana, $M_e = Q_2 = 3{,}5$ divide al rectángulo en dos partes.

El 25 % de los datos más bajos están más concentrados que el último 25 % de los datos. Los datos comprendidos entre el 25 % y el 50 % de la población están menos dispersos (el bigote es más corto a la izquierda) que entre el 50 % y el 75 % (el bigote es más largo a la derecha). El 50% de la población está comprendido entre 3 y 6. Aparecen los puntos atípicos 9 y 10.



EVALUACIÓN

1 La variable «tiempo de espera para revisar un coche en la ITV» es del tipo:

a. Cualitativa.

b. Cuantitativa discreta, porque el tiempo de espera se puede expresar numéricamente en horas y minutos.

c. Cuantitativa continua.

d. Ninguna de las anteriores.

2 Indica la frecuencia relativa del dato C en esta secuencia:

B C C A B D A B C A D C B A D C A C C B C D

a. 0,25

b. $f_i = \frac{n_i}{N} = \frac{8}{22} = 0{,}36$

c. 0,8

d. 0,42

3 ¿Cuál es el tamaño de la población sobre la que se ha realizado este

gráfico?

a. 5

b. La población es la suma de las frecuencias de todos los datos, es decir, $N = 3 + 6 + 4 + 7 + 5 = 25$

c. 7

d. 16

4 **Si en un diagrama de sectores a un dato le corresponde un sector de 36° de amplitud, ¿qué porcentaje tiene ese dato?**

a. La amplitud de un sector se obtiene multiplicando la frecuencia relativa por 360º, es decir, si el sector es de 36º significa que 36º = $p_i \cdot 360º \Longrightarrow p_i = 0{,}1$, o, expresado como porcentaje, p_i = 10 %.

b. 36 %

c. 25 %

d. 12 %

5 **Indica cuál es la mediana de los siguientes datos:**

9 2 3 8 5 7 9 3 10

a. 5

b. 6

c Se ordenan los datos de menor a mayor y se elige el dato central:

2 3 3 5 7 8 9 9 10

La mediana es 7.

d. 8

6 **Indica cuál es el tercer cuartil en esta distribución de datos:**

5 , 4 , 9 , 2 , 9 , 4 , 8 , 5 , 7 , 3 , 3 , 10

a. 5

b. 8,5

Se genera la tabla de datos:

x_i	Frecuencia absoluta, n_i	Frecuencia absoluta acumulada
2	1	1
3	2	3
4	2	5
5	2	7
7	1	8
8	1	9
9	2	11
10	1	12
	12	

Se observa que $\frac{N}{4} \cdot 3 = \frac{12}{4} \cdot 3 = 9$, y que en la columna de frecuencias acumuladas el valor 9 está entre los valores 8 y 9 de la columna x_i, de modo que se debe hacer la media de estos dos valores:

$$Q_3 = \frac{8+9}{2} = 8{,}5$$

c. 7

d. 4

7 **Se han registrado 20 mediciones en litros por metro cuadrado de las precipitaciones de lluvia caídas en toda España:**

12,5 16,1 20,3 6,2 18,8 35,4 27,0 29,6 17,3 10,1

13,7 22,3 15,6 31,4 19,9 11,3 25,5 26,8 23,4 26,7

¿Cuál es la media si agrupamos los datos en cinco intervalos comenzando en 6, de amplitud 6 cada uno?

a. 19,65

b. 15

c. Se genera la tabla de datos:

Datos	c_i	n_i	$n_i \cdot c_i$
[6 ;12)	9	3	27
[12 ; 18)	15	5	75
[18 ; 24)	21	5	105
[24 ; 30)	27	5	135
[30 ; 36)	33	2	66
		20	408

Se calcula la media acorde a esos intervalos:

porque $\bar{x} = \frac{\sum c_i \cdot n_i}{N} = \frac{27+75+105+135+66}{20} = \frac{408}{20} = 20{,}4$

d. 17,86