EJERCICIOS PROPUESTOS

13.1 Clasifica los siguientes caracteres estadísticos:

a) El récord del mundo de 100 metros lisos.

b) El número de hermanos que tienes.

c) Tu asignatura favorita.

d) La producción de hierro de una mina.

a) Cuantitativo continuo.

b) Cuantitativo discreto.

c) Cualitativo.

d) Cuantitativo continuo.

13.2 Propón un ejemplo en cada caso.

a) Carácter estadístico cualitativo.

b) Variable estadística discreta.

c) Variable estadística continua.

a) El medio de transporte que utilizas para ir al instituto.

b) El número de películas que ves al mes.

c) Tu altura.

13.3 Para realizar un sondeo sobre intención de voto en unas elecciones autonómicas, una empresa de demoscopia ha realizado entrevistas mediante llamadas telefónicas a 1000 personas, mediante llamadas telefónicas con el fin de conocer las expectativas de voto de un candidato.

a) ¿Te parece que éste es un buen procedimiento?

b) ¿Qué ocurre si una persona tiene tres líneas telefónicas a su nombre?

c) ¿Qué sucede si alguien no tiene teléfono?

a) No, ya que no todos tienen la misma probabilidad de ser elegidos para la muestra.

b) Una persona con tres líneas telefónicas a su nombre tiene el triple de posibilidades de ser elegida para la muestra que una que solo tiene una línea.

c) Si una persona no tiene teléfono, nunca podrá ser seleccionada para formar parte de la muestra.

13.4 El departamento de control de calidad de una fábrica realiza con cierta periodicidad mediciones sobre los productos que salen de su cadena de fabricación para comprobar si cumplen las características que figuran en el etiquetado del envase.

a) ¿Sería bueno probar con los 100 primeros productos que salen del proceso de fabricación a primera hora de la mañana?

b) ¿Estaría mejor examinar con los 100 últimos productos fabricados al final del día?

c) ¿Cómo podrías mejorar la elección de la muestra?

a) No, pues al comienzo del día puede estar la máquina bien, estropearse a la segunda hora y seguir estropeada todo el día.

b) No, porque la máquina puede funcionar bien a lo largo de toda la jornada y estropearse al final del día.

c) Tomando uno o dos productos cada hora para formar parte de la muestra.

13.5 Si sabemos que una muestra de 2500 personas es representativa de una población de 300 000, y que de esas 2500 personas, 500 son menores de edad, ¿cuántos menores de edad hay en la población?

Si de un total de 2500 personas de la muestra representativa, 500 son menores de edad, de los 300 000 de la población deberá haber:

$$\frac{2500}{500} = \frac{300\,000}{x} \Rightarrow x = 60\,000 \text{ menores de edad}$$

13.6 Un 65% de los 15 000 socios colaboradores de una ONG son mujeres. La dirección de la ONG quiere conocer la opinión de sus socios y realiza telefónicamente una encuesta a 1000 de ellos, de los cuales 450 son mujeres. ¿Es suficientemente representativa la muestra elegida?

En la muestra debe haber un 65% de mujeres.

65% de 1000 $= 650 > 450$, por lo que la muestra no es representativa.

Para que fuera representativa debería haber 650 mujeres.

13.7 En el estudio de una variable *X* se obtuvo la siguiente distribución de frecuencias.

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

Construye la tabla completa de frecuencias.

x_i	f_i	h_i	F_i	H_i
2	1	0,05	1	0,05
3	7	0,35	8	0,4
4	3	0,15	11	0,55
5	4	0,2	15	0,75
6	5	0,25	20	1
	20	1		

13.8 Las puntuaciones conseguidas en un test de cultura general realizado a 45 estudiantes fueron:

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

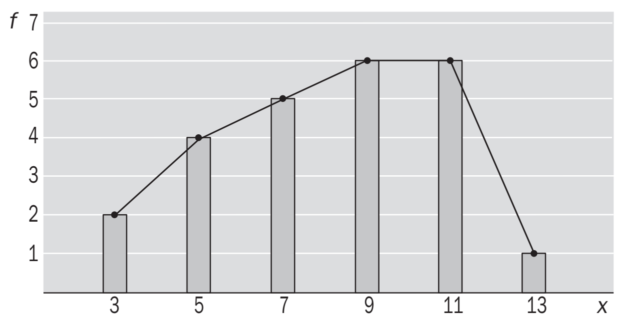
Construye la tabla completa de frecuencias.

Puntuación	f_i	h_i	F_i	H_i
1	6	0,133	6	0,133
2	4	0,089	10	0,222
3	6	0,133	16	0,356
4	4	0,089	20	0,444
5	6	0,133	26	0,578
6	2	0,044	28	0,622
7	4	0,089	32	0,711
8	8	0,178	40	0,889
9	5	0,111	45	1
	45	1		

13.9 En el estudio de una variable *X* se obtuvo la siguiente distribución de frecuencias.

x_i	3	5	7	9	11	13
f_i	2	4	5	6	6	1

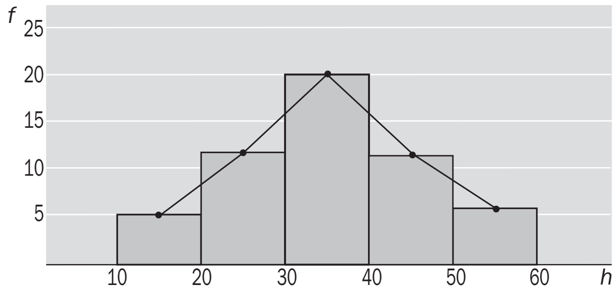
Construye el diagrama de barras y su polígono de frecuencias.



13.10 En el estudio de una variable *X* se obtuvo la siguiente distribución de frecuencias.

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

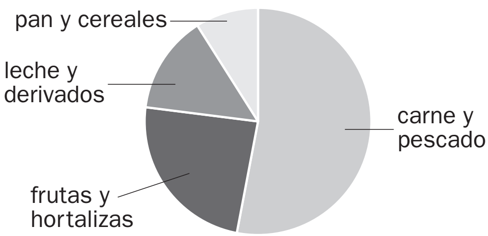
Construye el histograma y su polígono de frecuencias.



13.11 Este diagrama de sectores refleja, en porcentajes, la composición de la cesta de la compra en los hogares españoles.

a) ¿Qué tipo de alimento es el más consumido?

b) ¿Y el que menos?



a) La carne y el pescado.

b) El pan y los cereales.

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

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

Calcula la media aritmética, la moda y los cuartiles de la distribución.

$$\bar{x} = \frac{0 \cdot 10 + 1 \cdot 7 + 2 \cdot 6 + 3 \cdot 2 + 4 \cdot 1 + 5 \cdot 4}{10 + 7 + 6 + 2 + 1 + 4} \approx 1,63 \text{ faltas}$$

$M_o = 0$ faltas

N.° de faltas	f_i	F_i
0	10	$10 > 7,5$
1	7	$17 > 15$
2	6	$23 > 22,5$
3	2	25
4	1	26
5	4	30
	30	

$Q_1 = 0$ faltas

$Q_2 = 1$ falta

$Q_3 = 2$ faltas

13.13 La tabla adjunta muestra los resultados de unos alumnos en la prueba de salto con pértiga.

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

Calcula la media aritmética, la moda y los cuartiles de la distribución.

Número de alumnos	Marca de clase	h_i	F_i
[2; 2,5)	2,25	6	6
[2,5; 3)	2,75	12	18 > 9,25
[3; 3,5)	3,25	15	33 > 27,75
[3,5; 4)	3,75	4	37
		37	

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

$M_o = 3{,}25$ m

$Q_1 = [2{,}5; 3)$

$Q_2 = Q_3 = [3; 3{,}5)$

13.14 Calcula la media aritmética, la varianza y la desviación típica de la distribución dada por la siguiente tabla.

x_i	3	4	5	6	7	8	9
f_i	2	5	11	15	10	6	2

$$\bar{x} = \frac{3 \cdot 2 + 4 \cdot 5 + 5 \cdot 11 + 6 \cdot 15 + 7 \cdot 10 + 8 \cdot 6 + 9 \cdot 2}{2 + 5 + 11 + 15 + 10 + 6 + 2} \approx 6{,}019$$

$$s^2 = \frac{2 \cdot 3^2 + 5 \cdot 4^2 + 11 \cdot 5^2 + 15 \cdot 6^2 + 10 \cdot 7^2 + 6 \cdot 8^2 + 2 \cdot 9^2}{2 + 5 + 11 + 15 + 10 + 6 + 2} - 6{,}02^2$$

$$s = +\sqrt{1{,}96} = 1{,}4$$

13.15 Dados los datos de la siguiente tabla:

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

Calcula la media aritmética, la varianza y la desviación típica de la distribución asociada.

$$\bar{x} = \frac{15 \cdot 5 + 25 \cdot 12 + 35 \cdot 20 + 45 \cdot 11 + 55 \cdot 6}{5 + 12 + 20 + 11 + 6} = 35{,}185$$

$$s^2 = \frac{15^2 \cdot 5 + 25^2 \cdot 12 + 35^2 \cdot 20 + 45^2 \cdot 11 + 55^2 \cdot 6}{5 + 12 + 20 + 11 + 6} - 35{,}185^2 \approx 124{,}05$$

$$s = +\sqrt{124{,}05} \approx 11{,}137$$

13.16 Dado el siguiente conjunto de datos:

63, 62, 60, 20, 65, 80, 82, 110, 70, 75,
73, 72, 108, 84, 78, 67, 19, 60, 61, 63

a) Calcula la media aritmética.

b) ¿Hay algún atípico? Si los hay, descartalos y vuelve a calcular la media aritmética.

a) Hay 20 datos:

$$\bar{x} = \frac{2\cdot 63+62+2\cdot 60+20+65+80+82+110+70+75+73+72+108+84+78+67+19+61}{20} = \frac{1372}{20} \approx 68{,}6$$

b) 19, 20, 108 y 110 son datos atípicos. Los eliminamos y volvemos a calcular la media aritmética:

$$\bar{x} = \frac{2\cdot 63+62+2\cdot 60+65+80+82+70+75+73+72+84+78+67+61}{16} = \frac{1115}{16} = 69{,}69$$

13.17 Calcula la media aritmética truncada al 20% del siguiente conjunto de datos:

600, 396, 490, 1604, 8,
606, 604, 594, 1246, 42

Hay 10 datos. Para calcular la media truncada al 20% debemos descartar el 20% de los datos superiores y el 20% de los inferiores. Es decir, hay que eliminar el 20% de 10 = 2 datos superiores y 2 datos inferiores. Por tanto, el conjunto de datos queda de la siguiente forma:

600, 396, 490, 606, 604, 594

La media truncada es:

$$\bar{x} = \frac{600 + 396 + 490 + 606 + 604 + 594}{6} \approx 548{,}33$$

13.18 Para los datos del ejercicio anterior, calcula los valores atípicos, aplicando el criterio dado en el epígrafe.

$Q_3 = 606$; $Q_1 = 396$
$Q_3 + 1{,}5 \cdot (Q_3 - Q_1) = 921$
$Q_1 - 1{,}5 \cdot (Q_3 - Q_1) = 81$

Entonces:

x es atípico por la derecha si $x > Q_3 + 1{,}5 \cdot (Q_3 - Q_1) = 921$. Los valores atípicos por la derecha son 1246 y 1604.
x es atípico por la izquierda si $x < Q_1 - 1{,}5 \cdot (Q_3 - Q_1) = 81$. Los valores atípicos por la izquierda son 8 y 42.

13.19 Una distribución viene dada por la siguiente tabla:

x_i	3	4	5	6	7	8	9
f_i	2	5	11	15	10	6	2

Calcula la media y la desviación típica y halla el porcentaje de datos incluidos en los intervalos $(\bar{x} - s, \bar{x} + s)$; $(\bar{x} - 2s, \bar{x} + 2s)$; $(\bar{x} - 3s, \bar{x} + 3s)$.

Por el ejercicio 13.14 se sabe que $\bar{x} \approx 6{,}019$ y $s = 1{,}4$.

Intervalo $(\bar{x} - s, \bar{x} + s) = (4{,}619; 7{,}419)$
En este intervalo hay 11 + 15 + 10 = 36 datos. 36 de 51 datos totales representan el 70,6%.

Intervalo $(\bar{x} - 2s, \bar{x} + 2s) = (3{,}219; 8{,}819)$
En este intervalo hay 5 + 11 + 15 + 10 + 6 = 47 datos. 47 de 51 datos totales representan el 92,2%.

Intervalo $(\bar{x} - 3s, \bar{x} + 3s) = (1{,}819; 10{,}219)$
En este intervalo hay 2 + 5 + 11 + 15 + 10 + 6 + 2 = 51 datos. 51 de 51 datos totales representan el 100%.

13.20 En la tabla se recogen los datos correspondientes a una distribución estadística:

Clases	[1, 2)	[2, 3)	[3, 4)	[4, 5)	[5, 6)	[6, 7)	[7, 8)
f_i	6	12	15	20	16	11	6

Calcula la media y la desviación típica y halla el porcentaje de datos incluidos en los intervalos $(\bar{x} - s,\ \bar{x} + s)$; $(\bar{x} - 2s,\ \bar{x} + 2s)(\bar{x} - 3s,\ \bar{x} + 3s)$.

Clases	[1, 2)	[2, 3)	[3, 4)	[4, 5)	[5, 6)	[6, 7)	[7, 8)
Marca	1,5	2,5	3,5	4,5	5,5	6,5	7,5
f_i	6	12	15	20	16	11	6

$$\bar{x} = \frac{1{,}5 \cdot 6 + 2{,}5 \cdot 12 + 3{,}5 \cdot 15 + 4{,}5 \cdot 20 + 5{,}5 \cdot 16 + 6{,}5 \cdot 11 + 7{,}5 \cdot 6}{6 + 12 + 15 + 20 + 16 + 11 + 6} \approx 4{,}44$$

$$s = +\sqrt{\frac{1{,}5^2 \cdot 6 + 2{,}5^2 \cdot 12 + 3{,}5^2 \cdot 15 + 4{,}5^2 \cdot 20 + 5{,}5^2 \cdot 16 + 6{,}5^2 \cdot 11 + 7{,}5^2 \cdot 6}{6 + 12 + 15 + 20 + 16 + 11 + 6} - 4{,}44^2} = 1{,}77$$

Intervalo $(\bar{x} - s, \bar{x} + s) = (2{,}67; 6{,}21)$
En este intervalo hay 15 + 20 + 16 = 51 datos. 51 de 86 datos totales representan el 59%.

Intervalo $(\bar{x} - 2s, \bar{x} + 2s) = (0{,}9; 7{,}98)$
En este intervalo hay 6 + 12 + 15 + 20 + 16 + 11 = 80 datos. 80 de 86 datos totales representan el 93%.

Intervalo $(\bar{x} - 3s, \bar{x} + 3s) = (-0{,}87; 9{,}75)$
En este intervalo hay 6 + 12 + 15 + 20 + 16 + 11 + 6 = 86 datos. 86 de 86 datos totales representan el 100%.

13.21 Calcula el coeficiente de variación de la siguiente distribución.

x_i	1	2	3	4	5	6	7
f_i	5	12	18	11	7	4	1

$$\bar{x} = \frac{1 \cdot 5 + 2 \cdot 12 + 3 \cdot 18 + 4 \cdot 11 + 5 \cdot 7 + 6 \cdot 4 + 7 \cdot 1}{5 + 12 + 18 + 11 + 7 + 4 + 1} \approx 3{,}33$$

$$s = +\sqrt{\frac{1^2 \cdot 5 + 2^2 \cdot 12 + 3^2 \cdot 18 + 4^2 \cdot 11 + 5^2 \cdot 7 + 6^2 \cdot 4 + 7^2 \cdot 1}{5 + 12 + 18 + 11 + 7 + 4 + 1} - 3{,}33^2} \approx 1{,}41$$

$$CV = \frac{1{,}41}{3{,}33} = 0{,}42 = 42\%$$

13.22 Compara las dispersiones de las siguientes distribuciones.

x_i	6	3	4	3	7	5	6	8
y_i	63	39	64	55	66	70	65	62

$$\bar{x} = \frac{6 + 3 + 4 + 3 + 7 + 5 + 6 + 8}{8} = 5{,}25$$

$$s_x = +\sqrt{\frac{6^2 + 3^2 + 4^2 + 3^2 + 7^2 + 5^2 + 6^2 + 8^2}{8} - 5{,}25^2} \approx 1{,}71 \qquad CV_x = \frac{1{,}71}{5{,}25} = 0{,}33 \approx 33\%$$

$$\bar{y} = \frac{63 + 39 + 64 + 55 + 66 + 70 + 65 + 62}{8} = 60{,}5$$

$$s_y + \sqrt{\frac{63^2 + 39^2 + 64^2 + 55^2 + 66^2 + 70^2 + 65^2 + 62^2}{8} - 60{,}5^2} \approx 9{,}04 \qquad CV_y = \frac{9{,}04}{60{,}5} = 0{,}15 \approx 15\%$$

$CV_y < CV_x \Rightarrow$ La distribución Y es menos dispersa que la distribución X.

13.23 En ese mismo examen, Alberto (de 4.º A) ha sacado un 7, y Borja (de 4.º B), otro 7. Compara sus resultados.

Ordenando las notas de ambas clases, veremos que en la de Alberto hay cinco notas iguales o superiores a 7, y que en la de Borja hay 12.

Para comparar sus notas podemos calcular la puntuación típica teniendo en cuenta los parámetros que se han calculado en el problema resuelto. Restamos en cada caso la media del grupo y dividimos entre la desviación típica, obteniendo los siguientes valores:

$$\text{Alberto: } \frac{7 - 5{,}5}{1{,}23} = 1{,}22 \qquad \text{Borja: } \frac{7 - 5{,}5}{2{,}43} = 0{,}62$$

La puntuación de Alberto ha sido mejor, teniendo en cuenta los resultados del grupo.

13.24 En el siguiente examen de 4.º A, la nota media ha sido 5,8, y la desviación típica, 1,5. Ana ha vuelto a sacar un 6. Compara esta nota con la del primer examen.

Ahora, al no conocer las notas del resto de la clase, no podemos ordenarlas. Vamos a calcular su puntuación típica.

$$\text{Ana: } \frac{6 - 5{,}8}{1{,}5} = 0{,}13$$

La nota fue mejor en el primer examen.

ACTIVIDADES

EJERCICIOS PARA ENTRETENERSE

Caracteres, muestreo y recuento de datos

13.25 Di si los siguientes caracteres son cualitativos o cuantitativos, y relaciónalos con sus posibles modalidades o valores.

Carácter	Tipo	Modalidades/Valores
Estilo de veraneo		**1200, 2300, 800, 1000...**
N.º de mensajes de móvil		**Montaña, playa, aventura...**
Sueldo mensual de los trabajadores		**3, 15, 8, 12, 1, 7, 23, 9...**
N.º de libros que utiliza un alumno		**2, 6, 5, 3, 4, 1, 2, 4, 5...**

Carácter	Tipo	Modalidades/Valores
Estilo de veraneo	Cualitativo	Montaña, playa, aventura...
N.º de mensajes de móvil	Cuantitativo discreto	3, 15, 8, 12, 1, 7, 23, 9...
Sueldo mensual de los trabajadores	Cuantitativo continuo	1200, 2300, 800, 1000...
N.º de libros que utiliza un alumno	Cuantitativo discreto	2, 6, 5, 3, 4, 1, 2, 4, 5...

13.26 El consejo escolar de un centro desea realizar una encuesta para conocer el grado de satisfacción de los padres de los alumnos con el funcionamiento general del centro. ¿Cuál crees que es el mejor y más rápido procedimiento, y por qué?

A. Hacer el estudio con una clase elegida al azar.

B. En la salida del centro, elegir a los 30 primeros alumnos que salgan.

C. Seleccionar al azar 4 alumnos por grupo.

La opción C es la más adecuada, porque la muestra será más representativa del centro, al elegir alumnos de todos los grupos y niveles.

13.27 La siguiente tabla informa del uso de la sala de musculación de un polideportivo municipal.

Sexo	Renta
H 75%	Baja: 30%
	Media: 60%
	Alta: 10%
M 25%	Baja: 40%
	Media: 50%
	Alta: 10%

Se quiere realizar una encuesta a 40 personas sobre las instalaciones. ¿Cuántos usuarios de cada modalidad tendría que haber en la muestra?

- Hombres: 75% de 40 = 30 hombres, de los cuales:
  - Renta baja: 30% de 30 = 9
  - Renta media: 60% de 30 = 18
  - Renta alta: 10% de 30 = 3
- Mujeres: 25% de 40 = 10 mujeres, de las cuales:
  - Renta baja: 40% de 10 = 4
  - Renta media: 50% de 10 = 5
  - Renta alta: 10% de 10 = 1

13.28 Completa los valores de la tabla.

x_i	f_i	h_i	F_i	H_i
1	2			
2	4			
3	3			
4	7			
5	3			
6	1			
Total				

x_i	f_i	h_i	F_i	H_i
1	2	$\frac{1}{10}$	2	$\frac{1}{10}$
2	4	$\frac{1}{5}$	6	$\frac{3}{10}$
3	3	$\frac{3}{20}$	9	$\frac{9}{20}$
4	7	$\frac{7}{20}$	16	$\frac{4}{5}$
5	3	$\frac{3}{20}$	19	$\frac{19}{20}$
6	1	$\frac{1}{20}$	20	$\frac{20}{20}$
Total	20	1		

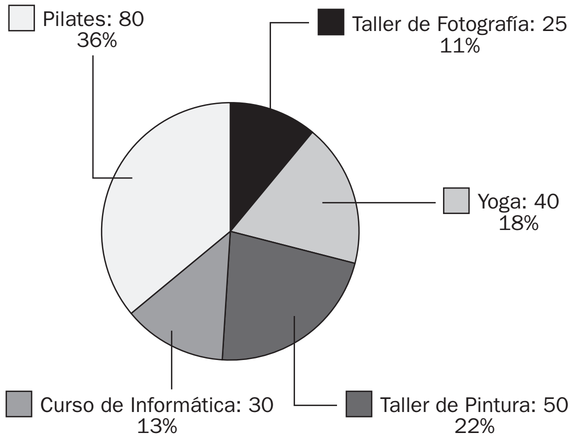
13.29 El número de llamadas perdidas al día que realizan 30 alumnos de 4.º ESO de un centro es:
30, 10, 17, 62, 57, 48, 74, 32, 47, 34, 12, 16, 15, 65, 22, 44, 38, 13, 36, 28, 38, 40, 61, 53, 52, 31, 27, 25, 20, 19
Agrupa los datos en clases y elabora una tabla de frecuencias.

Clases	f_i	h_i	F_i	H_i
[5, 15)	3	0,1	3	0,1
[15, 25)	6	0,2	9	0,3
[25, 35)	7	$0,2\overset{\frown}{3}$	16	$0,5\overset{\frown}{3}$
[35, 45)	5	$0,1\overset{\frown}{6}$	21	0,7
[45, 55)	4	$0,1\overset{\frown}{3}$	25	$0,8\overset{\frown}{3}$
[55, 65)	3	0,1	28	$0,9\overset{\frown}{3}$
[65, 75)	2	$0,0\overset{\frown}{6}$	30	1

Gráficos y parámetros

13.30 La siguiente tabla muestra las actividades ofertadas por un centro cultural y el número de vecinos que las realizan. Representa el diagrama de sectores asociado.

Actividades	N.º de vecinos
Yoga	**40**
Taller de pintura	**50**
Curso de informática	**30**
Pilates	**80**
Taller de fotografía	**25**



13.31 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 faltas	N.º de alumnos
[0, 4)	**8**
[4, 8)	**10**
[8, 12)	**8**
[12, 16)	**4**

a) Halla la media, la moda, la mediana y los otros dos cuartiles.

b) Calcula el rango, la varianza y la desviación típica.

c) Representa el histograma y el polígono de frecuencias.

a)

Marcas	f_i	F_i
2	8	$8 > 7{,}5$
6	10	$18 > 15$
10	8	$26 > 22{,}5$
14	4	30
	30	

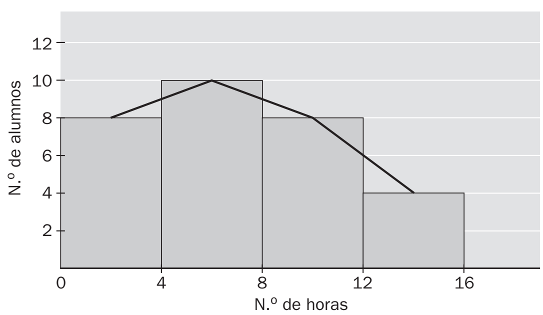
$$\bar{x} = \frac{2 \cdot 8 + 6 \cdot 10 + 102 \cdot 8 + 14 \cdot 4}{30} \approx 7{,}07$$

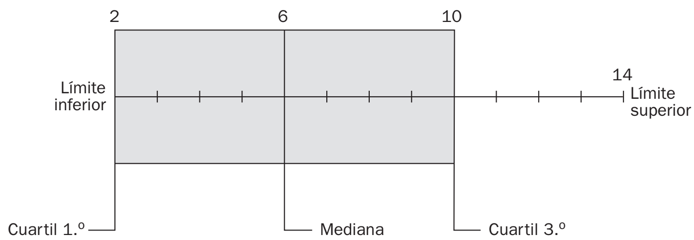
$M_o = 6$; $M = 6$; $Q_1 = 2$; $Q_3 \quad 10$

b) Rango $= 14 - 2 = 12$

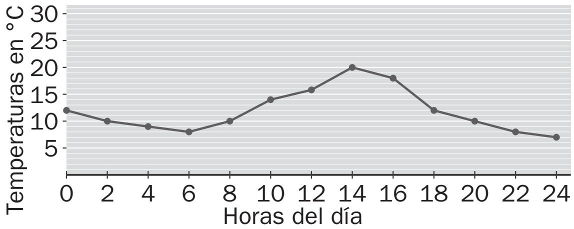
$$s^2 = \frac{2^2 \cdot 8 + 6^2 \cdot 10 + 10^2 \cdot 8 + 14^2 \cdot 4}{30} - 7{,}07^2 \approx 15{,}88 \Rightarrow s = +\sqrt{15{,}88} \approx 3{,}99$$

c)





13.32 La evolución de la temperatura a lo largo de un día de otoño en una ciudad del centro del país viene reflejada por el siguiente diagrama lineal.



Realiza otros dos diagramas lineales de lo que podría ser la evolución de la temperatura en un día de invierno y otro de verano en el mismo lugar.

Invierno

Verano

Valores atípicos

13.33 Los resultados obtenidos en un test de inteligencia en una población concreta han sido:

135, 125, 80, 140, 210, 156, 92, 141, 130, 128,
110, 230, 162, 60, 137, 118, 136, 205, 100, 143,
136, 128, 90, 143, 156, 157, 140, 125, 215, 130

a) Ordénalos en clases de amplitud 20.

b) Calcula sus cuartiles.

c) Halla los valores atípicos.

d) Calcula la media aritmética truncada al 20% y compárala con la media aritmética normal.

a)

Clases	Marca de clase	f_i	F_i
[60, 80)	70	1	1
[80, 100)	90	3	4
[100, 120)	110	3	7
[120, 140)	130	10	$17 > 7{,}5$; 15
[140, 160)	150	8	$25 > 22{,}5$
[160, 180)	170	1	26
[180, 200)	190	0	26
[200, 220)	210	3	29
[220, 240)	230	1	30

b) $Q_1 = Q_2 = 130$; $Q_3 = 150$

c) $Q_3 + 1{,}5 \cdot (Q_3 - Q_1) = 180 \cdot Q_1 - 1{,}5 \cdot (Q_3 - Q_1) = 100$

x es atípico por la derecha si $x > Q_3 + 1{,}5 \cdot (Q_3 - Q_1) = 180$. Valores atípicos por la derecha: 205, 210, 230 y 215.

x es atípico por la izquierda si $x < Q_1 - 1{,}5 \cdot (Q_3 - Q_1) = 100$. Valores atípicos por la izquierda: 60, 80, 90 y 92.

d) El 20% de 30 = 6.

Eliminamos seis datos superiores (157, 162, 205, 210, 215 y 230) y seis datos inferiores (60, 80, 90, 92 y 100).

La media es: $\bar{x} = \dfrac{70 \cdot 1 + 90 \cdot 3 + 110 \cdot 3 + 130 \cdot 10 + 150 \cdot 8 + 170 \cdot 1 + 210 \cdot 3 + 230 \cdot 1}{30} \approx 140$

La media truncada al 20% es: $\bar{x}_{truncada} = \dfrac{110 \cdot 2 + 130 \cdot 10 + 150 \cdot 7}{18} \approx 142{,}78$.

La media truncada es más representativa de la población.

Utilización conjunta de parámetros

13.34 Se ha realizado un estudio con el fin de averiguar la cantidad de toneladas de papel que recicla cada uno 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
64	1
65	2
67	3
68	6
72	3
74	3
80	1
85	1
	20

$$\bar{x} = \frac{64 + 65 \cdot 2 + 67 \cdot 3 + 68 \cdot 6 + 72 \cdot 3 + 74 \cdot 3 + 80 + 85}{20} = 70{,}3$$

$$s = +\sqrt{\frac{64^2 + 65^2 \cdot 2 + 67^2 \cdot 3 + 68^2 \cdot 6 + 72^2 \cdot 3 + 74^2 \cdot 3 + 80^2 + 85^2}{20} - 70{,}3^2} = 5{,}1$$

b) $(\bar{x} - 2s, \bar{x} + 2s) = (60{,}1;\ 80{,}5)$. En este intervalo hay 19 distritos, lo que supone un porcentaje del 95%.

13.35 La tabla recoge las temperaturas máximas alcanzadas en dos ciudades durante 10 días consecutivos del mes de agosto.

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

a) ¿Qué ciudad ha tenido una temperatura media más alta a lo largo de esos 10 días?

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

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

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

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

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

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.

13.36 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) La nota media de 4.º A es:

$$\overline{x}_A = \frac{0 \cdot 5 + 1 \cdot 4 + 2 \cdot 1 + 3 \cdot 0 + 4 \cdot 0 + 5 \cdot 0 + 6 \cdot 0 + 7 \cdot 0 + 8 \cdot 1 + 9 \cdot 4 + 10 \cdot 5}{5 + 4 + 1 + 1 + 4 + 5} = 5$$

La nota media de 4.º B es:

$$\overline{x}_A = \frac{2 \cdot 2 + 3 \cdot 2 + 4 \cdot 3 + 5 \cdot 6 + 6 \cdot 3 + 7 \cdot 2 + 8 \cdot 2}{20} = 5$$

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

$$s_A = +\sqrt{\frac{1^2 \cdot 4 + 2^2 \cdot 1 + 8^2 \cdot 1 + 9^2 \cdot 4 + 10^2 \cdot 5}{20} - 5^2} \approx 4{,}45$$

$$s_B = +\sqrt{\frac{2^2 \cdot 2 + 3^2 \cdot 2 + 4^2 \cdot 3 + 5^2 \cdot 6 + 6^2 \cdot 3 + 7^2 \cdot 2 + 8^2 \cdot 2}{20} - 5^2} \approx 1{,}7$$

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.

13.37 La media y la desviación típica de dos distribuciones *A* y *B* son las siguientes.

$$\overline{x}_A = 100, \quad \overline{x}_B = 200, \quad s_A = 1{,}2, \quad s_B = 2{,}1$$

Indica cuál de ellas es más dispersa.

Calculamos el coeficiente de variación de ambas:

$$CV_A = \frac{1{,}2}{100} = 0{,}012 \qquad CVB = \frac{2{,}1}{200} = 0{,}0105$$

Como $CV_A > CV_B$, la distribución *A* es más dispersa.

13.38 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 ($\overline{x} - 2s$, $\overline{x} + 2s$). Sustituyendo los datos del enunciado obtenemos el siguiente sistema de dos ecuaciones con dos incógnitas para la altura:

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

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

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

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

13.39 ¿Puede ser que la media no coincida con ningún valor de la variable? ¿Y la moda? Razona tus respuestas.

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.

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

7 12 15 22 23 28 32

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

13.41 ¿Qué habría de ocurrir para que la media aritmética de una variable estadística fuese cero?
Pon un ejemplo de una distribución que verifique la propiedad anterior.

Como $\bar{x} = \dfrac{x_1 \cdot f_1 + x_2 \cdot f_2 + ... + x_n \cdot f_n}{N}$, para que $\bar{x} = 0$ se ha de verificar que $x_1 \cdot f_1 + \dots + x_n \cdot f_n = 0$.

Las frecuencias son siempre mayores o iguales que cero, por lo que debe ocurrir que la variable tome valores positivos y negativos.

Ejemplo: Las temperaturas mínimas recogidas en una ciudad a lo largo de una semana del mes de enero han sido: 0°, −3°, 2°, 0°, 1°, −1°, 1°. Por tanto, $\bar{x} = 0$, es decir, la temperatura media a lo largo de esa semana fue de 0°.

13.42 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°.

13.43 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.

13.44 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.

13.45 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.

13.46 Fíjate en la tabla que recoge la distribución de cierta variable estadística y responde:

x_i	F_i
0	**1**
5	**9**
7	**18**
9	**28**
11	**39**
18	**40**

a) ¿Cuántos valores atípicos hay?

b) ¿Cuáles son los valores? Justifica tu respuesta.

c) Calcula la media truncada sin esos valores atípicos.

a) Calculamos los tres cuartiles.

x_i	F_i
0	1
5	9
7	$18 > 10$
9	$28 > 20$
11	$39 > 30$
18	40

$Q_1 = 7$; $Q_2 = 9$; $Q_3 = 11$

$Q_3 + 1{,}5 \cdot (Q_3 - Q_1) = 17 \cdot Q_1 - 1{,}5 \cdot (Q_3 - Q_1) = 1$

Entonces:

x es atípico por la derecha si $x > 17$. Valores atípicos por la derecha: 18. Hay 40 valores.

x es atípico por la izquierda si $x < 1$. Valores atípicos por la izquierda: 0. Hay 1 valor.

b) Del apartado a, los valores son 18 y 0.

c) Se calcula primero la frecuencia de cada dato:

x_i	f_i
0	1
5	8
7	9
9	10
11	11
18	1

La media truncada es:

$$\bar{x}_{truncada} = \frac{5 \cdot 8 + 7 \cdot 9 + 9 \cdot 10 + 11 \cdot 11}{38} = 8{,}26$$

PROBLEMAS PARA APLICAR

13.47 En una ciudad, 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 sobre los hábitos de la población de esa ciudad, se elige una muestra de 1200 habitantes. ¿Cuál de las tres muestras siguientes es la más representativa de la población?

	Muestra 1		Muestra 2		Muestra 3	
	H	M	H	M	H	M
Niños	**63**	**156**	**58**	**160**	**70**	**145**
Adultos	**105**	**195**	**100**	**182**	**95**	**170**
Mayores	**252**	**429**	**262**	**438**	**255**	**465**

Para que la muestra sea representativa de la población total se deben distribuir de la siguiente manera:

- Hombres: 35% de 1200 = 420, de los cuales:
  - Niños: 15% de 420 = 63
  - Adultos: 25% de 420 = 105
  - Mayores: 60% de 420 = 252
- Mujeres: 65% de 1200 = 780, de las cuales:
  - Niñas: 20% de 780 = 156
  - Adultos: 25% de 780 = 195
  - Mayores: 55% de 780 = 429

Por lo que la muestra más representativa es la 1.

13.48 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?

Variable *A*	
x_i	f_i
138	**2**
254	**1**
351	**1**
2	**30**

Variable *B*	
y_i	f_i
3	**8**
7	**6**
12	**7**
21	**5**

Variable *A*

$$\bar{x} = \frac{138 \cdot 2 + 254 \cdot 1 + 351 \cdot 1 + 2 \cdot 30}{2 + 1 + 1 + 30} = \frac{941}{34} \approx 27{,}68 \qquad s = +\sqrt{\frac{2 \cdot 138^2 + 254^2 \cdot 1 + 351^2 \cdot 1 + 2^2 \cdot 30}{2 + 1 + 1 + 30} - 27{,}68^2} \approx 76{,}67$$

$$M_o = 2$$

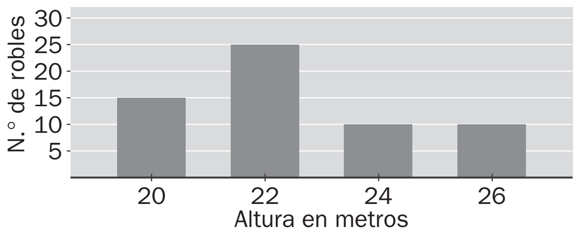
En esta variable, la varianza es muy grande, lo cual demuestra que la media no es muy representativa. Si nos fijamos en la tabla, se observa que de 34 datos hay 30 que son 2. Con lo cual la moda es mucho más representativa de la distribución.

Variable *B*

$$\bar{y} = \frac{3 \cdot 8 + 7 \cdot 6 + 12 \cdot 7 + 21 \cdot 5}{8 + 6 + 7 + 5} \approx 9{,}81 \qquad s = +\sqrt{\frac{3^2 \cdot 8 + 7^2 \cdot 6 + 12^2 \cdot 7 + 21^2 \cdot 5}{8 + 6 + 7 + 5} - 9{,}81^2} \approx 6{,}44 \qquad M_o = 3$$

Observando la varianza y la tabla, en esta variable, la media es bastante representativa, más que la moda.

13.49 En el Parque Nacional de los Picos de Europa se ha realizado un estudio sobre la altura de sus robles, y para ello se ha tomado una muestra en una superficie de 15 kilómetros cuadrados, obteniéndose los siguientes resultados gráficos.



a) Halla la media, la moda y la mediana.

b) Calcula el intervalo que contenga el 95% de los robles.

a) De la gráfica obtenemos la siguiente tabla:

x_i	f_i	F_i	$x_i f_i$	$x_i^2 f_i$
20	15	15	300	6000
22	25	40 > 30	550	12 100
24	10	50	240	5760
26	10	60	260	6760
	60		1350	30 620

x_i = Altura de los robles

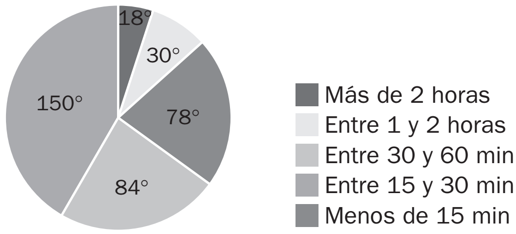
f_i = Número de robles

$$\bar{x} = \frac{20 \cdot 15 + 22 \cdot 25 + 24 \cdot 10 + 26 \cdot 10}{60} = 22{,}5; \quad ; M_o = 22; M_e = 22$$

b) En el intervalo ($\bar{x} - 2s$, $\bar{x} + 2s$) se encuentra al menos el 95% de los datos.

$$s = +\sqrt{\frac{20^2 \cdot 15 + 22^2 \cdot 25 + 24^2 \cdot 10 + 26^2 \cdot 10}{60} - 22{,}5^2} \approx 2{,}02 \Rightarrow (\bar{x} - 2s, \bar{x} + 2s) = (18{,}46; 26{,}54)$$

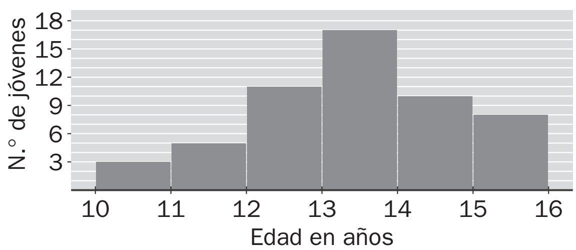
13.50 **Este diagrama recoge el tiempo que dedican diariamente al ordenador 60 alumnos de Secundaria.**



¿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

13.51 **Según un estudio realizado a nivel nacional, la edad media en que la población realiza su primer viaje al extranjero es de 13,6 años. En un estudio realizado en un centro de Secundaria a 54 jóvenes 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.° de alumnos	Marcas x_i	F_i
[10, 11)	3	10,5	3
[11, 12)	5	11,5	8
[12, 13)	11	12,5	19
[13, 14)	17	13,5	36
[14, 15)	10	14,5	46
[15, 16)	8	15,5	54
	54		

$$\bar{x} = \frac{3 \cdot 10{,}5 + 5 \cdot 11{,}5 + 11 \cdot 12{,}5 + 17 \cdot 13{,}5 + 10 \cdot 14{,}5 + 8 \cdot 15{,}5}{54} \approx 13{,}43$$

Mediana: $M = 13{,}5$

b) La media nacional es de 13,6, así que el estudio ratifica que la muestra realiza su primer viaje al extranjero a menor edad que la media de la población.

13.52 En cierto centro escolar cuyo número total de alumnos es de 600, el 40% son chicos, de los cuales solo el 25% se encuentra en Bachillerato, mientras que el 50% de las chicas cursan ESO.

a) Detalla cómo ha de ser la estructura de una muestra significativa que conste de la quinta parte del alumnado.

b) En cierta muestra aleatoria simple de 50 alumnos de 2.º de ESO, se ha estudiado el número de sobresalientes por alumno en la primera evaluación, obteniéndose los siguientes resultados.

Número de sobresalientes	Número de alumnos
0	4
1	7
2	10
3	15
4	5
5	6
6	1
7	1
8	1

Calcula la media truncada al 10% de esta distribución.

a) El tamaño de la muestra ha de ser la quinta parte del total: $n = 120$.

Para que la muestra sea representativa tiene que haber:

- Chicos: 40% de 120 = 48, de los cuales:
  - Bachillerato: 25% de 48 = 12
  - ESO: 75% de 48 = 36
- Chicas: 60% de 120 = 72, de las cuales:
  - Bachillerato: 50% de 72 = 36
  - Adultos: 50% de 72 = 36

Resumiendo, la muestra debe tener la siguiente distribución:

	ESO	Bachillerato	Total
Chicos	36	12	48
Chicas	36	36	72
Total	72	48	120

b) Hay 50 datos en total. Como el 10% de 50 = 5, para calcular la media truncada al 10% hay que eliminar 5 datos superiores (5, 5, 6, 7 y 8) y 5 datos inferiores (0, 0, 0, 0 y 1).

$$\bar{x} = \frac{1 \cdot 6 + 2 \cdot 10 + 3 \cdot 15 + 4 \cdot 5 + 5 \cdot 4}{40} \approx 2{,}78$$

13.53 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:

a) La desviación típica de la distribución.

b) El número de habitantes que miden entre 178 y 190 centímetros.

a) 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:

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

b) 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

13.54 La siguiente tabla muestra las edades de los jóvenes que acuden a un bibliobús 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

a) Halla la media, la moda y el tercer cuartil.

b) Calcula la desviación típica.

c) Representa el histograma y el polígono de frecuencias.

a)

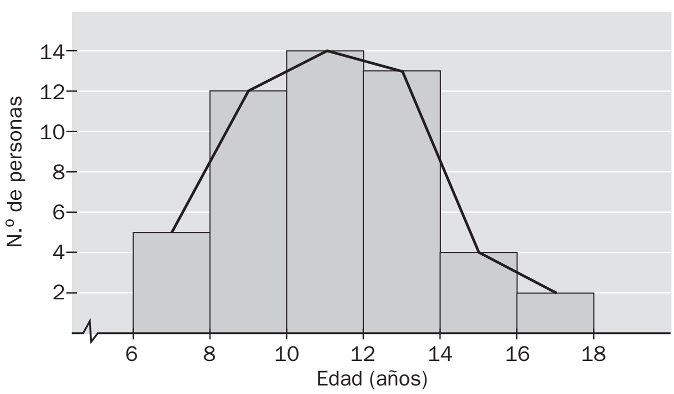
Marcas	*f*	*F*
7	5	5
9	12	17
11	14	31
13	13	44
15	4	48
17	2	50
	50	

$$\bar{x} = \frac{7 \cdot 5 + 9 \cdot 12 + 11 \cdot 14 + 13 \cdot 13 + 15 \cdot 4 + 17 \cdot 2}{50} = 11{,}2$$

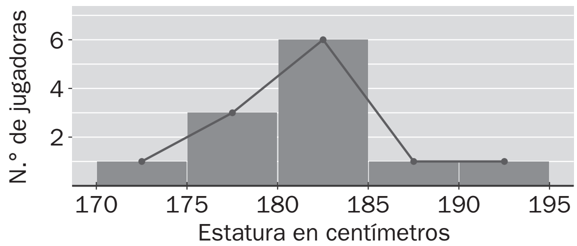
$M_o = 11$ años. $Q_3 = 13$ años

b) $$s = +\sqrt{\frac{7^2 \cdot 5 + 9^2 \cdot 12 + 11^2 \cdot 14 + 13^2 \cdot 13 + 15^2 \cdot 4 + 17^2 \cdot 2}{50} - 11{,}2^2} \approx 2{,}5$$

c)



13.55 El siguiente polígono de frecuencias 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)	Número de jugadoras
172,5	1
177,5	3
182,5	6
187,5	1
192,5	1
	12

$$\overline{x} = \frac{172{,}5 + 177{,}5 \cdot 3 + 182{,}5 \cdot 6 + 187{,}5 + 192{,}5}{12} \approx 181{,}67 \text{ cm}$$

$$s = +\sqrt{\frac{172{,}5^2 + 177{,}5^2 \cdot 3 + 182{,}5^2 \cdot 6 + 187{,}5^2 + 192{,}5^2}{12} - 181{,}67^2} \approx 4{,}81 \text{ cm}$$

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

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

¿Qué calificación debe obtener en el último examen que le queda para aprobar la asignatura con un 5?

Para aprobar necesita que su media sea 5, es decir:

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

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

Utilización conjunta de $\overline{x}$ y s

13.57 La siguiente tabla recoge 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) Equipo 1:

Número de goles	24	25	26	27
Número de partidos	3	2	1	2

$$\overline{x}_1 = \frac{24 \cdot 3 + 25 \cdot 2 + 26 + 27 \cdot 2}{8} = 25{,}25 \text{ goles}$$

Equipo 2:

Número de goles	20	21	22	27	28	30
Número de partidos	1	1	1	1	2	2

$$\overline{x}_2 = \frac{20 + 21 + 22 + 27 + 28 \cdot 2 + 30 \cdot 2}{8} = 25{,}75 \text{ goles}$$

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

$$s_1 = +\sqrt{\frac{24^2 \cdot 3 + 25^2 \cdot 2 + 26^2 + 27^2 \cdot 2}{8} - 25{,}25^2} \approx 1{,}2 \Rightarrow CV_1 = \frac{1{,}2}{25{,}25} \approx 0{,}05$$

$$s_2 = +\sqrt{\frac{20^2 + 21^2 + 22^2 + 27^2 + 28^2 \cdot 2 + 30^2 \cdot 2}{8} - 25{,}75^2} \approx 3{,}83 \Rightarrow CV_2 = \frac{3{,}83}{25{,}75} \approx 0{,}15$$

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

13.58 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:

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

AMPLIACIÓN

13.59 Completa la distribución 4, 6, 7, 7, 9, 9 con dos datos más, de forma que:

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 11.

b) Para que se conserve la media y disminuya la desviación típica, podemos añadir los datos 7 y 7, ya que la media es 7.

13.60 Dadas las distribuciones:

Variable *A*

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

Variable *B*

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

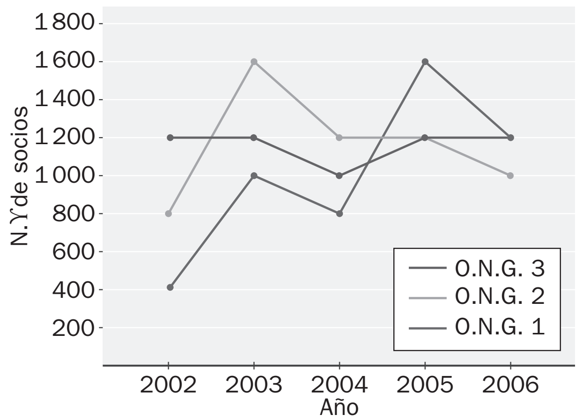
Variable *C*

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.

13.61 Los siguientes diagramas lineales muestran la evolución del número de socios de tres ONG en los últimos cinco años.



a) Calcula el coeficiente de variación de cada una de ellas.

b) ¿Cuál tiene mayor variabilidad en el número de socios? ¿Por qué?

a) De la gráfica se deduce la tabla de frecuencias para cada ONG.

ONG 1

x_i	f_i
400	1
800	1
1000	1
1200	1
1600	1
	5

$$\bar{x} = \frac{400 + 800 + 1000 + 1200 + 1600}{5} = 1000$$

$$s = +\sqrt{\frac{400^2 + 800^2 + 1000^2 + 1200^2 + 1600^2}{5} - 1000^2} = 400$$

$$CV = \frac{400}{1000} = 0,4 = 40\%$$

ONG 2

x_i	f_i
800	1
1000	1
1200	2
1600	1
	5

$$\bar{x} = \frac{800 + 1000 + 2 \cdot 1200 + 1600}{5} = 1160$$

$$s = +\sqrt{\frac{800^2 + 100^2 + 2 \cdot 1200^2 + 1600^2}{5} - 1160^2} = 265,32$$

$$CV = \frac{265,32}{1160} = 0,23 = 23\%$$

ONG 3

x_i	f_i
1000	1
1200	4
	5

$$\bar{x} = \frac{1000 + 4 \cdot 1200}{5} = 1160$$

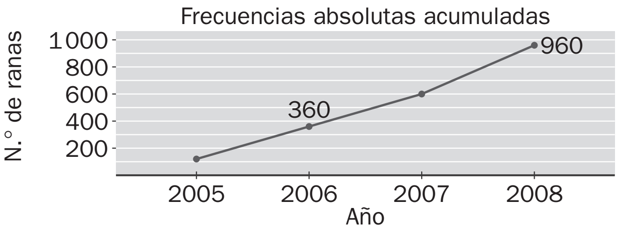
$$s = +\sqrt{\frac{1000^2 + 4 \cdot 1200^2}{5} - 1160^2} = 80$$

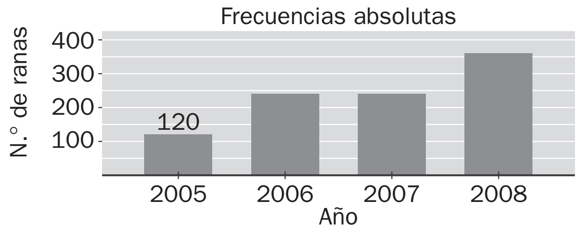
$$CV = \frac{80}{1160} = 0,07 = 7\%$$

b) La ONG 1 tiene mayor variabilidad, pues su coeficiente de variación es superior al de las otras dos ONG.

13.62 **Población de ranas**

Los siguientes gráficos estadísticos representan la misma distribución: la evolución de la población de ranas en un estanque durante los cuatro veranos que van desde el año 2005 hasta el 2008.





Completa la tabla con ayuda de los gráficos.

	F. absoluta	F. acumulada
2005	120	120
2006	240	360
2007	240	600
2008	360	960
Total	960	

13.63 **Media aritmética y media geométrica**

Como sabes, la media aritmética de dos números *a* y *b* se calcula mediante la expresión:

$$\bar{x} = \frac{a + b}{2}$$

Pero este valor no es el único parámetro estadístico de centralización. La media geométrica es también otro valor que sirve para representar los datos. Se calcula mediante la expresión:

$$x_g = \sqrt{a \cdot b}$$

a) Completa la siguiente tabla:

Datos		Media aritmética	Media geométrica
a	***b***		
1	**9**		
6	**24**		
5	**5**		
4	**0**		
−2	**18**		

b) Indica alguna desventaja que aprecies para la utilización de la media geométrica.

c) Recordando que el cuadrado de cualquier número es positivo o nulo, y el desarrollo de la expresión $\left(\sqrt{a} - \sqrt{b}\right)^2$, demuestra que la media geométrica de dos números es inferior o igual que la media aritmética de los mismos.

a)

Datos		Media aritmética	Media geométrica
a	***b***		
1	**9**	5	3
6	**24**	15	12
5	**5**	5	5
4	**0**	2	0
−2	**18**	8	No definida

b) La media geométrica no está definida si uno de los dos valores es negativo. Además, si uno de los datos es nulo, su valor siempre será cero.

c) $\left(\sqrt{a} - \sqrt{b}\right)^2 \geqslant 0 \Rightarrow \sqrt{a^2} + \sqrt{b^2} - 2\sqrt{ab} \geqslant 0 \Rightarrow a + b \geqslant 2\sqrt{ab} \Rightarrow \frac{a + b}{2} \geqslant \sqrt{ab}$

13.A1 **Esta tabla recoge los resultados de un atleta en las diez últimas pruebas de salto de longitud en que ha participado.**

Longitud de 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 los datos en un histograma.

a)

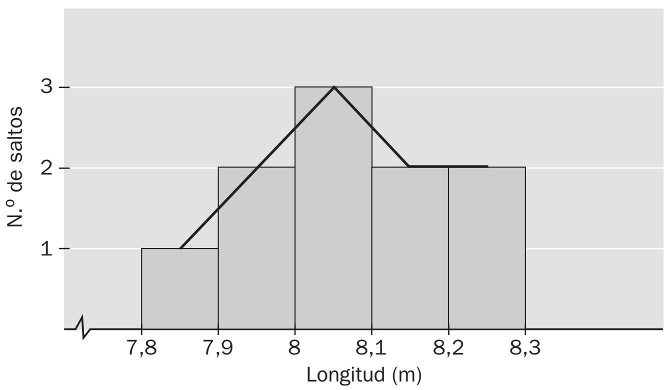
Longitud de salto (m)	Marca de clase	f_i	F_i
[7,8; 7,9)	7,85	1	1
[7,9; 8,0)	7,95	2	$3 > 2{,}5$
[8,0; 8,1)	8,05	3	$6 > 5$
[8,1; 8,2)	8,15	2	8
[8,2; 8,3)	8,25	2	10

$$\bar{x} = \frac{7{,}85 + 7{,}95 \cdot 2 + 8{,}05 \cdot 3 + 8{,}15 \cdot 2 + 8{,}25 \cdot 2}{1 + 2 + 3 + 2 + 2} = 8{,}07 \text{ m} \quad M_o = 8{,}05 \text{ m} \quad M = 8{,}05 \text{ m} \quad Q_1 = 7{,}95 \text{ m}$$

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

$$s = +\sqrt{\frac{7{,}85^2 + 7{,}95^2 \cdot 2 + 8{,}05^2 \cdot 3 + 8{,}15^2 \cdot 2 + 8{,}25^2 \cdot 2}{10} - 8{,}07^2} \approx 0{,}125$$

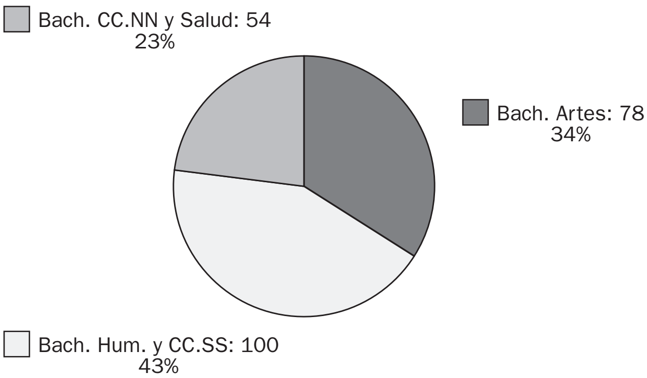
c) El histograma que se obtiene es el siguiente:



13.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.

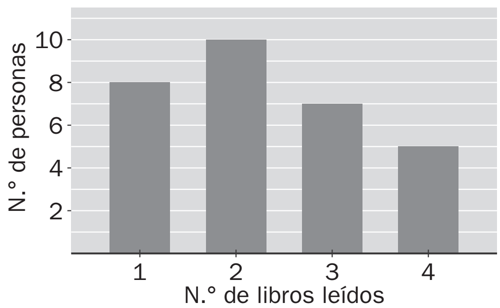
13.A3 En las aulas de Bachillerato de un centro escolar hay 100 alumnos en la modalidad de Humanidades y CC. SS., 54 en la de CC. NN. y de la Salud, y 78 en la de Artes.
Representa estos datos en un diagrama de sectores.



13.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)

Libros leídos x_i	N.° de personas f_i
1	8
2	10
3	7
4	5
	30

$$\bar{x} = \frac{1 \cdot 8 + 2 \cdot 10 + 3 \cdot 7 + 4 \cdot 5}{30} = 2{,}3$$

$$s = +\sqrt{\frac{1^2 \cdot 8 + 2^2 \cdot 10 + 3^2 \cdot 7 + 4^2 \cdot 5}{30} - 2{,}3^2} \approx 1{,}04$$

b) $CV = \dfrac{1{,}04}{2{,}3} \approx 0{,}45 = 45\%$

M A T E T I E M P O S

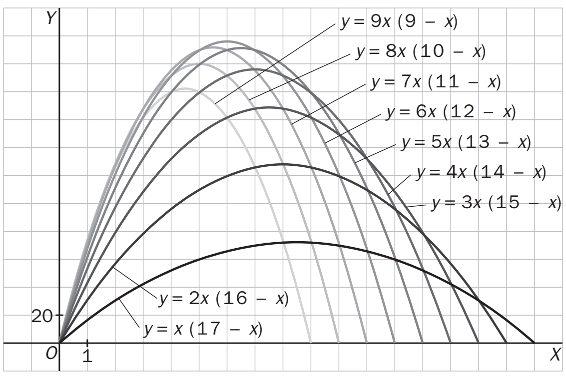
El equipaje de vuelo

Una compañía aérea tiene esta norma sobre la dimensión del equipaje que se puede transportar en sus aviones: "El equipaje de los pasajeros debe cumplir que la suma de sus tres dimensiones (largo, ancho y alto) no exceda de 1,80 metros". ¿Qué tamaño pueden tener las maletas para ser admitidas por esta compañía aérea?

El problema plantea la búsqueda de las dimensiones óptimas de un prisma sabiendo que la suma de las tres es de 1,80 m. Por comodidad, trabajaremos con decímetros, luego $a + b + c = 18 \Rightarrow c = 18 - a - b$.

De todos los prismas posibles, buscamos el de mayor volumen. Este será: $V = abc = ab(18 - a - b) = 18ab - a^2b - ab^2$.

b	1	2	3	4	5	...
$v(a)$	$a(17 - a)$	$2a(16 - a)$	$3a(15 - a)$	$4a(14 - a)$	$5a(13 - a)$	...



Si construimos una gráfica (con la ayuda de una calculadora gráfica) para cada valor de b entre 1 y 9, tenemos todas las posibilidades:

El máximo se produce cuando $a = 6$, $b = 6$ y $c = 18 - 6 - 6 = 6$. Luego las dimensiones óptimas serán:

$a = 6$ dm $= 60$ cm

$b = 6$ dm $= 60$ cm

$c = 6$ dm $= 60$ cm