HUMANIDADES Y CIENCIAS SOCIALES

SOLUCIONARIO

BLOQUE III

ESTADÍSTICA Y PROBABILIDAD

MATEMÁTICAS 1 aplicadas a las CIENCIAS SOCIALES

JOSE RAMÓN VIZMANOS

JOAQUÍN HERNÁNDEZ

FERNANDO ALCAIDE

sm

BACHILLERATO

PROYECTO EDITORIAL
Equipo de Educación Secundaria de Ediciones SM

AUTORES
José Ramón Vizmanos
Joaquín Hernández
Fernando Alcaide
María Moreno
Esteban Serrano

EDICIÓN
Rafaela Arévalo
Diana Cano
Miguel Ángel Ingelmo
Yolanda África Zárate

ILUSTRACIÓN
Félix Anaya
Juan Francisco Cobos
Modesto Arregui

DISEÑO DE CUBIERTAS E INTERIORES
Alfonso Ruano
Maritxu Eizaguirre

MAQUETACIÓN
Grafilia, SL

COORDINACIÓN EDITORIAL
Josefina Arévalo

DIRECCIÓN EDITORIAL
Aída Moya

ISBN:

Índice

ESTADÍSTICA Y PROBABILIDAD

Este solucionario se ha dividido en varios libros, uno por cada bloque de contenidos, para facilitar su uso por el profesor. En el primero de ellos se ha incluido una nueva versión de las hojas de soluciones que aparecen al final del libro del alumno, en la que se han corregido algunos errores y erratas que han sido detectados. El profesor puede distribuir estas páginas a sus alumnos si así lo considera conveniente.

(*) Una pequeña cantidad de ejercicios o apartados de ejercicios han sido marcados porque contienen alguna corrección en su enunciado respecto del que aparece en el libro del alumno.

11 Análisis estadístico de una variable

ACTIVIDADES INICIALES

11.I. Considera la población formada por los alumnos de tu clase. Para esta población, da cuatro ejemplos de caracteres estadísticos cualitativos, indicando algunas de sus modalidades.

Color del pelo: rubio, moreno, negro.

Profesión del padre: bombero, empresario.

Género de películas favorito: romántico, bélico, de acción.

Optativa que cursa: francés, Informática.

11.II. De una población formada por niños de dos años se extrae una muestra. Enumera cuatro caracteres estadísticos cuantitativos, indicando algunas de sus posibles valores.

Altura: 70 cm, 65 cm.

Peso: 15 kg, 20 kg.

Perímetro craneal: 40 cm, 45 cm.

Número de muelas: 3, 4.

EJERCICIOS PROPUESTOS

11.1. En el año 2003, el número de matrimonios de españoles en el extranjero, en los distintos continentes, viene dado por la siguiente tabla:

	N.º de matrimonios
Europa	5265
África	238
América	4804
Asia y Oceanía	299

Teniendo en cuenta estos datos, realiza los siguientes apartados:

a) Construye una tabla estadística que incluya frecuencias absolutas, frecuencias relativas y porcentajes.

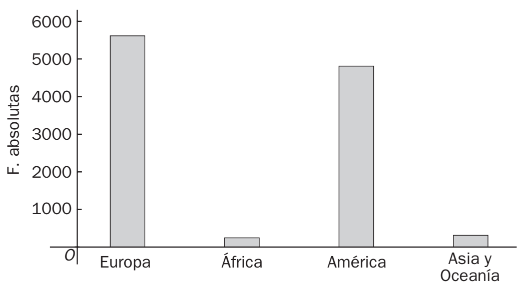
b) Dibuja el diagrama de barras.

c) Representa el diagrama de sectores para esta distribución, indicando el porcentaje de cada modalidad.

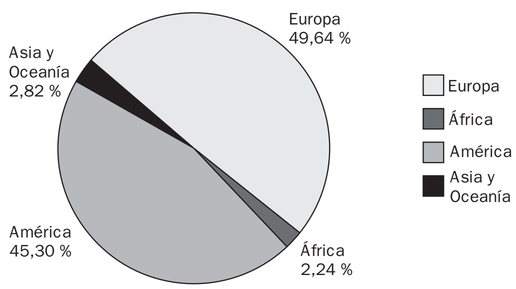
a)

Continente	f_i	h_i	p_i
Europa	5265	0,49642	49,64%
África	238	0,02244	2,24%
América	4804	0,45295	45,30%
Asia y Oceanía	299	0,02819	2,82%
	10 606		100

b) Diagrama de barras



c) Diagrama de sectores



11.2. El número de centros de salud en 20 ciudades es:

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

a) Construye la tabla de distribución de frecuencias de estos datos.

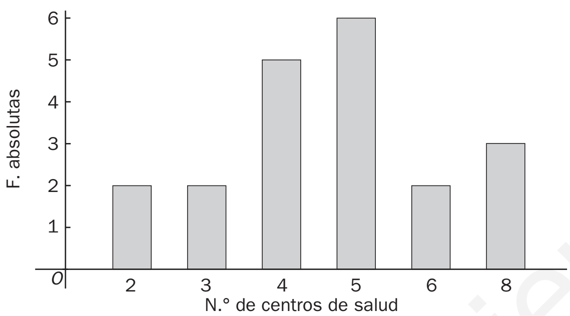
b) Representa el diagrama de barras de las frecuencias absolutas.

c) Representa el polígono de frecuencias absolutas acumuladas.

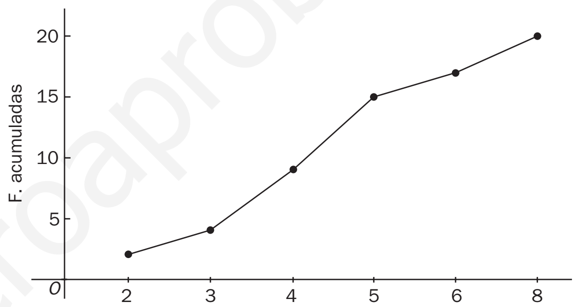
a)

N.º de centros de salud	f_i	F_i	h_i	H_i
2	2	2	0,1	0,1
3	2	4	0,1	0,2
4	5	9	0,25	0,45
5	6	15	0,3	0,75
6	2	17	0,1	0,85
7	3	20	0,15	1
	20		1	

b) Diagrama de barras



c) Polígono de frecuencias

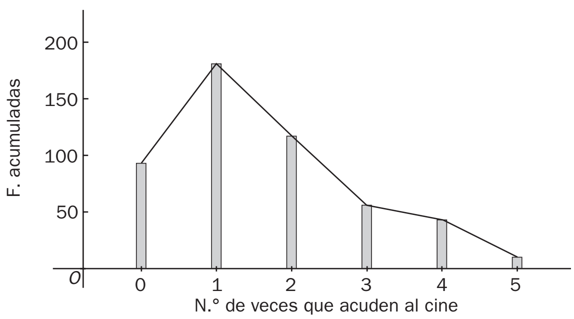


11.3. Se ha preguntado a 500 personas por el número de veces que van al cine durante un mes:

N.º de veces que acuden al cine	Frecuencias
0	93
1	181
2	117
3	56
4	43
5	10

Representa el diagrama de barras y el polígono de frecuencias absolutas.

Diagrama de barras + polígono de frecuencias



11.4. Los siguientes datos representan las alturas, en cm, de 20 personas:

165	171	154	165	149	159	151	171	191	163
173	193	176	152	188	169	171	184	152	183

a) Construye una tabla de distribución de frecuencias, agrupando los datos en intervalos de amplitud 10.

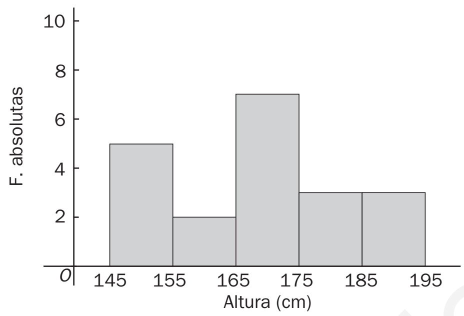
b) Representa el histograma de frecuencias absolutas.

c) Representa el polígono de frecuencias absolutas acumuladas

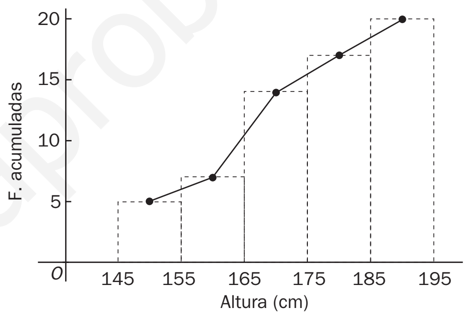
a)

$[L_i, L_s)$	x_i	f_i	F_i	h_i	H_i
[145, 155)	150	5	5	0,25	0,25
[155, 165)	160	2	7	0,1	0,35
[165, 175)	170	7	14	0,35	0,7
[175, 185)	180	3	17	0,15	0,85
[185, 195)	190	3	20	0,15	1
		20		1	

b) Histograma frecuencias absolutas



c) Polígono frecuencias absolutas acumuladas



11.5. El número de centros de salud en 20 ciudades es:

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

a) Calcula la media aritmética.

b) Halla la moda.

x_i	f_i	$x_i f_i$
2	2	4
3	2	6
4	5	20
5	6	30
6	2	12
8	3	24
	20	96

a) $\bar{x} = \dfrac{\sum_{i=1}^{n} x_i f_i}{N} = \dfrac{96}{20} = 4{,}8$

b) $M_o = 5$ centros de salud

11.6. Halla la mediana de las siguientes series estadísticas.

a) 1, 7, 3, 2, 4, 6, 2, 5, 6

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

a) 1, 2, 2, 3, $\boxed{4,}$ 5, 6, 6, 7 $M = 4$

b) 1, 1, 2, 3, $\boxed{3,\ 4,}$ 5, 6, 7, 8 La mediana es la media aritmética de los dos valores centrales, $M = 3{,}5$.

11.7. La tabla adjunta muestra las medidas, en cm, de unos bastones de esquí.

Medida en cm	[100, 105)	[105, 110)	[110, 115)	[115, 120)	[120, 125)
N.º de bastones	4	9	12	10	3

a) Halla la media aritmética y la moda.

b) Calcula la mediana.

c) Representa el histograma de frecuencias absolutas.

d) Dibuja el diagrama de sectores.

Formamos la siguiente tabla:

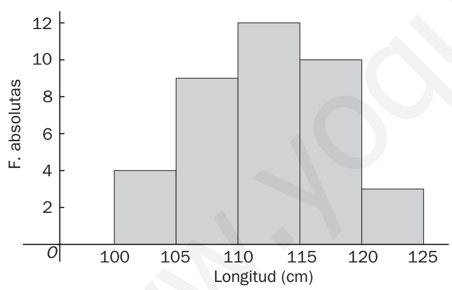
$[L_i, L_s)$	x_i	f_i	h_i	F_i	$x_i f_i$
[100, 105)	102,5	4	0,1053	4	410
[105, 110)	107,5	9	0,2368	13	967,5
[110, 115)	112,5	12	0,3158	25	13,50
[115, 120)	117,5	10	0,2632	35	11,75
[120, 125)	122,5	3	0,0789	38	367,5
		38	1		4270

a) $\bar{x} = \dfrac{\sum_{i=1}^{n} x_i f_i}{N} = \dfrac{4270}{38} = 112{,}37$ centímetros.

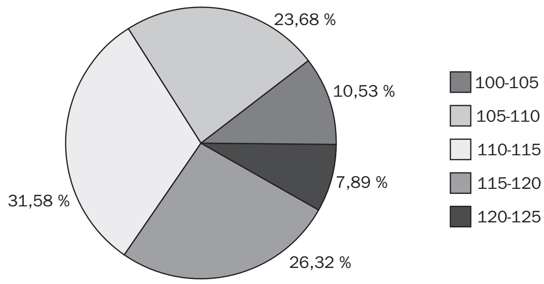
La moda estará en el intervalo [110, 115). Como primera aproximación de la moda se podría tomar la marca de la clase modal, es decir: $M_o = 112{,}5$ centímetros.

b) Como el número de datos es 38, la mitad de ellos es: $\dfrac{38}{2} = 19$. A la clase [110, 115) le corresponde la primera frecuencia absoluta mayor que 19; por tanto, la clase mediana es [110, 115). El valor aproximado de la clase mediana será la marca de la clase del intervalo [110, 115), es decir, $M = 112{,}5$ centímetros.

c) Histograma.



d) Diagrama de sectores.



11.8*. Con la variable edad, en años, de una muestra de 100 personas se forma la siguiente tabla de frecuencias absolutas.

Edad	[10, 30)	[30, 50)	[50, 70)	[70, 90)	[90, 110)
F_i	10	30	60	84	

Completa la tabla y calcula la desviación típica.

Para calcular las frecuencias absolutas, restamos las frecuencias absolutas acumuladas.

$[L_i, L_s)$	x_i	f_i	F_i	$x_i f_i$	$x_i^2 f_i$
[10, 30)	20	10	10	200	4000
[30, 50)	40	20	30	800	32 000
[50, 70)	60	30	60	1800	108 000
[70, 90)	80	24	84	1920	153 600
[90, 110)	100	16	100	1600	160 000
		100		6320	457 600

$\bar{x} = \dfrac{6320}{100} = 63{,}20$ años

$s^2 = \dfrac{457\,600}{100} - 63{,}20^2 = 581{,}76$

$s = \sqrt{581{,}76} = 24{,}11$ años

11.9. **Las calificaciones de Ana y Juan son las siguientes:**
Calificaciones de Ana: 4, 5, 6, 6, 7, 8
Calificaciones de Juan: 2, 3, 4, 4, 5, 6
¿Cuál de los dos alumnos tiene sus calificaciones más concentradas?

Hallamos las medias y las desviaciones típicas de ambas distribuciones:

$\bar{x}_A = \dfrac{4+5+6+6+7+8}{6} = 6 \quad s_A^2 = \dfrac{16+25+36+36+49+64}{6} - 36 = 1{,}67 \quad s_A = \sqrt{1{,}67} = 1{,}29$

$\bar{x}_J = \dfrac{2+3+4+4+5+6}{6} = 4 \quad s_J^2 = \dfrac{4+9+16+25+36}{6} - 16 = 1{,}67 \quad s_J = \sqrt{1{,}67} = 1{,}29$

Los coeficientes de variación son:

$$CV_A = \frac{1{,}29}{6} = 0{,}215 \qquad CV_J = \frac{1{,}29}{4} = 0{,}323$$

Como el coeficiente de variación de Ana es menor que el de Juan, se concluye que Ana tiene las notas más concentradas.

11.10. **Dada la siguiente distribución, calcula:**

Clase	[10, 15)	[15, 20)	[20, 25)	[25, 30)	[30, 35)
f_i	3	5	7	4	2

a) Los cuartiles Q_1 y Q_3.
b) El decil D_8.
c) El percentil P_{80}.

La tabla de frecuencias es:

$[L_i, L_s)$	x_i	f_i	F_i
[10, 15)	12,6	3	3
[15, 20)	17,5	5	8
[20, 25)	22,5	7	15
[25, 30)	27,5	4	19
[30, 35)	32,5	2	21

a) Como $21 \cdot \dfrac{25}{100} = 5{,}25$, Q_1 pertenece a la clase [15, 20).
Tomamos como valor aproximado de Q_1 la marca de clase: $Q_1 = 17{,}5$.
Como $21 \cdot \dfrac{75}{100} = 15{,}75$, Q_3 pertenece a la clase [25, 30).
Tomamos como valor aproximado de Q_3 la marca de clase: $Q_3 = 27{,}5$.

b) Como $21 \cdot \dfrac{8}{10} = 16{,}8$, D_8 pertenece a la clase [25, 30). Tomamos como valor aproximado de D_8 la marca de clase: $D_8 = 27{,}5$.

c) El percentil P_{80} coincide con el decil D_8; por tanto, $P_{80} = 27{,}5$.

EJERCICIOS

Variables cualitativas y cuantitativas

11.11. **Dadas las siguientes variables estadísticas:**
a) Número de hijos de los funcionarios de un ministerio.
b) Número de accidentes ferroviarios producidos cada mes durante un quinquenio.
c) Actividad de ocio preferida por un grupo de alumnos.
d) País de procedencia de un conjunto de inmigrantes.
e) Número de multas de tráfico que impone un policía al mes durante un año.
f) Grupo de rock preferido por un conjunto de alumnos.
g) Distancia recorrida por un autobús urbano durante un año.
Indica cuáles son variables cualitativas, cuantitativas discretas o cuantitativas continuas.

Variables cualitativas: c, d, f
Variables cuantitativas discretas: a, b, e
Variables cuantitativas continuas: g

11.12. Se preguntó a 40 personas por el medio de transporte utilizado para acudir al trabajo. Las posibles respuestas eran: a pie (*P*), en automóvil particular (*A*), en autobús (*B*), en metro (*M*) y otros (*O*).

Las respuestas obtenidas fueron:

A A A P A B M O M B A P A P O B M O M A

B B M A P A B M M P P A A B B M O M B O

a) Realiza la tabla de frecuencias absolutas y relativas de esta variable.

b) Representa el diagrama de barras y el de sectores.

c) Explica razonadamente qué conclusiones pueden extraerse del estudio.

a)

Transporte	f_i	h_i
Pie (*P*)	6	0,15
Automóvil (*A*)	11	0,275
Autobús (*B*)	9	0,225
Metro (*M*)	9	0,225
Otros (*O*)	5	0,125
	40	1

b) Diagrama de barras

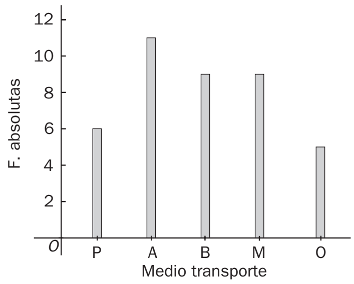
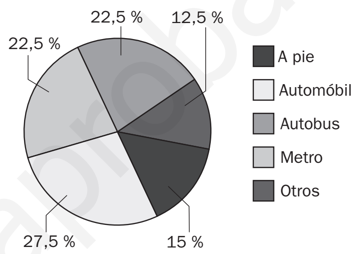


Diagrama de sectores



c) Aunque mayoritariamente se utilizan los transportes públicos, ninguno de ellos es mayoritario, ya que todos tienen una proporción muy similar.

Medidas de centralización

11.13. (PAU) a) Completa los datos que faltan en la siguiente tabla estadística, donde f_i, F_i y h_i representan, respectivamente, las frecuencias absoluta, absoluta acumulada y relativa:

x_i	f_i	F_i	h_i	$x_i f_i$
1	**4**	4	**0,08**	4
2	**4**	8	0,8	8
3	8	**16**	**0,16**	24
4	**7**	23	**0,14**	28
5	**5**	**28**	0,1	25
6	10	**38**	0,2	60
7	**7**	**45**	0,14	49
8	5	50	0,1	40
	50		1	238

b) Halla la media aritmética y la moda de esta distribución.

c) Calcula la mediana.

a) Calculamos primero el número de observaciones realizadas. Como $h_1 = 0{,}08$; $N = F_8 = \dfrac{4}{0{,}08} = 50$.

$F_1 = 4$ $\quad$ $F_2 = 4 + 4 = 8$ $\quad$ $f_3 = F_3 - F_2 = 16 - 8 = 8$

$F_4 = F_3 + f_4 = 23$ $\quad$ $f_6 = F_6 - F_5 = 38 - 28 = 10$ $\quad$ $f_8 = F_8 - F_7 = 50 - 45 = 5$

b) En la tabla ya hemos añadido la columna $x_i f_i$; por tanto, $\bar{x} = \dfrac{238}{50} = 4{,}76$.

El dato que más se repite es 6, $M_o = 6$.

c) La mitad de los datos es $\dfrac{50}{2} = 25$. El primer dato cuya frecuencia absoluta acumulada supera 25 es 5; por tanto, $M = 5$.

11.14. **(PAU) A un conjunto de datos de cinco números cuya media es 7,31 se le añaden los números 4,47 y 10,15. ¿Cuál es la media del nuevo conjunto de datos?**

Si la media de cinco datos es 7,31, la suma de los cinco datos es $5 \cdot 7{,}31 = 36{,}55$.

La media, por tanto, de los siete datos será: $\bar{x} = \dfrac{36{,}55 + 4{,}47 + 10{,}15}{7} = 7{,}31$.

11.15. **Un conjunto de cinco números naturales distintos tiene mediana 20 y media 17. ¿Cuál será el mayor número de esta serie?**

Sean los números ordenados de forma creciente: a, b, 20, c, d.

$$\frac{a + b + 20 + c + d}{5} = 17 \Rightarrow a + b + c + d = 65$$

Para que d tome el máximo valor posible, teniendo en cuenta que los números son naturales y distintos, se tiene que $a = 1$ y $b = 2$. Por tanto, $c + e = 62$.

Como el valor central de los datos es 20, el mínimo valor que puede tomar c es $c = 21$, que implica que el máximo valor que puede tomar d sea $d = 41$.

11.16. **La media de x, 3, $4x - 3$, $x + 4$, -16, 9 y $x - 4$ es 4. ¿Cuál es la mediana de estos 7 números?**

$$\bar{x} = \frac{x + 3 + 4x - 3 + x + 4 - 16 + 9 + x - 4}{7} = 4 \Rightarrow x = 5$$

Los números, ordenados de forma creciente, son: -16, 1, 3, 5, 9, 9, 17. La mediana es $M = 5$.

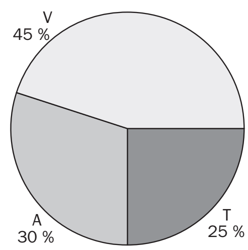
11.17. **(PAU) Los gastos anuales de cierta familia en los apartados de vivienda, alimentación y transportes están en las proporciones 9 – 6 – 5.**

a) Dibuja el gráfico de sectores que refleje la importancia de cada apartado en el total de gastos de los tres conceptos anteriores.

b) En el último año, los precios de los apartados anteriores subieron un 20%, un 5% y un 6%, respectivamente. ¿Cuál ha sido para esta familia el porcentaje de aumento anual del total de gastos en los tres apartados anteriores? (Utiliza la media ponderada.)

a) Si representamos por V, A y T los gastos en vivienda, alimentación y transportes, respectivamente, tenemos:

$$\frac{V}{9} = \frac{A}{6} = \frac{T}{5} = \frac{V + A + T}{9 + 6 + 5} = \frac{360}{20} = 18 \Rightarrow \begin{cases} V = 9 \cdot 18 = 162 \\ A = 6 \cdot 18 = 108 \\ T = 5 \cdot 18 = 90 \end{cases} \Rightarrow \begin{cases} V = 45\% \\ A = 30\% \\ T = 25\% \end{cases}$$



Siendo 162, 108 y 90 la amplitud en grados del sector circular que representa cada gasto.

b) Hay que hallar la media aritmética ponderada, en la que los valores 45, 30 y 25 están afectados de los pesos 0,2; 0,05 y 0,06:

$$\bar{x} = \frac{45 \cdot 0{,}2 + 30 \cdot 0{,}5 + 25 \cdot 0{,}06}{45 + 30 + 25} = \frac{9 + 1{,}5 + 1{,}5}{100} = \frac{12}{100} = 0{,}12$$. El aumento ha sido de un 12%.

Medidas de dispersión

11.18. **Se dan dos conjuntos de datos:**

***A*: 1, 3, 5, 7, 9** ***B*: 1, 5, 10, 15, 30**

Sin necesidad de hacer ningún cálculo, ¿cuál de los dos conjuntos tiene mayor dispersión?

Evidentemente, tiene mayor dispersión el conjunto de datos B, pues tratándose de igual número de datos, estos se encuentran más alejados entre ellos que el conjunto de datos A.

11.19. PAU) Dada la distribución estadística definida por la siguiente tabla:

x_i	[0, 5)	[5, 10)	[10, 15)	[15, 20)	[20, 25)	[25, 30)
F_i	3	5	7	8	2	5

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

b) Halla la varianza y la desviación típica

Formamos la siguiente tabla:

$[L_i, L_s)$	x_i	f_i	F_i	$x_i f_i$	$x_i^2 f_i$
[0, 5)	2,5	3	3	7,5	18,75
[5, 10)	7,5	5	8	37,5	281,25
[10, 15)	12,5	7	15	87,5	1093,75
[15, 20)	17,5	8	23	140	2450
[20, 25)	22,5	2	25	45	1012,5
[25, 30)	27,5	5	30	137,5	3781,25
		30		455	8637,5

a) $\bar{x} = \dfrac{455}{30} = 15{,}17$

La mitad de los datos es $\dfrac{30}{2} = 15$; por tanto, la clase mediana es [15, 20) y la mediana es $M = 17{,}5$.

La clase modal es [15, 20); por tanto, la moda será $M_o = 17{,}5$.

b) $s^2 = \dfrac{8637{,}5}{30} - 15{,}17^2 = 57{,}79$ $\qquad s = \sqrt{57{,}79} = 7{,}60$

11.20. Se tienen dos distribuciones cuyos datos son los siguientes:

Distribución *A*: 9, 5, 3, 2, 1, 2, 6, 4, 9, 8, 1, 3, 5, 4, 2, 6, 3, 2, 5, 6, 7

Distribución *B*: 1, 1, 3, 2, 5, 6, 7, 2, 5, 4, 3, 1, 2, 1, 5, 7, 8, 9, 9, 2, 1

a) Halla el rango de ambas distribuciones.

b) Halla la media aritmética y la desviación típica de ambas distribuciones.

c) Calcula el coeficiente de variación para discernir cuál de las dos distribuciones tiene los datos más concentrados.

a) Rango de *A*: $9 - 1 = 8$

Rango de *B*: $9 - 1 = 8$

b) DISTRIBUCIÓN *A*

x_i	f_i	$x_i f_i$	$x_i^2 f_i$
1	2	2	2
2	4	8	16
3	3	9	27
4	2	8	32
5	3	15	75
6	3	18	108
7	1	7	49
8	1	8	64
9	2	18	182
	21	93	535

$\bar{x}_A = \dfrac{93}{21} = 4{,}43$

$s_A^2 = \dfrac{535}{21} - 4{,}43^2 = 5{,}85$

$S_A = 2{,}42$

DISTRIBUCIÓN *B*

x_i	f_i	$x_i f_i$	$x_i^2 f_i$
1	5	5	5
2	4	8	16
3	2	6	18
4	1	4	16
5	3	15	75
6	1	6	36
7	2	14	98
8	1	8	64
9	2	18	182
	21	84	490

$\bar{x}_B = \dfrac{84}{4} = 4$

$s_B^2 = \dfrac{490}{21} - 42 = 7{,}33$

$s_B = 2{,}71$

c) $CV_A = \dfrac{2{,}42}{4{,}43} = 0{,}547$

$CV_B = \dfrac{2{,}71}{4} = 0{,}677$

La distribución *A* tiene los datos más concentrados.

11.21. (PAU) Considera los siguientes datos: 3, 8, 4, 10, 6, 2.

a) Halla la media y la varianza.

b) Si cada número se multiplica por 3, obtén su media y varianza a partir de los resultados anteriores.

a) $\bar{x} = \frac{3 + 8 + 4 + 10 + 6 + 2}{6} = \frac{33}{6} = 5{,}5$

$s^2 = \frac{3^2 + 8^2 + 4^2 + 10^2 + 6^2 + 2^2}{6} - (5{,}5)^2 = \frac{229}{6} - 30{,}25 = 7{,}92$

b) $\bar{y} = 3 \cdot 5{,}5 = 16{,}5$

$s_Y^2 = 9 \cdot 7{,}92 = 71{,}28$

Medidas de posición

11.22. (PAU) Se tiene el siguiente conjunto de datos:

10, 13, 4, 7, 8, 11, 10, 16, 18, 12, 3, 6, 9, 9, 4, 13, 20, 7, 5, 10, 17, 10, 16, 14, 8, 18

a) Obtén la mediana y los cuartiles.

b) Interpreta los resultados obtenidos.

a) Ordenamos los datos de menor a mayor:

3, 4, 4, 5, 6, 7, 7, 8, 8, 9, 9, 10, 10, 10, 10, 11, 12, 13, 13, 14, 16, 16, 17, 18, 18, 20

Como hay 26 valores, la mediana es la media de los dos valores centrales: $M = \frac{10 + 10}{2} = 10$.

La cuarta parte de los datos es $\frac{26}{4} = 6{,}5$. El primer cuartil es el valor que ocupa el séptimo lugar: $Q_1 = 7$.

El segundo cuartil coincide con la mediana; por tanto, $Q_2 = 10$.

Las tres cuartas partes de los datos son $\frac{26 \cdot 3}{4} = 19{,}5$. El tercer cuartil es el valor que ocupa el vigésimo lugar: $Q_3 = 14$.

b) El 25% de los datos son inferiores a 7.

El 50% de los datos son inferiores a 10.

El 75% de los datos son inferiores a 14.

11.23. a) Obtén la media y el segundo cuartil de la variable X cuya distribución de frecuencias es:

x_i	2	4	6	8	10
f_i	2	3	4	1	5

¿Qué otro nombre recibe el segundo cuartil?

b) Halla los cuartiles primero y tercero.

c) Halla los deciles 3 y 7.

d) Halla los percentiles P_{10} y P_{90}.

Formamos la tabla de frecuencias.

x_i	f_i	F_i	$x_i f_i$
2	2	2	4
4	3	5	12
6	4	9	24
8	1	10	8
10	5	15	50
	15		98

a) $\bar{x} = \frac{98}{15} = 6{,}53$

$\frac{N}{2} = \frac{15}{2} = 7{,}5 \Rightarrow Q_2 = M = 6$

Mediana,

b) $\frac{N}{4} = \frac{15}{4} = 3{,}75 \Rightarrow Q_1 = 4$

$\frac{3N}{4} = \frac{45}{4} = 11{,}25 \Rightarrow Q_3 = 10$

c) $\frac{3N}{10} = \frac{45}{10} = 4{,}5 \Rightarrow D_3 = 4$

$\frac{7N}{10} = \frac{105}{10} = 10{,}5 \Rightarrow D_7 = 10$

d) $\frac{10N}{100} = \frac{150}{10} = 1{,}5 \Rightarrow P_{10} = 2$

$\frac{90N}{100} = \frac{1350}{100} = 13{,}5 \Rightarrow P_{90} = 10$

PROBLEMAS

11.24. (PAU) Dada la siguiente tabla de ingresos:

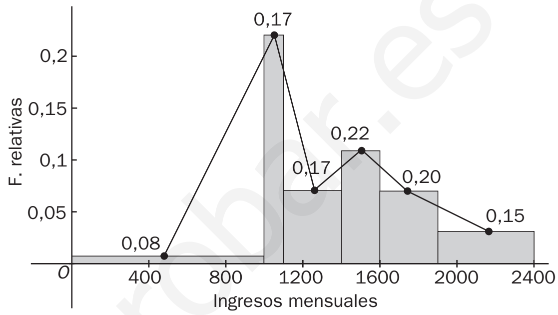
Ingresos mensuales	Frecuencia
Menos de 1000	35
[1000, 1100)	70
[1100, 1400)	70
[1400, 1600)	90
[1600, 1900)	85
Más de 1900	64

Construye el histograma de frecuencias relativas y el polígono de frecuencias relativas.

Calculamos las frecuencias relativas dividiendo las frecuencias entre el número total de casos, que es 414.

Como no todos los intervalos poseen la misma amplitud, debemos hallar la altura de cada rectángulo, teniendo en cuenta que el área del rectángulo sea proporcional a cada una de las frecuencias absolutas. Para ello, se divide cada frecuencia relativa entre el número de veces que es mayor el intervalo correspondiente al menor de todos.

Para la última clase supondremos que su límite superior es 2400.



11.25. (PAU) Las edades de un grupo de 19 personas aparecen en la siguiente tabla:

Edad	14	15	17	18	19	20	21
N.º de personas	3	1	2	3	5	3	2

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

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

c) ¿Cuántos años tiene la persona de mayor edad, de entre las que se encuentran en el 40% de las personas con menor edad?

Formamos la siguiente tabla:

x_i	f_i	f_i	$x_i f_i$	$x_i^2 f_i$
14	3	3	42	588
15	1	4	15	225
17	2	6	34	578
18	3	9	54	972
19	5	14	95	1805
20	3	17	60	1200
21	2	19	42	882
	19		342	6250

a) Media: $\bar{x} = \dfrac{342}{19} = 18$ años

Moda: $M_o = 19$ años

Mediana: La mitad de los datos es: $\dfrac{19}{2} = 9{,}5$;

por tanto, $M = 19$ años.

b) Rango: $21 - 14 = 7$ años

Varianza: $s^2 = \dfrac{6250}{19} - 18^2 = 4{,}95$ años^2

Desviación típica: $s = 2{,}22$ años

c) Hay que calcular el percentil 40. Por debajo de esos años estará el 40% de las personas más jóvenes.

$$\frac{40N}{100} = \frac{760}{100} = 7{,}6 \qquad\qquad P_{40} = 18$$

La persona de más edad entre las que se encuentran en el 40% de las más jóvenes tiene 18 años.

11.26. (PAU) Los pesos, en kg, de 20 estudiantes son:

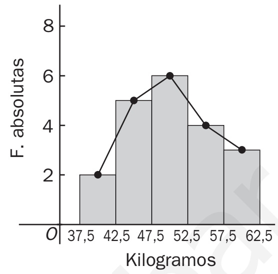
51, 47, 55, 53, 49, 47, 48, 50, 43, 60,
45, 54, 62, 57, 46, 49, 52, 42, 38, 61.

a) Agrupa los datos en cinco clases de igual amplitud.
b) Dibuja el histograma y el polígono de frecuencias absolutas correspondientes.
c) Halla la media de los datos.
d) Calcula los cuartiles primero y tercero.

a) Agrupamos los datos y formamos la tabla.

$[L_i, L_s)$	x_i	f_i	F_i	$x_i f_i$
[37,5 - 42,5)	40	2	2	80
[42,5 - 47,5)	45	5	7	225
[47,5 - 52,5)	50	6	13	300
[52,5 - 57,5)	55	4	17	220
[57,5 - 62,5)	60	3	20	180
		20		1005

b) Histograma y polígono

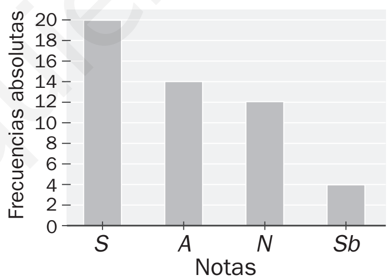


c) $\bar{x} = \dfrac{1005}{20} = 50,25$

d) $\dfrac{N}{4} = \dfrac{20}{4} = 5$. Por tanto, el cuartil primero está en la segunda clase. Tomaremos como valor aproximado su marca de clase $Q_1 = 45$.

$\dfrac{3N}{4} = \dfrac{60}{4} = 15$. Por tanto, el cuartil tercero está en la cuarta clase. Tomaremos como valor aproximado su marca de clase $Q_3 = 55$.

11.27. (PAU) El siguiente diagrama de barras muestra las calificaciones obtenidas por un grupo de 50 alumnos.

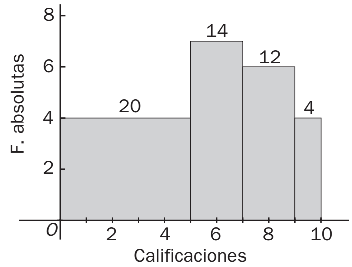


Construye el histograma correspondiente a las calificaciones numéricas y calcula la calificación media, teniendo en cuenta el siguiente cuadro de equivalencias:

Suspenso	Aprobado	Notable	Sobresaliente
[0, 5)	[5, 7)	[7, 9)	[9, 10)

Como no todos los intervalos poseen la misma amplitud, debemos hallar la altura de cada rectángulo, teniendo en cuenta que el área del rectángulo sea proporcional a cada una de las frecuencias absolutas.

$[L_i, L_s)$	x_i	f_i	$x_i f_i$	Base del rectángulo	Altura del rectángulo (f_i/base)
[0, 5)	2,5	20	50	5	4
[5, 7)	6	14	84	2	7
[7, 9)	8	12	96	2	6
[9, 10)	9,5	4	38	1	4
		50	268		



$\bar{x} = \dfrac{268}{50} = 5,36$

11.28. Una oficina bancaria ha tabulado las cantidades de dinero que retiran de sus cuentas 100 clientes en un determinado día.

Euros	[0, 120)	[120, 240)	[240, 360)	[360, 480)	[480, 600)
Cliente	33	27	19	14	7

Halla:

a) La cantidad media de dinero retirado por cliente.

b) ¿Qué porcentaje de clientes retiraron fondos por encima de la mediana?

c) Halla los cuartiles Q_1, Q_2 y Q_3.

Para los cálculos formamos la siguiente tabla:

Euros	x_i	f_i	F_i	$x_i f_i$
[0 - 120)	60	23	33	1980
[120 - 240)	180	27	60	4860
[240 - 360)	300	19	79	5700
[360 - 480)	420	17	93	5880
[480 - 600)	540	7	100	3780
		100		22 200

a) $\bar{x} = \dfrac{22\,200}{100} = 222$ €

b) Por definición de mediana, el porcentaje pedido es del 50%.

c) $\dfrac{N}{4} = \dfrac{100}{4} = 25 \Rightarrow Q_1 = 60$

$\dfrac{N}{2} = \dfrac{100}{2} = 50 \Rightarrow Q_2 = 180$

$\dfrac{3N}{4} = \dfrac{300}{4} = 75 \Rightarrow Q_3 = 300$

11.29. (PAU) Un especialista en pediatría obtuvo la siguiente tabla sobre los meses de edad de 50 niños de su consulta en el momento de andar por primera vez:

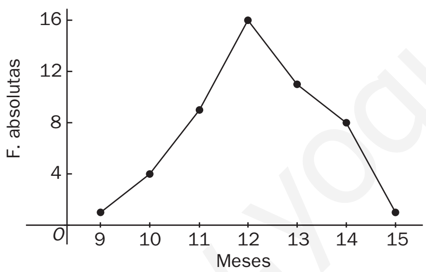
Meses	9	10	11	12	13	14	15
Niños	1	4	9	16	11	8	1

a) Dibuja el polígono de frecuencias.

b) Calcula la mediana, la moda y la varianza.

c) Halla el rango y el rango intercuartílico.

a)



Para los cálculos formamos la siguiente tabla:

x_i	f_i	F_i	x_i^2	$f_i x_i$	$f_i x_i^2$
9	1	1	81	9	81
10	4	5	100	40	400
11	9	14	121	99	1089
12	16	30	144	192	2304
13	11	41	169	143	1859
14	8	49	196	112	1568
15	1	50	225	15	225
				610	7526

b) $\dfrac{N}{2} = \dfrac{50}{2} = 25$; $M = 12$ meses

$M_o = 12$ meses

$\bar{x} = \dfrac{610}{50} = 12{,}2$;

$s^2 = \dfrac{7526}{50} - (12{,}2)^2 = 1{,}68$ meses^2

c) Rango: $15 - 9 = 6$ meses

$\dfrac{N}{4} = \dfrac{50}{4} = 12{,}5$; $Q_1 = 11$ meses

$\dfrac{3N}{4} = \dfrac{150}{4} = 37{,}5$; $Q_3 = 13$ meses

Rango intercuartílico: $13 - 11 = 2$ meses

11.30. (PAU) Se han lanzado dos dados 120 veces y cada vez se ha anotado su suma. Los resultados vienen reflejados en la siguiente tabla:

Sumas	2	3	4	5	6	7	8	9	10	11	12
N.º de veces	3	8	9	11	20	19	16	13	11	6	4

a) Calcula la media $\bar{x}$ y la desviación típica s.

b) Halla el porcentaje de valores comprendidos en los siguientes intervalos: $(\bar{x} - s, \bar{x} + s)$, $(\bar{x} - 2s, \bar{x} + 2s)$ y $(\bar{x} - 3s, \bar{x} + 3s)$.

Formamos la tabla:

x_i	f_i	$f_i x_i$	$f_i x_i^2$
2	3	6	12
3	8	24	72
4	9	36	144
5	11	55	275
6	20	120	720
7	19	133	931
8	16	128	1024
9	13	117	1053
10	11	110	1100
11	6	66	726
12	4	48	576
	120	843	6633

a) $\bar{x} = \dfrac{843}{120} = 7{,}025$

$s^2 = \dfrac{6633}{120} - 7{,}025^2 = 5{,}924$

$s = \sqrt{5{,}924} = 2{,}43$

b) $\bar{x} - s = 4{,}595$

$\bar{x} + s = 9{,}455$

Los valores comprendidos en (4,595; 9,455) son los correspondientes a las sumas 5, 6, 7, 8 y 9, que son un total de 79 valores.

$\dfrac{79}{120} = 0{,}6583 \Rightarrow$ El 65,83% de los datos está en el intervalo $(\bar{x} - s, \bar{x} + s)$

$\bar{x} - 2s = 2{,}165$

$\bar{x} + 2s = 11{,}885$

Los valores comprendidos en (2,165; 11,885) son correspondientes a las sumas 3, 4, 5, 6, 7, 8, 9, 10 y 11, que son un total de 113 valores.

$\dfrac{113}{120} = 0{,}9417 \Rightarrow$ El 94,17% de los datos está en el intervalo $(\bar{x} - 2s, \bar{x} + 2s)$

$\bar{x} - 3s = -0{,}265$

$\bar{x} + 3s = 14{,}315$

Todos los valores están comprendidos en (−0,265; 14,315).

El 100% de los datos está en el intervalo $(\bar{x} - 3s, \bar{x} + 3s)$.

11.31. (PAU) Se ha pasado un test de 79 preguntas a 600 personas. El número de respuestas correctas se refleja en la siguiente tabla:

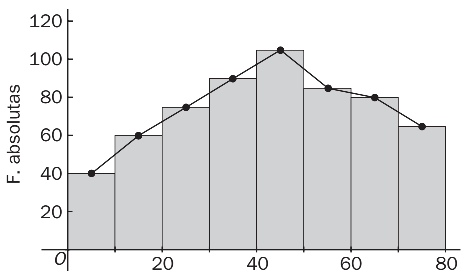
Respuestas	N.º de personas
[0, 10)	40
[10, 20)	60
[20, 30)	75
[30, 40)	90
[40, 50)	105
[50, 60)	85
[60, 70)	80
[70, 80)	65

a) Dibuja el histograma y el polígono de frecuencias de las frecuencias absolutas.

b) Halla la media y la desviación típica de respuestas correctas.

c) Calcula la mediana y el primer cuartil. ¿Qué miden estos parámetros?

a) Histograma + polígono



Formamos la tabla:

$[L_i, L_s)$	x_i	f_i	F_i	$f_i x_i$	$f_i x_i^2$
[0, 10)	5	40	40	200	1000
[10, 20)	15	60	100	900	13 500
[20, 30)	25	75	175	1875	146 875
[30, 40)	35	90	265	3150	110 250
[40, 50)	45	105	370	4725	212 625
[50, 60)	55	85	455	4675	257 125
[60, 70)	65	80	535	5200	338 000
[70, 80)	75	65	600	4875	365 625
				25 600	1 345 000

b) $\bar{x} = \dfrac{25\,600}{600} = 42{,}67$

$s^2 = \dfrac{1\,345\,000}{600} - 42{,}67^2 = 420{,}94$

$s = 420{,}94 = 20{,}52$

c) $\dfrac{600}{2} = 300$. El dato de orden 300 está en el quinto intervalo. Por tanto, tomaremos como valor aproximado de la mediana la marca de esta clase, es decir, $M = 45$.

$\dfrac{600}{4} = 150$. La clase que contiene al primer cuartil será la tercera y, en consecuencia, tomaremos como valor valor aproximado del primer cuartil su marca de clase, es decir, $Q_1 = 25$.

11.32. (PAU) En una encuesta sobre tráfico se ha preguntado a 1000 conductores sobre el número de multas recibidas que ha sido mayor o igual a cero y menor o igual a 5. Al efectuar la tabla correspondiente, algún número ha desaparecido, por lo que disponemos de la siguiente información:

N.º de conductores	–	260	150	190	100	90
Número de multas	0	1	2	3	4	5

Halla

a) La media.
b) La mediana.
c) La moda.
d) La desviación típica.
e) Los cuartiles primero y tercero.
f) El rango intercuartílico.

El dato desconocido es: $1000 - (260 + 150 + 190 + 100 + 90) = 210$

Formamos la tabla:

x_i	f_i	F_i	$x_i f_i$	$x_i^2 f_i$
0	210	210	0	0
1	260	470	260	260
2	150	620	300	600
3	190	810	570	1710
4	100	910	400	1600
5	90	1000	450	2250
	1000		1980	6420

a) $\bar{x} = \dfrac{1980}{1000} = 1{,}98$

b) $\dfrac{1000}{2} = 500$. $M = 2$

c) $M_o = 1$

d) $s^2 = \dfrac{6420}{1000} - 1{,}98^2 = 2{,}4996$

$s = 1{,}58$

e) $\dfrac{N}{4} = \dfrac{1000}{4} = 250$. $Q_1 = 1$

$\dfrac{3N}{4} = \dfrac{3000}{4} = 750$. $Q_3 = 3$

f) $RI = Q_3 - Q_1 = 3 - 1 = 2$

11.33. (PAU) Un dentista observa el número de caries de cada uno de los 100 niños de un colegio. La información resumida aparece en la siguiente tabla:

Número de caries	Frecuencia absoluta	Frecuencia relativa
0	25	0,25
1	20	0,20
2	x	z
3	15	0,15
4	y	0,05

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

b) Dibuja un diagrama de sectores.

c) Realiza un diagrama de barras.

d) Calcula el número medio de caries.

e) Calcula los cuartiles.

a) La suma de las frecuencias relativas ha de ser igual a 1:

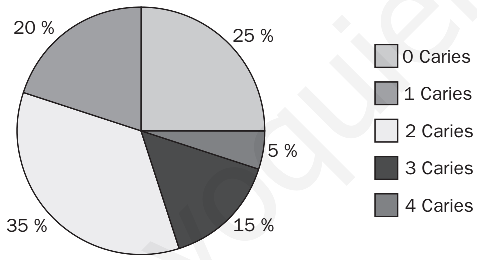
$$0{,}25 + 0{,}2 + z + 0{,}15 + 0{,}05 = 1 \Rightarrow 0{,}65 + z = 1 \Rightarrow z = 0{,}35$$

La frecuencia relativa de un dato es igual a su frecuencia absoluta dividida por la suma de las frecuencias absolutas, que el enunciado nos dice que es 100.

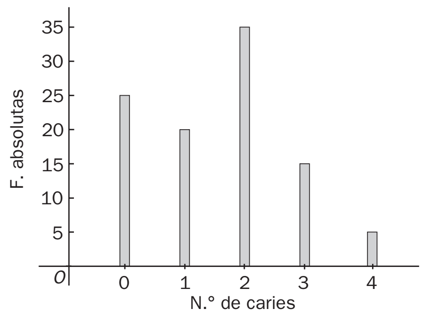
$$\frac{x}{100} = 0{,}35 \Rightarrow x = 35$$

$$\frac{y}{100} = 0{,}05 \Rightarrow y = 5$$

b) Diagrama de sectores



c) Diagrama de barras



Formamos la tabla:

x_i	f_i	F_i	x_if_i
0	25	25	0
1	20	45	20
2	35	80	70
3	15	95	45
4	5	100	20
		100	155

d) $\bar{x} = \dfrac{155}{100} = 1{,}55$ caries

e) $\dfrac{N}{4} = 25$. $Q_1 = 0$ caries

$\dfrac{N}{2} = 50$. $Q_2 = 2$ caries

$\dfrac{3N}{4} = 75$. $Q_3 = 2$ caries

11.34. El número de horas que 20 trabajadores perdieron por bajas médicas el año pasado es el siguiente:

0 3 4 8 10 12 12 15 15 17 19 21 21 23 25 26 32 33 40 60

a) Construye la tabla de frecuencias agrupando los datos en intervalos de amplitud 10, indicando también las frecuencias absolutas y relativas acumuladas.

b) Dibuja el histograma y el polígono de frecuencias de las frecuencias relativas.

c) Halla la media de días no trabajados por trabajador.

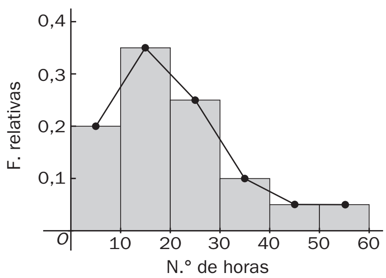
d) Halla el rango intercuartílico.

e) Calcula el coeficiente de variación.

a) En la tabla vamos a añadir las columnas correspondientes para el cálculo de la media y la desviación típica.

$[L_i, L_s)$	x_i	f_i	F_i	h_i	H_i	$x_i f_i$	$x_i^2 f_i$
[0, 10)	5	4	4	0,2	0,2	20	100
[10, 20)	15	7	11	0,35	0,55	105	1575
[20, 30)	25	5	16	0,25	0,8	125	3125
[30, 40)	35	2	18	0,1	0,9	70	2450
[40, 50)	45	1	19	0,05	0,95	45	2025
[70, 80)	55	1	20	0,05	1	55	3025
		20		1		420	12 300

b)



c) $\bar{x} = \frac{420}{20} = 21$

d) $\frac{N}{4} = 5$. El primer cuartil está en el intervalo [10, 20); por tanto, $Q_1 = 15$

$\frac{3N}{4} = 15$. El tercer cuartil está en el intervalo [20, 30); por tanto, $Q_3 = 25$

$RI = Q_3 - Q_1 = 25 - 15 = 10$

e) $s^2 = \frac{12\,300}{20} - 21^2 = 174$

$s = \sqrt{174} = 13,19$

$CV = \frac{s}{\bar{x}} = \frac{13,19}{20} = 0,7$

11.35. Las sumas de los puntos obtenidos al lanzar 20 veces dos dados son:

9 3 6 4 5 8 5 6 4 11 7 8 7 8 5 7 2 9 7 10

a) Calcula las frecuencias absolutas y relativas.

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

c) Calcula la varianza y la desviación típica.

a) Formamos la tabla:

x_i	f_i	F_i	h_i	$x_i f_i$	$x_i^2 f_i$
2	1	1	0,05	2	4
3	1	2	0,05	3	9
4	2	4	0,1	8	32
5	3	7	0,15	15	75
6	2	9	0,1	12	72
7	4	13	0,2	28	196
8	3	16	0,15	24	192
9	2	18	0,1	18	162
10	1	19	0,05	10	100
11	1	20	0,05	11	121
	20		1	131	963

b) $\bar{x} = \frac{131}{20} = 6,55$ $\frac{N}{2} = 10.$ $M = 7$ $M_o = 7$

c) $s^2 = \frac{963}{20} - 6,55^2 = 5,25$ $s = \sqrt{5,25} = 2,29$

11.36. **La tabla adjunta muestra el presupuesto dedicado a acción social de varios municipios de una determinada comunidad autónoma.**

Presupuesto (en miles de euros)	Número de municipios
[0, 30)	8
[30, 60)	12
[60, 90)	19
[90, 120)	21
[120, 150)	14
[150, 180)	6

a) Representa gráficamente la distribución mediante un histograma y su polígono de frecuencias.

b) Halla la mediana e interpreta este parámetro.

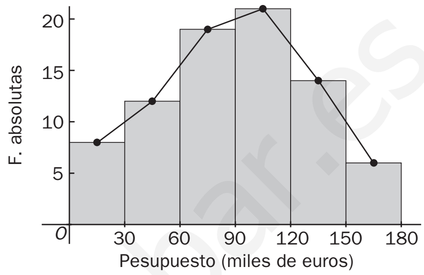
c) Calcula la media.

d) ¿Cuál es la proporción de municipios que dedican a acción social más de 30 000 euros y menos de 150 000?

Formamos la tabla:

$[L_i, L_s)$	x_i	f_i	F_i	$x_i f_i$
[0, 30)	15	8	8	120
[30, 60)	45	12	20	540
[60, 90)	75	19	39	1425
[90, 120)	105	21	50	2205
[120, 150)	135	14	64	1890
[150, 180)	165	6	70	990
		70		7170

a) Histograma + polígono



b) La mitad de los datos es 40. Como los 40 se alcanzan en la cuarta clase, tomaremos como aproximación a la mediana la marca de esta clase, es decir, $M = 105$. La mitad de los datos son inferiores a 105, y la otra mitad, superiores.

c) $\bar{x} = \dfrac{7170}{70} = 102{,}43$

d) Hemos de considerar las frecuencias correspondientes a las clases 2.ª, 3.ª, 4.ª y 5.ª: $12 + 19 + 21 + 14 = 66$.

Con una simple proporción, $x = \dfrac{66 \cdot 100}{90} = 73{,}33$. Es decir, aproximadamente el 73% de los municipios dedica a acción social más de 30 000 euros y menos de 150 000.

11.37. **Se ha realizado un test, compuesto de 10 preguntas, a 40 alumnos de un grupo, con los siguientes resultados:**

Respuestas	[0, 2)	[2, 4)	[4, 6)	[6, 8)	[8, 10)
Alumnos	4	9	15	7	5

a) Representa gráficamente la distribución.

b) Calcula el valor de la moda.

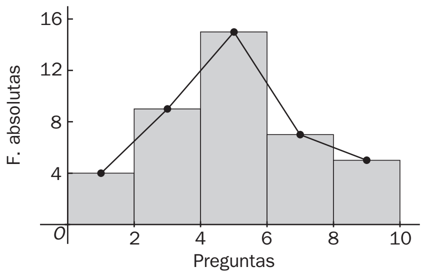
c) Halla la varianza y la desviación típica.

d) ¿A partir de qué dato se encuentra el 70% de los alumnos que han obtenido la mejor nota?

Formamos la tabla:

$[L_i, L_s)$	x_i	f_i	F_i	$x_i f_i$	$x_i^2 f_i$
[0, 2)	1	4	4	4	4
[2, 4)	3	9	13	27	181
[4, 6)	5	15	28	75	375
[6, 8)	7	7	35	49	343
[8, 10)	9	5	40	45	405
		40		200	1208

a)



b) La clase modal es [4, 6). Tomaremos como aproximación de la moda la marca de clase: $M_o = 5$.

c) $\bar{x} = \dfrac{200}{40} = 5$ $\qquad s^2 = \dfrac{1208}{40} - 5^2 = 5{,}2$ $\qquad s = \sqrt{5{,}2} = 2{,}28$

d) Hay que calcular el percentil 30.

$\dfrac{30N}{100} = \dfrac{1200}{100} = 12$. El percentil 30 se encuentra en la clase [2, 4). Tomamos como P_{30} la marca de clase; $P_{30} = 3$.

11.38 La tabla adjunta muestra las distancias, en cm, alcanzadas por un grupo de alumnos en salto de longitud.

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

b) Halla el rango y el rango intercuartílico, e interpreta este último parámetro.

Longitud del salto (cm)	Número de alumnos
[100, 110)	3
[110, 120)	7
[120, 130)	15
[130, 140)	16
[140, 150)	6
[150, 160)	2

Formamos la tabla:

$[L_i, L_s)$	x_i	f_i	F_i	$x_i f_i$
[100, 110)	105	3	3	315
[110, 120)	115	7	10	805
[120, 130)	125	15	25	1875
[130, 140)	135	16	41	2160
[140, 150)	145	6	47	870
[150, 160)	155	2	49	310
		49		6335

a) $\bar{x} = \dfrac{6335}{49} = 129{,}29$

$\dfrac{N}{2} = \dfrac{49}{2} = 24{,}5$. La mediana se encuentra en la clase [120, 130); por tanto, $M = 125$.

La clase modal es [130, 140); por tanto, $M_o = 135$.

b) $R = 160 - 100 = 60$

$\dfrac{N}{4} = \dfrac{49}{4} = 12{,}25$, el primer cuartil se alcanza en [120, 130); por tanto, $Q_1 = 125$ cm.

$\dfrac{3N}{4} = \dfrac{49 \cdot 3}{4} = 36{,}75$, el tercer cuartil se alcanza en [130, 140); por tanto, $Q_3 = 135$ cm.

$RI = Q_3 - Q_1 = 135 - 125 = 10$ cm. El 50% de los datos de la zona central están en tan solo 10 cm.

11.39. Se ha contado el número de oficinas municipales de información al consumidor abiertas al público en 25 ciudades. Estos son los datos:

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

a) Construye la tabla de frecuencias y representa el diagrama de barras.

b) Halla la media.

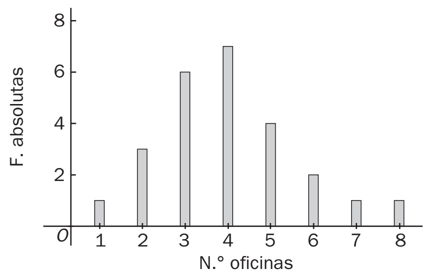
c) Halla la desviación media.

d) Calcula la varianza, la desviación típica y el coeficiente de variación.

e) Calcula la mediana y el rango intercuartílico.

a)

x_i	f_i	F_i	h_i	H_i	$x_i f_i$	$\lvert x_i - \bar{x} \rvert$	$\lvert x_i - \bar{x} \rvert f_i$	$x_i^2 f_i$
1	1	1	0,04	0,04	1	3	3	1
2	3	4	12	0,16	6	2	6	12
3	6	10	0,24	0,40	18	1	6	54
4	7	17	0,28	0,68	28	0	0	112
5	4	21	0,16	0,84	20	2	4	100
6	2	23	0,08	0,92	12	2	4	72
7	1	24	0,04	0,96	7	3	3	49
8	1	25	0,04	1	8	4	4	64
	25		1		100		30	464



b) $\bar{x} = \dfrac{100}{25} = 4$

c) $D_{\bar{x}} = \dfrac{30}{25} = 1{,}2$

d) $s^2 = \dfrac{464}{25} - 4^2 = 2{,}56$ $\qquad s = \sqrt{2{,}56} = 1{,}6$ $\qquad CV = \dfrac{s}{\bar{x}} = \dfrac{1{,}6}{4} = 0{,}4$

e) La mediana corresponde al dato número 13 de la serie ordenada, por tanto: $M = 4$.

$\dfrac{N}{4} = \dfrac{25}{4} = 6{,}25$, el primer cuartil corresponde al séptimo dato de la serie ordenada; por tanto, $Q_1 = 3$.

$\dfrac{3N}{4} = \dfrac{25 \cdot 3}{4} = 18{,}75$, el tercer cuartil corresponde al dato decimonoveno de la serie, por tanto, $Q_3 = 5$ cm.

$RI = Q_3 - Q_1 = 5 - 3 = 2$.

11.40. (PAU) Al estudiar la distribución de la edad en una población se obtuvieron los resultados siguientes:

Edad (en años)	[0, 20)	[20, 40)	[40, 60)	[60, 80)
Número de individuos	15	?	15	16

Como se ve, se ha extraviado el dato correspondiente al intervalo [20, 40).

a) ¿Cuál será el valor de este dato si la media es de 35 años?

b) ¿Cuál será el valor de este dato si la mediana es de 30 años?

c) ¿Cuál será la desviación típica si el dato es 16?

a) Formamos la tabla:

$[L_i, L_s)$	x_i	f_i	$x_i f_i$
[0, 20)	10	15	150
[20, 40)	30	f_2	$30 f_2$
[40, 60)	50	15	750
[60, 80)	70	16	1120
		$46 + f_2$	$2020 + 30 f_2$

Sea f_2 la frecuencia del intervalo [20, 40) si la edad media fuera 35 años:

$$35 = \frac{2020 + 30f_2}{46 + f_2} \Rightarrow 1610 + 35 f_2 = 2020 + 30 f_2$$

$f_2 = 82$

b) Formamos la tabla:

$[L_i, L_s)$	x_i	f_i	$x_i f_i$
[0, 20)	10	15	15
[20, 40)	30	f_2	$15 + f_2$
[40, 60)	50	15	$30 + f_2$
[60, 80)	70	16	$46 + f_2$
		$46 + f_2$	$2020 + 30 f_2$

La mediana es el dato central de la distribución. Como $M = 30$, se deduce que:

$$\frac{46 + f_2}{2} \geqslant 31 \Rightarrow 46 + f_2 \geqslant 62 \Rightarrow f_2 \geqslant 16$$

c) Formamos la tabla:

$[L_i, L_s)$	x_i	f_i	$x_i f_i$	$x_i^2 f_i$
[0, 20)	10	15	150	1500
[20, 40)	30	16	480	14 400
[40, 60)	50	15	750	37 500
[60, 80)	70	16	1120	78 400
		102	2500	131 800

$$\bar{x} = \frac{2500}{62} = 40{,}32$$

$$s^2 = \frac{131\,800}{62} - 40{,}32^2 = 500{,}1 \qquad s = 22{,}36$$

11.41. (PAU) Un centro de enseñanza tiene tres grupos de 1.º de Bachillerato. La nota media de los alumnos del grupo *A* es de 5,7 puntos, la de los del grupo *B* es de 5,6, y la de los del grupo *C* es de 5,5. En el grupo A hay 30 alumnos, y se sabe que en el grupo *C* hay 5 alumnos más que en el grupo *B*.

Si la nota media de todos los alumnos de 1.º de Bachillerato del centro es de 5,6 puntos, ¿cuántos alumnos de 1.º de Bachillerato hay en el mismo?

Sea *b* el número de alumnos del grupo *B*, $b + 5$ será el número de alumnos del grupo *C*.

$$5{,}6 = \frac{30 \cdot 5{,}7 + b \cdot 5{,}6 + (b+5) \cdot 5{,}5}{30 + b + (b+5)} \Rightarrow 5{,}6 = \frac{198{,}5 + 11{,}1 + b}{35 + 2b} \Rightarrow 196 + 11{,}2 \cdot b = 198{,}5 + 11{,}1 \cdot b \Rightarrow b = 25$$

El número de alumnos de 1.º de Bachillerato en el Instituto es: 30 + 25 + 30 = 85.

PROFUNDIZACIÓN

11.42. Eligiendo tres números al azar entre el 0 y el 9, se forma un número de tres cifras.

Si la media de las tres cifras es 5, y su moda, 7, ¿cuál es el mayor número que se puede formar de esta manera?

Sean las cifras elegidas *a*, *b* y *c*. Por ser la moda 7, la cifra 7 se repite dos o tres veces. Tres veces no se puede repetir, ya que implicaría que la media sería 7.

Si $a = b = 7 \Rightarrow \frac{14 + c}{3} = 15 \Rightarrow c = 1$. El mayor número que se puede formar con las cifras 1, 7 y 7 es 771.

11.43. Demuestra que las dos expresiones siguientes de la varianza son iguales.

$$s^2 = \frac{\Sigma f_i(x_i - \overline{x})^2}{\Sigma f_i} = \frac{\Sigma f_i x_i^2}{\Sigma f_i} - \overline{x}^2$$

$$s^2 = \frac{\sum f_i(x_i - \overline{x})^2}{\sum f_i} = \frac{\sum f_i(x_i^2 - 2\overline{x}\, x_i + \overline{x}^2)}{\sum f_i} = \frac{\sum f_i x_i^2}{\sum f_i} - 2\overline{x}\frac{\sum f_i x_i}{\sum f_i} + \overline{x}^2\frac{\sum f_i}{\sum f_i} = \frac{\sum f_i x_i^2}{\sum f_i} - 2\overline{x}^2 + \overline{x}^2 = \frac{\sum f_i x_i^2}{\sum f_i} - \overline{x}^2$$

11.44. En estadística unidimensional se dice que un valor x está alejado cuando se encuentra situado a la derecha del tercer cuartil más 1,5 veces el rango intercuartílico, o cuando está situado a la izquierda del primer cuartil menos 1,5 veces el rango intercuartílico. Es decir, x está alejado si:

$x > Q_3 + 1{,}5\ (Q_3 - Q_1)$ o bien $x < Q_1 - 1{,}5\ (Q_3 - Q_1)$

Para los siguientes conjuntos de datos, averigua si existe un valor alejado.

a) 30, 31, 31, 33, 34, 36, 37

b) 11, 78, 79, 81, 82, 83, 160

a) $Q_1 = 31$ y $Q_3 = 36$

Ningún valor está alejado, ya que para los valores extremos se tiene:

$31 - 1{,}5\ (36 - 31) = 23{,}5$ y $30 > 23{,}5$ $\quad 36 + 1{,}5\ (36 - 31) = 43{,}5$ y $36 < 43{,}5$

b) $Q_1 = 78$ y $Q_3 = 83$

$78 - 1{,}5\ (83 - 78) = 70{,}5$ y $11 < 70{,}5$ $\quad 83 + 1{,}5\ (83 - 78) = 90{,}5$ y $160 > 90{,}5$

Los valores extremos son valores alejados.

11.45. Se denomina sesgo la menor o mayor simetría de una distribución. El coeficiente de sesgo viene dado por la expresión $v = \dfrac{\overline{x} - M_o}{s}$.

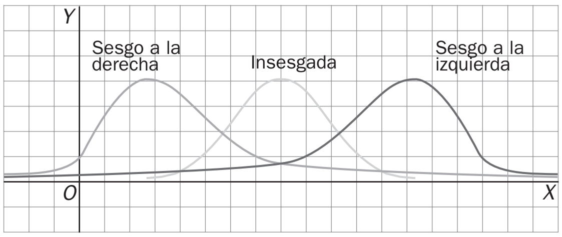
Dada la distribución definida por la siguiente tabla estadística:

x_i	6	7	8	9	10	11	12	13
f_i	1	2	3	4	7	15	13	4

a) Halla el coeficiente de sesgo.

b) Representa el polígono de frecuencias. ¿Qué observas?

c) ¿Cómo será el signo del coeficiente de sesgo de las siguientes distribuciones dadas por sus polígonos de frecuencias



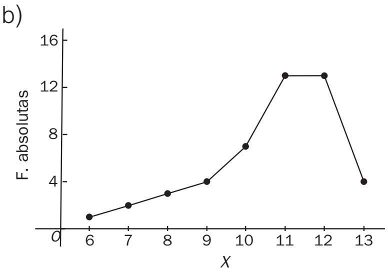
Formamos la tabla:

x_i	f_i	$x_i f_i$	$x_i^2 f_i$
6	1	6	36
7	2	14	98
8	3	24	192
9	4	36	324
10	7	70	700
11	15	165	1815
12	13	156	1872
13	4	52	676
	49	523	5713

a) $\overline{x} = \dfrac{587}{49} = 10{,}67 \quad M_o = 11$

$s^2 = \dfrac{5713}{49} - 10{,}67^2 = 2{,}74 \quad s = \sqrt{2{,}74} = 1{,}66 \quad v = \dfrac{10{,}67 - 11}{1{,}66} = -0{,}2$

b)



Se observa en el polígono de frecuencias una larga cola a la izquierda que corresponde a una distribución sesgada a la izquierda; por ello, el coeficiente de sesgo es negativo.

c) Sesgo a la derecha, $v > 0$

Insesgada; $v = 0$

Sesgo a la izquierda, $v < 0$