MATEMÁTICAS ORIENTADAS A LAS ENSEÑANZAS ACADÉMICAS
4.º ESO

somoslink

SOLUCIONES AL LIBRO DEL ALUMNO

Unidad 12. Estadística

Unidad 12. Estadística

SOLUCIONES PÁG. 267

1 Las emisiones de gases de efecto invernadero por habitante, en toneladas de CO_2, en diferentes países europeos han sido las siguientes:
22,6; 14,5; 12,8; 12,5; 11,5; 11,3; 10,7; 10,5; 9,3; 7,0; 6,6; 24,7; 24,0; 7,5; 21,3; 18,7; 19,4; 5,9; 12,6; 5,2; 13,1; 16,6; 8,8; 20,3.
Realiza una tabla estadística agrupando los datos en intervalos.

Emisiones, x_i	c_i	n_i	N_i	f_i	p_i
[5 , 9)	7	6	6	0,25	25 %
[9, 13)	11	8	14	0,33	33 %
[13, 17)	15	3	17	0,125	12,5 %
[17, 21)	19	3	20	0,125	12,5 %
[21, 25]	23	4	24	0,17	17 %
Total		**N = 25**		**1**	**100 %**

2 Elabora una tabla estadística con los siguientes datos:
82 81 83 82 81 80 83 84 82 83 81 84 82
84 81 82 83 84 82 82 81 82 83 80 82 82

Variable, x_i	n_i	N_i	f_i	p_i
80	2	2	0,08	8 %
81	5	7	0,19	19 %
82	10	17	0,38	38 %
83	5	22	0,19	19 %
84	4	26	0,15	15 %
Total	**N = 26**		**1**	**100 %**

3 Realiza las actividades de esta página de Internet:
http://www.ine.es/explica/explica_pasos_primera_encuesta.htm

Respuesta abierta.

4 En una encuesta realizada para conocer la edad media a la que se casan las parejas se han obtenido estos resultados:
30, 29, 31, 30, 29, 33, 31, 32, 30, 31, 29, 32, 30, 31, 30, 32, 29, 30, 31, 32, 30, 30, 29, 30, 31, 33, 30, 30, 29, 31, 29, 32, 30, 31, 30.
Realiza una tabla de frecuencias y calcula:

x_i	n_i	N_i	f_i	p_i	$n_i \cdot x_i$	x_i^2	$n_i \cdot x_i^2$
29	7	7	0,20	20 %	203	841	5 887
30	13	20	0,37	37 %	390	900	11 700
31	8	28	0,23	23 %	248	961	7 688
32	5	33	0,14	14 %	160	1 024	5 120
33	2	35	0,06	6 %	66	1 089	2 178
Total	**N = 35**		**1**	**100 %**	**1 067**	**4 815**	**32 573**

a. Los parámetros estadísticos de centralización.
Los parámetros de centralización son:
Mo = 30, porque es el valor de la variable que más se repite.
$\bar{x} = \frac{x_1 + \cdots + x_n}{\text{N}} = \frac{1\,067}{35}$ = 30,49
Me = 30 porque $\frac{35}{2} = 17{,}5$ y el primer valor de N_i que lo supera es 20 que se corresponde con X_i = 30

b. Los parámetros estadísticos de dispersión.
Los parámetros de dispersión son:
R = $X_{max} - X_{min}$ = 33 – 29 = 4
V (x) = $\frac{\sum_{i=1}^{n} x_i^2 \cdot n_i}{\text{N}} - \bar{x}^2 = \frac{32\,573}{35} - (30{,}49)^2$ = 930,657 – 929,64 = 1,02
S (x) = $+\sqrt{\text{V}(x)}$ = 1
CV = $\frac{\text{S}\,(x)}{\bar{x}} = \frac{1}{30{,}49}$ = 0,033 ⇒ CV = 3,3%

SOLUCIONES PÁG. 271

5 En un hospital se registra la temperatura (*X*), en grados Celsius, y el número de pulsaciones (*Y*) a un grupo de 13 pacientes. Los resultados obtenidos son los siguientes: (36 , 75) (36,5 , 90) (36 , 85) (37 , 90) (37 , 80) (36,5 , 85) (36,5 , 75) (37,5 , 90) (38 , 90) (37 , 80) (36,5 , 85) (37 , 90) (36,5 , 75)

a. Expresa los valores que toma cada una de las dos variables en una tabla simple.

x_i	y_j	n_{ij}
36	75	1
36	85	1
36,5	75	2
36,5	85	2
36,5	90	1
37	80	2
37	90	2
37,5	90	1
38	90	1

b. Indica los valores que adopta cada una de las dos variables en una tabla de doble entrada.

x_i \ y_j	75	80	85	90	Frecuencia absoluta de *X*
36	1	0	1	0	2
36,5	2	0	2	1	5
37	0	2	0	2	4
37,5	0	0	0	1	1
38	0	0	0	1	1
Frecuencia absoluta de *Y*	3	2	3	5	13

c. ¿Cuántas pulsaciones tienen los pacientes que han tenido una temperatura de 37,5º?

El único paciente que ha tenido una temperatura de 37,5 ºC tiene 90 pulsaciones.

d. Establece la distribución marginal de las variables *X* e *Y*.

Temperatura en ºC, *X*	Frecuencia
36	2
36,5	5
37	4
37,5	1
38	1

Número de pulsaciones, *Y*	Frecuencia
75	3
80	2
85	3
90	5

6 **En el marco de un estudio sobre educación se ha realizado una prueba a 24 personas para relacionar el número de faltas de ortografía cometidas en un dictado (*X*) y los años de estudio realizados (*Y*): (4 , 8) (6 , 9) (5 , 11) (5 , 10) (7 , 7) (4 , 12) (3 , 12) (5 , 9) (6 , 7) (7 , 7) (4 , 11) (3 , 12) (6 , 8) (3 , 11) (4 , 11) (5 , 9) (6 , 7) (7 , 8) (5 , 10) (3 , 12) (5 , 9) (4 , 11) (6 , 8) (5 , 9)**

a. Registra los valores que toma cada una de las dos variables en una tabla simple.

x_i	y_j	n_{ij}
3	11	1
3	12	3
4	8	1
4	11	3
4	12	1
5	9	4
5	10	2
5	11	1
6	7	2
6	8	2
6	9	1
7	7	2
7	8	1

b. Expresa los valores de cada una de las dos variables en una tabla de doble entrada.

x_i \ y_j	7	8	9	10	11	12	Frecuencia absoluta de *X*
3	0	0	0	0	1	3	4
4	0	1	0	0	3	1	5
5	0	0	4	2	1	0	7
6	2	2	1	0	0	0	5
7	2	1	0	0	0	0	3
Frecuencia absoluta de *Y*	4	4	5	2	5	4	24

c. Determina la distribución marginal de las variables *X* e *Y*.

X	Frecuencia
3	4
4	5
5	7
6	5
7	3

Y	Frecuencia
7	4
8	4
9	5
10	2
11	5
12	4

SOLUCIONES PÁG. 273

7 **En el estudio de una variable bidimensional se han recogido los siguientes datos:**
(6 , 3), (6 , 2), (5 , 1), (7 , 3), (7 , 0), (8 , 2) (9 , 4), (6 , 3), (8 , 1), (7 , 2), (10 , 4), (8 , 4) (6 , 3), (10 , 2), (6 , 3), (5 , 2), (6 , 4), (7 , 1) (5 , 3), (9 , 2), (5 , 4), (8 , 0), (6 , 1), (5 , 3)

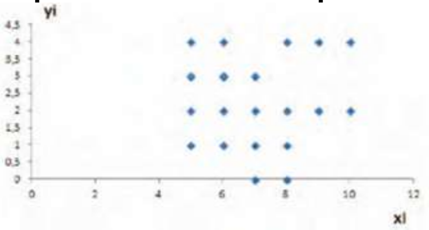
a. Construye una tabla simple con los datos.

x_i	y_j	n_{ij}
5	1	1
5	2	1
5	3	2
5	4	1
6	1	1
6	2	1
6	3	4
6	4	1
7	0	1
7	1	1
7	2	1
7	3	1
8	0	1
8	1	1
8	2	1
8	4	1
9	2	1
9	4	1
10	2	1
10	4	1

b. Elabora una tabla de doble entrada con los datos.

x_i \ y_j	0	1	2	3	4	Frecuencia absoluta de *X*
5	0	1	1	2	1	5
6	0	1	1	4	0	6
7	1	1	1	1	0	4
8	1	1	1	0	1	4
9	0	0	1	0	1	2
10	0	0	1	0	1	2
Frecuencia absoluta de *Y*	2	4	6	7	4	23

c. Representa la nube de puntos.



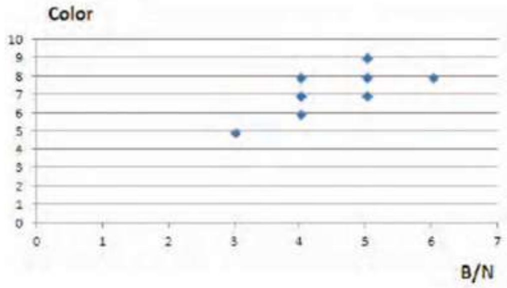
8 **Un centro escolar planea adquirir una fotocopiadora y ha comparado el precio en nueve modelos por cada mil fotocopias en blanco y negro y en color. La comparativa de las fotocopiadoras se muestra en la siguiente tabla:**

B/N	4	3	4	5	6	4	5	7	5
C	7	5	6	8	8	8	7	7	9

a. Elabora una tabla de doble entrada de los datos.

x_i \ y_j	5	6	7	8	9	Frecuencia absoluta de *X*
3	1	0	0	0	0	1
4	0	1	1	1	0	3
5	0	0	1	1	1	3
6	0	0	0	1	0	1
7	0	0	1	0	0	1
Frecuencia absoluta de *Y*	1	1	3	3	1	9

b. Representa la nube de puntos.



c. Halla la distribución marginal de las variables *X* e *Y*.

X	Frecuencia
3	1
4	3
5	3
6	1
7	1

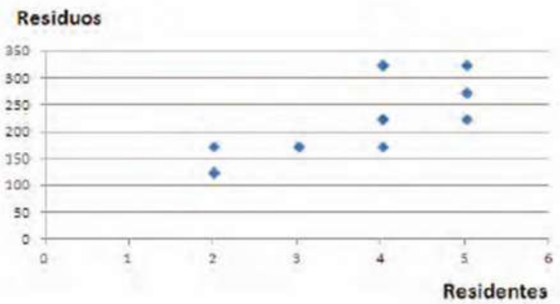
Y	Frecuencia
5	1
6	1
7	3
8	3
9	1

9 **Sobre una muestra de 20 hogares se ha realizado un estudio acerca del número de residentes que hay en la vivienda y la cantidad de residuos generados mensualmente, en kilogramos:**
(5 , 200) (4 , 220) (3 , 190) (5 , 250) (4 , 180) (3 , 150) (3 , 195) (2 , 150) (5 , 250) (4 , 300) (4 , 230) (4 , 220) (3 , 150) (5 , 300), (2 , 140) (5 , 270) (2 , 100) (4 , 310) (5 , 280), (4 , 230)

a. Elabora una tabla de doble entrada agrupando los datos de los residuos en intervalos.

y_j / x_i	[100 , 150)	[150 , 200)	[200 , 250)	[250 , 300)	[300 , 350]	Frecuencia absoluta de *X*
2	2	1	0	0	0	3
3	0	4	0	0	0	4
4	0	1	4	0	2	7
5	0	0	1	4	1	6
Frecuencia absoluta de *Y*	2	6	5	4	3	20

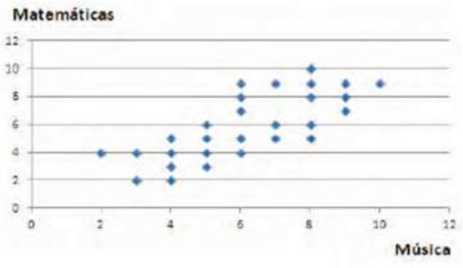
b. Representa la nube de puntos.



c. Halla la distribución marginal de las variables *X* e *Y*.

X	Frecuencia
2	3
3	4
4	7
5	6

10 **Las calificaciones obtenidas por 28 alumnos en las materias de Música y de Matemáticas, en este orden, han sido las siguientes: (7 , 9) (6 , 4) (8 , 10) (8 , 5) (6 , 5) (9 , 8) (4 , 2) (8 , 8) (5 , 4) (7 , 5) (8 , 9) (5 , 3) (6 , 7) (6 , 9) (7 , 6) (9 , 9) (6 , 8) (10 , 9) (3 , 4) (8 , 6) (5 , 6) (3 , 2) (2 , 4) (5 , 5) (4 , 5) (4 , 4) (9 , 7) (4 , 3)**
Representa el diagrama de dispersión de estos datos.



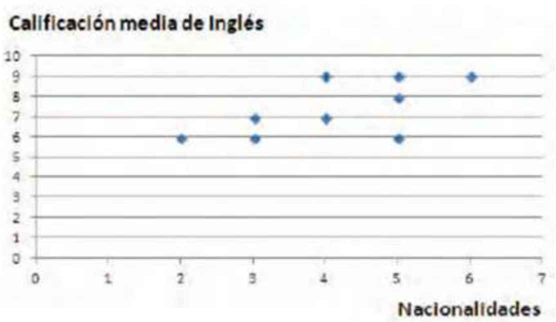
11 **En un centro escolar se ha realizado un recuento del número de nacionalidades diferentes que hay en cada clase (*X*) y la calificación media de la clase en la materia de Inglés (*Y*). Los resultados son:**

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

a. Agrupa los datos en una tabla simple.

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

b. Representa gráficamente los puntos en un diagrama de dispersión.



c. ¿En cuántas clases hay 5 nacionalidades?
En tres clases.

d. ¿En cuántas clases la nota media en Inglés es de un 9?
En tres clases.

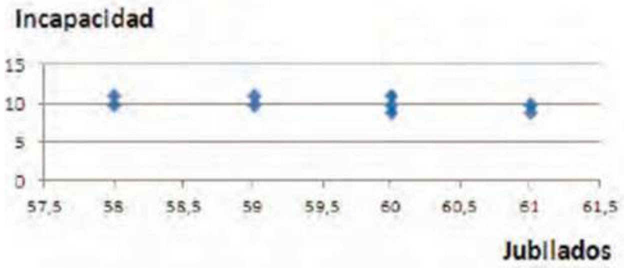
12 **El porcentaje del presupuesto de la Seguridad Social de varios países que se distribuye a personas jubiladas y a personas con incapacidad permanente es el siguiente:**

Jubilados	60	58	61	59	61	60	59	58	60
Incapacidad	10	11	9	10	10	9	11	10	11

a. Agrupa los datos en una tabla simple.

x_i	y_i
58	10
58	11
59	10
59	11
60	9
60	10
60	11
61	9
61	10

b. Representa gráficamente los puntos en un diagrama de dispersión.



c. ¿Cuántos países destinan un 60 % del presupuesto de la Seguridad Social a los jubilados?

Tres países.

d. ¿Cuántos países dedican más de un 9 % del presupuesto de la Seguridad Social a las personas con incapacidad permanente?

Siete países.

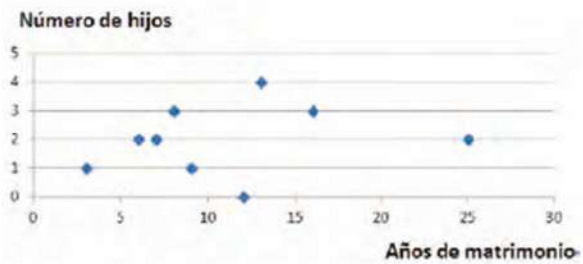
13 **Se ha preguntado a diversos matrimonios por el número de años que llevan casados y el número de hijos que han tenido desde que contrajeron matrimonio:**

Años	3	8	6	12	25	13	9	7	16
Hijos	1	3	2	0	2	4	1	2	3

a. Elabora una tabla simple de los datos.

x_i	y_i
3	1
6	2
7	2
8	3
9	1
12	0
13	4
16	3
25	2

b. Representa la nube de puntos.



c. ¿Cuántos matrimonios han tenido menos de 2 hijos?
Tres matrimonios.

d. ¿Cuántos años llevan casados los matrimonios que han tenido tres hijos?
Ocho y dieciséis.

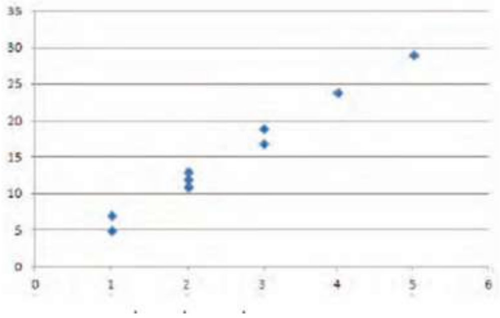
e. ¿Qué matrimonios han tenido mayor número de hijos: los que llevan casados más de 10 años o los que llevan menos de 10 años?
En ambos casos el número de hijos es el mismo.

SOLUCIONES PÁG. 275

14 Observa la siguiente distribución de esta variable bidimensional:

x_i	1	2	3	2	1	4	2	5	3
y_i	7	11	19	13	5	24	12	29	17

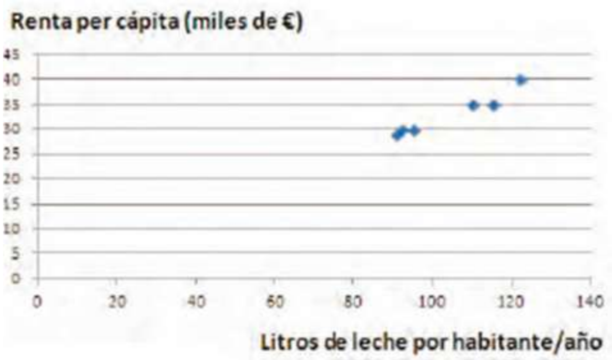
a. Dibuja la nube de puntos correspondiente.



b. Clasifica el tipo de dependencia que existe entre las variables unidimensionales.
Mantienen una dependencia aleatoria.

15 Se cree que el consumo de leche en la dieta de los países desarrollados está relacionado con la renta per cápita de cada uno. Analiza si esto se cumple y si existe algún tipo de relación entre estas variables a tenor de los siguientes resultados:

Litros de leche por habitante/año	92	110	122	91	115	95
Renta per cápita (miles €)	30	35	40	29	35	30



Existe una dependencia aleatoria o estadística entre las variables y, por tanto, una cierta influencia entre ellas.

16 Indica el tipo de dependencia que crees que puede existir en las siguientes variables bidimensionales:

a. El cambio de divisas entre libras esterlinas y euros.
Dependencia funcional.

b. Las horas de estudio y la calificación obtenida en una materia.
Dependencia aleatoria.

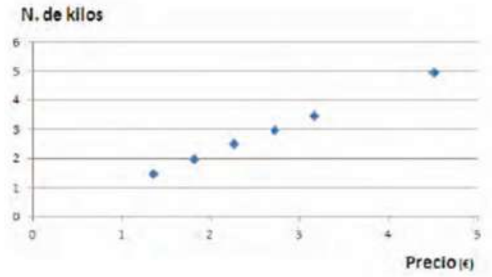
17 **Realiza el siguiente trabajo en grupo en clase: pregunta a tus compañeros la calificación obtenida en el último examen de Matemáticas (*X*) y el número de horas dedicadas a la preparación de la prueba durante la última semana (*Y*). Representa los datos en una nube de puntos y analiza el tipo de dependencia existente entre ambas variables.**

Respuesta abierta.

18 **Se ha comparado el precio y el número de kilos de naranjas que han adquirido los integrantes de un grupo de vecinos:**

Precio (€)	2,70	4,50	1,80	1,35	3,15	2,25
N.º de kilos	3	5	2	1,5	3,5	2,5

a. Representa gráficamente los datos mediante una nube de puntos.



b. Indica si existe dependencia entre los valores de las variables, y, en caso afirmativo, di de qué tipo es.

Sí, existe una dependencia de tipo funcional ya que los valores de la variable *X* determinan exactamente los de la *Y*.

19 **Fíjate en estos pares de variables bidimensionales y razona cuáles de ellos están integrados por variables unidimensionales que guardan una relación de dependencia y señala de qué tipo es:**

a. El plato preferido de una persona y su peso.

Independencia.

b. El número de horas dedicado a preparar una prueba de atletismo y la marca obtenida.

Dependencia aleatoria.

c. La modalidad de bachillerato elegida y el sexo del estudiante.

Independencia.

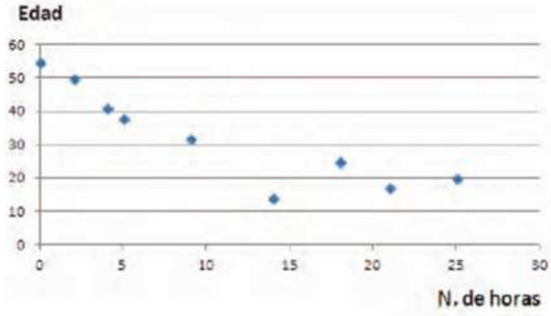
d. La fuerza con que una carga puntual fija atrae o repele a otras cargas colocadas a una distancia fija.

Dependencia funcional.

20 **Un estudio intenta determinar si existe alguna relación entre las siguientes variables: número de horas semanales dedicadas a jugar a videojuegos (*X*) y edad del jugador (*Y*). Ha arrojado los siguientes resultados:**

N.º de horas	14	21	18	9	25	4	0	2	5
Edad	14	17	25	32	20	41	55	50	38

a. Representa gráficamente los datos mediante un diagrama de dispersión.



b. ¿Qué conclusiones puedes extraer sobre la dependencia de ambas variables?

Las variables mantienen una dependencia aleatoria. Parecen seguir un rastro curvilíneo.

21 **Indica dos ejemplos de variables bidimensionales para cada uno de estos tipos de dependencia:**

a. Dependencia aleatoria.

Respuesta abierta, por ejemplo: Altura de una persona y tamaño del pie.

b. Dependencia funcional.

Respuesta abierta, por ejemplo: Grados Celsius y grados Kelvin.

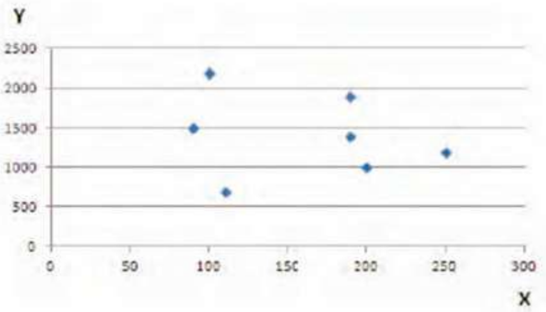
c. Independencia.

Respuesta abierta, por ejemplo: Peso de un alumno y calificación en matemáticas.

22 En la siguiente tabla se reflejan los datos recogidos en un estudio comparativo entre la variable precio del móvil de una persona (variable *X*) y sus ingresos medios mensuales (variable *Y*):

X	200	90	190	250	100	190	110	250	100
Y	1000	1500	1900	1200	2200	1400	700	1200	2200

a. Representa gráficamente los datos mediante un diagrama de dispersión.



b. A la vista de la nube de puntos obtenida, ¿qué puedes indicar sobre la dependencia de ambas variables?

Son independientes.

c. Elabora una tabla con unos nuevos valores que establezcan una dependencia diferente a la obtenida con la tabla actual.

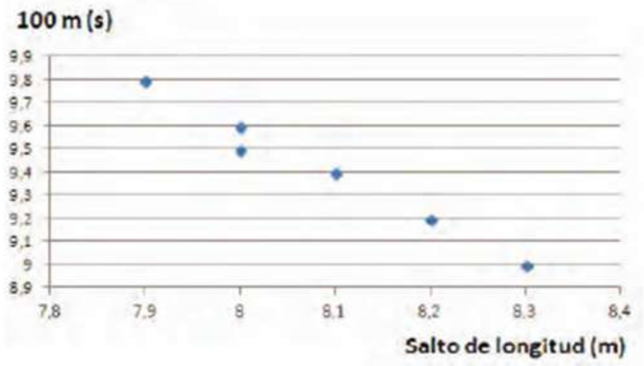
Respuesta abierta.

SOLUCIONES PÁG. 277

23 Un grupo de atletas de decatlón ha realizado sus dos últimas pruebas, salto de longitud y 100 metros lisos, con los siguientes registros:

Salto de longitud (m)	8,1	8,0	7,9	8,2	8	8,3
100 m (s)	9,4	9,6	9,8	9,2	9,5	9

Representa el diagrama de dispersión de los puntos e indica el tipo de correlación que hay entre las variables.



Hay una correlación lineal y negativa.

24 Actividad resuelta.

25 Halla la covarianza en la siguiente distribución de una variable bidimensional:

x_i	3	3	5	2	3	4	3	3	4
y_i	6	8	12	5	9	8	11	3	9

x_i	y_i	$x_i \cdot y_i$
3	6	18
3	8	24
5	12	60
2	5	10
3	9	27
4	8	32
3	11	33
3	3	9
4	9	36
30	71	249

$\bar{x} = \frac{x_1+\cdots+x_n}{\mathrm{N}} = \frac{30}{9} = 3{,}33$; $\bar{y} = \frac{y_1+\cdots+y_n}{\mathrm{N}} = \frac{71}{9} = 7{,}89$;

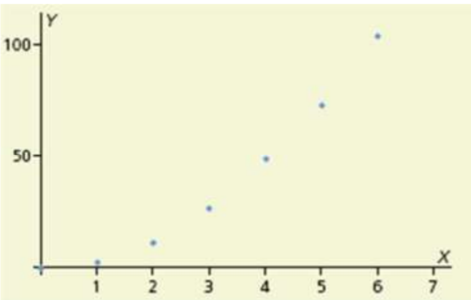
$S_{xy} = \frac{\sum_{i,j=1}^{k} n_{ij} \cdot x_i \cdot y_i}{\mathrm{N}} - \bar{x} \cdot \bar{y} = \frac{249}{9} - (3{,}33 \cdot 7{,}89) = 1{,}39$

26 Escribe un ejemplo de la distribución de una variable cuya covarianza sea nula.

Respuesta abierta.

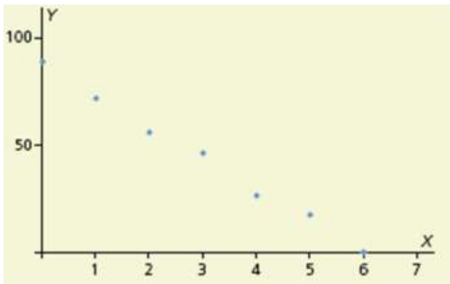
27 Clasifica los siguientes diagramas de dispersión en función del tipo de correlación que existe entre las variables:

a.



Correlación curvilínea.

b.

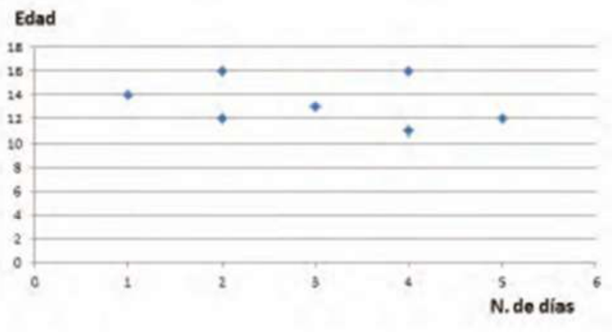


Correlación negativa.

28 **Para mejorar el plan de absentismo escolar en un centro educativo, se ha realizado un estudio con el objetivo de determinar la relación entre el número de días de ausencias semanales en el centro y la edad de los alumnos. Los resultados obtenidos son:**

N.º de días (x_i)	2	4	1	5	3	4	2
Edad (y_i)	12	16	14	12	13	11	16

a. Realiza la representación gráfica en una nube de puntos.



b. Determina el tipo de correlación que mantienen las variables.
La correlación es nula.

SOLUCIONES PÁG. 279

29 Durante el viaje de estudios de fin de etapa, un grupo de alumnos realizó un cambio de divisas, de euros a libras esterlinas, en una oficina bancaria:

Euros (x_i)	12	24	18	30	6	36
Libras (y_i)	10	20	15	25	5	30

Halla el coeficiente de correlación de los datos y razona de qué tipo de correlación se trata en función del valor obtenido.

x_i	y_i	$x_i \cdot y_i$	x_i^2	y_i^2
12	10	120	144	100
24	20	480	576	400
18	15	270	324	225
30	25	750	900	625
6	5	30	36	25
36	30	1 080	1 296	900
126	105	2 730	3 276	2 275

$\bar{x} = \frac{x_1 + \cdots + x_n}{N} = \frac{126}{6} = 21$; $\bar{y} = \frac{y_1 + \cdots + y_n}{N} = \frac{105}{6} = 17{,}5$

$S_{xy} = \frac{\sum_{i,j=1}^{k} n_{ij} \cdot x_i \cdot y_i}{N} - \bar{x} \cdot \bar{y} = \frac{2\,730}{6} - (21 \cdot 17{,}5) = 87{,}5$

$V(x) = \frac{\sum_{i=1}^{n} x_i^2 \cdot n_i}{N} - \bar{x}^2 = \frac{3\,276}{6} - 21^2 = 105$; $S(x) = 10{,}24$

$V(y) = \frac{\sum_{i=1}^{n} y_i^2 \cdot n_i}{N} - \bar{y}^2 = \frac{2\,275}{6} - 17{,}5^2 = 72{,}91$; $S(y) = 8{,}53$

$r = \frac{S_{xy}}{S_x \cdot S_y} = \frac{87{,}5}{10{,}24 \cdot 8{,}53} = 1$

Por tanto, tienen una correlación funcional lineal positiva.

30 La masa y la altura de un grupo de personas es:

Masa (kg), x_i	80	75	70	78	72
Altura (cm), y_i	185	170	165	171	168

a. Halla su coeficiente de correlación.

x_i	y_i	$x_i \cdot y_i$	x_i^2	y_i^2
80	185	14 800	6 400	34 225
75	170	12 750	5 625	28 900
70	165	11 550	4 900	27 225
78	171	13 338	6 084	29 241
72	168	12 096	5 184	28 224
375	859	64 534	28 193	147 815

$\bar{x} = \frac{x_1+\cdots+x_n}{\mathrm{N}} = \frac{375}{5} = 75$; $\bar{y} = \frac{y_1+\cdots+y_n}{\mathrm{N}} = \frac{859}{5} = 171{,}8$

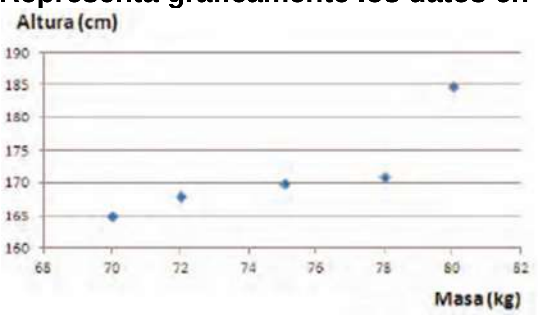
$S_{xy} = \frac{\sum_{i,j=1}^{k} n_{ij} \cdot x_i \cdot y_i}{\mathrm{N}} - \bar{x} \cdot \bar{y} = \frac{64\,534}{5} - (75 \cdot 171{,}8) = 21{,}8$

$V(x) = \frac{\sum_{i=1}^{n} x_i^2 \cdot n_i}{\mathrm{N}} - \bar{x}^2 = \frac{28\,193}{5} - 75^2 = 13{,}6$; $S(x) = 3{,}6$

$V(y) = \frac{\sum_{i=1}^{n} y_i^2 \cdot n_i}{\mathrm{N}} - \bar{y}^2 = \frac{147\,815}{5} - 171{,}8^2 = 47{,}76$; $S(y) = 6{,}91$

$r = \frac{S_{xy}}{S_x \cdot S_y} = \frac{21{,}8}{3{,}6 \cdot 6{,}91} = 0{,}87$

b. Representa gráficamente los datos en una nube de puntos.

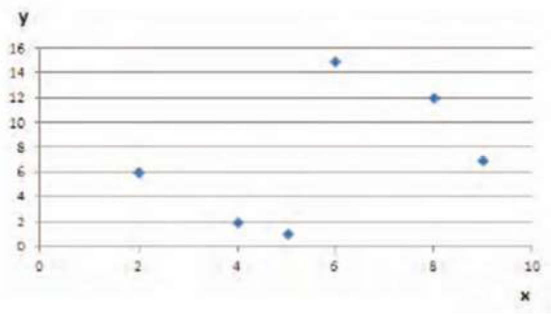


c. Describe el tipo de correlación existente entre las variables.
La correlación es positiva y fuerte.

31 Representa los siguientes datos mediante una nube de puntos e indica cuál de estos valores te parece más apropiado para el coeficiente de correlación:

0,9 0,1 –0,8 –0,6

x_i	2	4	5	6	8	9
y_i	6	2	1	15	12	7



El más apropiado es r = 0,1.

32 Calcula el coeficiente de correlación del siguiente conjunto de datos e interprétalo:

x_i	1	2	3	4	5	6	7
y_i	17	15	14	12	9	9	6

x_i	y_i	$x_i \cdot y_i$	x_i^2	y_i^2
1	17	17	1	289
2	15	30	4	225
3	14	42	9	196
4	12	48	16	144
5	9	45	25	81
6	9	54	36	81
7	6	42	49	36
28	82	278	140	1 052

$\bar{x} = \frac{x_1 + \cdots + x_n}{N} = \frac{28}{7} = 4$; $\bar{y} = \frac{y_1 + \cdots + y_n}{N} = \frac{82}{7} = 11,71$

$S_{xy} = \frac{\sum_{i,j=1}^{k} n_{ij} \cdot x_i \cdot y_i}{N} - \bar{x} \cdot \bar{y} = \frac{278}{7} - (4 \cdot 11,71) = -7,12$

$V(x) = \frac{\sum_{i=1}^{n} x_i^2 \cdot n_i}{N} - \bar{x}^2 = \frac{140}{7} - 4^2 = 4$; $S(x) = 2$

$V(y) = \frac{\sum_{i=1}^{n} y_i^2 \cdot n_i}{N} - \bar{y} = \frac{1\,052}{7} - 11,71^2 = 13,16$; $S(y) = 3,62$

$r = \frac{S_{xy}}{S_x \cdot S_y} = \frac{-7,12}{2 \cdot 3,62} = -0,98$

La distribución tiene coeficiente de correlación muy cercano a 1 en valor absoluto, lo cual indica que está muy próximo a tener una dependencia funcional negativa.

33 Se ha llevado a cabo un estudio sobre los tiempos realizados por unos nadadores en la prueba de 100 metros estilo libre y su estatura y se ha comprobado que no existe ninguna relación entre ambas variables. Responde razonadamente a las siguientes cuestiones:

a. ¿Cuál crees que es el valor del coeficiente de correlación lineal?
Cercano a 0.

b. Inventa una nube de puntos que represente la distribución de este conjunto de datos.
Respuesta abierta.

34 Explica el significado de los siguientes coeficientes de correlación en una distribución bidimensional:

a. r = 1
Hay una correlación funcional y positiva.

b. r = –0,1
La correlación es muy débil.

c. r = 0,25
Hay una correlación aleatoria leve y positiva.

d. r = 0,92
Hay una correlación aleatoria fuerte y positiva.

e. r = 0
La correlación es nula.

f. r = –1
Hay una correlación funcional y negativa.

35 Se conoce el coeficiente de correlación de dos variables bidimensionales cuyos valores son $r_1 = -0{,}7$ y $r_2 = 0{,}1$.

a. ¿En cuál de las dos distribuciones se ajustarán mejor los datos a una función lineal?
En la primera distribución, pues su coeficiente de correlación es más cercano al valor absoluto de 1.

b. Utilizando un diagrama de dispersión, representa gráficamente, de forma aproximada, cada conjunto de valores definido por los coeficientes de correlación anteriores.
Respuesta abierta.

36 En el caso de que entre los elementos de una variable estadística bidimensional haya independencia estadística, cabe concluir que la covarianza es nula. Sin embargo, ¿se puede afirmar el proceso contrario? Es decir, ¿implica el hecho de que la covarianza de las variables sea cero que haya una independencia estadística entre dichas variables?

Una covarianza nula no implica independencia, sino falta de relación lineal.

37 Interpreta el valor del coeficiente de correlación de:

X	1	2	4	6	7	9	10
Y	–1	–4	–10	–16	–19	–25	–28

x_i	y_i	$x_i \cdot y_i$	x_i^2	y_i^2
1	–1	–1	1	1
2	–4	–8	4	16
4	–10	–40	16	100
6	–16	–96	36	256
7	–19	–133	49	361
9	–25	–225	81	625
10	–28	–280	100	784
39	–103	–783	287	2 143

$\bar{x} = \frac{x_1+\cdots+x_n}{\mathrm{N}} = \frac{39}{7} = 5{,}57$; $\bar{y} = \frac{y_1+\cdots+y_n}{\mathrm{N}} = \frac{-103}{7} = -14{,}71$

$S_{xy} = \frac{\sum_{i,j=1}^{k} n_{ij} \cdot x_i \cdot y_i}{\mathrm{N}} - \bar{x} \cdot \bar{y} = \frac{-783}{7} - (5{,}57 \cdot (-14{,}71)) = -29{,}89$

$V(x) = \frac{\sum_{i=1}^{n} x_i^2 \cdot n_i}{\mathrm{N}} - \bar{x}^2 = \frac{287}{7} - 5{,}57^2 = 8{,}28$; $S(x) = 2{,}87$

$V(y) = \frac{\sum_{i=1}^{n} y_i^2 \cdot n_i}{\mathrm{N}} - \bar{y}^2 = \frac{2\,143}{7} - (-14{,}71)^2 = 89{,}75$; $S(y) = 9{,}47$

$r = \frac{S_{xy}}{S_x \cdot S_y} = \frac{-29{,}89}{2{,}87 \cdot 9{,}47} = -1{,}09$

Correlación negativa y lineal, dependencia funcional, los puntos se ajustan a una recta de pendiente negativa, r = –1.

SOLUCIONES PÁG. 281

38 El Ministerio de Medio Ambiente ha llevado a cabo un estudio sobre el número de incendios forestales que han tenido lugar en una comunidad y las temperaturas medias ambientales registradas durante esos días:

N.º de incendios (x_i)	2	6	9	5	3	1	6
Temperatura media, en °C (y_i)	30	35	39	34	32	29	36

a. Halla el coeficiente de correlación lineal.

x_i	y_i	$x_i \cdot y_i$	x_i^2	y_i^2
2	30	60	4	900
6	35	210	36	1 225
9	39	351	81	1 521
5	34	170	25	1 156
3	32	96	9	1 024
1	29	29	1	841
6	36	216	36	1 296
32	235	1 132	192	7 963

$\bar{x} = \frac{x_1+\cdots+x_n}{\mathrm{N}} = \frac{32}{7} = 4{,}57$; $\bar{y} = \frac{y_1+\cdots+y_n}{\mathrm{N}} = \frac{235}{7} = 33{,}57$

$S_{xy} = \frac{\sum_{i,j=1}^{k} n_{ij} \cdot x_i \cdot y_i}{\mathrm{N}} - \bar{x} \cdot \bar{y} = \frac{1\,132}{7} - (4{,}57 \cdot 33{,}57) = 8{,}3$

$V(x) = \frac{\sum_{i=1}^{n} x_i^2 \cdot n_i}{\mathrm{N}} - \bar{x}^2 = \frac{192}{7} - 4{,}57^2 = 6{,}54$; $S(x) = 2{,}56$

$V(y) = \frac{\sum_{i=1}^{n} y_i^2 \cdot n_i}{\mathrm{N}} - \bar{y}^2 = \frac{7963}{7} - 33{,}57^2 = 10{,}63$; $S(y) = 3{,}26$

$r = \frac{S_{xy}}{S_x \cdot S_y} = \frac{8{,}30}{2{,}56 \cdot 3{,}26} = 0{,}99$

b. Calcula la recta de regresión de *y* sobre *x*.

$y - \bar{y} = \frac{S_{xy}}{S_x^2} \cdot (x - \bar{x}); y - 33{,}57 = \frac{8{,}3}{2{,}56^2} \cdot (x - 4{,}57); y - 33{,}57 = 1{,}27x - 5{,}79;$

$y = 1{,}27x + 27{,}78$

c. Si se produjeran 12 incendios, ¿qué previsión de temperatura media se podría realizar?

$y = 1{,}27 \cdot 12 + 27{,}78$; $y = 43{,}02$

La previsión sería de 43,02 ºC

d. ¿Es la previsión fiable? Razona la respuesta.

Es bastante fiable, ya que el coeficiente de correlación lineal se aproxima mucho a 1.

39 **Dos conjuntos de datos tienen un coeficiente de correlación lineal de r_1 = 0,6 y r_2 = –0,8, respectivamente. ¿En cuál de ellos será mejor el ajuste de la recta de regresión? Razona tu respuesta.**

La estimación será más fiable en el segundo caso porque el coeficiente de correlación es más cercano, en valor absoluto a 1.

40 **En los siguientes apartados se da un coeficiente de correlación lineal; indica en cuáles tiene sentido calcular la recta de regresión.**

a. r = 0,9 **b. r = 0** **c. r = –0,1** **d. r = –1**

Únicamente tiene sentido calcular la recta de regresión tanto en el apartado a. como en el apartado d., ya que en ambos casos es un valor cercano o igual a 1 en valor absoluto. Sin embargo, en los otros dos apartados, el coeficiente de correlación es igual o está muy cercano al 0.

41 **Contesta razonadamente a las siguientes preguntas:**

a. ¿Qué punto tienen en común las dos rectas de regresión de una variable bidimensional?
$(\bar{x}, \bar{y})$, pues es el centro de gravedad de la nube de puntos.

b. Si el valor absoluto del coeficiente de correlación lineal se aproxima a 1, ¿a qué tienden las rectas de regresión de una variable bidimensional?
A que su ángulo de corte sea nulo, a que se igualen ambas rectas.

SOLUCIONES PÁG. 283

1 **Explica la regla para agrupar en intervalos dentro de una tabla de frecuencias los valores de una variable cuantitativa continua.**

El número de intervalos se aproxima a $\sqrt{\mathrm{N}}$. La amplitud de los intervalos se calcula como: $\frac{X_{\max}-X_{\min}}{\text{n.º intervalos}}$

2 **Enuncia las etapas que tienen lugar en el proceso de elaboración de un estudio estadístico.**

1. Elección de la población y del carácter a estudiar.
2. Diseño de la encuesta a realizar y del proceso de recogida de los datos.
3. Elección de la muestra, de forma que sea representativa de la población.
4. Recuento de los datos obtenidos y su organización en tablas.
5. Representación gráfica de la información obtenida.
6. Cálculo de los parámetros estadísticos.
7. Análisis de la información obtenida.

3 **¿Cuánto suman las frecuencias relativas de todos los valores de una variable unidimensional?**

Suman la unidad.

4 **Explica la diferencia entre los parámetros de centralización y los parámetros de dispersión de una variable estadística unidimensional.**

Los parámetros de centralización son valores alrededor de los cuales se distribuyen los datos y su labor es representar de forma global a toda la población, mientras que los parámetros de dispersión indican el grado de concentración de los datos en torno a los parámetros de centralización.

5 **¿En cuál de los parámetros estadísticos no influyen todos los datos?**

El rango o recorrido, pues solo depende de los valores extremos.

6 **¿En cuál de los parámetros de centralización influye el orden de los datos?**

En la mediana.

7 **¿Cuál es la diferencia entre una dependencia aleatoria y una dependencia funcional en una variable estadística bidimensional?**

En la dependencia aleatoria entre los valores de las variables existe una cierta influencia, pero los valores de la variable *X* no determinan los valores correspondientes de la variable *Y*. Mientras que en la funcional la influencia es absoluta y sí determinan los valores correspondientes de la otra variable.

8 **¿Qué se debe cumplir en la correlación de una variable estadística bidimensional para que esta sea positiva?**

Se debe cumplir que a medida que aumentan los valores de una variable, aumenten los valores de la otra variable.

9 **¿Qué determina el coeficiente de correlación lineal de una variable estadística bidimensional?**

El grado de correlación entre las variables.

10 **Prepara una presentación digital para tus compañeros sobre el origen de la estadística y su evolución a lo largo de la historia. Puedes hacer un documento PowerPoint, usar Glogster…**

Respuesta abierta.

SOLUCIONES PÁG. 284

ETAPAS DE UN ESTUDIO ESTADÍSTICO

1 Se desea extraer una muestra de 20 personas de un conjunto de 60 individuos en el que hay 3 grupos de edades: 12 niños, 30 adultos y 18 ancianos. Indica como realizarías la muestra si el tipo de muestreo fuera:

a. Aleatorio sistemático.

Se calcula k = $\frac{\mathrm{N}}{\mathrm{n}} = \frac{60}{20} =$ 3. Se elige al azar un número entre el 1 al 3, por ejemplo, el 2. Entonces las personas que integrarán la muestra serán los correspondientes a los números 2, 5, 8, 11, …, 59.

b. Aleatorio estratificado.

El tamaño muestral de cada estrato, niños, adultos y ancianos deberá ser proporcional a la presencia del mismo en la población:

$\frac{12}{60} = \frac{x}{20} \rightarrow x$ = 4 niños

$\frac{30}{60} = \frac{y}{20} \rightarrow y$ = 10 adultos

$\frac{18}{60} = \frac{z}{20} \rightarrow z$ = 6 ancianos

ESTADÍSTICA UNIDIMENSIONAL

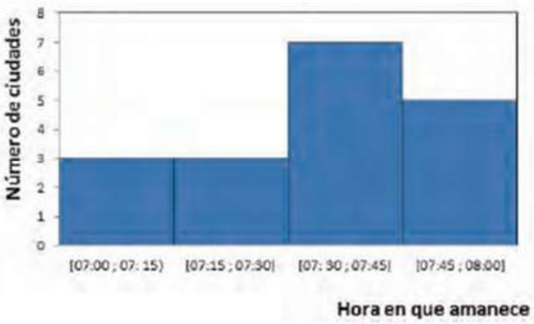
2 La hora a la que amanece en diferentes ciudades es:

07:32	**07:45**	**07:06**	**07:52**	**07:04**	**07:39**
07:27	**07:39**	**07:38**	**07:43**	**07:00**	**07:26**
07:35	**07:54**	**07:23**	**07:30**	**07:47**	**07:55**

a. Organiza los datos en una tabla estadística con cuatro intervalos de 15 min de amplitud.

x_i	c_i	n_i	N_i	f_i	p_i
[07:00 ; 07:15)	07:07, 30′	3	3	0,17	17 %
[07:15 ; 07:30)	07:22, 30′	3	6	0,17	17 %
[07: 30 ; 07:45)	07:37, 30′	7	13	0,39	39 %
[07:45 ; 08:00]	07:52, 30′	5	18	0,28	28 %
Total		**N = 18**		**1**	**100 %**

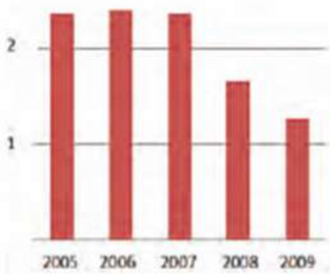
b. Represéntalos en un histograma.

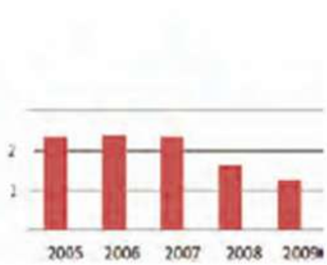


3 **En la siguiente tabla se registran las cifras de automóviles matriculados en cierta comunidad en función del año:**

Año	2005	2006	2007	2008	2009
Vehículos (millones)	2,37	2,40	2,37	1,67	1,27

a. Elabora dos diagramas de barras: uno que se ajuste a la regla de los tres cuartos y otro que la incumpla.





b. Describe la diferencia de información que se transmite en una primera impresión.

Respuesta abierta.

4 **La altura media de 6 alumnos es de 1,80 m, y la de 8 alumnas, de 1,74 m. Halla:**

a. La suma de las alturas de todos los alumnos.

Como $\overline{x} = \frac{x_1+\cdots+x_n}{N}$ = 1,80 → $\sum x_i$ = 6 · 1,80 = 10,8 m

b. La suma de las alturas de todas las alumnas.

Como $\overline{x} = \frac{x_1+\cdots+x_n}{N}$ = 1,74 → $\sum x_i$ = 8 · 1,74 = 13,92 m

c. La altura media de los 14 estudiantes.

$\overline{x} = \frac{x_1+\cdots+x_n}{N} = \frac{10,8+13,92}{6+8}$ = 1,77 m

5 Se ha contabilizado el número de centros educativos operativos en 30 localidades y se han obtenido estos datos:

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

a. Elabora una tabla estadística con estos datos.

Centros educativos, x_i	n_i	N_i	f_i	p_i
2	1	1	0,05	5 %
3	2	3	0,1	10 %
4	5	8	0,25	25 %
5	5	13	0,25	25 %
6	3	16	0,15	15 %
7	2	18	0,1	10 %
8	2	20	0,1	10 %
Total	N = 20		1	100 %

b. Represéntalos en un diagrama de sectores.

c. Halla los parámetros estadísticos.

Mo = 4 y 5 porque son los valores que más se repiten de la variable;

$\bar{x} = \frac{x_1+\cdots+x_n}{\mathrm{N}} = \frac{2\cdot1+3\cdot2+4\cdot5+5\cdot5+6\cdot3+7\cdot2+8\cdot2}{20} = \frac{101}{20} = 5{,}05$;

Me = 5 porque $\frac{20}{2} = 10$ y el valor que supera 10 en la columna de N_i se corresponde con el $X_i = 5$;

R = $X_{max} - X_{min} = 8 - 10 = 6$;

V(x) = $\frac{\sum_{i=1}^{n} x_i^2 \cdot n_i}{\mathrm{N}} - \bar{x}^2 = \frac{2^2\cdot1+3^2\cdot2+4^2\cdot5+5^2\cdot5+6^2\cdot3+7^2\cdot2+8^2\cdot2}{20} - (5{,}05)^2 = 2{,}55$

$S_x = +\sqrt{\mathrm{V}(x)} = 1{,}6$; CV = $\frac{\mathrm{S}\,(x)}{\bar{x}} = \frac{1{,}6}{5{,}05} = 0{,}32$

ESTADÍSTICA BIDIMENSIONAL

6 La temperatura media anual y la latitud de algunas ciudades europeas es, de forma aproximada: (13º , 51º) (16º , 49º) (17º , 47º) (17º , 48º) (16º , 47º) (18º , 47º) (14º , 49º) (18º , 46º) (16º , 48º) (16º , 49º) (15º , 49º) (14º , 48º)

a. Recoge los valores en una tabla de doble entrada.

x_i \ y_i	46	47	48	49	51	Frecuencia absoluta de *X*
13	0	0	0	0	1	1
14	0	0	1	1	0	2
15	0	0	0	1	0	1
16	0	1	1	2	0	4
17	0	1	1	0	0	2
18	1	1	0	0	0	2
Frecuencia absoluta de *Y*	1	3	3	4	1	12

b. Halla la distribución marginal de las variables *X* e *Y*.

X	Frecuencia
13	1
14	2
15	1
16	4
17	2
18	2

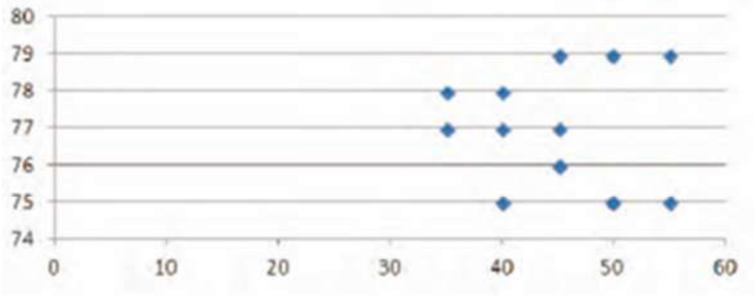
Y	Frecuencia
46	1
47	3
48	3
49	4
51	1

7 **Para comprobar el ahorro que se puede hacer en la compra diaria, se ha registrado el precio de una barra de pan y de un litro de leche, en céntimos de euro, en diferentes supermercados de una ciudad:**
(45 , 79) (40 , 77) (40 , 78) (45 , 77) (50 , 79) (35 , 77) (55 , 75) (45 , 76) (45 , 79) (40 , 75) (50 , 75) (45 , 76) (45 , 79) (35 , 78) (55 , 79)

a. Recoge los valores en una tabla de doble entrada.

x_i \ y_j	75	76	77	78	79	Frecuencia absoluta de *X*
35	0	0	1	1	0	2
40	1	0	1	1	0	3
45	0	2	1	0	3	6
50	1	0	0	0	1	2
55	1	0	0	0	1	2
Frecuencia absoluta de *Y*	3	2	3	2	5	15

b. Representa el diagrama de dispersión de los puntos.



c. Halla la distribución marginal de las variables *X* e *Y*.

X	Frecuencia
35	2
40	3
45	6
50	2
55	2

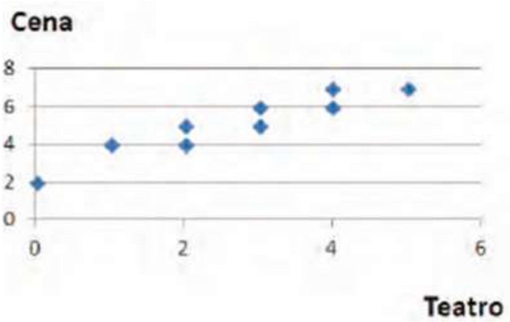
Y	Frecuencia
75	3
76	2
77	3
78	2
79	5

DEPENDENCIA ALEATORIA Y FUNCIONAL

8 El número de veces que un grupo de individuos de una muestra acude mensualmente al teatro y sale a cenar fuera de casa es:

Teatro	1	5	2	4	0	3	3	2	4
Cena	4	7	4	7	2	5	6	5	6

Dibuja la nube de puntos correspondiente y clasifica el tipo de dependencia que existe entre las variables.



Hay una dependencia aleatoria fuerte entre las variables.

CORRELACIÓN ENTRE DOS VARIABLES

9 **Se ha analizado en varias casas el porcentaje del sueldo que se destina al gasto de la vivienda (gas, agua, luz…) y al de ocio y espectáculos:**

x_i	30	31	30	32	28	31	33
y_i	6	7	7	5	9	6	5

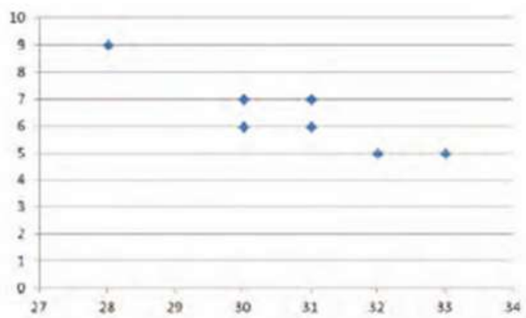
a. Halla la covarianza de esta distribución.

x_i	y_i	$x_i \cdot y_i$
30	6	180
31	7	217
30	7	210
32	5	160
28	9	252
31	6	186
33	5	165
215	45	1 370

$\bar{x} = \frac{x_1+\cdots+x_n}{\mathrm{N}} = \frac{215}{7} = 30{,}71$; $\bar{y} = \frac{y_1+\cdots+y_n}{\mathrm{N}} = \frac{45}{7} = 6{,}43$

$S_{xy} = \frac{\sum_{i,j=1}^{k} n_{ij} \cdot x_i \cdot y_i}{\mathrm{N}} - \bar{x} \cdot \bar{y} = \frac{1\,370}{7} - (30{,}71 \cdot 6{,}43) = -1{,}75$

b. Representa el diagrama de dispersión de los puntos e indica el tipo de correlación entre las variables.



Hay una dependencia aleatoria fuerte y negativa.

SOLUCIONES PÁG. 285

COEFICIENTE DE CORRELACIÓN LINEAL. INTERPRETACIÓN

10 **En el marco de un estudio sobre el fomento del aprendizaje del idioma inglés se ha relacionado el promedio de horas que un grupo de alumnos dedican a ver diariamente los programas de televisión en versión original con la calificación que obtienen en la materia de Inglés:**

Horas (x_i)	1	2	2	1	3	1	2	0	2
Nota (y_i)	7	9	8	5	10	6	7	4	8

a. Halla su coeficiente de correlación lineal.

x_i	y_i	$x_i \cdot y_i$	x_i^2	y_i^2
1	7	7	1	49
2	9	18	4	81
2	8	16	4	64
1	5	5	1	25
3	10	30	9	100
1	6	6	1	36
2	7	14	4	49
0	4	0	0	16
2	8	16	4	64
14	64	112	28	484

$\bar{x} = \frac{x_1+\cdots+x_n}{\mathrm{N}} = \frac{14}{9} = 1{,}56$; $\bar{y} = \frac{y_1+\cdots+y_n}{\mathrm{N}} = \frac{64}{9} = 7{,}11$

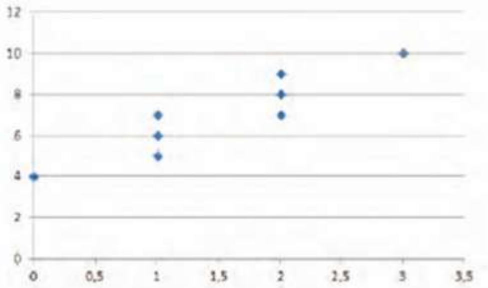
$S_{xy} = \frac{\sum_{i,j=1}^{k} n_{ij} \cdot x_i \cdot y_i}{\mathrm{N}} - \bar{x} \cdot \bar{y} = \frac{112}{9} - (1{,}56 \cdot 7{,}11) = 1{,}35$

$V(x) = \frac{\sum_{i=1}^{n} x_i^2 \cdot n_i}{\mathrm{N}} - \bar{x}^2 = \frac{28}{9} - 1{,}56^2 = 0{,}68$; $S(x) = 0{,}82$

$V(y) = \frac{\sum_{i=1}^{n} y_i^2 \cdot n_i}{\mathrm{N}} - \bar{y}^2 = \frac{484}{9} - 7{,}11^2 = 3{,}23$; $S(y) = 1{,}80$

$r = \frac{S_{xy}}{S_x \cdot S_y} = \frac{1{,}35}{0{,}82 \cdot 1{,}80} = 0{,}91$

b. Representa los datos en una nube de puntos y describe el tipo de correlación entre las variables.



Existe una dependencia aleatoria fuerte y positiva.

11 **Los datos de la tabla indican el PIB per cápita anual, (variable *X*), y la esperanza de vida general, (variable *Y*), en cada uno de los países indicados:**

	Esperanza de vida	PIB
Singapur	83	42 344 €
Luxemburgo	82	87 600 €
Venezuela	76	12 500 €
EEUU	79	41 100 €
Siria	71	1 215 €
Cuba	78,1	5 112 €
Uganda	52,7	538 €
Somalia	53	408 €

a. Halla el coeficiente de correlación de estas dos variables.

x_i	y_i	$x_i \cdot y_i$	x_i^2	y_i^2
83	42 344	3 514 552	6 889	1 793 014 336
82	87 600	7 183 200	6 724	7 673 760 000
76	12 500	950 000	5 776	156 250 000
79	41 100	3 246 900	6 241	1 689 210 000
71	1 215	86 265	5 041	1 476 225
78,1	5 112	399 247,2	6 099,61	26 132 544
52,7	538	283 52,6	2 777,29	289 444
53	408	21 624	2 809	166 464
574,8	190 817	15 430 140,8	42 356,9	11 340 299 013

$\bar{x} = \frac{x_1 + \cdots + x_n}{\mathrm{N}} = \frac{574{,}8}{8} = 71{,}85$; $\bar{y} = \frac{y_1 + \cdots + y_n}{\mathrm{N}} = \frac{190\,817}{8} = 23\,852{,}12$

$S_{xy} = \frac{\sum_{i,j=1}^{k} n_{ij} \cdot x_i \cdot y_i}{\mathrm{N}} - \bar{x} \cdot \bar{y} = \frac{15\,430\,140{,}8}{8} - (71{,}85 \cdot 23\,852{,}12) = 214\,992{,}42$

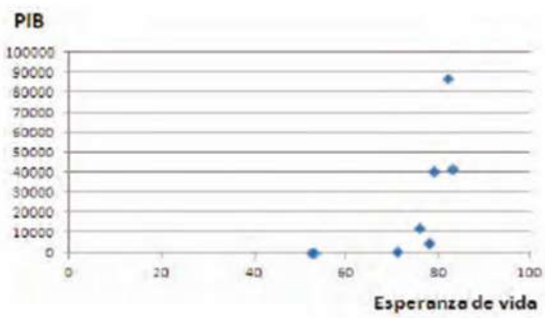
$V(x) = \frac{\sum_{i=1}^{n} x_i^2 \cdot n_i}{\mathrm{N}} - \bar{x}^2 = \frac{42\,356{,}9}{8} - 71{,}85^2 = 132{,}19$; $S(x) = 11{,}50$

$V(y) = \frac{\sum_{i=1}^{n} y_i^2 \cdot n_i}{\mathrm{N}} - \bar{y}^2 = \frac{11\,340\,299\,013}{8} - 23\,852{,}12^2 = 848\,613\,748{,}1$;

$S(y) = 29\,130{,}9$

$r = \frac{S_{xy}}{S_x \cdot S_y} = \frac{214\,992{,}42}{11{,}50 \cdot 29\,130{,}9} = 0{,}64$

b. Representa los datos en un diagrama de puntos e indica el tipo de correlación existente entre las variables.



Correlación positiva y débil.

REGRESIÓN LINEAL. RECTAS DE REGRESIÓN

12 Enrique quiere comprarse una raqueta y ha hecho un pequeño estudio de mercado que tiene en cuenta la masa, en gramos, y el precio, en euros, de varios modelos de raqueta:

Masa (x_i)	330	320	330	340	330
Precio (y_i)	180	200	170	150	190

Halla el coeficiente de correlación lineal y las dos rectas de regresión.

x_i	y_i	$x_i \cdot y_i$	x_i^2	y_i^2
330	180	59 400	108 900	32 400
320	200	64 000	102 400	40 000
330	170	56 100	108 900	28 900
340	150	51 000	115 600	22 500
330	190	62 700	108 900	36 100
1 650	890	293 200	544 700	159 900

$\bar{x} = \frac{x_1+\cdots+x_n}{\mathrm{N}} = \frac{1\,650}{5} = 330$; $\bar{y} = \frac{y_1+\cdots+y_n}{\mathrm{N}} = \frac{890}{5} = 178$

$S_{xy} = \frac{\sum_{i,j=1}^{k} n_{ij} \cdot x_i \cdot y_i}{N} - \bar{x} \cdot \bar{y} = \frac{293\,200}{5} - (330 \cdot 178) = -100$

$V(x) = \frac{\sum_{i=1}^{n} x_i^2 \cdot n_i}{\mathrm{N}} - \bar{x}^2 = \frac{544\,700}{5} - 330^2 = 40$; $S(x) = 6{,}32$

$V(y) = \frac{\sum_{i=1}^{n} y_i^2 \cdot n_i}{\mathrm{N}} - \bar{y}^2 = \frac{159\,900}{5} - 178^2 = 296$; $S(y) = 17{,}20$

$r = \frac{S_{xy}}{S_x \cdot S_y} = \frac{-100}{6{,}32 \cdot 17{,}20} = -0{,}92$

$x - \bar{x} = \frac{s_{xy}}{s_y^2} \cdot (y - \overline{y}) \rightarrow x - 330 = \frac{-100}{296} \cdot (y - 178)$

$y - \bar{y} = \frac{s_{xy}}{s_x^2} \cdot (x - \overline{x}) \rightarrow y - 178 = \frac{-100}{40} \cdot (y - 330)$

EVALUACIÓN

1 El muestreo de tipo probabilístico es:

a. Casual. **b. Aleatorio simple.**

c. De selección. **d. De bola de nieve.**

Solución b.

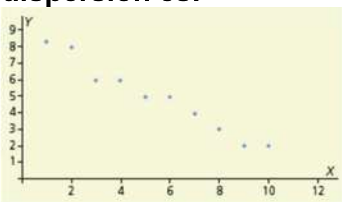
2 La media de estos datos es:

x_i	3	5	8	9	10
n_i	2	1	4	3	1

a. 6,2 **b. 16** **c. 7,27** **d. 18**

$\bar{x} = \frac{x_1+\cdots+x_n}{N} = \frac{6+5+32+27+10}{5} = 16$ Solución b.

3 El tipo de dependencia que mantienen las variables según su diagrama de dispersión es:



a. Positiva. **b. Funcional.** **c. Independiente.** **d. Negativa.**

Solución d.

4 La covarianza de la distribución de esta variable bidimensional es:

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

a. –0,57 **b. 0,57** **c. –0,82** **d. 0,82**

x_i	y_i	$x_i \cdot y_i$	x_i^2	y_i^2
1	4	4	1	16
2	1	2	4	1
3	2	6	9	4
5	1	5	25	1
6	2	12	36	4
7	1	7	49	1
8	3	24	64	9
32	14	60	188	36

$\bar{x} = \frac{x_1+\cdots+x_n}{N} = \frac{32}{7} = 4{,}57$; $\bar{y} = \frac{y_1+\cdots+y_n}{N} = \frac{14}{7} = 2$;

$S_{xy} = \frac{\sum_{i,j=1}^{k} n_{ij} \cdot x_i \cdot y_i}{N} - \bar{x} \cdot \bar{y} = \frac{60}{7} - (4{,}58 \cdot 2) = 0{,}57$

Solución b.

5 El coeficiente de correlación lineal de la distribución de esta variable bidimensional es:

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

a. 0,12 **b. 0,85** **c. –0,49** **d. –0,68**

x_i	y_i	$x_i \cdot y_i$	x_i^2	y_i^2
1	4	4	1	16
2	2	4	4	4
3	1	3	9	1
5	3	15	25	9
6	1	6	36	1
7	1	7	49	1
8	2	16	64	4
32	14	55	188	36

$\bar{x} = \frac{x_1+\cdots+x_n}{\mathrm{N}} = \frac{32}{7} = 4{,}57$; $\bar{y} = \frac{y_1+\cdots+y_n}{\mathrm{N}} = \frac{14}{7} = 2$;

$S_{xy} = \frac{\sum_{i,j=1}^{k} n_{ij} \cdot x_i \cdot y_i}{N} - \bar{x} \cdot \bar{y} = \frac{55}{7} - (4{,}57 \cdot 2) = -1{,}28$

$V(x) = \frac{\sum_{i=1}^{n} x_i^2 \cdot n_i}{N} - \bar{x}^2 = \frac{188}{7} - 4{,}57^2 = 5{,}98$; $S(x) = 2{,}44$

$V(y) = \frac{\sum_{i=1}^{n} y_i^2 \cdot n_i}{N} - \bar{y}^2 = \frac{36}{7} - 2^2 = 1{,}14$; $S(y) = 1{,}06$

$r = \frac{S_{xy}}{S_x \cdot S_y} = \frac{-1{,}28}{2{,}44 \cdot 1{,}06} = -0{,}49$

Solución c.