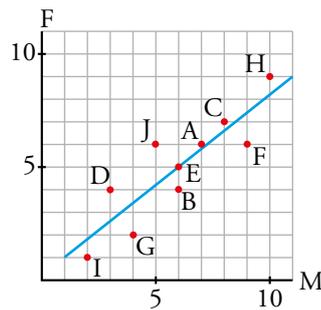


1 Dos variables relacionadas. Correlación

Página 185

1. Identifica los restantes puntos del diagrama de dispersión del ejemplo de las notas en matemáticas y en física.

A cada estudiante a, b, \dots , le corresponderá el punto A, B, ... en el diagrama de dispersión.



Página 187

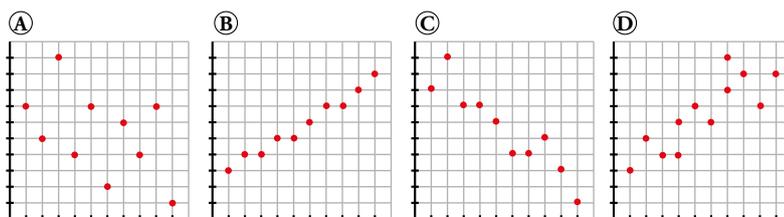
- 2. En las siguientes distribuciones bidimensionales referentes a tus compañeros y compañeras de clase, estima si la correlación será positiva o negativa, muy fuerte, fuerte, débil o casi nula:**
- a) Medida de un palmo - Medida del pie.
 - b) Número de horas semanales de estudio - Número de horas semanales viendo la televisión.
 - c) Número de horas semanales de estudio - Número de suspensos en la última evaluación.
 - d) Estatura - Peso.
 - e) Nota en matemáticas en el último examen - Número de asignaturas suspensas en la última evaluación.
 - f) Peso - Nota en matemáticas.
 - g) Estatura media de los padres - Estatura del alumno.
 - h) Distancia de su casa al centro de estudios - Tiempo medio que tarda en llegar.
 - i) Número de libros leídos al año - Número de asignaturas suspensas en la última evaluación.
- a) Positiva y fuerte.
- b) Habrá una correlación negativa y muy fuerte.
- c) Habrá una correlación negativa (a más horas semanales de estudio, menos número de suspensos) fuerte.
- d) Correlación positiva débil.
- e) Negativa y débil.
- f) No habrá correlación.
- g) Correlación positiva y fuerte.
- h) Correlación positiva y muy fuerte.
- i) Correlación negativa débil.

2 El valor de la correlación

Página 188

1. Los siguientes números son los valores absolutos de los coeficientes de correlación, r , de las distribuciones bidimensionales representadas a continuación:

0,75 0,47 0,92 0,97

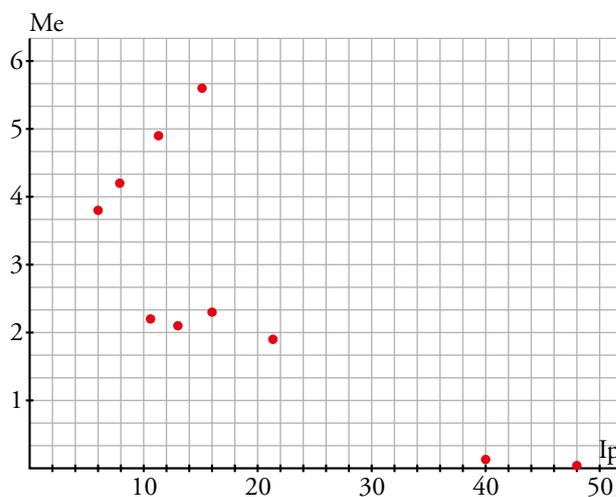


Asigna cada cual a la suya, cambiando el signo cuando convenga.

- Ⓐ → -0,47
- Ⓑ → 0,97
- Ⓒ → -0,92
- Ⓓ → 0,75

Página 189

- 2.** Representa la nube de puntos y la recta de regresión de la distribución bidimensional IP - Me del ejercicio resuelto anterior.



- 3.** Indica cuál de estos valores se ajusta mejor al valor de la correlación de la distribución del ejercicio 2.

0,5 -0,99 0,82 -0,77 0,99

$r = -0,77$

3 La recta de regresión para hacer estimaciones

Página 190

- 1. Estima, con los datos del ejemplo 1, el alargamiento correspondiente a una temperatura de 45 °C. ¿Consideras fiable la estimación?**

$$y = 0,12x \rightarrow \hat{y}(45) = 5,4 \text{ mm}$$

La estimación es muy fiable.

- 2. Estima, con los datos del ejemplo 2, el peso de un nuevo jugador cuya estatura sea de 180 cm. ¿Consideras fiable la estimación?**

Hallamos gráficamente el peso que corresponde a 180 cm: $\hat{y}(180) = 77 \text{ kg}$.

La estimación no será muy fiable puesto que, aunque la correlación es relativamente alta, 180 no está en el intervalo de datos considerados.

Página 191

- 3. Estima, mediante la recta de regresión, la presión correspondiente a 1 000 m. ¿Es fiable la estimación?**

Efectuamos la estimación con la ecuación de la recta de regresión:

$$\hat{y}(1\,000) = 760 - 0,0824 \cdot 1\,000 = 677,6 \text{ mm}$$

Es muy fiable la estimación, ya que la correlación es muy buena y 1 000 está dentro del intervalo de valores considerados.

- 4. Estima la presión correspondiente a una altura de 6 000 m. Comenta cómo de fiable es esa estimación.**

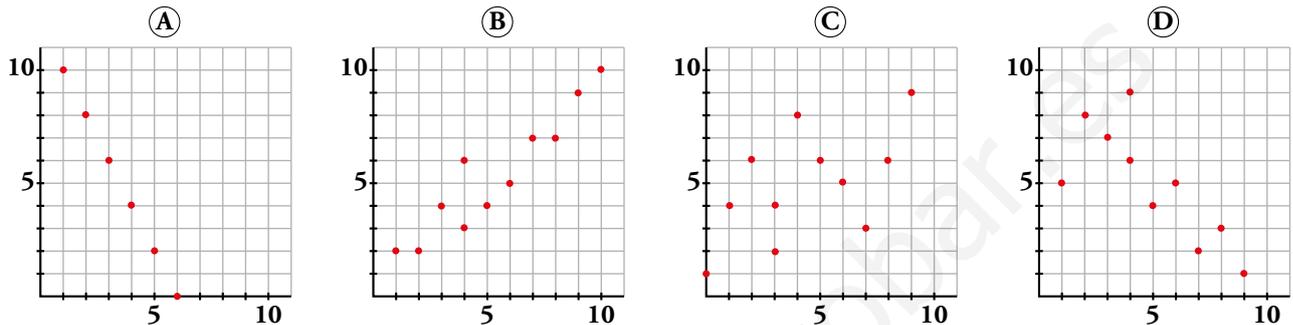
En este caso $\hat{y}(6\,000) = 760 - 0,0824 \cdot 6\,000 = 265,6 \text{ mm}$ y la estimación no es muy fiable porque aunque la correlación es muy buena, 6 000 está fuera del intervalo de datos disponibles.

Ejercicios y problemas

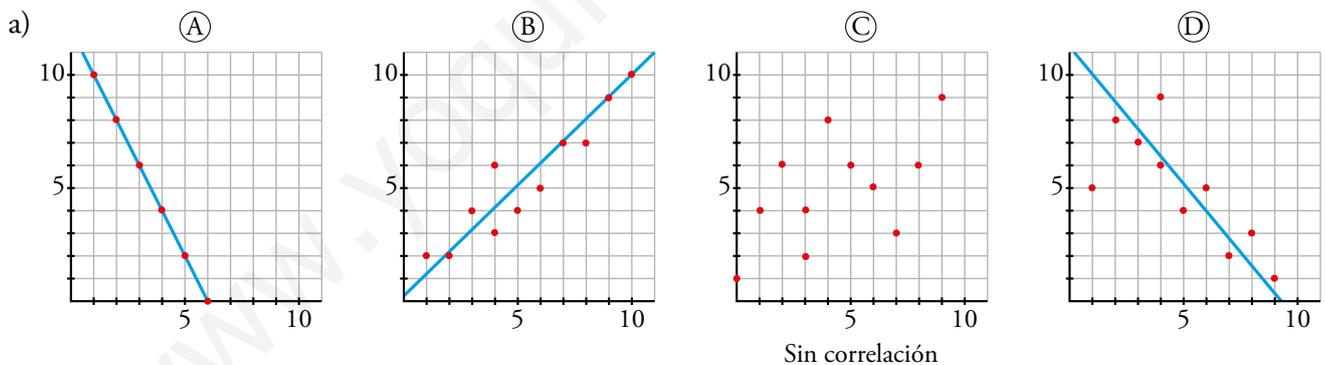
Página 192

Practica

1. a) Traza, a ojo, la recta de regresión en cada una de estas cuatro distribuciones bidimensionales:



- b) ¿Cuáles de ellas tienen correlación positiva y cuáles tienen correlación negativa?
 c) Una de ellas presenta relación funcional. ¿Cuál es? ¿Cuál es la expresión analítica de la función que relaciona las dos variables?
 d) Ordena de menor a mayor las correlaciones de las cuatro (en valor absoluto): en primer lugar, la que presenta correlación más débil, y, en último lugar, aquella cuya correlación es más fuerte.

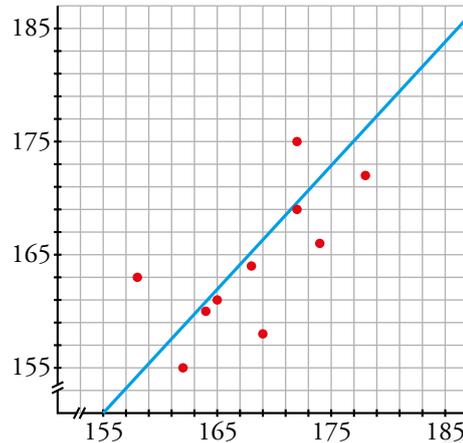


- b) (A) y (D) tienen correlación negativa y (B) correlación positiva. En el caso de (C) no se aprecia correlación.
 c) La (A) presenta una relación funcional:
 (6, 0) y (5, 2) pertenecen a la recta $\rightarrow m = \frac{2-0}{5-6} = -2 \rightarrow y = -2(x-6)$
 d) (C) < (D) < (B) < (A)

2. Las estaturas de 10 chicas (x_i) y las de sus respectivas madres (y_i) son:

x_i	158	162	164	165	168	169	172	172	174	178
y_i	163	155	160	161	164	158	175	169	166	172

Representa los valores sobre papel cuadrulado mediante una nube de puntos, traza a ojo la recta de regresión y di si la correlación es positiva o negativa y más o menos fuerte de lo que esperabas.



Se trata de una correlación positiva y fuerte.

3. Estos son los resultados que hemos obtenido al tallar y pesar a varias personas:

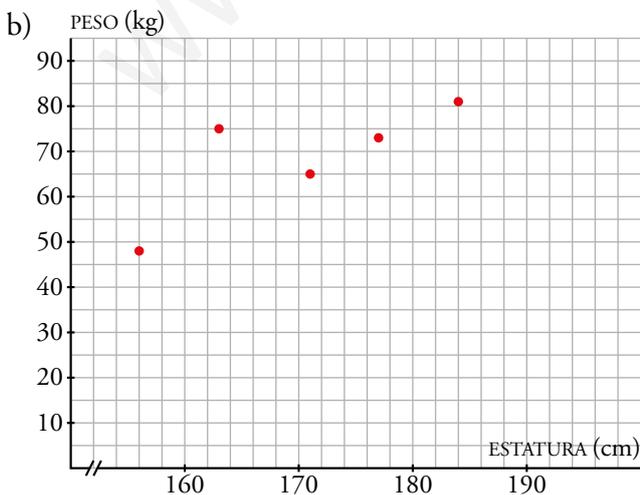
ESTATURA (cm)	156	163	171	177	184
PESO (kg)	48	75	65	73	81

a) ¿Es una distribución bidimensional? ¿Cuáles son las variables que se relacionan? ¿Cuáles son los individuos?

b) Representa la nube de puntos.

c) ¿Es una relación estadística o funcional?

a) Sí es una distribución bidimensional ya que a cada individuo (personas que pesamos y tallamos) tiene dos valores asociados correspondientes a las dos variables que se relacionan: estatura (cm) y peso (kg).

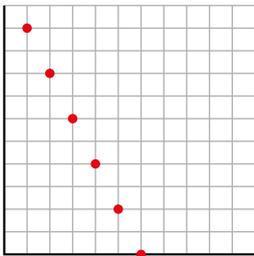


c) Es una relación estadística.

4. Representa el diagrama de dispersión correspondiente a la siguiente distribución y di cuál de estos tres valores puede ser su coeficiente de correlación:

$r = 1$ $r = -0,98$ $r = -1$

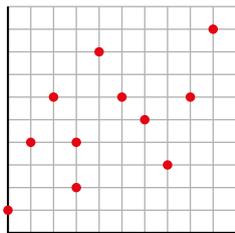
x	1	2	3	4	5	6
y	10	8	6	4	2	0



$r = -1$, la dependencia es funcional.

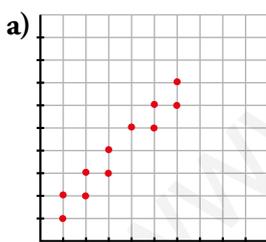
5. Representa la nube de puntos de la siguiente distribución y estima cuál de estos tres puede ser su coeficiente de correlación: $r = 0,98$; $r = -0,51$; $r = 0,57$.

x	0	1	2	3	3	4	5	6	7	8	9
y	1	4	6	2	4	8	6	5	3	6	9

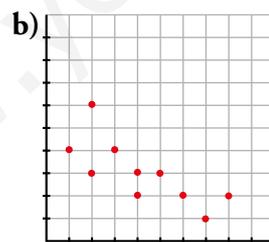


$r = 0,57$

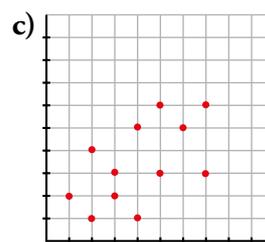
6. Los coeficientes de correlación de estas distribuciones bidimensionales son, en valor absoluto: 0,55; 0,75; 0,87 y 0,96. Asigna a cada una el suyo, cambiando el signo cuando proceda:



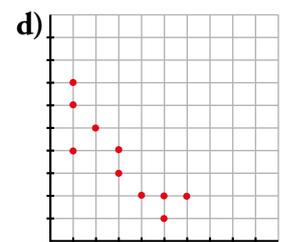
a) $r = 0,96$



b) $r = -0,75$

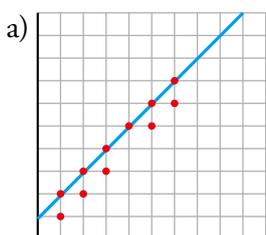


c) $r = 0,55$



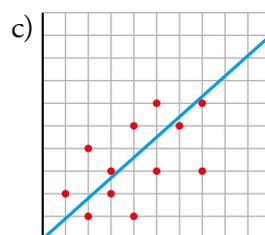
d) $r = -0,87$

7. Traza la recta de regresión de las distribuciones a) y c) del ejercicio anterior y estima, en cada una de ellas, los valores que corresponden a $x = 0$ y a $x = 10$. ¿En cuál son más fiables las estimaciones?



$x = 0 \rightarrow 0,9$

$x = 10 \rightarrow 11$



$x = 0 \rightarrow 0$

$x = 10 \rightarrow 9$

Serán más fiables las estimaciones de la distribución a).

Resuelve problemas

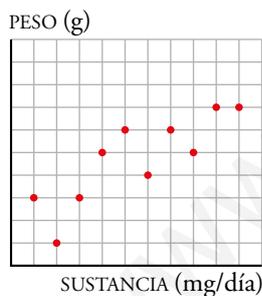
8.  Se ha hecho un estudio con ratones para ver los aumentos de peso (en g) mensuales que producen ciertas sustancias A, B y C (en mg diarios). Los datos obtenidos vienen dados en esta tabla:

SUSTANCIA	AUMENTO DE PESO SI LA SUSTANCIA ES A	AUMENTO DE PESO SI LA SUSTANCIA ES B	AUMENTO DE PESO SI LA SUSTANCIA ES C
1	3	2	3
2	1	2	3
3	3	1	2
4	5	3	0
5	6	0	1
6	4	3	-1
7	6	4	1
8	5	1	-2
9	7	3	-4
10	7	1	-2

Los resultados negativos quieren decir que en lugar de aumentar, el peso disminuye.

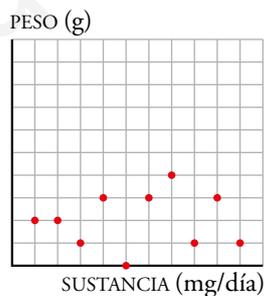
- Representa la nube de puntos de cada distribución.
- Indica si la correlación es positiva o negativa en cada una de ellas.
- Ordena las correlaciones de menos a más fuerte.

a) y b) Sustancia A



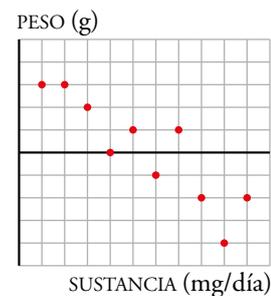
Correlación positiva.

Sustancia B



Correlación positiva.

Sustancia C



Correlación negativa.

c) Sustancia B < Sustancia A < Sustancia C.

9.  La correlación entre las temperaturas medias mensuales de una ciudad española y el tiempo que sus habitantes dedican a ver la televisión, es de $-0,89$. ¿Te parece razonable este valor? Explica su significado.

¿Será positiva o negativa la correlación entre la lluvia caída mensualmente y el consumo televisivo de sus habitantes?

Parece razonable que si las temperaturas aumentan, disminuye el tiempo que los habitantes dedican a ver la televisión, ya que parece lógico pensar que la gente pasa más tiempo en la calle.

La correlación entre la lluvia caída mensualmente y el consumo televisivo de sus habitantes será positiva porque si llueve, es lógico pensar que la gente pasará más tiempo en casa.

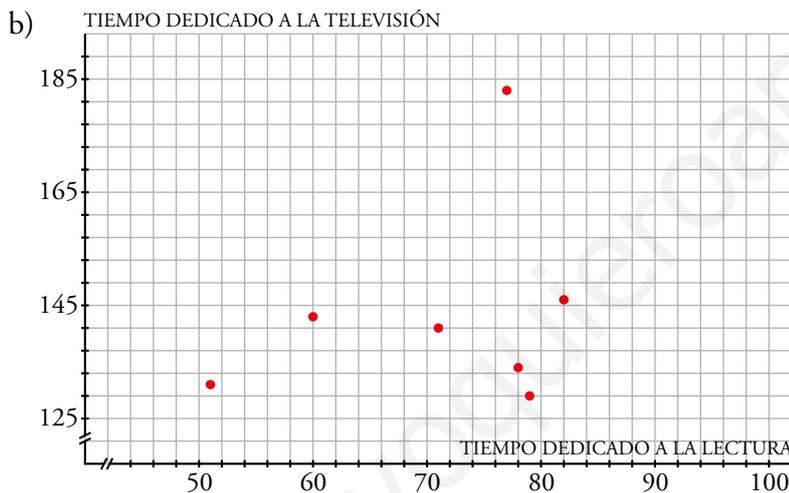
10. En una encuesta realizada en 7 países europeos, se obtuvieron los datos de este gráfico:



a) ¿Cómo crees que será la correlación entre los tiempos dedicados a la lectura y a la televisión?

b) Haz la nube de puntos correspondiente a estas dos variables y contrasta lo que observas en ella con tu respuesta del apartado a).

a) Esta respuesta depende de los alumnos y alumnas, que después comprobarán en el apartado b) lo acertado de su intuición.

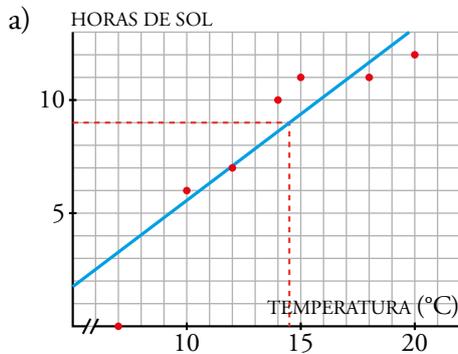


Se trata de una correlación positiva muy débil.

11.  Para realizar unos estudios sobre energía solar, se ha medido cada uno de los días de una semana la temperatura máxima y el número de horas de sol, obteniéndose los siguientes resultados:

S: N.º DE HORAS DE SOL	T: TEMPERATURA (°C)
7	12
10	14
0	7
6	10
11	15
12	20
11	18

- a) Traza a ojo la recta de regresión T-S.
 b) Si el lunes siguiente a la medición hubo 9 horas de sol, ¿qué temperatura máxima cabe esperar que hiciera? ¿Qué fiabilidad tiene tu predicción?



- b) Cabe esperar que hiciera una temperatura máxima de 14,5 °C.
 La fiabilidad es muy precisa porque la correlación es fuerte y la medición que nos piden está próxima a los valores que conocemos.

Curiosidades matemáticas

Encuentra explicaciones razonables a estos hechos

1. Suele decirse que la mayoría de los accidentes de automóvil se producen cerca de la casa del conductor. ¿Es más peligroso circular por nuestro barrio que a muchos kilómetros de nuestra residencia?
 2. Es fácil demostrar que los niños con pies grandes leen mejor que los que tienen pies pequeños. ¿Influye el tamaño del pie en la capacidad para la lectura?
 3. Se ha constatado que, en los pueblos de una cierta comarca, cuantos más nidos de cigüeña hay en sus tejados, más nacimientos de niños se producen. ¿Tienen que ver, pues, las cigüeñas con los nacimientos?
1. Los conductores se sienten más seguros en un entorno cercano y eso puede causar que bajen la guardia y tengan accidentes.
 2. Los niños de los cursos superiores leen mejor que los que empiezan el colegio. Estos niños son mayores y tienen los pies más grandes.
 3. Si hay más nidos, lo más probable es que haya más tejados y, por tanto, que el pueblo de la comarca sea más grande. Por tanto, habrá más nacimientos.