13 DISTRIBUCIONS BIDIMENSIONAIS

Páxina 331

REFLEXIONA E RESOLVE

Relación funcional e relación estatística

En cada un dos seguintes casos debes dicir se, entre as dúas variables que se citan, hai relación funcional ou relación estatística (correlación) e, neste último caso, indicar se é positiva ou negativa:

- **Nun conxunto de familias: *estatura media dos pais – estatura media dos fillos.***

  Correlación positiva.

- ***Temperatura á que quentamos unha barra de ferro – lonxitude alcanzada.***

  Funcional.

- **Entre os países do mundo respecto a España: *volume de exportación – volume de importación.***

  Correlación negativa.

- **Entre os países do mundo: *índice de mortalidade infantil – número de médicos por cada 1 000 habitantes.***

  Correlación negativa.

- **Nas vivendas dunha cidade: *kWh consumidos durante xaneiro – custo do recibo da luz.***

  Funcional.

- ***Número de persoas que viven en cada casa – custo do recibo da luz.***

  Correlación positiva.

- **Equipos de fútbol: *lugar que ocupan ao finalizar a liga – número de partidos perdidos.***

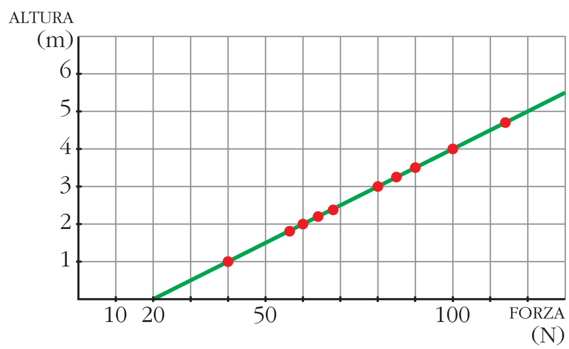
  Correlación positiva.

- **Equipos de fútbol: *lugar que ocupan ao finalizar a liga – número de partidos gañados.***

  Correlación negativa.

Exemplo de relación funcional

Distintas persoas lanzan cara arriba unha mesma pedra de 2 kg de masa, que alcanza máis ou menos altura segundo a forza con que foi impulsada. (A forza actúa nun tramo de 1 m).



a) Que altura, por riba da man, consideras que alcanzará a pedra se se impulsa cunha forza de 110 newton?

b) Poderiamos escribir unha fórmula que dea directamente a altura que alcanza a pedra, desde o momento en que se solta, en función da forza con que é impulsada cara arriba?

a) 4,5 m

b) Altura $= \frac{F}{20} - 1$ para $F \geq 20$

Obtención física de la fórmula:

La fórmula en la que se basa todo el desarrollo posterior es:

$$v = \sqrt{2ad}$$

donde v: Aumento de la velocidad en el tramo d.

a: Aceleración constante con la que se mueve el móvil.

d: Espacio que recorre con la aceleración a.

Así, la velocidad con que sale de la mano es:

$$v_s = \sqrt{2a\,1} = \sqrt{2a}$$

Además:

$$F = m(a + g) \rightarrow a = \frac{F}{m} - g = \frac{F}{2} - 10$$

Luego:

$$v_s = \sqrt{2\left(\frac{F}{2} - 10\right)} = \sqrt{F - 20}$$

Por otra parte, si se deja caer una piedra desde una altura h, adquiere una velocidad:

$$v_s = \sqrt{2gh}$$

O bien, si se empuja una piedra hacia arriba de modo que salga con una velocidad v_s, alcanza una altura h.

En este caso:

$$v_s = \sqrt{2 \cdot 10 \cdot h} = \sqrt{20h}$$

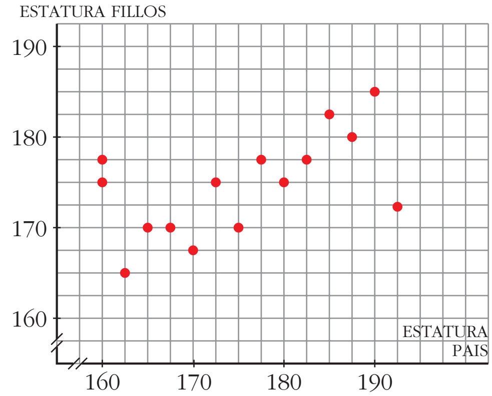
Igualando:

$$\sqrt{F - 20} = \sqrt{20h} \;\rightarrow\; h = \frac{F}{20} - 1$$

Para que $h \geq 0$, debe ser $F \geq 20$.

Exemplo de relación estatística

Na seguinte gráfica, cada punto corresponde a un rapaz. A abscisa é a estatura do pai, e a ordenada, a súa propia altura.



a) Identifica a Guille e Gabriel, irmáns de boa estatura, cuxo pai é baixiño.

b) Identifica a Serxio, de estatura normaliña, cuxo pai é un xigantón.

c) Podemos dicir que hai unha certa relación entre as estaturas destes 15 rapaces e as dos pais?

a) Guille y Gabriel están representados por los puntos (160, 175) y (160; 177,5)

b) Sergio está representado por el punto (192,5; 172,5).

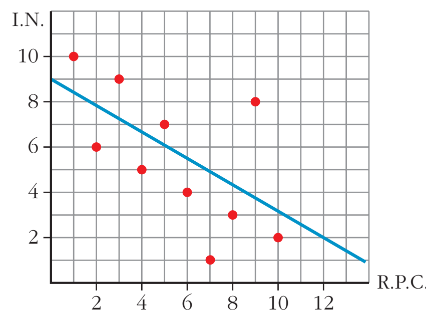
c) En general, sí.

Páxina 333

1. A táboa da dereita mostra como se ordenan entre si dez países, A, B, C…, segundo dúas variables, R.P.C. (*renda per cápita*) e I.N. (*índice de natalidade*). Representa os resultados nunha nube de puntos, traza a recta de regresión e di como che parece a correlación.

PAÍSES	A	B	C	D	E	F	G	H	I	J
R.P.C.	1	2	3	4	5	6	7	8	9	10
I.N.	10	6	9	5	7	4	1	3	8	2

La correlación es negativa y moderadamente alta (–0,62).



Páxina 335

1. Obtén mediante cálculos manuais os coeficientes de correlación das distribucións que está na páxina 332:

Matemáticas – Filosofía ***Distancia – Número de encestes***

Faino tamén cunha calculadora con MODO LR.

Matemáticas-Filosofía:

x_i	y_i	x_i^2	y_i^2	x_iy_i
2	2	4	4	4
3	5	9	25	15
4	2	16	4	8
4	7	16	49	28
5	5	25	25	25
6	4	36	16	24
6	6	36	36	36
7	6	49	36	42
7	7	49	49	49
8	5	64	25	40
10	5	100	25	50
10	9	100	81	90
72	63	504	375	411

$$\bar{x} = \frac{72}{12} = 6$$

$$\bar{y} = \frac{63}{12} = 5,25$$

$$\sigma_x = \sqrt{\frac{504}{12} - 6^2} = 2,45$$

$$\sigma_y = \sqrt{\frac{375}{12} - 5,25^2} = 1,92$$

$$\sigma_{xy} = \frac{411}{12} - 6 \cdot 5,25 = 2,75$$

Por tanto: $r = \dfrac{2,75}{2,45 \cdot 1,92} = 0,58$

Distancia-Número de encestes:

x_i	y_i	x_i^2	y_i^2	$x_i y_i$
1	9	1	81	9
2	10	4	100	20
3	6	9	36	18
4	4	16	16	16
5	2	25	4	10
6	0	36	0	0
7	1	49	1	7
8	0	64	0	0
36	32	204	238	80

$\bar{x} = \dfrac{36}{8} = 4,5 \qquad \bar{y} = \dfrac{32}{8} = 4$

$$\sigma_x = \sqrt{\frac{204}{8} - 4,5^2} = 2,29$$

$$\sigma_y = \sqrt{\frac{238}{8} - 4^2} = 3,71$$

$$\sigma_{xy} = \frac{80}{8} - 4,5 \cdot 4 = -8$$

Por tanto: $r = \dfrac{-8}{2,29 \cdot 3,71} = -0,94$

Páxina 344

EXERCICIOS E PROBLEMAS PROPOSTOS

PARA PRACTICAR

Sen fórmulas

1 Para cada un dos seguintes casos indica:

- **Cales son as variables que se relacionan.**
- **Se se trata dunha relación funcional ou dunha relación estatística e, nestes casos, o signo da correlación.**

a) Renda mensual dunha familia-gasto en electricidade.

b) Raio dunha esfera-volume desta.

c) Litros de chuvia recollidos nunha cidade-tempo dedicado a ver a televisión polos seus habitantes.

d) Lonxitude do traxecto percorrido nunha liña de proximidades-prezo do billete.

e) Peso dos alumnos de 1.° de Bacharelato-número de calzado que usan.

f) Toneladas de tomate recollidas nunha colleita-prezo do quilo de tomate no mercado.

a) Renta (€), gasto (€).

Correlación positiva.

b) Relación funcional.

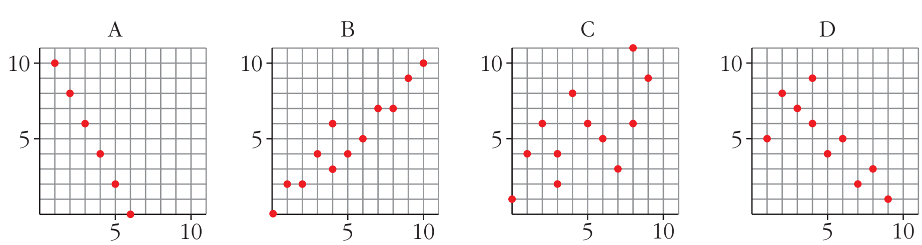
c) Relación estadística. Seguramente muy débil. Positiva (¿cabe pensar que cuanto más llueva más tiempo pasarán en casa y, por tanto, más verán la televisión?).

d) Aunque lo parezca a priori, seguramente la relación no es funcional. Es una correlación positiva fuerte.

e) Correlación positiva.

f) Correlación negativa (cuanto mayor sea la cosecha, más baratos están los tomates).

2 a) Traza, a ollo, a recta de regresión en cada unha destas distribucións bidimensionais:

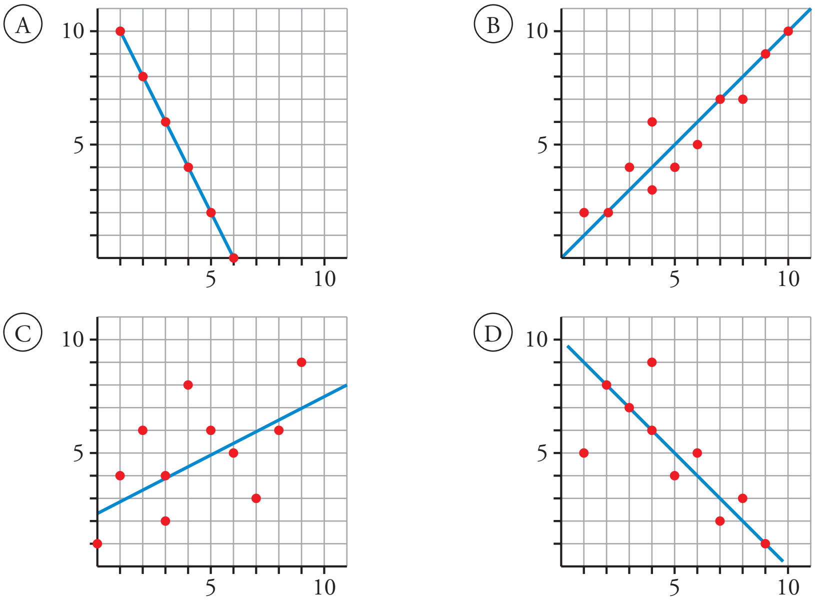


b) Cales delas teñen correlación positiva e cales teñen correlación negativa?

c) Unha delas presenta relación funcional. Cal é? Cal é a expresión analítica da función que relaciona as dúas variables?

d) Ordena de menor a maior as correlacións.

a)



b) B y C tienen correlación positiva; A y D, negativa.

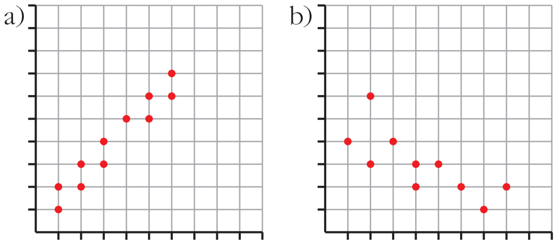
c) La A es relación funcional: $y = 12 - 2x$.

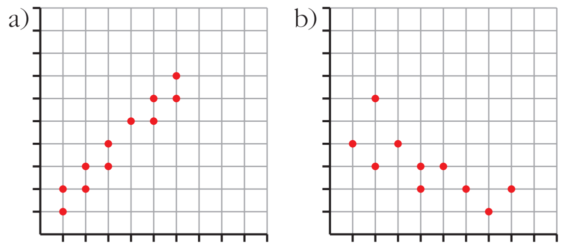
d) C, D, B, A (prescindiendo del signo).

3 **Os coeficientes de correlación das distribucións bidimensionais que aparecen a continuación son, en valor absoluto, os seguintes:**

0,55 0,75 0,87 0,96

Asígnalle a cada un o seu, cambiando o signo cando proceda:

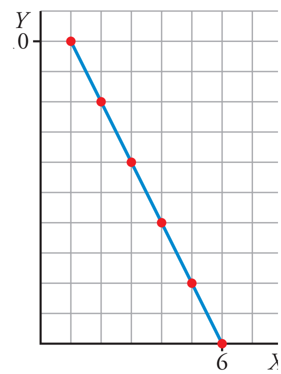




a) r = 0,96 b) r = –0,75 c) r = 0,55 d) r = –0,87

4 **Representa a nube de puntos correspondente a esta distribución e di canto vale o coeficiente de correlación.**

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



El coeficiente de correlación vale –1.

5 **Representa a nube de puntos desta distribución e estima cal destes tres pode ser o coeficiente de correlación:**

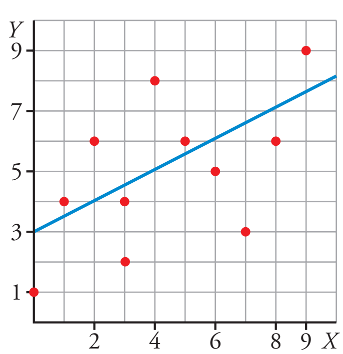
a) r **= 0,98**

b) r **= –0,87**

c) r **= 0,5**

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

c) r = 0,5



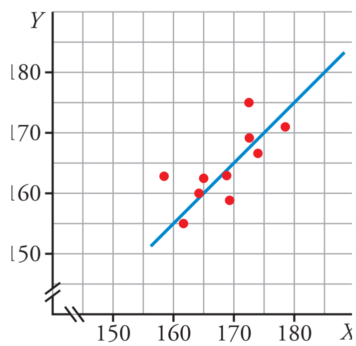
6 **As estaturas de 10 rapazas e as das súas respectivas nais son:**

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 os valores, sobre papel cuadriculado, mediante unha nube de puntos.

Traza a ollo a recta de regresión e di se a correlación é positiva ou negativa e se é máis ou menos forte do que esperabas.

La correlación es positiva y fuerte.



Páxina 345

Con fórmulas

7 **Esta é a distribución bidimensional dada no exercicio 2B) mediante unha nube de puntos:**

x	0	1	2	3	4	4	5	6	7	8	9	10
y	0	2	2	4	3	6	4	5	7	7	9	10

Determina:

a) $\bar{x}$, $\bar{y}$, σ_x, σ_y, σ_{xy}.

b) O coeficiente de correlación, *r*. Interprétao.

c) As dúas rectas de regresión.

$n = 12$, $\Sigma x = 59$ $\Sigma y = 59$

$\Sigma x^2 = 401$ $\Sigma y^2 = 389$ $\Sigma xy = 390$

a) $\bar{x} = 4{,}92$ $\bar{y} = 4{,}92$

$\sigma_x = 3{,}04$ $\sigma_y = 2{,}87$ $\sigma_{xy} = 8{,}33$

b) $r = \dfrac{\sigma_{xy}}{\sigma_x \sigma_y} = 0{,}95$. Se trata de una correlación fuerte y positiva.

c) Recta de regresión de Y sobre X:

$$\frac{\sigma_{xy}}{\sigma_x^2} = 0{,}90 \;\rightarrow\; y = 4{,}92 + 0{,}9(x - 4{,}92)$$

Recta de regresión de X sobre Y:

$$\frac{\sigma_{xy}}{\sigma_y^2} = 1{,}01 \;\rightarrow\; y = 4{,}92 + \frac{1}{1{,}01}(x - 4{,}92) \;\rightarrow\; y = 4{,}92 + 0{,}99(x - 4{,}92)$$

8 **Observa a distribución D do exercicio 2.**

a) Descríbea mediante unha táboa de valores.

b) Realiza os cálculos para obter o seu coeficiente de correlación.

c) Representa os puntos no caderno. Calcula a ecuación da recta de regresión de *Y* sobre *X* e represéntaa.

a)

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

b) $n = 10$ $\quad \Sigma x = 49 \quad \bar{x} = \dfrac{49}{10} = 4{,}9$

$\Sigma y = 50 \quad \bar{y} = \dfrac{50}{10} = 5$

$\Sigma x^2 = 301 \quad \sigma_x = \sqrt{\dfrac{301}{10} - 4{,}9^2} = 2{,}47$

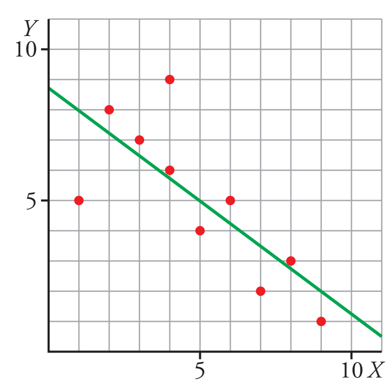
$\Sigma y^2 = 310 \quad \sigma_y = \sqrt{\dfrac{301}{10} - 5^2} = 2{,}45$

$\Sigma xy = 199 \quad \sigma_{xy} = \dfrac{199}{10} - 4{,}9 \cdot 5 = -4{,}6$

$$r = \frac{4{,}6}{2{,}47 \cdot 2{,}45} = -0{,}76$$

c) Recta de regresión de Y sobre X:

$$y = 5 - \frac{4{,}6}{6{,}1}(x - 4{,}9) \;\rightarrow\; y = 8{,}675 - 0{,}75x$$



9 **a) Representa a seguinte distribución bidimensional:**

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

b) Comproba coa calculadora que os seus parámetros son:

$\bar{x} = 4,4$ $\quad \bar{y} = 4,9 \sigma_{xy} = 3,67$

$\sigma_x = 2,77$ $\quad \sigma_y = 2,31$ $\quad r = 0,57$

c) Indica as ecuacións das dúas rectas de regresión, *X* sobre *Y* e *Y* sobre *X*, e represéntaas xunto coa nube de puntos.

a) Representada en el ejercicio 5.

b) Se comprueba.

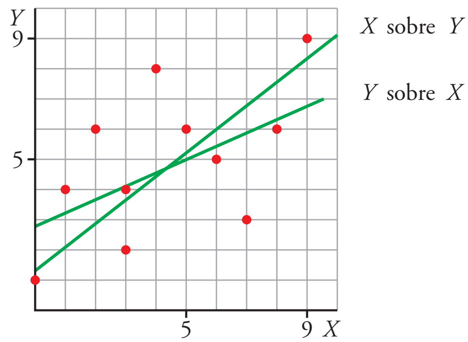
c) • Recta de regresión de Y sobre X:

$$m_{yx} = \frac{\sigma_{xy}}{\sigma_x^2} = \frac{3,67}{2,77^2} = 0,48 \rightarrow y = 4,9 + 0,48(x - 4,4) \rightarrow y = 0,48x + 2,79$$

• Recta de regresión de X sobre Y:

$$m_{xy} = \frac{\sigma_{xy}}{\sigma_y^2} = \frac{3,67}{2,31^2} = 0,69 \rightarrow \frac{1}{m_{xy}} = 1,45 \rightarrow y = 4,9 + 1,45(x - 4,4) \rightarrow$$

$$\rightarrow y = 1,45x - 1,48$$



10 **Unha distribución bidimensional na que os valores de *x* son 12, 15, 17, 21, 22 e 25, ten unha correlación $r = 0,99$ e a súa recta de regresión é $y = 10,5 + 3,2x$.**

Calcula $\hat{y}(13)$, $\hat{y}(20)$, $\hat{y}(30)$, $\hat{y}(100)$.

Cales das estimacións anteriores son fiables, cal pouco fiable e cal non se debe facer?

Expresa os resultados en termos adecuados. (Por exemplo: $\hat{y}(13) = 52,1$. Para $x = 13$ é moi probable que o valor correspondente de y sexa próximo a 52).

$$\hat{y}(13) = 52,1; \quad \hat{y}(20) = 74,5; \quad \hat{y}(30) = 106,5; \quad \hat{y}(100) = 330,5$$

Son fiables $\hat{y}(13)$ e $\hat{y}(20)$, porque 13 y 20 están en el intervalo de valores utilizados para obtener la recta de regresión.

$\hat{y}(30)$ es menos fiable, pues 30 está fuera del intervalo, aunque cerca de él.

$\hat{y}(100)$ es una estimación nada fiable, pues 100 está muy lejos del intervalo [12, 25].

PARA RESOLVER

11 **A seguinte táboa mostra o número de xermes patóxenos por centímetro cúbico dun determinado cultivo segundo o tempo transcorrido:**

N.° DE HORAS	0	1	2	3	4	5
N.° DE XERMES	20	26	33	41	47	53

a) Calcula a recta de regresión para predicir o número de xermes por centímetro cúbico en función do tempo.

b) Que cantidade de xermes por centímetro cúbico cabe esperar que haxa ás 6 horas? É boa esta estimación?

a) $y = 19,81 + 6,74x$, donde: $x \to$ número horas, $y \to$ número de gérmenes

b) $\hat{y}(6) = 60,25 \approx 60$ gérmenes.

Es una buena predicción, puesto que $r = 0,999$ (y 6 está cercano al intervalo de valores considerado).

12 **A media dos pesos dos individuos dunha poboación é de 65 kg, e a das súas estaturas, 170 cm. As súas desviacións típicas son 5 kg e 10 cm. A covarianza é 40 kg · cm. Determina:**

a) Coeficiente de correlación.

b) A recta de regresión dos pesos respecto das estaturas.

c) Estima o peso dun individuo de 180 cm de estatura pertencente a ese colectivo.

a) $r = 0,8$

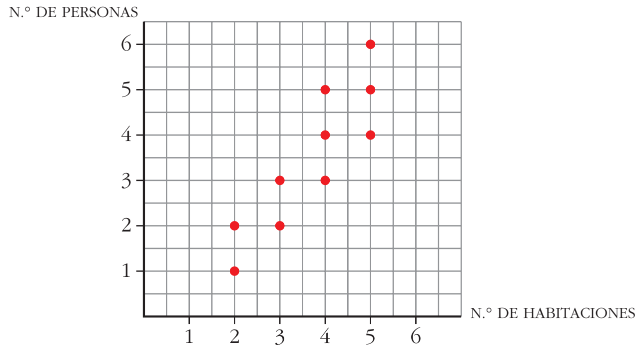
b) $y = 65 + 0,4(x - 170) = 0,4x - 3 \to \begin{cases} x: \text{ estaturas en cm} \\ y: \text{ pesos en kg} \end{cases}$

c) $\hat{y}(180) = 69$ kg

13 **Nunha zona residencial tomouse unha mostra para relacionar o *número de habitacións que ten cada piso* (*h*) co *número de persoas que viven nel* (*p*). Estes son os resultados que se obtiveron:**

h	2	2	3	3	4	4	4	5	5	5
p	1	2	2	3	3	4	5	4	5	6

Represéntaos mediante unha nube de puntos. Calcula o coeficiente de correlación e interprétao.



h: número de habitaciones

p: número de personas

$n = 10 \qquad \Sigma h = 37 \qquad \bar{h} = \dfrac{37}{10} = 3,7$

$\Sigma p = 35 \qquad \bar{p} = \dfrac{35}{10} = 3,5$

$\Sigma h^2 = 149 \qquad \sigma_h = \sqrt{\dfrac{149}{10} - 3,7^2} = 1,1$

$\Sigma p^2 = 145 \qquad \sigma_p = \sqrt{\dfrac{145}{10} - 3,5^2} = 1,5$

$\Sigma hp = 144 \qquad \sigma_{hp} = \dfrac{144}{10} - 3,7 \cdot 3,5 = 1,45$

$$r = \frac{1,45}{1,1 \cdot 1,5} = 0,88$$

Es una correlación positiva y fuerte (a más habitaciones, más personas en el piso).

14 **A táboa adxunta relaciona o número atómico de varios metais coa súa densidade:**

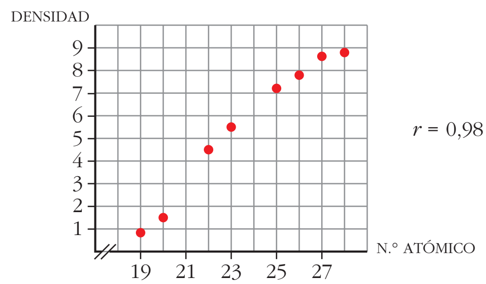
Elemento	K	Ca	Ti	V	Mn	Fe	Co	Ni
N.° atómico	19	20	22	23	25	26	27	28
Densidade	0,86	1,54	4,50	5,60	7,11	7,88	8,70	8,80

a) Representa os puntos e calcula o coeficiente de correlación.

b) Mediante unha recta de regresión, estima a densidade do cromo se o seu número atómico é 24: Cr (24).

c) Estima a densidade do escandio: Sc (21).

a)



$r = 0{,}98$

b) y c) $\hat{y} = -16{,}5 + 0{,}93x$ $\quad\quad$ $\hat{y}(24) = 5{,}86$ $\quad\quad$ $\hat{y}(21) = 3{,}06$

Las densidades del Cr y del Sc son, aproximadamente, 5,86 y 3,01. (Los valores reales de estas densidades son 7,1 y 2,9.)

Páxina 346

15 **Nunha confraría de pescadores, as capturas rexistradas de certa variedade de peixes, en quilogramos, e o prezo de poxa en lonxa, en euros/kg, foron os seguintes:**

x (kg)	**2 000**	**2 400**	**2 500**	**3 000**	**2 900**	**2 800**	**3 160**
y (euros/kg)	**1,80**	**1,68**	**1,65**	**1,32**	**1,44**	**1,50**	**1,20**

a) Cal é o prezo medio rexistrado?

b) Indica o coeficiente de correlación linear e interprétao.

c) Estima o prezo que alcanzaría en lonxa o quilo desa especie se se pescasen 2 600 kg.

a) $\bar{y} = 1{,}51$ euros

b) $r = -0{,}97$. La relación entre las variables es fuerte y negativa. A mayor cantidad de pescado, menor es el precio por kilo.

c) La recta de regresión es $y = 2{,}89 - 0{,}0005x$.

$\hat{y}(2\,600) = 1{,}59$ euros.

16 **Durante 10 días, realizamos medicións sobre o consumo dun coche (litros consumidos e quilómetros percorridos). Os datos obtidos foron os seguintes:**

x (km)	100	80	50	100	10	100	70	120	150	220
y (l)	6,5	6	3	6	1	7	5,5	7,5	10	15

a) Indica o coeficiente de correlación e a recta de regresión de Y sobre X.

b) Se queremos facer unha viaxe de 190 km, que cantidade de combustible debemos poñer?

a) $r = 0{,}99$; $y = 0{,}157 + 0{,}066x$

b) $\hat{y}(190) = 12{,}697$ litros. Debemos poner, como mínimo, unos 13 litros.

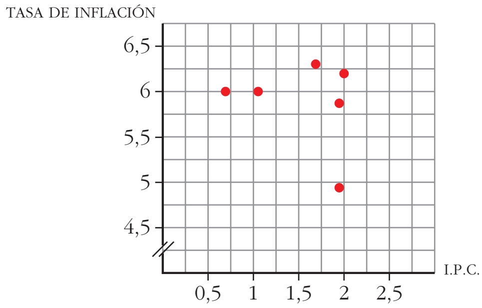
17 **A evolución do IPC (índice de prezos ao consumo) e da taxa de inflación en 1987 foi:**

	XANEIRO	FEBREIRO	MARZO	ABRIL	MAIO	XUÑO
IPC	0,7	1,1	1,7	2	1,9	1,9
TAXA DE INFLACIÓN	6	6	6,3	6,2	5,8	4,9

a) Representa a nube de puntos.

b) Calcula o coeficiente de correlación entre o IPC e a taxa de inflación.

c) Pódese estimar a taxa de inflación a partir do IPC?



$r = -0{,}24$. La nube de puntos es muy dispersa. No se puede estimar de forma fiable la tasa de inflación a partir del IPC (pues $|r|$ es muy bajo).

CUESTIÓNS TEÓRICAS

18 **O coeficiente de correlación dunha distribución bidimensional é 0,87.**

Se os valores das variables se multiplican por 10, cal será o coeficiente de correlación desta nova distribución?

El mismo, puesto que r no depende de las unidades; es adimensional.

19 **Calculamos a covarianza dunha certa distribución e resultou negativa.**

Xustifica por que podemos afirmar que tanto o coeficiente de correlación como as pendentes das dúas rectas de regresión son números negativos.

Hay que tener en cuenta que $r = \dfrac{\sigma_{xy}}{\sigma_x \, \sigma_y}$; $m_{yx} = \dfrac{\sigma_{xy}}{\sigma_x^2}$; $m_{xy} = \dfrac{\sigma_{xy}}{\sigma_y^2}$ y que $\sigma_x \geq 0$, $\sigma_y \geq 0$ siempre.

Luego r, m_{yx}, m_{xy} tienen el mismo signo que σ_{xy}. (Además, suponemos $\sigma_x \neq 0$ y $\sigma_y \neq 0$).

20 **Que punto teñen en común as dúas rectas de regresión?**

El centro de gravedad de la distribución, $(\bar{x}, \ \bar{y})$.

21 **Que condición debe cumprir r para que as estimacións feitas coa recta de regresión sexan fiables?**

$|r|$ debe estar próximo a 1.

22 **Proba que o produto dos coeficientes de regresión m_{yx} e m_{xy} é igual ao cadrado do coeficiente de correlación.**

$$m_{yx} \cdot m_{xy} = \frac{\sigma_{xy}}{\sigma_x^2} \cdot \frac{\sigma_{xy}}{\sigma_y^2} = \left(\frac{\sigma_{xy}}{\sigma_x \, \sigma_y} \right)^2 = r^2$$

23 **Dunha distribución bidimensional (x, y) coñecemos os seguintes resultados:**

- **Recta de regresión de Y sobre X:**

$$y = 8{,}7 - 0{,}76x$$

- **Recta de regresión de X sobre Y:**

$$y = 11{,}36 - 1{,}3x$$

a) Calcula o centro de gravidade da distribución.

b) Indica o coeficiente de correlación.

a) El centro de gravedad, $(\bar{x}, \ \bar{y})$, es el punto de corte entre las dos rectas:

$$\left. \begin{array}{l} y = 8{,}7 - 0{,}76x \\ y = 11{,}36 - 1{,}3x \end{array} \right\}$$

$8{,}7 - 0{,}76x = 11{,}36 - 1{,}3x$

$0{,}54x = 2{,}66$

$x = 4{,}93$

$y = 4{,}95$

El centro de gravedad es $(\bar{x}, \ \bar{y}) = (4{,}93; \ 4{,}95)$.

b) Para hallar r tenemos en cuenta el ejercicio anterior:

$$r^2 = m_{yx} \cdot m_{xy} = -0{,}76 \cdot \frac{1}{-1{,}3} = 0{,}58 \ \rightarrow \ r = 0{,}76$$

24 **A estatura media de 100 escolares de certo curso de ESO é de 155 cm cunha desviación típica de 15,5 cm.**

A recta de regresión da estatura respecto ao peso é:

$$y = 80 + 1{,}5x \quad (x\text{: peso; } y\text{: estatura})$$

a) Cal é o peso medio deses escolares?

b) Cal é o signo do coeficiente de correlación entre peso e estatura?

a) La recta de regresión es:

$$y = \bar{y} + m(x - \bar{x}) = 155 + 1{,}5(x - \bar{x}) = 155 + 1{,}5x - 1{,}5\bar{x} = (155 - 1{,}5\bar{x}) + 1{,}5x =$$
$$= 80 + 1{,}5x \ \rightarrow \ 155 - 1{,}5\bar{x} = 80 \ \rightarrow \ \bar{x} = 50 \text{ kg}$$

b) Positivo (igual que el signo de la pendiente de la recta de regresión).

Páxina 347

PARA AFONDAR

25 **Nunha mostra de 64 familias estudouse o número de membros que están en idade laboral, x, e o número deles que están en activo, y. Os resultados son os da táboa. Calcula o coeficiente de correlación linear entre ambas as variables e interprétao.**

x \ y	1	2	3
1	6	0	0
2	10	2	0
3	12	5	1
4	16	8	4

$r = 0{,}31$. La relación entre las variables es débil.

26 **Unha compañía discográfica recompilou a seguinte información sobre o número de concertos dados, durante o verán, por 15 grupos musicais e as vendas de discos destes grupos (expresados en miles de CD):**

CD (x) \ CONCERTOS (y)	10 - 30	30 - 40	40 - 80
1 - 5	3	0	0
5 - 10	1	4	1
10 - 20	0	1	5

a) Calcula o número medio de CD vendidos.

b) Cal é o coeficiente de correlación?

c) Obtén a recta de regresión de *Y* sobre *X*.

d) Se un grupo musical vende 18 000 CD, que número de concertos se prevé que dea?

$x \rightarrow$ CD; $y \rightarrow$ Conciertos

a) $\bar{x} = 9{,}6 \approx 10$

b) $r = 0{,}814$

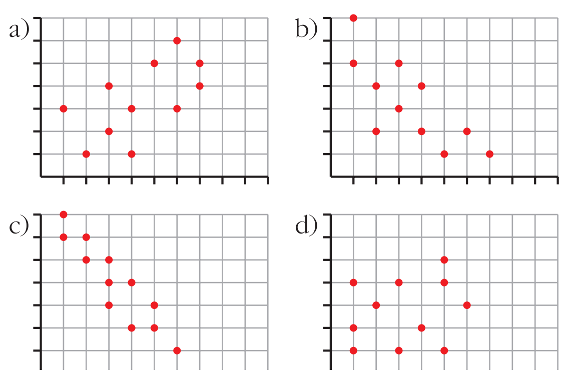
c) $y = 13{,}51 + 2{,}86x$

d) $\hat{y}(18) = 64{,}99 \approx 65$ conciertos

Páxina 347

AUTOAVALIACIÓN

1. Observa estas distribucións bidimensionais:



Asigna razoadamente un dos seguintes coeficientes de correlación a cada gráfica: 0,2; –0,9; –0,7; 0,6.

La correlación de a) es positiva, y las de b) y c), negativas. En d) no se aprecia correlación. La correlación de c) es más fuerte que la de b). Por tanto:

a) $\rightarrow$ 0,6

b) $\rightarrow$ –0,7

c) $\rightarrow$ –0,9

d) $\rightarrow$ 0,2

2. **Representa esta distribución bidimensional:**

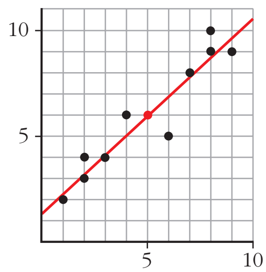
x	1	2	2	3	4	6	7	8	8	9
y	2	4	3	4	6	5	8	9	10	9

a) Calcula os parámetros $\bar{x}$, $\bar{y}$, σ_x, σ_y, σ_{xy}.

b) Indica o coeficiente de correlación.

c) Indica a recta de regresión de Y sobre X.

d) Estima o valor de y para $x = 5$ e para $x = 10$. Son "boas" estas estimacións?



a) $\bar{x} = 5$, $\bar{y} = 6$

$\sigma_x = 2{,}8$; $\sigma_y = 2{,}7$; $\sigma_{xy} = 7{,}1$

b) $r = 0{,}95$

c) $y = 0{,}91x + 1{,}45$

d) $\hat{y}(5) = 6$, $\hat{y}(10) = 10{,}55$

Las estimaciones son muy fiables porque $r = 0{,}95$ es un valor muy alto. Si se tratase de "notas" (de 0 a 10), la segunda estimación habría que "hacerla real" y darle el valor 10.

3. **A recta de regresión de Y sobre X dunha certa distribución bidimensional é $y = 1{,}6x - 3$. Sabemos que $\bar{x} = 10$ e $r = 0{,}8$.**

a) Calcula $\bar{y}$.

b) Estima o valor de y para $x = 12$ e para $x = 50$. Que estimación parece máis fiable?

c) Indica a recta de regresión de X sobre Y.

a) Puesto que la recta pasa por $(\bar{x}, \bar{y})$:

$\bar{y} = 1{,}6\bar{x} - 3 = 1{,}6 \cdot 10 - 3 = 13$

b) $\hat{y}(12) = 1{,}6 \cdot 12 - 3 = 16{,}2$

$\hat{y}(50) = 1{,}6 \cdot 50 - 3 = 77$

La primera estimación es aceptable por ser 12 próximo a $\bar{x} = 10$ (carecemos de información sobre los valores que toma x). La segunda estimación es muy poco significativa, pues 50 se separa demasiado de $\bar{x}$.

c) Conociendo $r = 0{,}8$ y el coeficiente de regresión de Y sobre X (pendiente de la recta), 1,6:

(Coef. Y sobre X) $\cdot$ (Coef. X sobre Y) $= r^2$

Coef. X sobre $Y = \frac{0,8^2}{1,6} = 0,4$

Por tanto, la pendiente de la recta de regresión de X sobre Y es $m_{xy} = \frac{1}{0,4} = 2,5$.

Ecuación de la recta de regresión de X sobre Y: $y = 6 + 2,5(x - 5)$

4. O consumo de enerxía per cápita *y* en miles de kWh e a renda per cápita *x* en miles de euros de seis países son:

	A	B	C	D	E	F
x	11,1	8,5	11,3	4,5	9,9	6,5
y	5,7	5,0	5,1	2,7	4,6	3,1

a) Calcula a recta de regresión de *Y* sobre *X*.

b) Determina o coeficiente de correlación entre o consumo e a renda.

c) Que predición podemos facer sobre o consumo de enerxía per cápita dun país cuxa renda per cápita é de 4,4 miles de euros?

$\bar{x} = 8,63$, $\bar{y} = 4,37$

$\sigma_x = 2,46$, $\sigma_y = 1,09$, $\sigma_{xy} = 2,51$

a) Recta de regresión de Y sobre X:

$$y = 4,37 + \frac{2,51}{2,46^2}(x - 8,63) \rightarrow y = 0,79 + 0,41x$$

b) Coeficiente de correlación:

$$r = \frac{2,51}{1,09 \cdot 2,46} = 0,93$$

c) Para $x = 4,4$, estimamos el valor de y:

$\hat{y}(4,4) = 0,79 + 0,41 \cdot 4,4 = 2,59$

Se le estima un consumo de energía de 2,59 miles de Kw/h por habitante.