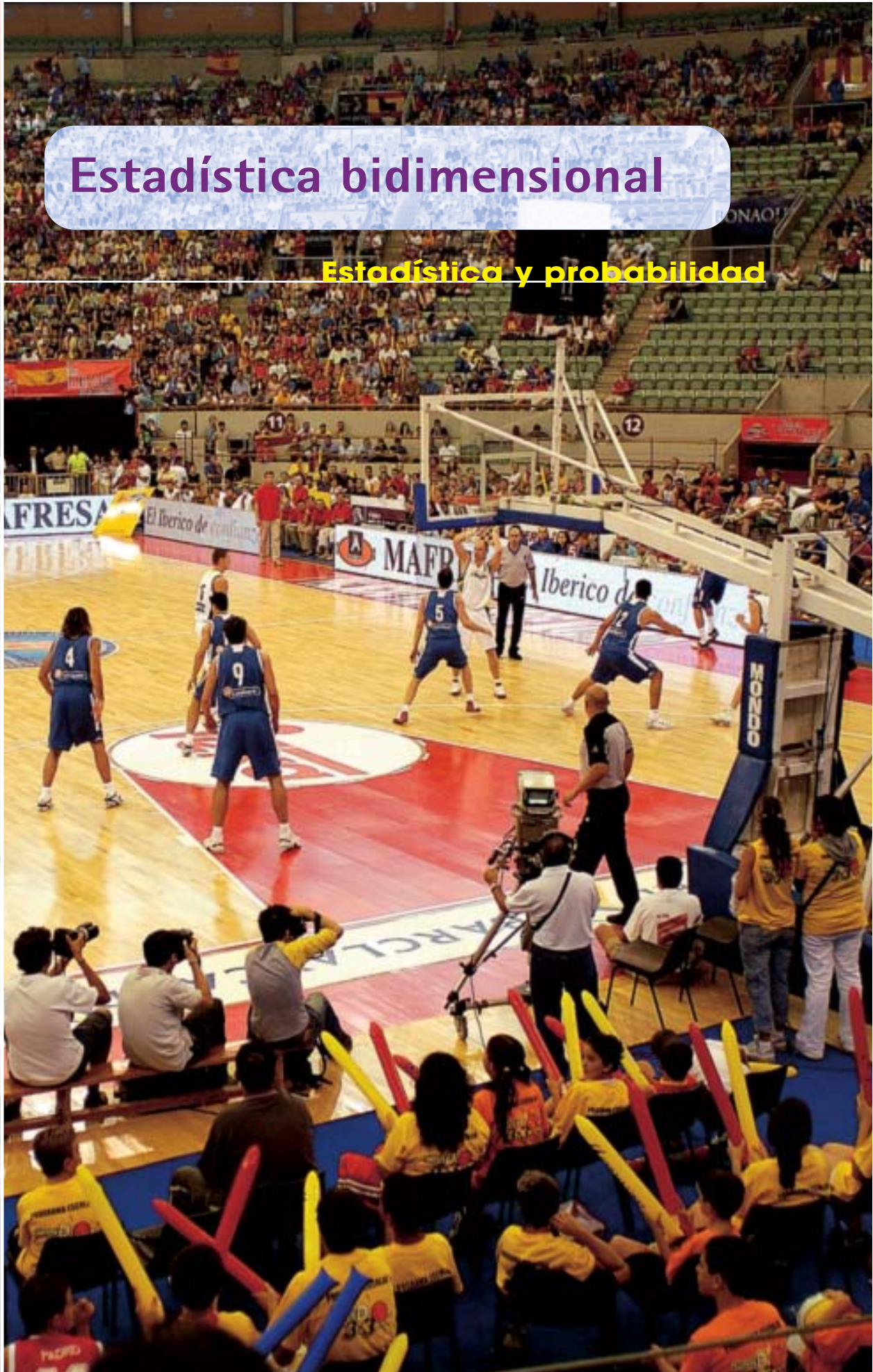


# 12

## Estadística bidimensional

### Estadística y probabilidad





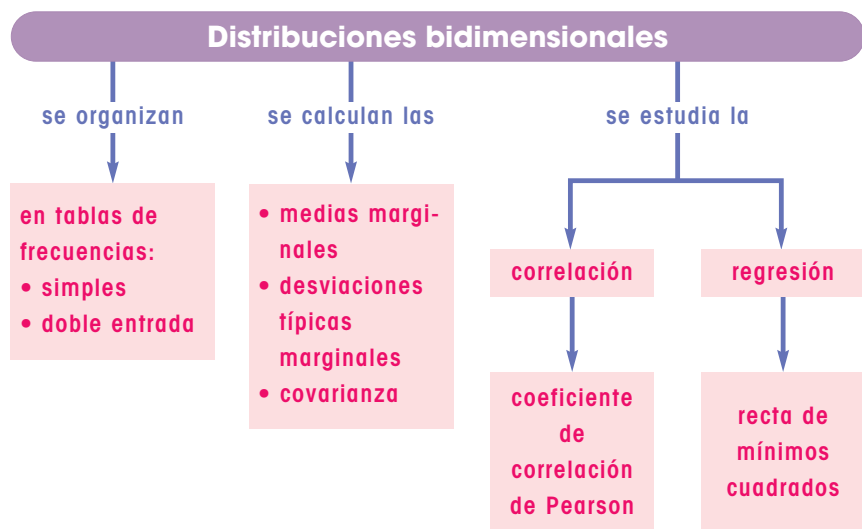


## Introducción

Las variables estadísticas pueden estudiarse de forma conjunta cuando se desea ver la relación entre dos o más caracteres del mismo individuo. Por ejemplo, se puede estudiar la relación que hay entre la estatura y el peso de unos jugadores de un equipo de baloncesto, o la relación que hay entre la cantidad de un medicamento y el tiempo que tarda en hacer reacción, etcétera.

En este tema se trabajan las distribuciones bidimensionales. En concreto, se estudia la forma de organizar la información en tablas de frecuencia, los parámetros que permiten interpretar dicha información, la correlación, que es la medida del grado de relación entre las variables, y la regresión, que estudia una variable condicionada al comportamiento de la otra.

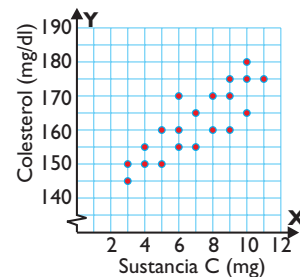
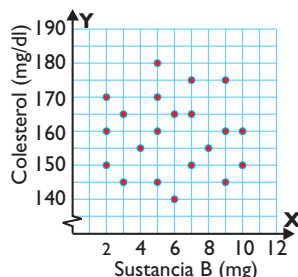
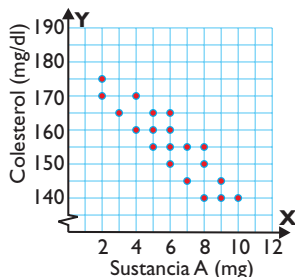
## Organiza tus ideas



# 1. Distribuciones bidimensionales

## Piensa y calcula

Se ha administrado una sustancia A, otra B y otra C a 20 individuos para estudiar su relación con los niveles de colesterol. Observando las gráficas, indica qué sustancia tiene mayor relación con la subida o bajada de colesterol.



### 1.1. Distribución bidimensional

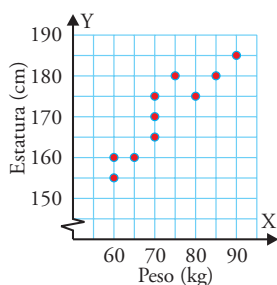
Una **distribución bidimensional** es la que se obtiene al estudiar un fenómeno respecto de dos variables estadísticas unidimensionales X e Y

Los **datos** de una distribución bidimensional son pares  $(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$ , donde  $x_1, x_2, \dots, x_n$  son los valores de la variable X, y donde  $y_1, y_2, \dots, y_n$  son los valores de la variable Y

#### Ejemplo

La distribución bidimensional que se obtiene al estudiar la estatura y el peso de 10 personas es:

<b>Peso (kg)</b>	70	65	85	60	70	75	90	80	60	70
<b>Estatura (cm)</b>	175	160	180	155	165	180	185	175	160	170

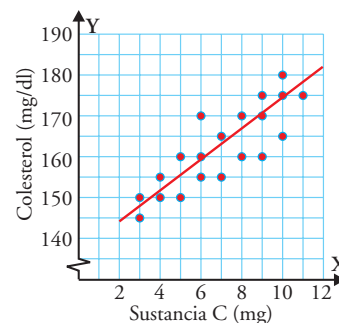
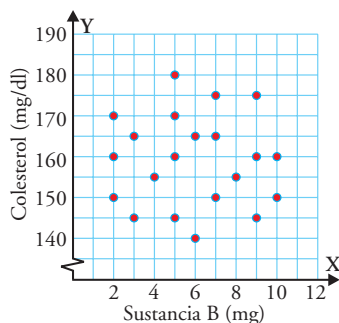
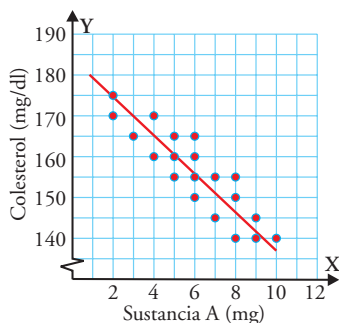


#### Nube de puntos o diagrama de dispersión

Una **nube de puntos** o **diagrama de dispersión** es la representación en unos ejes cartesianos de los datos  $(x_i, y_i)$  de una distribución bidimensional.

En una nube de puntos se puede apreciar de forma general la relación que existe entre las variables.

#### Ejemplo



En la gráfica de la izquierda se observa que a mayor cantidad de la sustancia A, los niveles de colesterol bajan. La gráfica de la derecha muestra lo contrario. Se puede decir que la sustancia A es buena para bajar el colesterol y que la sustancia C es perjudicial. En la gráfica central no hay relación.

## 1.2. Tabla de frecuencias

Cuando el número de datos de una distribución bidimensional es pequeño, se trabaja con los datos ordenados, pero cuando el número de datos es grande, se trabaja con tablas de frecuencias. Dichas tablas pueden darse de dos maneras:

- Tablas simples:** se recogen en fila o columna las frecuencias de los datos.
- Tablas de doble entrada:** se recoge en cada casilla la frecuencia correspondiente a cada fila y cada columna de los valores de cada variable.

### Ejemplo

Dada la siguiente distribución bidimensional:

Variable X	1	1	1	1	1	1	1	1	1	2	2	2	2	2	3	3	4	4	4	4
Variable Y	1	1	3	3	3	3	4	4	4	1	1	2	4	4	2	2	1	3	3	3

se obtienen las tablas de frecuencias.

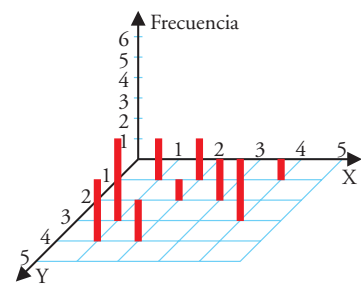
Tabla simple

X	Y	Frecuencia
1	1	2
1	3	4
1	4	3
2	1	2
2	2	1
2	4	2
3	2	2
4	1	1
4	3	3

Tabla de doble entrada

Y \ X	1	2	3	4	
1	2	2	0	1	5
2	0	1	2	0	3
3	4	0	0	3	7
4	3	2	0	0	5
	9	5	2	4	20

Diagrama de barras



### Aplica la teoría

- Las calificaciones de 30 estudiantes en dos exámenes han sido las siguientes:

1 <sup>er</sup> examen	4	5	6	7	7	9	10
2 <sup>o</sup> examen	5	5	7	6	7	8	10
Nº estudiantes	5	10	4	2	4	3	2

Haz la tabla de frecuencia de doble entrada.

- Dibuja el diagrama de barras correspondiente a la siguiente distribución bidimensional:

Y \ X	1	2	3	4	5
1	3	0	0	0	0
2	0	1	0	0	0
3	2	1	6	0	0
4	0	3	0	3	2
5	0	0	0	0	4

- Dibuja la nube de puntos de la siguiente distribución bidimensional:

X	Y
2	2
1	5
4	1
2	3
1	3
3	2
4	2
2	4
3	3
1	4

## 2. Parámetros

### ■ Piensa y calcula

La siguiente distribución recoge las calificaciones de Matemáticas y de Lengua de un grupo de 6 alumnos. Calcula mentalmente la media de cada asignatura:

<b>Matemáticas</b>	2	3	5	5	6	9
<b>Lengua</b>	4	4	5	6	7	10

### 2.1. Parámetros

Las **medias marginales** son las medias de las variables X e Y:

$$\bar{x} = \frac{\sum n_i \cdot x_i}{N} \qquad \bar{y} = \frac{\sum n_i \cdot y_i}{N}$$

El **centro de gravedad** es el par de valores de las medias marginales:  $G(\bar{x}, \bar{y})$

Las **desviaciones típicas marginales** son las desviaciones típicas de las variables X e Y:

$$s_x = \sqrt{\frac{\sum n_i \cdot (x_i - \bar{x})^2}{N}}$$

$$s_y = \sqrt{\frac{\sum n_i \cdot (y_i - \bar{y})^2}{N}}$$

$$s_x = \sqrt{\frac{\sum n_i \cdot x_i^2}{N} - \bar{x}^2}$$

$$s_y = \sqrt{\frac{\sum n_i \cdot y_i^2}{N} - \bar{y}^2}$$

#### Covarianza

La **covarianza** de una variable bidimensional (X, Y) es:

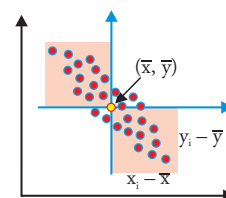
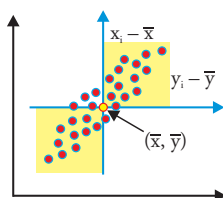
$$s_{xy} = \frac{\sum n_i (x_i - \bar{x})(y_i - \bar{y})}{N}$$

$$s_{xy} = \frac{\sum n_i x_i y_i}{N} - \bar{x} \cdot \bar{y}$$

#### Interpretación de la covarianza

Según sea el signo de la covarianza, se interpreta:

- Covarianza positiva:** al aumentar los valores de la variable X, aumentan los valores de la variable Y. La nube de puntos se orienta a la derecha y hacia arriba.
- Covarianza negativa:** al aumentar los valores de la variable X, disminuyen los valores de la variable Y. La nube de puntos se orienta a la derecha y hacia abajo.



Si se calcula el centro de gravedad  $G(\bar{x}, \bar{y})$  y se toman unos ejes con el origen en este centro, se observa:

- Si los puntos están en el 1<sup>er</sup> y 3<sup>er</sup> cuadrantes, mayoritariamente los productos  $(x_i - \bar{x})(y_i - \bar{y})$  son positivos.
- Si los puntos están en el 2<sup>o</sup> y 4<sup>o</sup> cuadrantes, mayoritariamente los productos  $(x_i - \bar{x})(y_i - \bar{y})$  son negativos.

**Ejemplo**

Calcula los parámetros de la distribución del número de años de antigüedad en una empresa y el salario diario que tienen 40 trabajadores. Interpreta los resultados.

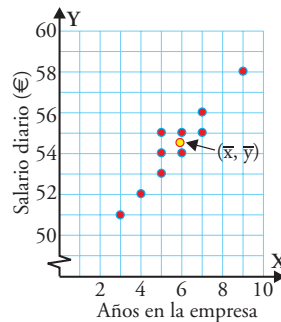
Años: $x_i$	Salario: $y_i$	$n_i$	$n_i \cdot x_i$	$n_i \cdot x_i^2$	$n_i \cdot y_i$	$n_i \cdot y_i^2$	$n_i \cdot x_i \cdot y_i$
4	52	3	12	48	156	8 112	624
5	54	4	20	100	216	11 664	1 080
7	55	5	35	245	275	15 125	1 925
6	54	7	42	252	378	20 412	2 268
5	53	3	15	75	159	8 427	795
7	56	5	35	245	280	15 680	1 960
5	55	4	20	100	220	12 100	1 100
9	58	3	27	243	174	10 092	1 566
3	51	2	6	18	102	5 202	306
6	55	4	24	144	220	12 100	1 320
<b>Total</b>		<b>40</b>	<b>236</b>	<b>1 470</b>	<b>2 180</b>	<b>118 914</b>	<b>12 944</b>

$$\bar{x} = \frac{236}{40} = 5,9 \quad \bar{y} = \frac{2 180}{40} = 54,5$$

$$s_x = \sqrt{\frac{1 470}{40} - 5,9^2} = 1,39$$

$$s_y = \sqrt{\frac{118 914}{40} - 54,5^2} = 1,61$$

$$s_{xy} = \frac{12 944}{40} - 5,9 \cdot 54,5 = 2,05$$



**Interpretación de los resultados:** al ser la covarianza positiva, la nube de puntos se orienta a la derecha y hacia arriba; es decir, al aumentar la antigüedad, aumenta el salario.

**Aplica la teoría**

4. Calcula la covarianza de la siguiente distribución bidimensional:

$x_i$	8	7	6	5	7	8	6	5
$y_i$	5	4	7	4	3	6	5	5
$n_i$	2	4	3	5	3	4	2	2

5. Calcula la covarianza de la siguiente distribución bidimensional:

Y \ X	2	4	6	8
1	1	3	0	2
2	2	5	1	0
3	3	1	4	6
4	0	2	0	0

6. La siguiente tabla recoge la distribución de la cilindrada de un motor y la velocidad máxima que puede generar:

Cilindrada (cm <sup>3</sup> )	Velocidad (km/h)
1 000	125
1 200	130
1 400	140
1 600	145
1 600	150
1 800	170
2 000	190
2 000	195

- Representa la nube de puntos.
- Representa el centro de gravedad.
- Calcula e interpreta la covarianza.

**Calculadora**

Se hace para la **CASIO fx-82MS**.

Para otras es análogo; ver el manual.

a) Se selecciona el modo **REG/Lin**

**[MODE]** (REG) **[3]** (Lin) **[1]**

b) Se borran todos los datos:

**[SHIFT]** **[CLR]** (Scl) **[1]** **[=]**

c) Se escribe el 1<sup>er</sup> dato de X, se pulsa

una coma **[,]**, se escribe el 1<sup>er</sup> dato

de Y y se pulsa **[DT]**

Si la frecuencia es mayor que uno,

se pulsa antes de la frecuencia **[;]**

como se indica:

**[4]** **[,]** **[52]** **[;]** **[3]** **[DT]** **[...]**

d) Si se introduce un dato erróneo, se

puede modificar utilizando **[▲]** o **[▼]**

para buscar el dato o la frecuencia;

se introduce y se pulsa la tecla **[=]**

e) Se obtienen los resultados:

• Media  $\bar{x}$

**[SHIFT]** **[S-VAR]** ( $\bar{x}$ ) **[1]** **[=]** **5,9**

• Media  $\bar{y}$

**[SHIFT]** **[S-VAR]** ( $\bar{y}$ ) **[1]** **[=]** **54,5**

• Desviación típica:  $s_x$

**[SHIFT]** **[S-VAR]** ( $s_x$ ) **[2]** **[=]** **1,39**

• Desviación típica:  $s_y$

**[SHIFT]** **[S-VAR]** ( $s_y$ ) **[2]** **[=]** **1,61**

• Covarianza:  $s_{xy}$

**[Σxy]** **[÷]** **[n]** **[-]** **[x]** **[x]** **[y]** **[=]** **2,05**

Para obtener:  $\Sigma xy$

**[SHIFT]** **[S-SUM]** ( $\Sigma xy$ ) **[3]**

Para obtener: n

**[SHIFT]** **[S-SUM]** (n) **[3]**

**Volver a modo normal la calculadora**

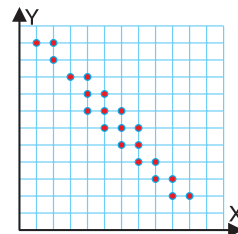
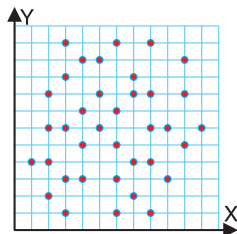
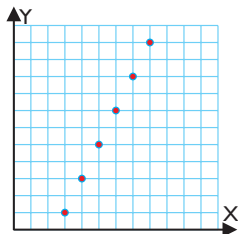
Se pulsas las teclas:

**[SHIFT]** **[CLR]** (Mode) **[2]** **[=]**

### 3. Correlación

#### Piensa y calcula

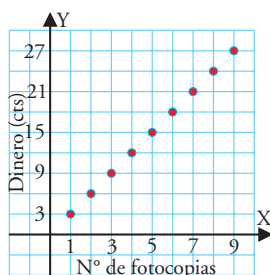
Indica el signo de la covarianza y si la relación entre las variables es funcional, fuerte o nula en los siguientes casos:



#### 3.1. Correlación

La **covarianza** indica cómo es la relación entre dos variables; es decir, cómo se orienta la nube de puntos, pero este parámetro no indica de una forma precisa la medida de esa relación. Para resolver este problema, se definen los conceptos de **correlación** y **coeficiente de correlación**.

**Correlación** es la relación que existe entre las dos variables que intervienen en una distribución bidimensional.



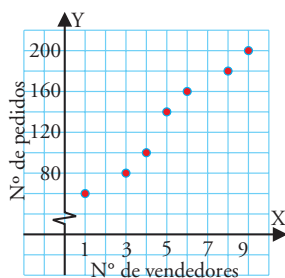
**a) Correlación funcional:** todos los puntos están situados sobre una recta o una curva. Existe una relación funcional entre las variables X e Y

##### Ejemplo

El precio de las fotocopias de una copistería es:

Nº copias: $x_i$	1	2	3	4	5	6
Dinero (cts): $y_i$	3	6	9	12	15	18

La función  $y = 3x$  da la relación entre las dos variables.

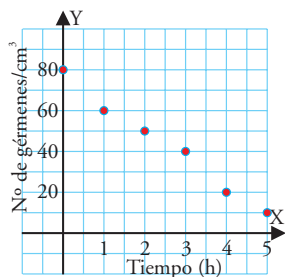


**b) Correlación directa:** al aumentar una variable, aumenta la otra.

##### Ejemplo

El número de pedidos que sirve un almacén y el número de vendedores que tiene contratados dicho almacén son:

Nº de vendedores: $x_i$	1	3	4	5	6	8	9
Nº de pedidos: $y_i$	60	80	100	140	160	180	200



**c) Correlación inversa:** al aumentar una variable, la otra disminuye.

##### Ejemplo

El número de gérmenes por  $cm^3$  y el tiempo transcurrido con un tratamiento específico son:

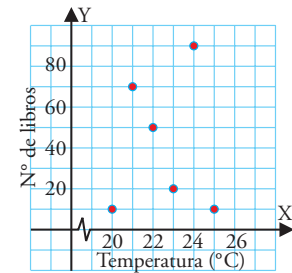
Tiempo (h): $x_i$	0	1	2	3	4	5
Nº de gérmenes: $y_i$	80	60	50	40	20	10

d) **Correlación nula:** no existe relación entre las variables.

**Ejemplo**

El número de libros vendidos en una librería y la temperatura del día es:

Temperatura (°C): $x_i$	20	21	22	23	24	25
Nº de libros: $y_i$	10	70	50	20	90	10



### 3.2. Coeficiente de correlación de Pearson

El **coeficiente de correlación de Pearson** es:

$$r = \frac{s_{xy}}{s_x s_y}$$

**Propiedades del coeficiente de correlación**

- a) El coeficiente de correlación de Pearson es un número. No depende de las unidades en las que están expresadas las variables  $x$  e  $y$
- b) Está comprendido entre  $-1$  y  $1$ 
  - Si  $r = -1$  o  $r = 1$ , la **correlación es perfecta o funcional**.
  - Si  $r$  está próximo a  $-1$  o a  $1$ , la **correlación es fuerte**.
  - Si  $r$  está próximo a cero, la **correlación es débil**.
  - El signo,  $r > 0$  o  $r < 0$ , indica si la correlación es directa o inversa, respectivamente.

**Ejemplo**

Calcula el coeficiente de correlación entre el número de pedidos que sirve un almacén y el número de vendedores que tiene contratados dicho almacén.

Nº de vendedores: $x_i$	2	4	5	6	7	9	10
Nº de pedidos: $y_i$	70	90	110	150	170	190	210

El coeficiente de correlación es:  $r = \frac{s_{xy}}{s_x s_y} = \frac{124,08}{2,59 \cdot 48,82} = 0,98$

La correlación es fuerte y directa.

**Interpretación**

El coeficiente de correlación de Pearson indica la correlación que existe entre las dos variables; es decir, si los puntos están muy próximos o alejados del centro de gravedad.

**Correlación fuerte**

Se considera que la correlación es fuerte si  $|r| > 0,85$

**Calculadora**

Se introducen los datos y se pulsa:



**Aplica la teoría**

7. Calcula el coeficiente de correlación e indica el tipo de correlación para la siguiente distribución bidimensional:

$x_i$	1	4	4	2	5	3	1
$y_i$	5	2	3	6	3	2	4

8. La temperatura media en los meses de invierno en varias ciudades y el gasto medio por habitante en calefacción han sido:

Temperatura (°C)	10	12	14	15	17	20
Gasto (€)	150	120	102	90	50	18

Calcula el coeficiente de correlación e interpreta el resultado.

9. Calcula el coeficiente de correlación e indica el tipo de correlación para la siguiente distribución bidimensional:

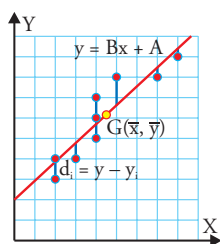
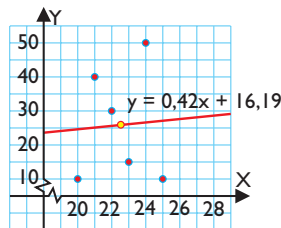
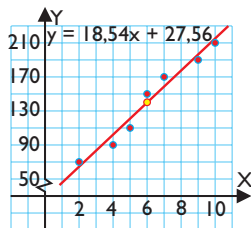
$Y \backslash X$	1	2	3	4
1	1	2	0	0
2	2	1	0	0
3	0	1	2	3
4	0	4	3	1



## 4. Regresión

### Piensa y calcula

Se han ajustado las nubes de puntos adjuntas según las rectas dadas. Calcula el valor de  $y$  para  $x = 20$  en la 1ª y  $x = 30$  en la 2ª. ¿Qué estimación crees que es más fiable?

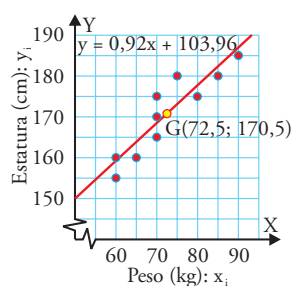


La ecuación de la **recta de regresión de x sobre y** es:

$$x - \bar{x} = \frac{s_{xy}}{s_y^2}(y - \bar{y})$$

Pendiente de la recta de **x sobre y**:

$$m_{xy} = \frac{s_{xy}}{s_y^2}$$



#### Calculadora

Se introducen los datos y se obtienen los parámetros de la recta:

$$y = Bx + A$$

SHIFT [S-VAR] [▶▶] (B) [2]

= [0,92]

SHIFT [S-VAR] [▶▶] (A) [1]

= [103,96]

### 4.1. Rectas de regresión de los mínimos cuadrados

La ecuación de la **recta de regresión de y sobre x** es:

$$y - \bar{y} = \frac{s_{xy}}{s_x^2}(x - \bar{x})$$

Estas rectas se pueden expresar de la forma  $y = Bx + A$ . Las letras **A** y **B** son las que usan, generalmente, las calculadoras.

Estas rectas se determinan haciendo que se cumplan las siguientes condiciones:

- Tienen que pasar por el centro de gravedad  $G(\bar{x}, \bar{y})$
- La suma de los cuadrados de las distancias,  $\sum d_i^2$ , debe ser mínima. Siendo  $d_i = y - y_i$ ;  $y$ , ordenada de la recta;  $y_i$ , ordenada de cada punto.

### Coefficientes de regresión

Los **coeficientes de regresión** son las pendientes de las rectas de regresión:

Pendiente de la recta de **y sobre x**:  $m_{yx} = \frac{s_{xy}}{s_x^2}$

#### Ejemplo

Calcula la recta de regresión de la altura sobre el peso en la distribución que se obtiene al estudiar la estatura y el peso de 10 personas.

<b>Peso (kg): <math>x_i</math></b>	70	65	85	60	70	75	90	80	60	70
<b>Estatura (cm): <math>y_i</math></b>	175	160	180	155	165	180	185	175	160	170

Se tiene:

Medias marginales:  $\bar{x} = 72,5$ ,  $\bar{y} = 170,5$

Desviaciones típicas marginales:  $s_x = 9,55$ ,  $s_y = 9,6$

Covarianza:  $s_{xy} = 83,75$

Pendiente:  $\frac{s_{xy}}{s_x^2} = \frac{83,75}{9,55^2} = 0,92$

La recta de regresión de **y sobre x** es:

$$y - 170,5 = 0,92(x - 72,5) \Rightarrow y = 0,92x + 103,96$$

## 4.2. Estimaciones con la recta de regresión

La recta de regresión se ajusta a la nube de puntos y describe, de una forma genérica, su tendencia. Se pueden hacer estimaciones con la recta de regresión sin olvidar que la estimación será fiable siempre que:

- el coeficiente de correlación esté próximo a 1 o a -1. Es decir, los puntos están muy cerca de la recta.
- los valores sobre los que se hacen las estimaciones no estén muy alejados de los datos utilizados.

### Ejemplo

Se han recogido los datos de la temperatura en °C y la presión en mm en distintas ciudades.

- Estima la presión que habría para una temperatura de 23°
- Estima la temperatura si la presión fuese de 830 mm

<b>Temperatura (°C): <math>x_i</math></b>	18	19	20	18	22	21
<b>Presión (mm): <math>y_i</math></b>	790	800	805	795	820	810

### Solución

- El coeficiente de correlación es:  $r = 0,98$

La correlación es fuerte y directa. Para hacer la estimación se calcula la recta de regresión:

$$y = 6,5x + 675,5$$

y con ella se calcula el valor de  $y$  para  $x = 23$

$$y = 6,5 \cdot 23 + 675,5 = 825$$

Es decir, aproximadamente 825 mm de presión.

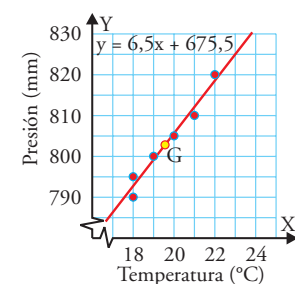
- Utilizando la recta de regresión de  $y$  sobre  $x$ , se tiene:

$$\text{Si } y = 830 \Rightarrow 830 = 6,5x + 675,5$$

Se obtiene:  $x = 23,77$  °C

### Evitar errores

Para hacer una estimación de  $x$  para un determinado valor de  $y$ , se puede utilizar la recta de regresión de  $y$  sobre  $x$  siempre que  $|r|$  sea próximo a uno. En este caso, las estimaciones son muy fiables.



### Recta de $x$ sobre $y$

La recta de regresión de  $x$  sobre  $y$  es:

$$x - 19,67 = 0,5(y - 803,33)$$

Para  $y = 830$  mm, se tiene:

$$x = 23,67$$
 °C

Se observa que la estimación realizada con la recta de  $y$  sobre  $x$  es muy parecida.

## Aplica la teoría

- Calcula la recta de regresión de la siguiente distribución bidimensional:

<b>X</b>	1	2	3	4	5
<b>Y</b>	26	30	27	31	28

- Un laboratorio ha experimentado, en 6 pacientes, con un medicamento para bajar la temperatura de los enfermos, observado el tiempo que tarda en desaparecer, y ha obtenido los resultados siguientes:

<b>Dosis (mg)</b>	100	200	300	400	500	600
<b>Tiempo (h)</b>	4	3,5	3	2	2,5	1,5

Calcula la recta de regresión y estima el tiempo que tardaría en normalizarse la temperatura para 650 mg

- En una empresa, la relación entre el número de piezas defectuosas que elaboran unos trabajadores y la antigüedad de éstos es:

<b>Antigüedad</b>	1	2	3	4	5	6
<b>Nº piezas</b>	7	8	6	4	3	2

- Calcula la recta de regresión.
- Estima el número de piezas defectuosas que haría un obrero con 7 años de antigüedad.
- Estima el tiempo que llevaría trabajando un obrero si no hiciese piezas defectuosas.

# Ejercicios y problemas

## 1. Distribuciones bidimensionales

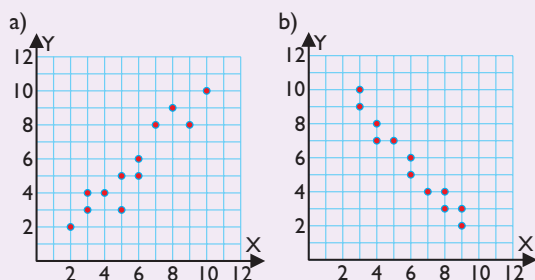
13. Haz la tabla de frecuencia de doble entrada de la siguiente distribución bidimensional:

<b>X</b>	14	16	16	17	17	19	20
<b>Y</b>	20	19	21	20	21	20	21
<b>n<sub>i</sub></b>	6	12	8	5	4	3	4

14. Dibuja el diagrama de barras correspondiente a la siguiente distribución bidimensional:

<b>Y \ X</b>	1	2	3	4
1	0	0	0	0
2	0	2	0	1
3	2	5	6	5
4	4	3	0	3

15. Haz la tabla de frecuencias de las siguientes nubes de puntos:



## 2. Parámetros

16. Calcula la covarianza de la siguiente distribución bidimensional:

<b>x<sub>i</sub></b>	5	7	4	5	7	8	4	6
<b>y<sub>i</sub></b>	11	11	10	10	12	11	9	10
<b>n<sub>i</sub></b>	2	4	3	3	4	2	3	5

17. Calcula la covarianza de la siguiente distribución bidimensional:

<b>Y \ X</b>	3	5	7	9	11
24	3	0	0	1	0
26	2	4	3	0	0
28	0	3	2	4	2
30	1	0	1	3	4

18. La siguiente tabla recoge el crecimiento de una planta según los gramos de abono que se le suministran. Abono (g): X; crecimiento (cm): Y

<b>X</b>	1	2	3	4	5	6	7	8
<b>Y</b>	5	5	6	7	9	10	12	12

- a) Representa la nube de puntos.  
b) Representa el centro de gravedad.  
c) Calcula e interpreta la covarianza.

## 3. Correlación

19. Calcula el coeficiente de correlación e indica el tipo de correlación para la siguiente distribución bidimensional:

<b>x<sub>i</sub></b>	14	15	16	19	17	15
<b>y<sub>i</sub></b>	80	81	80	82	81	78

20. La temperatura en grados y la presión atmosférica en milímetros en distintas ciudades son:

<b>Temp. (°C)</b>	12	13	14	17	15	13	16
<b>Presión (mm)</b>	800	805	803	810	805	800	810

Calcula el coeficiente de correlación e interpreta el resultado.

21. Calcula el coeficiente de correlación e indica el tipo de correlación para la siguiente distribución bidimensional:

<b>X</b>	3	4	5	6	6	7	8	8	10
<b>Y</b>	3	6	6	7	8	7	8	10	10
<b>n<sub>i</sub></b>	5	7	13	5	6	5	4	2	3

## 4. Regresión

22. Calcula la recta de regresión de y sobre x de la siguiente distribución bidimensional:

<b>X</b>	2	4	6	7	8
<b>Y</b>	16	20	25	34	34

Calcula el valor de y para x = 9 y el valor de x para y = 30

# Ejercicios y problemas

23. Calcula la recta de regresión de  $y$  sobre  $x$  de la siguiente distribución bidimensional:

Y \ X	0	1	2	3	4
0	5	2	0	0	0
1	1	8	6	7	0
2	2	5	10	10	0
3	0	1	4	6	2
4	0	0	6	0	0

24. Las 10 últimas cotizaciones de dos empresas dedicadas a dar servicios por Internet han sido:

X	8,20	8,15	8,40	8,50	8,88	8,81	8,87	8,75	8,87	8,99
Y	4,80	4,83	4,90	4,88	4,95	4,96	4,88	4,80	4,85	4,92

Calcula la recta de regresión de  $y$  sobre  $x$  y analiza si sería fiable hacer alguna estimación.

## Para ampliar

25. La tabla siguiente recoge los datos de un grupo de estudiantes con las horas dedicadas al estudio de un examen,  $X$ , y la calificación obtenida,  $Y$ :

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

Dibuja la nube de puntos e indica si sobre ella se puede deducir alguna relación.

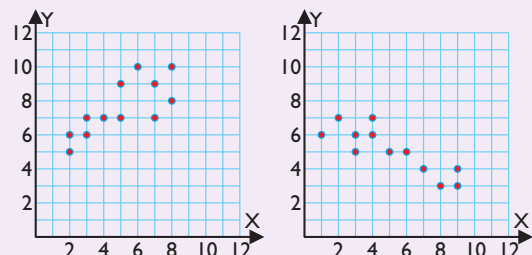
26. Haz la tabla de frecuencia de doble entrada de la siguiente distribución bidimensional:

X	6	6	5	4	4	3	2
Y	2	3	3	3	4	5	5
$n_i$	7	10	6	4	5	2	3

27. Dibuja el diagrama de barras correspondiente a la siguiente distribución bidimensional:

Y \ X	1	2	3	4
1	0	4	3	4
2	1	1	2	1
3	3	5	0	0
4	2	0	0	0

28. Haz la tabla de frecuencias de las siguientes nubes de puntos:



29. Calcula la covarianza de la siguiente distribución bidimensional:

$x_i$	1	1	3	4	6	6	8	9	9	10
$y_i$	0	4	2	5	8	5	4	8	10	8
$n_i$	2	3	2	4	3	2	3	4	2	5

30. Calcula la covarianza de la siguiente distribución bidimensional:

Y \ X	0	1	2	3
60	4	2	0	0
70	6	7	8	0
80	0	3	3	4
90	0	0	0	3



31. Calcula el coeficiente de correlación e indica el tipo de correlación para la siguiente distribución bidimensional:

$x_i$	65	63	67	64	68	62	70	66
$y_i$	68	66	68	65	69	66	68	65

32. Calcula el coeficiente de correlación e indica el tipo de correlación para la siguiente distribución bidimensional:

<b>X</b>	175	181	192	211	235	255	275	286	292
<b>Y</b>	169	185	202	219	240	266	295	329	357

33. Calcula la recta de regresión de  $y$  sobre  $x$  de la siguiente distribución bidimensional:

<b>X</b>	6	5	8	8	7	6	10
<b>Y</b>	8	7	10	9	8	8	11

34. Calcula la recta de regresión de  $y$  sobre  $x$  de la siguiente distribución bidimensional:

	<b>X</b>	4	5	6	7
<b>Y</b>	5	3	2	0	0
	6	4	6	0	0
	7	0	5	8	0
	8	0	0	2	6

## Problemas

35. La siguiente tabla recoge la estatura en centímetros de un grupo de padres (X) y sus respectivos hijos mayores (Y):

<b>X</b>	170	168	170	165	175	169	180	175	173
<b>Y</b>	173	170	173	170	178	170	179	172	180

Calcula:

- el coeficiente de correlación.
- la recta de regresión de  $y$  sobre  $x$
- Estima la estatura de un hijo cuyo padre mida 185 cm, e indica si la estimación es fiable.

36. La siguiente tabla muestra el cierre de los últimos días de los índices del IBEX35 (X) y Dow Jones 30 (Y):

<b>X</b>	8 236,9	8 164,7	8 236,1	8 202,1	8 241,2
<b>Y</b>	10 334,5	10 235,1	10 198,2	10 313,7	10 356,4

- Calcula el coeficiente de regresión.
- Calcula la recta de regresión del Dow Jones sobre el IBEX.

37. El rendimiento anual obtenido según la inversión realizada, en miles de euros, en una plantación agrícola es:

<b>Inversión</b>	12	14	16	15	18	20	21	15
<b>Rendimiento</b>	3	3,5	4,5	5	6	6,5	7,5	4,5

Calcula:

- el coeficiente de correlación.
- la recta de regresión del rendimiento sobre la inversión.
- Estima el rendimiento para una inversión de 22 000 €, e indica si la estimación es fiable.

38. Las estaturas y los pesos de un grupo de personas son:

<b>Estatura</b>	175	180	180	185	183	180	190	175
<b>Peso</b>	77	79	80	82	80	80	85	75

Calcula:

- el coeficiente de correlación.
- la recta de regresión del peso sobre la estatura.
- Estima el peso para una persona que mida 195 cm, e indica si la estimación es fiable.

39. En un taller de artesanía se ha registrado el número de piezas acabadas que unos artesanos hacen según las horas de trabajo:

<b>Horas</b>	8	7,5	8	8,5	6	7	8	9
<b>Nº piezas</b>	3	4	4	5	2	3	5	4

Calcula:

- el coeficiente de correlación.
- la recta de regresión del número de piezas sobre el número de horas.
- Estima el número de piezas para 10 h de trabajo, e indica si la estimación es fiable.

40. De una goma se cuelgan distintos pesos en gramos y se mide el alargamiento en centímetros producido; se obtienen los siguientes resultados:

<b>Peso (g)</b>	10	20	30	40	50	60	70	80
<b>Alargamiento (cm)</b>	2	4	7	10	12	15	18	20

Calcula:

- el coeficiente de correlación.
  - la recta de regresión del alargamiento sobre el peso.
  - Estima el alargamiento que se producirá para un peso de 90 g, e indica si la estimación es fiable.
41. Calcula el centro de gravedad, las desviaciones típicas marginales, la covarianza y el coeficiente de correlación de la siguiente distribución:

<b>Cilindrada (cm<sup>3</sup>)</b>	1000	1200	1400	1600	1600	1800	2000	2000
<b>Velocidad (km/h)</b>	125	130	140	145	150	170	190	195

- Representa la nube de puntos y calcula la recta de regresión de  $y$  sobre  $x$ , e interpreta los resultados.
  - Un coche tiene 1900 cm<sup>3</sup> de cilindrada. ¿Qué velocidad máxima alcanzará?
  - Un coche tiene una velocidad máxima de 150 km/h. ¿Qué cilindrada tendrá?
42. Calcula el centro de gravedad, las desviaciones típicas marginales, la covarianza y el coeficiente de correlación de la siguiente distribución:

<b>Nº de vendedores: <math>x_i</math></b>	2	4	5	6	7	9	10
<b>Nº de pedidos: <math>y_i</math></b>	70	90	110	150	170	190	210

- Representa la nube de puntos y calcula la recta de regresión de  $y$  sobre  $x$ , e interpreta los resultados.
  - Si hubiese 12 vendedores, ¿cuántos pedidos se esperarían?
  - Para obtener 250 pedidos, ¿cuántos vendedores harían falta?
43. El número de bacterias por centímetro cúbico que hay en un cultivo, según el paso del tiempo, es:

<b>Tiempo (h)</b>	0	1	2	3	4	5	6
<b>Nº de bacterias</b>	10	17	24	32	40	48	52

Calcula:

- el coeficiente de correlación.
- la recta de regresión del número de bacterias sobre el tiempo.
- Estima el número de bacterias que habrá después de 7 horas, e indica si la estimación es fiable.

44. En una compañía telefónica, han registrado en una muestra los siguientes datos sobre el número de teléfonos y el número de llamadas interurbanas realizadas:

<b>Nº de teléfonos</b>	550	600	650	700	750	800
<b>Nº de llamadas</b>	50	55	58	62	65	70

Calcula:

- el coeficiente de correlación.
- la recta de regresión del número de llamadas sobre el número de teléfonos.
- Estima el número de llamadas para 850 teléfonos, e indica si es fiable la estimación.

### Para profundizar

45. Se ha medido experimentalmente la presión del vapor del agua en centímetros de mercurio según la temperatura, y se han obtenido los siguientes resultados:

<b>Temperatura</b>	0	10	20	30	40	50
<b>Presión</b>	0,5	0,8	1,6	3	5,5	9

- Calcula el coeficiente de correlación.
- Dibuja la nube de puntos.
- ¿Qué tipo de curva crees que se ajustará mejor a estos puntos?

46. Las calificaciones de un grupo de estudiantes en Matemáticas y en Física se distribuyen así:

$Y_{(Física)}$ \ $X_{(Matem.)}$	0 a 2	2 a 4	4 a 6	6 a 8	8 a 10
0 a 2	6	2			
2 a 4	8	14	1		
4 a 6	1	3	12	1	
6 a 8		2	4	12	1
8 a 10		1	1	1	10

Calcula:

- el coeficiente de correlación.
- la recta de regresión de  $y$  sobre  $x$
- Estima la calificación en Física para un alumno que haya sacado un 7,5 en Matemáticas.
- ¿Se debería hacer la recta de regresión de  $x$  sobre  $y$  para estimar la calificación en Matemáticas de un alumno que haya obtenido un 6,5 en Física? Haz dicha estimación.

# Tema 12. Estadística bidimensional

## Paso a paso



Aumentar decimales



Combinar y centrar



Color de relleno



Disminuir decimales



Bordes



Color de fuente

Ajustar siempre los resultados a dos decimales.

### Autoajustar el ancho de una columna

Para autoajustar el ancho de una columna al contenido, se coloca el ratón en la cabecera de las columnas, entre la columna que se desea autoajustar y la siguiente; cuando el cursor se transforma en se hace *doble-clic*.

47. Calcula el centro de gravedad, las desviaciones típicas marginales, la covarianza y el coeficiente de correlación de la siguiente distribución:

Peso (kg)	70	65	85	60	70	75	90	80	60	70
Estatura (cm)	175	160	180	155	165	180	185	175	160	170

Representa la nube de puntos y calcula la recta de regresión de  $y$  sobre  $x$ , e interpreta los resultados.

- Una persona pesa 95 kg. ¿Cuánto medirá?
- Una persona mide 177 cm. ¿Cuánto pesará?

#### Solución:

Abre **Microsoft Excel** y, en la **Hoja 1**, copia los datos iniciales que hay en la tabla siguiente.

Observa que el rango **A1:C1** está combinado, y lo mismo **B15:C15** y **B16:C16**

	A	B	C
1	Relación Peso-Estatura		
2		Peso (kg)	Estatura (cm)
3		70	175
4		65	160
5		85	180
6		60	155
7		70	165
8		75	180
9		90	185
10		80	175
11		60	160
12		70	170
13		Centro de gravedad	
14		Desviaciones típicas marginales	
15		Covarianza	
16		Coefficiente de correlación	
17		Predecir resultados	
18		Buscar objetivos	

#### Centro de gravedad

- Sitúa el cursor en la celda **B13** y elige **Insertar función**
- En el cuadro de texto **O seleccionar una categoría**, elige **Estadísticas**
- Busca la función **PROMEDIO**, y en el rango **selecciona con el ratón B3:B12**; debes obtener: **72,5**
- Arrastra el **Controlador de relleno** de la celda **B13** hasta **C13**; debes obtener: **170,5**

#### Desviaciones típicas marginales

- Sitúa el cursor en la celda **B14**, busca la función **DESVESTP**, y **selecciona con el ratón** en el rango **B3:B12**, disminuye a dos decimales; debes obtener: **9,55**

- Arrastra el **Controlador de relleno** de la celda **B14** hasta **C14**; debes obtener: **9,60**

#### Covarianza

- Sitúa el cursor en la celda **B15**, busca la función **COVAR**, en **Matriz1** **selecciona con el ratón** el rango **B3:B12**, y en **Matriz2**, el rango de datos **C3:C12**; debes obtener: **83,75**

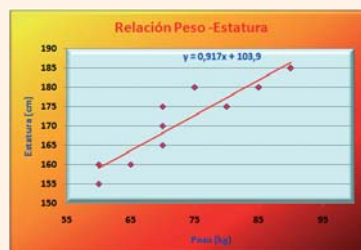
**Interpretación de la covarianza:** al ser positiva, el peso y la estatura se relacionan de forma directa; es decir, al aumentar los valores del peso, aumentan los valores de la estatura, por lo que la nube de puntos se orientará a la derecha y arriba.

#### Coefficiente de correlación

- Sitúa el cursor en la celda **B16**, busca la función **COEF.DE.CORREL**, en **Matriz1** **selecciona con el ratón** el rango **B3:B12**, y en **Matriz2**, el rango de datos **C3:C12**; debes obtener: **0,91**

**Interpretación del coeficiente de correlación:** al ser un número cercano a 1, la correlación es fuerte.


#### Nube de puntos y recta de regresión




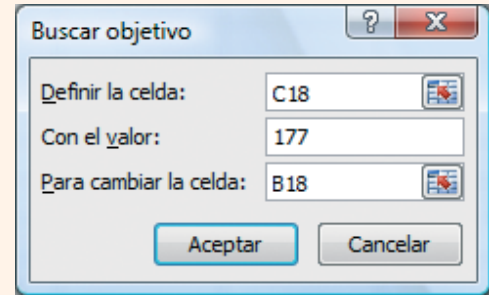
- Selecciona con el ratón el rango **B3:C12**
- En el menú **Insertar**, elige
- Elimina la leyenda de la parte derecha.
- Selecciona los puntos del gráfico haciendo *click* sobre uno de ellos. En su menú *Contextual* elige **Agregar línea de tendencia...**; elige **Lineal**; y activa la casilla de verificación **Presentar ecuación en el gráfico**

- e) En el menú *Contextual* del eje **X**, elige **Dar formato a eje...**, en **Mínima** activa el botón de opción **Fija** y escribe **55**
- f) Ponle al eje **Y** de mínimo **150**
- g) Mejora la presentación del gráfico para que quede como el del libro, o mejor. Para poner los títulos, elige el menú **Presentación**

### Predecir resultados

- Una persona pesa 95 kg. ¿Cuánto medirá?
  - a) En la celda **B17** escribe **95**
  - b) En la celda **C17** introduce la fórmula obtenida  $=0,917*B17 + 103,9$ . Debes obtener **191,15**
- Una persona mide 177 cm. ¿Cuánto pesará?
  - a) Arrastra el  **Controlador de relleno** de la celda **C17** hasta **C18**

- b) En el menú **Datos**, elige  **/Buscar objetivo...** En la ventana que aparece, escribe en **Definir la celda: C18**, **Con el valor: 177**, **Para cambiar la celda: B18**. Debes obtener **79,72**



- 48. **Internet.** Abre la web: [www.editorial-bruno.es](http://www.editorial-bruno.es), elige **Matemáticas**, curso y tema.

## Así funciona

### Funciones de estadística bidimensional utilizadas

**PROMEDIO:** media o media aritmética.

**COVAR:** covarianza.

**DESVESTP:** desviación típica.

**COEF.DE.CORREL:** coeficiente de correlación.

## Practica

Elimina las hojas: **Hoja2** y **Hoja3**. Los problemas 49, 50 y 51 son muy parecidos al 47; para hacerlos, en la **Hoja1** se elige en el menú *Contextual* de la pestaña de la hoja **Mover o copiar**, se selecciona (**mover al final**) y se activa la casilla de verificación **Crear una copia**. Para terminar, se hacen los cambios oportunos.

- 49. Calcula el centro de gravedad, las desviaciones típicas marginales, la covarianza y el coeficiente de correlación de la siguiente distribución:

<b>Cilindrada (cm<sup>3</sup>)</b>	1000	1200	1400	1600	1600	1800	2000	2000
<b>Velocidad (km/h)</b>	125	130	140	145	150	170	190	195

- a) Representa la nube de puntos y calcula la recta de regresión de **y** sobre **x**, e interpreta los resultados.
- b) Un coche tiene de cilindrada 1 900 cm<sup>3</sup>. ¿Qué velocidad máxima alcanzará?
- c) Un coche tiene una velocidad máxima de 150 km/h. ¿Qué cilindrada tendrá?

- 50. Calcula el centro de gravedad, las desviaciones típicas marginales, la covarianza y el coeficiente de correlación de la siguiente distribución:

<b>Nº de vendedores: x<sub>i</sub></b>	2	4	5	6	7	9	10
<b>Nº de pedidos: y<sub>i</sub></b>	70	90	110	150	170	190	210

- a) Representa la nube de puntos y calcula la recta de regresión de **y** sobre **x**, e interpreta los resultados.
- b) Si hubiese 12 vendedores, ¿cuántos pedidos se esperarían?
- c) Para obtener 250 pedidos, ¿cuántos vendedores harían falta?

- 51. Calcula el centro de gravedad, las desviaciones típicas marginales, la covarianza y el coeficiente de correlación de la siguiente distribución:

<b>Temperatura (°C): x<sub>i</sub></b>	18	19	20	18	22	21
<b>Presión (mm): y<sub>i</sub></b>	790	800	805	795	820	810

- a) Representa la nube de puntos y calcula la recta de regresión de **y** sobre **x**, e interpreta los resultados.
- b) Para una temperatura de 23 °C, ¿qué presión habrá?
- c) Para una presión de 900 mm, ¿qué temperatura habrá?



# Tema 12. Estadística bidimensional

## Paso a paso

Añadir decimal

Combinar celdas

Color de fondo

Borde

Eliminar decimal

Centrar

Color de fuente

Negrita

Ajustar siempre los resultados a dos decimales.

### Autoajustar el ancho de una columna

Para autoajustar el ancho de una columna al contenido, se coloca el ratón en la cabecera de las columnas, entre la columna que se desea autoajustar y la siguiente; cuando el cursor se transforma en doble flecha horizontal, se hace *doble-clic*.

47. Calcula el centro de gravedad, las desviaciones típicas marginales, la covarianza y el coeficiente de correlación de la siguiente distribución:

Peso (kg)	70	65	85	60	70	75	90	80	60	70
Estatura (cm)	175	160	180	155	165	180	185	175	160	170

Representa la nube de puntos y calcula la recta de regresión de  $y$  sobre  $x$ , e interpreta los resultados.

La recta de regresión es:  $y = 0,917x + 103,9$

- Una persona pesa 95 kg. ¿Cuánto medirá?
- Una persona mide 177 cm. ¿Cuánto pesará?

#### Solución:

Abre **Calc** y, en la **Hoja1**, copia los datos iniciales que hay en la tabla siguiente.

Observa que el rango **A1:C1** está combinado, y lo mismo **B15:C15** y **B16:C16**

	A	B	C
1	Relación Peso-Estatura		
2		Peso (kg)	Estatura (cm)
3		70	175
4		65	160
5		85	180
6		60	155
7		70	165
8		75	180
9		90	185
10		80	175
11		60	160
12		70	170
13		Centro de gravedad	
14		Desviaciones típicas marginales	
15		Covarianza	
16		Coefficiente de correlación	
17		Predicir resultados	
18		Buscar objetivos	

#### Centro de gravedad

- Sitúa el cursor en la celda **B13** y elige **Asistente: Funciones**
- En el cuadro de texto **Categoría**, elige **Estadística**
- Busca la función **PROMEDIO**, y en el rango **selecciona con el ratón B3:B12**; debes obtener: **72,5**
- Arrastra el **Controlador de relleno** de la celda **B13** hasta **C13**; debes obtener: **170,5**

#### Desviaciones típicas marginales

- Sitúa el cursor en la celda **B14**, busca la función **DESVESTP**, y **selecciona con el ratón** el rango **B3:B12**, disminuye a dos decimales; debes obtener: **9,55**

- Arrastra el **Controlador de relleno** de la celda **B14** hasta **C14**; debes obtener: **9,60**

#### Covarianza

- Sitúa el cursor en la celda **B15**, busca la función **COVAR**, en **Datos\_1 selecciona con el ratón** el rango **B3:B12**, y en **Datos\_2**, el rango de datos **C3:C12**; debes obtener: **83,75**

**Interpretación de la covarianza:** al ser positiva, el peso y la estatura se relacionan de forma directa; es decir, al aumentar los valores del peso, aumentan los valores de la estatura, por lo que la nube de puntos se orientará a la derecha y arriba.

#### Coefficiente de correlación


- Sitúa el cursor en la celda **B16**, busca la función **COEF.DE.CORREL**, en **Datos\_1 selecciona con el ratón** el rango **B3:B12**, y en **Datos\_2**, el rango de datos **C3:C12**; debes obtener: **0,91**

**Interpretación del coeficiente de correlación:** al ser un número mayor que 0,85, es decir, cercano a 1, la correlación es fuerte.

#### Nube de puntos y recta de regresión




- Selecciona con el ratón el rango **B3:C12**
- Elige **Insertar diagrama** y haz *clic* en cualquier lugar y *arrastra*.
- Haz *clic* en **Siguiente**
- Selecciona el gráfico **Diagrama XY**, activa la casilla de verificación **Representación de texto en previsualización**

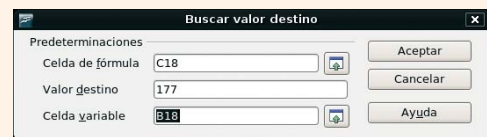
- e) Elige **Avanzar/Avanzar**, desactiva la leyenda, escribe los títulos y elige **Crear**
- f) Haz *doble-clic* sobre el gráfico y selecciona los puntos haciendo *doble-clic* sobre uno de ellos. En la ficha **Estadística**, elige la opción  **Regresión lineal**
- g) En el menú *Contextual* del eje X, **Propiedades de objeto**, en la ficha **Escala** en **Mínimo** desactiva el botón de opción **Automático** y escribe **55**, en **Máximo** **95**
- h) En el eje Y, mínimo **150** y máximo **190**
- i) Mejora la presentación del gráfico para que quede como el del libro, o mejor.

### Predecir resultados

- Una persona pesa 95 kg. ¿Cuánto medirá?
  - a) En la celda **B17** escribe **95**

- b) En la celda **C17** introduce la fórmula obtenida **=0,917\*B17 + 103,9**. Debes obtener **191,15**

- Una persona mide 177 cm. ¿Cuánto pesará?
  - a) *Arrastra* el  **Controlador de relleno** de la celda **C17** hasta **C18**
  - b) En el menú **Herramientas**, elige **Búsqueda del valor destino...** En la ventana que aparece, escribe en **Celda de fórmula: C18**, **Valor destino: 177**, **Celda variable: B18**. Debes obtener **79,92**



- 48. **Internet.** Abre la web: [www.editorial-bruno.es](http://www.editorial-bruno.es), elige **Matemáticas**, curso y tema.

## Así funciona

### Funciones de estadística bidimensional utilizadas

**PROMEDIO:** media o media aritmética.

**COVAR:** covarianza.

**DESVESTP:** desviación típica.

**COEF.DE.CORREL:** coeficiente de correlación.

## Practica

Elimina las hojas: **Hoja2** y **Hoja3**. Los problemas 49, 50 y 51 son muy parecidos al 47; para hacerlos, en la **Hoja1** se elige en el menú *Contextual* de la pestaña de la hoja **Mover/copiar hoja...**, se selecciona **desplazar a la última posición** y se activa la casilla de verificación **Copia**. Para terminar, se hacen los cambios oportunos. En el gráfico hay que seleccionar el nuevo rango de datos.

- 49. Calcula el centro de gravedad, las desviaciones típicas marginales, la covarianza y el coeficiente de correlación de la siguiente distribución:

<b>Cilindrada (cm<sup>3</sup>)</b>	1000	1200	1400	1600	1600	1800	2000	2000
<b>Velocidad (km/h)</b>	125	130	140	145	150	170	190	195

- a) Representa la nube de puntos y calcula la recta de regresión de **y** sobre **x**, e interpreta los resultados.
- b) Un coche tiene de cilindrada 1 900 cm<sup>3</sup>. ¿Qué velocidad máxima alcanzará?
- c) Un coche tiene una velocidad máxima de 150 km/h. ¿Qué cilindrada tendrá?

- 50. Calcula el centro de gravedad, las desviaciones típicas marginales, la covarianza y el coeficiente de correlación de la siguiente distribución:

<b>Nº de vendedores: x<sub>i</sub></b>	2	4	5	6	7	9	10
<b>Nº de pedidos: y<sub>i</sub></b>	70	90	110	150	170	190	210

- a) Representa la nube de puntos y calcula la recta de regresión de **y** sobre **x**, e interpreta los resultados.
- b) Si hubiese 12 vendedores, ¿cuántos pedidos se esperarían?
- c) Para obtener 250 pedidos, ¿cuántos vendedores harían falta?

- 51. Calcula el centro de gravedad, las desviaciones típicas marginales, la covarianza y el coeficiente de correlación de la siguiente distribución:

<b>Temperatura (°C): x<sub>i</sub></b>	18	19	20	18	22	21
<b>Presión (mm): y<sub>i</sub></b>	790	800	805	795	820	810

- a) Representa la nube de puntos y calcula la recta de regresión de **y** sobre **x**, e interpreta los resultados.
- b) Para una temperatura de 23 °C, ¿qué presión habrá?
- c) Para una presión de 900 mm, ¿qué temperatura habrá?