

Estadística I

TEMA 2. DISTRIBUCIÓN DE FRECUENCIAS BIDIMENSIONAL



Marta Guijarro Garvi

DEPARTAMENTO DE ECONOMÍA

Este material se publica bajo licencia:

[Creative Commons BY-NC-SA 4.0](https://creativecommons.org/licenses/by-nc-sa/4.0/)



Tema 2: Distribución de frecuencias bidimensional

Índice de contenidos

- Principales conceptos: variable bidimensional, frecuencia absoluta y relativa conjunta.
- Distribución de frecuencias bidimensional: distribuciones marginales y condicionadas.
- Independencia estadística.
- Regresión lineal: covarianza, rectas de regresión y bondad del ajuste.

Tema 2: Distribución de frecuencias bidimensional

Resultados de aprendizaje

- Comprender el concepto de distribución de frecuencias bidimensional.
- Saber obtener las distribuciones marginales y las condicionadas.
- Entender la naturaleza unidimensional de las distribuciones condicionadas.
- Saber calcular medidas de posición y dispersión de distribuciones condicionadas.
- Conocer el concepto de independencia.
- Saber hallar las rectas de regresión y calcular e interpretar la bondad del ajuste.

Principales conceptos: variable bidimensional

Al observar conjuntamente las variables X e Y en las unidades de una población se obtienen pares de datos. Los valores de la **variable bidimensional** (X, Y) son los pares de valores (x_i, y_j) donde x_i ($i = 1, \dots, h$) son los valores de X e y_j ($j = 1, \dots, k$) los valores de Y .

Ejemplo 2.1 Se considera la población formada por los estudiantes matriculados en una asignatura de cierta titulación. A partir de las variables X , edad, en años, e Y , altura, en centímetros, se define la variable bidimensional (X, Y) , cuyos valores son pares en los que la primera y segunda componente se corresponden, respectivamente, con los valores de X e Y obtenidos al observar ambas variables en las unidades de la población.

Principales conceptos: frecuencias absolutas y relativas conjuntas

La frecuencia absoluta o **frecuencia absoluta conjunta** n_{ij} del valor (x_i, y_j) es el número de observaciones iguales a dicho valor.

Se cumple:
$$\sum_{i=1}^h \sum_{j=1}^k n_{ij} = N$$

La frecuencia relativa o **frecuencia relativa conjunta** f_{ij} del valor (x_i, y_j) es la proporción de observaciones iguales a dicho valor.

Se cumple: $f_{ij} = \frac{n_{ij}}{N}$ y, por tanto,
$$\sum_{i=1}^h \sum_{j=1}^k f_{ij} = 1$$

Tema 2. Distribución de frecuencias bidimensional

Distribución de frecuencias bidimensional

La distribución de frecuencias de la variable (X,Y) o **distribución de frecuencias bidimensional** es el conjunto de valores de la variable con sus respectivas frecuencias absolutas o relativas. Se denota por $(x_i, y_j; n_{ij})$ o por $(x_i, y_j; f_{ij})$ según el tipo de frecuencia utilizado.

La distribución de frecuencias bidimensional se presenta en una tabla de doble entrada llamada *tabla de correlación*:

$X \backslash Y$	y_1	$\cdots$	y_k
x_1	n_{11}	$\cdots$	n_{1k}
$\vdots$	$\vdots$		$\vdots$
x_h	n_{h1}	$\cdots$	n_{hk}

$X \backslash Y$	y_1	$\cdots$	y_k
x_1	f_{11}	$\cdots$	f_{1k}
$\vdots$	$\vdots$		$\vdots$
x_h	f_{h1}	$\cdots$	f_{hk}

Cuando las distribuciones de frecuencias están agrupadas en intervalos, las marcas de clase desempeñan el papel de representantes del intervalo.

Distribución de frecuencias bidimensional

Ejemplo 2.2 La siguiente tabla recoge la distribución bidimensional del ingreso semanal y del gasto en transporte (ambas en euros) de 20 familias:

Gasto	30-60	60-90
Ingreso		
1000-1500	5	4
1500-2200	3	8

Las frecuencias de la tabla de correlación muestran el comportamiento conjunto de ambas variables. Así, por ejemplo, hay 4 familias con un ingreso semanal entre 1000 y 1500 euros y un gasto en transporte entre 60 y 90 euros ($n_{12} = 4$), mientras que hay 5 familias que tienen un ingreso semanal entre 1000 y 1500 euros y un gasto en transporte entre 30 y 60 euros ($n_{11} = 8$).

A partir de las frecuencias absolutas pueden calcularse las frecuencias relativas. Por ejemplo, el 40 % de las familias tienen un ingreso semanal entre 1500 y 2200 euros y un gasto en transporte entre 60 y 90 euros:

$$f_{22} = \frac{n_{22}}{N} = \frac{8}{20} = 0,4$$

Distribución de frecuencias bidimensional: distribuciones marginales

A partir de la distribución de frecuencias bidimensional se obtienen las **distribuciones de frecuencias marginales** de las variables X e Y .

Distribución marginal de la X : $(x_i; n_{i.})$ o $(x_i; f_{i.})$

X	$n_{i.}$
x_1	$n_{1.}$
$\vdots$	$\vdots$
x_h	$n_{h.}$

$$n_{i.} = \sum_{j=1}^k n_{ij}$$

$$f_{i.} = n_{i.} / N$$

X	$f_{i.}$
x_1	$f_{1.}$
$\vdots$	$\vdots$
x_h	$f_{h.}$

$$f_{i.} = \sum_{j=1}^k f_{ij}$$

Distribución marginal de la Y : $(y_j; n_{.j})$ o $(y_j; f_{.j})$

Y	$n_{.j}$
y_1	$n_{.1}$
$\vdots$	$\vdots$
y_k	$n_{.k}$

$$n_{.j} = \sum_{i=1}^h n_{ij}$$

$$f_{.j} = n_{.j} / N$$

Y	$f_{.j}$
y_1	$f_{.1}$
$\vdots$	$\vdots$
y_k	$f_{.k}$

$$f_{.j} = \sum_{i=1}^h f_{ij}$$

Las distribuciones de frecuencias marginales son distribuciones unidimensionales, por lo que resulta aplicable todo lo estudiado en el Tema 1.

Distribución de frecuencias bidimensional: distribuciones marginales

Ejemplo 2.3 (continuación del Ejemplo 2.2) De la tabla del Ejemplo 2.2 se obtienen las distribuciones marginales de cada una de las variables. Así, sumando cada valor de la primera columna con el correspondiente de la segunda, se obtienen las frecuencias de la variable X , ingreso semanal, en euros, recogidas en la segunda columna de la tabla siguiente. Por ejemplo, el número de familias con ingresos semanales entre 1000 y 1500 euros es $n_{1.} = n_{11} + n_{12} = 5 + 4 = 9$. La distribución de frecuencias de la variable X es, por tanto:

La tercera columna de la tabla recoge las frecuencias relativas de esta distribución marginal.

Ingreso	$n_{i.}$	$f_{i.}$
1000-1500	9	0,45
1500-2200	11	0,55

Análogamente, sumando cada frecuencia absoluta conjunta de la primera fila de la tabla del Ejemplo 2.2 con la correspondiente de la segunda, se obtienen las frecuencias de la variable Y , gasto semanal en transporte, en euros. Así, por ejemplo, el número de familias con un gasto semanal en transporte entre 60 y 90 euros es $n_{.2} = n_{12} + n_{22} = 4 + 8 = 12$. En consecuencia, las distribuciones de frecuencias de la variable Y es:

La tercera columna de la tabla muestra las frecuencias relativas de esta distribución marginal.

Gasto	$n_{.j}$	$f_{.j}$
30-60	8	0,4
60-90	12	0,6

Distribución de frecuencias bidimensional: distribuciones marginales

Ejemplo 2.3 (continuación) Para calcular, por ejemplo, el gasto semanal en transporte más frecuente, esto es, la moda de la distribución marginal de Y se calcula, en primer lugar, el intervalo modal. Como los intervalos tienen la misma amplitud, el intervalo modal es 60-90, ya que es el de mayor frecuencia.

Gasto	$n_{\cdot j}$
30-60	8
60-90	12

En consecuencia, la moda es:

$$Mo_Y = L_{j-1} + \frac{n_{\cdot j+1}}{n_{\cdot j-1} + n_{\cdot j+1}} \cdot c = 60 + \frac{0}{8 + 0} \cdot 30 = 60$$

Para hallar, por ejemplo, el mínimo ingreso de la mitad de las familias con mayores ingresos, es decir, la mediana de la distribución marginal de X , se halla el intervalo mediano. Así, se compara $0,5 \cdot N = 10$ con las frecuencias absolutas acumuladas (tercera columna de la tabla). Como $N_{2\cdot} = 20 > 10$, entonces, el intervalo mediano es 1500-2200.

Ingreso	$n_{i\cdot}$	$N_{i\cdot}$
1000-1500	9	9
1500-2200	11	20

En definitiva, la mediana es:

$$Me_X = L_{i-1} + \frac{0,5 \cdot N - N_{i-1\cdot}}{n_{i\cdot}} \cdot c_i = 1500 + \frac{10 - 9}{11} \cdot 700 = 1563,64$$

Distribución de frecuencias bidimensional: distribuciones condicionadas

A partir de la distribución de frecuencias bidimensional se obtienen las **distribuciones de frecuencias condicionadas** de la variable X por los valores de Y , y de la variable Y por los valores de X .

Distribución de X condicionada por el valor y_j de Y

Notación: $(x_i / Y = y_j; n_{i/j})$ o bien $(x_i / Y = y_j; f_{i/j})$

$x_i / Y = y_j$	$n_{i/j}$	$f_{i/j}$
x_1	$n_{1/j}$	$f_{1/j}$
$\vdots$	$\vdots$	$\vdots$
x_h	$n_{h/j}$	$f_{h/j}$

$$n_{i/j} = n_{ij} \quad f_{i/j} = n_{i/j} / n_{\cdot j} = n_{ij} / n_{\cdot j}$$

Distribución de Y condicionada por el valor x_i de X

Notación: $(y_j / X = x_i; n_{j/i})$ o bien $(y_j / X = x_i; f_{j/i})$

$y_j / X = x_i$	$n_{j/i}$	$f_{j/i}$
y_1	$n_{1/i}$	$f_{1/i}$
$\vdots$	$\vdots$	$\vdots$
y_k	$n_{k/i}$	$f_{k/i}$

$$n_{j/i} = n_{ij} \quad f_{j/i} = n_{j/i} / n_{i\cdot} = n_{ij} / n_{i\cdot}$$

Cuando las frecuencias de la distribución bidimensional son relativas, entonces:

$$f_{i/j} = f_{ij} / f_{\cdot j}$$

$$f_{j/i} = f_{ij} / f_{i\cdot}$$

Distribución de frecuencias bidimensional: distribuciones condicionadas

Las distribuciones de frecuencias condicionadas son distribuciones unidimensionales, por lo que resulta aplicable todo lo estudiado en el Tema 1.

Así, la media y la varianza de la distribución condicionada $(x_i / Y = y_j; f_{i/j})$ son, respectivamente,

$$\bar{x} / (Y = y_j) = \sum_{i=1}^h x_i \cdot f_{i/j} \qquad S_{X/Y=y_j}^2 = \sum_{i=1}^h [x_i - \bar{x} / (Y = y_j)]^2 f_{i/j}$$

o, equivalentemente, cuando las frecuencias son absolutas $(x_i / Y = y_j; n_{i/j})$:

$$\bar{x} / (Y = y_j) = \frac{1}{n_{\cdot j}} \sum_{i=1}^h x_i \cdot n_{ij} \qquad S_{X/Y=y_j}^2 = \frac{1}{n_{\cdot j}} \sum_{i=1}^h [x_i - \bar{x} / (Y = y_j)]^2 n_{ij}$$

Y, la media y la varianza de la distribución condicionada $(y_j / X = x_i; f_{j/i})$ son, respectivamente,

$$\bar{y} / (X = x_i) = \sum_{j=1}^k y_j \cdot f_{j/i} \qquad S_{Y/X=x_i}^2 = \sum_{j=1}^k [y_j - \bar{y} / (X = x_i)]^2 f_{j/i}$$

o, equivalentemente, cuando las frecuencias son absolutas $(y_j / X = x_i; n_{j/i})$:

$$\bar{y} / (X = x_i) = \frac{1}{n_{i \cdot}} \sum_{j=1}^k y_j \cdot n_{ij} \qquad S_{Y/X=x_i}^2 = \frac{1}{n_{i \cdot}} \sum_{j=1}^k [y_j - \bar{y} / (X = x_i)]^2 n_{ij}$$

Distribución de frecuencias bidimensional: distribuciones condicionadas

Ejemplo 2.4 (continuación del Ejemplo 2.2) A partir de la información proporcionada por la tabla de correlación del Ejemplo 2.2, se desea conocer el gasto medio semanal en transporte de las familias con ingresos comprendidos entre 1500 y 2200 euros semanales, así como una medida de la dispersión de la distribución.

La siguiente tabla muestra la distribución del gasto semanal en transporte condicionada por un valor del ingreso igual a 1850 euros, marca de clase del intervalo, esto es, la distribución $(y_j / X = x_2; n_{j/2})$:

Aunque no son necesarias para resolver el problema, en la cuarta columna de la tabla se muestran las frecuencias relativas de la distribución condicionada. En la segunda columna se recogen las marcas de clase.

Gasto	y_j	$n_{j/2}$	$f_{j/2}$
30-60	45	3	0,27
60-90	75	8	0,73

La media y la varianza de la distribución condicionada son, respectivamente,

$$\bar{y} / (X = x_2) = \frac{1}{n_{2.}} \sum_{j=1}^k y_j \cdot n_{2j} = \frac{45 \cdot 3 + 75 \cdot 8}{11} = 66,82$$

$$S_{Y/X=x_2}^2 = \frac{1}{n_{2.}} \sum_{j=1}^k [y_j - \bar{y} / (X = x_2)]^2 n_{2j} = \frac{1}{n_{2.}} \sum_{j=1}^k y_j^2 \cdot n_{2j} - [\bar{y} / (X = x_2)]^2 = \frac{45^2 \cdot 3 + 75^2 \cdot 8}{11} - 66,82^2 = 178,27$$

Distribución de frecuencias bidimensional: distribuciones condicionadas

Ejemplo 2.4 (continuación) Para calcular, por ejemplo, el gasto semanal en transporte más frecuente de las familias con ingreso semanal comprendido entre 1500 y 2200 euros, esto es, la moda de la distribución se halla, en primer lugar, el intervalo modal. Como los intervalos tienen la misma amplitud, el intervalo modal es 60-90, intervalo con mayor frecuencia.

Gasto	$n_{j/2}$	$N_{j/2}$
30-60	3	3
60-90	8	11

En consecuencia, la moda de la distribución condicionada es:

$$Mo_{Y/X=x_2} = L_{j-1} + \frac{n_{j+1/2}}{n_{j-1/2} + n_{j+1/2}} \cdot c = 60 + \frac{0}{3+0} \cdot 30 = 60$$

Asimismo, para calcular el mínimo gasto semanal en transporte de la mitad de las familias con mayores gastos dentro de las que tienen un ingreso semanal comprendido entre 1500 y 2200 euros, es decir, la mediana de esta distribución condicionada es necesario determinar el intervalo mediano. Para ello, se compara la tercera columna de la tabla anterior con $0,5 \cdot n_2 = 7$, resultando que el primer intervalo cuya frecuencia absoluta acumulada es mayor que 7 es 60-90: $N_{2/2} = 11 > 7$.

En definitiva, la mediana de la distribución condicionada es:

$$Me_{Y/X=x_2} = L_{j-1} + \frac{0,5 \cdot n_2 - N_{j-1/2}}{n_{j/2}} \cdot c = 60 + \frac{7-3}{8} \cdot 30 = 75$$

Independencia estadística

Dada una distribución de frecuencias bidimensional $(x_i, y_j; f_{ij})$, se dice que las variables X e Y son **independientes** cuando, para cualesquiera i y j , se cumple:

$$f_{ij} = f_{i\cdot} \cdot f_{\cdot j}$$

La relación entre frecuencias absolutas y relativas permite expresar la definición anterior usando frecuencias absolutas. Así, dada una distribución de frecuencias bidimensional $(x_i, y_j; n_{ij})$, las variables X e Y son independientes cuando, para cualesquiera i y j , se cumple:

$$n_{ij} = \frac{n_{i\cdot} \cdot n_{\cdot j}}{N}$$

Sea una distribución de frecuencias bidimensional $(x_i, y_j; f_{ij})$. La condición necesaria y suficiente para que las variables X e Y sean independientes es que, para cualesquiera i y j ,

$$f_{i/j} = f_{i\cdot}$$

y

$$f_{j/i} = f_{\cdot j}$$

Independencia estadística

Ejemplo 2.5 La distribución conjunta del salario mensual, en miles de euros, y del gasto semanal en comida precocinada, en euros, de un grupo de familias, así como las distribuciones marginales, se recogen en la tabla siguiente:

Gasto Salario	0-100	100-150	150-300	
2-2,5	60	90	30	180
2,5-3,5	40	60	20	120
	100	150	50	300

Como:

$$\frac{60}{100} = \frac{90}{150} = \frac{30}{50} = \frac{180}{300} \quad y \quad \frac{40}{100} = \frac{60}{150} = \frac{20}{50} = \frac{120}{300}, \text{ entonces, las variables son independientes.}$$

Para estudiar la independencia de estas variables, comprobamos la proporcionalidad de filas y columnas de la tabla de correlación, esto es, comprobamos si se cumple o no las sucesiones de igualdades siguientes¹:

$$f_{1/1} = f_{1/2} = f_{1/3} = f_1. \quad y \quad f_{2/1} = f_{2/2} = f_{2/3} = f_2.$$

O, equivalentemente:

$$\frac{n_{11}}{n_{.1}} = \frac{n_{12}}{n_{.2}} = \frac{n_{13}}{n_{.3}} = \frac{n_{1.}}{N} \quad y \quad \frac{n_{21}}{n_{.1}} = \frac{n_{22}}{n_{.2}} = \frac{n_{23}}{n_{.3}} = \frac{n_{2.}}{N}$$

¹Si se cumple la proporcionalidad entre los elementos de las filas, también se cumple entre los de las columnas.

Regresión lineal: covarianza

Sea una distribución de frecuencias bidimensional $(x_i, y_j; f_{ij})$. La **covarianza** entre X e Y es

$$S_{X,Y} = \sum_{i=1}^h \sum_{j=1}^k (x_i - \bar{x}) \cdot (y_j - \bar{y}) \cdot f_{ij} = \frac{1}{N} \sum_{i=1}^h \sum_{j=1}^k (x_i - \bar{x}) \cdot (y_j - \bar{y}) \cdot n_{ij}$$

según las frecuencias sean relativas o absolutas.

Se cumple:

$$S_{X,Y} = \sum_{i=1}^h \sum_{j=1}^k x_i \cdot y_j \cdot f_{ij} - \bar{x} \cdot \bar{y} = \frac{1}{N} \sum_{i=1}^h \sum_{j=1}^k x_i \cdot y_j \cdot n_{ij} - \bar{x} \cdot \bar{y}$$

Regresión lineal: rectas de regresión

La recta que mejor explica¹ la variable Y en función de la variable X es la **recta de regresión de Y sobre X** :

$$y - \bar{y} = \frac{S_{X,Y}}{S_X^2} (x - \bar{x})$$

Análogamente, la recta que mejor explica la variable X en función de la variable Y es la **recta de regresión de X sobre Y** :

$$x - \bar{x} = \frac{S_{X,Y}}{S_Y^2} (y - \bar{y})$$

¹El criterio utilizado para obtener las rectas de regresión es el de los mínimos cuadrados.

Los coeficientes de regresión de las rectas de regresión son, respectivamente,

$$b_{Y/X} = S_{X,Y} / S_X^2 \text{ y } b_{X/Y} = S_{X,Y} / S_Y^2.$$

Regresión lineal: bondad de ajuste

La evaluación de la calidad de la explicación lineal obtenida mediante las rectas de regresión, esto es, el análisis de la **bondad del ajuste** se realiza con el **coeficiente de correlación**:

$$r = \frac{S_{X,Y}}{S_X \cdot S_Y} \quad (-1 \leq r \leq 1)$$

y con su cuadrado, el **coeficiente de determinación lineal**:

$$r^2 = \frac{S_{X,Y}^2}{S_X^2 \cdot S_Y^2} \quad (0 \leq r^2 \leq 1)$$

- Si r es igual a -1 o a 1 ($r^2 = 1$), el ajuste es perfecto (negativo o positivo, respectivamente), existiendo una relación lineal exacta (negativa o positiva) entre las variables.
- Si r es igual a 0 ($r^2 = 0$), el ajuste es pésimo, no existiendo relación lineal entre las variables. En este caso, las variables están *incorrelacionadas*.
- El signo de las pendientes de las rectas de regresión, de los coeficientes de regresión y del coeficiente de correlación es el signo de la covarianza.
- Si dos variables son independientes están incorrelacionadas, pero el recíproco no es cierto.

Regresión lineal: rectas de regresión y bondad del ajuste

Ejemplo 2.6 La distribución conjunta y las distribuciones marginales de X e Y se presentan en la siguiente tabla:

Las medias y varianzas de X e Y , así como la covarianza entre ellas son:

$X \setminus Y$	20-60	60-70	70-170	
6-10	90	1	0	91
10-14	4	30	1	35
14-18	1	0	73	74
	95	31	74	200

$$\bar{x} = \frac{1}{N} \sum_{i=1}^h x_i \cdot n_{i.} = 11,66$$

$$S_X^2 = \frac{1}{N} \sum_{i=1}^h (x_i - \bar{x})^2 \cdot n_{i.} = 13,08$$

$$\bar{y} = \frac{1}{N} \sum_{j=1}^k y_j \cdot n_{.j} = 73,47$$

$$S_Y^2 = \frac{1}{N} \sum_{j=1}^k (y_j - \bar{y})^2 \cdot n_{.j} = 1344,30$$

$$S_{XY} = \frac{1}{N} \sum_{i=1}^h \sum_{j=1}^k (x_i - \bar{x}) \cdot (y_j - \bar{y}) \cdot n_{ij} = 127,68$$

En definitiva, las rectas de regresión de Y sobre X y de X sobre Y son, respectivamente,

$$y = 9,76 \cdot x - 40,31 \quad \text{y} \quad x = 0,09 \cdot y + 4,68.$$

Además, los coeficientes de correlación y determinación lineal son:

$$r = \frac{S_{XY}}{S_X \cdot S_Y} = 0,96 \quad \text{y} \quad r^2 = 0,93.$$

Como los coeficientes de bondad del ajuste son muy próximos a uno, existe una fuerte relación lineal entre las variables. Además, como la covarianza es positiva, esta relación es creciente.

Regresión lineal: rectas de regresión y bondad del ajuste

Ejemplo 2.7 La distribución conjunta de X e Y se presenta en la siguiente tabla:

En este ejemplo, cada observación de la variable X se corresponde con una observación de la variable Y (frecuencias unitarias). Esto dota de mayor sencillez a los cálculos.

X	4,5	4,6	4,7	4,8	4,9	5	5,1	5,2	5,3	5,4
Y	80	79	72	65	70	64	61	50	45	43

En consecuencia, las medias y varianzas de X e Y , así como la covarianza entre ellas son:

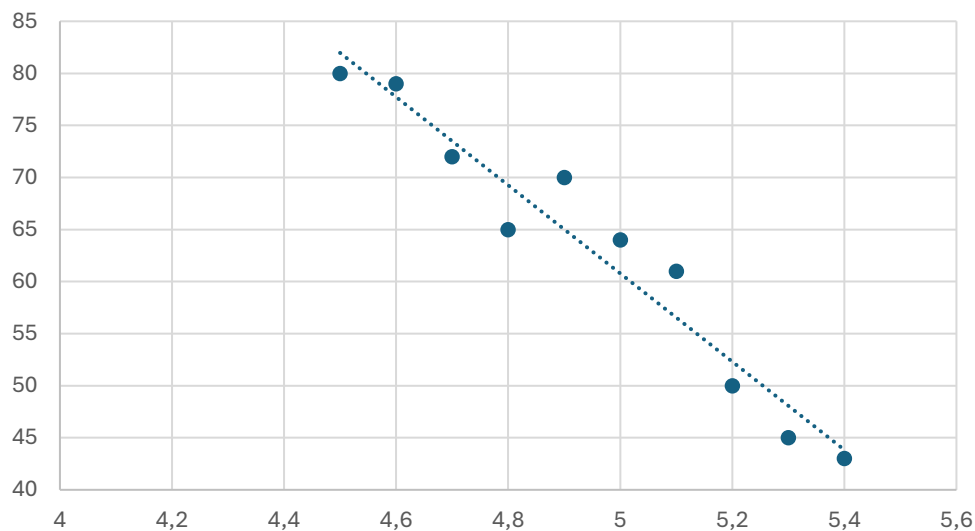
$$\begin{aligned}\bar{x} &= \frac{1}{N} \sum_{i=1}^N x_i = 4,95 & \bar{y} &= \frac{1}{N} \sum_{i=1}^N y_i = 62,90 & S_X^2 &= \frac{1}{N} \sum_{i=1}^N (x_i - \bar{x})^2 = 0,08 & S_Y^2 &= \frac{1}{N} \sum_{i=1}^N (y_i - \bar{y})^2 = 157,69 \\ S_{XY} &= \frac{1}{N} \sum_{i=1}^N (x_i - \bar{x}) \cdot (y_i - \bar{y}) = -3,49\end{aligned}$$

En definitiva, las rectas de regresión de Y sobre X y de X sobre Y son, respectivamente,

$$y = -42,36 \cdot x + 272,6 \quad \text{y} \quad x = -0,02 \cdot y + 6,34.$$

Regresión lineal: rectas de regresión y bondad del ajuste

Ejemplo 2.7 (continuación) Como se ve en la gráfica siguiente (*diagrama de dispersión*), existe una fuerte correlación decreciente entre las variables



La información mostrada en la gráfica se corrobora con los valores de los coeficientes de correlación y determinación lineal:

$$r = \frac{S_{XY}}{S_X \cdot S_Y} = -0,97 \quad \text{y} \quad r^2 = 0,94.$$

En definitiva, como los coeficientes de bondad del ajuste son muy cercanos a la unidad, existe una fuerte relación lineal entre las variables. Además, el signo negativo de la covarianza y, por tanto, del coeficiente de correlación indican que esta relación es decreciente.