2.- Variables estadísticas bidimensionales.

2.1 Representaciones numéricas

2.2 Distribuciones marginales y condicionadas

2.3 Independencia de variables estadísticas

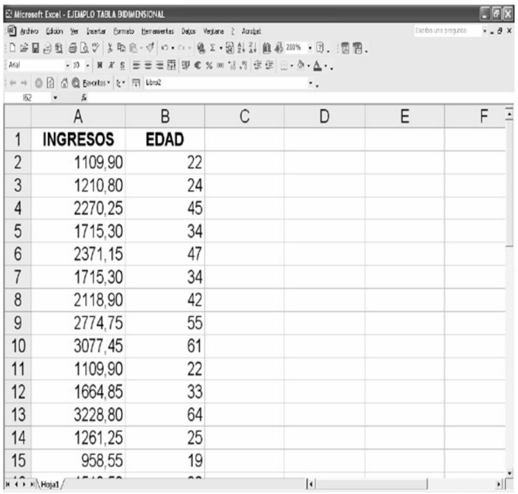
2.4 Covarianza. Coeficiente de correlación lineal

2.5 Recta de regresión de mínimos cuadrados

1

REPRESENTACIONES NUMÉRICAS. DOS COLUMNAS.

x_i	y_i
x_1	y_1
x_2	y_2
...	...
x_n	y_n



	A	B	C	D	E	F
1	INGRESOS	EDAD				
2	1109,90	22				
3	1210,80	24				
4	2270,25	45				
5	1715,30	34				
6	2371,15	47				
7	1715,30	34				
8	2118,90	42				
9	2774,75	55				
10	3077,45	61				
11	1109,90	22				
12	1664,85	33				
13	3228,80	64				
14	1261,25	25				
15	958,55	19				

2

REPRESENTACIONES NUMÉRICAS. TABLA DE CONTINGENCIA O DE CORRELACIÓN.

La distribución según salarios,Y, y edades, X, de un grupo de 100 jóvenes es

X/Y	**500-1.000**	**1.000-1.500**	**1.500-2.000**	$n_{i\bullet}$
20	10	3	2	15
21	5	15	5	25
22	2	20	15	37
23	0	13	10	23
$n_{\bullet j}$	17	51	32	100

REPRESENTACIONES NUMÉRICAS. TABLA DE CONTINGENCIA O DE CORRELACIÓN.

X/Y	B_1 $L_0^*-L_1^*$ y_1	...	B_j $L_{j-1}^*-L_j^*$ y_j	...	B_p $L_{p-1}^*-L_p^*$ y_p	$n_{i\bullet}$
A_1 L_0-L_1 x_1	n_{11}	...	n_{1j}	...	n_{1p}	$n_{1\bullet}$
...	...	...	...	...	...	...
A_i $L_{i-1}-L_i$ x_i	n_{i1}	...	n_{ij}	...	n_{ip}	$n_{i\bullet}$
...	...	...	...	...	...	...
A_k $L_{k-1}-L_k$ x_k	n_{k1}	...	n_{kj}	...	n_{kp}	$n_{k\bullet}$
$n_{\bullet j}$	$n_{\bullet 1}$	...	$n_{\bullet j}$	...	$n_{\bullet p}$	n

n_{ij} $\quad n=\sum_{i=1}^{k}\sum_{j=1}^{p}n_{ij}$ $\quad f_{ij}=\frac{n_{ij}}{n}$ $\quad 1=\sum_{i=1}^{k}\sum_{j=1}^{p}f_{ij}$

Distribuciones marginales.

La distribución marginal de las edades en el ejemplo

x_i	$n_{i\bullet}$
20	15
21	25
22	37
23	23
	100

La distribución marginal de los salarios en el ejemplo

y_j	$n_{\bullet j}$
500-1.000	17
1.000-1.500	51
1.500-2.000	32
	100

$$n_{i\bullet} = \sum_{j=1}^{p} n_{ij} \qquad n_{\bullet j} = \sum_{i=1}^{k} n_{ij} \qquad n = \sum_{i=1}^{k} n_{i\bullet} = \sum_{j=1}^{p} n_{\bullet j}$$

$$f_{i\bullet} = \frac{n_{i\bullet}}{n} \qquad f_{\bullet j} = \frac{n_{\bullet j}}{n} \qquad \sum_{i=1}^{k} f_{i\bullet} = \sum_{j=1}^{p} f_{\bullet j} = 1$$

5

Distribuciones Marginales

x_i
x_1
x_2
...
x_n

6

Distribuciones condicionadas

Distribución de los salarios condicionada a que el joven tiene 22 años

y_j	n_j
500-1.000	2
1.000-1.500	20
1.500-2.000	15
	37

Distribución de las edades de los jóvenes que tienen un salario mayor que 1.000€ pero menor que 1.500€

x_i	n_i
20	3
21	15
22	20
23	13
	51

INDEPENDENCIA ESTADÍSTICA

Comprobar si hay independencia estadística entre el salario, X, y el sexo, Y, en el grupo de 49 trabajadores representados en la siguiente tabla

X/Y	hombre	mujer
500-1.000	20	8
1.000-1.500	10	4
1.500-2.000	5	2

X/Y	hombre	mujer	TOTAL
500-1.000	$\frac{20}{35}=0,5714$	$\frac{8}{14}=0,5714$	$\frac{28}{49}=0,5714$
1.000-1.500	$\frac{10}{35}=0,2857$	$\frac{4}{14}=0,2857$	$\frac{14}{49}=0,2857$
1.500-2.000	$\frac{5}{35}=0,1429$	$\frac{2}{14}=0,1429$	$\frac{7}{49}=0,1429$

$$n_{ij} = \frac{n_{i\bullet} n_{\bullet j}}{n}$$

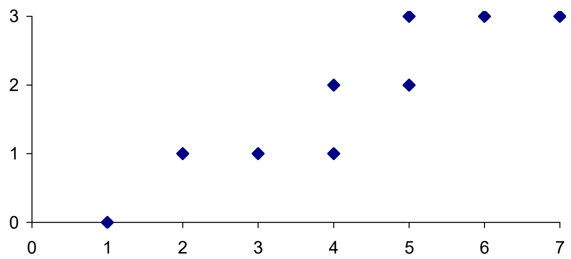
X/Y	hombre	mujer	$n_{i\bullet}$
500-1.000	$20 = \frac{28 \times 35}{49}$	$8 = \frac{28 \times 14}{49}$	28
1.000-1.500	$10 = \frac{14 \times 35}{49}$	$4 = \frac{14 \times 14}{49}$	14
1.500-2.000	$5 = \frac{7 \times 35}{49}$	$2 = \frac{7 \times 14}{49}$	7
$n_{\bullet j}$	35	14	$n = 49$

9

REPRESENTACIONES GRÁFICAS

Nube de puntos o **diagrama de dispersión.**

X	Y
3	1
5	2
4	1
6	3
7	3
5	3
4	2
2	1
1	0



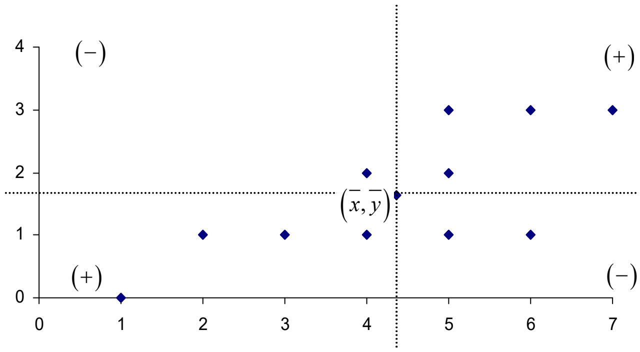
10

COVARIANZA

$$S_{xy}=\frac{1}{n}\sum_{i=1}^{k}\sum_{j=1}^{p}\left(x_i-\bar{x}\right)\left(y_j-\bar{y}\right)n_{ij} \qquad S_{xy}=\frac{1}{n}\sum_{i=1}^{n}\left(x_i-\bar{x}\right)\left(y_i-\bar{y}\right)$$

Calcule la covarianza para la siguiente tabla de datos e interprete su valor de acuerdo a la representación gráfica.

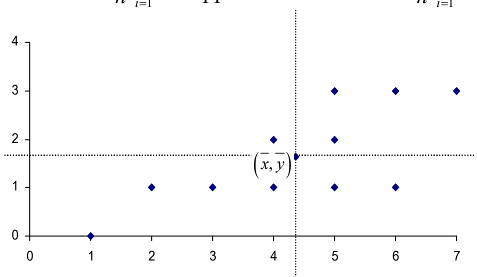
	X	Y
	3	1
	5	2
	4	1
	6	3
	7	3
	5	3
	4	2
	2	1
	1	0
	5	1
	6	1
medias	4,36	1,64



11

	x_i	y_i	$x_i-\bar{x}$	$y_i-\bar{y}$	$\left(x_i-\bar{x}\right)\left(y_i-\bar{y}\right)$
	3	1	-1,36	-0,64	0,87
	5	2	0,64	0,36	0,23
	4	1	-0,36	-0,64	0,23
	6	3	1,64	1,36	2,23
	7	3	2,64	1,36	3,60
	5	3	0,64	1,36	0,87
	4	**2**	-0,36	0,36	**-0,13**
	2	1	-2,36	-0,64	1,50
	1	0	-3,36	-1,64	5,50
	5	**1**	0,64	-0,64	**-0,40**
	6	**1**	1,64	-0,64	**-1,04**
totales	**48**	**18**			**13,46**

$$\bar{x}=\frac{1}{n}\sum_{i=1}^{n}x_i=\frac{48}{11}=4,36 \qquad \bar{y}=\frac{1}{n}\sum_{i=1}^{n}y_i=\frac{18}{11}=1,64 \qquad S_{xy}=\frac{1}{n}\sum_{i=1}^{n}\left(x_i-\bar{x}\right)\left(y_i-\bar{y}\right)=\frac{13,46}{11}=1,22$$



12

$$S_{xy} = \frac{1}{n}\sum_{i=1}^{n}\left(x_i - \bar{x}\right)\left(y_i - \bar{y}\right) = \frac{1}{n}\sum_{i=1}^{n}\left(x_i y_i - x_i \bar{y} - \bar{x} y_i + \bar{x}\bar{y}\right) =$$

$$= \left(\frac{1}{n}\sum_{i=1}^{n} x_i y_i\right) - \left(\bar{y}\frac{1}{n}\sum_{i=1}^{n} x_i\right) - \left(\bar{x}\frac{1}{n}\sum_{i=1}^{n} y_i\right) + \left(\bar{x}\bar{y}\frac{1}{n}\sum_{i=1}^{n} 1\right) = \left(\frac{1}{n}\sum_{i=1}^{n} x_i y_i\right) - \bar{x}\bar{y} - \bar{x}\bar{y} + \bar{x}\bar{y} = \left(\frac{1}{n}\sum_{i=1}^{n} x_i y_i\right) - \bar{x}\bar{y}$$

$$S_{xy} = \left(\frac{1}{n}\sum_{i=1}^{n} x_i y_i\right) - \bar{x}\,\bar{y}$$ *(datos en dos columnas)*

$$S_{xy} = \left(\frac{1}{n}\sum_{i=1}^{k}\sum_{j=1}^{p} x_i y_j n_{ij}\right) - \bar{x}\,\bar{y}$$ *(datos en tablas de contingencia)*

13

Obtenga la covarianza ulilizando esta última expresión

x_i	y_i	$x_i y_i$
3	1	3
5	2	10
4	1	4
6	3	18
7	3	21
5	3	15
4	2	8
2	1	2
1	0	0
5	1	5
6	1	6
48	**18**	**92**

$$n = 11 \quad \bar{x} = \frac{48}{11} = 4{,}36 \quad \bar{y} = \frac{18}{11} = 1{,}64$$

$$S_{xy} = \frac{1}{n}\sum_{i=1}^{n} x_i y_i - \bar{x}\,\bar{y} = \frac{92}{11} - (4{,}36 \times 1{,}64) = 1{,}21$$

14

Si realizamos un *cambio de origen y escala* sobre las variables originales X e Y

$$X^* = eX + c \qquad Y^* = hY + d$$

la covarianza de las nuevas variables es

$$S_{x^*y^*} = ehS_{xy}$$

$$S_{x^*y^*} = \frac{1}{n}\sum_{i=1}^{k}\sum_{j=1}^{p}\left(ex_i + c - e\overline{x} - c\right)\left(hy_j + d - h\overline{y} - d\right)n_{ij} = \frac{1}{n}\sum_{i=1}^{k}\sum_{j=1}^{p}\left(ex_i - e\overline{x}\right)\left(hy_j - h\overline{y}\right)n_{ij} =$$

$$= \frac{1}{n}\sum_{i=1}^{k}\sum_{j=1}^{p} e\left(x_i - \overline{x}\right)h\left(y_j - \overline{y}\right)n_{ij} = eh\frac{1}{n}\sum_{i=1}^{k}\sum_{j=1}^{p}\left(x_i - \overline{x}\right)\left(y_j - \overline{y}\right)n_{ij} = ehS_{xy}$$

15

En caso de ***independencia*** estadística $\Rightarrow$ $\boxed{S_{xy} = 0}$

(sin embargo, $S_{xy} = 0$ no implica necesariamente que se dé la condición de independencia estadística)

Para demostrarlo nos apoyamos en la caracterización de la independencia estadística

$$n_{ij} = \frac{n_{i\bullet}n_{\bullet j}}{n}$$

$$S_{xy} = \left(\frac{1}{n}\sum_{i=1}^{k}\sum_{j=1}^{p} x_i y_j n_{ij}\right) - \overline{x}\,\overline{y} = \left(\frac{1}{n}\sum_{i=1}^{k}\sum_{j=1}^{p} x_i y_j \frac{n_{i\bullet}n_{\bullet j}}{n}\right) - \overline{x}\,\overline{y} = \left(\frac{1}{n}\sum_{i=1}^{k} x_i n_{i\bullet}\sum_{j=1}^{p} y_j \frac{n_{\bullet j}}{n}\right) - \overline{x}\,\overline{y} =$$

$$= \left(\frac{1}{n}\sum_{i=1}^{k} x_i n_{i\bullet}\,\frac{1}{n}\sum_{j=1}^{p} y_j n_{\bullet j}\right) - \overline{x}\,\overline{y} = \left(\overline{x}\,\overline{y}\right) - \overline{x}\,\overline{y} = 0$$

16

X/Y	5	15	25
10	4	6	4
20	3	9	3
40	2	12	2

$x_i y_j n_{ij}$	**5**	**15**	**25**	$n_{i\bullet}$	$x_i\, n_{i\bullet}$
10	200	900	1000	14	140
20	300	2700	1500	15	300
40	400	7200	2000	16	640
$n_{\bullet j}$	9	27	9	n=45	1080
$y_j n_{\bullet j}$	45	405	225	675	

$$\sum_{i=1}^{k}\sum_{j=1}^{p} x_i y_j n_{ij} = 200+900+\ldots+7200+2000=16200$$

$$\bar{x}=\frac{1}{n}\sum_{i=1}^{k} x_i n_{i\bullet}=\frac{1080}{45}=24 \qquad \bar{y}=\frac{1}{n}\sum_{j=1}^{p} y_j n_{\bullet j}=\frac{675}{45}=15$$

$$S_{xy}=\left(\frac{1}{n}\sum_{i=1}^{k}\sum_{j=1}^{p} x_i y_j n_{ij}\right)-\bar{x}\,\bar{y}=\frac{16200}{45}-\left(24\times 15\right)=360-360=0$$

17

X/Y	5	15	25	$n_{i\bullet}$
10	4	6	4	14
20	3	9	3	15
40	2	12	2	16
$n_{\bullet j}$	9	27	9	n=45

$$n_{ij}=\frac{n_{i\bullet} n_{\bullet j}}{n} \quad \forall i,j$$

$$n_{11}\neq\frac{n_{1\bullet}\, n_{\bullet 1}}{n} \qquad 4\neq\frac{14\times 9}{45}=2,8$$

18

Coeficiente de correlación lineal.

$$r_{xy} = \frac{S_{xy}}{S_x S_y} \qquad \left(-1 \leq r_{xy} \leq 1\right)$$

El cambio de origen y escala no afecta al valor de este coeficiente

$$X^* = eX + c \qquad Y^* = hY + d \qquad S_{x^*y^*} = ehS_{xy} \qquad S_{x^*} = eS_x \qquad S_{y^*} = hS_y$$

$$r_{x^*y^*} = \frac{S_{x^*y^*}}{S_{x^*}S_{y^*}} = \frac{ehS_{xy}}{eS_x hS_y} = \frac{S_{xy}}{S_x S_y} = r_{xy}$$

19

Calcule el coeficiente de correlación lineal

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
3	1	3	9	1
5	2	10	25	4
4	1	4	16	1
6	3	18	36	9
7	3	21	49	9
5	3	15	25	9
4	2	8	16	4
2	1	2	4	1
1	0	0	1	0
5	1	5	25	1
6	1	6	36	1
48	**18**	**92**	**242**	**40**

$$n = 11 \quad \bar{x} = \frac{48}{11} = 4,364 \quad \bar{y} = \frac{18}{11} = 1,636 \qquad S_{xy} = \frac{1}{n}\sum_{i=1}^{n} x_i y_i - \bar{x}\,\bar{y} = \frac{92}{11} - (4,364 \times 1,636) = 1,22$$

$$S_x^2 = \left(\frac{1}{n}\sum_{i=1}^{n} x_i^2\right) - \bar{x}^2 = \frac{242}{11} - 4,364^2 = 2,9555 \quad \Rightarrow \quad S_x = 1,72$$

$$S_y^2 = \left(\frac{1}{n}\sum_{i=1}^{n} y_i^2\right) - \bar{y}^2 = \frac{40}{11} - 1,636^2 = 0,9599 \quad \Rightarrow \quad S_y = 0,98$$

$$r_{xy} = \frac{S_{xy}}{S_x S_y} = \frac{1,22}{1,72 \times 0,98} = 0,7238$$

20

X/Y	**500-1.000** **750**	**1.000-1.500** **1250**	**1.500-2.000** **1750**	$n_{i\bullet}$	$x_i n_{i\bullet}$	$x_i^2 n_{i\bullet}$
20	10 $\boxed{150000}$	3 $\boxed{75000}$	2 $\boxed{70000}$	15	300	6000
21	5 $\boxed{78750}$	15 $\boxed{393750}$	5 $\boxed{183750}$	25	525	11025
22	2 $\boxed{33000}$	20 $\boxed{550000}$	15 $\boxed{577500}$	37	814	17908
23	0 $\boxed{0}$	13 $\boxed{373750}$	10 $\boxed{402500}$	23	529	12167
$n_{\bullet j}$	17	51	32	n=100	2168	47100
$y_j n_{\bullet j}$	12750	63750	56000	132500		
$y_j^2 n_{\bullet j}$	9562500	79687500	98000000	187250000		

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{k} x_i n_{i\bullet} = \frac{2168}{100} = 21,68 \qquad \bar{y} = \frac{1}{n}\sum_{j=1}^{p} y_j n_{\bullet j} = \frac{132500}{100} = 1325$$

$$S_{xy} = \left(\frac{1}{n}\sum_{i=1}^{k}\sum_{j=1}^{p} x_i y_j n_{ij}\right) - \bar{x}\,\bar{y} = \frac{150000+\ldots+402500}{100} - \left(21,68\times 1325\right) = \frac{2888000}{100} - 28726 = 154$$

$$S_x^2 = \frac{1}{n}\sum_{i=1}^{k}\left(x_i - \bar{x}\right)^2 n_{i\bullet} = \frac{1}{n}\sum_{i=1}^{k} x_i^2 n_{i\bullet} - \bar{x}^2 = \frac{47100}{100} - 21,68^2 = 0,9776 \qquad \Rightarrow \quad S_x = 0,9887$$

$$S_y^2 = \frac{1}{n}\sum_{j=1}^{p}\left(y_j - \bar{y}\right)^2 n_{\bullet j} = \frac{1}{n}\sum_{j=1}^{p} y_j^2 n_{\bullet j} - \bar{y}^2 = \frac{187250000}{100} - 1325^2 = 116875 \qquad \Rightarrow \quad S_y = 341,87$$

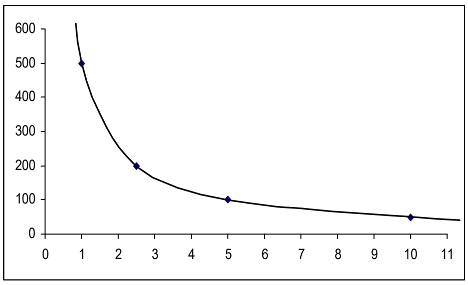
$$r_{xy} = \frac{S_{xy}}{S_x S_y} = \frac{154}{0,9887\times 341,87} = 0,4556$$

21

RECTA DE REGRESIÓN DE MINIMOS CUADRADOS

t (horas)	v (Km./h)
5	100
10	50
2.5	200
1	500

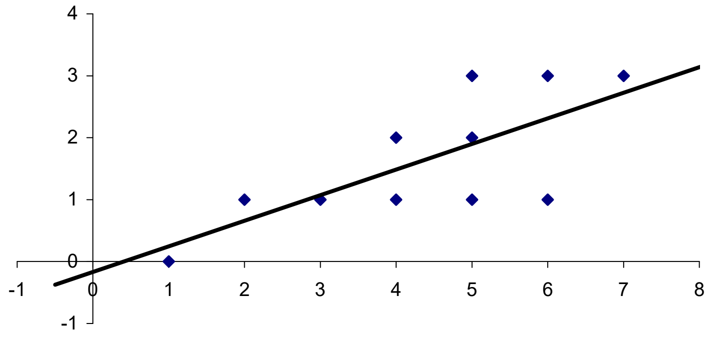
$$vt = 500 \qquad v = \frac{500}{t}$$



22

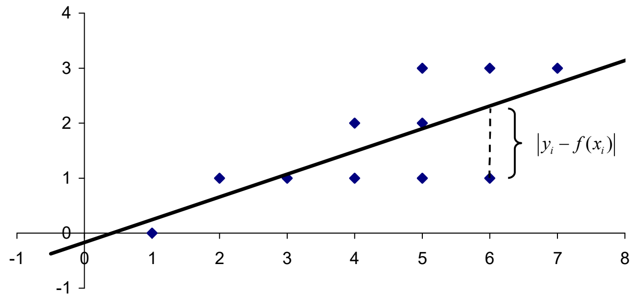
dependencia estadística **dependencia funcional**

X	Y
3	1
5	2
4	1
6	3
7	3
5	3
4	2
2	1
1	0
5	1
6	1



23

AJUSTE POR MÍNIMOS CUADRADOS



$$\sum_{i=1}^{n} \left(y_i - f(x_i)\right)^2 \qquad\qquad f(x_i) = a + bx_i$$

$$b = \frac{S_{xy}}{S_x^2} \qquad\qquad a = \overline{y} - b\overline{x} = \overline{y} - \frac{S_{xy}}{S_x^2}\overline{x}$$

24

$$y = a + bx \qquad y = \overline{y} - b\overline{x} + bx \qquad y - \overline{y} = bx - b\overline{x} \qquad y - \overline{y} = b\left(x - \overline{x}\right)$$

$$y - \overline{y} = \frac{S_{xy}}{S_x^2}\left(x - \overline{x}\right)$$

$$r = \frac{S_{xy}}{S_x S_y} \qquad \frac{S_{xy}}{S_x^2} = \frac{S_{xy}}{S_x S_x}\frac{S_y}{S_y} = \frac{S_{xy}}{S_x S_y}\frac{S_y}{S_x} = r\frac{S_y}{S_x}$$

Recta de regresión de mínimos cuadrados de Y/X

$$y - \overline{y} = \frac{S_{xy}}{S_x^2}\left(x - \overline{x}\right) \qquad y - \overline{y} = r\frac{S_y}{S_x}\left(x - \overline{x}\right)$$

Recta de regresión de mínimos cuadrados de X/Y

$$x - \overline{x} = \frac{S_{xy}}{S_y^2}\left(y - \overline{y}\right) \qquad x - \overline{x} = r\frac{S_x}{S_y}\left(y - \overline{y}\right)$$

En la práctica se trabaja sólo con la recta de regresión de *Y/X,* así haremos en esta asignatura y también en cursos próximos en *Econometría.* Antes de realizar los correspondientes cálculos, llamaremos *Y* a la variable dependiente o explicada y denominaremos *X* a la variable independiente o explicativa (véase ejemplo 2.16 y ejercicio resuelto 3).

►***EJEMPLO 2.16.***

Se conoce que el consumo de pan en la dieta de los países desarrollados está relacionado con la renta per cápita. Ajuste una recta que explique dicha relación e interprete los coeficientes.

PAIS	Kg/habitante	Renta per cápita (miles de euros)
Alemania	40	14,8
Bélgica	41	15,1
Francia	57	13,5
Grecia	71	7,3
Holanda	61	12
Italia	42	14
Portugal	69	7,2

27

Y=consumo de pan (kg) *X*= Renta per cápita (miles de euros)

y_i	x_i	$x_i y_i$	x_i^2
40	14,8	592	219,04
41	15,1	619,1	228,01
57	13,5	769,5	182,25
71	7,3	518,3	53,29
61	12	732	144
42	14	588	196
69	7,2	496,8	51,84
381	83,9	4315,7	1074,43

$$\bar{y} = \frac{381}{7} = 54,4286 \qquad \bar{x} = \frac{83,9}{7} = 11,9857 \qquad S_x^2 = \left(\frac{1}{n}\sum_{i=1}^{n} x_i^2\right) - \bar{x}^2 = \frac{1074,43}{7} - 11,9857^2 = 9,833$$

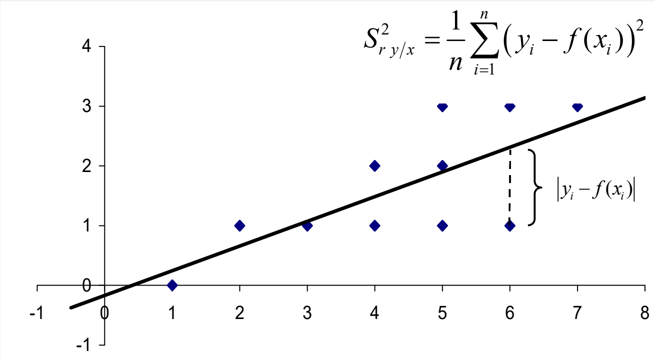
$$S_{xy} = \frac{1}{n}\sum_{i=1}^{n} x_i y_i - \bar{x}\,\bar{y} = \frac{4315,7}{7} - 54,4286 \times 11,9857 = -35,836$$

$$y - \bar{y} = \frac{S_{xy}}{S_x^2}\left(x - \bar{x}\right) \qquad y - 54,4286 = \frac{-35,836}{9,833}\left(x - 11,9857\right) \Leftrightarrow y = -3,64x + 98,1$$

28

Correlación lineal simple

Varianza residual. Coeficiente de determinación.



29

Correlación lineal simple

Varianza residual. Coeficiente de determinación.

$$S^2_{r\,y/x} = \frac{1}{n}\sum_{i=1}^{n}\left(y_i - f(x_i)\right)^2$$

$$R^2 = 1 - \frac{S^2_{ry}}{S^2_y} \qquad 0 \le R^2 \le 1$$

$$R^2 = r^2$$

30

►***EJEMPLO 2.17.***

Dé una medida de la bondad del ajuste de la recta de los ejemplo 2.13 y 2.15 y obtenga la varianza residual.

El coeficiente de determinación que mide la bondad del ajuste de la mencionada recta es $R^2 = r^2$.

$$r_{xy} = \frac{S_{xy}}{S_x S_y} = \frac{1,22}{1,72 \times 0,98} = 0,7238 \quad \Rightarrow \quad r_{xy}^2 = 0,5239$$

Varianza residual: $r^2 = 1 - \frac{S_{ry}^2}{S_y^2} \quad \Rightarrow \quad S_{r\,y/x}^2 = \left(1 - r^2\right) S_y^2 = \left(1 - 0,5239\right) 0,9599 = 0,457$

31

►***EJEMPLO***

Una fábrica de cerveza ha tomado al azar 8 semanas del año, observando la temperatura media en cada una de ellas y los miles de litros de cerveza demandados

Temperatura	Cerveza
10	26
28	82
30	98
35	103
20	60
12	35
27	68
25	71

¿Puede la fábrica planificar la cantidad de producción en función de la temperatura esperada? ¿De qué forma?

32

Según la relación entre ambas variables, la variable dependiente es *Y=demanda de cerveza* y la variable independiente *X=temperatura media*.

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
10	26	260	100	676
28	82	2296	784	6724
30	98	2940	900	9604
35	103	3605	1225	10609
20	60	1200	400	3600
12	35	420	144	1225
27	68	1836	729	4624
25	71	1775	625	5041
187	543	14332	4907	42103

33

$$\bar{x} = \frac{187}{8} = 23,375 \qquad \bar{y} = \frac{543}{8} = 67,875$$

$$S_x^2 = \left(\frac{1}{n}\sum_{i=1}^{n} x_i^2\right) - \bar{x}^2 = \frac{4907}{8} - 23,375^2 = 66,98$$

$$S_y^2 = \left(\frac{1}{n}\sum_{i=1}^{n} y_i^2\right) - \bar{y}^2 = \frac{42103}{8} - 67,875^2 = 655,86$$

$$S_{xy} = \frac{1}{n}\sum_{i=1}^{n} x_i y_i - \bar{x}\,\bar{y} = \frac{14332}{8} - 23,375 \times 67,875 = 204,92$$

$$r_{xy}^2 = \frac{S_{xy}^2}{S_x^2 S_y^2} = \frac{(204,92)^2}{66,98 \times 655,86} = 0,956$$

34

$$y-\overline{y}=\frac{S_{xy}}{S_x^2}\left(x-\overline{x}\right)$$

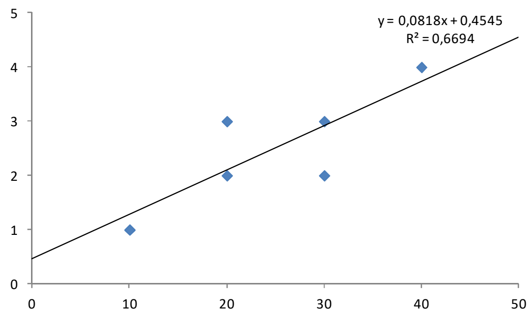
$$y-67,875=\frac{204,92}{66,98}\left(x-23,375\right) \qquad \Leftrightarrow \qquad y=3,06x-3,65$$

Por ejemplo, para una temperatura media de 33º se esperará una demanda de $y=\left(3,06\times 33\right)-3,65=97,33$ (97330 litros)

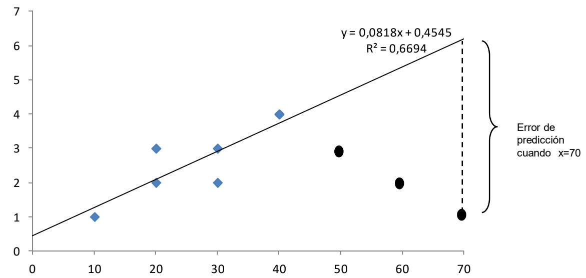
35

PREDICCIONES

x_i	y_i
10	1
20	2
20	3
30	2
30	3
40	4



36


y = 0,0818x + 0,4545
R² = 0,6694
Error de predicción cuando x=70
0
1
2
3
4
5
6
7
0
10
20
30
40
50
60
70

37