e-mail: **imozas@elx.uned.es**
www.telefonica.net/web/imm

EJERCICIOS DEL TEMA 5, DISTRIBUCIONES DE FRECUENCIAS BIDIMENSIONALES PROPUESTOS EN EXÁMENES

1.- a) Definición de varianza residual y coeficiente de determinación. **b)** Calcular la varianza residual y el coeficiente de correlación de la distribución:

x_i	y_i
2	3
4	2
8	4
9	7
11	10

(Febrero 99)

Respuesta.-

b)

x_i	y_i	x_i^2	$x_i \cdot y_i$	y_i^2
2	3	4	6	9
4	2	16	8	4
8	4	64	32	16
9	7	81	63	49
11	10	121	110	100
34	26	286	219	178

$a_{10} = 68$; $a_{01} = 5{,}2$; $a_{20} = 57{,}2$; $a_{02} = 35{,}6$; $a_{11} = 43{,}8$, de donde se obtiene:
$S_{xy} = m_{11} = 8{,}44$; $S_x^2 = m_{20} = 10{,}96$ y $S_y^2 = m_{02} = 8{,}56$. Así pues:

$$S_{ry}^2 = S_y^2 - \frac{S_{xy}^2}{S_x^2} = 8{,}56 - \frac{71{,}2336}{10{,}96} \cong 2{,}06; \quad R = \frac{S_{xy}}{S_x S_y} = \frac{8{,}44}{\sqrt{10{,}96 \cdot 8{,}56}} \cong 0{,}871$$

2.- Relación existente entre la varianza de la variable dependiente, la varianza explicada por la regresión y la varianza residual. Significado de cada una de ellas.*(Feb 2001)*

Respuesta.-

Supongamos que la recta de regresión de Y/X es y = a + bx. La relación que se pide es: $\mathbf{S_y^2 = S_{t_y}^2 + S_{r_y}^2}$, donde $S_y^2 = m_{02}$ es la varianza de la variable dependiente, $S_{t_y}^2 = \frac{m_{11}^2}{m_{20}}$ es la varianza explicada por la regresión (varianza de la variable $a + bx_i$) y $S_{r_y}^2 = m_{02} - \frac{m_{11}^2}{m_{20}}$ la varianza residual (varianza de la variable $r_i = y_i - a - bx_i$).

3.- Campo de variación del coeficiente de correlación lineal. Interpretación del valor nulo de este coeficiente, ¿qué se puede decir de las rectas de regresión en este caso? Razone la respuesta. *(Sept 2001 y Feb 2002)*

Respuesta.-

El coeficiente de correlación lineal $R = \frac{m_{11}}{\sqrt{m_{20} \cdot m_{02}}}$ cumple que $\mathbf{-1 \le R \le 1}$.

Las rectas de regresión:

- de Y/X: $y - a_{01} = \frac{m_{11}}{m_{20}}(x - a_{10})$

- de X/Y: $x - a_{10} = \frac{m_{11}}{m_{02}}(y - a_{01})$

e-mail: **imozas@elx.uned.es**
www.telefonica.net/web/imm

Si R = 0, no existe correlación entre las variables y, por ser $m_{11} = 0$, las rectas de regresión serían:

- de Y/X: $\mathbf{y - a_{01} = 0}$
- de X/Y: $\mathbf{x - a_{10} = 0}$

que son dos rectas perpendiculares, respectivamente paralelas a los ejes de coordenadas y que se cortan en el punto $(a_{10}, a_{01}) = (\overline{x}, \overline{y})$

4.- La Universidad Central de un cierto país ha realizado un estudio de la relación existente entre la exposición a un elemento contaminante y el número de personas que han desarrollado una nueva enfermedad. Esta investigación concluye que sí existe dicha relación, con una recta de regresión estimada de y = –2 + 1,2x, siendo "y" el porcentaje de personas afectadas, "x" los años de exposición a este elemento y el coeficiente de correlación lineal igual a 0,8. *(Sept 2001)*

a) Explíquese el significado de los valores –2 y 1,2 en la recta de regresión. **b)** ¿Qué porcentaje de enfermos puede esperarse para personas que han estado en contacto con el elemento contaminante durante 30 años? **c)** Si el coeficiente de correlación lineal hubiera sido igual a 1 ¿podríamos decir que el elemento contaminante es la única causa de la enfermedad?

Solución.-

a) El valor –2 de la recta de regresión es la ordenada en el origen. Carece de significado estadístico pues sería el porcentaje de personas afectadas, expuestas "cero" años al elemento contaminante. Esto implica además que, hasta que no pasen $\frac{5}{3}$ años (1 año y 8 meses) de exposición, no comenzará a haber enfermos ya que $-2 + 1{,}2 \cdot \frac{5}{3} = 0$

El valor 1,2 es la pendiente, esto es, el porcentaje que aumenta el número de personas afectadas, por año de exposición.

b) Haciendo x = 30 en la recta de regresión: $y = -2 + 1{,}2 \cdot 30 = 34$, es decir, el 34%.

c) Si R = 1 existe entre las variables una dependencia funcional exacta; también el coeficiente de determinación $R^2 = 1$, lo cual significa que el tiempo de exposición determina al 100% el porcentaje de enfermos. Ahora bien, no podríamos asegurar que fuese la única causa de la enfermedad por que desconocemos otras características de los enfermos (por ejemplo la edad, o alguna insuficiencia –conocida o desconocida- etc...)

5.- En el marco de un estudio sobre el comportamiento de la bolsa tras los acontecimientos de septiembre de 2.001 se decide incluir una encuesta a cien profesionales y empresarios. Su titulación universitaria es la siguiente: 50 economistas, 40 ingenieros y 10 abogados. El 20% de los economistas opina que subirá, mientras que el 40% de ellos piensa que bajará; el 50% de los ingenieros se inclina por la subida y tan sólo el 10% cree que bajará; por último, el 40% de los abogados se decanta por la estabilidad y el 60% que bajará. ¿Existe relación entre los pronósticos sobre la evolución del mercado bursátil y la profesión del encuestado? *(Feb 2002)*

Solución.-

En el pronóstico deberemos incluir también a aquellos que no contestan.

Valores observados

	Subirá	Bajará	NS/NC	
Economista	10	20	20	50
Ingeniero	20	4	16	40
Abogado	4	6	0	10
	34	30	36	100

Valores esperados

Subirá	Bajará	NS/NC
17	15	18
13,6	12	14,4
3,4	3	3,6

e-mail: **imozas@elx.uned.es**
www.telefonica.net/web/imm

Efectuando los cálculos correspondientes se obtiene un cuadrado de contingencia $\chi^2 = 20$ y un coeficiente de contingencia $C \cong 0{,}41$, es decir, los pronósticos sobre la evolución del mercado bursátil no son independientes de la profesión del encuestado, existiendo un grado de asociación que podemos cifrar en un 41%.

6.- Significado del cuadrado de contingencia. Utilización *(Sep 2002)*

7.- Asociación entre variables cualitativas. Coeficiente de contingencia de Pearson .*(Feb 2003)*

Respuesta.-

Consideremos un atributo M del que existen r modalidades y un atributo M' del que existen s modalidades. Sea n_{ij} el número de individuos con la modalidad i (i = 1, 2, ..., r) del atributo M y la modalidad j (j = 1, 2, 3, ..., s) del atributo M'. Sea $N = \sum_{i=1}^{r}\sum_{j=1}^{s} n_{ij}$ el total de individuos de la población, $n_{i\cdot}$ el total de individuos de la modalidad i de M y $n_{\cdot j}$ el total de individuos de la modalidad j de M'. Los atributos son independientes si se cumple que $\frac{n_{ij}}{N} = \frac{n_{i\bullet}}{N}\cdot\frac{n_{\bullet j}}{N}$ que equivale a que $n_{ij} = \frac{n_{i\bullet}\cdot n_{\bullet j}}{N}$, que es también equivalente a $n_{ij} = n'_{ij}$, habiendo llamado $n'_{ij} = \frac{n_{i\bullet}\cdot n_{\bullet j}}{N}$.

Se llama **cuadrado de contingencia** al coeficiente $\chi^2 = \sum_{i=1}^{r}\sum_{j=1}^{s}\frac{\left(n'_{ij} - n_{ij}\right)^2}{n'_{ij}}$

Se llama **coeficiente de contingencia de Pearson** a $C = \sqrt{\frac{\chi^2}{N+\chi^2}}$.

Se cumple que $0 \le C < 1$. Si los atributos son independientes entonces $C = 0$; a medida que aumenta el grado de asociación, C se aproxima a 1.

8.- Covarianza de dos variables estadísticas. Interpretación de sus posibles valores.*(Feb 2003 res)*

Respuesta.-

Consideremos la variable aleatoria bidimensional $(X,Y) = \{(x_i, y_i), i = 1,2,3...,n\}$ donde cada punto (x_i, y_i) tiene frecuencia 1. Sean $a_{10} = \frac{1}{n}\sum_{i=1}^{n} x_i$ y $a_{01} = \frac{1}{n}\sum_{i=1}^{n} y_i$ las medias marginales. Se define la covarianza $m_{11} = \frac{1}{n}\sum_{i=1}^{n}(x_i - a_{10})(y_i - a_{01})$. Puede adoptar cualquier valor y es quien proporciona el signo a la pendiente de las rectas de regresión y al coeficiente de correlación. Si es positiva, hay correlación directa y al aumentar los valores de X aumentan los de Y; si es negativa, hay correlación inversa y al aumentar X disminuye Y; si es cero, el coeficiente de correlación también y en ese caso no existe correlación entre las variables.

9.- Una determinada empresa está analizando la posibilidad de lanzar un nuevo producto "NP" al mercado, manteniendo el producto clásico "P"; "NP" ofrece mejores prestaciones y es un 10% más caro que "P". Con carácter previo a su fabricación ha encargado a una consultora un estudio de mercado entre sus consumidores. Esta consultora ha clasificado la población en cuatro grandes grupos profesionales, G_1, G_2, G_3 y G_4. A los seleccionados se les pregunta únicamente si estarían dispuestos a sustituir el producto "P" por "NP". El análisis proporciona el coeficiente de correlación de las variables (grupo profesional y decisión adoptada). Comente la conclusión de la encuesta. *(Sep 2003)*

e-mail: **imozas@elx.uned.es**
www.telefonica.net/web/imm

Respuesta.-

Es imposible proporcionar el coeficiente de correlación de variables cualitativas.

10.- Problema de la multicolinealidad en el ajuste de un plano. *(Sep 2003, 2005 y 2006)*

Respuesta.-

Se presenta este problema cuando el coeficiente de correlación R_{12} entre las variables exógenas X_1 y X_2 es próximo a 1 ó –1, por ejemplo $0{,}8 < |R_{12}| < 1$. En ese caso, los coeficientes de regresión parcial que se calculen no resultan fiables. Si fuese $|R_{12}| = 1$, los coeficientes de regresión parcial no se pueden calcular pues en ambos casos se obtiene $\frac{0}{0}$.

11.- El coeficiente de correlación lineal simple. Interpretación de sus posibles valores. *(Sep 2003)*

Respuesta.-

$R = \frac{m_{11}}{\sqrt{m_{20}\cdot m_{02}}}$, donde m_{11} es la covarianza y m_{20} y m_{02} son las varianzas de las distribuciones marginales de la x y de la y, respectivamente. Se cumple que $-1 \le R \le +1$.

Su valor se usa para determinar el grado de dependencia de la y (variable endógena) respecto de la x (variable exógena)

Si $R = \pm 1$, la nube de puntos forma una línea recta con pendiente positiva (si R = +1) o negativa (si R = –1) que a su vez coincide con ambas rectas de regresión. La dependencia es exacta o funcional.

Si R = 0, no existe correlación entre las variables y las rectas de regresión son respectivamente $y = a_{01}$; $x = a_{10}$. No existe dependencia de tipo lineal entre las variables.

Si $|R| \ge 0{,}75$, se considera que la dependencia es fuerte o aceptable y en ese caso las rectas de regresión pueden usarse para hacer predicciones.

12.- Mediante un ajuste por mínimos cuadrados, y dada la siguiente tabla

t_i	1.986	1.987	1.988	1.989	1.990	1.991	1.992
y_i	3	6	7	8	10	11	12

hágase una previsión de la renta para este año y calcúlese el coeficiente de correlación lineal expresando el grado de bondad del ajuste efectuado. *(Sep 2003)*

Solución.-

Construimos la siguiente tabla, haciendo un cambio de escala en la variable t_i para facilitar los cálculos:

t_i	$t'_i = t_i - 1989$	y_i	$t'_i \cdot y_i$	t'^2_i	y_i^2
1986	-3	3	-9	9	9
1987	-2	6	-12	4	36
1988	-1	7	-7	1	49
1989	0	8	0	0	64
1990	1	10	10	1	100
1991	2	11	22	4	121
1992	3	12	36	9	144
	0	**57**	**40**	**28**	**523**

$\mathbf{a_{10}} = 0$ $\quad$ $\mathbf{a_{11}} = \frac{57}{7}$ $\quad$ $\mathbf{m_{11}} = \frac{57}{7}$

$\mathbf{a_{01}} = \frac{57}{7}$ $\quad$ $\mathbf{a_{20}} = 4$ $\quad$ $\mathbf{m_{20}} = 4$

$\mathbf{a_{02}} = \frac{523}{7}$ $\quad$ $\mathbf{m_{02}} = \frac{523}{7} - \left(\frac{57}{7}\right)^2 = \frac{412}{49}$

De donde la recta de regresión de Y/T':

$$y - \frac{57}{7} = \frac{10}{7}t' \leftrightarrow y\ \frac{10t'+57}{7}$$ o deshaciendo el cambio de escala:

$$y = \frac{10(t-1989)+57}{7} = \frac{10t-19833}{7}$$

La previsión de la renta para este año: $y_{2003} = \frac{197}{7} \cong 28{,}14$

El coeficiente de correlación (es invariante ante los cambios de escala):

$$R = \frac{40}{7\sqrt{4\cdot\frac{412}{49}}} \cong 0{,}985$$

de donde se deduce que la previsión efectuada es aceptable.

13.- Se les pregunta a 500 personas con empleo cuál es, en su opinión, el problema económico más importante en España durante el año 2.001, obteniéndose el siguiente resultado *(Sep 2003)*

Categoría profesional	Paro	Inflación	Total
Asalariados	n_{11}=236	n_{12}=86	$n_{1\cdot}$=322
Profesionales libres	n_{21}=53	n_{22}=125	$n_{2\cdot}$=178
Total	$n_{\cdot 1}$=289	$n_{\cdot 2}$=211	N=500

¿Existe alguna relación entre la ocupación laboral y el problema que más preocupa?.

Solución.-

Construimos la tabla de las frecuencias teóricas $n'_{ij} = \frac{n_{i\cdot}n_{\cdot j}}{N}$.

Categoría profesional	Paro	Inflación	Total
Asalariados	n'_{11} =186,12	n'_{12} =135,88	322
Profesionales libres	n'_{21} =102,88	n'_{22} = 75,12	178
Total	289	211	500

De aquí obtenemos los elementos de la χ^2: $\frac{\left(n'_{ij}-n_{ij}\right)^2}{n'_{ij}}$

Categoría profesional	Paro	Inflación
Asalariados	13,37	18,31
Profesionales libres	24,19	33,13

y el cuadrado de contingencia: $\chi^2 = \sum_{i=1}^{2}\sum_{j=1}^{2}\frac{\left(n'_{ij}-n_{ij}\right)^2}{n'_{ij}} = 89.$

(Si los atributos considerados fuesen independientes este coeficiente sería cero.)

El coeficiente de contingencia de Pearson:

$$C = \sqrt{\frac{\chi^2}{N+\chi^2}} = \sqrt{\frac{89}{500+89}} \cong 0{,}39$$

Existe por tanto cierto grado de asociación entre los atributos dados.

14.- Finalizadas las elecciones, un determinado partido político de un país X realiza un análisis de los gastos de propaganda de las campañas y su rentabilidad política. En las últimas cinco elecciones celebradas el coste de la publicidad y el número de diputados obtenidos (que es la cuantificación que realiza de la rentabilidad política citada) son los siguientes:

Gastos en propaganda (millones de u.m)	Diputados elegidos
1,50	3
1,75	4
3,25	4
4,00	6
5,00	8

En ese país existe una Gabinete Electoral Central que está estudiando la posibilidad de un presupuesto de propaganda electoral para las elecciones de diez millones de unidades monetarias. a) ¿Cuál será el número de diputados pertenecientes al partido X que se podría esperar fueran elegidos con ese presupuesto en las próximas elecciones? b) ¿Cómo cuantificaría, con una banda entre 0 y 1, la confianza con la que se puede esperar ese resultado? *(Feb 2004)*

Solución.- Siendo los "gastos en propaganda" y el "número de diputados elegidos" las variables X e Y respectivamente, calcularemos la recta de regresión de Y/X:

x_i	y_i	x_i^2	y_i^2	x_iy_i
1,5	3	2,25	9	4,5
1,75	4	3,0625	16	7
3,25	4	10,5625	16	13
4	6	16	36	24
5	8	25	64	40
15,5	25	56,875	141	88,5

De la tabla obtenemos los valores:

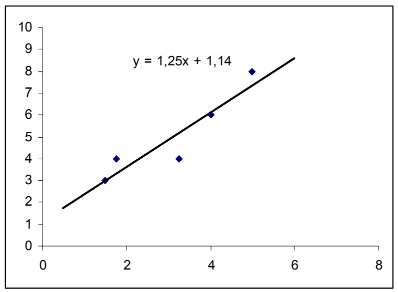
a_{10}=3,1 $\quad$ m_{11}=2,2

a_{01}=5 $\quad$ m_{20}=1,765

a_{11}=17,7 $\quad$ m_{02}=3,2

a_{20}=11,375

a_{02}=28,2



y de aquí la recta de regresión: $y \cong 1{,}25x + 1{,}14$ y el coeficiente de determinación $r^2 \cong 0{,}857$. Así pues:

a) Haciendo x = 10 en la recta de regresión, obtenemos y = 13,6 ≅ 14 diputados

b) La confianza, que cuantificamos con el coeficiente de determinación, es del 85,7 %.

15.- Tablas de contingencia. Utilización. *(Feb 2004)*

16.- El precio de un determinado producto X de la cesta de la compra en los últimos cinco años, así como el consumo total de productos alimenticios en ese mismo periodo, y para una determinada demarcación geográfica, han sido los siguientes:

Precio unitario producto (u.m.)	Consumo total (miles de millones u.m.)
12	42
16	48
22	50
24	55
30	62

Si se ha estimado que el consumo en el próximo año va a ser de 70.000 millones u.m., compruébese que los ingresos estimados a obtener por la venta del producto X (para un producción de 200.000 unidades) pueden financiar los costes, que se han presupuestado en 7 millones de u. m. Cuantifique la medida de la fiabilidad del resultado. Observación: suponga para la resolución del problema que el precio del producto X depende del consumo total. *(Feb 2004)*

Solución.-

Efectuemos los cálculos para hallar la recta de regresión y el coeficiente de correlación:

x_i	y_i	x_i^2	y_i^2	x_iy_i
12	42	144	1764	504
16	48	256	2304	768
22	50	484	2500	1100
24	55	576	3025	1320
30	62	900	3844	1860
104	257	2360	13437	5552

De donde obtenemos:

$a_{10}=20,8$ $m_{11}=41,28$
$a_{01}=51,4$ $m_{20}=39,36$
$a_{11}=1110,4$ $m_{02}=45,44$
$a_{20}=472$
$a_{02}=2687,4$

Puesto que se supone que el precio depende del consumo, calculamos la recta de regresión de X sobre Y, para la que se obtiene: **x ≅ 0,91y – 25,89.** Haciendo y = 70, se obtiene x ≅ 37,70 u.m de precio unitario. El precio de 200.000 unidades sería:

200000· 37,70 = **7.539.436,62 u.m**.

que supera a los 7.000.000 u.m. en que se han presupuestado los costes.

El coeficiente de determinación, cuyo valor es $r^2 \cong 0,953$ nos indica que la fiabilidad del resultado supera el 95%.

17.- Relación existente entre la varianza de la variable dependiente, la varianza explicada por la regresión y la varianza residual. Significado de cada una de ellas. *(Sep 2004)*

Respuesta.-

$$\boxed{S_y^2 = S_{y_t}^2 + S_{r_y}^2}$$

S_y^2 es la varianza de la variable independiente y_i, que es el momento de segundo orden respecto de la media $m_{02} = \frac{1}{N}\sum_{i=1}^{N}(y_i - a_{01})^2$; $S_{y_t}^2$ es la varianza de la variable $y_{t_i} = a + bx_i$, explicada por la regresión cuyo valor es $S_{y_t}^2 = \frac{m_{11}^2}{m_{20}}$, donde m_{11} es la covarianza; $S_{r_y}^2$ es la varianza de la variable residual $e_i = y_i – a – bx_i$.

18.- Dadas las rectas y = 2x + 1

x = 5y + 10

Compruebe si son, respectivamente, las rectas de regresión mínimo-cuadráticas de Y sobre X y de X sobre Y de una misma serie de observaciones. *(Feb 2005)*

Solución.-

Si lo fuesen, debería ser: $\left.\begin{array}{l}\frac{m_{11}}{m_{20}} = 2\\ \frac{m_{11}}{m_{02}} = 5\end{array}\right\}$. Multiplicando miembro a miembro se obtiene $\frac{m_{11}^2}{m_{20}\cdot m_{02}} = 10$. Pero esto no puede ser porque $\frac{m_{11}^2}{m_{20}\cdot m_{02}}$ es el coeficiente de determinación R^2 que debe ser ≤ 1.

19.- Coeficiente de determinación: definición y expresión analítica. Campo de variación y significado. *(Sep 2005)*

Respuesta.-

Puesto que la varianza de la variable dependiente (S_y^2) es igual a la suma de la varianza explicada por la regresión (S_{yt}^2) más la varianza residual (S_{ry}^2), se tiene que $1 = \frac{S_{yt}^2}{S_y^2} + \frac{S_{ry}^2}{S_y^2}$. Se denomina coeficiente de determinación R^2 al cociente $\frac{S_{yt}^2}{S_y^2}$, es decir, a la porción de varianza que es explicada por la regresión. Se deduce de lo anterior que $0 \le R^2 \le 1$ y su valor proporciona una medida de la bondad del ajuste, es decir, la mayor o menor aproximación de la recta de regresión a la nube de puntos. El ajuste será mejor cuanto más próximo a 1 esté R^2.

20.- La empresa Maquimport se dedica a la importación de un cierto tipo de maquinaria y tiene una cuota de mercado del 4% del sector. Teniendo en cuenta que, en los últimos seis años el volumen de importación de maquinaria y la producción industrial de los sectores que han absorbido estas importaciones han sido:

Año	Importación (M€)	Producción (M€)
X_1	22	195
X_2	33	120
X_3	45	125
X_4	50	130
X_5	65	140
X_6	67	154

Determínese: **a)** El volumen de importación de esta empresa en un año en el que la producción estimada es de 200 M€ (suponiendo que se mantenga en dicho año la relación inicial entre las magnitudes); **b)** Fiabilidad de dicha estimación; **c)** Calcúlese la varianza debida a la regresión y la varianza residual.

Solución.-

a) Tomaremos la producción como variable independiente y la importación como variable dependiente. De la tabla:

Producción x_i	Importación y_i	x_i^2	y_i^2	x_iy_i
195	22	38025	484	4290
120	33	14400	1089	3960
125	45	15625	2025	5625
130	50	16900	2500	6500
140	65	19600	4225	9100
154	67	23716	4489	10318
864	282	128266	14812	39793

Obtenemos los momentos:

$a_{10} = 144$ $m_{20} = 641,67$
$a_{01} = 47$ $m_{02} = 259,67$
$a_{20} = 21377,67$ $m_{11} = -135,83$
$a_{02} = 2468,67$
$a_{11} = 6632,17$

de donde obtenemos la recta de regresión de Y/X:

$$y - 47 = \frac{-135,83}{259,67}(x - 144) \leftrightarrow \mathbf{y = -0,2117x + 77,483}$$

haciendo x = 200, obtenemos el volumen de importación: 35,15 M€.

b) Para medir la fiabilidad de la estimación usaremos el coeficiente de determinación, cuyo valor es $R^2 = \frac{m_{11}^2}{m_{20}m_{02}} = \frac{(-135,83)^2}{641,67 \cdot 259,67} \cong 0,11$. Un valor tan pequeño significa que la estimación no es fiable.

c) Varianza debida a la regresión $= \frac{m_{11}^2}{m_{20}} = \frac{(-135,83)^2}{641,67} \cong 28,75$

Varianza residual $= m_{02} - \frac{m_{11}^2}{m_{20}} = 259,67 - \frac{(-135,83)^2}{641,67} \cong 230,91$

21.- Dada la distribución bidimensional

x_i	y_i
10	200
20	180
30	150
40	120
50	100

a)Ajústese una recta por el procedimiento de los mínimos cuadrados. **b)**Calcúlese el coeficiente de correlación lineal. *(Sep 2005 Res)*

Solución.-

Efectuamos los cálculos:

x_i	y_i	x_i^2	y_i^2	x_iy_i
10	200	100	40000	2000
20	180	400	32400	3600
30	150	900	22500	4500
40	120	1600	14400	4800
50	100	2500	10000	5000
150	750	5500	119300	19900

de donde obtenemos:

$a_{10}=30$
$a_{01}=150$ $m_{11}=-520$
$a_{11}=3980$ $m_{20}=200$
$a_{20}=1100$ $m_{02}=1360$
$a_{02}=23860$

y de aquí la recta de regresión de Y/X:

$$y - 150 = \frac{-520}{200}(x - 30) \leftrightarrow y = -2,6x + 228$$

El coeficiente de correlación: $R = \frac{m_{11}}{\sqrt{m_{20} \cdot m_{02}}} = \frac{-520}{\sqrt{200 \cdot 1360}} \cong -0,997$

22.- *(Feb 2006 y, parcialmente Feb 2004)*

Cite los ajustes no lineales por mínimos cuadrados que conozca, expresando analíticamente, en cada caso, el modelo que se pretende ajustar.

Respuesta.-

Ajuste de una parábola: .... $y = a_0 + a_1x + a_2x^2$.

Ajuste de una hipérbola equilátera: ... $y = a_0 + \frac{a_1}{x}$

Ajuste potencial: .... $y = a_0 x^{a_1}$.

Ajuste exponencial: .... $y = a_0 a_1^x$