

EJERCICIOS REGRESIÓN LINEAL

1) Se presume que los accidentes en las empresas son directamente proporcionales a los incidentes. Para comprobarlo, hacer una Regresión Lineal con los datos de Accidentes e Incidentes tomados de las siguientes empresas:

EMPRESA	1	2	3	4	5
INCIDENTES AL AÑO	120	128	144	168	176
ACCIDENTES AL AÑO	13	14	16	19	20

Rta.: Como $R = 1$ hay correlación perfecta, por ende SI, son directamente proporcionales.

Luego $a = 1/8$ y $b = -2$ hacen la recta de regresión: $Y' = \frac{1}{8} \cdot X - 2$

2) Se presume que los accidentes en las empresas son inversamente proporcionales a la inversión realizada por las empresas para evitarlos. Para comprobarlo, hacer una Regresión Lineal con los datos de Accidentes e Inversión tomados de las siguientes empresas y calcule que accidentes habría en una empresa que invierte \$21.000 y en otra que invierte \$36.000:

EMPRESA	1	2	3	4	5	6
INVERSION EN HYS	\$ 15.000	\$ 16.500	\$ 18.000	\$ 19.500	\$ 24.000	\$ 28.500
ACCIDENTES AL AÑO	39	38	37	36	33	30

Rta.: Como $R = 1$ hay correlación perfecta, por ende SI, son inversamente proporcionales.

Luego $a = -1/1500$ y $b = 49$ hacen la recta de regresión: $Y' = -\frac{1}{1500} \cdot X + 49$

Para \$21000 de inversión, reemplazo en la recta y me da: $Y' = 35$ Accidentes

Para \$36000 de inversión, reemplazo en la recta y me da: $Y' = 25$ Accidentes

3) Se presume que la concentración de un contaminante en el aire esta ligada de alguna forma con el uso de un producto llamado CONTAMINOL. Indique:

a- qué grado de correlación hay

b- qué concentración se espera si se utilizan 5 lts. de CONTAMINOL

c- cual es el error medio en la recta de regresión

CONCENTRACION ppm	600	650	670	680	720
USO CONTAMINOL (lts.)	1,92	2,18	2,35	2,40	2,65

Rta. Problema 3:

CONCENTRACION ppm	600	650	670	680	720	Y Media =	664 ppm
USO CONTAMINOL (lts.)	1,92	2,18	2,35	2,40	2,65	X Media =	2,30 lts.

$(X_i - X)(Y_i - Y) =$	24,32	1,68	0,3	1,6	19,6	Suma =	47,5
$(X_i - X)^2 =$	0,1444	0,0144	0,0025	0,01	0,1225	Suma =	0,2938
$(Y_i - Y)^2 =$	4096	196	36	256	3136	Suma =	7720

a =	161,675	sale de 47,5 / 0,2938
b =	292,148	sale de 664 - 161,675 x 2,30

R =	0,99738	sale de 47,5 / raiz(0,2938x7720)
------------	----------------	----------------------------------

- a) hay una gran correlación (cercana a 1), por ende ambos están ligados, es decir, la concentración del contaminante depende en gran medida del uso del producto CONTAMINOL
- b) con la recta $Y' = 161,675.X + 292,15$ calculo la concentración esperada para un uso de 5 lts. de Contaminol, es decir: $Y' = 161,675.5 + 292,15 = 1100 \text{ ppm} = \text{Conc. Esperada}$

c) Debo calcular los Y' (estimados para c/u de los puntos) y luego aplico la fórmula: $Err\% = \frac{\sum \frac{|Y' - Y|}{Y}}{n}$

CONC. REAL ppm (Y)	600	650	670	680	720
CONC. CALC. ppm (Y')	602,56	644,60	672,08	680,17	720,59

Err % =	0,004273	0,008309	0,003110	0,000246	0,000814
Suma Err% =	0,016752				
Promedio Err% =	0,003350	Es decir, el error porcentual promedio cometido es del 0,33%			

CURVA DE REGRESIÓN

