

Ejercicios de Regresión Lineal

1º) En un determinado proceso industrial, se piensa que la producción "Y" en toneladas, está linealmente relacionada con la temperatura "X". Se toman dos observaciones de producción para cada nivel de temperatura, obteniendo los siguientes datos:

Temperatura (°C)	0	0.5	1.5	2.0	2.5
Producción	7.2 ; 6.9	8.1 ; 8.4	9.8 ; 10.1	11.3 ; 11.7	12.9 ; 13.2

- Obtenga la ecuación de regresión y su coeficiente correlación lineal
- ¿Es significativa la regresión?. Use $\alpha = 0,01$
- Obtenga un intervalo del 90% de confianza para un valor particular de la producción, cuando la temperatura sea de 1°C .

Solución : a) $y = 6,96 + 2,31 x$ b) $F = 376,66$ Si es significativa

2º) El número de bacterias por unidad de volumen en un cultivo tras "X" horas, viene dado en la siguiente tabla:

X(Horas)	0	1	2	3	4	5	6
Y (Bac/Vol)	32	47	65	92	132	190	275

- Ajústeles una curva de la forma $Y = a b^X$.
- Estime el valor de Y cuando $X = 7$

Solución: $a = 32,14$, $b = 1,427$ $Y = 387$

3º) Doce láminas de acero trabajadas en frío, que tienen diferente contenido de Cobre (%) , y diferentes temperaturas de recocido (°F) , tuvieron la siguiente Dureza (Rockwell) :

Dureza	78.8	65.1	55.4	56.2	80.9	69.5	57.4	55.2	85.6	71.8	60.2	58.7
%Cobre	0.02	0.02	0.02	0.02	0.10	0.10	0.10	0.10	0.18	0.18	0.18	0.18
Temp	1000	1100	1200	1300	1000	1100	1200	1300	1000	1100	1200	1300

Estos datos fueron procesados con el programa SPSS opción Regresión Lineal Múltiple.

ANOVA

	Sum of Squares	df	Mean Square	F	Sig.
Regression	?	?	40,607	?	
Residual	?	14,432			
Total	?				

a Predictors: (Constant), TEMP, COBRE

b Dependent Variable: DUREZA

	Coefficients		t	Sig.
	Beta	Std. Error		
(Constant)	?	11,457	14,163	?
COBRE	?	16,789	?	,085
TEMP	-8,633E-02	,010	?	?

a Dependent Variable: DUREZA

- a) Complete los valores que faltan, y calcule el coeficiente de determinación del modelo.
 b) Redacte un informe con sus conclusiones. ¿Cuáles son las variables más influyentes, y como lo hacen?. ¿cómo es la precisión del modelo?
 c) Construya un intervalo del 95% de confianza, para el incremento en resistencia, cuando la temperatura de recocido se incrementa en 1°F ?.

Solución:

ANOVA					
	Sum of Squares	df	Mean Square	F	Sig.
Regression	1172,097	2	586,048	40,607	,000
Residual	129,890	9	14,432		
Total	1301,987	11			

	Coefficients		t	Sig.
	Coefficients	Std. Error		
(Constant)	162,267	11,457	14,163	,000
COBRE	32,500	16,789	1,936	,085
TEMP	-8,633 E-02	,010	-8,802	,000

4º) Suponga que al aplicar un modelo de regresión lineal simple, Ud. obtuvo la siguiente ecuación : $Y = 3.72 + 0,25 X$; en donde "Y" representa una longitud expresada en pulgadas, y "X" una temperatura expresada en ° F . , y Ud. desea pasar dicha ecuación a centímetros y ° C .

¿Considera Ud., que al hacer esta conversión, se altera el coeficiente de correlación del modelo ? . Justifique y demuestre su respuesta .

Solución: No se altera

5º) En la regresión lineal simple : $Y = \beta_0 + \beta_1 X$, cuando se quiere probar la Hipótesis $H_0: \beta_1 = 0$ vs. $H_1: \beta_1 \neq 0$, el procedimiento es mediante una t- Student con $(n - 2)$ grados de libertad, dada por la siguiente expresión:

$t = \frac{\hat{\beta}_1}{S_e} \cdot \sqrt{\sum_{i=1}^{i=n} (X_i - \bar{X})^2}$; y cuando se quiere probar la Hipótesis: $H_0: \rho = 0$ vs.

$H_1: \rho \neq 0$, se aplica otra t- Student , también con $(n-2)$ grados de libertad, dada

por la expresión: $t = \frac{r \cdot \sqrt{n-2}}{\sqrt{1-r^2}}$.

Explique la importancia práctica de dichas pruebas, y diga justificadamente si son equivalentes.

Solución : Son equivalentes.

6º) Se tiene un conjunto de puntos: (X_1, Y_1) (X_2, Y_2), (X_n, Y_n) a los cuales se les quiere ajustar una línea recta.

Suponga que la ecuación de la recta de regresión de "Y" sobre "X" , obtenida por el método de mínimos cuadrados es de la forma : $Y = a_1 + b_1 X$, y que la recta de regresión de "X" sobre "Y" , es de la forma : $X = a_2 + b_2 Y$

a) ¿ Considera Ud. que ambas rectas son idénticas ? . Justifique .

b) ¿ Tienen ambas rectas el mismo coeficiente de correlación ? . Justifique .

c) Demuestre que el coeficiente de determinación viene dado por: $r^2 = b_1 b_2$ para ambos casos .

Solución: Las rectas no son idénticas, mientras que "r" si.

7º) El objetivo de un estudio es identificar cuales son los factores más influyentes en el consumo de combustible por los vehículos.

En el estudio en cuestión, se consideraron 25 modelos diferentes, y se seleccionaron las siguientes variables:

X ₁	X ₂	X ₃	X ₄	X ₅	X ₆	X ₇	Y	
350	165	260	4	200.3	69.9	3910	18.90	X ₁ = Cilindraje (in ³)
250	105	185	1	196.7	72.2	3510	20.00	
351	143	255	2	199.9	74.0	3890	18.25	X ₂ = Caballos de fuerza
225	95	170	1	194.1	71.8	3365	20.07	
440	215	330	4	184.5	69.0	4215	11.20	X ₃ = Momento de torsión
231	110	175	2	179.3	65.4	3020	22.12	
89.7	70	81	2	155.7	64.0	1905	34.70	X ₄ = Carburador (gargantas)

Ejercicios de Estadística para Ingenieros
 Profesor: Angel Francisco Arvelo L.

96.9	75	83	2	165.2	65.0	2320	30.40	
350	155	250	4	195.4	74.4	3885	16.50	X ₅ = Longitud (pies)
85.3	80	83	2	160.6	62.2	2009	36.50	
171	109	146	2	170.4	66.9	2655	21.50	X ₆ = Ancho (pies)
258	110	195	1	171.5	77.0	3375	19.70	
302	129	220	2	199.9	74.0	3890	17.80	X ₇ = Peso (libras)
500	190	360	4	224.1	79.8	5290	14.39	
440	215	330	4	231.0	79.7	5185	14.89	Y = Consumo de combustible (millas /galón)
350	155	250	4	196.7	72.2	3910	17.80	
231	110	175	2	179.3	65.4	3050	23.54	
360	180	290	2	214.2	76.3	4250	21.47	
400	185	300	4	196.0	73.0	3850	16.59	
96.9	75	83	2	165.2	61.8	2275	31.90	
460	223	366	4	228.0	79.8	5430	13.27	
133.6	96	120	2	171.5	63.4	2535	23.90	
318	140	255	2	215.3	76.3	4370	19.73	
351	148	243	2	215.5	78.5	4540	13.90	
350	165	255	4	185.2	69.0	3660	16.50	

Esta información fue procesada con el programa SPSS, obteniendo los siguientes resultados:

SOLUCION

Model Summary

Model	R Square	Std. Error of the Estimate
	?	?

ANOVA

	Sum of Squares	df	Mean Square	F	Sig.
Regression	?	?	?	?	?
Residual	135,530	?	?		
Total	1032,898	?			

a Predictors: (Constant), X7, X4, X6, X2, X5, X1, X3

b Dependent Variable: Y

Coefficients

	Coefficients		t	Sig.
	B	Std. Error		
(Constant)	7,114	?	,350	?
X1	?	,048	-2,426	?
X2	-6,571E-02	?	-,839	,413
X3	,140	,090	?	?
X4	?	1,111	?	,092
X5	?	,109	?	,188
X6	,206	?	?	,493

Ejercicios de Estadística para Ingenieros
 Profesor: Angel Francisco Arvelo L.

X7 -6,360E-03 ,005 ? ?

a Dependent Variable: Y

a) Complete los valores que faltan.

b) Redacte un informe detallado con las conclusiones del modelo.

c) Encuentre un intervalo del 95% de confianza para β_1 , e interprete el resultado.

d) Pruebe la hipótesis: $H_0 : \beta_0 = 0$ contra $H_1 : \beta_0 \neq 0$

SOLUCION

Model Summary

Model	R	R Square	Std. Error of the Estimate
	,932	,869	2,8235

ANOVA

	Sum of Squares	df	Mean Square	F	Sig.
Regression	897,368	7	128,195	16,080	,000
Residual	135,530	17	7,972		
Total	1032,898	24			

a Predictors: (Constant), X7, X4, X6, X2, X5, X1, X3

b Dependent Variable: Y

Coefficients

	Coefficients		t	Sig.	95% Confidence Interval for B	
	B	Std. Error			Lower Bound	Upper Bound
(Constant)	7,114	20,304	,350	,730	-35,723	49,951
X1	-,117	,048	-2,426	,027	-,218	-,015
X2	-6,571E-02	,078	-,839	,413	-,231	,100
X3	,140	,090	1,558	,138	-,049	,329
X4	1,983	1,111	1,786	,092	-,360	4,326
X5	,150	,109	1,372	,188	-,081	,380
X6	,206	,294	,701	,493	-,415	,827
X7	-6,360E-03	,005	-1,340	,198	-,016	,004

a Dependent Variable: Y

8º) Suponga que se quiere construir un modelo de regresión múltiple, con una superficie de respuesta de ecuación : $Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_1 X_2 + e$.

Se dispone de un conjunto de "n" datos $\{(x_{11}, x_{21}, y_1) \{ (x_{12}, x_{22}, y_2) \dots \dots \{ (x_{1n}, x_{2n}, y_n) \}$.

Aplique el método de mínimos cuadrados, para obtener las ecuaciones normales de la regresión, que permitan despejar los estimadores para los coeficientes: β_0 , β_1 , β_2 y β_3 .

9º) En un modelo de regresión lineal simple $Y = \beta_0 + \beta_1 X + e.$, suponga que se realizan "k" observaciones de Y cuando $X = X_1$, "k" observaciones de Y cuando $X = X_2$, ..., "k" observaciones de Y cuando $X = X_m$.

Demuestre que la recta de regresión es idéntica a la determinada por los puntos (X_1, \bar{Y}_1) (X_2, \bar{Y}_2) (X_m, \bar{Y}_m) .

¿Se puede decir lo mismo si el número de observaciones para cada valor de "X" es diferente ? .

10º) La presión "P" de un gas que corresponde a varios volúmenes V, se registra como sigue:

V (cm ³)	50	60	70	90	100
P (Kg/cm ²)	64.7	51.3	40.5	25.9	7.8

La ley de los gases ideales establece: $PV^\gamma = C$, donde " γ " y "C" son constantes para cada gas.

Estime las constantes " γ " y "C" de este gas.