Estadística – I 2025

Regresión Lineal Univariada

IC para parámetros

Isaías 26:7

Hipótesis de regresión

*Si para cada valor x_i de X, se tiene que la variable Y_{x_i} cumple que (**hipótesis de***

a) Y_{x_i} *sigue una distribución normal*

b) $E(Y_{x_i}) = \alpha + \beta x_i$

c) $Var(Y_{x_i}) = \sigma^2$ *(note que no depende del valor de x_i)*

d) *Las variables Y_{x_i} son independientes*

$$\boxed{Y_{x_i} \sim N(\alpha + \beta x_i, \sigma^2)}$$

$\boxed{Y_{x_i} \text{ son independientes,}}$

Hipótesis de regresión

1. *Se tiene que* $\varepsilon_i \sim N\left(0, \sigma^2\right)$ *y los* ε_i *son independientes.*

2. $Var\left(B\right) = \dfrac{\sigma^2}{S_{xx}}$

3. $Var\left(A\right) = \sigma^2\left(\dfrac{1}{n} + \dfrac{\overline{x}^2}{S_{xx}}\right) = \dfrac{\sigma^2 \sum\limits_{i=1}^{n} x_i^2}{nS_{xx}}.$

4. $B \sim N\left(\beta, \dfrac{\sigma^2}{S_{xx}}\right).$

5. $A \sim N\left(\alpha, \sigma^2\left(\dfrac{1}{n} + \dfrac{\overline{x}^2}{S_{xx}}\right)\right)$

Intervalos de confianza para coeficientes

Teorema 49 *Bajo las hipótesis de regresión, se tiene que IC de* $(1-\delta)\,100\%$ *para*

1. α *tiene extremos* $a \pm t_{\delta/2,n-2}\left(s\sqrt{\dfrac{\sum x^2}{nS_{xx}}}\right)$.

2. β *tiene extremos* $b \pm t_{\delta/2,n-2}\left(s\sqrt{\dfrac{1}{S_{xx}}}\right)$.

$$s^2 = \frac{\sum_{i=1}^{n} (e_i)^2}{n-2} = \frac{SCE}{n-2} = \frac{S_{yy} - bS_{xy}}{n-2}$$

Intervalos de confianza y predicción.

Teorema 51 *Bajo las hipótesis de regresión, se tiene que IC de* $(1-\delta)\,100\%$ *para* $E\left(Y_{x_0}\right)$ *tiene extremos*

$$\underbrace{(a+bx_0)}_{\widehat{y}_0} \pm t_{\delta/2,n-2}\left(s\sqrt{\frac{1}{n}+\frac{(x_0-\overline{x})^2}{S_{xx}}}\right).$$

Definición 36 *Bajo las hipótesis de regresión, se tiene que un intervalo de predicción* (IP) *de* $(1-\delta)\,100\%$ *para el valor desconocido* y_0 *asociado a* x_0 *tiene extremos*

$$\widehat{y}_0 \pm t_{\delta/2,n-2}\; s\sqrt{1+\frac{1}{n}+\frac{(x_0-\overline{x})^2}{S_{xx}}}.$$

$]4.09575, \quad 14.53005[$

Ejemplo

Una cooperativa de ahorro ha sufrido un cierre técnico debido a una deficiente administración. Esta entidad financiera estima que el tiempo (en días) para la devolución de dinero depende linealmente de la cantidad de dinero (millones de colones) que el cliente tiene en su cuenta. Los siguientes datos se refieren a información de 10 clientes sobre el dinero en su cuenta (x) y el tiempo que han esperado para su devolución (y).

$$\sum_{i=1}^{10} x_i^2 = 433 \qquad \sum_{i=1}^{10} x_i = 51 \qquad \sum_{i=1}^{10} y_i^2 = 1545 \qquad \sum_{i=1}^{10} y_i = 95 \qquad \sum_{i=1}^{10} x_i y_i = 808$$

c) Construya un intervalo de predicción del 95 % para el tiempo que debe esperar un cliente que tiene 5 millones de colones en su cuenta. **(5 puntos)**

Ejercicio

. La tabla siguiente presenta las notas de aprobación en matemática discreta y programación de 6 estudiantes elegidos al azar de Ingeniería en Computación.

Matemática discreta:	35	60	93	65	87	71
Programación:	50	57	95	73	91	80

(c) Determine un intervalo de predicción del 95% para la nota en programación de un estudiante que obtuvo un 80 en matemática discreta. $R/ \quad]66.04931621, 101.7813495[$

Examen

Un empresa de servicio de *delivery* quiere determinar un modelo lineal para estimar los ingresos diarios, en miles de colones, a partir ya sea del número de colaboradores que hacen las entregas o bien el número de clientes. La siguiente tabla resume los resultados obtenidos en una muestra aleatoria.

Ingresos	867	1210	995	760	678
Colaboradores	18	25	20	15	17
Clientes	230	345	324	280	150

c) [**5 puntos**] Realice una estimación por intervalo, con una confianza del 95 %, para el parámetro β calculado en la parte b) de este ejercicio.

$]16.84797248, 86.32444132[$

Gracias por la amable atención!!