Práctica 5: Regresión lineal simple

Ejercicio 1:

Se utiliza regresión lineal para analizar los datos de un estudio donde se investigó la relación que existe entre la temperatura de la superficie de una carretera (x) y la deformación del pavimento (y). El resumen de las cantidades es el siguiente:

$n = 20 \qquad \sum y_i = 12{,}75 \qquad \sum y_i^2 = 8{,}86 \qquad \sum x_i = 1478$

$\sum x_i^2 = 143215{,}8 \qquad \sum x_i y_i = 1083{,}67$

a) Calcular las estimaciones de mínimos cuadrados de la pendiente y la ordenada al origen. Hacer un gráfico de la recta de regresión y estimar σ^2.

b) Utilice la ecuación de la recta ajustada para predecir la deformación del pavimento observada cuando la temperatura de la superficie sea 85°F.

Resolución a)

Las estimaciones de la pendiente y la ordenada al origen son:

$$\hat{\beta}_1 = \frac{S_{xy}}{S_{xx}} \quad \text{con} \quad S_{xx} = \sum x_i^2 - \frac{1}{n}\left(\sum x_i\right)^2 \quad y \quad S_{xy} = \sum x_i y_i - \frac{1}{n}\left(\sum x_i\right)\left(\sum y_i\right)$$

$$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \cdot \bar{x} \quad \text{con} \quad \bar{y} = \frac{1}{n}\sum y_i \quad y \quad \bar{x} = \frac{1}{n}\sum x_i$$

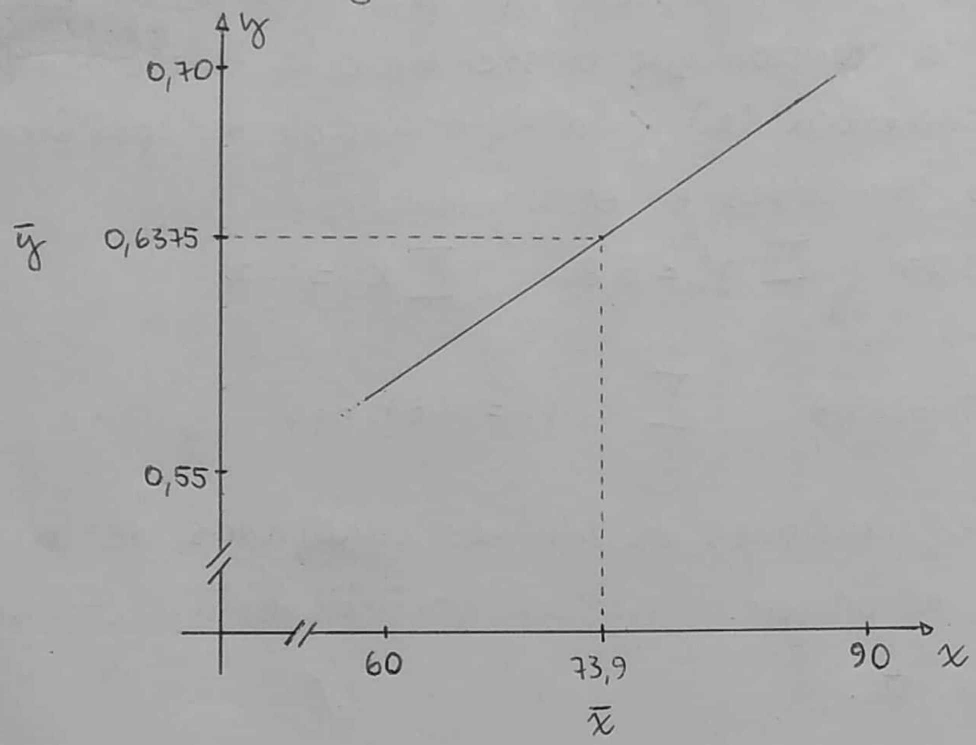
Reemplazando los datos:

$$\left.\begin{aligned} S_{xx} &= 143215{,}8 - \frac{1}{20}\cdot 1478^2 = 33991{,}6 \\ S_{xy} &= 1083{,}67 - \frac{1}{20}\cdot 1478 . 12{,}75 = 141{,}445 \end{aligned}\right\} \Rightarrow \hat{\beta}_1 = \frac{141{,}445}{33991{,}6} = \boxed{4{,}161 . 10^{-3}}$$

Escaneado con CamScanner

$$\left.\begin{array}{l} \bar{y} = \frac{1}{20} \cdot 12,75 = 0,6375 \\ \\ \bar{x} = \frac{1}{20} \cdot 1478 = 73,9 \end{array}\right\} \Longrightarrow \hat{\beta}_0 = 0,6375 - 4,161 \cdot 10^{-3} \cdot 73,9 = \boxed{0,3300}$$

Luego, la recta de regresión estimada es $y(x) = 0,3300 + 4,161 \cdot 10^{-3} \cdot x$



Nótese que como no conocemos el rango de x no podemos determinar el rango de validez de la recta estimada. Sin embargo, sabemos por cómo está construida que el rango tiene que estar entorno al valor de $\bar{x}$.

Por otro lado, la estimación de σ^2 es:

$$\hat{\sigma}^2 = \frac{SS_R}{n-2} \quad \text{con } SS_R = S_{yy} - \frac{S_{xy}^2}{S_{xx}} \quad \text{y} \quad S_{yy} = \sum y_i^2 - \frac{1}{n}\left(\sum y_i\right)^2$$

Reemplazando los datos:

$$S_{yy} = 8,86 - \frac{1}{20} \cdot 12,75^2 = 0,7319$$

$$SS_R = 0,7319 - \frac{141,445^2}{33991,6} = 0,1433$$

Luego:

$$\hat{\sigma}^2 = \frac{0,1433}{18} = \boxed{7,961 \cdot 10^{-3}}$$

Escaneado con CamScanner

b) Reemplazando $x = 85$ en la ecuación de la recta estimada:

$$y_{(85)} = 0{,}3300 + 4{,}161 . 10^{-3} . 85 = 0{,}6837$$

Así, la deformación del pavimento predicha para una temperatura de 85°F es 0,6837, en su unidad correspondiente.

Escaneado con CamScanner