Ejercicio 2

Equipo

2024-03-26

2. Regresión lineal simple.

Considere el modelo de regresión $y_i = \beta_0 + \beta_1 x_i + \epsilon_i$, donde $E(\epsilon_i) = 0$, $V(\epsilon_i) = \sigma^2$ y $Cov(\epsilon_i, \epsilon_j) = 0$, $\forall i \neq j$; $i, j = 1, ..., n$.

Calcular $V(e_i)$, donde $e_i = y_i - \hat{y}_i$ y $\hat{y}_i = \hat{\beta}_0 + \hat{\beta}_1 x_i$, con $\hat{\beta}_0$ y $\hat{\beta}_1$ los estimadores de los parámetros del modelo.

Hint: Se puede usar que $V(A - B) = V(A) + V(B) - 2Cov(A, B)$ y que $\hat{y}_i$ se puede escribir como una combinación lineal de las $y_{i's}$.

SOLUCIÓN

Como $V(y_i) = V(\beta_0 + \beta_1 x_i + \epsilon_i) = V(\epsilon_i) = \sigma^2$ por ser β_0, β_1 y x_i constantes.

Como $V(\hat{y}_i) = V(\hat{\beta}_0 + \hat{\beta}_1 x_i) = V(\beta_0) + V(\beta_1 x_i) + 2Cov(\beta_0, \beta_1 x_i) = V(\hat{\beta}_0) + x_i^2 V(\hat{\beta}_1) + 2x_i Cov(\hat{\beta}_0, \hat{\beta}_1) = \sigma^2(\frac{1}{n} + \frac{\bar{X}^2}{SSx}) + x_i^2(\frac{\sigma^2}{SSx}) + 2x_i(\frac{-\bar{X}\sigma^2}{SSx}) = \sigma^2(\frac{SSx + n\bar{X}^2}{nSSx} + \frac{x_i^2}{SSx} - \frac{2x_i\bar{X}}{SSx}) = \sigma^2(\frac{1}{n} + \frac{(x_i - \bar{X})^2}{SSx})$, con $SS_x = \sum_{i=1}^n (x_i - \bar{X})^2$.

Como $Cov(y_i, \hat{y}_i) = Cov(y_i, \bar{Y} + \hat{\beta}_1 x_i - \hat{\beta}_1 \bar{X}) = Cov(y_i, \bar{Y}) + Cov(y_i, \hat{\beta}_1 x_i) + Cov(y_i, -\hat{\beta}_1 \bar{X}) = Cov(y_i, \hat{\beta}_0 + \hat{\beta}_1 \bar{X}) + x_i Cov(y_i, \hat{\beta}_1) - \bar{X} Cov(y_i, \hat{\beta}_1) = Cov(y_i, \hat{\beta}_0) + \bar{X} Cov(y_i, \hat{\beta}_1) + x_i Cov(y_i, \hat{\beta}_1) - \bar{X} Cov(y_i, \hat{\beta}_1) = Cov(y_i, \hat{\beta}_0) + x_i Cov(y_i, \hat{\beta}_1) = (\frac{1}{n} - \frac{\bar{X}(x_i - \bar{X})}{SSx})\sigma^2 + x_i(\frac{x_i - \bar{X}}{SSx})\sigma^2 = \sigma^2(\frac{1}{n} - \frac{\bar{X}(x_i - \bar{X})}{SSx} + x_i(\frac{x_i - \bar{X}}{SSx})$.

Entonces:

$$V(e_i) = V(y_i - \hat{y}_i) = V(y_i) + V(\hat{y}_i) - 2Cov(y_i, \hat{y}_i) = \sigma^2 + \sigma^2(\frac{1}{n} + \frac{(x_i - \bar{X})^2}{SS_x}) - 2\sigma^2(\frac{1}{n} - \frac{\bar{X}(x_i - \bar{X})}{SSx} + \frac{x_i(x_i - \bar{X})}{SSx})$$

$$= \sigma^2 + \frac{\sigma^2}{n} + \frac{\sigma^2(x_i - \bar{X})^2}{SSx} - \frac{-2\sigma^2}{n} + \frac{2\sigma^2\bar{X}(x_i - \bar{X})}{SSx} - \frac{2\sigma^2 x_i(x_i - \bar{X})}{SSx} = \sigma^2 + \frac{\sigma^2}{n} + (\frac{-\sigma^2 x_i^2 - \sigma^2\bar{X}^2 + 2\sigma^2\bar{X}x_i}{SSx})$$

$= \sigma^2 + \frac{\sigma^2}{n} - \frac{\sigma^2}{SSx}(x_i - \bar{X})^2$

Además se usaron los siguientes resultados:

$\bar{Y} = \hat{\beta}_0 + \hat{\beta}_1 \bar{X}$

$\hat{y}_i = \hat{\beta}_0 + \hat{\beta}_1 x_i = (\bar{Y} - \bar{X}\hat{\beta}_1) + \hat{\beta}_1 x_i = \bar{Y} + \hat{\beta}_1 x_i - \hat{\beta}_1 \bar{X}$

$V(\hat{\beta}_0) = Cov(\hat{\beta}_0, \hat{\beta}_0) = Cov(\sum_{i=1}^n k_{i_0} y_i, \sum_{j=1}^n k_{j_0} y_j) = \sigma^2 \sum_{i=1}^n k_{i_0}^2 = \sigma^2 \sum_{i=1}^n (\frac{1}{n} - \frac{\bar{X}(x_i - \bar{X})}{SSx})^2 = \sigma^2(\frac{1}{n} + \frac{\bar{X}^2}{SSx})$,

$V(\hat{\beta}_1) = Cov(\hat{\beta}_1, \hat{\beta}_1) = Cov(\sum_{i=1}^n k_{i_1} y_i, \sum_{j=1}^n k_{j_1} y_j) = \sigma^2 \sum_{i=1}^n k_{i_1}^2 = \sigma^2 \sum_{i=1}^n (\frac{x_i - \bar{X}}{SSx})^2 = \frac{\sigma^2}{(SSx)^2} \sum_{i=1}^n (x_i - \bar{X})^2 = \frac{\sigma^2}{SSx}$,

$Cov(\hat{\beta}_0, \hat{\beta}_0) = Cov(\sum_{i=1}^n k_{i_0} y_i, \sum_{j=1}^n k_{j_1} y_j) = \sigma^2 \sum_{i=1}^n k_{i_0} k_{i_1} = \sigma^2 \sum_{i=1}^n (\frac{1}{n} - \frac{\bar{X}(x_i - \bar{X})}{SSx})(\frac{x_i - \bar{X}}{SSx}) = -\frac{\bar{X}\sigma^2}{SSx}$,

$Cov(y_i, \hat{\beta}_0) = Cov(y_i, \sum_{i=1}^n k_{i_0} y_i) = k_{i_0} Cov(y_i, y_i) = k_{i_0} V(y_i) = k_{i_0}\sigma^2 = (\frac{1}{n} - \frac{\bar{X}(x_i - \bar{X})}{SSx})\sigma^2$,

$Cov(y_i, \hat{\beta}_1) = Cov(y_i, \sum_{i=1}^{n} k_{i_1} y_i) = k_{i_1} Cov(y_i, y_i) = k_{i_1} V(y_i) = k_{i_1} \sigma^2 = (\frac{x_i - \bar{X}}{SSx})\sigma^2$,

$\frac{SSx}{(x_i - \bar{X})^2} = \frac{\sum_{i=1}^{n} (x_i - \bar{X})^2}{(x_i - \bar{X})^2} = \sum_{i=1}^{n} 1 = n$