SME0820 - Modelos de Regressão e Aprendizado Supervisionado I - Exercício I

Brenda da Silva Muniz 11811603 Francisco Rosa Dias de Miranda 4402962
Heitor Carvalho Pinheiro 11833351 Mônica Amaral Novelli 11810453

Setembro 2021

Exercício 2

Queremos mostrar que os estimadores $\hat{\beta}_0$ e $\hat{\beta}_1$ são não enviesados.

$$E(\hat{\beta}_1) = E\left(\frac{S_{XY}}{S_{XX}}\right) = E\left(\frac{\sum_{i=1}^n (x_i - \bar{x})y_i}{\sum_{i=1}^n (x_i - \bar{x})x_i}\right) = \frac{\sum_{i=1}^n (x_i - \bar{x})E(y_i)}{\sum_{i=1}^n (x_i - \bar{x})x_i}$$

Temos que $y_i = \beta_0 + \beta_1 x_i + \epsilon_i,\ i = 1, ..., n.$ Assim:

$$E(y_i) = E(\beta_0 + \beta_1 x_i + \epsilon_i) = E(\beta_0) + E(\beta_1 x_i) + E(\epsilon_i) = \beta_0 + \beta_1 x_i$$

Voltando a expressão original, temos que:

$$E(\hat{\beta}_1) = \frac{\sum_{i=1}^n (x_i - \bar{x})(\beta_0 + \beta_1 x_i)}{\sum_{i=1}^n (x_i - \bar{x})x_i} = \frac{\beta_0 \sum_{i=1}^n (x_i - \bar{x})}{\sum_{i=1}^n (x_i - \bar{x})x_i} + \frac{\beta_1 \sum_{i=1}^n (x_i - \bar{x})x_i}{\sum_{i=1}^n (x_i - \bar{x})x_i} = \beta_1.$$

Procedendo de forma análoga para β_0, temos que:

$$E(\hat{\beta}_0) = E(\bar{y} - \hat{\beta}_1 \bar{x}) = E(\bar{y}) - \bar{x}E(\hat{\beta}_1)$$

Resta-nos obter

$$E(\bar{y}) = E\left(\frac{1}{n}\sum_{i=1}^n y_i\right) = \frac{1}{n}\sum_{i=1}^n E(y_i) = \frac{1}{n}\sum_{i=1}^n (\beta_0 + \beta_1 x_i) =$$

$$\frac{n\beta_0}{n} + \beta_1 \frac{\sum_{i=1}^n x_i}{n} = \beta_0 + \beta_1 \bar{x}$$

Segue que:

$$E(\hat{\beta}_0) = E(\bar{y}) - \bar{x}E(\hat{\beta}_1) = (\beta_0 + \beta_1 \bar{x}) - \beta_1 \bar{x} = \beta_0$$

Portanto, $\hat{\beta}_1$ e $\hat{\beta}_0$ são não enviesados.

Exercício 3

Queremos obter o **EMMQ** para o modelo linear simples sem intercepto β_0.

Nesse caso, temos a equação:

$$y_i = \beta_1 x_i + \epsilon_i, \quad i = 1, 2, ..., n.$$

Definindo um estimador $\hat{\beta}_1$ para β_1:

Queremos que $\sum_{i=1}^{n} \epsilon_i^2 = 0$, dado que $\epsilon_i = y_i - \beta_1 x_i$

Teremos que:

$$S(\hat{\beta}_1) = \sum_{i=1}^{n} (y_i - \hat{\beta}_1 x_i)^2 = \sum_{i=1}^{n} y_i^2 - 2\hat{\beta}_1 \sum_{i=1}^{n} y_i x_i + \hat{\beta}_1^2 \sum_{i=1}^{n} x_i^2$$

Minimizando o valor de $S(\hat{\beta}_1)$ e igualando a zero:

$$\frac{\partial S(\hat{\beta}_1)}{\partial \hat{\beta}_1} = 0 \Rightarrow \frac{\partial S(\hat{\beta}_1)}{\partial \hat{\beta}_1} = -2 \sum_{i=1}^{n} y_i x_i + 2\hat{\beta}_1 \sum_{i=1}^{n} x_i^2 = 0$$

Este ponto é, de fato, o de mínimo global, pois:

$$\frac{\partial^2 S(\hat{\beta}_1)}{\partial \hat{\beta}_1^2} = 2 \sum_{i=1}^{n} x_i^2 > 0$$

Logo:

$$\hat{\beta}_1 = \sum_{i=1}^{n} \left(\frac{y_i x_i}{x_i^2} \right)$$

Para mostrar que $\hat{\beta}_1$ é não-viesado, precisamos mostrar que $E(\hat{\beta}_1) = \beta_1$.

Segue que:

$$E(\hat{\beta}_1) = \frac{1}{\sum_{i=1}^{n} x_i^2} \sum_{i=1}^{n} x_i E(y_i) = \frac{1}{\sum_{i=1}^{n} x_i^2} \sum_{i=1}^{n} x_i (\beta_1 x_i) = \frac{\sum_{i=1}^{n} x_i^2}{\sum_{i=1}^{n} x_i^2} \beta_1 = \beta_1$$

$$\therefore \; E(\hat{\beta}_1) = \beta_1$$

Desse modo, o **EMMQ** de $\hat{\beta}_1$ para o modelo linear simples sem o intercepto é não-viesado.

Exercício 4

Queremos mostrar que $SQ_{total} = \sum_{i=1}^n Y_i^2$ no modelo linear simples sem o intercepto pode ser escrito como $SQ_{total} = \beta_1 \sum_{i=1}^n X_i Y_i + SQ_{res}$.

Temos que $\hat{Y}_i = \hat{\beta}_1 X_i$, para $i = 1, 2, ..., n$. Assim:

$$SQ_{res} = \sum_{i=1}^{n}(Y_i - \hat{Y}_i)^2 = \sum_{i=1}^{n}(Y_i^2 - 2Y_i\hat{Y}_i + \hat{Y}_i^2) = \sum_{i=1}^{n} Y_i^2 - 2\sum_{i=1}^{n} Y_i(\hat{\beta}_1 X_i) + \sum_{i=1}^{n}(\hat{\beta}_1 X_i)^2$$

Temos que $SQ_{total} = \sum_{i=1}^n Y_i^2$. Assim, nossa equação fica:

$$\begin{aligned} SQ_{res} &= SQ_{total} - 2\sum_{i=1}^{n} Y_i(\hat{\beta}_1 X_i) + \sum_{i=1}^{n}(\hat{\beta}_1 X_i)^2 \\ &= SQ_{total} - 2\sum_{i=1}^{n} \hat{\beta}_1 X_i Y_i + \sum_{i=1}^{n} \hat{\beta}_1^2 X_i^2 \\ &= SQ_{total} - 2\hat{\beta}_1 \sum_{i=1}^{n} X_i Y_i + \hat{\beta}_1^2 \sum_{i=1}^{n} X_i^2 \end{aligned}$$

Isolando SQ_{total}, temos que:

$$\begin{aligned} SQ_{total} &= SQ_{res} + 2\hat{\beta}_1 \sum_{i=1}^{n} X_i Y_i - \hat{\beta}_1^2 \sum_{i=1}^{n} X_i^2 \\ &= SQ_{res} + \hat{\beta}_1(2\sum_{i=1}^{n} X_i Y_i - \hat{\beta}_1 \sum_{i=1}^{n} X_i^2) \\ &= SQ_{res} + \hat{\beta}_1(2\hat{\beta}_1 \sum_{i=1}^{n} X_i^2 - \hat{\beta}_1 \sum_{i=1}^{n} X_i^2) \\ &= SQ_{res} + \hat{\beta}_1(\hat{\beta}_1 \sum_{i=1}^{n} X_i^2) \\ &= SQ_{res} + \hat{\beta}_1 \sum_{i=1}^{n} X_i Y_i. \end{aligned}$$