

Considere o modelo de regressão linear simples

$$Y_i = \beta_0 + \beta_1 X_i + \epsilon_i, i = 1, \dots, n$$

com  $E(\epsilon_i) = 0$  e  $Var(\epsilon_i) = \sigma^2$ , com  $\epsilon_i$  não correlacionado com  $\epsilon_j$ ,  $i, j = 1, \dots, n$  e  $i \neq j$ .

1. Mostre que

i.  $S_{XX} = \sum_{i=1}^n (X_i - \bar{X})^2 = \sum_{i=1}^n (X_i - \bar{X})X_i = \sum_{i=1}^n X_i^2 - n\bar{X}^2$

ii.  $S_{YY} = \sum_{i=1}^n (Y_i - \bar{Y})^2 = \sum_{i=1}^n (Y_i - \bar{Y})Y_i = \sum_{i=1}^n Y_i^2 - n\bar{Y}^2$

iii.  $S_{XY} = \sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y}) = \sum_{i=1}^n (X_i - \bar{X})Y_i = \sum_{i=1}^n X_i Y_i - n\bar{X}\bar{Y}$

(Dica: Ver Draper & Smith, pág 24-25)

2. Ajuste o modelo de regressão linear simples aos dados de tempo de reação (Dados acuidade, disponíveis na CoteiaWiki) com Y o tempo de reação e X a idade, apresentando todos os cálculos (utilize apenas a calculadora). Em seguida, utilize um pacote estatístico para obter o mesmo ajuste. Se optar pelo R, veja o arquivo Exercicio\_MQ.r). Qual a interpretação das estimativas dos parâmetros obtidas?

3. No modelo de regressão linear simples, mostre que  $Cov(\hat{\beta}_0, \hat{\beta}_1) = -\frac{\sigma^2 \bar{X}}{S_{XX}}$ .

4. Considerando o modelo de regressão linear simples, mostre que  $E(\hat{Y}_i) = E(Y_i)$  e que  $Var(\hat{Y}_i) = \sigma^2 \left[ \frac{1}{n} + \frac{(X_i - \bar{X})^2}{S_{XX}} \right]$ .

5. Mostre que se  $\epsilon_i \sim N(0, \sigma^2)$  para  $i = 1, \dots, n$  e  $\epsilon_i$  e  $\epsilon_j$  são não correlacionados, os estimadores de máxima verossimilhança de  $\beta_0$  e  $\beta_1$  no modelo de regressão linear simples coincidem com os estimadores de mínimos quadrados e são dados por  $\hat{\beta}_0 = \bar{Y} - \hat{\beta}_1 \bar{X}$  e  $\hat{\beta}_1 = \frac{S_{XY}}{S_{XX}}$ .

6. Obtenha intervalos com 95% de confiança para  $\beta_0$  e  $\beta_1$  considerando os dados de tempo de reação com Y o tempo de reação e X a idade. Avalie (separadamente) se  $\beta_1 = 0$  ou se  $\beta_0 = 0$ , apresentando todos os cálculos (utilize apenas a calculadora). Programe em R os intervalos de confiança e os testes de hipóteses acima.