Lista 1 - Gabarito

Prof. Erica Castilho Rodrigues

Disciplina: Análise de Regressão

16 de setembro de 2014

1. Da primeira equação temos $\sum_{i=1}^{n}(Y_i - \hat{\beta}_0 - \hat{\beta}_1 X_i) = 0$

$$\sum_{i=1}^{n}(Y_i - \hat{\beta}_0 - \hat{\beta}_1 X_i) = \sum_{i=1}^{n} Y_i - \hat{\beta}_0 n - \hat{\beta}_0 \sum_{i=1}^{n} 1 + \hat{\beta}_1 \sum_{i} X_i$$

$$= n\frac{\sum_{i=1}^{n} Y_i}{n} - n\hat{\beta}_0 + n\hat{\beta}_1 \frac{\sum_i X_i}{n} = n\overline{Y} - n\hat{\beta}_0 - n\hat{\beta}_1 \overline{X} = 0$$

logo

$$\hat{\beta}_0 = \overline{Y} - \hat{\beta}_1 \overline{X}$$

Da segunda equação temos que $\sum_{i=1}^{n}(Y_i - \hat{\beta}_0 - \hat{\beta}_1 X_i)X_i = 0$

$$\sum_{i=1}^{n}(Y_i - \hat{\beta}_0 - \hat{\beta}_1 X_i)X_i = \sum_{i} Y_i X_i - \hat{\beta}_0 \sum_{i} X_i - \hat{\beta}_1 \sum_{i} X_i^2$$

substituindo $\hat{\beta}_0$

$$= \sum_{i} Y_i X_i - (\overline{Y} - \hat{\beta}_1 \overline{X}) \sum_{i} X_i - \hat{\beta}_1 \sum_{i} X_i^2 = 0$$

logo

$$\hat{\beta}_1 = \frac{\sum_i Y_i X_i - n\overline{Y}\overline{X}}{\sum_i X_i^2 - n\overline{X}^2}$$

2. Resolvido em sala de aula.

3. **a)** $\sum_{i=1}^{n} e_i = 0$

$$\sum_{i=1}^{n} e_i = \sum_{i=1}^{n}((Y_i - \hat{Y}_i) = \sum_{i=1}^{n}(Y_i - \hat{\beta}_0 - \hat{\beta}_1 X_i) = 0$$

pois essa é a primeira equação normal.

b) $\sum_{i=1}^n X_i e_i = 0$

$$\sum_{i=1}^n X_i e_i = \sum_{i=1}^n X_i(Y_i - \hat{Y}_i) = \sum_{i=1}^n X_i(Y_i - \hat{\beta}_0 - \hat{\beta}_1 X_i) = 0$$

pois essa é a segunda equação normal.

c) $\sum_{i=1}^n \hat{Y}_i e_i = 0$

$$\sum_{i=1}^n \hat{Y}_i e_i = \sum_{i=1}^n (\hat{\beta}_0 + \hat{\beta}_1 X_i) e_i = \hat{\beta}_0 \sum_{i=1}^n e_i + \hat{\beta}_1 \sum_{i=1}^n X_i e_i = 0$$

pois dos items anteriores temos que $\sum_{i=1}^n e_i = 0$ e $\sum_{i=1}^n X_i e_i = 0$.

d) $\overline{\hat{Y}} = \overline{Y}$, com $\overline{\hat{Y}} = \sum_i \hat{Y}_i / n$

$$\overline{\hat{Y}} = \sum_i \hat{Y}_i / n = \frac{1}{n} \sum_{i=1}^n (Y_i - e_i) = \frac{\sum_i Y_i}{n} - \frac{\sum_i e_i}{n} = \overline{Y}$$

pois vimos que $\sum_i e_i = 0$.

e) A reta de mínimos quadrados passa pelos pontos $(\overline{X}, \overline{Y})$. Fazendo $X = \overline{X}$ na equação da reta estimada $Y = \hat{\beta}_0 + \hat{\beta}_1 X$ e substituindo a expressão de $\hat{\beta}_0$ temos que

$$\hat{\beta}_0 + \hat{\beta}_1 \overline{X} = (\overline{Y} - \hat{\beta}_1 \overline{X}) + \hat{\beta}_1 \overline{X} = \overline{Y} \; .$$

f) $\sum_{i=1}^n (Y_i - \overline{Y})(X_i - \overline{X}) = \sum_i Y_i (X_i - \overline{X})$

$$\sum_{i=1}^n (Y_i - \overline{Y})(X_i - \overline{X}) = \sum_i Y_i (X_i - \overline{X}) - \overline{Y} \sum_i (X_i - \overline{X})$$

mas temos que $\sum_i (X_i - \overline{X}) = \sum_i X_i - n\overline{X} =$ e ficamos com

$$\sum_{i=1}^n (Y_i - \overline{Y})(X_i - \overline{X}) = \sum_i Y_i (X_i - \overline{X}) \; .$$

g) $\sum_{i=1}^n (\hat{Y}_i - \overline{Y})(Y_i - \hat{Y}_i) = 0$

$$\sum_{i=1}^n (\hat{Y}_i - \overline{Y})(Y_i - \hat{Y}_i) = \sum_{i=1}^n (\hat{Y}_i - \overline{Y}) e_i = \sum_{i=1}^n \hat{Y}_i e_i - \overline{Y} \sum_i e_i = 0$$

pois do item (c) temos que $\sum_{i=1}^n \hat{Y}_i e_i = 0$ e do item (a) temos que $\sum_i e_i = 0$.

h) $\sum_{i=1}^{n}(\hat{Y}_i - \overline{Y})^2 = (\hat{\beta}_1)^2 S_{XX}$

$$\sum_{i=1}^{n}(\hat{Y}_i - \overline{Y})^2 = \sum_{i=1}^{n}(\hat{\beta}_0 + \hat{\beta}_1 X_i - \overline{Y})^2$$

substituindo $\hat{\beta}_0$ ficamos com

$$\sum_{i=1}^{n}(\overline{Y} - \hat{\beta}_1\overline{X} + \hat{\beta}_1 X_i - \overline{Y})^2 = \sum_{i=1}^{n}(-\hat{\beta}_1\overline{X} + \hat{\beta}_1 X_i)^2 = \hat{\beta}_1 \sum_{i=1}^{n}(X_i - \overline{X})^2 = \hat{\beta}_1 S_{XX}$$

4. As constantes β_0 e β_1 representam o intercepto (coeficiente linear) e a inclinação (coeficiente angular) da equação da reta de regressão, respectivamente. O valor de β_1 é interpretado como a mudança esperada na resposta Y, quando aumenta uma unidade da regressora X . O valor de β_0 é interpretado como o valor esperado de Y quando a regressora X assume o valor zero, o que pode não ter sentido na prática. A variável ϵ representa o erro atribuído ao modelo, considerado aleatório.

5. Sabemos que

$$E(Y|x) = \beta_0 + \beta_1 x \quad Var(Y|x) = \sigma^2 \quad Y|x \sim N(\beta_0 + \beta_1 x, \sigma^2)$$

Para $x = 4$

$$E(Y|x) = \beta_0 + \beta_1 4 \quad Var(Y|x) = 9 \quad Y|x \sim N(\beta_0 + \beta_1 x, \sigma^2)$$

ou seja

$$E(Y|x) = -2 + (0,5)(4) \quad Var(Y|x) = 9 \quad Y|x = 4 \sim N(0,9)\,.$$

Para $x = 3$

$$E(Y|x) = \beta_0 + \beta_1 3 \quad Var(Y|x) = 9 \quad Y|x \sim N(\beta_0 + \beta_1 x, \sigma^2)$$

ou seja

$$E(Y|x) = -2 + (0,5)(3) \quad Var(Y|x) = 9 \quad Y|x = 4 \sim N(-0,5;9)\,.$$

Para $x = 5$

$$E(Y|x) = \beta_0 + \beta_1 5 \quad Var(Y|x) = 9 \quad Y|x \sim N(\beta_0 + \beta_1 x, \sigma^2)$$

ou seja

$$E(Y|x) = -2 + (0,5)(5) \quad Var(Y|x) = 9 \quad Y|x = 4 \sim N(0,5; 9)\ .$$

Obervamos que a média de Y aumenta quando x aumenta e a variância não depende de x.