Regressão Linear

Projeto 3 - 2ª entrega - Thiago Carletti
Martim Ferreira } 2A

a) Sabemos que:

$$\sum_i \epsilon_i^2 = \sum_i [y_i - (\beta_0 + \beta_1 x_i)]^2 =$$

$$= \sum_i y_i^2 - 2y_i(\beta_0 + \beta_1 x_i) + (\beta_0 + \beta_1 x_i)^2 =$$

$$= \sum_i y_i^2 - 2\beta_0 y_i - 2\beta_1 y_1 x_1 + \beta_0^2 + 2\beta_0 \beta_1 x_i + \beta_1^2 x_1^2$$

Considerando que as médias de 'x' e 'y' estão contidas na reta, temos:

$$\sum y_i^2 = y_1 + y_2 + \dots + y_n \rightarrow \frac{\sum_i y_i^2}{n} = \overline{y^2} \rightarrow \sum y_i^2 = \overline{y^2} \cdot n$$

$$\sum x_i y_i = \overline{xy} \cdot n$$

$$\sum_i x_i = \overline{x} \cdot n$$

$$\sum y_i = \overline{y} \cdot n$$

$$\sum x^2 = \overline{x^2} \cdot n$$

Portanto

$$(n\overline{y^2}) - 2\beta_1 n \overline{xy} - 2\beta_0 n \overline{y} + \beta_1^2 n \overline{x^2} + 2\beta_1 \beta_0 n \overline{x} + n\beta_0^2 = SQ$$

Considerando o ponto de menor erro como o vértice das parábolas, temos para β_1:

$$\frac{\partial SQ}{\partial \beta_1} = -2n\overline{xy} + 2\beta_1 n\overline{x^2} + 2\beta_0 n\overline{x} = 0 \quad \div(2n)$$

$$-\overline{xy} + \beta_1\overline{x^2} + \beta_0\overline{x} = 0$$

$$\beta_1\overline{x^2} + \beta_0\overline{x} = \overline{xy} \quad \div(\overline{x})$$

$$\beta_1\frac{\overline{x^2}}{\overline{x}} + \beta_0 = \frac{\overline{xy}}{\overline{x}} \quad ①$$

Já para β_0:

$$\frac{\partial SQ}{\partial \beta_0} = -2n\overline{y} + 2\beta_1 n\overline{x} + 2n\beta_0 = 0 \quad \div(2n)$$

$$-\overline{y} + \beta_1\overline{x} + \beta_0 = 0 \quad ②$$

$$\beta_1\overline{x} + \beta_0 = \overline{y}$$

Relacionando ① e ②, temos:

$$\beta_1\left(\overline{x} - \frac{\overline{x^2}}{\overline{x}}\right) = \overline{y} - \frac{\overline{xy}}{\overline{x}} \qquad \beta_0 = \overline{y} - \beta_1\overline{x}$$

$$\beta_1 = \frac{\overline{y} - \frac{\overline{xy}}{\overline{x}}}{\overline{x} - \frac{\overline{x^2}}{\overline{x}}} \cdot \left(\frac{\overline{x}}{\overline{x}}\right) \qquad \beta_0 = \overline{y} - \left(\frac{\overline{x}.\overline{y} - \overline{xy}}{(\overline{x})^2 - \overline{x^2}}\right)\cdot\overline{x}$$

$$\beta_1 = \frac{\overline{x}\cdot\overline{y} - \overline{xy}}{(\overline{x})^2 - \overline{x^2}}$$

b) Assumindo a regressão como uma normal, a média é zero e a variância é dada por σ^2, que podem ser verificadar por meio de um teste f.

c) $H_0: \beta_1 = 0$
$H_A: \beta_1 \neq 0$

No caso, recusar β_0 demonstra que há uma relação entre 'x' e 'y'.

d) Realizar uma regressão múltipla é possível, realizando as mesmas suposições feitas no item 'b', com um modelo dado por:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \varepsilon,$$

por exemplo. Já o teste de hipóteses seria dado por:

$H_0: \beta_j \neq 0$; $j = 0, 1, 2, \ldots, n$
$H_j: \beta_j = 0$