Projeto 3 – 2ª Etapa

Pedro de la Peña e Luigi Noronha

a)

Entrega 2 22/11'16

a) $\hat{y}_i = \hat{\beta}_0 + \hat{\beta}_i X_i$ / $y_i = \hat{\beta}_0 + \hat{\beta}_1 X_i + \varepsilon_i$

$\varepsilon_i = y_i - \hat{y}_i$ → Somatória dos erros Quadrados = SEQ

$SEQ \Rightarrow \sum_{i=1}^{n} \varepsilon_i^2 = \sum_{i=1}^{n} (y_i - \hat{y}_i)^2 = \sum_{i=1}^{n} (y_i - \hat{\beta}_0 - \hat{\beta}_1 X_i)^2$

$\boxed{\frac{dSEQ}{d\hat{\beta}_0} = 0}$ (A) $\boxed{\frac{dSEQ}{d\hat{\beta}_i} = 0}$ (B)

A $\frac{dSEQ}{d\hat{\beta}_0} = -2 \cdot \sum_{i=1}^{n} (y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i)$

$\frac{dSEQ}{d\hat{\beta}_0} = -2 \left[\sum_{i=1}^{n} y_i - n \hat{\beta}_0 - \hat{\beta}_1 \sum_{i=1}^{n} X_i \right] \div n$

$= -2n \left\{ \frac{\sum y_i}{n} - \frac{n \hat{\beta}_0}{n} - \hat{\beta}_1 \cdot \frac{\sum x}{n} \right\} = -2 \cdot n \left\{ \bar{y} - \hat{\beta}_0 - \hat{\beta}\right.$

$= 0 \rightarrow \hat{\beta}_0 = -\hat{\beta}_1 \bar{X} + \bar{y}$ (1)

B $\frac{dSEQ}{d\beta_1} = -2 \sum_{i=1}^{n} x_i (y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i) = 0$

$= -2 \left\{ \sum x_i y_i - \hat{\beta}_0 \sum x_i - \hat{\beta}_1 \sum x_i^2 \right\} = 0$ (2)

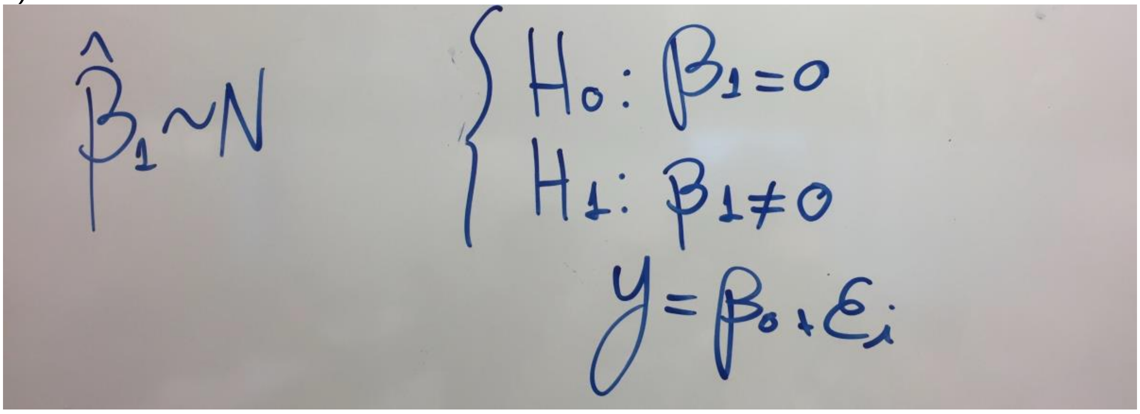
Substituindo 1 em 2:

$\sum_{i=1}^{n} x_i y_i + (\hat{\beta}_1 \bar{x} - \bar{y}) \cdot \sum_{i=1}^{n} x_i - \hat{\beta}_1 \cdot \sum_{i=1}^{n} X_i^2 = 0$

$\hat{\beta}_i = \bar{y} \sum x_i - \sum x_i y_i = \sum x_i y_i - n \bar{y} \bar{x} = \frac{cov(x, y)}{var(x)}$

b)Os erros seguem distribuição Normal, com valor esperado igual a zero e com variância constante. Além disso, os erros são todos independentes entre si. Na prática, todas essas suposições podem ser verificadas através da visualização gráfica das duas variáveis.

c)



A não rejeição de H_0 implica a ausência de relação entre as duas variáveis, enquanto sua rejeição implica a existência dessa relação.

d) É possível. O modelo para regressão múltipla fica como demonstrado na imagem abaixo e as suposições feitas para regressões simples são todas validas para ele. Entretanto, se a correlação entre as duas variáveis explicativas for muito forte, as estimativas dos parâmetros do modelo são prejudicadas. O número de testes de hipóteses que devem ser realizados corresponde ao número de variáveis explicativas do modelo, e a interpretação da rejeição ou não de H_0 continua a mesma.