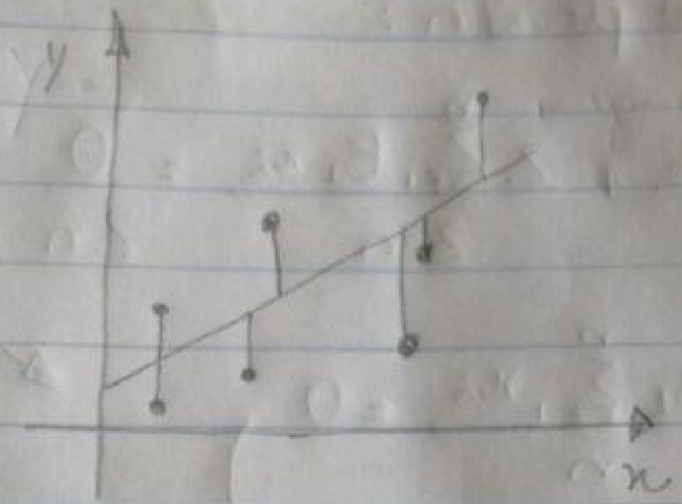
ETAPA 2

a) 

$$Y = \beta_0 + \beta_1 \cdot x_i + \varepsilon_i$$

- Metodo dos mínimos quadrados minimiza $\sum_{i=1}^{n} \varepsilon_i^2$

- Substituindo ε_i por $y - B_0 - B_1 \cdot x_i$

$$S(B_0, B_1) = \sum_{i=1}^{n} (y - B_0 - B_1 x_i)^2$$

$$S(B_0, B_1) = \sum_{i=1} (y - B_0 - B_1 x_i)$$

- A fim de minimizar o erro, devemos derivar $S(B_0, B_1)$ em relação a B_0 e B_1:

- $\dfrac{\partial S}{\partial B_0} = \dfrac{\partial S}{\partial x} \times \dfrac{\partial x}{\partial \alpha}$

- $\dfrac{\partial S}{\partial x} = 2 \sum_{i=1}^{n} (y_i - B_0 - B_1 x_i)$

- $\dfrac{\partial x}{\partial B_0} = -1$

- $\dfrac{\partial S}{\partial B_0} = -2 \sum_{i=1}^{n} (y_i - B_0 - B_1 x_i) = 0$

- $\dfrac{\partial S}{\partial B_1} = -2 \sum_{i=1}^{n} x_i (y_i - B_0 - B_1 x_i) = 0$

Dividindo a expressão por 2n:

$$\frac{-2\sum_{i=1}^{n} y_i}{2n} + \frac{2\sum_{i=1}^{n} B_0}{2n} + \frac{2\sum_{i=1}^{n} B_1 x_i}{2n} = \frac{0}{2n}$$

$$-\frac{\sum_{i=1}^{n} y_i}{n} + \frac{\sum_{i=1}^{n} B_0}{n} + \frac{B_1\sum_{i=1}^{n} x_i}{n} = 0$$

$$-\bar{y} + B_0 + B_1\bar{x} = 0$$

$B_0 = \bar{y} - B_1\bar{x} \longrightarrow$ média amostral x

média amostral y

y

Substituindo esse resultado na segunda expressão:

$$-2\sum_{i=1}^{n} x_i \left(y_i - \bar{y} + B_1 \bar{x} - B_1 x_i\right) = 0$$

$$\sum_{i=1}^{n} \left[x_i \left(y_i - \bar{y}\right) + x_i B_1 \left(\bar{x} - x_i\right)\right] = 0$$

$$\sum_{i=1}^{n} x_i \left(y_i - \bar{y}\right) + B_1 \sum_{i=1}^{n} x_i \left(\bar{x} - x_i\right) = 0$$

$$B_1 = \frac{\sum_{i=1}^{n} x_i \left(y_i - \bar{y}\right)}{\sum_{i=1}^{n} x_i \left(x_i - \bar{x}\right)} \quad \text{ou} \quad B_1 = \frac{\sum_{i=1}^{n} \left(x_i - \bar{x}\right)\left(y_i - \bar{y}\right)}{\sum_{i=1}^{n} \left(x_i - \bar{x}\right)^2}$$

$$B_0 = \bar{y} - B_1 \cdot \bar{x}$$

b) A distribuição, a qual os erros seguem é uma distribuição normal. Seu valor esperado é zero, com uma variância constante.

Para checar essas suposições na prática, utiliza-se a análise ...... do modelo.

c) $H_0 : (\beta_1 = 0)$ → significa que não há correlação entre X e Y

$H_1 : (\beta_1 \neq 0)$ → significa que há correlação entre X e Y

x e y

d) É possível realizar uma regressão com mais do que uma variável explicativa. Sabe-se que na regressão simples, existe apenas o β_0 e o β_1 da equação que é a única possível.

Quando a regressão é múltipla, somente a quantidade de β' irão aumentar na função, que ficará assim: $y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots \beta_n x_n$; sendo n a quantidade de variáveis na regressão.

Já na regressão multipla, a suposição do modelo continuaria a mesma, no entanto, o teste de hipótese irá aumentar de acordo com a quantidade de variáveis.