Gabriel Moreira e Matheus Campbell 2°B

Entrega 2

A)

a)

$$y = \hat{\beta}_1 x + \hat{\beta}_0$$

$$y_i = \hat{\beta}_1 x_i + \beta_0 + \varepsilon$$

$$sq = \sum_{i=1}^{n} \varepsilon_i^2 = \sum_{i=2}^{n} (y_i - \hat{y}_i)^2$$

$$\frac{d\,sq}{da} = 0$$

① $$\frac{d\,sq}{d\hat{\beta}_1} = -2 \sum_{i=1}^{n} (y_i - \hat{\beta}_1 - \beta_0 x_i) = 0$$

$$\frac{-2 \sum_{i=n}^{n} y_i + 2 \sum_{i=n}^{n} \hat{\beta}_1 + 2 \sum_{i=n}^{n} \hat{\beta}_0 x_i}{2n} = \frac{0}{2n}$$

$$-\frac{\sum_{i=n}^{n} y_i}{n} + \frac{\sum_{i=n}^{n} \hat{\beta}_1}{n} + \frac{\sum_{i=n}^{n} \hat{\beta}_0 x_i}{n} = 0$$

$$-\bar{y} + \hat{\beta}_1 + \hat{\beta}_0 \bar{x} = 0$$

$\bar{y}$ = média amostral

$\bar{x}$ = média amostral

$$\boxed{\hat{\beta}_1 = \bar{y} - \hat{\beta}_0 \bar{x}}$$

① $$\frac{d\,sq}{da} = -2 \sum_{i=1}^{n} x_i (y_i - \bar{y} + \hat{\beta}_0 \bar{x} - \hat{\beta}_0 x_i) = 0$$

$$\sum_{i=1}^{n} \left(x_i (y_i - \bar{y}) + x_i \hat{\beta}_0 (\bar{x} - x_i) \right) = 0$$

$$\sum_{i=1}^{n} x_i (y_i - \bar{y}) + \hat{\beta}_0 \sum_{i=1}^{n} x_i (\bar{x} - x_i) = 0$$

$$\boxed{\hat{\beta}_0 = \frac{\sum_{i=1}^{n} x_i (y_i - \bar{y})}{\sum_{i=1}^{n} x_i (\bar{x} - x_i)}}$$

B) É possível supor que os erros e_i na resolução acima possui uma distribuição normal padrão de média ***μ*** / valor esperado igual à 0. Além disso, os erros possuem variância ***σ²*** desconhecida e constante (existe homoscedasticidade)e não possuem correlação, ou seja, $Corr(X,Y)=Cov\frac{(X,Y)}{(DP(X)DP(Y))}=0$. É possível checar, na prática, que tais suposições são verdadeiras fazendo a análise dos resíduos, verificando se esta possui ***σ²*** constante, é linear, não possui pontos fora da curva (outlier), ou que possui alguma independência.

C) Os testes de hipótese ficam desta forma:

$$H_0: \beta_1=0 \rightarrow não\,há\,relação\,entre\,x\,e\,y$$
$$H_1: \beta_1 \neq 0 \rightarrow há\,relação\,entre\,x\,e\,y$$

A rejeição de H_0 mostra que existe uma relação entre *x* e *y*. A não rejeição, por outro lado, mostra que não existe uma relação entre *x* e *y*.

D) Sim, é possível fazer uma regressão múltipla (regressão com mais de duas variáveis explicativas). A equação que descreve a regressão múltipla é:

$$\sum_{i=1}^{n} \varepsilon^2{}_i=\sum_{i=1}^{n}(Y_i-\beta_0-\beta_1 x_{i1}-\beta_2 x_{i2}\ldots\beta_p x_{ip})^2$$

Quanto as suposições: elas são exatamente iguais às da regressão linear simples.

No caso da regressão múltipla, é feito um teste de hipótese para cada variável explicativa.