

## Gabarito da Lista 5 – Tópico 3

a) 1. Hipóteses estatísticas

$$\begin{cases} H_0 : \beta_1 = 0 \\ H_1 : \beta_1 \neq 0 \end{cases}$$

2. Taxa de erro:  $\alpha = 0,05$

3. Cálculo das somas de quadrados

$$SQ_{\text{Total}} = SQY = 3713,88$$

$$SQ_{\text{Reg}} = \beta_1^2 \cdot SQX = 0,9036^2 \times 4152,18 = 3390,55$$

$$SQ_{\text{Resíduo}} = SQ_{\text{Total}} - SQ_{\text{Reg}} = 3713,88 - 3390,55 = 323,33$$

Tabela da análise da variação:

Fontes	GL	SQ	QM	F
Regressão	1	3390,55	3390,55	325,08
Resíduo	31	323,33	10,43	-
Total	32	3713,88	-	-

Como  $f_{\alpha(1; 31)} = 4,17 < f = 325,08$ , rejeitamos  $H_0$ .

Concluimos ao nível de 5% de significância que o coeficiente de regressão populacional difere de zero. Portanto, existe relação linear significativa entre a redução de ST e a redução da DQO.

b) O coeficiente de determinação, denotado por  $r^2$ , expressa a proporção da variação total de Y que é explicada pela regressão, ou seja, pelo efeito linear de X sobre Y. Assim, o  $r^2$  informa sobre a qualidade do ajustamento do modelo aos dados observados. O coeficiente de determinação é o quadrado do coeficiente de correlação linear de Pearson.

c)  $r^2 = 3390,55 / 3713,88 = 0,9129$

Verificamos que 91,29% da variação a redução da DQO é explicada pela variação da redução de ST. Portanto, o modelo de regressão apresenta um ótimo ajustamento aos dados observados.

d)  $IC(\beta_0; 0,95): [0,22; 7,44]$

Concluimos, com 95% de confiança, que o intervalo de 0,22 a 7,44 contém o intercepto populacional, ou seja, quando a redução de ST é zero a redução da DQO populacional está entre 0,22 e 7,44.

e)  $IC(\beta_1; 0,95): [0,80; 1,00]$

Concluimos, com 95% de confiança, que o coeficiente de regressão populacional é coberto pelos limites 0,80 e 1,00. Ou seja, a taxa de redução da DQO para cada unidade que aumenta na redução de ST está no intervalo de 0,8 a 1,00.

f)  $IC(\mu_{y|x=30}; 0,95): 30,94 \pm 1,20$   
 $IC(\mu_{y|x=30}; 0,95): [29,74; 32,14]$

Concluimos, com 95% de confiança, que o intervalo de 29,74 a 32,14 contém a média de Y para  $X=30$ . Isto significa que temos 95% de confiança de que quando a redução de ST é igual a 30, a redução média de DQO está no intervalo de 29,74 a 32,14.

- g)  $IC(y_{x=30}; 0,95): 30,94 \pm 6,7$   
 $IC(y_{x=30}; 0,95): [24,24; 37,64]$

Concluimos, com 95% de confiança, que o intervalo de 24,24 a 37,64 contém uma observação de Y para X=30. Isto significa que temos 95% de confiança de que para uma redução de ST igual a 30 a redução de DQO para uma observação qualquer está entre 24,24 e 37,64.

- h) A amplitude do intervalo na predição é menor que na previsão, ou seja, a precisão é maior na predição.
- i) A extrapolação da equação de regressão para além dos limites dos dados amostrais utilizados na estimativa dos parâmetros do modelo de regressão linear não é recomendada por dois motivos: (1) a amplitude do intervalo de confiança sobre a linha de regressão aumenta à medida que os valores da variável X se afastam da média e (2) a relação entre as variáveis X e Y pode não ser linear para valores que extrapolam os dados utilizados na regressão.