Gabarito da Lista 3 – Tópicos 2 e 3

a) Redução da DQO: variável numérica contínua, escala de razão

Redução de ST: variável numérica contínua, escala de razão

b) Tabela auxiliar.

i	x	y	x^2	y^2	xy
1	3	5	9	25	15
2	7	11	49	121	77
3	11	21	121	441	231
4	15	16	225	256	240
5	18	16	324	256	288
6	27	28	729	784	756
7	29	27	841	729	783
8	30	25	900	625	750
9	30	35	900	1225	1050
10	31	30	961	900	930
11	31	40	961	1600	1240
12	32	32	1024	1024	1024
13	33	34	1089	1156	1122
14	33	32	1089	1024	1056
15	34	34	1156	1156	1156
16	36	37	1296	1369	1332
17	36	38	1296	1444	1368
18	36	34	1296	1156	1224
19	37	36	1369	1296	1332
20	38	38	1444	1444	1444
21	39	37	1521	1369	1443
22	39	36	1521	1296	1404
23	39	45	1521	2025	1755
24	40	39	1600	1521	1560
25	41	41	1681	1681	1681
26	42	40	1764	1600	1680
27	42	44	1764	1936	1848
28	43	37	1849	1369	1591
29	44	44	1936	1936	1936
30	45	46	2025	2116	2070
31	46	46	2116	2116	2116
32	47	49	2209	2401	2303
33	50	51	2500	2601	2550
Soma	1104	1124	41086	41998	41355
Média	33,45	34,06			

c) r = 0,9555

Interpretação: Correlação **forte** e **positiva** entre as variáveis redução da DQO e redução de ST, ou seja, existe uma forte tendência de valores altos de redução da DQO estarem associados a valores altos de redução de ST, e vice-versa.

d) 1. Pressuposição

- Distribuição da variável (X,Y) é normal bivariada;

2. Hipóteses estatísticas

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

3. Taxa de erro

$\alpha = 0{,}05$

4. Estatística do teste

$$t = \frac{0{,}9555}{\sqrt{\dfrac{1 - 0{,}9555^2}{33 - 2}}} = 18{,}04$$

5. Decisão e conclusão

Como $t_{\alpha/2(31)} = 2{,}042 < t = 18{,}04$, rejeitamos H_0.

Concluímos, ao nível de 5% de significância, que o coeficiente de correlação populacional difere de zero. Portanto, existe uma correlação linear positiva entre redução da DQO e redução de ST, Valores altos de redução da DQO estão associados a valores altos de redução de ST.

e)

$$IC(\rho;\ 0{,}95): \left[\frac{exp\left\{2\left[\frac{1}{2}\,ln\left(\frac{1+0{,}9555}{1-0{,}9555}\right) - \frac{1{,}96}{\sqrt{32-3}}\right]\right\} - 1}{exp\left\{2\left[\frac{1}{2}\,ln\left(\frac{1+0{,}9555}{1-0{,}9555}\right) - \frac{1{,}96}{\sqrt{32-3}}\right]\right\} + 1};\frac{exp\left\{2\left[\frac{1}{2}\,ln\left(\frac{1+0{,}9555}{1-0{,}9555}\right) + \frac{1{,}96}{\sqrt{32-3}}\right]\right\} - 1}{exp\left\{2\left[\frac{1}{2}\,ln\left(\frac{1+0{,}9555}{1-0{,}9555}\right) + \frac{1{,}96}{\sqrt{32-3}}\right]\right\} + 1}\right]$$

IC (ρ; 0,95): [0,9110; 0,9780]

Concluímos com 95% de confiança que os limites 0,9110 e 0,9780 compreendem o coeficiente de correlação populacional.

f) $y_i = \beta_0 + \beta_1 x_i + e_i$

onde:

y_i : redução da DQO é a variável resposta;

x_i : redução de ST é a variável preditora;

β_0 : é o intercepto, ou seja, o valor redução da DQO quando a redução de ST é zero;

β_1 : é o coeficiente de regressão, ou seja, a taxa de variação da redução da DQO para cada unidade de redução de ST que aumenta;

e_i: é o erro (variação aleatória não controlável)

Se Y é uma variável aleatória, então, está sujeita a um erro de observação. Este erro (e_i) deverá ser adicionado ao modelo, desde que se admitam como verdadeiras as seguintes pressuposições:

1. Os erros são aleatórios, têm média zero e variância constante, ou seja, $E(e_i)=0$ e $V(e_i)=\sigma^2$.
2. Os erros têm distribuição normal e são independentes entre si.

3. O modelo é adequado para todas as observações, não podendo haver nenhum valor de X que produza um valor de Y discrepante dos demais.

4. A variável X é fixa (não aleatória).

g) É o método de estimação de parâmetros que obtém estimativas de tal forma que a soma de quadrados dos erros seja o menor valor possível. Os estimadores dos coeficientes do modelo de regressão linear simples são obtidos por este método.

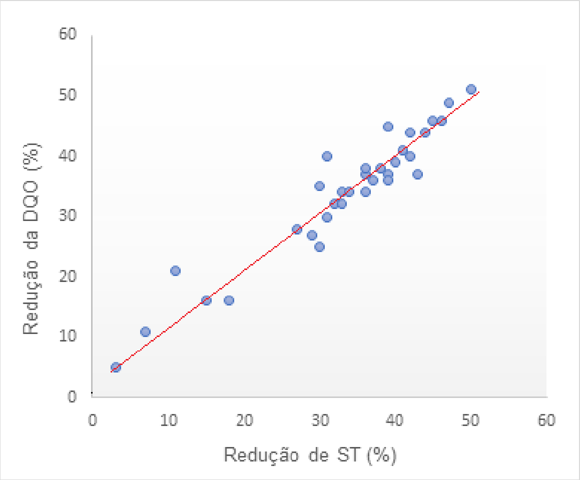
h) $\hat{\beta}_1 = 0{,}9036$

$\hat{\beta}_0 = 3{,}830$

i) Equação da reta ajustada: $\hat{\mu}_i = 3{,}83 + 0{,}9036\, x_i$

$\hat{\mu}_5 = 3{,}83 + 0{,}9036 \times 18 = 20{,}09$

$\hat{e}_5 = 16 - 20{,}09 = -4{,}09$



j) No modelo de regressão simples, o coeficiente de regressão é **proporcional** e tem o **mesmo sinal** do coeficiente de correlação.

$$\hat{\beta}_1 = r_{xy} \frac{\sqrt{SQX.\,SQY}}{SQX}$$

k) 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. Estatística do teste

$$t = \frac{\hat{\beta}_1}{\sqrt{\frac{\frac{\sum \hat{e}_i^2}{n-2}}{SQX}}} = 18{,}03$$

4. Decisão e conclusão

Como $t_{\alpha/2(31)} = 2{,}042 < t = 18{,}03$, rejeitamos H_0.

Concluímos 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.

l) $\hat{\mu}_{y|x=30} = 3{,}83 + 0{,}9036 \times 30 = 30{,}94$