

# ANÁLISE DE DADOS

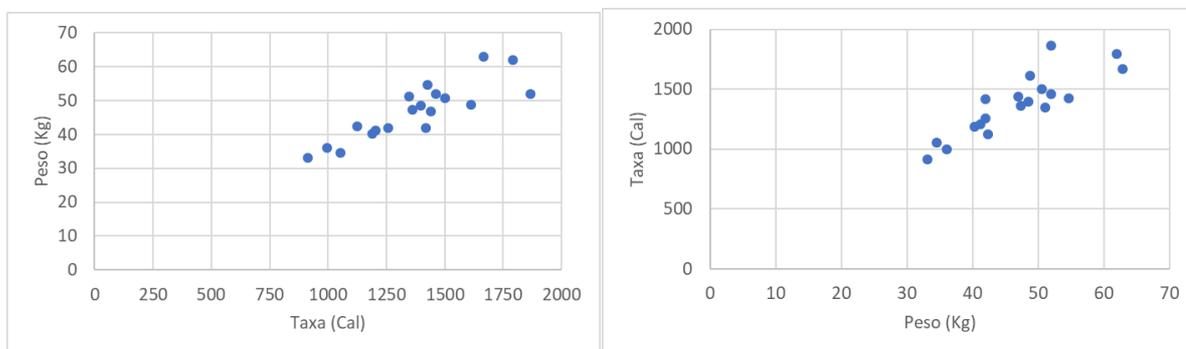
Ano Lectivo 2019/2020

## Exercícios – Regressão Linear Simples

1. Para um mesmo indivíduo, considere as variáveis Peso (Kg) e a Taxa de Metabolismo (calorias consumidas em 24 horas, importante no estudo da variação de peso com a dieta e o exercício físico). Suspeita-se que o peso tem grande influência no consumo de energia. Na tabela seguinte estão representados valores observados daquelas variáveis para 19 indivíduos de ambos os sexos:

| Peso (kg) | Taxa (cal) | Peso (kg) | Taxa (cal) |
|-----------|------------|-----------|------------|
| 62.0      | 1792       | 40.3      | 1189       |
| 62.9      | 1666       | 33.1      | 913        |
| 36.1      | 995        | 51.9      | 1460       |
| 54.6      | 1425       | 42.4      | 1124       |
| 48.5      | 1396       | 34.5      | 1052       |
| 42.0      | 1418       | 51.1      | 1347       |
| 47.4      | 1362       | 41.2      | 1204       |
| 50.6      | 1502       | 51.9      | 1867       |
| 42.0      | 1256       | 46.9      | 1439       |
| 48.7      | 1614       |           |            |

- a) O que verifica na representação gráfica dos dados? Calcule o coeficiente de correlação empírico.

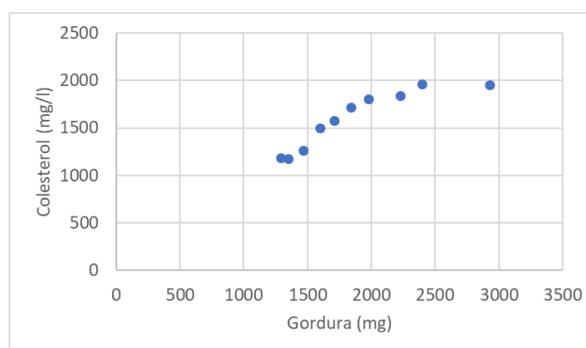


- b) Adapte o modelo que lhe pareça mais conveniente e justifique.

$$\sum_i \text{Peso}_i = 888.1; \sum_i \text{Peso}_i^2 = 42747.03; \sum_i \text{Taxa}_i = 26021; \sum_i \text{Taxa}_i^2 = 36829995; \sum_i \text{Peso}_i \times \text{Taxa}_i = 1249481.2$$

2. Os dados da tabela seguinte representam o nível de colesterol (mg/l) e a média diária de gordura saturada (mg) ingerida por 10 atletas Olímpicos. Pretende averiguar-se da possível influência da gordura saturada ingerida por dia no nível de colesterol:

| Gordura-mg | Colesterol-mg/l |
|------------|-----------------|
| 1290.0     | 1182.0          |
| 1350.0     | 1172.0          |
| 1470.0     | 1264.0          |
| 1600.0     | 1493.0          |
| 1710.0     | 1571.0          |
| 1840.0     | 1711.0          |
| 1980.0     | 1804.0          |
| 2230.0     | 1840.0          |
| 2400.0     | 1956.0          |
| 2930.0     | 1954.0          |



- a) Ajuste aos dados a melhor recta de regressão. O que pode dizer sobre a qualidade do ajustamento? Teste a significância do modelo ajustado.

$$\sum_i \text{Gord}_i = 18800; \sum_i (\text{Gord}_i - \overline{\text{Gord}})^2 = 2411400; \sum_i \text{Col}_i = 15947; \sum_i (\text{Col}_i - \overline{\text{Col}})^2 = 846402.1;$$

$$\sum_i (\text{Gord}_i - \overline{\text{Gord}})(\text{Col}_i - \overline{\text{Col}}) = 1302890; SSE = 142444.917$$

- b) Calcule e represente graficamente os resíduos. Comente. Estime a variância  $\sigma^2$ .  
 c) De acordo com o modelo anterior qual o nível de colesterol esperado para uma quantidade de gordura saturada igual a 1600 mg?

# ANÁLISE DE DADOS

Ano Lectivo 2019/2020

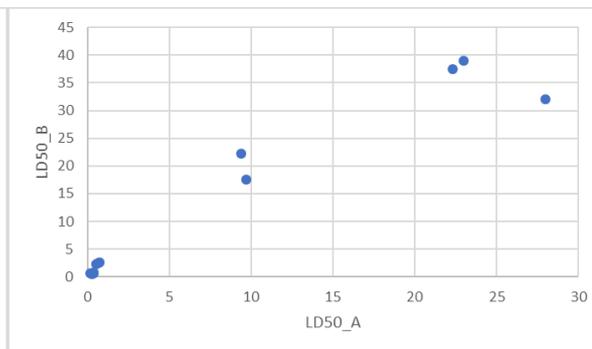
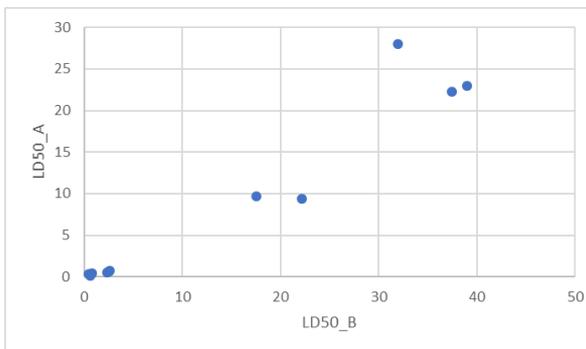
## Exercícios – Regressão Linear Simples

3. Segundo a lei de Hooke, uma força  $x$  aplicada a um corpo cujo comprimento é grande, quando comparado com a sua área transversal, produz uma mudança  $y$  no seu comprimento, proporcional à força, isto é,  $y = \beta x$ , sendo  $\beta$  constante. Os resultados obtidos por um grupo de alunos numa aula de laboratório, para verificarem a lei de Hooke, estão apresentados na tabela seguinte:

| Força (N) | Alteração do comprimento (mm) |
|-----------|-------------------------------|
| 29.4      | 4.25                          |
| 39.2      | 5.25                          |
| 49.0      | 6.50                          |
| 58.8      | 7.85                          |
| 68.6      | 8.75                          |
| 78.4      | 10.00                         |

- a) Represente graficamente os dados e adapte-lhes o modelo  $y = \beta x + \varepsilon$ , usando o método de mínimos quadrados.
- b) Calcule os resíduos e represente-os graficamente. Estime a variância  $\sigma^2$ .
4. Com o intuito de analisar o efeito de certos produtos tóxicos nos peixes, realizaram-se experiências laboratoriais para medir o LD50 (LD significa *Lethal Dose*; LD50 é a quantidade de uma substância, administrada de uma só vez, que causa a morte a 50% (metade) dos animais de um grupo em teste) utilizando dois métodos distintos, A e B. Pretende-se agora encontrar um modelo que relacione os dois tipos de observações, tendo por base os valores que constam da tabela transcrita em baixo.

| Produto Tóxico | LD50 – método A | LD50 – método B |
|----------------|-----------------|-----------------|
| 1              | 23.00           | 39.00           |
| 2              | 22.30           | 37.50           |
| 3              | 9.40            | 22.20           |
| 4              | 9.70            | 17.50           |
| 5              | 0.15            | 0.64            |
| 6              | 0.28            | 0.45            |
| 7              | 0.75            | 2.62            |
| 8              | 0.51            | 2.36            |
| 9              | 28.00           | 32.00           |
| 10             | 0.39            | 0.77            |
| 11             | 0.62            | 2.50            |
| 12             | 0.36            | 0.65            |



Tomando para variável dependente a medição segundo o método A,

- a) estime o modelo e analise-o;

$$\overline{LD50\_A} = 7.96; \overline{LD50\_B} = 13.18; s_{PLD50\_A}^2 = 112.2678; s_{PLD50\_B}^2 = 244.1193;$$

$$r_{PLD50\_A, PLD50\_B} = 0.9636; SSR = 1146.6675$$

- b) Calcule a melhor estimativa do desvio-padrão da variável dependente.
- c) usando o modelo estimado em a), estime agora o valor que se espera obter com o método A, no caso de o valor obtido com o método B ser 14.

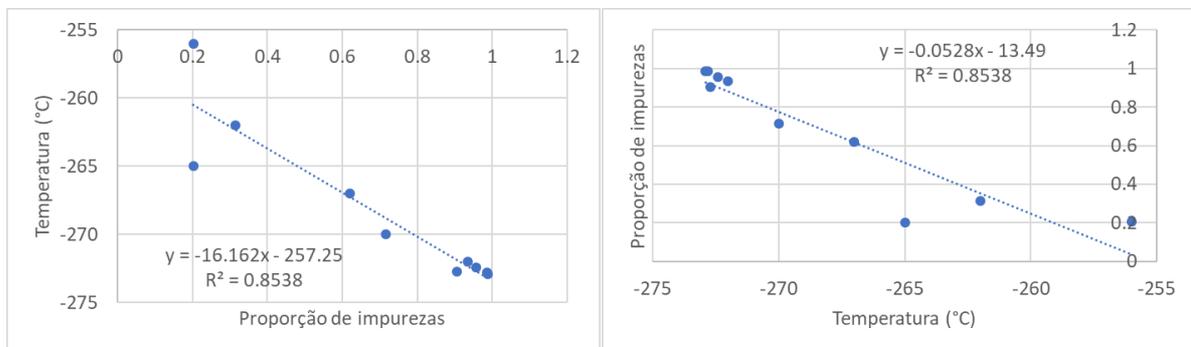
## ANÁLISE DE DADOS

Ano Lectivo 2019/2020

### Exercícios – Regressão Linear Simples

5. Efetuou-se uma experiência com hélio sólido a várias temperaturas próximas do 0 absoluto ( $-273^{\circ}\text{C}$ ), colocando-o junto a uma substância sólida impura e registando a fracção (em peso) de impurezas que passaram através do hélio sólido. Os dados obtidos constam da tabela seguinte:

| Temperatura ( $^{\circ}\text{C}$ ) | Proporção de impurezas que passaram através do hélio sólido |
|------------------------------------|---|
| -262.0                             | 0.315   |
| -265.0                             | 0.202   |
| -256.0                             | 0.204   |
| -267.0                             | 0.620   |
| -270.0                             | 0.715   |
| -272.0                             | 0.935   |
| -272.4                             | 0.957   |
| -272.7                             | 0.906   |
| -272.8                             | 0.985   |
| -272.9                             | 0.987   |



- a) Indique a expressão da recta dos mínimos quadrados que se ajusta a estes dados e pronuncie-se sobre a sua qualidade.
- b) Determine a proporção de impurezas que se estima que passem através do hélio sólido a uma temperatura de  $-250^{\circ}\text{C}$ .

FIM