
MÉTODOS ESTATÍSTICOS - L.EIC

EXERCÍCIOS – 2023/2024

Folha 1

(Estatística Descritiva/Análise Exploratória dos Dados)

Nota: O ficheiro ‘*DadosFolha1.csv*’, está disponível no Moodle, na página da UC, e contém dados necessários à resolução de exercícios desta folha.

1. Num estudo estatístico sobre uma doença neurológica foram obtidos dados relativos a sexo, peso, tipo de tratamento, número de convulsões e classificação da doença (leve, moderada e severa) de diversos doentes na mesma faixa-etária.

Classificar os diversos tipos de dados.

2. [Adaptado de (*)] Relativamente a cada um dos cenários seguintes, identificar a(s) variável(eis) em estudo, a dimensão da amostra e o tipo de dados:

- (a) Um paleontólogo mediu a largura do molar superior em 36 espécimes do extinto *Acropithecus rigidus*.
- (b) Foram registados, o peso à nascença, o dia de nascimento e a nacionalidade da mãe de 65 bebés.
- (c) Foram registados o tipo de sangue e o nível de colesterol em 125 adultos.
- (d) Um biólogo registou o número de folhas em cada uma de 25 plantas.

3. [Adaptado de (**)]

Os dados seguintes dizem respeito a observações da magnitude (na escala de Richter) de 30 sismos na Califórnia [DadosFicha1]:

1.0	8.3	3.1	1.1	5.1
1.2	1.0	4.1	1.1	4.0
2.0	1.9	6.3	1.4	1.3
3.3	2.2	2.3	2.1	2.1
1.4	2.7	2.4	3.0	4.1
5.0	2.2	1.2	7.7	1.5

$$\left[\text{Dados auxiliares: } \sum_{i=1}^{30} x_i = 86.1 \quad \sum_{i=1}^{30} x_i^2 = 357.61 \right]$$

- (a) Representar os dados através de um diagrama de pontos e interpretar.
- (b) Calcular a média, desvio padrão e mediana da amostra.

4. [Adaptado de (**)]

Os dados seguintes dizem respeito a observações do perímetro cefálico (em cm) de 35 recém nascidos do sexo masculino [DadosFicha1]:

33.1	33.4	34.8	33.8	34.7	34.3	35.6
34.5	34.6	34.1	33.9	33.6	34.6	35.2
33.7	35.8	34.2	34.0	34.7	35.2	34.3
33.4	36.0	34.5	36.1	35.1	35.1	34.6
33.7	34.9	34.2	34.2	34.2	35.3	34.2

$$\left[\text{Dados auxiliares: } \sum_{i=1}^{35} x_i = 1207.6 \quad \sum_{i=1}^{35} x_i^2 = 41684.32 \right]$$

- (a) Representar os dados através de um diagrama de pontos e um histograma adequado.
- (b) Calcular a média e desvio padrão da amostra.

5. [Adaptado de (*)]

Um paleontologista mediu a largura do último molar superior em 36 indivíduos da extinta espécie de mamíferos *Acropithecus Rigidus*. Os resultados foram os seguintes [DadosFicha1]:

6.1	5.7	6.0	6.5	6.0	5.7
6.1	5.8	5.9	6.1	6.0	6.2
6.3	6.2	6.1	6.2	6.0	5.7
6.2	5.8	5.7	6.3	6.2	5.7
6.2	6.1	5.9	6.5	5.4	6.7
5.9	6.1	5.9	5.9	6.1	6.1

- (a) Representar graficamente os dados.
- (b) Descrever a forma da distribuição.

6. [Adaptado de (**)]

Num estudo sobre a esquizofrenia, foi medida a atividade de uma determinada enzima nas plaquetas sanguíneas de 18 pacientes. Os resultados (em determinadas unidades), foram os seguintes [DadosFicha1]:

6.8	8.4	8.7	11.9	14.2	18.8
9.9	4.1	9.7	12.7	5.2	7.8
7.8	7.4	7.3	10.6	14.5	10.7

- (a) Construir um histograma para esta amostra, considerando 5 classes.
- (b) Calcular a mediana e os quartis.
- (c) Calcular o intervalo inter-quartis.
- (d) Construir o diagrama de caixa e bigodes dos dados.

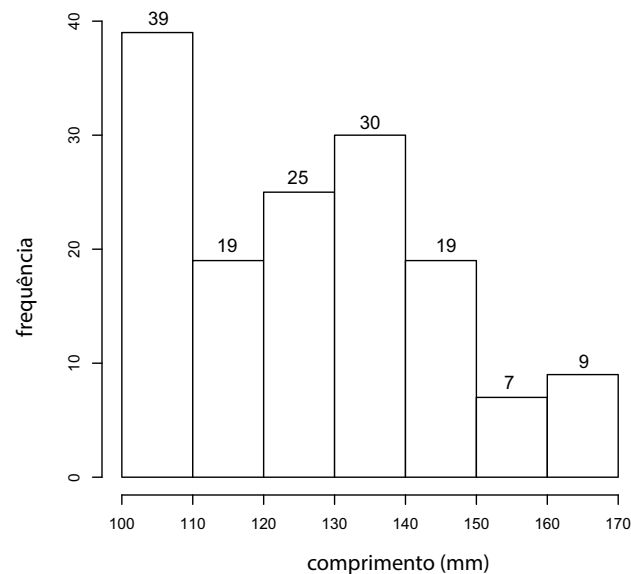
7. [Adaptado de (*)]

Tripanossomas são parasitas causadores de doenças em seres humanos e em animais. Num dos primeiros estudos da morfologia dos tripanossomas, os investigadores mediram os comprimentos de 500 tripanossomas obtidos a partir do sangue de um rato. Os resultados da distribuição de frequências foram os seguintes [DadosFicha1]:

Comprimento (μm)	Frequência (n de indivíduos)	Comprimento (μm)	Frequência (n de indivíduos)
15	1	27	36
16	3	28	41
17	21	29	48
18	27	30	28
19	23	31	43
20	15	32	27
21	10	33	23
22	15	34	10
23	19	35	4
24	21	36	5
25	34	37	1
26	44	38	1

- Construir um histograma dos dados usando 24 classes, isto é, uma classe para cada valor inteiro do comprimento (de 15 até 38).
- Que característica do histograma sugere a interpretação de que os 500 indivíduos são um misto de dois tipos distintos?
- Construir um histograma dos dados usando apenas 6 classes. Comentar o facto deste histograma permitir uma interpretação qualitativamente diferente da inferida na primeira alínea.

8. Os resultados obtidos num estudo sobre o comprimento de uma certa espécie de peixes estão apresentados no seguinte histograma:



- Qual a dimensão da amostra utilizada nesse estudo?
- Quantos peixes apresentaram um comprimento superior ou igual a 15 cm ?
- Estimar a percentagem de peixes na amostra com comprimento inferior a 13.5 cm?
- Alterar a escala das ordenadas para frequências relativas.
- Alterar a escala das ordenadas para densidades.

9. [Adaptado de (**)]

Os dados seguintes dizem respeito à idade (em anos) no momento em que foi feito o primeiro diagnóstico de diabetes tipo 2, de 20 diabéticos selecionados aleatoriamente [DadosFicha1]:

35.5	40.1	47.3	48.9	52.4	$\left[\sum_{i=1}^{20} x_i = 880.4 \quad \sum_{i=1}^{20} x_i^2 = 40854.56 \right]$
39.8	39.3	55.6	40.3	60.9	
30.5	59.8	44.5	36.8	36.6	
42.1	26.2	33.3	65.4	45.1	

- Representar os dados graficamente.
- Calcular a média e o desvio padrão dos dados.

10. [Adaptado de (*)]

Uma bióloga mediu um certo pH em cada um de 24 sapos, obtendo valores típicos:

7.43 7.16 7.51 ...

Calculou uma média de 7.373 e um desvio padrão de 0.129 para estas medidas originais de pH. A seguir, transformou os dados, subtraindo 7 a cada observação e depois multiplicando por 100. Por exemplo 7.43 foi transformado em 43. Quais são a média e o desvio padrão dos dados transformados?

11. [Adaptado de (*)]

A tabela seguinte mostra o tamanho da ninhada (número de leitões que sobrevivem 21 dias), para cada uma de 36 porcas [DadosFicha1].

Tamanho da ninhada	Frequência (no. de porcas)
5	1
6	0
7	2
8	3
9	3
10	9
11	8
12	5
13	3
14	2
Total	36

- Representar os dados usando um gráfico adequado.
- Calcular as medidas de localização da amostra.
- Representar a caixa de bigodes e decidir acerca do enviesamento da distribuição dos dados.

12. [Adaptado de (*)]

Um botânico plantou 15 plantas de pimenta numa estufa. Vinte e um dias depois mediu a altura total (em cm) do caule das plantas e obteve os valores seguintes [DadosFicha1]:

12.4	12.2	13.4	12.1	12.2
11.8	13.5	12.0	10.9	13.2
12.6	11.9	13.1	14.1	12.7

- (a) Calcular a altura média do caule das plantas da amostra, a variância, e o desvio padrão.
 (b) Representar os dados através de um histograma.

13. [Adaptado de (*)]

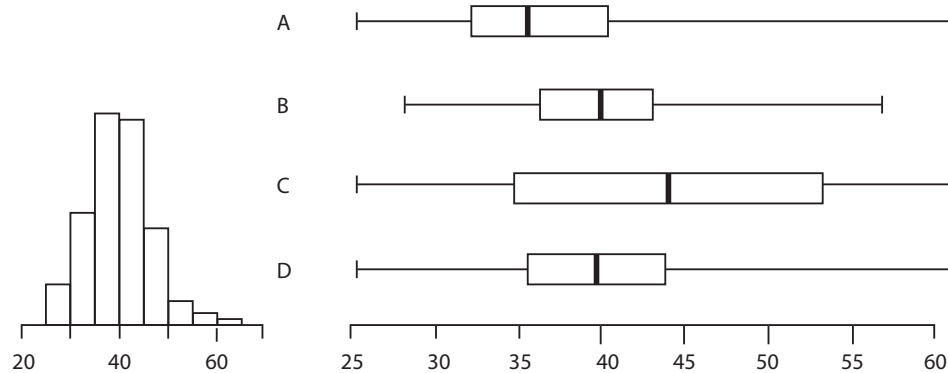
Dez pacientes hipertensos participaram num estudo para avaliar a eficácia de um medicamento para reduzir a tensão arterial. A tabela abaixo mostra a tensão sistólica medida antes e depois de duas semanas de tratamento.

- (a) Comparar as duas distribuições usando um diagrama de caixa e bigodes.
 (b) Representar a amostra num diagrama de dispersão complementando-o com o valor do coeficiente de correlação amostral.

Paciente	1	2	3	4	5	6	7	8	9	10
Antes	172	186	170	205	174	184	178	156	190	168
Depois	159	157	163	207	164	141	182	171	177	138

14. [Adaptado de (*)]

O histograma seguinte representa os mesmos dados do que um dos diagramas de extremos e quartis. Qual?

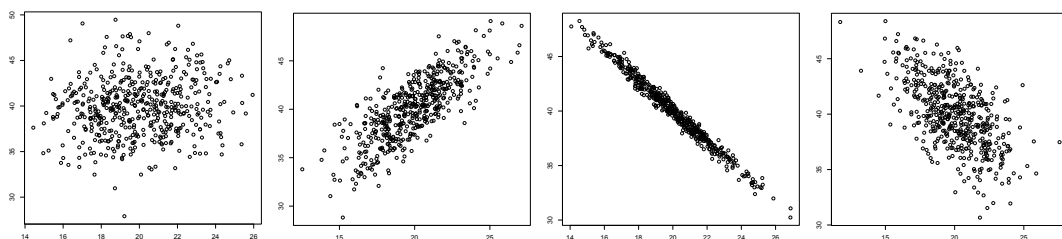


15. Considerar a amostra bivariada:

x	4.6	5.2	3.6	5.0	3.5	4.2	3.9	3.8	4.2	3.9	5.1
y	3.8	4.4	3.2	4.0	4.5	3.5	3.5	3.4	3.9	3.3	4.5

- (a) Representar os dados através de um diagrama de dispersão
 (b) Calcular o coeficiente de correlação de Pearson da amostra.

16. Os valores 0.1, 0.77, -0.6 e -0.99 são os coeficientes de correlação de Pearson de 4 amostras bivariadas representadas abaixo por diagramas de dispersão. Associar os coeficientes aos respetivos diagramas.



17. O aumento de peso, em gramas, em dois conjuntos de animais submetidos a diferentes dietas foi:

Grupo A	56	67	42	48	55	61	52	39	47	58	50	40	59	62	44	57
Grupo B	78	34	37	72	58	68	27	55	65	40	75	33	66			

Utilizar os diagramas de caixa e bigodes no estudo da diferença entre os dois tipos de dieta.

(*) *Statistics for the Life Sciences*, Samuels, Witmer & Schaffner, PrenticeHall, (2012)

(**) *Introduction to Statistics*, Milton, McTeer, Corbet, McGraw-Hill, (1997)
