

Licenciatura em Engenharia Informática (LEI)

2024/2025

Matemática Computacional (MATCP)

CAPÍTULO 2

Estatística Descritiva

EXERCÍCIOS

CARACTERIZAÇÃO, ORGANIZAÇÃO E REPRESENTAÇÃO DOS DADOS. MEDIDAS DESCRITIVAS.

1. Realizou-se um estudo de uma amostra de 104 doentes renais, registando-se, para cada um, o tempo (meses) de hemodiaise antes da realização do transplante, obtendo-se a tabela de frequências de dados classificados:

Tempo hemodiálise	Freq. abs.	Freq. rel.
0 – 15	9	8.7%
15 – 30	35	33.7%
30 – 45	20	19.2%
45 – 60	20	19.2%
60 – 75	7	6.7%
75 – 90	4	3.8%
90 – 105	5	4.8%
105 – 120	1	1%
120 – 135	1	1%
135 – 150	2	1.9%
Total	104	100%

Faça um programa que calcule, em média, o número de meses que os pacientes têm de esperar para a realização do transplante. O programa deverá ler,

- do teclado, o limite inferior e superior da primeira classe e o número de classes;
- de um ficheiro *Excel* (.csv) as frequências absolutas para cada classe.

O programa deverá ter como saída a média pedida (com uma casa decimal).

2. A tabela mostra a distribuição de cargas máximas (em toneladas) suportadas por um tipo de cabos fabricados por uma companhia. Sem recorrer à programação em Python, responda à perguntas.

Carga máxima (toneladas)	Número de cabos
20 – 20.5	25
20.5 – 21	48
21 – 21.5	60
21.5 – 22	53
22 – 22.5	28
22.5 – 23	18
23 – 23.5	9
23.5 – 24	6
24 – 24.5	2
24.5 – 25	1
Total	250

2.1 Determine:

- A média, a mediana e a classe modal.
- A variância, o desvio padrão e o coeficiente de variação.

2.2 A partir da análise da tabela de frequências e sem efetuar qualquer cálculo, indique o sinal do coeficiente de assimetria e justifique.

2.3 Determine o valor do coeficiente de curtose e interprete o resultado obtido.

3. Uma empresa pretende determinar se o grau de satisfação dos seus clientes é aceitável. Para isso, encomendou um estudo para analisar, entre outros parâmetros, o tempo (em segundos) entre reclamações que chegam à sua central telefónica. Os dados recolhidos estão no ficheiro "*Dados_Cap2_Exercicio3_csv.csv*".
 - 3.1 Classifique a variável em estudo quanto ao tipo: qualitativa/quantitativa, discreta/contínua.
 - 3.2 Determine o valor máximo e o valor mínimo do conjunto de dados.
 - 3.3 Escolha o número de classes pretendido.
 - 3.4 Estime a amplitude das classes e decida o seu valor.
 - 3.5 Proceda à organização dos dados construindo uma tabela frequências (dados classificados) contendo as frequências absolutas, relativas e absolutas acumuladas.
 - 3.6 Para o número de classes obtido, obtenha um histograma que representa graficamente os dados da tabela.
 - 3.7 Determine:
 - i. A média, a mediana e os quartis q_1 e q_3 .
 - ii. A variância, o desvio padrão e o coeficiente de variação.
 - 3.8 Determine os valores dos coeficientes de assimetria e de curtose e interprete os resultados obtidos.
 - 3.9 Crie uma *boxplot* associada aos dados e retire as principais conclusões sobre o conjunto de dados que são evidentes na *boxplot*.

4. Um governo europeu, estando a pensar reformular a política de Saúde Pública do seu país, comparou os custos médios diários (em euros) do tratamento de doentes internados, em 100 dos seus maiores hospitais. O registo dos dados obtidos foi feito no ficheiro "*Dados_Cap2_Exercicio4_csv.csv*".
 - 4.1 Obtenha histogramas que representam graficamente os dados da tabela, para 4, 8 e 16 classes. Mostre os histogramas em linha com o título identificativo do número de classes.
 - 4.2 Determine:
 - i. A média, a mediana, a moda e a amplitude interquartil.
 - ii. A variância, o desvio padrão e o coeficiente de variação.
 - 4.3 Determine os valores dos coeficientes de assimetria e de curtose e interprete os resultados obtidos.
 - 4.4 Crie uma *boxplot* associada aos dados e retire as principais conclusões sobre o conjunto de dados que são evidentes na *boxplot*.
5. O ficheiro "*Dados_Cap2_Exercicio5_csv.csv*" contém as variações de cotação das ações de um grupo de 100 empresas cotadas numa Bolsa de Valores, de uma cidade europeia.
 - 5.1 Obtenha um histograma que representa graficamente os dados da tabela (com número de classes automático).
 - 5.2 Determine a média, a mediana e a moda.
 - 5.3 Determine o percentil p_{10} e interprete o resultado.
 - 5.4 Crie uma *boxplot* associada aos dados e retire as principais conclusões sobre o conjunto de dados que são evidentes na *boxplot*.

6. Na tabela seguinte mostra-se o número de filhos que os 46 presidentes dos Estados Unidos tiveram (de George Washington - 1789 a 1797, sem filhos a Joe Biden - desde 2021, com quatro filhos).

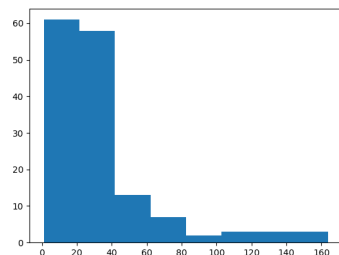
0	5	6	0	2	4	0	4	10	14
0	6	2	3	0	4	5	4	8	7
3	5	3	5	2	6	3	3	0	2
2	6	1	2	3	2	2	4	4	4
6	1	2	2	5	4				

- 6.1 Classifique a variável em estudo quanto ao tipo: qualitativa/quantitativa, discreta/contínua.
- 6.2 Suponha que se opta por 15 classes (0, 1,...,14).
- Crie uma tabela de frequências absolutas e relativas.
 - Obtenha o gráfico de barras que representa os dados da tabela.
7. Uma conhecida marca de automóveis encomendou um estudo de mercado para se concluir qual a cor preferida dos portugueses para os seus automóveis de entre as cinco cores: {branco, preto, vermelho, verde, cinzento}. O resultado do inquérito a 200 portugueses foi registado no ficheiro "*Dados_Cap2_Exercicio7_csv.csv*".
- 7.1 Classifique a variável em estudo quanto ao tipo: qualitativa/quantitativa, discreta/contínua.
- 7.2 Crie uma tabela de frequências.
- 7.3 Obtenha o gráfico circular que representa os dados da tabela.

SOLUÇÕES DOS EXERCÍCIOS

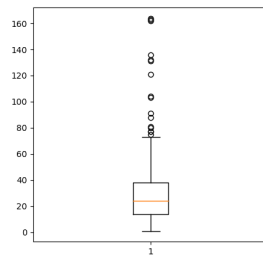
1. $\bar{x} = 43.0$ meses; **2.1.i** $\bar{x} \approx 21.55$, $\tilde{x} \approx 21.43$,
 classe modal: $[21, 21.5[$; **2.1.ii** $s^2 \approx 0.80$, $s \approx 0.90$, $c_v \approx 0.04$;
2.2 Como a distribuição de frequências é enviesada à direita ou
 assimétrica positiva, $a_3 > 0$; **2.3** $a_4 \approx 3.49$, valor superior a 3,
 logo a distribuição é mais esguia, caudas mais pesadas do que a
 distribuição Normal; **3.1** Variável quantitativa contínua;
3.2 Valor máximo = 164, valor mínimo = 1; **3.3** O número de
 classes = 8; **3.4** Amplitude = 20.4; **3.5**

Classes	Freq. abs.	Freq. rel.	Freq. rel. acum.
1.0–21.4	61	40.7%	61
21.4–41.8	58	38.7%	119
41.8–62.2	13	8.7%	132
62.2–82.6	7	4.7%	139
82.6–103.0	3	2.0%	142
103.0–123.4	2	1.3%	144
123.4–143.8	3	2.0%	147
143.8–164.2	3	2.0%	150



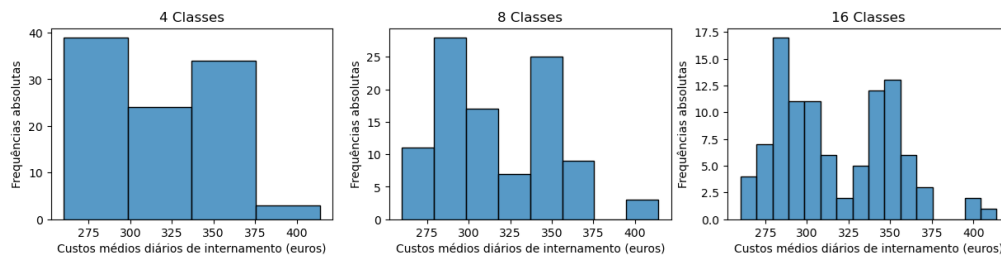
3.6

3.7.i $\bar{x} \approx 33.5$, $\tilde{x} \approx 24.0$,
 $q_1 = 14.0$, $q_3 = 38.2$; **3.7.ii** $s^2 \approx 1051.6$, $s \approx 32.4$, $c_v \approx 1.0$;
3.8 $a_3 = 2.1645 > 0$, assim a distribuição de frequências é en-
 viesada à direita (como é evidente no histograma),
 $a_4 = 8.0360 > 3$, assim a distribuição de frequências é mais
 achatada do que a distribuição Normal;

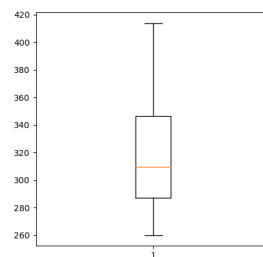


3.9 A análise do gráfico confirma que 50% das observações estão entre $q_1 = 14.0$ e $q_3 = 38.2$, confirma o valor da mediana e identifica um número significativo de pontos isolados (*outliers*), este facto também é confirmado pelo valor elevado do desvio padrão;

4.1

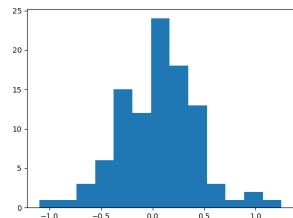


4.2.i $\bar{x} \approx 317.6$, $\tilde{x} \approx 309.5.0$, $\text{Moda} = 283$, amplitude interquartil $q_3 - q_1 = 59.8$; **4.2.ii** $s^2 \approx 1179.9$, $s \approx 34.4$, $c_v \approx 0.1$; **4.3** $a_3 = 0.4162 > 0$, assim a distribuição de frequências é ensiesada à direita, $a_4 = 2.3924 < 3$, assim a distribuição de frequências é mais esguia do que a distribuição Normal;



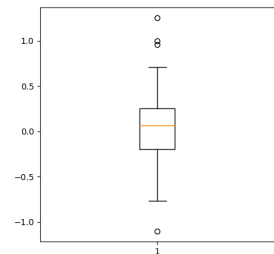
4.4

5.1



5.2 $\bar{x} \approx 0.1$, $\tilde{x} \approx 0.065$, $\text{Moda} = 0.25$,

5.3 $p_{10} = -0.409$, significa que 10% das observações são menores que -0.409 ;

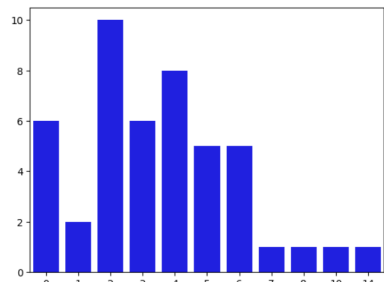


5.4

6.1 Variável quantitativa discreta;

6.2.i

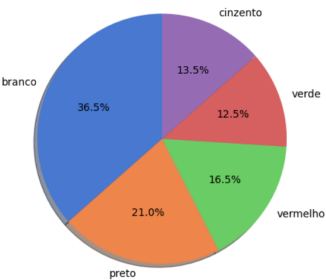
Nº Filhos	Freq. abs.	Freq. rel.
0	6	13.0%
1	2	4.3%
2	10	21.7%
3	6	13.0%
4	8	17.4%
5	5	10.9%
6	5	10.9%
7	1	2.2%
8	1	2.2%
10	1	2.2%
14	1	2.2%



6.2.ii
inal; 7.2.i

7.1 Variável qualitativa nom-

Cores	Freq. abs.	Freq. rel.
branco	73	36.5%
preto	42	21.0%
vermelho	33	16.5%
verde	25	12.5%
cinzento	27	13.5%



7.2.ii