

Gabarito da Lista 2 – Tópicos 2.3 e 2.4

1. a) Média = $129,05 / 40 = 3,23$ reais

$$n \times p = 40 \times 0,5 = 20,0$$

$$Md = x_{.5} = (x_{(20)} + x_{(21)}) / 2 = (3,22 + 3,22) / 2 = 3,22 \text{ reais}$$

$$\text{Moda} = 3,22 \text{ reais}$$

b) Média é uma medida não resistente, ou seja, é influenciada por valores discrepantes.

Mediana e Moda são medidas resistente, não sendo influenciadas por valores discrepantes

c)

- De acordo com a segunda propriedade matemática da média aritmética, se todos os preços sofrerem um aumento de 50 centavos, a média aumentará em 50 centavos.

- De acordo com a terceira propriedade matemática da média, se todos os preços dobrarem, a média também dobrará.

- Se o maior valor for substituído por 5,50, a média aumentará, mas a mediana não sofrerá alteração.

d) Os desvios medem a diferença (ou afastamento) de cada observação em relação à média aritmética.

Uma propriedade da média relativa ao desvio afirma que a soma de todos os desvios em relação à média de um conjunto de valores é nula.

2.

i	x_i	$(x_i - \bar{x})^2$
1	9	16
2	1	16
3	5	0
4	6	1
5	4	1
6	4	1
7	6	1
Σ	35	36

$$\bar{x} = 35 / 7 = 5$$

$$s^2 = 36 / 6 = 6$$

$$s = 2,449$$

Interpretação: A variação média dos dados em torno da média aritmética é de 2,449.

$$CV = 2,449 / 5 \times 100 = 48,98\%$$

3. a) $\bar{x} = 836 / 30 = 27,87$ terminais

$$s^2 = 511,47 / 29 = 17,64 \text{ terminais}^2$$

$$m_2 = \frac{\sum (x_i - \bar{x})^2}{n} = 511,47 / 30 = 17,05 \text{ terminais}^2$$

$$m_3 = \frac{\sum (x_i - \bar{x})^3}{n} = 242,66 / 30 = 8,088 \text{ terminais}^3$$

$$m_4 = \frac{\sum (x_i - \bar{x})^4}{n} = 16414,85 / 30 = 547,16 \text{ terminais}^4$$

$$a_3 = \frac{m_3}{m_2 \sqrt{m_2}} = 0,125$$

$$a_4 = \frac{m_4}{m_2^2} = 1,882$$

b)

$$a_3 = 0,125 \quad -0,5 \leq a_3 \leq 0,5 \rightarrow \text{Simétrica}$$

$$a_4 = 1,882 \quad a_4 < 3 \rightarrow \text{Platicúrtica}$$

4. a) Dados ordenados:

7, 7, **8, 8**, 8, 8, 9, **9, 9**, 9, 9, 9, 9, 9, **10, 10**, 10, 10, 10, 10, 10, 10, **10, 10**, 11, 11, 11, 11, **12, 12**

$$30 \times 0,1 = 3,0$$

$$x_{.1} = (x_{(3)} + x_{(4)}) / 2 = (8 + 8) / 2 = 8$$

$$30 \times 0,25 = 7,5$$

$$x_{.25} = x_{(8)} = 9$$

$$30 \times 0,5 = 15,0$$

$$x_{.5} = (x_{(15)} + x_{(16)}) / 2 = (10 + 10) / 2 = 10$$

$$30 \times 0,75 = 22,5$$

$$x_{.75} = x_{(23)} = 10$$

$$30 \times 0,95 = 28,5$$

$$x_{.95} = x_{(29)} = 12$$

b) Em pelo menos 10% dos grupos o número de pessoas com diabetes é menor ou igual a 8.

Em pelo menos 95% dos grupos o número de pessoas com diabetes é menor ou igual a 12.

c) $a_t = ES - EI = 12 - 7 = 5$

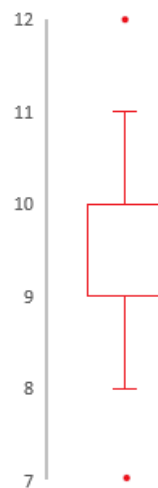
Interpretação: Todos os grupos diferem entre si, no máximo, em 5 pessoas com diabetes.

$$a_q = Q_3 - Q_1 = 10 - 9 = 1$$

Interpretação: 50% dos grupos (mais centrais) diferem, no máximo, em 1 pessoa com diabetes.

d) $a_q = 1$
 $CI = Q_1 - 1,5 a_q = 9 - 1,5 \times 1 = 7,5$
 $CS = Q_3 + 1,5 a_q = 10 + 1,5 \times 1 = 11,5$

e)



$Md = Q_3$, logo $DS < DI \rightarrow$ Assimétrica negativa

5. (C)
 (E)
 (E)
 (C)
 (C)
 (C).
 (C)
 (C)
 (C)
 (C)

6. a)

3	5	9	9						
4	3	3	4	4					
4	6	7	7	8	8				
5	2	2	3	3	4				
5	5	5	5	6	7	7	8	9	9
6	1	1	2	2	3				
6	5	6							
7	2	3							
7	9								

Chave: 3|5 = 35

b) Aproximadamente simétrica.

c) O diagrama indica que a distribuição do número de peças defeituosas é aproximadamente simétrica, com maior frequência de valores no centro da distribuição.

d) Sim, pois o gráfico é simétrico e há maior frequência no centro.

e) O diagrama de ramo e folhas dá ideia do formato da distribuição sem perder informação, pois todos os dados são apresentados no diagrama.

7. a)

j	Classes	c_j	F_j	F'_j	f_j	f'_j
1	3,5 — 5,5	4,5	1	1	0,01	0,01
2	5,5 — 7,5	6,5	3	4	0,03	0,04
3	7,5 — 9,5	8,5	8	12	0,08	0,12
4	9,5 — 11,5	10,5	17	29	0,17	0,29
5	11,5 — 13,5	12,5	19	48	0,19	0,48
6	13,5 — 15,5	14,5	19	67	0,19	0,67
7	15,5 — 17,5	16,5	11	78	0,11	0,78
8	17,5 — 19,5	18,5	16	94	0,16	0,94
9	19,5 — 21,5	20,5	3	97	0,03	0,97
10	21,5 — 23,5	22,5	2	99	0,02	0,99
11	23,5 — 25,5	24,5	1	100	0,01	1
Total		-	100	-	1	-

b) Classes modais = 5 e 6, pois ambas apresentam a maior frequência absoluta.

A classe 4 compreende o Q_1 .

A classe 6 compreende o Q_2 .

A classe 7 compreende o Q_3 .

c)
$$\bar{x}_p = (4,5 \times 1 + 6,5 \times 3 + 8,5 \times 8 + \dots + 24,5 \times 1) / 100$$
$$= 1392 / 100 = 13,92 \text{ placas}$$

8.

a) O conjunto de técnicas que compreende este gráfico é denominado Análise Exploratória de dados.

b) Essas técnicas foram propostas por John Tukey com o objetivo de contornar as limitações da média e da variância na representação de distribuições assimétricas. Para isso Tukey propôs o uso de medidas resistentes.

c) Assimétrica positiva.

d)
$$Md = x_{.5} = (x_{(20)} + x_{(21)}) / 2 = (443 + 443) / 2 = 443 \text{ dólares}$$
$$40 \times 0,5 = 20,0$$

Pelo menos 50% dos alunos da disciplina ministrada na Fundação Getúlio Vargas em São Paulo têm salário menor ou igual 443 dólares.

- e) A maior diferença entre quaisquer dois valores desse conjunto de dados é expressa pela amplitude total: $a_t = 1.241 - 142 = 1.099$ dólares
- f) Sim, os salários 1064 dólares e 1241 dólares parecem ser muito altos e, portanto, discrepantes dos demais.
- g) Quando se encontra um valor discrepante no conjunto de dados, a primeira atitude que se deve tomar é investigar a sua origem e certificar-se de que ele não é proveniente de erro. Se ficar comprovado que o valor resulta de algum erro, ele deve ser eliminado; caso contrário, deve permanecer no conjunto de dados.
- h) Não, pois a distribuição é assimétrica.