



CENTRO DE CIÊNCIA E TECNOLOGIA
LABORATÓRIO DE CIÊNCIAS MATEMÁTICAS
UNIVERSIDADE ESTADUAL DO NORTE FLUMINENSE

Curso: Ciência da Computação **Disciplina:** Estatística e Probabilidade

Data: 20./03/ 2024

Lista de exercícios 1 – Estatística Descritiva

1.- Dados os números que aparecem na seguinte tabela:

6,3	8,8	7,9	9,2	8,6	8,7	8,3	7,8	4,1	6,7
6,8	7,6	4,6	8,1	9,2	7,7	8,4	8,6	7,0	6,6
7,7	7,5	9,8	8,1	8,2	8,1	8,7	7,8	7,0	6,0
9,4	7,9	5,2	8,2	7,7	8,1	7,7	7,0	7,4	6,1

- Defina intervalos de classe de maneira a agrupar esses dados. Para isso, defina o número de classes e especifique os limites: inferior e superior de cada classe de maneira que cada intervalo possua igual amplitude.
- Construa uma tabela com a distribuição de frequência, frequência acumulada e ponto médio do intervalo.

2.- Considere o conjunto de dados a seguir:

Observações	X	Y	Z
1	1	2	4
2	2	4	5
3	3	6	6
4	4	8	7
5	5	10	8

- Para cada uma das variáveis X, Y, Z, calcule: média, variância, desvio padrão, coeficiente de variação (cv), mediana (Q_2), quartis, Q_1 e Q_3 .

Resposta 2.-

Realizando os calculos para dados não agrupados temos:

Média: Para o conjunto X Para o conjunto Y Para o conjunto Z

$$\bar{x} = \frac{\sum_{i=1}^n x_i}{n} = \frac{15}{5} = 3 \quad \bar{x} = \frac{\sum_{i=1}^n x_i}{n} = \frac{30}{5} = 6 \quad \bar{x} = \frac{\sum_{i=1}^n x_i}{n} = \frac{30}{5} = 6$$

Variância:**Para o conjunto X**

$$s^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n-1} = \frac{(1-3)^2 + (2-3)^2 + (3-3)^2 + (4-3)^2 + (5-3)^2}{5-1}$$

$$= \frac{4+1+0+1+4}{4} = \frac{10}{4} = 2,5$$

Para o conjunto Y

$$s^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n-1} = \frac{(2-6)^2 + (4-6)^2 + (6-6)^2 + (8-6)^2 + (10-6)^2}{5-1}$$

$$= \frac{16+4+0+4+16}{4} = \frac{40}{4} = 10$$

Para o conjunto Z

$$s^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n-1} = \frac{(4-6)^2 + (5-6)^2 + (6-6)^2 + (7-6)^2 + (8-6)^2}{5-1}$$

$$= \frac{4+1+0+1+4}{4} = \frac{10}{4} = 2,5$$

Desvio Padrão: Para o conjunto X Para o conjunto Y Para o conjunto Z

$$s = \sqrt{2,5} = 1,5811 \quad s = \sqrt{10} = 3,1623 \quad s = \sqrt{2,5} = 1,5811$$

Coef. Variação:

Para o conjunto X $CV = \frac{s}{\bar{x}} \cdot 100 = \frac{1,5811}{3} \cdot 100 = 52,703$

Para o conjunto Y $CV = \frac{s}{\bar{x}} \cdot 100 = \frac{3,1623}{6} \cdot 100 = 52,705$

Para o conjunto Z $CV = \frac{s}{\bar{x}} \cdot 100 = \frac{1,5811}{6} \cdot 100 = 26,352$

Mediana: Para o conjunto X Para o conjunto Y Para o conjunto Z

$$M_e = 3 \quad M_e = 6 \quad M_e = 6$$

Quartil (Q₁): Para o conjunto X Para o conjunto Y Para o conjunto Z

$$Q_1 = \frac{1+2}{2} = 1,5 \quad Q_1 = \frac{2+4}{2} = 3 \quad Q_1 = \frac{4+5}{2} = 4,5$$

Quartil (Q₃): Para o conjunto X Para o conjunto Y Para o conjunto Z

$$Q_3 = \frac{4+5}{2} = 4,5 \quad Q_3 = \frac{8+10}{2} = 9 \quad Q_3 = \frac{7+8}{2} = 7,5$$

- 3.- Dois fabricantes de refrigeradoras anunciam que a vida “média” de seus produtos é de sete anos. No entanto, ao se avaliar uma amostra aleatória referente à duração dos produtos, um fiscal encontra os seguintes dados:

A vida em anos do produto do fabricante A é:

5, 5, 5, 6, 6, 6, 6, 7, 7, 7, 7, 7, 7, 8, 8, 8, 8, 9, 9, 9

A vida em anos do produto do fabricante B é:

2, 3, 4, 5, 5, 5, 5, 6, 6, 6, 7, 7, 7, 7, 7, 8, 8, 20, 20, 20

- Qual medida de tendência central utilizou realmente cada fabricante?
- Qual produto representa um melhor investimento?
- Com qual produto você sentiria mais segurança ao afirmar que sua vida “promédio” é de sete anos.

Resposta 3.-

- a) Ao se calcular a média aritmética temos:

$$\bar{x}_A = \frac{\sum_{i=1}^n x_i}{n} = \frac{140}{20} = 7 \qquad \bar{x}_B = \frac{\sum_{i=1}^n x_i}{n} = \frac{158}{20} = 7,9$$

Ao se calcular a mediana temos:

$$M_{e(A)} = \frac{7+7}{2} = 7 \qquad M_{e(B)} = \frac{6+7}{2} = 6,5$$

Ao se calcular a moda temos:

$$M_{o(A)} = 7 \qquad M_{o(B)} = 7$$

No caso do fabricante A, todas as medidas são iguais a 7. No caso do fabricante B, a moda tem valor 7.

- Tendo como referência a mediana, observa-se que metade das observações tem duração abaixo de 7 no primeiro fabricante e duração abaixo de 6,5 no segundo fabricante. O primeiro fabricante parece um investimento mais confiável.
- O fabricante A fornece mais segurança já que a duração de seus produtos varia menos em torno da média aritmética. Teria sido interessante calcular o desvio padrão em cada caso.

- 4.- Em um mês dado, uma pessoa compra três dúzias de ovos por \$0,50; uma dúzia por \$0,45; duas dúzias por \$0,55 e uma dúzia por \$0,40. Qual é o preço médio por dúzia?

Resposta 4.-

Podemos agrupar os dados considerando o preço e a sua frequência.

Preço (\$) x_i	f_i	x_i f_i
0,40	1	0,40
0,45	1	0,45
0,50	3	1,50
0,55	2	1,10
Totais	7	3,45

Calcula-se a média para dados agrupados:

$$\bar{x} = \frac{\sum_{i=1}^k x_i f_i}{\sum_{i=1}^k f_i} = \frac{3,45}{7} = 0,4928 \text{ que é o preço médio por dúzia.}$$

5.- Supor que um comerciante vende o seguinte número de maçãs de segunda-feira a **sábado**:

a) 30, 30, 30, 30, 30, 30

Qual é a média?, A mediana?, A moda?

b) 25, 30, 35, 30, 35, 25

Qual é a média?, A mediana?, A moda?

c) 10, 25, 30, 35, 25, 30

Qual é a média?, A mediana?, A moda?

6.- Considere a seguinte tabela com a distribuição de frequências do tempo de permanência na última internação referente a uma amostra de pacientes do Hospital Espírita de Porto Alegre nos quatro primeiros meses de 1996. Determine a medidas de tendência central e de dispersão: Média, Médiana, Moda. Desvio Padrão. Primeiro Quartil e Terceiro Quartil. Estenda a tabela caso seja necessário.

Permanência (dias)	Frequência Simples (f _i)
0 a 10	70
10 a 20	69
20 a 30	90
30 a 40	43
40 a 50	43
50 a 60	31
60 a 70	16
70 a 80	7
80 a 90	0
90 a 100	1
Total	370

Resposta 6.-

Como comentado em aula, a extensão refere-se apenas a adição de colunas para realizar os cálculos necessários para o cálculo das estatísticas de tendência central e dispersão para o caso de dados agrupados.

Intervalos de Classe	Ponto Médio (x_i)	Freq. Simples (f_i)	$x_i f_i$	Freq. Acum. (f_{ac})	$(x_i - \bar{x})^2$	$f_i (x_i - \bar{x})^2$
[0, 10 >	5	70	350	70	538,983	37728,81
[10, 20 >	15	69	1.035	139	174,663	12051,747
[20, 30 >	25	90	2.250	229	10,343	930,87
[30, 40 >	35	43	1.505	272	46,023	1978,989
[40, 50 >	45	43	1.935	315	281,703	12113,229
[50, 60 >	55	31	1.705	346	717,383	22238,873
[60, 70 >	65	16	1.040	362	1353,063	21649,008
[70, 80 >	75	7	525	369	2188,743	15321,201
[80, 90 >	85	0	0	369	3224,423	0
[90, 100 >	95	1	95	370	4460,103	4460,103
Totais		370	10.440			128472,83

Tratam-se os dados como valores contínuos.

$$\text{A média neste caso é: } \bar{x} = \frac{\sum_{i=1}^k x_i f_i}{\sum_{i=1}^k f_i} = \frac{10.440}{370} = 28,216 \text{ dias}$$

A posição da mediana $P = 370/2 = 185$

$$\text{A mediana neste caso é: } M_e = LI_e + \left(\frac{P - f'_{ac}}{f_{M_e}} \right) \cdot h = 20 + \left(\frac{185 - 139}{90} \right) \cdot 10 = 25,111$$

A classe modal coincide com a classe mediana.

$$\text{A moda neste caso é: } M_o = LI_o + \left(\frac{\Delta_1}{\Delta_1 + \Delta_2} \right) \cdot h = 20 + \left(\frac{21}{21 + 47} \right) \cdot 10 = 23,088$$

$$\text{O desvio padrão é: } s^2 = \sqrt{\frac{\sum_{i=1}^k f_i (x_i - \bar{x})^2}{n - 1}} = \sqrt{\frac{128472,83}{369}} = 18,659$$

Deixo os quartis pra vocês.

- 7.- Uma pessoa compara os preços de um produto em vários supermercados. Foram observados os seguintes preços por quilograma em reais:

56, 65, 48, 73, 59, 72, 63, 65, 60, 63, 44, 79, 63, 61, 66, 69, 64, 71, 58, 63.

- Calcule a média
- Calcule a amplitude, o desvio médio
- Calcule o desvio padrão e a variância.

8.- A tabela a seguir contém para cada um de vários países de Europa, estimativas do consumo diário de proteína em gramas per capita por tipo de alimento.

País	Carne	Porco e Aves	Ovos	Leite	Peixe	Cereais	Amido	Nozes	Frutas e Vegetais
Albânia	10,1	1,4	0,5	8,9	0,2	42,3	0,6	5,5	1,7
Áustria	8,9	14,0	4,3	19,9	2,1	28,0	3,6	1,3	4,3
Bélgica	13,5	9,3	4,1	17,5	4,5	26,6	5,7	2,1	4,0
Bulgária	7,8	6,0	1,6	8,3	1,2	56,7	1,1	3,7	4,2
Tcheco	9,7	11,4	2,8	12,5	2,0	34,3	5,0	1,1	4,0
Dinamarca	10,6	10,8	3,7	25,0	9,9	21,9	4,8	0,7	2,4
Alemania Oriental	8,4	11,6	3,7	11,1	5,4	24,6	6,5	0,8	3,6
Finlândia	9,5	4,9	2,7	33,7	5,8	26,3	5,1	1,0	1,4
França	18,0	9,9	3,3	19,5	5,7	28,1	4,8	2,4	6,5
Grécia	10,2	3,0	2,8	17,6	5,9	41,7	2,2	7,8	6,5
Hungria	5,3	12,4	2,9	9,7	0,3	40,1	4,0	5,4	4,2
Irlanda	13,9	10,0	4,7	25,8	2,2	24,0	6,2	1,6	2,9
Itália	9,0	5,1	2,9	13,7	3,4	36,8	2,1	4,3	6,7
Holanda	9,5	13,6	3,6	23,4	2,5	22,4	4,2	1,8	3,7
Noruega	9,4	4,7	2,7	23,3	9,7	23,0	4,6	1,6	2,7
Polônia	6,9	10,2	2,7	19,3	3,0	36,1	5,9	2,0	6,6
Portugal	6,2	3,7	1,1	4,9	14,2	27,0	5,9	4,7	7,9
Romênia	6,2	6,3	1,5	11,1	1,0	49,6	3,1	5,3	2,8
Espanha	7,1	3,4	3,1	8,6	7,0	29,2	5,7	5,9	7,2
Suécia	9,9	7,8	3,5	24,7	7,5	19,5	3,7	1,4	2,0
Suíça	13,1	10,1	3,1	23,8	2,3	25,6	2,8	2,4	4,9
Reino Unido	17,4	5,7	4,7	20,6	4,3	24,3	4,7	3,4	3,3
Rússia	9,3	4,6	2,1	16,6	3,0	43,6	6,4	3,4	2,9
Alemania Occidental	11,4	12,5	4,1	18,8	3,4	18,6	5,2	1,5	3,8
Iugoslávia	4,4	5,0	1,2	9,5	0,6	55,9	3,0	5,7	3,2

Determine para cada uma das nove variáveis descritas nas colunas da tabela:

A média, a mediana, o desvio padrão, o primeiro quartil, o terceiro quartil e o coeficiente de variação.