

Lista de exercícios 2
Gabarito e chave de respostas
Estatística
Prof.: Nielsen

1) Calcule Q1, Q2 e Q3 para o seguinte conjunto de valores.

$$A\{4,1,8,0,11,10,7,8,6,2,9,12\}$$

O número que corresponde a 25% do rol é o valor 2.

O número que corresponde a 50% do rol é o valor 7

O número que corresponde a 75% do rol é o valor 9

2) Determinação dos quartís para o conjunto:

$$B\{10,15,18,13,19,25,27,16,15,21,22,29,14,18,20,24\}$$

O número que corresponde a 25% do rol é o valor 15

O número que corresponde a 50% do rol é o valor 18

O número que corresponde a 75% do rol é o valor 22

3) Calcular os valores do Q_1 , Q_2 e Q_3 da tabela seguinte:

Tabela 1 – Números de acidentes / mês no Cruzamento X em CG/07

N° de acidentes / mês	f_i
0	4
1	6
2	9
3	5
4	4
$\sum f = 28$	

A variável que corresponde à posição do 7° elemento é 1 (na segunda classe).

25% da pesquisa mostrou que este cruzamento teve 1 acidente / mês.

A variável que corresponde à posição do 14° elemento é 2 (na terceira classe).

50% da pesquisa mostrou que este cruzamento teve 2 acidentes / mês.

A variável que corresponde à posição do 21° elemento é 3 (na quarta classe).

75% da pesquisa mostrou que este cruzamento teve 3 acidentes / mês.

4) Encontre os valores dos quartís de dados agrupados com intervalos de classe, consideramos a distribuição dos pesos de um grupo de turistas que visita um parque temático.

Tabela2 - Pesos de um grupo de turistas do Parque Temático

i	Pesos (kg)	Frequência (f_i)
---	---------------	----------------------

1	10 30	10
2	30 50	24
3	50 70	57
4	70 90	44
5	90 110	29
6	110 130	16

Primeiro, calcula-se a classe a que pertence o quartil Q_1 ($k=1$), ou seja, a posição:

$$P_{Q_1} = \frac{1 \cdot \sum f_i}{4} = \frac{180}{4} = 45$$

Observando a coluna de frequência acumulada, verificamos que o quadragésimo quinto termo pertence à terceira classe (a frequência acumulada da terceira classe abrange do 35º termo ao 91º termo). Sabendo que a classe do primeiro quartil é a terceira classe, podemos verificar qual o valor numérico do primeiro quartil utilizando a expressão:

$$Q_1 = l_{Q_1} + \left[\frac{1 \sum f_i - F_{ant}}{f_{Q_1}} \right] \cdot a_{Q_1} = 50 + \left[\frac{45 - 34}{57} \right] \cdot 20 \cong 53,9 \text{ kg}$$

Os cálculos para os quartis Q_2 e Q_3 processam-se de forma análoga ao cálculo do primeiro quartil.

$$2^\circ \text{ quartil} \rightarrow \frac{2 \sum f_i}{4} = \frac{2 \times 180}{4} = 90 \text{ (o segundo quartil pertence à terceira classe).}$$

$$Q_2 = l_{Q_2} + \left[\frac{2 \sum f_i - F_{ant}}{f_{Q_2}} \right] \cdot a_{Q_2} = 50 + \left[\frac{90 - 34}{57} \right] \cdot 20 \cong 69,7 \text{ kg}$$

$$3^\circ \text{ quartil} \rightarrow \frac{3 \sum f_i}{4} = \frac{3 \times 180}{4} = 135 \text{ (o terceiro quartil pertence à quarta classe)}$$

$$Q_3 = l_{Q_3} + \left[\frac{3 \sum f_i - F_{ant}}{f_{Q_3}} \right] \cdot a_{Q_3} = 70 + \left[\frac{135 - 91}{44} \right] \cdot 20 = 90,0 \text{ kg}$$

Assim temos: $Q_1 = 53,9$ kg; $Q_2 = 69,7$ kg e $Q_3 = 90,0$ kg

5) Calcular os valores do Q_1 , Q_2 e Q_3 da tabela seguinte:

Tabela 3 – Números de faltas de acadêmicos do primeiro semestre.

Nº de faltas	f	fa
1	6	8
3	8	14
4	13	27
6	7	34
7	4	38
		$\sum f = 38$

a) Vamos calcular inicialmente Q_1

1º Passo: Determinar a posição do 1º quartil (25%)

$$P_{Q_1} = \frac{1 \cdot 38}{4} = 9,5 \Rightarrow \text{posição do 1º quartil}$$

2º Passo: Procurar na coluna da *fa* a posição do 9,5º elemento

3º Passo: A variável que corresponde à posição do 9,5º elemento é 3 (na segunda classe).

25% dos acadêmicos tiveram 3 faltas

b) Vamos calcular o Q_2

1º Passo: Determinar a posição do 2º quartil (50%)

$$P_{Q_2} = \frac{2 \cdot 38}{4} = 19 \Rightarrow \text{posição do 2º quartil}$$

2º Passo: Procurar na coluna da *fa* a posição do 19º elemento

3º Passo: A variável que corresponde à posição do 19º elemento é 4 (na terceira classe).

50% dos acadêmicos tiveram 4 faltas.

c) Vamos calcular o Q_3

1º Passo: Determinar a posição do 3º quartil (75%)

$$P_{Q_3} = \frac{3 \cdot 38}{4} = 28,5 \Rightarrow \text{posição do 3º quartil}$$

2º Passo: Procurar na coluna da *fa* a posição do 28,5º elemento

3º Passo: A variável que corresponde à posição do 28,5º elemento é 6 (na quarta classe).

75% dos acadêmicos tiveram 6 faltas.

6) Calcular D3 e D7 usando a seguinte tabela:

Tabela 4 Quantidade de filhos dos funcionários de uma pequena empresa.

filhos	f	fa
0	18	18
1	35	53
2	46	99
3	28	127
4	25	152
5	10	162
6	5	167
7	3	170
		$\sum f = 170$

a) Cálculo do D₃

1° Passo: Calcula-se a posição do D₃

$$D_3 = \frac{3 \sum f}{10} = \frac{3 \cdot 170}{10} = 51 \text{ (posição)}$$

2° passo: Procura-se a posição do D₃ pela coluna da frequência acumulada, o D₃ está na 2° classe (fa 53)

3° Passo: O valor da variável na segunda classe é 1 filho, que corresponde a 30% da pesquisa.

b) Cálculo do D₈

1° Passo: Calcula-se a posição do D₈

$$D_8 = \frac{8 \sum f}{10} = \frac{8 \cdot 170}{10} = 144 \text{ (posição)}$$

2° passo: Procura-se a posição do D₈ pela coluna da frequência acumulada, o D₈ está na 5° classe (fa 152)

3° Passo: O valor da variável na segunda classe é 4 filhos, que corresponde a 80% da pesquisa.

7) Calcular o percentil, o P₂₈ e P₈₂ do conjunto $B\{15,2,4,6,10,12,13,7,21,18,20\}$

Devemos inicialmente ordenar os valores:

X ₁	X ₂	X ₃	X ₄	X ₅	X ₆	X ₇	X ₈	X ₉	X ₁₀	X ₁₁
2	4	6	7	10	12	13	15	18	20	21

a) Cálculo do P₂₈

1° Passo: Determinar a posição do P₂₈ $P_{28} = \frac{28 \cdot n}{100} = \frac{28 \cdot 11}{100} = 3,08$

2° Passo: procura-se no rol o valor da posição do 3° elemento;

3° Passo: A variável que corresponde à posição desejada é o número 6

b) Cálculo do P₈₂

- 1° Passo: Determinar a posição do P_{82} $P_{28} = \frac{82 \cdot n}{100} = \frac{28 \cdot 11}{100} = 9,02$
 2° Passo: procura-se no rol o valor da posição do 9° elemento;
 3° Passo: A variável que corresponde à posição desejada é o número 18

8) Calcular P_{45} e P_{93} da tabela

Tabela 5. Número de quartos/chalés em Bonito/MS/07

Número de quartos/chalés	f	fa
1	15	15
2	30	45
3	20	65
4	12	77
5	10	87
6	8	95
		$\sum f_i = 95$

a) Calcular P_{45}

1° Passo: Determinar a posição do P_{45} $P_{45} = \frac{45 \cdot n}{100} = \frac{45 \cdot 95}{100} = 42,75$

2° Passo: Procurar a posição do 43 elemento pela coluna da frequência acumulada, podemos observar que o elemento de posição 43 está na segunda classe;

3° Passo O valor da variável que corresponde a 45% da pesquisa revelou que os pesquisados preferem até dois quartos por chalé.

b) Calcular P_{93}

1° Passo: Determinar a posição do P_{93} $P_{93} = \frac{93 \cdot n}{100} = \frac{93 \cdot 95}{100} = 88,35$

2° Passo: Procurar a posição do 88° elemento pela coluna da frequência acumulada, podemos observar que o elemento de posição 88 está na sexta classe;

3° Passo O valor da variável que corresponde a 93% da pesquisa revelou que os pesquisados preferem até seis quartos por chalé.

9) Determine o desvio padrão dos dados a seguir, esses dados foram obtidos em indivíduos contaminados pelo veneno de um certo tipo de inseto e submetidos a tratamento. A variável de interesse Recup é definida como o tempo (em horas) entre a administração do tratamento e a recuperação do indivíduo. Os valores de Recup são os seguintes: 3, 90, 23, 46, 2, 42, 47, 37, 12, 51, 11, 1, 3, 3, 45, 3, 4, 11, 2, 8, 56, 39, 22, 16, 5 e 52.

$$\text{Média} = \bar{X} = \frac{X_1 + X_2 + \dots + X_{26}}{26} = \frac{\sum_{i=1}^{26} X_i}{26} = (3 + 90 + \dots + 52)/26 = 24,38$$

Mediana: valor que deixa 50% das observações ordenadas à sua esquerda.

Com os dados ordenados, a observação que ocupa a "posição" $(26+1)/2=13,5$ será a mediana. Portanto será a média entre os elementos que ocupam as posições 13 e 14.

$$\text{Mediana} = (12+16)/2=14$$

Desvio padrão:

Temos que a variância pode ser obtida pela seguinte fórmula:

$$Var = S^2 = \frac{\sum_{i=1}^n X_i^2 - n(\bar{X})^2}{n-1} = \frac{(3^2 + \dots + 52^2) - 26(24,38)^2}{25} = 555,61$$

Assim, desvio padrão = $S = \sqrt{S^2} = \sqrt{555,61} = 23,57$

10) Separe o conjunto de dados (questão 9) em três grupos denominados cura rápida, com valor de Recup menor ou igual a 12, cura normal, se o valor de Recup for maior do que 12 e menor ou igual a 45, e cura lenta, se o valor de Recup estiver acima de 45. Compare a variabilidade desses três grupos através de seus coeficientes de variação.

	Média	DP	CV (%)
Cura rápida	5,23	3,88	74,19
Cura normal	32,00	11,40	35,62
Cura lenta	57,00	16,56	29,05

Os cálculos foram obtidos da seguinte maneira:

Médias:

Media da Cura rápida = $(1+2+2+3+3+3+3+4+5+8+11+11+12)/13 = 5,23$

Média da Cura normal = $(16+22+23+37+39+42+45)/7 = 32$

Média da Cura lenta = $(46+47+51+52+56+90)/6 = 57$

Desvio Padrão (s)

$$\text{Desvio Padrão da Cura rápida} = \sqrt{\frac{\sum_{i=1}^n X_i^2 - n(\bar{X})^2}{n-1}} = \sqrt{\frac{(1^2 + \dots + 12^2) - 13(5,23)^2}{12}} = 3,88$$

$$\text{Desvio Padrão da Cura normal} = \sqrt{\frac{\sum_{i=1}^n X_i^2 - n(\bar{X})^2}{n-1}} = \sqrt{\frac{(16^2 + \dots + 45^2) - 7(32)^2}{6}} = 11,40$$

$$\text{Desvio Padrão da Cura leta} = \sqrt{\frac{\sum_{i=1}^n X_i^2 - n(\bar{X})^2}{n-1}} = \sqrt{\frac{(46^2 + \dots + 90^2) - 6(57)^2}{5}} = 16,56$$

Coefficiente de Variação (CV)

$$\text{CV da cura rápida} = \frac{s}{\bar{X}} * 100 = \frac{3,88}{5,23} * 100 = 74,19\%$$

$$\text{CV da cura normal} = \frac{s}{\bar{X}} * 100 = \frac{11,40}{32} * 100 = 35,62\%$$

$$\text{CV da cura lenta} = \frac{s}{\bar{X}} * 100 = \frac{16,56}{57} * 100 = 29,05\%$$

Em relação às médias, a variabilidade do grupo cura rápida é maior que a dos demais grupos. O grupo cura lenta apresentou a menor variabilidade.

11) Um órgão do governo do estado está interessado em determinar padrões sobre o investimento em educação, por habitante, realizado pelas prefeituras. De um levantamento de dez cidades, foram obtidos os valores (codificados) da tabela abaixo:

Cidade	A	B	C	D	E	F	G	H	I	J
Investimento	25	16	14	10	19	15	19	16	19	18

Receberão um programa especial as cidades com valores de investimento inferiores à média menos duas vezes o desvio padrão. Alguma cidade receberá o programa?

$$\text{Média} = \bar{X} = \frac{X_1 + X_2 + \dots + X_{10}}{10} = \frac{\sum_{i=1}^{10} X_i}{10} = \bar{X} = \frac{25 + \dots + 18}{10} = 17,1$$

Receberão um programa especial as cidades com valores de investimento inferiores à média menos duas vezes o desvio padrão. Alguma cidade receberá o programa?

Média	DP	2DP	Média - 2DP	Média + 2DP
17,1	3,96	7,92	9,18	25,02

Nenhuma cidade receberá o programa especial, pois nenhuma cidade apresentou valor de investimento inferior a 9,18.

O desvio padrão foi calculado da seguinte maneira:

$$s = \sqrt{s^2} = \sqrt{\frac{\sum_{i=1}^n X_i^2 - n(\bar{X})^2}{n-1}} = \sqrt{\frac{(25^2 + \dots + 18^2) - 10(17,1)^2}{9}}$$

12) As concentrações de óxido de nitrogênio e hidrocarbono (em µg/m³) foram determinadas em uma área urbana, em locais e horários específicos. Os dados são mostrados a seguir.

Dia	Óxido de Nitrogênio (O)	Hidrocarbono (H)	DIF= O - H
1	104	108	-4
2	116	118	-2
3	84	89	-5
4	77	71	6
5	61	66	-5
6	84	83	1
7	81	88	-7
8	72	76	-4
9	61	68	-7
10	97	96	1
11	84	81	3

Calcule média e desvio padrão para cada variável

Classifique as variáveis em estudo.

Óxido de Nitrogênio: quantitativa contínua

Hidrocarbono de Nitrogênio: quantitativa contínua

Realize uma análise descritiva dos dados. Calcule média e desvio padrão para cada variável e para a variável DIF = O - H (diferença entre as concentrações dos poluentes).

Variável	Média (\bar{X})	Desvio Padrão (s)
O	83,73	16,89
H	85,82	16,44
DIF = O - H	-2,09	4,28

Os cálculos das médias e dos desvios padrões estão apresentados a seguir.

Médias:

$$\text{Média de O} = (104 + \dots + 84) / 11 = 83,73$$

$$\text{Média de H} = (108 + \dots + 81) / 11 = 85,82$$

$$\text{Média de DIF} = \text{O} - \text{H} = (-4 + \dots + 3) / 11 = -2,09$$

Desvio Padrão (s)

$$\text{Desvio Padrão de O} = \sqrt{\frac{\sum_{i=1}^n X_i^2 - n(\bar{X})^2}{n-1}} = \sqrt{\frac{(104^2 + \dots + 84^2) - 11(83,73)^2}{10}} = 16,89$$

$$\text{Desvio Padrão de H} = \sqrt{\frac{\sum_{i=1}^n X_i^2 - n(\bar{X})^2}{n-1}} = \sqrt{\frac{(108^2 + \dots + 81^2) - 11(85,82)^2}{10}} = 16,44$$

$$\text{Desvio Padrão de DIF} = \sqrt{\frac{\sum_{i=1}^n X_i^2 - n(\bar{X})^2}{n-1}} = \sqrt{\frac{((-4)^2 + \dots + 3^2) - 11(-2,09)^2}{10}} = 4,28$$

Considerando a variável DIF, pode-se dizer que as duas classes de poluentes estão presentes nas mesmas concentrações?

Considerando a variável DIF, pode-se dizer que a variável O apresenta uma concentração média menor que a variável H.