

1. Classifique as variáveis abaixo quanto ao tipo em: qualitativa nominal (N), qualitativa ordinal (O), quantitativa discreta (D) e quantitativa contínua (C).

- () Número de filhos por casal
- () Teor de gordura no leite (%)
- () Tempo de vida de uma lâmpada (dias)
- () Quantidade diária de quedas na internet
- () Temperatura de um escritório às 12 horas (°C)
- () Satisfação de um cliente (bom à ruim)
- () Número de e-mails recebidos por hora
- () Intervalo de tempo entre e-mails (min)
- () Conceito sobre a temperatura (frio à quente)
- () Voto com relação a uma proposta

2. A tabela de distribuição de frequências do número de DVDs alugados pela locadora Star Video nos finais de semana do ano de 2008 está representada abaixo. Complete as cédulas vazias fazendo uso das cédulas presentes por meio da relação entre os tipos de frequência.

k	Classe	f_{abs}	f_{rel}	F_{abs}	F_{rel}
1	[20,25)		0,04	2	
2	[25,30)	1	0,02	3	0,06
3	[30,35)		0,17	12	
4	[35,40)	11			0,44
5	[40,45)			42	0,81
6	[45,50)	8	0,15		
7	[50,55)			51	0,98
8	[55,60)	1	0,02	52	

*números arredondados para 2 casas decimais

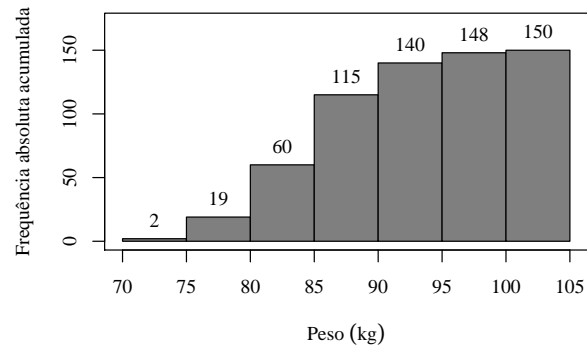
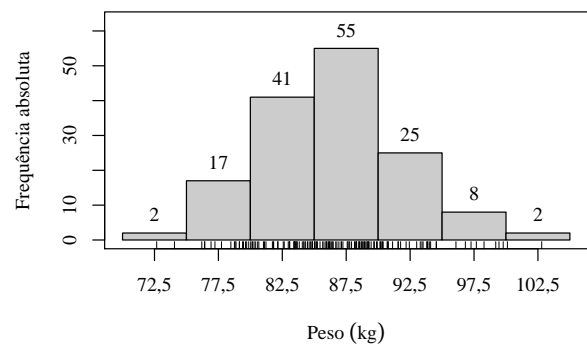
3. O número de acidentes de trabalho mensais em dois setores de uma indústria com 131 operários está apresentado abaixo (em ordem crescente).

Setor	Número de acidentes mensais									
A	2	2	3	3	4	5	7	7	8	9
B	3	3	4	4	5	5	6	6	7	7

Calcule a média (\bar{x}), a mediana (m_d), a variância (σ^2 , populacional), o desvio-padrão (σ , populacional),

o desvio médio (d_m) e o coeficiente de variação (CV) dos valores para cada setor. Em qual setor houve mais acidentes? Qual setor apresenta série de eventos mais variável? Indique os cálculos.

4. Uma empresa interessada em avaliar as condições de saúde dos seus colaboradores realizou uma amostra que considerou diversas variáveis. O peso dos indivíduos amostrados, por exemplo, está representado no histograma de frequências absolutas e frequências absolutas acumuladas a seguir.



Faça a avaliação da informação contida nos gráficos e responda: a) qual o número de indivíduos amostrados, b) qual a amplitude de classe de peso usada, c) qual a média de peso da amostra (\bar{x}), d) qual a moda de peso (m_o), e) qual a mediana de peso (m_d), f) quantos indivíduos tem menos de 90 kg, g) qual a porcentagem de indivíduos entre 80 e 90 kg. Indique os cálculos.

Expressões úteis

$$\bar{x} = \frac{\sum_{i=1}^n x_i}{n}, \quad \bar{x} = \frac{\sum_{i=1}^k f_i \cdot x_i}{n}, \quad \sigma^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n}, \quad \sigma^2 = \frac{\sum_{i=1}^n x_i^2 - \frac{(\sum_{i=1}^n x_i)^2}{n}}{n}, \quad \sigma = \sqrt{\sigma^2}, \quad CV = 100 \cdot \frac{\sigma}{\bar{x}}$$

$$d_m = \frac{\sum_{i=1}^n |x_i - \bar{x}|}{n}, \quad m_d = \begin{cases} X_{(\frac{n+1}{2})} \\ (X_{(\frac{n}{2})} + X_{(\frac{n+2}{2})})/2 \end{cases}, \quad m_d = l_{md} + \frac{n/2 - F_{esq}}{f_{md}} \cdot h_{md}, \quad m_o = l_{mo} + \frac{d_1}{d_1 + d_2} \cdot h_{mo}$$

Cada questão da prova vale 2,5 pontos, que são igualmente distribuídos nos subitens da questão. Dessa forma, temos 10, 13, 14, e 7 subitens para as questões 1, 2, 3 e 4 respectivamente, como assinalado pelos textos em vermelho dentro do gabarito¹.

1. A classificação das variáveis quanto ao tipo foi

D; C; C; D; C; O; D; C; O; N.

2. $f_{abs}(k)$ refere-se a frequência absoluta da classe k e o mesmo vale para os outros tipos de frequência, assim

$$\begin{aligned}
 f_{abs}(1) &= F_{abs}(1) = 2 \\
 F_{rel}(1) &= F_{abs}(1)/F_{abs}(8) = f_{rel}(1) = 0,04 \\
 f_{abs}(3) &= F_{abs}(3) - F_{abs}(2) = 12 - 3 = 9 \\
 F_{rel}(3) &= F_{abs}(3)/F_{abs}(8) = 12/52 = 0,23 \\
 f_{rel}(4) &= f_{abs}(4)/F_{abs}(8) = 11/51 = 0,21 \\
 F_{abs}(4) &= F_{abs}(3) + f_{abs}(4) = 12 + 11 = 23 \\
 f_{abs}(5) &= F_{abs}(5) - F_{abs}(4) = 42 - 23 = 19 \\
 f_{rel}(5) &= f_{abs}(5)/F_{abs}(8) = 19/52 = 0,37 \\
 F_{abs}(6) &= F_{abs}(5) + f_{abs}(6) = 42 + 8 = 50 \\
 F_{abs}(6) &= F_{abs}(5)/F_{abs}(8) = 50/52 = 0,96 \\
 f_{abs}(7) &= F_{abs}(7) - F_{abs}(6) = 51 - 50 = 1 \\
 f_{rel}(7) &= f_{abs}(7)/F_{abs}(8) = 1/51 = 0,02 \\
 F_{rel}(8) &= F_{abs}(8)/F_{abs}(8) = 1/1 = 1
 \end{aligned}$$

k	Classe	f_{abs}	f_{rel}	F_{abs}	F_{rel}
1	[20,25)	2	0,04	2	0,04
2	[25,30)	1	0,02	3	0,06
3	[30,35)	9	0,17	12	0,23
4	[35,40)	11	0,21	23	0,44
5	[40,45)	19	0,37	42	0,81
6	[45,50)	8	0,15	50	0,96
7	[50,55)	1	0,02	51	0,98
8	[55,60)	1	0,02	52	1

*números arredondados para 2 casas decimais

3. Aconselha-se construir a seguinte tabela auxiliar

i	x_i	Setor A		Setor B		
		$ x_i - \bar{x} $	$(x_i - \bar{x})^2$	x_i	$ x_i - \bar{x} $	$(x_i - \bar{x})^2$
1	2	3	9	3	2	4
2	2	3	9	3	2	4
3	3	2	4	4	1	1
4	3	2	4	4	1	1
5	4	1	1	5	0	0
6	5	0	0	5	0	0
7	7	2	4	6	1	1
8	7	2	4	6	1	1
9	8	3	9	7	2	4
10	9	4	16	7	2	4
total	50	22	60	50	12	20
total/n	5	2,2	6	5	1,2	2

¹Notas e faltas disponíveis em <http://www.leg.ufpr.br/~walmes/ensino/notas-1.pdf>

Por meio da tabela auxiliar obtém-se a média (\bar{x}), o desvio-médio (d_m) e a variância (σ^2 , populacional) como resultado da última linha da tabela. Poderia ter-se calculado por meio das expressões, que para o setor A, seria

$$\begin{aligned}\bar{x} &= \frac{\sum_{i=1}^n x_i}{n} = \frac{2 + 2 + 3 + \dots + 9}{10} = \frac{50}{10} = 5 \\ d_m &= \frac{\sum_{i=1}^n |x_i - \bar{x}|}{n} = \frac{|2 - 5| + |2 - 5| + \dots + |9 - 5|}{10} = \frac{22}{5} = 2,2 \\ \sigma^2 &= \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n} = \frac{(2 - 5)^2 + (2 - 5)^2 + \dots + (9 - 5)^2}{10} = \frac{60}{5} = 6 \\ \sigma &= \sqrt{\sigma^2} = \sqrt{6} = 2,45 \\ CV &= 100 \cdot \frac{\sigma}{\bar{x}} = 100 \cdot \frac{2,45}{5} = 48,98 \\ m_d &= (x_{(\frac{n}{2})} + x_{(\frac{n+2}{2})})/2 = (x_{(5)} + x_{(6)})/2 = (4 + 5)/2 = 4,5\end{aligned}$$

Para o setor B essas quantidades são $\bar{x} = 5$, $d_m = 1,2$, $\sigma^2 = 2$, $\sigma = 1,41$, $CV = 28,28$, $m_d = 5$. Em ambos setores houve a mesma quantidade de acidentes, no caso 50 acidentes. O setor A apresentou uma série de eventos mais variável como indicado pelo coeficiente de variação. O uso de outra medida dispersão, que não fosse relativa a média como é o caso do CV, levaria a mesma conclusão uma vez que as médias de acidentes mensais foram iguais entre os setores.

4. a) A amostra contém 150 indivíduos como indicado pela número da última barra do gráfico que representa as frequências absolutas acumuladas, b) a amplitude classe é de 5 kg uma vez que essa é a diferença entre os limites superior e inferior de cada classe, c) a média da amostra é obtida usando-se as frequências absolutas e os pontos médios de cada classe, assim

$$\bar{x} = \frac{\sum_{i=1}^k f_i \cdot x_i}{n} = \frac{2 \cdot 72,5 + 17 \cdot 77,5 + \dots + 2 \cdot 102,5}{150} = 86,36,$$

d) a moda de peso é obtida via interpolação gráfica por meio da expressão para a moda de dados agrupados em classes, baseado nos seguintes valores: limite inferior da classe modal $l_{mo} = 85$, diferença em frequências entre a classe modal e a classe à sua esquerda $d_1 = 55 - 41 = 14$, diferença em frequências entre a classe modal e a classe à sua direita $d_2 = 55 - 25 = 30$ e amplitude de classe $h_{mo} = 5$, com isso e a expressão temos que a moda é

$$m_o = l_{mo} + \frac{d_1}{d_1 + d_2} \cdot h_{mo} = 85 + \frac{14}{14 + 30} \cdot 5 = 86,59,$$

e) a mediana de dados agrupados em classe também é obtida por interpolação gráfica que requer os seguintes valores: limite inferior da classe medianal $l_{md} = 85$, posição do valor medianal $n/2 = 150/2 = 75$, frequência acumulada da classe à esquerda da classe medianal $F_{esq} = 60$, frequência absoluta da classe medianal $f_{md} = F_{md} - F_{esq} = 115 - 60 = 55$ e amplitude de classe $h_{md} = 5$, com isso e a expressão temos que a mediana é

$$m_d = l_{md} + \frac{n/2 - F_{esq}}{f_{md}} \cdot h_{md} = 85 + \frac{75 - 60}{55} \cdot 5 = 86,36,$$

f) 115 indivíduos têm menos de 90 kg pelo gráfico de frequências acumuladas, g) as classes de peso de [80,85) e [85,90) têm juntas $41 + 55 = 96$ indivíduos num total de 150, o que em termos percentuais representa $100 \cdot 96/150 = 64\%$ dos indivíduos da amostra.

