

Prova 1 - Estatística Básica (CE301)

Prof. Lineu Alberto Cavazani de Freitas

Prof. Silvia Emiko Shimakura

Nome: _____

Data: ____ / ____ / _____ GRR: _____ Assinatura: _____

Um questionário foi aplicado a uma amostra de alunos do primeiro ano de uma escola fornecendo as seguintes informações: turma, sexo, altura em cm e peso em kg. Os dados da amostra estão descritos na tabela.

Turma	Sexo	Altura	Peso
A	M	185	81
B	F	170	60
A	F	160	60
B	F	167	58
A	F	157	49
B	M	176	75
A	M	178	68
A	M	180	85
B	F	165	56
A	F	170	60

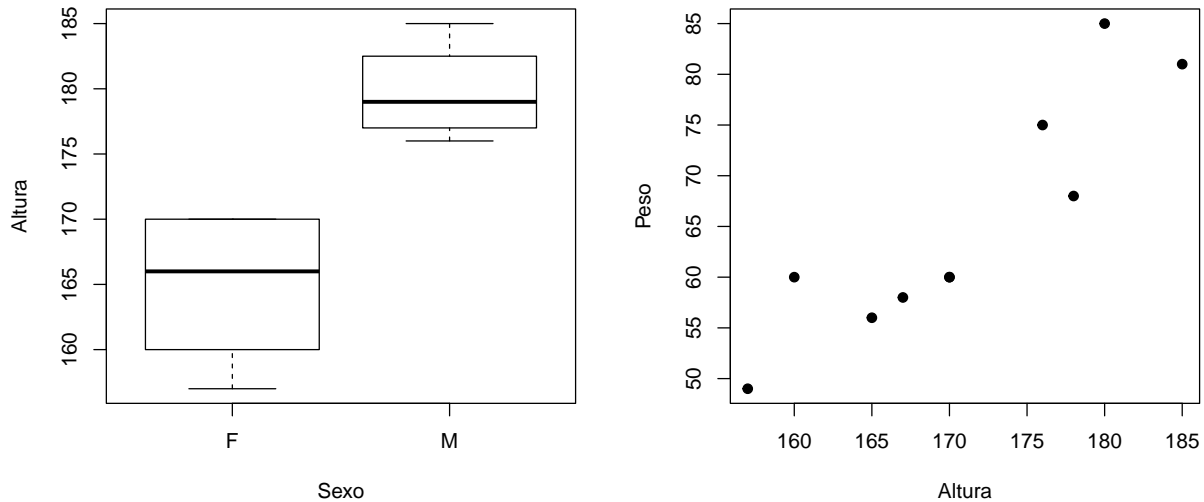
Com base nos dados, responda as questões de 1 até 11.

Nas respostas use pelo menos 2 casas decimais.

Nos gráficos atente-se para a legenda e escalas.

-
- 1) Quais são os tipos das variáveis coletadas? Classifique-as em qualitativa nominal, qualitativa ordinal, quantitativa discreta ou quantitativa contínua. (0,5 ponto)
 - 2) Considere que existia um cadastro de alunos alocados em turmas. Em cada turma os alunos foram selecionados por meio de um sorteio em que todos os elementos tinham a mesma probabilidade de fazer parte da amostra. A amostra final foi composta pela junção das amostras tomadas em cada turma. Qual o nome deste método de amostragem? Este plano de amostragem corresponde a um método probabilístico ou não probabilístico? (0,5 ponto)
 - 3) Monte uma tabela de frequências para a variável turma. Use frequências absolutas e relativas. Qual seria o gráfico mais adequado para representar esta tabela? (0,5 ponto)
 - 4) Monte uma tabela de frequências para a variável peso. Use faixas de tamanho 10, partindo de 40 até 90. Qual é a faixa modal? (0,5 ponto)
 - 5) Obtenha média e desvio padrão para as variáveis altura e peso. (1,0 ponto)
 - 6) Altura e peso são variáveis em diferentes escalas, qual delas apresenta variabilidade maior? Utilize uma medida de comparação adequada. (1,0 ponto)

- 7) Com base na tabela do item (4), esboce o histograma dos pesos. O que você conclui a respeito da simetria? (1,0 ponto)
- 8) Obtenha as quantidades necessárias e esboce o box-plot das alturas. Coloque nos eixos os valores utilizados para o esboço. O que você conclui a respeito da simetria e da presença de valores atípicos? (1,0 ponto)
- 9) Monte uma tabela de dupla entrada usando frequências absolutas para sexo e turma. O que você conclui? (1,0 ponto)
- 10) Avalie os gráficos abaixo. O que você conclui? (1,0 ponto)



- 11) Obtenha uma medida de associação entre turma e sexo. O que você conclui? (1,0 ponto)
- 12) Responda de forma sucinta: (1,0 ponto)
- Qual a diferença entre amostragem estratificada e por conglomerados?
 - Por que é melhor evitar o gráfico de setores?
 - Por que não podemos afirmar que dois conjuntos de valores com mesma média são iguais?
 - O que é um ponto atípico ou outlier?
 - Correlação implica em causalidade?

$$\bar{y} = \frac{\sum_{i=1}^n y_i}{n}$$

$$\bar{y} = \frac{\sum_{i=1}^k f_i \cdot y_i}{\sum_{i=1}^k f_i}$$

$$A = \max(y) - \min(y)$$

$$DAM_{média} = \frac{1}{n} \sum_{i=1}^n |(y_i - \bar{y})|$$

$$DAM_{mediana} = \frac{1}{n} \sum_{i=1}^n |(y_i - md)|$$

$$s^2 = \text{Var}(y) = \frac{1}{n-1} \sum_{i=1}^n (y_i - \bar{y})^2 = \frac{1}{n-1} \left(\sum_{i=1}^n y_i^2 - \frac{(\sum_{i=1}^n y_i)^2}{n} \right)$$

$$s = \sqrt{s^2}$$

$$CV = 100 \cdot \frac{s}{\bar{y}}$$

$$z = \frac{y_i - \bar{y}}{s}$$

$$H = - \sum_{i=1}^S f_r \times \ln(f_r)$$

$$Q = \sum_{i=1}^r \sum_{j=1}^s \frac{(o_{ij} - e_{ij})^2}{e_{ij}}$$

$$\text{Cov}(y_1, y_2) = \frac{1}{n-1} \sum_{i=1}^n (y_{1i} - \bar{y}_1) \cdot (y_{2i} - \bar{y}_2)$$

$$r = \frac{\sum_{i=1}^n (y_{1i} - \bar{y}_1) \cdot (y_{2i} - \bar{y}_2)}{\sqrt{\sum_{i=1}^n (y_{1i} - \bar{y}_1)^2} \cdot \sqrt{\sum_{i=1}^n (y_{2i} - \bar{y}_2)^2}} = \frac{\text{Cov}(y_1, y_2)}{\sqrt{V(y_1) \cdot V(y_2)}}$$