

IME - PPGEst/UFRGS
Prova I - EST0204 - Estatística Multivariada
Trimestre 2020/03.

Nome completo:

OBS: Nas questões 1 e 2 os cálculos e desenvolvimentos devem aparecer explícitos no arquivo de entrega.

Exercício 1. (2.0 pontos) Considere a matriz de dados bivariada abaixo.

$$\mathbf{X} = \begin{bmatrix} \mathbf{y}_1 & \mathbf{y}_2 \\ 1 & 1 \\ 2 & 2 \\ 3 & 4 \\ 4 & 9 \end{bmatrix}$$

- (a) Obtenha os vetores coluna de diferenças. A partir deles, obtenha a matriz \mathbf{S} de covariâncias e a matriz \mathbf{R} de correlações amostrais.
- (b) Calcule a medida de variância generalizada a partir de \mathbf{S} e \mathbf{R} e mostre a relação entre elas.
- (c) Calcule a variância total utilizando a matriz \mathbf{S} . Comente as diferenças entre as duas medidas.

Exercício 2. (2.0 pontos) Seja $\mathbf{X} \sim N_3(\mu, \Sigma)$ um vetor aleatório com

$$\Sigma = \begin{bmatrix} 1 & -2 & 0 \\ -2 & 5 & 0 \\ 0 & 0 & 2 \end{bmatrix} \quad \text{e} \quad \mu' = [-3, 1, 4]$$

- (a) Escreva genericamente a forma quadrática $d^2 = (\mathbf{x} - \mu)' \Sigma^{-1} (\mathbf{x} - \mu)$ decomposta em função de uma soma envolvendo autovetores e autovalores de Σ .
- (b) Encontre o comprimento dos eixos do elipsoide formado por d^2 , com 0.95 de probabilidade (Encontre os autovalores no R usando a função *eigen*).
- (c) Encontre a distribuição do vetor aleatório $\mathbf{Y}' = (Y_1, Y_2)$, onde $Y_1 = (X_1 + X_2)/2$ e $Y_2 = 2X_1 - X_2 + X_3$.

OBS: As questões 3 a 6 utilizam bancos de dados que podem ser acessados diretamente através do arquivo R **Prova1-EST0204**. Este arquivo também contém as informações sobre os bancos.

Exercício 3. (1.5 pontos) Banco de dados *carData::Anscombe*.

- (a) Aplique o teste de Shapiro de normalidade univariada e bivariada. Informe o valor p de cada teste e conclua a respeito.
- (b) Teste a hipótese a 5% de que o vetor de média populacional seja $\mu' = [210, 3225, 360, 665]$. Informe o valor da estatística de teste, o valor crítico e conclua a respeito.
- (c) Para cada variável construa os intervalos T e bonferroni e conclua a respeito.

Exercício 4. (1.5 pontos) Um programa de reforço de estudos foi avaliado a partir de uma amostra de 15 estudantes. Cada um deles realizou provas com conteúdos de matemática, física e química antes e depois de serem submetidos ao programa de reforço.

- (a) Aplique o teste de Shapiro de normalidade multivariada. Informe o valor p de cada teste e conclua a respeito.
- (b) O programa de reforço teve influência nas notas? Teste com 1% de significância. Informe o valor da estatística de teste, o valor crítico e conclua a respeito.
- (c) Construa os intervalos T e bonferroni e conclua a respeito.

Exercício 5. (1.5 pontos) Banco de dados *mtcars*.

- (a) Aplique o teste de Shapiro de normalidade multivariada dentro de cada grupo (motores em V e convencionais). Informe o valor p de cada teste e conclua a respeito. Em caso de não normalidade, verifique se os dados transformados (via logaritmo natural) aderem à hipótese de normalidade. Neste caso realizar as análises subsequentes utilizando os dados transformados.
- (b) Existe a 5% diferença significativa nas variáveis analisadas entre os 2 tipos de motores? Informe o valor da estatística de teste, o valor crítico e conclua a respeito.
- (c) Construa os intervalos T e bonferroni e conclua a respeito.

Exercício 6. (1.5 pontos) Banco de dados *carData::Salaries*.

- (a) Realize a MANOVA a 5% e conclua a respeito. Informe o valor da estatística de Wilks e seu valor p .
- (b) Realize a ANOVA em cada variável a 5% e conclua a respeito. Informe o valor da estatística F e seu valor p .
- (c) Faça uma conclusão geral sobre os resultados.