Universidade Federal do Ceará

Departamento de Estatística e Matemática Aplicada

Prof. : Juvêncio S. Nobre

CC290 - Modelos de Regressão I - 2022.2

Lista de exercícios (de boas vindas) # 0: Revisão de teoria das matrizes, probabilidade e inferência

Distribuição: 17/08/2022

Entrega: $\aleph_1$

Parte 1: Teoria das Matrizes

1. Considere $\mathbf{x} = (x_1, \ldots, x_n)^\top$ um vetor de dimensão $n \times 1$. Podemos definir a norma do vetor $\mathbf{x}$ por

$$\|\mathbf{x}\| = \sqrt{\mathbf{x}^\top \mathbf{x}} = \sqrt{\sum_{i=1}^{n} x_i^2}.$$

Mostre que:

i) $\|\mathbf{x}\| \geq 0, \forall \mathbf{x} \in \mathbb{R}^n$.

ii) $\|\mathbf{x}\| = 0 \Leftrightarrow \mathbf{x} = \mathbf{0}$.

iii) $\|c\mathbf{x}\| = |c| \, \|\mathbf{x}\|$.

iv) $\|\mathbf{x} + \mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|, \forall \mathbf{x}, \mathbf{y} \in \mathbb{R}^n$.

2. Considere $\mathbf{x}$ e $\mathbf{y}$ dois vetores reais de dimensão $n \times 1$. Prove a desigualdade de Cauchy-Scharwz:

$$(\mathbf{x}^\top \mathbf{y})^2 \leq (\mathbf{x}^\top \mathbf{x})(\mathbf{y}^\top \mathbf{y}).$$

Em que ocasiões a igualdade é válida?

Sugestão: Use o fato de que $\|\mathbf{x} + \lambda \mathbf{y}\| \geq 0, \forall \mathbf{x}, \mathbf{y} \in \mathbb{R}^n, \lambda \in \mathbb{R}$.

3. O traço de uma matriz quadrada $\mathbf{A}$ de ordem n, denotado por $\mathrm{tr}(\mathbf{A})$, é definido como a soma dos elementos da diagonal principal de $\mathbf{A}$, isto é,

$$\mathrm{tr}(\mathbf{A}) = \sum_{i=1}^{n} a_{ii}.$$

Considere $\mathbf{A}$ e $\mathbf{B}$ duas matrizes quadradas de ordem n e $\mathbf{C}$ e $\mathbf{D}$ duas matrizes de ordens $m \times n$ e $n \times m$, respectivamente. Mostre que:

i) $\text{tr}(\mathbf{A}^\top) = \text{tr}(\mathbf{A})$.

ii) $\text{tr}(\mathbf{A} + \mathbf{B}) = \text{tr}(\mathbf{A}) + \text{tr}(\mathbf{B})$.

iii) $\text{tr}(\mathbf{CD}) = \text{tr}(\mathbf{DC})$.

iv) $\text{tr}(\mathbf{A}^\top\mathbf{A}) = \text{tr}(\mathbf{A}\mathbf{A}^\top) = \sum_{i=1}^n \sum_{j=1}^n a_{ij}^2$.

v) Considerando $\mathbf{x}$ um vetor de dimensão $n \times 1$, mostre que $\|\mathbf{x}\|^2 = \text{tr}(\mathbf{x}\mathbf{x}^\top)$.

4. Faça um resumo sobre:

i) Decomposição espectral.

ii) Decomposição de `Cholesky`.

iii) Decomposição em valor singular (decomposição `SVD`).

iv) Decomposição de `Schur`.

Adicionamente, mostre que as decomposições espectral, de `Cholesky` e `SVD` são casos particulares da decomposição de `Schur`. Apresente os comandos no `R` (com exemplos), para obter tais decomposições.

5. Considere $\mathbf{A}$ uma matriz quadrada de ordem n cujos respectivos auto-valores são dados por $\lambda_1, \ldots, \lambda_n$. Prove que

$$\text{tr}(\mathbf{A}) = \sum_{i=1}^n \lambda_i \;\text{ e }\; |\mathbf{A}| = \prod_{i=1}^n \lambda_i.$$

Sugestão: Utilize a decomposicão de `Schur` da matriz $\mathbf{A}$.

6. Uma matriz quadrada $\mathbf{A}$ é dita ser *idempotente* se $\mathbf{A}^2 = \mathbf{A}$. Considere $\mathbf{A}$ uma matriz idempotente. Mostre que:

i) $|\mathbf{A}| = 0$ ou 1.

ii) Os autovalores de $\mathbf{A}$ são iguais a zero ou 1.

iii) $\mathbf{I}_n - \mathbf{A}$ é idempotente e $\mathbf{A}(\mathbf{I}_n - \mathbf{A}) = (\mathbf{I}_n - \mathbf{A})\mathbf{A} = \mathbf{0}_n$, em que $\mathbf{I}_n$ e $\mathbf{0}_n$ representam, respectivamente a matriz identidade de ordem n e a matriz quadrada de ordem n com todos elementos iguais a zero.

iv) Prove que:

a) $\text{posto}(\mathbf{A}) = \text{tr}(\mathbf{A})$.

b) Se $\text{posto}(\mathbf{A}) = n$, então $\mathbf{A} = \mathbf{I}_n$.

Sugestão: Utilize a decomposição de `Schur` da matriz $\mathbf{A}$.

7. Mostre que se $\mathbf{A}^\top \mathbf{A} = \mathbf{A}$, então $\mathbf{A}$ é simétrica e idempotente.

8. Considere $\mathbf{A}$ e $\mathbf{B}$ duas matrizes de ordens $n \times m$ e $p \times q$, respectivamente. O produto de *Kronecker* entre as matrizes $\mathbf{A}$ e $\mathbf{B}$, denotado por $\mathbf{A} \otimes \mathbf{B}$, é dado pela seguinte matriz de ordem $np \times mq$:

$$\mathbf{A} \otimes \mathbf{B} = \begin{pmatrix} a_{11}\mathbf{B} & a_{12}\mathbf{B} & \cdots & a_{1m}\mathbf{B} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1}\mathbf{B} & a_{n2}\mathbf{B} & \cdots & a_{nm}\mathbf{B} \end{pmatrix}.$$

Mostre que (assuma que as ordens das matrizes são tais que as operações sejam bem definidas) :

i) $(a\mathbf{A}) \otimes (b\mathbf{B}) = ab(\mathbf{A} \otimes \mathbf{B})$, para quaisquer escalares a, b.

ii) $(\mathbf{A} + \mathbf{C}) \otimes \mathbf{B} = \mathbf{A} \otimes \mathbf{B} + \mathbf{C} \otimes \mathbf{B}$.

iii) $(\mathbf{A} \otimes \mathbf{B}) \otimes \mathbf{C} = \mathbf{A} \otimes (\mathbf{B} \otimes \mathbf{C})$

iv) $(\mathbf{A} \otimes \mathbf{B})^\top = \mathbf{A}^\top \otimes \mathbf{B}^\top$.

v) $(\mathbf{A} \otimes \mathbf{B})(\mathbf{C} \otimes \mathbf{D}) = \mathbf{AC} \otimes \mathbf{BD}$.

vi) Para matrizes quadradas $\mathbf{A}$ e $\mathbf{B}$: $\text{tr}(\mathbf{A} \otimes \mathbf{B}) = \text{tr}(\mathbf{A})\text{tr}(\mathbf{B})$.

vii) Se $\mathbf{A}$ e $\mathbf{B}$ são duas matrizes quadradas de ordens n e m, respectivamente, mostre que $|\mathbf{A} \otimes \mathbf{B}| = |\mathbf{A}|^m |\mathbf{B}|^n$.

9. A soma direta de matrizes $\mathbf{A}_i$, com dimensão $(n_i \times m_i)$, $i = 1, \ldots, n$ é a matriz $(\sum_{i=1}^n n_i \times \sum_{i=1}^n m_i)$ definida por:

$$\bigoplus_{i=1}^n \mathbf{A}_i = \begin{pmatrix} \mathbf{A}_1 & \mathbf{0} & \ldots & \mathbf{0} \\ \mathbf{0} & \mathbf{A}_2 & \ldots & \mathbf{0} \\ \vdots & \vdots & \vdots & \vdots \\ \mathbf{0} & \mathbf{0} & \ldots & \mathbf{A}_n \end{pmatrix}.$$

Mostre que:

i) $\text{tr}\left(\mathbf{A} \bigoplus \mathbf{B}\right) = \text{tr}(\mathbf{A}) + \text{tr}(\mathbf{B})$, com $\mathbf{A}$ e $\mathbf{B}$ representando matrizes quadradas de mesma dimensão.

ii) $\left(\bigoplus_{i=1}^{n} \mathbf{A}_i\right)\left(\bigoplus_{i=1}^{n} \mathbf{B}_i\right) = \bigoplus_{i=1}^{n} \mathbf{A}_i\mathbf{B}_i$, com $\mathbf{A}_i$ e $\mathbf{B}_i$ representando matrizes de mesma dimensão.

iii) $\mathbf{I}_k \otimes \mathbf{A} = \bigoplus_{i=1}^{k} \mathbf{A}$.

iv) $|\mathbf{A} \bigoplus \mathbf{B}| = |\mathbf{A}||\mathbf{B}|$.

10. A operação de vetorização de uma matriz $\mathbf{A} = (\mathbf{a}_1 \ldots \mathbf{a}_n)$, denotada por $\text{vec}(\mathbf{A})$, consiste em "empilhar" seus elementos na forma de um vetor

$$\text{vec}(\mathbf{A}) = (\mathbf{a}_1^\top, \ldots, \mathbf{a}_n^\top)^\top.$$

Para uma matriz simétrica $\mathbf{A}$, o operador $\text{vech}(\mathbf{A})$ consiste em empilhar todos os elementos **distintos** de $\mathbf{A}$ em um vetor.

Sejam $\mathbf{A}$, $\mathbf{B}$ matrizes de mesma dimensão e $\mathbf{a}$ e $\mathbf{b}$ vetores de ordens $n \times 1$ e $m \times 1$, respectivamente. Prove as seguintes propriedades:

i) $\text{vec}(\mathbf{a}^\top) = \text{vec}(\mathbf{a})$.

ii) $\text{vec}(\mathbf{a}\mathbf{b}^\top) = \mathbf{b} \otimes \mathbf{a}$.

iii) $\text{tr}(\mathbf{A}^\top\mathbf{B}) = \text{vec}(\mathbf{A})^\top\text{vec}(\mathbf{B})$.

iv) Considere $\mathbf{A}$ uma matriz simétrica 3×3, com elementos $a_{11}, \cdots, a_{33}$. Obtenha $\text{vec}(\mathbf{A})$ e $\text{vech}(\mathbf{A})$.

11. Considere $\mathbf{A}$ uma matriz simétrica de ordem n. A norma de *Frobenius* (também denotada por norma de Hilbert−Schmidt) da matriz $\mathbf{A}$, denotada por $\|\mathbf{A}\|_F$, é definida por

$$\|\mathbf{A}\|_F = \sqrt{\sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}^2}.$$

Mostre que

$$\|\mathbf{A}\|_F = \|\text{vec}(\mathbf{A})\| = \sqrt{\text{tr}(\mathbf{A}^\top\mathbf{A})}.$$

12. Considere $\mathbf{A}$ e $\mathbf{B}$ duas matrizes simétricas de ordem n e c um escalar. Mostre que:

i) $\|\mathbf{A}\|_F \geq 0, \forall \mathbf{A} \in \mathcal{M}_{\mathbb{R}}(n,n) := \mathcal{M}_{\mathbb{R}}(n)$.

ii) $\|\mathbf{A}\|_F = 0 \Leftrightarrow \mathbf{A} = \mathbf{0}$.

iii) $\|c\mathbf{A}\|_F = |c| \, \|\mathbf{A}\|_F$.

iv) $\|\mathbf{A} + \mathbf{B}\|_F \leq \|\mathbf{A}\|_F + \|\mathbf{B}\|_F, \forall \mathbf{A}, \mathbf{B} \in \mathcal{M}_{\mathbb{R}}(n)$.

13. Considere $\mathbf{A}$ uma matriz quadrada de ordem n. A matriz $\mathbf{A}$ é dita ser positiva (não-negativa) definida, denotada por $\mathbf{A} \succ (\succeq) 0$, se

$$\mathbf{x}^\top \mathbf{A}\mathbf{x} > (\geq) 0, \ \forall \mathbf{x} \neq \mathbf{0}.$$

Considere $\mathbf{X}$ uma matriz $n \times p$ $(n \leq p)$ de posto completo. Mostre que

i) $\mathbf{X}^\top\mathbf{X}$ é simétrica.

ii) $\mathbf{X}^\top\mathbf{X}$ é positiva definida.

iii) Usando o fato de que uma matriz é positiva definida, se e somente se, todos seus autovalores são positivos, mostre que $\mathbf{X}^\top\mathbf{X}$ é inversível.

iv) $\mathbf{H} = \mathbf{X}(\mathbf{X}^\top\mathbf{X})^{-1}\mathbf{X}^\top$ é simétrica e idempotente.

v) $\mathbf{I}_n - \mathbf{H}$ é simétrica e idempotente.

vi) posto$(\mathbf{I}_n - \mathbf{H}) = n - p$.

14. Considere a matriz

$$\mathbf{A} = \begin{pmatrix} X & 1 \\ 1 & X \end{pmatrix}.$$

i) Considerando X um escalar, determine para quais valores de X a matriz $\mathbf{A}$ é positiva definida.

ii) Se $X \sim \mathcal{U}(-2, 2)$, calcule a probabilidade da matriz $\mathbf{A}$ ser positiva definida.

iii) Repita o item ii), considerando que $X \sim \mathcal{N}(0, 1)$.

15. Considere $f(\mathbf{X}) : \mathbb{R}^{m\times n} \to \mathbb{R}$ uma função real de uma matriz $\mathbf{X} = (x_{ij})$ de dimensão $n \times m$. A derivada de f com respeito a $\mathbf{X}$ é definida como sendo a matriz $n \times m$ de derivadas $\partial f/\partial x_{ij}$, *i.e.*,

$$\frac{\partial f(\mathbf{X})}{\partial \mathbf{X}} := \begin{pmatrix} \frac{\partial f(\mathbf{X})}{\partial x_{11}} & \frac{\partial f(\mathbf{X})}{\partial x_{21}} & \cdots & \frac{\partial f(\mathbf{X})}{\partial x_{1m}} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f(\mathbf{X})}{\partial x_{n1}} & \frac{\partial f(\mathbf{X})}{\partial x_{n2}} & \cdots & \frac{\partial f(\mathbf{X})}{\partial x_{nm}} \end{pmatrix}.$$

Considere $\mathbf{x} = (x_1, \ldots, x_n)^\top$ e $\mathbf{a} = (a_1, \ldots, a_n)^\top$ vetores reais de dimensão $n \times 1$ e $\mathbf{A}$ uma matriz quadrada de ordem n. Prove que

$$\begin{aligned} \frac{\partial \mathbf{a}^\top \mathbf{x}}{\partial \mathbf{x}} &= \frac{\partial \mathbf{x}^\top \mathbf{a}}{\partial \mathbf{x}} = \mathbf{a} \\ \frac{\partial \mathbf{x}^\top \mathbf{A}\mathbf{x}}{\partial \mathbf{x}} &= (\mathbf{A} + \mathbf{A}^\top)\mathbf{x} \\ \frac{\partial \exp(-\frac{1}{2}\mathbf{x}^\top \mathbf{A}\mathbf{x})}{\partial \mathbf{x}} &= -\exp\left(-\frac{1}{2}\mathbf{x}^\top \mathbf{A}\mathbf{x}\right)\mathbf{A}\mathbf{x}\text{ , se } \mathbf{A} \text{ é simétrica.} \end{aligned}$$

16. Considere a função

$$g(\boldsymbol{\beta}) = (\mathbf{y} - \mathbf{X}\boldsymbol{\beta})^\top \mathbf{V}^{-1}(\mathbf{y} - \mathbf{X}\boldsymbol{\beta}),$$

em que $\mathbf{y}$ e $\boldsymbol{\beta}$ são vetores de dimensões $n \times 1$ e $p \times 1$ $(n > p)$, respectivamente, $\mathbf{X}$ uma matriz de dimensão $n \times p$ de posto completo e $\mathbf{V}$ uma matriz simétrica de ordem n positiva definida funcionalmente independente de $\boldsymbol{\beta}$. Obtenha:

i) $\frac{\partial g(\boldsymbol{\beta})}{\partial \boldsymbol{\beta}}$.

ii) $\frac{\partial^2 g(\boldsymbol{\beta})}{\partial \boldsymbol{\beta}\partial \boldsymbol{\beta}^\top}$.

17. Mostre que a matriz $\mathbf{I}_n - \mathbf{J}_n/n$ é simétrica, idempotente e não-negativa definida.

18. Faça um resumo sobre **matriz de projeção**.

19. Considere $\mathbf{X}$ uma matriz $n \times p$ $(n \leq p)$ de posto completo e considere a matriz de **projeção** $\mathbf{H} = \mathbf{X}(\mathbf{X}^\top\mathbf{X})^{-1}\mathbf{X}^\top$ com elementos denotados por h_{ij}, $i, j \in \{1, ..., n\}$. Mostre que para $i = 1, \cdots, n$:

i) $h_{ii} = h_{ii}^2 + \sum_{j\neq i} h_{ij}^2$.

ii) $0 \leq h_{ii} \leq 1$.

20. Considere a matriz $\mathbf{X}$ definida por

$$\mathbf{X}^\top = \begin{pmatrix} 1 & 1 & \cdots & 1 \\ x_1 & x_2 & \cdots & x_n \end{pmatrix}$$

e a matriz de projeção $\mathbf{H}$ definida na questão anterior.

i) Que condição deve ser satisfeita para que a matriz $\mathbf{X}$ seja de posto completo?

ii) Obtenha algebricamente o valor de $h_{ij}, i, j = 1, \cdots, n$.

iii) Com base no item ii), mostre que $n^{-1} \leq h_{ii} \leq 1, i = 1, \cdots, n$

21. Considere $\mathbf{x} = (x_1, \ldots, x_n)^\top$ um vetor de n observações. Mostre que a média amostral e a variância amostral podem escritas na seguinte forma matricial:

$$\begin{aligned} \bar{x} &= \frac{1}{n}\sum_{i=1}^{n} x_i = \frac{1}{n}\mathbf{1}_n^\top \mathbf{x} \\ S_x^2 &= \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2 = \frac{1}{n-1}(\mathbf{x} - \mathbf{1}_n\bar{x})^\top(\mathbf{x} - \mathbf{1}_n\bar{x}) \\ &= \frac{1}{n-1}\mathbf{x}^\top(\mathbf{I}_n - \mathbf{J}_n)\mathbf{x}, \end{aligned}$$

em que $\mathbf{1}_n$ representa um vetor de dimensão $n \times 1$ com todos elementos iguais a 1 e $\mathbf{J}_n = \mathbf{1}_n\mathbf{1}_n^\top$.

22. Mostre que a variância amostral pode ser reescrita como uma forma quadrática (Questão # 21).

23. Mostre que a covariância amostral pode ser reescrita como uma forma bilinear.

24. Faça um resumo sobre o método dos **multiplicadores de Lagrange**.

25. Apresente dois exemplos de aplicação do método dos **multiplicadores de Lagrange**.

Sugestões de referências:

1. Aggarwal, C.C. (2020). *Linear Algebra and Optimization for Machine Learning*. New York: Springer.

2. Banerjee, S. and Roy, A. (2014). *Linear Algebra and Matrix Analysis for Statistics*. Boca Raton: CRC Press.

3. Gentle, J.A. (2017). *Matrix Algebra: Theory, Computations and Applications in Statistics*, 2nd edition. New York: Springer.

4. Harville, D.A. (2000). *Matrix Algebra from Statistician's Perspective*. New York: Springer.

5. Harville, D.A. (2018). *Linear Models and the Relevant Distributions and Matrix Algebra*. Boca Raton: CRC Press.

6. Magnus, J.R. and Neudecker, H. (1999). *Matrix Differential Calculus with Applications in Statistics and Econometrics*, 2nd Edition. New York: John Wiley & Sons.

7. Puntanen, S., Styan, G.P.H. and Isotalo, J. (2011). *Matrix Tricks for Linear Statistical Models: Our Personal Top Twenty*. New York: Springer.

8. Rao, C.R. and Rao, M.B. (1998). *Matrix Algebra and Its Applications to Statistics and Econometrics*. New York: World Scientific.

9. Schott, J.R. (2017). *Matrix Analysis for Statistics*, 3rd edition. New York: John Wiley & Sons.

10. Searle, S.R. and Khuri, A. I. (2017). *Matrix Algebra Useful for Statistics*, 2nd edition. New York: John Wiley & Sons.

Parte 2: Probabilidade e Inferência

26. Considere X_1 e X_2 V.A's independentes. Encontre a distribuição condicional de $X_1|X_1+X_2 = y$, quando:

i) $X_i \sim \mathrm{P}(\lambda_i),\ i = 1, 2;$

ii) $X_i \sim \mathrm{B}(n_i, p),\ i = 1, 2;$

27. Usando o resultado da questão anterior e as propriedades da esperança e variância condicional, determine $\mathbb{E}[X_1]$ e $\mathrm{Var}[X_1]$.

28. Um mineiro está preso numa mina contendo 3 portas. A porta 1 o conduz à saída após 2 horas de caminhada. A porta 2 o conduz a um túnel que o retorna ao mesmo lugar da mina após 3 horas. A porta 3 o conduz a um túnel que o retorna ao mesmo lugar após 5 horas. Se o mineiro escolhe qualquer porta aleatoriamente todas as vezes, qual é o tempo esperado que ele vai levar para sair da mina? Calcule a variância desse tempo de saída.

29.(Soma aleatória de variáveis aleatórias) Considere $\{X_n\}_{n\geq 1}$ uma sequencia de variáveis aleatórias iid e N uma variável aleatória discreta não-negativa, independente de X_i, $\forall i \in \mathbb{N}$. Defina $S_N := \sum_{i=1}^{N} X_i$. Mostre que

$$\mathbb{E}[S_N] = \mathbb{E}[N]\mathbb{E}[X_1] \text{ e } \mathrm{Var}[S_N] = \mathrm{Var}[X_1]\mathbb{E}[N] + (\mathbb{E}[X_1])^2\mathrm{Var}[N].$$

30. Considere $X_1, \ldots, X_n$ variáveis aleatórias definidas em um mesmo espaço de probabilidades e $\mathrm{Cov}(\cdot,\cdot)$ o operador covariância. Prove que:

i) $\mathrm{Cov}(X_i, X_i) = \mathrm{Var}[X_i]$.

ii) $\mathrm{Cov}(X_i, X_j) = \mathrm{Cov}(X_j, X_i)$.

iii) $\mathrm{Cov}(aX_i, bX_j) = ab\mathrm{Cov}(X_i, X_j)$, para quaisquer constantes a e b.

iv) Para qualquer constante a, $\mathrm{Cov}(X_i, a) = 0$.

v) $\mathrm{Cov}(X_1, X_3 + X_4) = \mathrm{Cov}(X_1, X_3) + \mathrm{Cov}(X_1, X_4)$.

vi) $\mathrm{Cov}(X_1 + X_2, X_3 + X_4) = \mathrm{Cov}(X_1, X_3) + \mathrm{Cov}(X_1, X_4) + \mathrm{Cov}(X_2, X_3) + \mathrm{Cov}(X_2, X_4)$.

vii) $\mathrm{Var}(X_1 + X_2) = \mathrm{Var}(X_1) + \mathrm{Var}(X_2) + 2\mathrm{Cov}(X_1, X_2)$.

viii) $\mathrm{Cov}\left(\sum_{i=1}^{n} a_i X_i, \sum_{j=1}^{n} b_j Y_j\right) = \sum_{i=1}^{n}\sum_{j=1}^{n} a_i b_j \mathrm{Cov}(X_i, Y_j)$.

31. Exiba um vetor aleatório discreto (X_1, X_2) tal que:

i) $\mathrm{Cov}(X_1, X_2) = 2022$.

ii) $\mathrm{Cov}(X_1, X_2) = 2022$, com a restrição de que de X_i só assuma dois valores (com probabilidade positiva), $i = 1, 2$.

32. Seja $\mathbf{X} = (X_1, X_2, X_3)$ um vetor aleatório cujas componentes são iids com distribuição $\mathcal{N}(\mu, \sigma^2)$. Defina $\mathbf{Y} = (Y_1, Y_2)$ em que $Y_1 := X_2 - X_1$ e $Y_2 := X_3 - X_1$. Determinar a covariância e o coeficiente de correlação das componentes de $\mathbf{Y}$.

33. Forneça um exemplo de um vetor aleatório absolutamente contínuo bivariado com componentes dependentes e não correlacionadas.

34. Forneça um exemplo em que correlação nula implica **independência**.

35. Considere um vetor aleatório $\mathbf{X} = (X_1, X_2)$ cuja respectiva função geradora de momentos é dada por

$$M_{\mathbf{X}}(\mathbf{t}) = \left(1 - \sum_{i=1}^{2} p_i + \sum_{i=1}^{2} p_i e^{t_i}\right)^n.$$

i) Identifique as marginais de X_1 e X_2.

ii) Usando $M_{\mathbf{X}}(\mathbf{t})$, $M_{X_1}(t_1)$ e $M_{X_2}(t_2)$, discuta a independência das componentes de $\mathbf{X}$.

iii) Calcule ρ_{12} e o interprete.

36. Considere um vetor aleatório $\mathbf{X} = (X_1, X_2)$ cuja respectiva função geradora de momentos é dada por

$$M_{\mathbf{X}}(\mathbf{t}) = \exp\left\{\mu_1 t_1 + \mu_2 t_2 + \frac{1}{2}\left(\sigma_1^2 t_1^2 + \sigma_2^2 t_2^2 + 2\rho\sigma_1\sigma_2 t_1 t_2\right)\right\},$$

com $\mu_1, \mu_2 \in \mathbb{R}$, $\sigma_1, \sigma_2 \in \mathbb{R}^+$ e $\rho \in (-1, 1)$.

i) Identifique as marginais de X_1 e X_2.

ii) Usando $M_{\mathbf{X}}(\mathbf{t})$, $M_{X_1}(t_1)$ e $M_{X_2}(t_2)$, discuta a independência das componentes de $\mathbf{X}$.

iii) Forneça uma condição necessária e suficiente para garantir a independência das componentes de $\mathbf{X}$.

iv) Calcule $\rho_{12} := \mathrm{Corr}(X_1, X_2)$ e o interprete.

37. Considere $\mathbf{X} = (X_1, X_2)$ um vetor aleatório absolutamente contínuo com densidade dada por

$$f_{(X_1,X_2)}(x_1, x_2) = \frac{1}{2} x_1 x_2 \mathbb{1}_{(0,2)}(x_1) \mathbb{1}_{(0,x_1)}(x_2).$$

i) Encontre as marginais de X_1 e X_2.

ii) As variáveis X_1 e X_2 são independentes?

ii) Determine $\mathrm{Cov}(X_1, X_2)$.

38. Considere $U \sim \mathcal{U}(0, 1)$ e $Z \sim \exp(1)$, independentes. Mostre que

$$\begin{aligned} X &= \sqrt{2Z}\cos(2\pi U) \\ Y &= \sqrt{2Z}\mathrm{sen}(2\pi U), \end{aligned}$$

são iid com distribuição $\mathcal{N}(0,1)$.

39. Considere $\mathbf{X} = (X_1, X_2) \sim \mathcal{N}_2(0,0,1,1,1/2)$. Encontre a distribuição conjunta de $Y_1 := X_1 + X_2$ e $Y_2 = X_1 - X_2$.

40. Seja $\mathbf{X} = (X_1, X_2) \sim \mathcal{N}_2(3,1,16,25,6/10)$. Determinar:

i) $\mathbb{P}(3 \leq X_2 \leq 8)$.

ii) $\mathbb{P}(3 \leq X_2 \leq 8 | X_1 = 7)$.

iii) $\mathbb{P}(-3 \leq X_1 \leq 3)$.

iv) $\mathbb{P}(-3 \leq X_1 \leq 3 | X_2 = -4)$.

v) Obtenha as distribuições condicionais de X_2 dado X_1 e de X_1 dado X_2.

vi) Obtenha a distribuição conjunta de $Y_1 = 2X_1 - X_2$ e $Y_2 = -X_1 + 3X_2$.

41. Considere $\mathbf{X} \sim \mathcal{N}_2(\boldsymbol{\mu}, \boldsymbol{\Sigma})$.

i) Mostre que $f_{\mathbf{X}}(\mathbf{x})$ é uma legítima fdp.

ii) Usando $f_{\mathbf{X}}(\mathbf{x})$, determine as distribuições marginais.

iii) Obtenha as distribuições condicionais.

iv) Prove que $\rho = \text{Corr}(X_1, X_2) = 0$ é uma condição necessária e suficiente para garantir a independência das componentes de $\mathbf{X}$.

v) Encontre $M_{\mathbf{X}}(\mathbf{t}), \forall \mathbf{t} \in \mathbb{R}^2$.

42. Usando um software de sua preferência, plote a função densidade de probabilidade e a respectiva curva de nível (mapa de contorno) da $\mathcal{N}_2(0,0,1,1,\rho)$, para $\rho = 0, \pm 0.1, \pm 0.5, \pm 0.9$. Interprete os gráficos.

43. Apresente as 3 definições equivalentes da distribuição normal multivariada (sem a necessidade da fdp) e suas propriedades apresentadas de forma matricial.

44. Faça um breve ensaio sobre distribuição t-Student multivariada e suas propriedades apresentadas de forma matricial.

45. Usando um software de sua preferência, plote a função densidade de probabilidade e a respectiva curva de nível (mapa de contorno) da $t_\nu(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, para $\boldsymbol{\mu} = \mathbf{0}$, $\boldsymbol{\Sigma} = \rho \mathbf{I}_2$, $\rho = 0, \pm 0.1, \pm 0.5, \pm 0.9$ e $\nu = 1, 2, 10, 30$ e 100. Interprete os gráficos.

46. Apresente um resumo sobre:

i) Distribuições χ^2 e F não centrais.

ii) Distribuições de formas lineares e quadráticas (sob suposição de normalidade).

47.(Família de localização/posição) Seja $\mathcal{F}_\theta = \{f(\cdot;\theta), \theta \in \mathbb{R}\}$ uma família de densidades. O parâmetro θ é definido ser um *parâmetro de localização/posição* se e somente se a densidade $f(x;\theta)$ poder ser escrita da forma $f(x;\theta) = g(x-\theta)$, com g uma função conhecida, denominada de função *geradora*. Adicionalmente, diz-se que a densidade $f(\cdot;\theta)$ (ou a V.A. associada) é um membro da família de localização. Mostre que as seguintes V.A's pertencem à família de localização, identificando o parâmetro de localização e a respectiva função geradora:

i) $X \sim \mathcal{N}(\mu, \sigma^2)$, com σ conhecido.

ii) $X \sim \text{Cauchy}(\mu, \sigma)$, com σ conhecido, i.e.,

$$f_X(x) = \frac{\sigma}{\pi\left[\sigma^2 + (x-\mu)^2\right]} \mathbb{1}_{\mathbb{R}}(x).$$

iii) $X \sim t_k(\mu, \sigma)$, com k e σ conhecidos, i.e.,

$$f_X(x) = \frac{1}{\sqrt{k} B(1/2, k/2)\sigma} \left(1 + \frac{\left(\frac{x-\mu}{\sigma}\right)^2}{k}\right)^{-(k+1)/2} \mathbb{1}_{\mathbb{R}}(x).$$

iv) $X \sim \text{Laplace}(\mu, \sigma)$, com σ conhecido, i.e.,

$$f_X(x) = \frac{1}{2\sigma} e^{-\frac{|x-\mu|}{\sigma}} \mathbb{1}_{\mathbb{R}}(x).$$

48.(Família de escala) Seja $\mathcal{F}_\theta = \{f(\cdot;\theta), \theta \in \mathbb{R}\}$ uma família de densidades. O parâmetro θ é definido ser um *parâmetro de escala* se e somente se a densidade $f(x;\theta)$ poder ser escrita da forma $f(x;\theta) = \theta^{-1} g(x/\theta)$, com g uma função conhecida, denominada de função *geradora*. Adicionalmente, diz-se que a densidade $f(\cdot;\theta)$ (ou a V.A. associada) é um membro da família de escala. Mostre que as seguintes V.A's pertencem à família de escala, identificando o parâmetro de escala e a respectiva função geradora:

i) $X \sim \mathcal{U}(-\theta, \theta)$.

ii) $X \sim \exp(\lambda)$.

iii) $X \sim \mathcal{N}(\mu, \sigma^2)$, com μ conhecido.

iv) $X \sim t_k(\mu, \sigma)$, com k e μ conhecidos.

v) $X \sim \text{Laplace}(\mu, \sigma)$, com μ conhecido.

49.(Família de localização-escala) Seja $\mathcal{F}_\theta = \{f(\cdot;\theta_1, \theta_2), \theta_1 \in \mathbb{R},\ \theta_2 > 0\}$ uma família de densidades. Os parâmetros θ_1 e θ_2 são definidos, respectivamente, parâmetros de *localização* e de *escala* se e somente se a densidade $f(x;\theta_1,\theta_2)$ poder ser escrita da forma $f(x;\theta) = \theta_2^{-1} g((x-\theta_1)/\theta_2)$, com g uma função conhecida, denominada de função *geradora*. Adicionalmente, diz-se que a densidade $f(\cdot;\theta_1,\theta_2)$ (ou a V.A. associada) é um membro da família de localização-escala. Mostre que as seguintes V.A's pertencem à família de localização-escala, identificando os parâmetros de localização e de escala, e a respectiva função geradora:

i) $X \sim \mathcal{N}(\mu, \sigma^2)$.

ii) $X \sim t_k(\mu, \sigma)$, com k conhecido.

iii) $X \sim \text{Laplace}(\mu, \sigma)$.

iv) $X \sim \text{Logística}(\mu, \sigma)$, i.e.,

$$f_X(x) = \frac{e^{-\left(\frac{x-\mu}{\sigma}\right)}}{\sigma\left\{1 + e^{-\left(\frac{x-\mu}{\sigma}\right)}\right\}^2} \mathbb{1}_{\mathbb{R}}(x)$$

50.(Distribuições simétricas) Considere X uma variável aleatória com suporte em $\mathbb{R}$, com parâmetro de localização $\mu \in \mathbb{R}$ e de escala $\phi > 0$. Dizemos que X pertence a família de distribuições simétricas, se sua densidade é escrita da seguinte forma

$$f_X(x) = \frac{1}{\sqrt{\phi}} g\left\{\frac{(x-\mu)^2}{\phi}\right\} \mathbb{1}_{\mathbb{R}}(x),$$

para alguma função $g(\cdot)$ denominada função geradora de densidades, com $g(u) > 0$ se $u > 0$ e $\int_0^\infty u^{-1/2} g(u) d_u = 1$. Mostre que as seguintes V.A's pertencem à família de de distribuições simétricas, identificando o parâmetro de localização o de escala e a respectiva função geradora:

i) $X \sim \mathcal{N}(\mu, \sigma^2)$.

ii) $X \sim \text{Cauchy}(\mu, \sigma)$.

iii) $X \sim t_k(\mu, \sigma)$, com k conhecido.

iv) $X \sim \text{Logística}(\mu, \sigma)$.

Sugestão de leitura: Para maiores detalhes sobre esta classe de distribuições, bem como de aplicações em modelos de regressão veja: Cysneiros, F.J., Paula, G.A. e Galea, M. (2007). *Modelos simétricos aplicados*. ABE, 9ª Escola de Modelos de Regressão, Águas de São Pedro-SP, Brasil.

51.(Família exponencial) Considere X uma variável aleatória com suporte em $\mathcal{A} \subset \mathbb{R}$ funcionalmente independente de θ. Dizemos que X pertence a família exponencial unidimensional de parâmetro θ , denotada por $X \sim \text{FE}(\theta)$, se sua densidade é escrita da seguinte forma

$$f_X(x; \theta) = \exp\{\eta(\theta)T(x) - A(\theta)\}h(x)\mathbb{1}_{\mathcal{X}}(x),$$

em que $\mathcal{X}$ não depende de θ e $\eta(\theta)$ e $T(x)$ são funções contínuas não-triviais. Mostre que as seguintes variáveis aleatórias pertencem à família exponencial, identificando o parâmetro θ e as funções $T(x)$, $A(\theta)$ e $h(x)$:

i) $X \sim \text{B}(p)$, $0 < p < 1$.

ii) $X \sim \text{B}(n, p)$ com $n \in \mathbb{N}$ conhecido e $0 < p < 1$.

iii) $X \sim \text{P}(\lambda)$, $\lambda > 0$.

iv) $X \sim \text{G}(p)$, $0 < p < 1$.

v) $X \sim \text{BN}(r, p)$ com $n \in \mathbb{N}$ conhecido e $0 < p < 1$.

vi) $X \sim \text{Logarítmica}(p)$ com $0 < p < 1$, i.e., com respectiva função de probabilidade

$$\mathbb{P}(X = x) = \frac{p^x}{-\log(1-p)x}\mathbb{1}_{\{1,2,\ldots\}}(x).$$

vii) $X \sim \Gamma(r, \lambda)$ com $r, \lambda > 0$, r conhecido.

viii) $X \sim \mathcal{N}(\mu, 1)$ com $\mu \in \mathbb{R}$ e $\sigma > 0$ conhecido.

ix) $X \sim \mathcal{N}(\mu, \sigma^2)$ com $\mu \in \mathbb{R}$ conhecido e $\sigma > 0$.

x) $X \sim \beta(a, b)$ com um dos parâmetros conhecido.

xi) $X \sim \text{Weibull}(b, a)$, com $a > 0$ e $b > 0$ conhecido, i.e., com respectiva densidade

$$f_X(x) = abx^{b-1}e^{-ax^b}\mathbb{1}_{\mathbb{R}^+}(x).$$

xii) X possuindo distribuição secante hiperbólica generalizada de parâmetro $\theta > 0$, i.e., com respectiva densidade

$$f_X(x) = \frac{1}{2}\exp(\theta x + \ln\cos(\theta))\cosh\left(\frac{\pi x}{2}\right)\mathbb{1}_{\mathbb{R}}(y).$$

52. Mostre que se $X \sim \text{FE}(\theta)$, então $\mathbb{E}[T(X)] = \frac{\partial A(\eta(\theta))}{\partial \eta(\theta)}$ e $\text{Var}[T(X)] = \frac{\partial^2 A(\eta(\theta))}{\partial^2 \eta(\theta)}$.
Sugestão: Usando o fato que $f_X(x;\theta)$ é uma legítima função densidade de probabilidade, i.e.:

$$\int_{\mathcal{A}} \exp\{\eta(\theta)T(x) + c(x)\}d_x = \exp\{A(\theta)\} = \exp\{A(\eta(\theta))\},$$

e use essa identidade para mostrar que a função geradora de momentos (sob condições de regularidade) de $Y = T(X)$ é dada por

$$M_Y(t) = \exp\{A(\eta(\theta) + t) - A(\eta(\theta))\}.$$

53. Os dados abaixo referem-se a resistência à flexão (MPa) de 27 vigas de um tipo de concreto de alto desempenho obtidos pela utilização de superplásticos e determinados adesivos.

Tabela 1: Resistência à flexão (em MPa) de 27 vigas.

5,9	7,2	7,3	6,3	8,1	6,8	7,0
7,6	6,8	6,5	7,0	6,3	7,9	9,0
8,2	8,7	7,8	9,7	7,4	7,7	9,7
7,8	7,7	11,6	11,3	11,8	10,7	

i) Defina a variável de interesse (X).

ii) Calcule uma estimativa pontual do valor médio (μ) da resistência para a população conceitual de todas as vigas fabricadas dessa forma e diga qual estimador você utilizou.

iii) Calcule uma estimativa pontual do valor da resistência que separa as 50% mais fracas de todas as vigas 50% mais fortes ($\widetilde{\mu}$) e diga qual estimador você utilizou.

iv) Calcule e interprete uma estimativa pontual do desvio-padrão da população σ. (Sugestão: $\sum_{i=1}^{27} x_i^2 = 1860,94$.)

v) Calcule uma estimativa pontual da proporção de todas as vigas cuja resistência à flexão exceda a 10 MPa.

vi) Calcule uma estimativa pontual do coeficiente de variação σ/μ da população e diga qual estimador você utilizou.

54. Os pesos das peças produzidas por uma máquina (produção de 5.000 peças/dia) seguem distribuição normal com uma média de 22g e desvio padrão de 1,25g. Foi coletada 50 amostras, de 16 peças cada uma.

i) Determine a média e o desvio padrão da distribuição das médias amostrais.

ii) Em quantas amostras pode-se esperar que a média se encontre entre 19,3 e 20,5g? e abaixo de 19g?

iii) Qual a probabilidade de encontrarmos uma peça escolhida dessa produção com dimensão entre 19,3g e 20,5g?

55. Considere X uma variável aleatória com distribuição normal, com média 100 e desvio-padrão 10.

i) Qual a $\mathbb{P}(90 < X < 110)$?

ii) Se $\overline{X}$ representar a média de uma amostra aleatória de 16 elementos retirados dessa população, calcule $\mathbb{P}(90 < \overline{X} < 110)$.

iii) Denotando a variância amostral por S_X^2, calcule $\mathbb{E}[S_X^2]$.

iv) Represente, num único gráfico, as distribuições de X e $\bar{X}$.

v) Que tamanho deveria ter a amostra para que $\mathbb{P}(90 < \overline{X} < 110) = 0,95$?

56. Encontre a distribuição amostral da diferença de médias nas seguintes situações:

i) Populações normais com variâncias desconhecidas iguais e amostras independentes.

ii) Populações normais com variâncias desconhecidas diferentes e amostras independentes.

iii) Populações normais com variâncias desconhecidas diferentes e amostras dependentes.

57. Considere dois estimadores ($\hat{\theta}_1$ e $\hat{\theta}_2$) para um determinado parâmetro populacional θ. Para ajudar a escolher o *melhor*, simulou-se uma situação em que $\theta = 100$. Dessa população retiraram-se 1.000 amostras de dez unidades cada uma, e obtemos ambas as estimativas usando às dez unidades de cada amostra. Desse modo obtêm-se 1.000 estimativas baseadas em $\hat{\theta}_1$ e outras 1.000 estimativas baseadas em $\hat{\theta}_2$, cujos estudos descritivos estão resumidos abaixo. Qual dos dois estimadores você acha mais conveniente para estimar θ. Por quê?

	$\hat{\theta}_1$	$\hat{\theta}_2$
Média	102	100
Variância	5	10
Mediana	100	100
Moda	98	100

58. Um pesquisador está em dúvida sobre dois possíveis estimadores $\hat{\theta}_1$ e $\hat{\theta}_2$, para um parâmetro θ. Assim, ele decidiu usar simulação para uma situação hipotética, procurando encontrar pistas que o ajudassem a decidir qual o melhor estimador. Partindo de uma população fictícia, em que $\theta = 10$, ele retirou 1000 amostras de 20 elementos, e para cada amostra calculou o valor das duas estimativas. Em seguida, construiu a distribuição de frequências, segundo o quadro abaixo. Com base nesses resultados, escolha (justificando o motivo) qual dos dois estimadores deve ser preferível.

Classes	% de $\hat{\theta}_1$	% de $\hat{\theta}_2$
$[5, 7)$	10	5
$[7, 9)$	20	30
$[9, 11)$	40	35
$[11, 13)$	20	25
$[13, 15)$	10	5

59. Sejam $X_1, \ldots, X_n$ uma amostra aleatória de tamanho n da distribuição de uma variável aleatória $X \sim \mathcal{U}(0, \theta), \theta > 0$. Considere os estimadores $\hat{\theta}_1 = c_1\overline{X}_n$ e $\hat{\theta}_2 = c_2X_{(n)}$, em que $X_{(n)} := \max\{X_1, \cdots, X_n\}$.

i) Determine o espaço paramétrico referente ao modelo estatístico em questão.

ii) Encontre c_1 e c_2 que tornam os estimadores não viciados.

iii) Considere a seguinte amostra aleatória de tamanho 20: `3.59,3.13,3.30,0.30,1.54,0.40, 3.94,1.33,3.36,0.18,3.29,1.50,0.32,1.03,0.95,0.37,3.65,3.44,1.05,2.60`. Baseado nos estimadores não viciados obtidos no item ii), obtenha as estimativas pontuais e dos seus respectivos erros-padrão.

60. Para estimar a média μ de uma população, foram propostos dois estimadores não-viesados independentes $\hat{\mu}_1$ e $\hat{\mu}_2$, de tal sorte que $\text{Var}[\hat{\mu}_1] = \text{Var}[\hat{\mu}_2]/3$. Considere os seguintes estimadores ponderados de μ:

a) $T_1 = \frac{\hat{\mu}_1+\hat{\mu}_2}{2}$.

b) $T_2 = \frac{4\hat{\mu}_1+\hat{\mu}_2}{5}$.

c) $T_3 = \hat{\mu}_1$.

i) Quais estimadores são não-viesados?

ii) Dispor esses estimadores em ordem de eficiência.

iii) Considerando a consistência de que $\hat{\mu}_1$ e $\hat{\mu}_2$ são estimadores consistentes, mostre que os 3 estimadores acima também são.

61. Considere $X \sim \text{P}(\theta)$ e $\phi(\theta) = \exp(-3\theta)$. Mostre que $T = (-2)^X$ é um estimador não viciado para $\phi(\theta)$. Ele é um **bom** estimador?

62. Para estimar a média μ de uma população, foram propostos dois estimadores não-viesados $\hat{\mu}_{1n}$ e $\hat{\mu}_{2n}$, de tal sorte que $\text{Var}[\hat{\mu}_{1n}] = \text{Var}[\hat{\mu}_{2n}]/2$ e $\text{Cov}(\hat{\mu}_{1n}, \hat{\mu}_{2n}) = (2n)^{-1}$. Considerando que $\hat{\mu}_{2n}$ é um estimador consistente de μ.

i) Mostre que o estimador $\hat{\mu}_{1n}$ também é consistente para μ.

ii) Mostre que o estimador $T_n(\alpha) = \alpha\hat{\mu}_{1n} + (1-\alpha)\hat{\mu}_{2n}$, $\alpha \in (0, 1)$, é consistente para μ.

iii) Determine o $\text{EQM}(T_n(\alpha))$ e desenhe o seu gráfico em função de α e encontre o valor de α que minimiza $\text{EQM}(T_n(\alpha))$.

63. Seja $X_1, \dots, X_n$ uma a.a. de uma variável aleatória a.c. que possui densidade simétrica em torno de μ. É possível mostrar que a distribuição assintótica da mediana amostral $\tilde{X}_n$ é tal que

$$\tilde{X}_n \sim \mathcal{AN}\left(\mu, \frac{1}{4n\{f(\mu)\}^2}\right). \tag{1}$$

i) Compare $\mathrm{Var}(\overline{X}_n)$ e $\mathrm{Var}(\tilde{X}_n)$ quando a distribuição subjacente for a normal. Você esperava este resultado? Comente.

ii) Quando a distribuição subjacente for a Cauchy, a média amostral será um **bom** estimador? Quanto vale a variância assintótica de $\tilde{X}_n$?

64. Através de um estudo de simulação, verifique a validade, variando o tamanho da amostra, do resultado (1) para as distribuições: Normal, Cauchy, t-Student (variando os g.l.) e Laplace.

65. Considere $X_1, \dots, X_n \overset{\text{iid}}{\sim} X$. Obtenha as distribuições assintóticas da média e mediana amostrais, quando:

i) $X \sim \mathcal{N}(\mu, \sigma^2)$.

ii) $X \sim t_k(\mu, \sigma)$, com $k > 2$ conhecido.

iii) $X \sim \mathrm{Laplace}(\mu, \sigma)$.

iv) $X \sim \mathrm{Logística}(\mu, \sigma)$.

Para cada um dos itens acima, discuta qual estimador deve ser preferível para estimar o parâmetro de localização.

66. Sejam $X_1, \dots, X_n \overset{\text{iid}}{\sim} \mathrm{P}(\theta)$. Sabemos que $\overline{X}_n$ e S_n^2 são dois estimadores não viciados para θ. Encontre um outro estimador não viciado para θ que seja função de $\overline{X}_n$ e S_n^2 **simultaneamente**.

67*. Sejam $(X_1, Y_1) \dots, (X_n, Y_n)$ uma a.a. de uma distribuição normal bivariada com parâmetros $\mu_1, \mu_2, \sigma_1^2, \sigma_2^2$ e ρ. Assuma que $\mu_1 = \mu_2 = \mu$ e considere a seguinte classe de estimadores

$$\hat{\mu}(\alpha) := \alpha\overline{X}_n + (1-\alpha)\overline{Y}_n.$$

i) O estimador $\hat{\mu}(\alpha)$ é consistente?

ii) Determine o valor $\alpha = \alpha_0$ que minimiza a variância do estimador $\hat{\mu}(\alpha)$ e considere o estimador

$$\hat{\mu}(\alpha_0) := \alpha_0\overline{X}_n + (1-\alpha_0)\overline{Y}_n.$$

Mostre que se $\sigma_1 = \sigma_2$, o BLUE (melhor estimador linear não viciado) de μ é dado por

$$\hat{\mu} = \frac{\overline{X}_n + \overline{Y}_n}{2}.$$

68. Sejam $X_1, \ldots, X_n \overset{\text{iid}}{\sim} (\mu, \sigma^2)$. Mostre que o desvio-padrão amostral é um estimador **viciado** de σ. Note que é preciso **apenas** mostrar que $\mathbb{E}[S_n] \neq \sigma$, não importando o valor **exato**.

69. Considere $X_1, \ldots, X_n$ uma a.a. da distribuição exponencial truncada à esquerda de θ

$$f_X(x) = e^{-(x-\theta)} \mathbb{1}_{[\theta,\infty)}(x).$$

Mostre que $X_{(1)} = \min\{X_1, \ldots, X_n\}$ é um estimador consistente de θ.

70. Considere a questão anterior. Determine um estimador não viciado de θ que seja função de $X_{(1)}$.

71. Sejam $X_1, \ldots, X_n$ uma amostra aleatória de tamanho n da distribuição $\mathcal{N}(0, \sigma^2)$. Seja $S_{n*}^2 := \sum_{i=1}^{n} X_i^2$. Considere os estimadores da forma:

$$\hat{\sigma}_c^2 := cS_{n*}^2.$$

i) Determine o espaço paramétrico referente ao modelo estatístico em questão.

ii) Mostre que $\frac{S_{n*}^2}{\sigma^2} \sim \chi_n^2$.

iii) Encontre o EQM$(\hat{\sigma}_c^2)$.

iv) Encontre o valor de c que minimiza o EQM em iii).

72. Sejam $X_1, \ldots, X_n$ uma amostra aleatória de tamanho n da distribuição $\mathcal{N}(\mu, \sigma^2)$. Seja $S_n^2 := \sum_{i=1}^{n} (X_i - \bar{X}_n)^2$, em que $\bar{X}_n = n^{-1} \sum_{i=1}^{n} X_i$. Considere os estimadores da forma:

$$\hat{\sigma}_c^2 := cS_n^2.$$

i) Determine o espaço paramétrico referente ao modelo estatístico em questão.

ii) Encontre o EQM$(\hat{\sigma}_c^2)$.

iii) Encontre o valor de c que minimiza o EQM em ii).

73*. Sejam $Y_1,\ldots,Y_n$ variáveis aleatórias independentes, tais que $Y_i \sim \mathcal{N}(\alpha+\beta x_i,\sigma^2)$, em que os valores de x_i não são estocásticos, satisfazendo $\sum_{i=1}^{n}(x_i-\bar{x}_n)^2>0$.

i) Mostre que

$$\hat{\beta}=\frac{\sum_{i=1}^{n}(x_i-\bar{x}_n)Y_i}{\sum_{i=1}^{n}(x_i-\bar{x}_n)^2} \quad \text{e} \quad \hat{\alpha}=\overline{Y}_n-\hat{\beta}\bar{x}_n.$$

são estimadores não viciados de β e α, respectivamente.

ii) Mostre que $\hat{\alpha}$ pode ser reescrito como uma combinação linear das variáveis Y_i's, i.e.

$$\hat{\alpha}=\sum_{i=1}^{n}c_iY_i.$$

iii) Sob a suposição de normalidade, encontre a distribuição conjunta dos estimadores.

iv) Mostre que $\hat{\beta}(\hat{\alpha})$ é o BLUE de $\beta(\alpha)$.

v) Forneça uma condição **suficiente** para que os estimadores $\hat{\alpha}$ e $\beta(\alpha)$ sejam consistentes.

74. Considere $X_1,\ldots,X_n\overset{\text{iid}}{\sim}(\mu,\sigma^2)$. Mostre que

$$\hat{\mu}:=\frac{2}{n(n+1)}\sum_{i=1}^{n}iX_i,$$

é um estimador não viciado e fracamente consistente de μ.

75. Seja $X_1,\ldots,X_n\overset{\text{iid}}{\sim}X$, em que

$$f_X(x)=e^{-(x-\theta)}\mathbb{1}_{(\theta,\infty)}(x).$$

Considere a classe dos estimadores

$$T_b(X_{(1)})=X_{(1)}+b, b\in\mathbb{R}.$$

Mostre que o estimador que possui o menor EQM nesta classe é $T^*=X_{(1)}-1/n$.

76. Considere que $x_1,\ldots,x_n$ são valores fixos e que as observações de interesse $y_1,\ldots,y_n$ seguem o seguinte modelo de regressão: $y_i=\beta_0+\beta_1x_i+e_i, i=1,\ldots,n$ com $e_1,\ldots,e_n$ representando uma amostra aleatória da distribuição $\mathcal{N}(0,\sigma^2)$. Determine o espaço paramétrico e vetor de estatísticas suficientes para os parâmetros do modelo, quando:

i) σ é conhecido.

ii) σ é desconhecido.

77. Seja $X_1, \ldots, X_n \overset{\text{iid}}{\sim} X$, com X representando uma variável aleatória com média μ e variância σ^2.

i) Mostre que $\mathbb{E}\left[\sum_{i=1}^{n} a_i X_i\right] = \mu$, para quaisquer conjunto de constantes $\{a_1, \ldots, a_n\}$, satisfazendo $\sum_{i=1}^{n} a_i = 1$.

ii) Se $\sum_{i=1}^{n} a_i = 1$, mostre que $\text{Var}\left[\sum_{i=1}^{n} a_i X_i\right]$ é minimizada quando $a_i = 1/n$, $i = 1, \ldots, n$, i.e., $\bar{X}_n$ é o ENVVUM de μ.

Sugestão: Mostre que $\sum_{i=1}^{n} a_i^2 = \sum_{i=1}^{n} (a_i - 1/n)^2 + 1/n$ quando $\sum_{i=1}^{n} a_i = 1$.

78. Sejam $X_1, \ldots, X_n$ e $Y_1, \ldots, Y_m$ variáveis aleatórias independentes com $X_i \overset{\text{iid}}{\sim} \mathcal{N}(\theta, \sigma^2)$ e $Y_j \overset{\text{iid}}{\sim} \mathcal{N}(\lambda, \tau^2)$, $\theta, \lambda \in \mathbb{R}$, $\sigma, \tau > 0$. Encontre estatísticas suficientes minimais para os casos abaixo:

i) Todos os parâmetros desconhecidos.

ii) $\theta = \lambda$ e σ, τ são desconhecidos.

iii) $\sigma = \tau$ e θ, λ são desconhecidos.

iv) $\theta = \lambda$ e $\sigma = \tau$ são desconhecidos.

Em quais das situações acima a estatística encontrada também é **completa**?

79. Admita que $(X_1, Y_1), \ldots (X_n, Y_n)$ são vetores aleatórios iid com distribuição normal bivariada com $\mathbb{E}[X_1] = \mathbb{E}[Y_1] = 0$, $\text{Var}[X_1] = \text{Var}[Y_1] = 1$ e $\text{Cov}(X_1, Y_1) = \rho$.

i) Encontre uma estatística suficiente minimal para o modelo.

ii) Mostre que as estatísticas $T_1 = \overline{X}_n$ e $T_2 = \overline{Y}_n$ são ancilares, mas o vetor $(T_1, T_2)^\top$ não é.

80. Considere $X_1, \ldots, X_n \overset{\text{aa}}{\sim} \mathcal{U}(0, \theta)$, $\theta > 0$. Considere $X_{(n)} = \max\{X_1, \ldots, X_n\}$.

i) Dentre todos os estimadores da forma $aX_{(n)}$, com a representando uma constante que pode depender de n, determine o estimador que possui o menor EQM.

ii) Encontre o ENVVUM de θ. Nesse caso, o ENVVUM de θ pode ser eficiente? Justifique.

iii) Qual dos dois estimadores deve ser preferível? Justifique.

81. Considere $X_1, \ldots, X_n \overset{\text{aa}}{\sim} X$, tal que $f_X(x) = \frac{2x}{\theta^2}\mathbb{1}_{(0,\theta)}(x)$, $\Theta = \mathbb{R}^+$.

i) Encontre uma estatística suficiente e completa para θ.

ii) Existe alguma função de θ para o qual existe um estimador não viciado cuja variância coincide com o LICR(θ)?

iii) Determine LICR(θ).

Sugestão: Verifique se as condições de regularidade são válidas para esse modelo.

iv) Determine o ENVVUM de θ.

82. Considere $X_1, \ldots, X_n \overset{\text{aa}}{\sim} X$, com X representando uma variável aleatória qualquer. Considere que o interesse é estimar a função de distribuição de X em um ponto $x \in \mathbb{R}$, i.e., estimar $F_X(x)$. Mostre que

i) A função distribuição empírica, $\hat{F}_n(x) := \frac{1}{n}\sum_{i=1}^{n} \mathbb{1}(X_i \leq x)$ é tal que $\mathbb{E}[\hat{F}_n(x)] = F_X(x)$, $\forall x \in \mathbb{R}$.

ii) Assuma que a estatística $T(\mathbf{x}) = (X_{(1)}, \ldots, X_{(n)})$ é suficiente e completa para o modelo em questão. Prove que $\hat{F}_n(x)$ é o ENVVUM de $F_X(x)$, $\overline{X}_n$ é o ENVVUM de $\mu = \mathbb{E}[X]$ e S_X^2 é o ENVVUM de $\sigma^2 = \text{Var}[X]$.

83. Sejam $(X_1, Y_1), \ldots, (X_n, Y_n)$ uma amostra aleatória das distribuições consideradas a seguir. Encontre estatísticas suficientes não triviais para os parâmetros indicados.

i) $(X, Y) \sim$ Trinomial(n, p_1, p_2), i.e.,

$$\mathbb{P}(X = x, Y = y) = \binom{n}{x, y} p_1^x p_2^y (1 - p_1 - p_2)^{n-x-y} \mathbb{1}_{\{0,\ldots,n\}}(x) \mathbb{1}_{\{0,\ldots,n-x\}}(y),$$

$n \in \mathbb{N}$ conhecido, $0 < p_1, p_2 < 1$ e $0 < p_1 + p_2 < 1$, em que $\binom{n}{x,y} = \frac{n!}{x!y!(n-x-y)!}$.

ii) $f(x, y/\boldsymbol{\theta}) = \frac{\beta^{\alpha+\gamma}}{\Gamma(\alpha)\Gamma(\gamma)} x^{\alpha-1}(y - x)^{\gamma-1} e^{-\beta y} \mathbb{1}(0 < x < y)$, $\boldsymbol{\theta} = (\alpha, \beta, \gamma)$.

iii) $(X,Y) \sim \text{Dirichlet}(\alpha_1, \alpha_2, \alpha_3)$, i.e.,

$$f(x, y/\alpha_1, \alpha_2, \alpha_3) = \frac{\Gamma\left(\sum_{i=1}^{3} \alpha_i\right)}{\prod_{i=1}^{3} \Gamma(\alpha_i)} x^{\alpha_1 - 1} y^{\alpha_2 - 1} (1 - x - y)^{\alpha_3 - 1} \mathbb{1}_{(0,1)}(x) \mathbb{1}_{(0,1-x)}(y),$$

$\alpha_i > 0$, $i = 1, 2, 3$.

iv) $(X,Y) \sim \mathcal{N}_2$, i.e.,

$$\begin{aligned} f(x, y/\boldsymbol{\theta}) &= \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \exp\left\{ -\frac{1}{2(1-\rho^2)} \left[\left(\frac{x-\mu_1}{\sigma_1}\right)^2 - 2\rho \left(\frac{x-\mu_1}{\sigma_1}\right) \right.\right. \\ &\times \left.\left. \left(\frac{y-\mu_2}{\sigma_2}\right) + \left(\frac{y-\mu_2}{\sigma_2}\right)^2 \right] \right\} \mathbb{1}_{\mathbb{R}^2}(x, y), \end{aligned}$$

em que $\boldsymbol{\theta} = (\mu_1, \mu_2, \sigma_1, \sigma_2, \rho)^\top$.

84. Considere que $x_1, \ldots, x_n$ são valores fixos e que as observações de interesse $y_1, \ldots, y_n$ seguem o seguinte modelo de regressão: $y_i = \beta_0 + \beta_1 x_i + e_i, i = 1, \ldots, n$ com $e_1, \ldots, e_n$ representando uma amostra aleatória da distribuição $\mathcal{N}(0, \sigma^2)$. Determine o espaço paramétrico e vetor de estatísticas suficientes minimais e completas para os parâmetros do modelo, quando:

i) σ é conhecido.

ii) σ é desconhecido.

85. Faça um resumo sobre o método dos momentos (MM) e o método de máxima verossimilhança (MV), evidenciando vantagens e desvantagens, além das principais propriedades. Por qual razão o EMV é o mais utilizado?

86. Considere $X_1, \ldots, X_n \overset{\text{aa}}{\sim} \mathcal{N}(0, \theta)$, $\theta \in \mathbb{R}^{++}$.

i) Calcule $I_F(\theta)$.

ii) Mostre que o EMV de θ é eficiente.

87. (Prova seleção IME-USP, 2016) Considere X uma v.a. a.c. com função densidade de probabilidade dada por

$$f_X(x|\theta) = \frac{\theta}{x^2} \mathbb{1}_{[\theta, \infty)}(x), \quad \theta > 0,$$

e assuma que $X_1, \ldots, X_n \overset{\text{iid}}{\sim} X$.

i) Determine uma estatística suficiente minimal para θ.

ii) Determine o EMM e EMV de θ.

iii) Qual dos dois estimadores deve ser preferido? Justifique.

iv) Existe alguma função de θ para o qual existe um estimador não viciado cuja variância coincide com o LICR ? Justifique sua resposta.

88. Considere $X_1, \ldots, X_n$ uma a.a. da distribuição $\mathcal{N}(\mu, 1)$, $\mu \geq 0$. Encontre o EMV de μ.

89. Considere $X_1, \ldots, X_n$ uma a.a. da distribuição $\mathcal{N}(\mu, \sigma^2)$.

i) Mostre que os parâmetros são ortogonais. O que isso facilita no processo de estimação?

ii) Determine a distribuição assintótica conjunta de $(\hat{\mu}, \hat{\sigma}^2)^\top$, com $\hat{\theta}$ representando o EMV de θ.

90. Considere $Y_1, \ldots, Y_n$ independentes, tais que $Y_i \sim \mathcal{N}(\beta x_i, \sigma^2)$, x_i conhecido.

i) Determine os EMV de β e σ^2.

ii) Os parâmetros são ortogonais? Você esperaria esse resultado? Justifique.

ii) Determine a distribuição assintótica conjunta de $(\hat{\beta}, \hat{\sigma}^2)^\top$.

91*. Admita que $(X_1, Y_1), \ldots (X_n, Y_n)$ são vetores aleatórios iid com distribuição normal bivariada com $\mathbb{E}[X_1] = \mathbb{E}[Y_1] = 0$, $\mathrm{Var}[X_1] = \mathrm{Var}[Y_1] = 1$ e $\mathrm{Cov}(X_1, Y_1) = \rho$.

i) Determine o EMM e o EMV de ρ.

ii) Repita o item i) considerando uma amostra de vetores aleatórios iid com distribuição normal bivariada de vetor de média $\boldsymbol{\mu}$ e matriz de variância covariância $\boldsymbol{\Sigma}$.

92. Defina:

i) Estatística, estimador, estimativa, estimação e parâmetro.

ii) Faça uma comparação entre intervalo de confiança e teste de hipóteses, abordando os objetivos de cada técnica. Qual a grande vantagem de se fazer um teste de hipóteses com relação a construção de um IC?

93. Faça um breve resumo sobre:

i) O método da quantidade pivotal para obtenção de intervalos de confiança.

ii) A obtenção de intervalos de comprimento mínimo.

94. Considere $X_1, \ldots, X_n$ uma a.a. de uma densidade $f(\cdot)$. Mostre que se $f(\cdot)$ pertence a família de:

i) Localização, então $T = \overline{X} - \theta$ é uma quantidade pivotal.

ii) Escala, então $T = \overline{X}/\theta$ é uma quantidade pivotal.

iii) Localização (θ_1) e escala (θ_2), então $T = (\overline{X} - \theta_1)/\theta_2$ é uma quantidade pivotal.

95. Sejam $X_1, \ldots, X_n$ uma amostra aleatória da distribuição exponencial de média θ, $\theta > 0$.

i) Usando o fato de que $Y = 2\theta^{-1} \sum_{i=1}^{n} X_i \sim \chi^2_{(2n)}$, construa um intervalo de confiança para θ com coeficiente de confiança $1 - \alpha$ $(0 < \alpha < 1)$ baseado na quantidade pivotal Y.

ii) Mostre que

$$\left(\frac{\sqrt{n}\bar{X}_n}{\sqrt{n} + z_{1-\alpha/2}}, \frac{\sqrt{n}\bar{X}_n}{\sqrt{n} - z_{1-\alpha/2}} \right),$$

é um intervalo de confiança assintótico para θ com coeficiente de confiança aproximadamente igual a $1 - \alpha$ $(0 < \alpha < 1)$. Aqui, $z_{1-\alpha/2}$ representa o quantil de ordem $1 - \alpha/2$ da normal padrão, i.e., $z_{1-\alpha/2} = \Phi^{-1}(1 - \alpha/2)$.

96. Os dados abaixo referem-se as notas de uma amostra de 45 estudantes da *aprazível* disciplina de Inferência II no ano de 2017.

i) Faça uma análise descritiva dos dados.

ii) Construa um intervalo de confiança de 95% para a nota mediana da turma.

Tabela 2: Notas de Introdução à Estatística.

5,70	5,60	3,00	2,40	1,90	2,60	3,20	1,30	3,00	6,70
2,40	1,20	5,90	9,00	3,30	3,40	2,90	1,70	5,60	2,70
5,50	8,70	1,90	0,80	1,50	1,80	5,80	1,50	3,50	3,30
1,90	5,80	2,20	2,90	0,80	4,30	1,00	1,70	2,60	2,80
2,70	1,00	5,50	4,60	4,80					

iii) Obtenha e desenhe o gráfico da função distribuição empírica das notas.

iv) Construa um intervalo de confiança para a probabilidade do aluno não ser reprovado direto, *i.e.*, tenha nota maior ou igual a 4.

v) Através do gráfico de quantis-quantis, verifique se é razoável supor normalidade (use a média e a variância amostral, como os verdadeiros parâmetros da distribuição normal).

vi) Para responder os itens ii) e iv) você fez alguma suposição? Caso tenha feito, é razoável considerá-la verdadeira para o problema em questão? Discuta.

97. Considere $X_1, \ldots, X_n \overset{\text{iid}}{\sim} \mathcal{N}(\mu, \sigma^2)$, com $\sigma > 0$ conhecido. Encontre o **melhor** IC de nível $(1-\alpha)$ usando o método da quantidade pivotal.

98. Repita o exercício anterior na situação em que σ é **desconhecido**.

99. $X_1, \ldots, X_n \overset{\text{iid}}{\sim} \mathcal{N}(\mu, \sigma^2)$, com $\sigma > 0$ desconhecido. Encontre um IC para σ^2 de nível $(1-\alpha)$ usando o método da quantidade pivotal quando:

i) μ é conhecido.

ii) μ é desconhecido.

100. Sejam $X_{i_1}, \ldots, X_{i_{n_i}}$ observações independentes de distribuições normais de médias μ_i e variância σ_i^2, para $i = 1, 2$, $\mu_1, \mu_2 \in \mathbb{R}$, $\sigma_1, \sigma_2 > 0$. Obtenha o intervalo de confiança de comprimento mínimo para $\mu_1 - \mu_2$ com coeficiente de confiança confiança γ $(0 < \gamma < 1)$ supondo que as variâncias são conhecidas.

101. Na questão anterior, considerando que as variâncias são desconhecidas, obtenha um intervalo de confiança para σ_1^2/σ_2^2 com coeficiente de confiança confiança γ $(0 < \gamma < 1)$.

102. Repita o Exercício 100 supondo agora que as variâncias são desconhecidas mas iguais, i.e. $\sigma_1^2 = \sigma_2^2 = \sigma^2$.

103. Sejam $X_1, \ldots, X_n$ observações independentes de distribuições exponenciais de médias $\alpha\beta^i$,$i = 1, \ldots, n$, respectivamente, $\alpha, \beta > 0$.

i) Mostre que $\frac{X_i}{\alpha\beta^i} \sim \exp(1)$.

ii) Supondo que β é conhecido, construa um intervalo de confiança para α com coeficiente de confiança confiança γ $(0 < \gamma < 1)$.

104. Sejam $X_1, \ldots, X_n$ observações independentes de distribuições uniformes no intervalo $[0, \theta]$, em que $\theta > 0$.

i) Mostre que $\frac{X_{(n)}}{\theta}$ é uma quantidade pivotal.

ii) Mostre que

$$\left[\frac{X_{(n)}}{(1-\alpha/2)^{1/n}}, \frac{X_{(n)}}{(\alpha/2)^{1/n}}\right],$$

é um intervalo de confiança para θ com coeficiente de confiança $1 - \alpha$ $(0 < \alpha < 1)$.

iii) Encontre o IC ótimo para θ de nível $(1 - \alpha)$ usando o método da quantidade pivotal com base na estatística suficiente completa e minimal $X_{(n)}$.

105. Considere X uma única observação da densidade

$$f_X(x; \theta) = \theta x^{\theta-1} \mathbb{1}_{(0,1)}(x), \theta > 0.$$

i) Determine uma quantidade pivotal e com base nela, determine um IC para θ com coeficiente de confiança $1 - \alpha$ $(0 < \alpha < 1)$.

ii) Considere $Y := -1/\ln X$. Mostre que $(Y/2, Y)$ é um intervalo de confiança para θ. Determine o coeficiente de confiança associado a este intervalo.

iii) Encontre o IC ótimo para θ de nível $(1 - \alpha)$ com base em Y.

106. $X_1, \ldots, X_n \overset{\text{iid}}{\sim} \text{Cauchy}(\mu, 1)$. Encontre um IC para μ de nível $(1 - \alpha)$:

i) Baseado na distribuição exata de $\overline{X}_n$. Neste caso, encontre o intervalo de comprimento mínimo **exato**.

ii) Baseado na distribuição assintótica do EMV de μ, encontrando o intervalo de comprimento mínimo **assintótico**.

iii) Em que situação cada intervalo deve ser preferível ao outro? Por quê?

iv) Realize um estudo de simulação para ajudar a responder a pergunta do item anterior.

107. Explique os erros tipo I e tipo II que podem ocorrer nas seguinte situações:

i) Um júri decide condenar ou não o réu.

ii) O juiz marcou penalidade máxima contra o time azul.

108. (**Gráfico do poder do teste**) Define-se como gráfico do poder do teste para um determinado teste de hipóteses, a curva que expressa o comportamento do poder $(1-\beta)$ em função das diversas hipóteses alternativas $\mathcal{H}_1$, fixando-se o nível de significância α. Desenhe o gráfico do poder do teste para $\mathcal{H}_0 : \mu = 10$ versus $\mathcal{H}_1 : \mu \neq 10$, em que $n = 9$, $\sigma^2 = 4$ e $\alpha = 0,05$. Para simplificar, assuma normalidade da população e admita: 7,0; 7,5; 8,0; 8,5; 9,0; 9,5; 10,00; 10,5; 11,0; 11,5; 12,0; 12,5 e 13,0 como possíveis valores de μ.

109. Refaça a questão anterior, considerando diferentes tamanhos de amostras, por exemplo, $n = 20$ e $n = 30$. Desenhe as três curvas em um mesmo gráfico e comente.

110. Considere $X_1, \ldots, X_n$ uma a.a. da distribuição Laplace$(\mu, 1)$. Obtenha o teste da razão de verossimilhanças generalizada para $\mathcal{H}_0 : \mu = 0$ vs $\mathcal{H}_1 : \mu \neq 0$ ao nível α. Estime empiricamente o tamanho e o poder do teste utilizando os valores $\mu = 0, \pm 0.2, \pm 0.4, \pm 0.6, \pm 0.8, \pm 1.2, \pm 1.4, \pm 1.6, \pm 1.8$. Plote a curva da função poder empírica. Considere $n = 5, 10, 15, 20, 30$ e 50 e discuta os resultados.

111. Considere que o interesse é testar a hipótese $\mathcal{H}_0 : \mu = \mu_0$, com σ conhecido e sob suposição de normalidade. Mostre que a probabilidade do erro tipo II para um teste de nível α, para os diferentes tipos de hipóteses alternativas, são dadas por:

Hipótese	**Probabilidade do erro tipo II**
$\mathcal{H}_1 : \mu = \mu' > \mu_0$	$\Phi\left(z_\alpha + \frac{\mu_0 - \mu'}{\sigma/\sqrt{n}}\right)$
$\mathcal{H}_1 : \mu = \mu' < \mu_0$	$1 - \Phi\left(-z_\alpha + \frac{\mu_0 - \mu'}{\sigma/\sqrt{n}}\right)$
$\mathcal{H}_1 : \mu = \mu' \neq \mu_0$	$\Phi\left(z_{\alpha/2} + \frac{\mu_0 - \mu'}{\sigma/\sqrt{n}}\right) - \Phi\left(-z_{\alpha/2} + \frac{\mu_0 - \mu'}{\sigma/\sqrt{n}}\right)$,

em que $\Phi(z) = \mathbb{P}(Z \leq z)$ e $z_\alpha = \Phi^{-1}(\alpha)$ representam, respectivamente, a função distribuição acumulada e o quantil de ordem $\alpha \in (0,1)$ da distribuição $\mathcal{N}(0,1)$. Adicionalmente, mostre que $\lim_{n\to\infty} \beta(\mu') = 0, \forall \mu' \neq \mu_0$. Você consegue explicar este resultado?

112. De forma análoga a questão anterior, encontre a probabilidade do erro tipo II, quando o interesse é testar $\mathcal{H}_0 : \mu = \mu_0$, com σ desconhecido e sob suposição de normalidade. Deixe a resposta em termos da função distribuição acumulada da distribuição t com k graus de liberdade: $\Phi_k(\cdot)$.

113.(Determinação do tamanho da amostra) Em situações práticas é desejável controlar a probabilidade de se cometer ambos os tipos de erros. Fixado α, temos que a probabilidade de se cometer o erro tipo II é uma função decrescente do tamanho da amostra (Questões # 31 e # 32). Baseado no resultado que você obteve na Questão # 33, mostre que o tamanho da amostra para o qual um teste de nível α possui probabilidade tipo II igual a $\beta(\mu')$ (perceba que usamos o valor μ' especificado) é dado por

Hipótese	**Tamanho da amostra**
$\mathcal{H}_1 : \mu = \mu' > \mu_0 (\mu' < \mu_0)$	$n = \left\lfloor \left(\frac{\sigma(z_\alpha + z_\beta)}{\mu_0 - \mu'} \right)^2 \right\rfloor,$

em que $\lfloor x \rfloor$, representa o menor número inteiro maior ou igual a x.

114. Para investigar a influência do tipo de ensino (Particular e Público, referente ao ultimo ano do ensino médio) sobre a média no curso de Introdução à Estatística de recém-ingressos na UFC, obteve-se a seguinte amostra:

Tabela 3: Notas de 20 alunos no curso de Introdução à Estatística.

Particular	Público
2,5	8,0
7,8	4,2
3,5	7,4
8,3	4,0
5,0	4,1
3,9	5,6
6,0	5,5
10,0	6,0
9,1	5,2
2,5	3,3

i) Denotando as notas dos estudantes oriundos de escolas particulares (públicas) por $x_1, \ldots, x_{10}$ $(y_1, \cdots, y_{10})$ obtenha $\bar{x}, \bar{y}, s_x^2$ e s_y^2.

ii) Calcule $\hat{\sigma}^2 = ((n_x - 1)s_x^2 + (n_y - 1)s_y^2)/(n_x + n_y - 2)$, em que n_x e n_y, representam, respectivamente, o número de alunos na amostra oriundos de escolas particulas e públicas.

iii) Recalcule $\bar{x}, \bar{y}, s_x^2, s_y^2$ e σ^2, usando as formas matriciais apresentadas na Questão # 21.

iv) Considerando válida as suposições de Normalidade, independência e igualdade de variâncias, utilize o teste t para verificar se existe evidência a favor da hipótese de que os alunos oriundos de escolas particulares apresentam um melhor desempenho na disciplina.

115. (Uso de simulação para avaliar um teste de hipóteses) As seguintes hipóteses sobre a média são consideradas $\mathcal{H}_0 : \mu = 1$ vs $\mathcal{H}_1 : \mu \neq 1$. Fixado um nível de significância α, o teste irá a rejeitar a hipótese nula α% das vezes, quando esta for verdadeira. Para estimar α empiricamente, utilize o seguinte procedimento de simulação:

i) Gerar 10.000 amostras de tamanho 20 da distribuição $\mathcal{N}(1, 2)$.

ii) Para cada amostra gerada determinar as estatísticas apropriadas para o teste e realizar o teste usando $\alpha = 0.05$. Determine se a hipótese nula é ou não rejeitada.

iii) Determinar a porcentagem de amostras em que a hipótese nula é rejeitada, digamos $\hat{\alpha}$) e comparamos com o verdadeiro valor α.

Você espera que $\hat{\alpha} = 0.05$? Justifique.

Adicionalmente, para avaliar a função poder do teste no ponto $\mu \neq 1$, você pode repetir o processo acima, trocando o item i) por

i') Gerar 10.000 amostras de tamanho 100 da distribuição $\mathcal{N}(\mu', 2)$.

Estime empiricamente o poder do teste para $\mu = 0.2, 0.4, 0.6, 0.8, 1.2, 1.4, 1.6, 1.8$. Comente os resultados.

116. Estime empiricamente o nível de significância do teste de hipóteses $\mathcal{H}_0 : \mu = 1$ vs $\mathcal{H}_1 : \mu \neq 1$, quando você gera valores de uma distribuição χ^2_1 (média 1 e variância 2) e use o teste obtido sob normalidade. Comente os resultados.