

**Exame de Qualificação – Doutorado em Estatística**  
**Parte de Probabilidade**  
**04 de fevereiro de 2025**

**Instruções:**

1. Leia atentamente as questões.
2. A prova é composta de 3 exercícios, que devem ser respondidos de forma clara, completa e detalhada.
3. A duração da prova é de 3 horas.
4. Não é permitido consulta.
5. Inicie cada questão em uma folha separada e coloque o seu nome completo em cada folha.

**Exercício 1:** (Teorema de Egoroff) Seja  $(E, \mathcal{A}, \mu)$  um espaço de probabilidade. Seja  $(f_n)_{n \geq 1}$  uma sequência de funções mensuráveis de  $E$  em  $\mathbb{R}$ . Assuma que  $(f_n)_{n \geq 1}$  converge pontualmente para uma função real  $f$ . O objetivo deste exercício é mostrar que para todo  $\varepsilon > 0$  existe um conjunto  $A \in \mathcal{A}$  tal que  $\mu(A^c) < \varepsilon$  e a convergência de  $(f_n)_{n \geq 1}$  para  $f$  é uniforme no conjunto  $A$ , isto é,  $\sup_{x \in A} |f_n(x) - f(x)| \rightarrow 0$  quando  $n \rightarrow \infty$ .

(a) [8 pts.] Considere para todo  $k, n \geq 1$  o conjunto

$$E_{k,n} = \bigcup_{j \geq n} \left\{ x \in E : |f_j(x) - f(x)| > \frac{1}{k} \right\}.$$

Mostrar que para  $k$  fixo a sequência  $(E_{k,n})_{n \geq 1}$  é decrescente e precisar o conjunto limite.

(b) [11 pts.] Deduzir que para todo  $\varepsilon > 0$  e  $k \geq 1$ , existe um inteiro  $n_k = n_k(\varepsilon)$  tal que  $\mu(E_{k,n_k}) < \frac{\varepsilon}{2^k}$ .

(c) [11 pts.] Seja  $B = \bigcup_{k \geq 1} E_{k,n_k}$ . Mostrar que a convergência de  $(f_n)_{n \geq 1}$  é uniforme em  $B^c$  e deduzir o Teorema de Egoroff.

**Exercício 2:** Sejam  $X$  e  $Y$  variáveis aleatórias independentes e identicamente distribuídas, com distribuição *Bernoulli*( $p$ ),  $p \in (0, 1)$ , ou seja,

$$P(X = 1) = p = 1 - P(X = 0).$$

Definimos  $Z = \mathbf{1}_{\{X+Y \geq 1\}}$  e  $\mathcal{G} = \sigma(Z)$ .

(a) [10 pts.] Descrever os eventos da  $\sigma$ -álgebra  $\mathcal{G}$ .

(b) [12 pts.] Mostrar que

$$E[X | \mathcal{G}] = \frac{Z}{2-p} \text{ q.c.}$$

(c) [8 pts.] Note, por simetria, que  $E[Y | \mathcal{G}] = E[X | \mathcal{G}]$  q.c. Concluir que as variáveis aleatórias  $E[X | \mathcal{G}]$  e  $E[Y | \mathcal{G}]$  não são independentes.

**Exercício 3:** Seja  $\lambda > 0$  um número conhecido. Suponha que  $X = X_\lambda$  é uma variável aleatória com distribuição *Exponencial*( $\lambda$ ) e que  $Y = Y_\lambda$  é uma variável aleatória tal que

$$P(Y = n | X) = \frac{e^{-X} X^n}{n!}, \quad n = 0, 1, 2, \dots$$

Em outras palavras, dado que  $X = x$ , temos que  $Y \sim \text{Poisson}(x)$ . Mostrar que:

(a) [12 pts.] A função de probabilidade marginal de  $Y$  é dada por:

$$P(Y = n) = \lambda \left( \frac{1}{1 + \lambda} \right)^{n+1}, \quad n = 0, 1, 2, \dots$$

(b) [12 pts.] A função característica de  $Y$  é:

$$\phi_{Y_\lambda}(t) = \frac{\lambda}{1 + \lambda - e^{it}}, \quad t \in \mathbb{R}.$$

(c) [8 pts.] A média e a variância de  $Y$  são, respectivamente,

$$E(Y_\lambda) = \frac{1}{\lambda} \quad \text{e} \quad \text{Var}(Y_\lambda) = \frac{1 + \lambda}{\lambda^2}.$$

(d) [8 pts.] A variável aleatória

$$Z_\lambda = \frac{Y_\lambda - E(Y_\lambda)}{\sqrt{\text{Var}(Y_\lambda)}}$$

converge em lei para  $N(0, 1)$  quando  $\lambda \rightarrow \infty$ .

*Dica:* No item (d), defina a função  $\psi_\lambda(t) = \log \phi_{Y_\lambda}(t)$  e considere a expansão de Taylor de  $\psi_\lambda$  em torno do ponto  $t = 0$ .

**Exame de Qualificação – Doutorado em Estatística**  
**Parte II**

**Seis de fevereiro de 2025**

**Importante, ler com cuidado:** Todo tipo de documento, livro etc é proibido. A solução das questões é de caráter individual e, portanto, consultas, conversas, troca de idéias ou materiais constituem violação desse caráter. Para recebimento de crédito total, desenvolva suas respostas. **As soluções têm que ser apresentadas em ordem; por exemplo, 1a → 1b → 1c → 1d. Soluções fora de ordem serão consideradas apenas em sua primeira subsequência ordenada. Por exemplo: 1a → 1b → 1c ⇒ corrigem-se 1a,1b e 1c; e 1c → 1b → 1a ⇒ corrige-se apenas 1c.** Boa Sorte.

**Exercício 1:** Sejam  $a$  e  $b$  números inteiros não-negativos, fixos e conhecidos, sendo  $a + b \geq 1$ . Sejam  $0 < p < 1$  um número real, fixo e conhecido e  $q = 1 - p$ .

- (a) [5 pts.] Seja  $m$  um número inteiro positivo, fixo e conhecido. Defina a distribuição binomial negativa com parâmetros  $p$  e  $m$ , a partir do esquema de amostragem binomial, também conhecido como **amostragem binomial inversa**. Mostre que

$$P(Y = y) = \binom{m + y - 1}{m - 1} p^m q^y, \quad (1)$$

para  $y = 0, 1, \dots$ , em que  $Y + m$  é o número de experimentos de Bernoulli realizados.

- (b) [5 pts.] Considere  $Y$  definida no item (a). Prove que o estimador UMVU para  $p^r$  ( $r < m$ ) é dado por

$$\delta(Y) = \frac{(m - r + Y - 1)(m - r + Y - 2) \cdots (m - r)}{(m + Y - 1)(m + Y - 2) \cdots m}. \quad (2)$$

Para os itens (c)-(d), considere o esquema de amostragem binomial até que pelo menos  $a$  sucessos e  $b$  fracassos sejam observados.

- (c) [5 pts.] Discuta as possíveis configurações determinadas por esse esquema amostral, de acordo com os valores de  $a$  de  $b$ . Identifique as distribuições associadas e suas potenciais aplicações.

(d) [10 pts.] Ache um estimador razoável para  $\log(p/q)$

Dica: (i) Na solução do item (d), considere a fórmula alternativa da binomial negativa, em que se usam as combinações de Newton do tipo  $\binom{-m}{k}$  e obtenha um estimador não-viciado para  $\log(p)$ .

**Exercício 2:** Considere  $X_1, \dots, X_n$  uma amostra i.i.d. da distribuição uniforme no intervalo  $(\theta_1 - \theta_2, \theta_1 + \theta_2)$ , com  $\theta_1 \in \mathbb{R}$ ,  $\theta_2 > 0$ , e  $n > 2$ .

(a) [11 pts.] Encontre os ENVUMV de  $\theta_1$  e  $\theta_2$ .

(b) [11 pts.] Encontre o ENVUMV de  $\theta_1/\theta_2$ .

**Exercício 3:** Seja  $X_1, \dots, X_n$  uma amostra i.i.d. da exponencial truncada de parâmetros  $a \in \mathbb{R}$  e  $\theta > 0$ , isto é, com densidade

$$f(x; a, \theta) = \frac{1}{\theta} \exp\left\{-\frac{x-a}{\theta}\right\} \mathbf{1}_{[a, \infty)}(x). \quad (3)$$

(a) [7 pts.] Encontre um teste UMP de tamanho  $\alpha$  para testar  $H_0 : a = a_0$  contra  $H_1 : a \neq a_0$ , quando  $\theta$  é conhecido.

(b) [7 pts.] Para testar  $H_0 : a = a_0$  contra  $H_1 : a = a_1 < a_0$ , mostre que qualquer teste UMP  $T_*$  de tamanho  $\alpha$  satisfaz  $\beta_{T_*}(a_1) = 1 - (1 - \alpha)e^{-n(a_0 - a_1)/\theta}$ .

(c) [7 pts.] Encontre um teste UMP de tamanho  $\alpha$  para testar  $H_0 : a = a_0$  contra  $H_1 : a \neq a_0$ .

(d) [7 pts.] Encontre um teste UMP de tamanho  $\alpha$  para testar  $H_0 : a = a_0, \theta = \theta_0$  contra  $H_1 : a < a_0, \theta < \theta_0$ .

**Exercício 4:** Considere duas distribuições de probabilidade contínuas  $F_1$  e  $F_2$ . Tomem-se amostras  $x_1, \dots, x_m$  e  $y_1, \dots, y_n$  de  $F_1$  e  $F_2$ , respectivamente. Sejam  $m_1$  e  $n_1$  o número de  $x$ 's e  $y$ 's que excedem a  $k$ -ésima estatística de ordem da amostra combinada  $x_1, \dots, x_m, y_1, \dots, y_n$ .

(a) [5 pts.] Mostre que, se  $F_1 \equiv F_2$ , temos

$$P(m_1, n_1) = \frac{\binom{m}{m_1} \binom{n}{n_1}}{\binom{m+n}{k}}, \quad (4)$$

para  $m_1 + n_1 = m + n - k$ .

- (b) [10 pts.] Use (4) para propor um teste (não-paramétrico) para  $H_0 : F_1 \equiv F_2$ . Apresente o teste completamente e discuta suas principais propriedades.
- (c) [10 pts.] Suponha agora que  $F_1 = N(\mu_1, \sigma_1^2)$  e  $F_2 = N(\mu_2, \sigma_2^2)$ . Apresente testes (paramétricos) para  $F_1 = F_2$  para as diferentes configurações de  $(\mu_1, \sigma_1^2, \mu_2, \sigma_2^2)$ . Discuta, em termos gerais (e particulares, quando as especificidades forem relevantes), as vantagens e desvantagens dos testes paramétricos frente ao teste não-paramétrico proposto em (b).

# Exame de Qualificação – Doutorado em Estatística

## Parte I

22 de julho de 2025

### Instruções:

1. Leia atentamente as questões.
2. A prova é composta de 3 exercícios, que devem ser respondidos de forma clara, completa e detalhada.
3. A duração da prova é de 3 horas.
4. Não é permitido consulta.
5. Inicie cada questão em uma folha separada e coloque o seu nome completo em cada folha.

**Exercício 1:** Sejam  $(X_n)_{n \geq 1}$  uma sequência de variáveis aleatórias não negativas sobre um espaço de probabilidade  $(\Omega, \mathcal{F}, P)$  e  $(\mathcal{F}_n)_{n \geq 1}$  uma sequência de sub- $\sigma$ -álgebras de  $\mathcal{F}$ . Supomos que  $E[X_n | \mathcal{F}_n]$  converge em probabilidade para 0. O objetivo deste exercício é mostrar que  $X_n$  converge em probabilidade para 0.

**Por contradição,** assumamos que para algum  $\varepsilon > 0$ ,  $P[X_n > \varepsilon] \geq \varepsilon$  para uma infinidade de  $n$ 's. A partir de agora, consideramos somente estes  $n$ 's. Seja  $A_n := \{E[X_n | \mathcal{F}_n] > \frac{\varepsilon^2}{3}\}$ .

(a) [10 pts.] Mostrar que para  $n$  suficientemente grande,

$$P[\{X_n > \varepsilon\} \setminus A_n] \geq \frac{\varepsilon}{2}.$$

(b) [10 pts.] Mostrar que para  $n$  suficientemente grande,

$$E[X_n \mathbf{1}_{A_n^c}] \leq \frac{\varepsilon^2}{3}.$$

(c) [15 pts.] Usando o item (a), mostrar que para  $n$  suficientemente grande,

$$E[X_n \mathbf{1}_{A_n}] \geq \frac{\varepsilon^2}{2},$$

e concluir o objetivo.

*Dica:* No item (b), use a Propriedade Característica da Esperança Condicional.

**Exercício 2:** Sejam  $X$  e  $Y$  variáveis aleatórias independentes e identicamente distribuídas, com distribuição uniforme no intervalo  $[-1/2, 1/2]$ .

(a) [10 pts.] Calcular a função característica de  $X$  e daí obter a função característica de  $X + Y$ .

(b) [10 pts.] Seja  $Z$  uma variável aleatória com densidade de Lebesgue em  $\mathbb{R}$  dada por:

$$f(z) = \begin{cases} 1 - |z| & \text{se } -1 \leq z \leq 1, \\ 0 & \text{caso contrário.} \end{cases}$$

Calcular a função característica de  $Z$ .

(c) [10 pts.] Deduzir a lei de  $X + Y$ .

**Exercício 3:** Sejam  $(X_n)_{n \geq 1}$  uma sequência de variáveis aleatórias e  $X$  uma variável aleatória definidas no mesmo espaço de probabilidade. Queremos mostrar a seguinte implicação: se  $X_n \geq 0$ ,  $E(X_n) \rightarrow E(X) < \infty$  e  $X_n \rightarrow X$  q.c., então  $X_n \rightarrow X$  em  $L^1$ .

Para isto, propomos o seguinte roteiro:

(a) [10 pts.] Mostrar que

$$(X - X_n)_+ \leq X, \text{ q.c.}$$

onde  $(X - X_n)_+$  é a parte positiva de  $X - X_n$ .

(b) [10 pts.] Deduzir que  $E[(X - X_n)_+] \rightarrow 0$  quando  $n \rightarrow \infty$ .

(c) [15 pts.] Concluir que  $E[|X - X_n|] \rightarrow 0$  quando  $n \rightarrow \infty$ .

**Exame de Qualificação – Doutorado em Estatística**  
**24 de julho de 2025**

**Importante, ler com cuidado:** Todo tipo de documento, livro etc. é proibido. A solução das questões é de caráter individual e, portanto, consultas, conversas, troca de ideias ou materiais constituem violação desse caráter. Para recebimento de crédito total, desenvolva suas respostas. **As soluções têm que ser apresentadas em ordem; por exemplo, 4a → 4b.**

Boa Sorte.

**Exercício 1:** Seja  $(X_1, Y_1), (X_2, Y_2), \dots, (X_n, Y_n)$  uma amostra i.i.d. da distribuição uniforme no disco de raio  $\theta > 0$  centrado em  $(a, b) \in \mathbb{R}^2$ , isto é, com densidade

$$f(x, y) = \begin{cases} \frac{1}{2\pi\theta^2} & \text{se } (x - a)^2 + (y - b)^2 \leq \theta^2, \\ 0 & \text{caso contrário,} \end{cases}$$

para  $(x, y) \in \mathbb{R}^2$ .

- (a) (15 pontos) Se  $a = 0, b = 0$ , encontre uma estatística  $T$  completa e suficiente para  $\theta$ , e sua distribuição.
- (b) (10 pontos) No caso geral com  $a, b$ , e  $\theta$  desconhecidos, mostre que o invólucro convexo (*convex hull*) dos pontos  $(X_1, Y_1), (X_2, Y_2), \dots, (X_n, Y_n)$ , denotado por  $H(\mathbf{X}, \mathbf{Y})$ , é uma estatística suficiente para  $\theta$ . Dica: mostre primeiro que os elementos  $(x_{i^*}, y_{i^*}), i^* = 1, \dots, m \leq n$ , que maximizam  $\|(x_i, y_i) - (a, b)\|$ , para  $i = 1, 2, \dots, n$ , estão em  $H(\mathbf{x}, \mathbf{y})$ .

**Exercício 2:** Seja  $Y_1, \dots, Y_n$  i.i.d. da  $U(0, \theta)$ , com  $\theta \in (1, \infty)$ . Suponha que observamos apenas  $X_1, \dots, X_n$  onde

$$X_i = \begin{cases} Y_i & \text{se } Y_i \geq 1, \\ 1 & \text{se } Y_i < 1, \end{cases}$$

para  $i = 1, 2, \dots, n$ .

- (a) (5 pontos) Observe que a medida de probabilidade  $\mathbb{P}_X$  é absolutamente contínua com respeito a  $m + \delta_1$ , onde  $m$  é a medida de Lebesgue e  $\delta_1$  é uma massa pontual em 1. Encontre a densidade de  $X_1$  com respeito a  $\mathbb{P}_X$ . Escreva a função de verossimilhança de  $\theta$ .
- (b) (5 pontos) Mostre que  $X_{(n)}$  é suficiente para  $\theta$ .
- (c) (15 pontos) Encontre o teste UMP de tamanho  $\alpha \in (0, 1)$  para  $H_0 : \theta \leq \theta_0$  contra  $H_1 : \theta > \theta_0$ , onde  $\theta_0$  é conhecido e  $\theta_0 > (1 - \alpha)^{-1/n}$ .

**Exercício 3:**

- (a) (5 pontos) Sejam  $X_1, \dots, X_n$  *i.i.d.* com alguma distribuição contínua  $F$ . Prove que o vetor de estatísticas de ordem,  $(X_{(1)}, \dots, X_{(n)})$  é suficiente para  $F$ , em que  $X_{(1)} \leq \dots \leq X_{(n)}$  é uma permutação ordenada da amostra de tamanho  $n$ .
- (b) (10 pontos) Sejam  $X_1, \dots, X_n$  *i.i.d.* com distribuição  $F$  tal que  $E|X|^\alpha < +\infty$  para algum  $\alpha > 0$ . Suponha que:

$$r \leq \alpha \min(k, n - k + 1). \quad (1)$$

Prove que (1) implica em:

$$E|X_{(k)}|^r < +\infty. \quad (2)$$

- (c) (10 pontos) Suponha a distribuição Cauchy, ou seja,  $X \sim C(a, b)$ :

$$f(x) = \frac{b}{\pi} \frac{1}{b^2 + (x - a)^2}.$$

Estude o erro quadrático médio (EQM) da média truncada como estimador do parâmetro de posição em função da proporção de truncamento. Ao final, faça um gráfico e comente.

**Exercício 4:** Considere uma amostra *i.i.d.*,  $\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_n$ , Pareto com parâmetros  $\theta > 0$  e  $\nu > 0$ , cuja densidade é dada por

$$f(x|\theta, \nu) = \frac{\theta \nu^\theta}{x^{\theta+1}} \quad \text{para } x > \nu.$$

- (a) (10 pontos) Mostre que o TRV para  $H_0 : \theta = 1, \nu$  desconhecido vs  $H_1 : \theta \neq 1, \nu$  desconhecido pode ser escrito como  $\phi(\mathbf{x}) = 1$  para  $\mathbf{x} \in A$ ,  $A = \{\mathbf{x} : T(\mathbf{x}) \leq c_1 \text{ ou } T(\mathbf{x}) \geq c_2\}$ ,  $0 < c_1 < c_2$ , em que

$$T(\mathbf{x}) = \ln \left( \prod_{i=1}^n X_i / (\min_{1 \leq i \leq n} X_i)^n \right).$$

- (b) (15 pontos) Mostre que, sob  $H_0$ ,  $2T$  tem distribuição qui-quadrado (DICA: Trabalhe com a distribuição condicional de  $T$  dado  $\min_{1 \leq i \leq n} X_i$ ).