Exame de Qualificação – Doutorado em Estatística
Parte de Probabilidade
24 de Janeiro de 2023

Instruções:

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

Exercício 1: Seja $\{\mu_n\}_{n\geq 1}$ uma sequência de medidas de probabilidade em um espaço mensurável $(\Omega, \mathcal{F})$ (ou seja, $\mu_n(\Omega) = 1$ para todo $n \geq 1$). Definimos a função de conjuntos μ em $\mathcal{F}$ por

$$\mu(A) = \sum_{n=1}^{\infty} \frac{\mu_n(A)}{2^n},\ A \in \mathcal{F}.$$

(a) [*10 pts.*] Mostrar que μ é uma medida de probabilidade em $\mathcal{F}$.

(b) [*15 pts.*] Provar que para qualquer função mensurável $f : \Omega \to \mathbb{R}$ não negativa ou μ-integrável,

$$\int f\, d\mu = \sum_{n=1}^{\infty} \frac{1}{2^n} \int f\, d\mu_n.$$

Dica: considerar primeiro $f = I_A$, $A \in \mathcal{F}$.

Exercício 2: [*25 pts.*] Seja $\{X_n\}_{n\geq 1}$ uma sequência de variáveis aleatórias em $\mathcal{L}^2(\Omega, \mathcal{F}, P)$, e defina

$$\mu_n = E(X_n) \quad \text{e} \quad \sigma_n^2 = \mathrm{Var}(X_n),\ n \geq 1.$$

Mostrar que se $\lim\limits_{n\to\infty} \mu_n = \mu$ e $\sum_{n=1}^{\infty} \sigma_n^2 < \infty$, então $X_n \xrightarrow{\text{q.c.}} \mu$.

Exercício 3: Sejam $\{X_n\}_{n\geq 1}$ e X, variáveis aleatórias definidas no mesmo espaço de probabilidade.

(a) [*15 pts.*] Mostrar que $X_n \to X$ em probabilidade se e somente se $E\left(\frac{|X_n - X|}{1+|X_n - X|}\right) \to 0$ quando $n \to \infty$. Dica: a função $g(x) = x/(1+x)$ é crescente em $[0, \infty)$.

(b) [*10 pts.*] Supomos que as variáveis aleatórias X_n, $n \geq 1$, têm valores no intervalo $[-M,M]$ onde M é um real positivo. Mostrar que $X_n \to 0$ em probabilidade se e somente se $E(|X_n|) \to 0$ quando $n \to \infty$.

Exercício 4: Sejam X e Y duas variáveis aleatórias independentes e identicamente distribuídas, com variância finita σ^2. Supomos também que $\frac{X+Y}{\sqrt{2}}$ tem mesma lei que X e Y.

(a) [*5 pts.*] Calcular $E(X)$.

(b) [*10 pts.*] Seja $\{X_i, i \geq 1\}$ uma sequência de variáveis aleatórias i.i.d. com mesma lei que X. Para $n \geq 1$, seja

$$Z_n = \frac{1}{\sqrt{2^n}} \sum_{i=1}^{2^n} X_i.$$

Mostrar por indução que, para todo $n \geq 1$, Z_n tem mesma lei que X.

(c) [*10 pts.*] Usando o Teorema Central do Limite, determinar a lei de X.

EXAME DE QUALIFICAÇÃO PARA O DOUTORADO EM ESTATÍSTICA

Parte: Inferência

26 de janeiro de 2023 - 09:00-13:00 hs - Sala 225 (IMECC)

Importante, ler com cuidado:

- Todo tipo de documento, livro, notas, etc.é proibido.
- A solução das questões é de caráter individual e, portanto, consultas, conversas, troca de ideias ou materiais constituem violação desse caráter.
- Escolha 4 das 5 questões apresentadas, indique claramente quais são as questões selecionadas.
- Cada questão tem o mesmo peso (25 pontos de 100 pontos).
- Para recebimento de crédito total, desenvolva suas respostas, justificando os passos.
- **As soluções têm que ser apresentadas em ordem; por exemplo, 4a → 4b.**

Boa Sorte.

Exercício 1: Considere uma amostra aleatória $X_1, X_2, ..., X_n$ da função densidade de probabilidade Normal com média θ e variância 1, $\mathrm{N}(\theta, 1)$, $\theta \in \mathbb{R}$.

(a) [*7 pts.*] Prove que o procedimento dado por $T(X_1, ..., X_n) = \mathcal{I}_{(c_1,c_2)}(\overline{X})$ é Uniformemente Mais Poderoso ao nível α (UMP(α)) para

$$H_0 : |\theta| \geq \theta_0 \quad \text{vs} \quad H_1 : |\theta| < \theta_0,$$

onde c_1 e c_2 são constantes dependendo de α, $\mathcal{I}_A(\cdot)$ é a função indicadora do conjunto A, e quando a regra de decisão toma o valor 1 indica-se a rejeição de H_0, sendo $\overline{X} = \frac{\sum_{i=1}^n X_i}{n}$.

(b) [*4 pts.*] Mostre que no item anterior $c_1 = -c$ e $c_2 = c$ e identifique c em termos do valor de $\alpha \in (0, \frac{1}{2})$.

(c) [*7 pts.*] Assuma uma distribuição à priori em θ dada por uma distribuição Normal com média μ_0 e variância σ_0^2, $\theta \sim \mathrm{N}(\mu_0, \sigma_0^2)$. (i) Identifique a distribuição à posteriori de θ, (ii) identifique o estimador Bayesiano de θ sob perda quadrática, (iii) analise os resultados (i) e (ii) quando $\sigma_0^2 \to \infty$.

(d) [*7 pts.*] Utilize a distribuição à posteriori identificada no item anterior e estabeleça a regra 0-1 de aceitação de H_0.

Exercício 2: Considere uma amostra aleatória $X_1, X_2, ..., X_n$ da função densidade de probabilidade dada pela equação (1)

$$f(x; a) = \frac{1}{\theta} e^{-\frac{(x-a)}{\theta}} \mathcal{I}_{(a,\infty)}(x), \quad a > 0. \tag{1}$$

Sendo θ conhecido e positivo, onde $\mathcal{I}_A(\cdot)$ é a função indicadora do conjunto A.

(a) [*8 pts.*] Identifique uma estatística suficiente e completa para a, baseada na amostra $X_1, X_2, ..., X_n$. Justifique.

(b) [*6 pts.*] Ache um estimador não viciado de variância uniformemente mínima (ENVVUM) para a. Justifique.

(c) [*5 pts.*] Dê um intervalo de confiança $100(1-\alpha)\%$ exato para a baseado no ENVVUM de a.

(d) [*6 pts.*] Determine um ENVVUM para $P(X_1 \geq t), \ \forall t$. Justifique.

Exercício 3: Estimadores de Máxima Verossimilhança

(a) [*12.5 pts.*] Considere uma amostra aleatória $(X_1, Y_1), (X_2, Y_2), ..., (X_n, Y_n)$ gerada pela lei

$$f(x, y; \theta) = \begin{cases} (\theta+1)(1-x)^{\theta}, & \text{se } x > y \\ (\theta+1)(1-y)^{\theta}, & \text{se } x < y \\ \theta(1-x)^{\theta+1}, & \text{se } x = y. \end{cases}$$

Defina T como o total de $(X_i, Y_i), i = 1, ...n$ tais que $X_i = Y_i$ e defina $Z_i = \max\{X_i, Y_i\}, i = 1, ..., n$.

i. (7pts) Apresente a forma da função de verossimilhança em termos de T e $Z_i, i = 1, ..., n$.

ii. (5.5pts) Encontre o estimador de Máxima Verossimilhança para θ. Justifique cada etapa do processo.

(b) [*12.5 pts.*] Considere uma amostra aleatória $X_1, X_2, ..., X_n$ da função densidade de probabilidade Uniforme $U(\theta, \theta + |\theta|)$, sendo $\theta \neq 0$.

i. (7.5pts) Determine o estimador de Máxima Verossimilhança de θ (para $\theta \neq 0$). Justifique

ii. (5pts) Dê um intervalo de confiança $100(1-\alpha)\%$ exato para θ baseado no estimador de Máxima Verossimilhança de θ.

Exercício 4: Seja $(X_1, Y_1), \ldots, (X_n, Y_n)$ vetores aleatorios i.i.d.; assuma que $X_1 \sim \text{Cauchy}(0,1)$ e, condicionado em $X_1 = x$, $Y \sim \text{Cauchy}(\beta x, 1)$, com $\beta \in \mathbb{R}$. Seja $\bar{X}$ e $\bar{Y}$ as médias amostrais baseadas em X_i e Y_i, respectivamente. Use o fato que $Z \sim \text{Cauchy}(z_0, \gamma)$ se Z tem densidade de Lebesgue e função de distribuição acumulada, respectivamente,

$$f(z) = \frac{1}{\pi\gamma} \frac{\gamma^2}{\gamma^2 + (z - z_0)^2}, \text{ e } F(z) = \frac{1}{\pi} \arctan\left(\frac{z - z_0}{\gamma}\right) + \frac{1}{2},$$

ou função característica

$$\varphi_Z(t) = \exp\{z_0 it - \gamma|t|\}.$$

(a) [*5 pts.*] Prove o seguinte fato cautelar para a Cauchy: se $Z \sim \text{Cauchy}(0,1)$ e $W = Z$, ainda assim $\varphi_{Z+W}(t) = \varphi_Z(t)\varphi_W(t)$.

(b) [*10 pts.*] Obtenha as distribuições marginais de $\bar{Y}$ e $\bar{Y} - \beta\bar{X}$. Dica: encontre a distribuição conjunta de $Y_1 - \beta X_1$ e X_1.

(c) [*10 pts.*] Encontre a distribuição da estatística $\bar{Y}/\bar{X}$ e argumente que ela não é estimador consistente de β.

Exercício 5: Desenvolva com cuidado, identificando claramente os passos:

(a) [*9 pts.*] Seja $Y_n = \prod_{i=1}^n X_i$, onde $X_1, X_2, \ldots$ são variáveis aleatorias i.i.d. não-negativas, com $0 < \text{Var}(X_1) < \infty$ e $\mathbb{E}(X_1) = 1$. Prove que $Y_n \xrightarrow{\mathbb{P}} 0$ quando $n \to \infty$.

(b) Seja $\hat{\zeta}$ o estimador de máxima verossimilhança de $\zeta = p(1-p)$ baseado em uma amostra de tamanho n de variáveis aleatórias i.i.d. Bernoulli(p), com $0 < p < 1$. Seja p_0 o valor verdadeiro de p.

- [*7 pts.*] Mostre que quando $p_0 \neq 1/2$ a distribuição assintótica de $\hat{\zeta}$ é normal (com a normalização adequada) e indique seus parâmetros.
- [*9 pts.*] Encontre uma distribuição assintótica de $\hat{\zeta}$ quando $p_0 = 1/2$, com normalização apropriada. Dica: tente reescrever $\sqrt{n}(\hat{\zeta} - \zeta)$ como uma transformação de algo conhecido. A distribuição assintótica não precisa ser normal.

Exame de Qualificação – Doutorado em Estatística
Parte de Probabilidade
9 de agosto de 2023

Instruções:

1. Leia atentamente as questões.

2. A prova é composta de 4 questões, que devem ser respondidas 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 e RA em cada folha.

Exercício 1: [*25 pts.*] Sejam $\{Z_n\}_{n\geq 1}$ uma sequência de variáveis aleatórias independentes e identicamente distribuídas, com distribuição normal padrão, e $\{\alpha_n\}_{n\geq 1}$ uma sequência de números reais. Para $n \geq 1$, definimos

$$Y_n = \sum_{i=1}^{n} \alpha_i Z_i \quad \text{e} \quad W_n = \exp\left\{\sum_{i=1}^{n} Z_i - \frac{n}{2}\right\}.$$

(a) [*12 pts.*] Mostre que se $\sum_{n=1}^{\infty} \alpha_n^2 < \infty$, então $\{Y_n\}$ converge quase certamente e em $\mathcal{L}^2$.

(b) [*13 pts.*] Prove que $\{W_n\}$ converge quase certamente. A sequência $\{W_n\}$ é uniformemente integrável?

Dica: Use que a função geradora de momentos de uma variável aleatória $Z \sim N(0,1)$ é $M_Z(t) = e^{t^2/2}$, $t \in \mathbb{R}$.

Exercício 2: [*25 pts.*]

(a) [*6 pts.*] Considere uma variável aleatória $X \sim \text{Binomial}(n,p)$, $n \geq 1$, $p \in (0,1)$, com função de probabilidade

$$P(X = x) = \binom{n}{x} p^x \, (1-p)^{n-x}, \; x = 0, 1, \ldots, n.$$

Mostre que a função característica de X é dada por:

$$\phi_X(t) = \left[p\, e^{it} + 1 - p\right]^n, \; t \in \mathbb{R}.$$

(b) [*6 pts.*] Considere uma variável aleatória $Y \sim \text{Poisson}(\lambda)$, $\lambda > 0$, com função de probabilidade

$$P(Y = y) = \frac{e^{-\lambda}\, \lambda^y}{y!}, \; y = 0, 1, \ldots$$

Prove que a função característica de Y é dada por:

$$\phi_Y(t) = \exp\left\{\lambda(e^{it} - 1)\right\}, \; t \in \mathbb{R}.$$

(c) [*13 pts.*] Seja $\{X_n\}_{n\geq 1}$ uma sequência de variáveis aleatórias com $X_n \sim \text{Binomial}(n, p_n)$, onde $n \geq 1$ e $p_n \in (0,1)$. Suponha que $\lim_{n\to\infty} n\, p_n = \lambda > 0$, e seja $Y \sim \text{Poisson}(\lambda)$. Demonstre que:

$$X_n \xrightarrow{\mathcal{D}} Y.$$

Exercício 3: [*25 pts.*]

(a) Seja X uma variável aleatória não negativa definida em um espaço de probabilidade $(\Omega, \mathcal{F}, P)$.

(i) [*7 pts.*] Prove que:

$$\sum_{n=1}^{\infty} P(X \geq n) \leq E(X) \leq 1 + \sum_{n=1}^{\infty} P(X \geq n).$$

(ii) [*6 pts.*] Mostre que para qualquer $\alpha > 0$, temos a seguinte equivalência:

$$E(X) < \infty \iff \sum_{n=1}^{\infty} P(X \geq \alpha n) < \infty.$$

(b) Seja $\{X_n\}_{n\geq 1}$ uma sequência de variáveis aleatórias não negativas, independentes e identicamente distribuídas. Deduza a seguinte dicotomia:

(i) [*6 pts.*] Se $E(X_1) < \infty$, então $\limsup_{n\to\infty} \frac{X_n}{n} = 0$, q.c.

(ii) [*6 pts.*] Se $E(X_1) = \infty$, então $\limsup_{n\to\infty} \frac{X_n}{n} = \infty$, q.c.

Dica: Use o Lema de Borel–Cantelli.

Exercício 4: [*25 pts.*] Sejam X, Y e Z variáveis aleatórias em um espaço de probabilidade $(\Omega, \mathcal{F}, P)$. Suponha que X e Z são independentes, $X \in \mathcal{L}^1$, $Z \in \mathcal{L}^1$ e $Y = h(Z)$, com $h : \mathbb{R} \to \mathbb{R}$ uma função Borel mensurável. Prove que:

$$E(XZ \,|\, Y) = E(X)\, E(Z \,|\, Y) \text{ q.c.}$$

Dica: Use que $\{Y \in B\} = \{Z \in h^{-1}(B)\}$ para $B \in \mathcal{B}(\mathbb{R})$.

Exame de Qualificação – Doutorado em Estatística
Parte II
11 de agosto de 2023

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.

Importante: você deve escolher quatro questões, dentre as cinco, para entregar. Somente quatro questões serão corrigidas. Todas as questões têm peso igual, correspondendo a 2.5 pontos.

Boa Sorte.

Exercício 1: Seja $X_1,...,X_n$, $n > 1$, uma amostra aleatória (variáveis i.i.d) proveniente da distribuição cuja densidade é dada pela equação (1)

$$f(x) = \theta a^{\theta} x^{-(\theta+1)} \mathbf{1}_{[a,\infty)}(x), \tag{1}$$

sendo θ conhecido: $\theta > 2$, onde $\mathbf{1}_A(\cdot)$ é a função indicadora do conjunto A.

(a) (0.625 pontos) Postule uma estatística suficiente (diferente de $(X_1,...,X_n)$) e completa para o parâmetro a. Justifique a sua escolha.

(b) (0.625 pontos) Determine a distribuição da estatística proposta no item anterior.

(c) (0.625 pontos) Identifique (caso exista) um estimador não viciado de variância uniformemente mínima (ENVUMV) para a. Justifique a sua proposta.

(d) (0.625 pontos) Ache uma quantidade pivotal baseada na estatística identificada no primeiro item e determine o intervalo de confiança exato para a, considerando uma confiança igual a $1 - \alpha$.

Exercício 2: Suponha que $n \geq 2$, $0 \leq \theta \leq 1$, e seja $X_1, \ldots, X_n$ variáveis aleatórias independentes e identicamente distribuídas, com função massa de probabilidade

$$\mathbb{P}_\theta(X_1 = 0) = 1 - \theta, \quad \mathbb{P}_\theta(X_1 = j) = \frac{\theta}{5}, \qquad \text{para } j = 1, \ldots, 5.$$

(a) (1.25 pontos) Desenvolva um contraexemplo para mostrar que $T = \sum_{i=1}^{n} X_i$ não é uma estatística suficiente para θ. Encontre uma estatística suficiente não-trivial (isto é, diferente de $\mathbf{T} = (X_1, \ldots, X_n)$) para θ.

(b) (1.25 pontos) Existe um ENVUMV para θ? Se sim, encontre-o e mostre que é ENVUMV. Caso contrário, explique o porquê.

Exercício 3: Seja X_1,...,X_n, $n > 1$, uma amostra aleatória (variáveis i.i.d) proveniente da distribuição cuja densidade é dada pela equação (2)

$$f(x) = \frac{1}{2K}(1 + \theta x)\mathbf{1}_{[-K,K])}(x), \; K > 0 \; (K \text{ fixo}). \tag{2}$$

Sendo $\theta : \theta \in [-\frac{1}{K}, \frac{1}{K}]$, onde $\mathbf{1}_A(\cdot)$ é a função indicadora do conjunto A.

(a) (0.625 pontos) Ache o estimador de θ pelo método dos momentos, $\hat{\theta}_M$. Dica: use o primeiro momento.

(b) (0.625 pontos) Determine a variância de $\hat{\theta}_M$ e verifique se $\hat{\theta}_M$ é consistente para estimar θ. Apresente claramente a noção de consistência que irá considerar.

(c) (1.25 pontos) Desenvolva o teste $H_0 : \theta = -\frac{1}{K}$ vs $H_1 : \theta = \frac{1}{K}$. Para $n = 1$,

- **(i)** [*0.625 pts.*] identifique a região de rejeição do teste Mais Poderoso ao nível α (MP(α)) para testar H_0 vs H_1;

- **(ii)** [*0.625 pts.*] identifique explicitamente o quantil envolvido na região anterior em função de α.

Exercício 4: Seja $X \geq 0$ variável aleatória com $\mu = \mathbb{E}(X)$ e seja X' uma v.a. independente e com a mesma distribuição de X. Definimos o coeficiente de Gini γ da distribuição de X por

$$\gamma = \frac{1}{2}\frac{\mathbb{E}(|X - X'|)}{\mu}.$$

Note que $0 \leq \gamma \leq 1$. Suponha que queremos investigar o coeficiente de Gini γ quando $X \sim$ Pareto(θ), isto é, com densidade dada por

$$f_\theta(x) = \frac{\theta}{x^{\theta+1}}\mathbf{1}\{x > 1\}.$$

Assuma que $\theta > 1$, embora a densidade exista para $\theta > 0$.

(a) (0.5 pontos) Mostre que para $X \sim$ Pareto(θ), temos $\gamma = (2\theta - 1)^{-1}$. A seguinte identidade pode ser útil: se $a > 0$ e $b > 0$, $|a - b| = a + b - 2\min\{a,b\}$.

(b) (1 ponto) Se $X_1, \ldots, X_n$ é uma amostra i.i.d. de $X \sim$ Pareto(θ), encontre $\hat{\gamma}$ pelo método de máxima verossimilhança.

(c) (1 ponto) Mostre que a variância assintótica de $\hat{\gamma}$ é $\frac{\gamma^2(1+\gamma)^2}{n}$. Dica: pode usar sem mostrar que nas condições do exercício $\mathbb{E}(\log(X_1)) = \dfrac{1}{\theta}$.

Exercício 5: Seja $X_1, \ldots, X_n$ i.i.d. com densidade

$$f_\theta(x) = \frac{1}{2}(1 - \theta^2)e^{\theta x - |x|},$$

onde $\theta \in (-1,1)$. Considere as hipóteses $H_0 : \theta \leq 0$ contra $H_1 : \theta > 0$.

(a) (1.25 pontos) Mostre que existe teste UMP para H_0 de tamanho $\alpha \in (0, 1/2)$ quando $n = 1$ da forma $Y(X) > c$, e encontre Y e c explicitamente.

(b) (1.25 pontos) Generalize o teste do item (a) para $n > 1$ e indique como obter $\{c_n\}$ tal que $\lim_{n\to\infty} \mathbb{P}_{\theta=0}(Y(X_1, \ldots, X_n) > c_n) = \alpha$.