CC0288 - Inferência Estatística I

Lista Especial 1 - 26/04/2023.

Prof. Maurício

Vamos fazer uma lista com as questões de Inferência que caíram na prova de Seleção do Mestrado do IME-USP

1. (Novembro de 2015) Seja X variável aleatória com função densidade de probabilidade
$$f(x|\theta) = \frac{2x}{\theta^2}, \quad 0 \leq x \leq \theta,$$
em que $\theta > 0$ é um parâmetro desconhecido. Seja $(X_1, ..., X_n)$ uma amostra aleatória de X.
   (a) Encontre o estimador do método dos momentos de θ.
   (b) Encontre o estimador de máxima verossimilhança de θ.
   (c) Algum dos dois estimadores obtidos nos itens acima pode fornecer estimativas não plausíveis para θ?
   (d) Calcule os vícios dos dois estimadores. São não-viciados?

2. (Fevereiro 2016) Seja $(X_1, ..., X_n)$ uma amostra aleatória de uma distribuição Normal de média zero e variância $\sigma^2 > 0$ desconhecida.
   (a) Mostre que a distribuição de $(X_1, ..., X_n)$ faz parte da família exponencial unidimensional e mostre que a $\sum_{i=1}^n X_i^2$ é uma estatística suficiente para σ^2.
   (b) Construa um intervalo de confiança para σ^2 com coeficiente de confiança γ $(0 < \gamma < 1)$ que dependa dos dados apenas através da estatística suficiente $\sum_{i=1}^n X_i^2$.

3. (Fevereiro 2016) Seja X variável aleatória com função densidade de probabilidade
$$f(x|\theta) = \frac{\theta}{x^2}, \quad \theta \leq x < \infty,$$
em que $\theta > 0$ é um parâmetro desconhecido. Seja $(X_1, ..., X_n)$ uma amostra aleatória de X.
   (a) Encontre o estimador de máxima verossimilhança para θ.
   (b) Obtenha o viés do estimador de máxima verossimilhança. Este estimador é não-viciado?

4. (Novembro 2016) Seja $X = (X_1, \ldots, X_{n_1})$ uma amostra aleatória de tamanho n_1 da variável aleatória $X \sim N(\mu, \sigma^2)$ e seja $Y = (Y_1, \ldots, Y_{n_2})$ uma amostra aleatória de tamanho n_2 da variável aleatória $Y \sim N(\mu, \frac{\sigma^2}{4})$ sendo X e Y independentes.
   Se o interesse é estimar μ (admitindo σ^2 conhecido) responda:
   (a) Qual deve ser a relação entre os tamanhos das amostras n_1 e n_2 para que os estimadores
$$\bar{X}_{n_1} = \frac{\sum_{i=1}^{n_1} X_i}{n_1} \quad e \quad \bar{Y}_{n_2} = \frac{\sum_{i=1}^{n_2} Y_i}{n_2} \quad ,$$
   tenham a mesma variância?

(b) Sendo $n_1 = n_2 = n$, obtenha o estimador de máxima verossimilhança de μ baseado na amostra completa, com $2n$ observações.

(c) Sendo $n_1 = n_2 = n$, compare o estimador de máxima verossimilhança ($\hat{\mu}_1$) e o estimador

$$\hat{\mu}_2 = \frac{\bar{Y}_n + \bar{X}_n}{2},$$

sob o ponto de vista de viés e erro quadrático médio. Qual dos dois estimadores é mais indicado? Justifique.

5. (Novembro 2016) Seja $(X_1, ..., X_n)$ uma amostra aleatória de X com função densidade de probabilidade

$$f(x\ \theta) = \frac{1}{\theta}\, x^{-(1/\theta+1)}, \quad x > 1, \ \ \theta > 0.$$

(a) Mostre que a distribuição de X pertence a uma família exponencial unidimensional e mostre que $\sum_{i=1}^{n} log\ X_i$ é uma estatística suficiente para θ.

(b) Mostre que a

$$\frac{\sum_{i=1}^{n} log\ X_i}{\theta},$$

é uma quantidade pivotal e utilize esta quantidade pivotal para construir um intervalo de confiança para θ com coeficiente de confiança γ,$(0 < \gamma < 1)$.

6. (Fevereiro 2017)

(a) Dê a definição de **estatística suficiente**. Interprete do ponto de vista de inferência estatística. Justifique bem sua resposta.

(b) Considere o problema de se fazer inferência sobre p, a probabilidade desconhecida de ocorrência de cara de uma moeda. Um experimento é realizado da seguinte forma: a moeda é lançada até o aparecimento da primeira cara e conta-se o número de coroas obtidas; repete-se o procedimento independentemente n vezes. Seja X_i o número de coroas observadas na i-ésima repetição. A partir da **definição de estatística suficiente**, mostre que $\sum_{i=1}^{n} X_i$ é suficiente.

7. (Fevereiro 2017) Seja $\mathbf{X} = (X_1, ..., X_n)$ uma amostra aleatória de tamanho n da variável aleatória X que tem distribuição de Rayleigh com função densidade de probabilidade

$$f(x\ ;\sigma^2) = \frac{x}{\sigma^2}\, exp\{-\frac{x^2}{2\sigma^2}\}, \quad x > 0,$$

em que $\sigma^2 > 0$ é desconhecido.
(Obs. $E(X) = \sigma\sqrt{\pi/2}$, $Var(X) = \sigma^2(4-\pi)/2$, e $Var(X^2) = 4\sigma^4$.)

(a) Encontre o estimador de máxima verossimilhança e o estimador de método dos momentos (baseado na média de X) de σ^2 e verifique se são não viciados.

(b) Usando aproximação normal para o estimador de máxima verossimilhança, obtenha um intervalo de confiança para σ^2 com coeficiente de confiança aproximado de 95%.

8. (Novembro 2017) Suponha que n componentes eletrônicos serão colocados em teste e seja T_i o tempo de vida do componente i, para $i = 1, 2 \ldots, n$, Admita que $T1_1, T_2 \ldots, T_n$ sejam independentes e que T_i tenha uma distribuição exponencial de média $c_i\ \lambda$, em que $\lambda > 0$ é desconhecido e $c_i > 0$, para $i = 1, 2, \ldots, n$, são números

   fixados (conhecidos)

   (a) Mostre que a distribuição de $(T1, T_2, \ldots, T_n$ faz parte da família exponencial unidimensional.

   (b) A estatística

   $$\sum_{i=1}^{n} \frac{T_I}{c_i}$$

   é suficiente? Justifique.

   (c) Encontre o estimador de máxima verossimilhança de λ e mostre que é não viciado.

   OBS. Dizemos que uma variável aleatória X tem distribuição exponencial de média $\theta > 0$, se sua função densidade de probabilidade é da forma

   $f(x;\theta) = (1/\theta)\ \exp(-x/\theta)$, para $x > 0$.

9. (Novembro 2017) A distribuição de uma variável aleatória X, que depende do parâmetro θ é dada pela tabela

x	-1	0	2
$P(X = x)$	θ	$2\theta - 0,2$	$1,2 - 3\theta$

   temos a amostra $x_1 = 0, x_2 = 2$.

   (a) Qual é o espaço paramétrico mais amplo possível para esse problema?
   (b) Encontre estimativa $\hat{\theta}^{ML}$ pelo método de máxima verossimilhança.
   (c) Encontre estimativa $\hat{\theta}^{MM}$ pelo método dos momentos.

10. (Fevereiro 2018) Sejam X, Y e Z variáveis aleatórias independentes com distribuição de Bernoulli, com parâmetros θ_1, θ_2 e $\dfrac{\theta_1+\theta_2}{2}$. Para estimar $\gamma = \dfrac{\theta_1+\theta_2}{2}$ dois procedimentos foram propostos:

(i) selecionar uma amostra aleatória $X_1, \ldots, X_n$ de tamanho n de X e calcular $\bar{X} = \dfrac{1}{n} \displaystyle\sum_{i=1}^{n} X_i$.

Selecionar uma amostra aleatória $Y_1, \ldots, Y_n$ de tamanho n de Y e calcular $\bar{Y} = \dfrac{1}{n} \displaystyle\sum_{i=1}^{n} Y_i$.

Usar $\dfrac{\bar{X} + \bar{Y}}{2}$ para estimar γ.

(ii) selecionar uma amostra aleatória $Z_1, \ldots, Z_{2n}$ de tamanho $2n$ de Z e calcular $\bar{Z} = \dfrac{1}{2n} \displaystyle\sum_{i=1}^{2n} Z_i$. Usar $\bar{Z}$ para estimar γ.

a. Verifique que os estimadores propostos em (i) e (ii) são não viciados.

b. Baseando-se no erro quadrático médio, determine qual dos dois estimadores é mais indicado para estimar γ. Por quê?

11. (Fevereiro 2018) Seja $X_1, \ldots, X_n$ uma amostra da variável aleatória X com função densidade de probabilidade

$$f(x; \theta) = \exp(-(x - \theta)), \; \theta < x < \infty; \; \theta > 0.$$

a Defina **quantidade pivotal**. Verifique se $Q = X_1 - \theta$ é uma quantidade pivotal, sendo $X_1 = \text{mínimo}(X_1, \ldots, X_n)$.

b Utilize a quantidade pivotal Q, e mostre que qualquer intervalo da forma $(X_1 - b, X_1 - a)$ com $0 < a < b$ satisfazendo $\exp(-na) - \exp(-nb) = 1 - \alpha$ é um intervalo de confiança para θ com coeficiente de confiança $1 - \alpha, \;\; 0 < \alpha < 1$.

c Use (b) para mostrar que

$$\left(X_1 + \frac{\log(\alpha/2)}{n} \;\; , \;\; X_1 + \frac{\log(1 - \alpha/2)}{n} \right),$$

é um intervalo de confiança para θ com coeficiente de confiança $1 - \alpha, \;\; 0 < \alpha < 1/2$.

12. (Novembro 2018) Sejam n variáveis aleatórias $Y_1, \ldots, Y_n$ independentes tais que Y_i tem distribuição Normal com média βx_i, em que $\beta \in (-\infty, \infty)$ e $x_i, \; i = 1, 2, \ldots, n$, são valores conhecidos e não aleatórios., e variância conhecida $\sigma^2 = 1$ para $i = 1, 2, \ldots, n$.

a) Obtenha o estimador de máxima verossimilhança para β. Calcule seu viés e erro quadrático médio.

b) Apresente condições para as quais o estimador seja consistente. Construa o intervalo de confiança de 95% para o parâmetro β.

13. (Novembro 2018) Sejam $X_1, \ldots, X_n$ variáveis aleatórias independentes e identicamente distribuídas tais que

$$X_1 \sim Poisson(\theta), \;\; \theta > 0.$$

Defina

$$S = \mathbb{I}_{\{0\}} \left(X_1 \right) \quad e \quad T = \sum_{i=1}^{n} \; X_i.$$

a Verifique se T é uma estatística suficiente para o modelo estatístico em questão.(utilize a definição de estatística suficiente)

b Encontre $T_1 = \mathbb{E}_\theta(S|T)$ e verifique de T_1 é um estimador eficiente.

c Mostre que $Var_\theta(T) \leq Var_\theta(S)$ e calcule

$$\lim_{n\to\infty} n\, Var_\theta(T_1).$$

14. (Novembro 2019) Seja $X_1, X_2, \ldots, X_n$ amostra aleatória de X cuja função densidade de probabilidade é $f_\theta(x)$, em que θ é um parâmetro desconhecido.

    $T(X_1, X_2, \ldots, X_n)$ é uma função da amostra $X_1, X_2, \ldots, X_n$ que será usada para estimar θ.

    Defina em termos matemáticos o que significa cada termo abaixo:

    a) $T(X_1, X_2, \ldots, X_n)$ é um estimador não viesado de θ.
    b) $T(X_1, X_2, \ldots, X_n)$ é um estimador consistente de θ.
    c) $T(X_1, X_2, \ldots, X_n)$ é uma estatística suficiente para o modelo estatístico em questão.

15. (Novembro 2019) Seja $(X_1, X_2, \ldots, X_n)$ uma amostra aleatória de X cuja função de probabilidade é dada por

$$P_\theta(X = x) = \frac{(x+1)\,\theta^2}{(\theta+1)^{x+2}}\ I_N\ (x), \theta > 0,$$

    em que θ é um parâmetro desconhecido do modelo estatístico.

    (a) Mostre que $P_\theta(X = x)$ é de fato uma função de probabilidade para cada $\theta > 0$.
    (b) Obtenha $T : \mathbb{N}^n \to \mathbb{N}$ suficiente para o modelo estatístico acima.
    (c) Encontre o estimador de máxima verossimilhança para $g(\theta) = P_\theta(X = 0)$.

16. (Fevereiro 2020) Uma caixa contém 10 bolas, das quais θ são brancas e $10 - \theta$ são verdes, $\theta \in \{0, 1, 2, \ldots, 10\}$.

    Duas bolas são extraídas, uma a uma, sem reposição, da urna.

    Seja $X_i = 1$ se a i-ésima bola retirada da urna é branca e $X_i = 0$ se verde , $i = 1, 2$.

    (a) Obtenha o estimador de máxima verossimilhança para θ.

    (b) Verifique se o estimador obtido em (a) é não viciado para θ.

17. (Fevereiro 2020) Sejam X_1, X_2 e X_3 variáveis aleatórias tais que

$$X_1 \sim Poi(\theta),\ \theta > 0,\ \ X_2|X_1 = x_1 \sim Poi(\theta(1+x_1))\ \ e\ \ \ X_3|X_1 = x_1, X_2 = x_2 \sim Poi(\theta(1+x_2)).$$

(a) Mostre que $\sum_{i=1}^{3} X_i$ não é suficiente para θ.

(b) Exiba $T : \mathbb{N}^3 \to \mathbb{N}^2$ suficiente para θ.

(c) Obtenha o estimador de máxima verossimilhança para θ.