Universidade Federal de Lavras

Departamento de Estatística

Prof. Daniel Furtado Ferreira

6ª Lista de Exercícios — Teoria da Estimação pontual e intervalar

1) Marcar como verdadeira ou falsa as seguintes afirmativas:

a) ( ) O método dos momentos consiste em igualar os momentos amostrais como a média, a variância, o momento amostral de assimetria, etc. com os valores esperados correspondentes, como a média (momento de ordem 1 não centrado), a variância (momento de ordem 2 centrado na média) e momento de assimetria (momento de ordem 3 centrado na média), etc.

b) ( ) Parâmetro é uma característica populacional, sendo uma constante, como por exemplo a média μ, a variância σ^2, entre outras possibilidades.

c) ( ) Estimador é uma expressão que permite estimar algum parâmetro e não possui erro associado ao seus resultados.

d) ( ) Estimador é uma função dos elementos amostrais e, portanto, é uma variável aleatória, como a média amostral $\bar{X}$ e a variância amostral S^2, entre outros casos.

e) ( ) A estimativa de máxima verossimilhança é o valor de θ que minimiza o valor de $L(\theta; x_1, x_2, \cdots, x_n)$.

f) ( ) Estimativa é valor que representa o parâmetro e é constante para todas as amostras aleatórias extraídas de uma mesma população.

g) ( ) Estimativas são constantes de uma amostra específica e representa o resultado de um estimador particular. As estimativas de amostras para amostras variam.

h) ( ) O viés ou vício de um estimador é a propriedade em que quando o valor médio deste estimador (esperança matemática) em todas as amostras possíveis de um determinado tamanho n de uma população (na distribuição de amostragem) é igual ao parâmetro, dizemos que o estimador é não viesado ou não viciado.

i) ( ) São exemplos de estimadores não viesados, a média amostral $\bar{X}$ e a variância amostral S^2 e de estimador viesado o desvio padrão amostral S.

j) ( ) Um estimador mais eficiente é aquele dentre os estimadores não viesados de um mesmo parâmetro que possui a menor variância.

k) ( ) Um estimador eficiente é aquele que sua variância vai para zero, quando n tende para infinito.

l) ( ) A mediana amostral é menos eficiente que média amostral para estimar a média μ em populações normais, pois possui maior variância.

m) ( ) Um estimador é consistente, se sua variância vai para ∞, quando n tende a zero.

n) ( ) Estimador viciado ou viesado é aquele que sua distribuição apresenta média igual ao parâmetro que ele está estimando.

o) ( ) Um estimador é consistente se sua variância tende a zero, na medida que o tamanho amostral n cresce, tendendo para ∞.

p) ( ) São métodos de estimação pontual os métodos de momentos, da máxima verossimilhança e dos quadrados mínimos.

q) ( ) O método dos momentos é obtido tomando-se o produto das funções densidades de cada observação amostral X_j, f_{X_j} e maximizando esta função resultante deste produto em relação aos parâmetros.

r) ( ) Assim, a função de verossimilhança, L, é dada por

$$L(\theta; x_1, x_2, \cdots, x_n) = f(x_1;\theta)f(x_2;\theta) \times \cdots \times f(x_n;\theta) = \prod_{i=1}^{n} f(x_i;\theta).$$

s) ( ) A estimativa de máxima verossimilhança é o valor de θ que maximiza o valor de $L(\theta; x_1, x_2, \cdots, x_n)$.

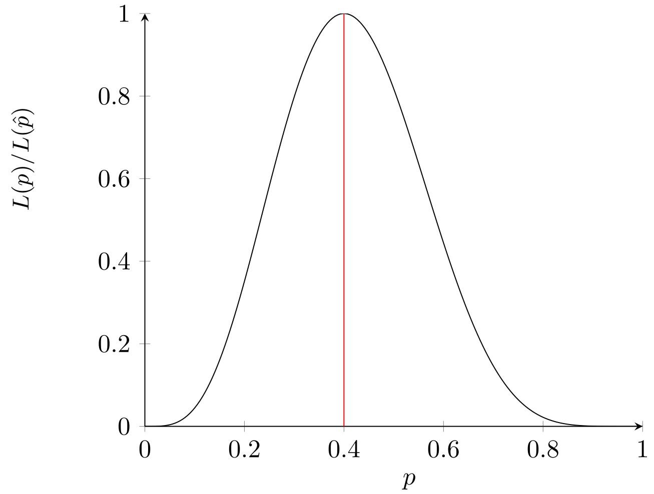
t) ( ) O método dos quadrados mínimos, para estimar a média em um modelo $Y_i = \mu + \epsilon_i$, corresponde a minimização de $Q = \sum_{i=1}^{n} \epsilon_i^2 = \sum_{i=1}^{n}(X_i - \mu)^2$, em que estimador resultante é $\hat{\mu} = \bar{X}$.

u) ( ) O método dos quadrados mínimos, para estimar a média em um modelo $Y_i = \mu + \epsilon_i$, corresponde a minimização de $Q = \sum_{i=1}^{n} \epsilon_i^2 = \sum_{i=1}^{n}(X_i - \mu)^2$, em que estimador resultante é $\hat{\mu} = m_d$.

v) ( ) O método dos quadrados mínimos exige normalidade de X_i.

w) ( ) Os modelos probabilísticos são, em geral, dependentes de parâmetros e o objetivo de se realizarem experimentos e/ou amostragem é buscar obter informações a respeito destes parâmetros.

x) ( ) A estimativa de máxima verossimilhança, em um experimento específico, para p de um modelo binomial $L(p;\, X, n) = p^X(1-p)^{n-X}$ é $\hat{p} = 0{,}4$, como pode ser observado da figura da função de verossimilhança relativa (função de verossimilhança dividida pelo seu máximo) a seguir:



Verossimilhança relativa

y) ( ) Nos processos de estimação intervalar de médias de populações normais com variâncias conhecidas, o intervalo de confiança resultante é derivado a partir da afirmativa probabilística:

$$P\left(\bar{X} - Z_{\alpha/2}\frac{\sigma}{\sqrt{n}} \leq \mu \leq \bar{X} + Z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\right) = 1 - \alpha.$$

z) ( ) A função de verossimilhança é dada pelo produto das funções densidades de cada X_j's, $j = 1, 2, \cdots, n$, porque a independência é garantida pelo uso de sorteios ou casualizações nas amostragens ou experimentos.

2) Com base em uma população finita fictícia $X = \{1, 2, 4\}$ obter todas as amostras com reposição de tamanho $n = 3$. Estimar em cada uma delas a média, a variância e o desvio padrão.

a) Obter a distribuição de amostragem de $\bar{X}$ e calcular $E(\bar{X})$, $E(S^2)$ e $E(S)$. Que conclusões podem ser retiradas confrontando as respectivas esperanças com μ, σ^2 e σ, em relação ao viés?

b) Calcular a variância das médias $\sigma^2_{\bar{X}} = \text{Var}(\bar{X})$ no sistema de amostragem. Confrontar com o resultado teórico em que $\sigma^2_{\bar{X}} = \sigma^2/n$.

3) Supondo-se que a amostra de $n = 10$ elementos a seguir é proveniente de uma distribuição normal com média desconhecida μ e variância $\sigma^2 = 3$, obter o intervalo de 95% de confiança para μ. Interpretar o intervalo gerado.

7,15	12,60	10,03	9,13	8,93
11,18	13,39	8,61	8,18	12,38

Resolução

1) Marcar como verdadeira ou falsa as seguintes afirmativas:

   a) (V) O método dos momentos consiste em igualar os momentos amostrais como a média, a variância, o momento amostral de assimetria, etc. com os valores esperados correspondentes, como a média (momento de ordem 1 não centrado), a variância (momento de ordem 2 centrado na média) e momento de assimetria (momento de ordem 3 centrado na média), etc.

   b) (V) Parâmetro é uma característica populacional, sendo uma constante, como por exemplo a média μ, a variância σ^2, entre outras possibilidades.

   c) (F) Estimador é uma expressão que permite estimar algum parâmetro e não possui erro associado ao seus resultados.

   d) (V) Estimador é uma função dos elementos amostrais e, portanto, é uma variável aleatória, como a média amostral $\bar{X}$ e a variância amostral S^2, entre outros casos.

   e) (F) A estimativa de máxima verossimilhança é o valor de θ que minimiza o valor de $L(\theta; x_1, x_2, \cdots, x_n)$.

   f) (F) Estimativa é valor que representa o parâmetro e é constante para todas as amostras aleatórias extraídas de uma mesma população.

   g) (V) Estimativas são constantes de uma amostra específica e representa o resultado de um estimador particular. As estimativas de amostras para amostras variam.

   h) (V) O viés ou vício de um estimador é a propriedade em que quando o valor médio deste estimador (esperança matemática) em todas as amostras possíveis de um determinado tamanho n de uma população (na distribuição de amostragem) é igual ao parâmetro, dizemos que o estimador é não viesado ou não viciado.

   i) (V) São exemplos de estimadores não viesados, a média amostral $\bar{X}$ e a variância amostral S^2 e de estimador viesado o desvio padrão amostral S.

   j) (V) Um estimador mais eficiente é aquele dentre os estimadores não viesados de um mesmo parâmetro que possui a menor variância.

   k) (F) Um estimador eficiente é aquele que sua variância vai para zero, quando n tende para infinito.

   l) (V) A mediana amostral é menos eficiente que média amostral para estimar a média μ em populações normais, pois possui maior variância.

   m) (F) Um estimador é consistente se sua variância vai para ∞ quando n tende a zero.

   n) (F) Estimador viciado ou viesado é aquele que sua distribuição apresenta média igual ao parâmetro que ele está estimando.

   o) (V) Um estimador é consistente se sua variância tende a zero na medida que o tamanho amostral n cresce, tendendo para ∞.

   p) (V) São métodos de estimação pontual os métodos de momentos, da máxima verossimilhança e dos quadrados mínimos.

   q) (F) O método dos momentos é obtido tomando-se o produto das funções densidades de cada observação amostral X_j, f_{X_j} e maximizando esta função resultante deste produto em relação aos parâmetros.

   r) (V) A função de verossimilhança L é dada por:

$$L(\theta; x_1, x_2, \cdots, x_n) = f(x_1; \theta) f(x_2; \theta) \times \cdots \times f(x_n; \theta) = \prod_{i=1}^{n} f(x_i; \theta).$$

   s) (V) A estimativa de máxima verossimilhança é o valor de θ que maximiza o valor de $L(\theta; x_1, x_2, \cdots, x_n)$.

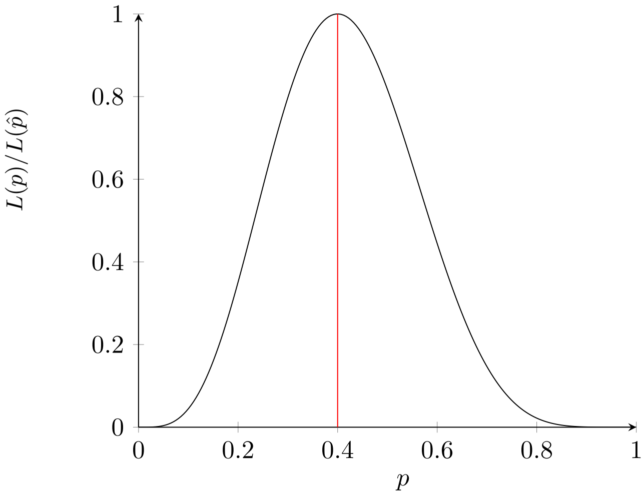
   t) (V) O método dos quadrados mínimos, para estimar a média em um modelo $Y_i = \mu + \epsilon_i$, corresponde a minimização de $Q = \sum_{i=1}^{n} \epsilon_i^2 = \sum_{i=1}^{n} (X_i - \mu)^2$, em que estimador resultante é $\hat{\mu} = \bar{X}$.

   u) (F) O método dos quadrados mínimos, para estimar a média em um modelo $Y_i = \mu + \epsilon_i$, corresponde a minimização de $Q = \sum_{i=1}^{n} \epsilon_i^2 = \sum_{i=1}^{n} (X_i - \mu)^2$, em que estimador resultante é $\hat{\mu} = m_d$.

   v) (F) O método dos quadrados mínimos exige normalidade de X_i.

   w) (V) Os modelos probabilísticos são, em geral, dependentes de parâmetros e o objetivo de se realizarem experimentos e/ou amostragem é buscar obter informações a respeito destes parâmetros.

   x) (V) A estimativa de máxima verossimilhança, em um experimento específico, para p de um modelo binomial $L(p;\, X, n) = p^X (1-p)^{n-X}$ é $\hat{p} = 0{,}4$, como pode ser observado da figura da função de verossimilhança relativa (função de verossimilhança dividida pelo seu máximo) a seguir:



Verossimilhança relativa

y) (V) Nos processos de estimação intervalar de médias de populações normais com variâncias conhecidas, o intervalo de confiança resultante é derivado a partir da afirmativa probabilística:

$$P\left(\bar{X} - Z_{\alpha/2}\frac{\sigma}{\sqrt{n}} \leq \mu \leq \bar{X} + Z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\right) = 1 - \alpha.$$

z) (V) A função de verossimilhança é dada pelo produto das funções densidades de cada X_j's, $j = 1, 2, \cdots, n$, porque a independência é garantida pelo uso de sorteios ou casualizações nas amostragens ou experimentos.

2) Considerando o exercício relativo a distribuição de amostragem temos: a) As amostras e as respectivas médias, variâncias e desvios padrões estão apresentadas a seguir:

Amostras	$\bar{X}$	S^2	S
1 1 1	1,0000	0,0000	0,0000
1 1 2	1,3333	0,3333	0,5774
1 1 4	2,0000	3,0000	1,7321
1 2 1	1,3333	0,3333	0,5774
1 2 2	1,6667	0,3333	0,5774
1 2 4	2,3333	2,3333	1,5275
1 4 1	2,0000	3,0000	1,7321
1 4 2	2,3333	2,3333	1,5275
1 4 4	3,0000	3,0000	1,7321
2 1 1	1,3333	0,3333	0,5774
2 1 2	1,6667	0,3333	0,5774
2 1 4	2,3333	2,3333	1,5275
2 2 1	1,6667	0,3333	0,5774
2 2 2	2,0000	0,0000	0,0000
2 2 4	2,6667	1,3333	1,1547
2 4 1	2,3333	2,3333	1,5275
2 4 2	2,6667	1,3333	1,1547
2 4 4	3,3333	1,3333	1,1547
4 1 1	2,0000	3,0000	1,7321
4 1 2	2,3333	2,3333	1,5275
4 1 4	3,0000	3,0000	1,7321
4 2 1	2,3333	2,3333	1,5275
4 2 2	2,6667	1,3333	1,1547
4 2 4	3,3333	1,3333	1,1547
4 4 1	3,0000	3,0000	1,7321
4 4 2	3,3333	1,3333	1,1547
4 4 4	4,0000	0,0000	0,0000
Média	2,3333	1,5556	1,1093
Variância	0,5185	1,2099	0,3251

As esperanças matemáticas, neste caso, são as médias dos estimadores na distribuição de amostragem, ou seja, nas 27 amostras possíveis apresentadas na tabela anterior. Assim, $E(\bar{X}) = 2{,}3333$, $E(S^2) = 1{,}5556$ e $E(S) = 1{,}1093$. A população apresenta os seguintes parâmetros: $\mu = 2{,}3333$, $\sigma^2 = 1{,}5556$ e $\sigma = 1{,}2472$. Se compararmos os valores médios na distribuição de amostragem, esperanças com os respectivos parâmetros verificamos que os valores são iguais para os estimadores média e variância amostrais, mas diferentes para o desvio padrão amostral. Isto significa que $\bar{X}$ e S^2 são não viciados ou não viesados e S é um estimador viciado ou viesado.

b) A variância das médias, na distribuição de amostragem é, $\sigma^2_{\bar{X}} = 0{,}5185$ e a variância teórica é:

$$\sigma^2_{\bar{X}} = \frac{\sigma^2}{n} = \frac{1{,}5556}{3} = 0{,}5185.$$

Assim, concluímos que $\sigma^2_{\bar{X}} = \sigma^2/n$ é igual a variância dos valores de $\bar{X}$ na distribuição de amostragem. Assim, não precisamos fazer toda a distribuição de amostragem para saber que:

$$\sigma^2_{\bar{X}} = \frac{\sigma^2}{n} \qquad \text{e} \qquad \mu_{\bar{X}} = \mu.$$

Sabemos ainda que se $n > 30$, a distribuição de $\bar{X}$ é aproximadamente normal, pelo teorema do limite central, o que é uma dos resultados mais importantes da estatística.

3) A média e a variância amostrais são:

$$\bar{X} = 10{,}158 \qquad \text{e} \qquad S^2 = 4{,}480996.$$

O quantil superior $Z_{\alpha/2} = Z_{0,025} = 1{,}96$ é obtido por `qnorm(0.975)` no programa R.

Assim, o intervalo de confiança é:

$$\begin{aligned} IC_{0,95}(\mu) &= \bar{X} \pm Z_{\alpha/2}\frac{\sigma}{\sqrt{n}} \\ &= 10{,}158 \pm 1{,}96 \times \frac{\sqrt{3}}{\sqrt{10}} \\ &= 10{,}158 \pm 1{,}0735 \\ &= [9{,}0845,\ 11{,}2315]. \end{aligned}$$

Assim, a verdadeira média μ é um valor do intervalo [9,0845, 11,2315], com 95% de confiança.

Observação: Matérias para as duas últimas provas

```
Prova 2: Capítulo 8: seções 8.1, 8.2 (8.2.1, 8.2.2, 8.2.3), 8.3,
         Capítulo 9, seções 9.1 (9.1.1, 9.1.2, 9.1.3, 9.1.4, 9.1.8, 9.1.9),
                           9.2 (9.2.1, 9.2.3),9.4 (9.4.1)
         Capítulo 11, seções 11.1 (11.1.1)

Prova 3: Capítulo 8, secão 8.4
         Capítulo 9, Seções 9.1 (9.1.5), 9.2 (9.2.2), 9.4(9.4.3)
         Capítulo 11, Seções 11.1 (11.1.3)
```