Estimação - Capítulo 7 - Magalhães e Lima

Inferência Estatística - processo que permite construirmos afirmações sobre características de uma população com base nos resultados da amostra.

População - Conjunto de elementos (indivíduos, objetos, animais) tendo pelo menos uma variável de interesse em comum.

Amostra - Qualquer subconjunto da população.

Exemplo 1

População - 1.000 alunos de uma escola do ensino médio

Objetivo - estudar a proporção de alunos que pretende fazer vestibular.

Toma-se uma amostra de 20 alunos.

Variável de Interesse

Intenção de fazer vestibular

- 0 não deseja fazer vestibular
- 1 deseja fazer vestibular

Exemplo 2

Duração de lâmpadas de um novo tipo.

População - todas as possíveis lâmpadas fabricadas pelo novo processo.

É registrado o tempo de vida em horas de uma amostra de 100 lâmpadas

Nesse caso, não poderíamos obter informações sobre a população toda pois inutilizaria todas as lâmpadas.

X - tempo de vida das lâmpadas desse novo tipo

X_1 - tempo de vida da 1ª lâmpada da amostra
X_2 - tempo de vida da 2ª lâmpada da amostra
$\vdots$
X_{100} - tempo de vida da 100ª lâmpada da amostra

Se X é a variável de interesse e n é o tamanho da amostra, então, a amostra será indicada por

$(X_1, X_2 \dots X_n)$.

No exemplo 1, $n = 20$ e

$$X_i = \begin{cases} 0 & \text{se o aluno não deseja fazer o vestibular} \\ 1 & \text{se o aluno deseja fazer o vestibular} \end{cases}$$

$i = 1, 2 \dots 20$

Formalmente, admite-se que o comportamento probabilístico da característica de interesse populacional é representado por uma v. a. X da qual sorteamos uma amostra de tamanho n denotada por $(X_1, X_2 \ldots X_n)$.

Em geral, serão consideradas amostras aleatórias simples com reposição.

Esse procedimento garante que $X_1, X_2 \ldots X_n$ sejam independentes todos com a mesma distribuição de probabilidades da variável original X.

Definição 1: Parâmetro

Característica populacional desconhecida sobre as quais se tem interesse e deseja-se obter informação.

Ex: μ, σ, p, θ.

μ - duração média das lâmpadas do novo tipo (exemplo 2)

p - proporção de alunos no colégio que deseja fazer vestibular (exemplo 1)

Definição 2 - Estimador e Estimativa

Qualquer função dos elementos amostrais, construída com o objetivo de obter informações sobre o parametro desconhecido é denominada estimador.

Notação: $\hat{\mu}$, $\hat{\sigma}$, $\hat{p}$, $\hat{\theta}$

Aos valores numéricos assumidos pelos estimadores denominamos estimativas pontuais ou somente estimativas.

$\hat{\mu}$ - estimador da média populacional μ

De modo geral, $\hat{\theta} = T(X_1, X_2 \dots X_n)$

Obs:

- $X_1, X_2 \dots X_n$ são variáveis aleatórias
  $\Longrightarrow$ estimador é uma variável aleatória
  (seus valores variam de amostra para amostra)

- A distribuição de probabilidades do estimador é denominada distribuição amostral (fornece todas as propriedades do estimador)

Exemplo 3

Altura média de jovens entre 15 e 18 anos nascidos na região sudeste do país.

População - todos os jovens com idade entre 15 e 18, nascidos na região sudeste.

Tomada uma amostra de tamanho $n = 10$, ao acaso.

Variável: X - altura

Parâmetro de interesse: μ - altura média dos jovens

$$\mu = E(X)$$

amostra: $(X_1, X_2 \dots X_{10})$

altura do 1º jovem da amostra

altura do 10º jovem da amostra

Possíveis estimadores de μ:

$$\hat{\mu}_1 = T_1(X_1, X_2 \dots X_{10}) = \frac{\min\{X_i\} + \operatorname{Max}\{X_i\}}{2}$$

$$\hat{\mu}_2 = T_2(X_1, X_2 \ldots X_{10}) = X_1$$

$$\hat{\mu}_3 = T_3(X_1, X_2 \ldots X_{10}) = \frac{X_1 + X_2 + \ldots + X_{10}}{10}$$

→ média amostral

Exemplo 1 (Continuação)

No exemplo do vestibular

População: alunos da escola de ensino médio

Parâmetro de interesse: p - proporção de alunos da escola que desejam fazer vestibular

amostra de $n = 20$ alunos

$$X_i = \begin{cases} 0 & \text{se o aluno não deseja fazer vestibular} \\ 1 & \text{se o aluno deseja fazer vestibular} \end{cases}$$

$i = 1, 2 \ldots 20$

Estimador Natural: Proporção amostral de alunos que desejam fazer vestibular

$$\hat{p} = \frac{\text{número de alunos na amostra que desejam fazer vestibular}}{20}$$

$$= \frac{X_1 + X_2 + \ldots + X_{20}}{20}$$

Exemplo 4

Para analisar o nível de colesterol (em mg/dL) em uma população de esportistas, coletou-se uma amostra de 10 atletas, obtendo-se os seguintes valores:

180, 196, 185, 165, 190, 195, 180, 176, 165 e 195

Parâmetro de interesse: μ - nível médio de colesterol da população de esportistas

Estimativa:

$$\bar{x} = \frac{180 + 196 + 185 + \dots + 176 + 165 + 195}{10} = 182{,}7$$

O limite de colesterol para pessoas sadias é 200.
No caso, $\bar{x} < 200$.

Deseja-se estimar a proporção p de atletas com taxa acima de 190. Se X_i - nível de colesterol do i-ésimo atleta na amostra

$$Y_i = \begin{cases} 1 & \text{se } X_i > 190 \\ 0 & \text{se } X_i \leq 190 \end{cases} \qquad i = 1, 2, \dots, 10$$

$$\hat{p} = \frac{Y_1 + Y_2 + \dots + Y_{10}}{10} = \frac{3}{10} = 0{,}3 \rightarrow \text{estimativa de } p$$

Estimadores Naturais

$\hat{p}$ - proporção amostral
estimador da proporção populacional p

$\hat{\mu} = \bar{X} = \dfrac{\sum_{i=1}^{n} X_i}{n}$ média amostral

estimador da média populacional μ

$\hat{\sigma}^2 = \dfrac{1}{n} \sum_{i=1}^{n} (X_i - \bar{X})^2$ ou $\hat{\sigma}^2 = S^2 = \dfrac{1}{n-1} \sum_{i=1}^{n} (X_i - \bar{X})^2$

Variância amostral estimadores da variância populacional σ^2

Outros possíveis estimadores

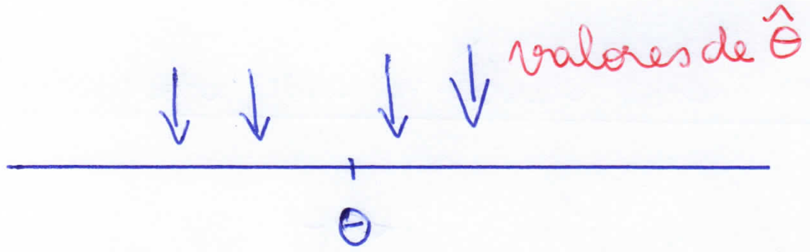
Para μ : $\hat{\mu} = Md\{X_i,\ i = 1, 2, \dots n\}$

Para σ^2 : Amplitude amostral $= Max\{X_i\} - Min\{X_i\}$

Propriedades dos Estimadores

Θ - Parâmetro desconhecido

Estimador é uma v.a. pois é função de $(X_1, X_2 \dots X_n)$ e varia de amostra para amostra.



Definição 3 - Estimador não viciado

Um estimador $\hat{\Theta}$ é não viciado ou não viesado para o parâmetro Θ se

$$E(\hat{\Theta}) = \Theta.$$

O vício ou viés de um estimador é dado por

$$\text{Vício} = E(\hat{\Theta}) - \Theta.$$

Definição 4 - Eficiência

Dados dois estimadores $\hat{\theta}_1$ e $\hat{\theta}_2$, não viciados para um parâmetro θ, dizemos que $\hat{\theta}_1$ é mais eficiente do que $\hat{\theta}_2$ se

$$Var(\hat{\theta}_1) < Var(\hat{\theta}_2).$$

Lembrando: $Var(X) = E\left[(X-\mu)^2\right]$

onde $\mu = E(X)$.

$$E(\hat{\theta}_1) = \theta \qquad E(\hat{\theta}_2) = \theta$$

$$\therefore \; Var(\hat{\theta}_1) = E\left[(\hat{\theta}_1 - \theta)^2\right]$$

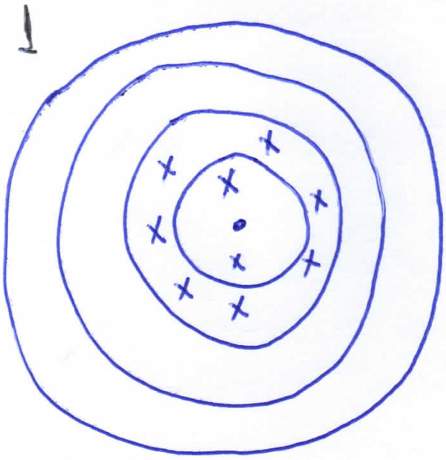
$$Var(\hat{\theta}_2) = E\left[(\hat{\theta}_2 - \theta)^2\right]$$

$$Var(\hat{\theta}_1) < Var(\hat{\theta}_2) \Rightarrow E\left[(\hat{\theta}_1 - \theta)^2\right] < E\left[(\hat{\theta}_2 - \theta)^2\right]$$

$\hat{\theta}_1$ é mais concentrado em torno de θ que $\hat{\theta}_2$

Obs: Variância de um estimador só será uma medida de sua eficiência se o estimador for não viciado.

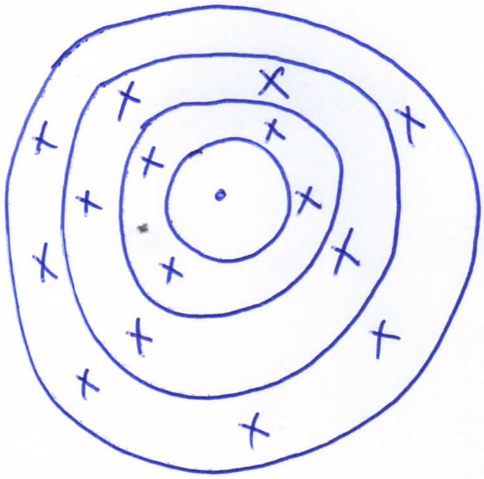
1



Não viciado

Variância pequena

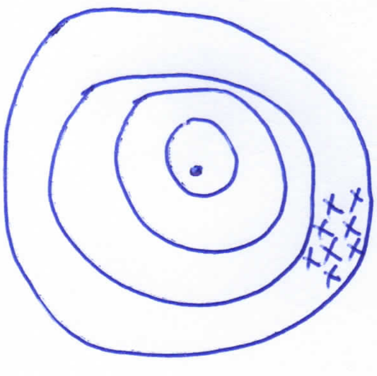
2



Não viciado

Variância grande

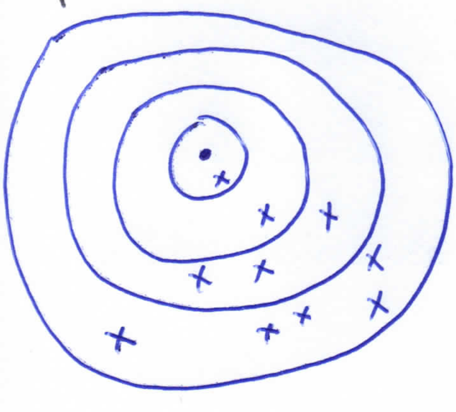
3



Viciado

Variância pequena

4



Viciado

Variância grande

Na figura 3

O estimador tem variância pequena mas é viciado
∴ altamente concentrado em torno de sua média,
mas esta média não é o valor do parâmetro.

A Medida de Precisão de um estimador será sempre

$$* = E\left\{ [\hat{\theta} - \theta]^2 \right\}$$

Lembrando $\operatorname{Var}(\hat{\theta}) = E\left\{ [\hat{\theta} - E(\hat{\theta})]^2 \right\}$

Se $E(\hat{\theta}) = \theta$ ($\hat{\theta}$ não viciado) então

$$* = E\left\{ [\hat{\theta} - \theta]^2 \right\} = \operatorname{Var}(\hat{\theta})$$

e nesse caso $\operatorname{Var}(\hat{\theta})$ é uma medida da precisão do estimador.

Se $\hat{\theta}$ é viciado, $E(\hat{\theta}) \neq \theta$,

$$\operatorname{Var}(\hat{\theta}) = E\left\{ [\hat{\theta} - \underbrace{E(\hat{\theta})}_{\neq \theta}]^2 \right\}$$

não mede a variabilidade desejada.

Nesse caso
- não usar $\operatorname{Var}(\hat{\theta})$
- usar $* = E\left\{ [\hat{\theta} - \theta]^2 \right\}$

Definição 5 - Erro Quadrático Médio

A quantidade *, $E\{[\hat{\theta}-\theta]^2\}$, é denominada Erro Quadrático Médio do Estimador $\hat{\theta}$.

$$EQM(\hat{\theta}) = E\{[\hat{\theta}-\theta]^2\}.$$

notação

Obs:

1) Se $\hat{\theta}$ é não viciado, $E(\hat{\theta})=\theta$ e $EQM(\hat{\theta}) = Var(\hat{\theta})$

2) Prova-se que

$$EQM(\hat{\theta}) = Var(\hat{\theta}) + Vício^2$$

$$Vício = E(\hat{\theta}) - \theta$$

3) Referência para Erro Quadrático Médio: Bussab e Morettin, pag 302.

Exemplo 5 - pag 229 Magalhães e Lima

Considere a v. a. X associada a uma certa população, com a seguinte distribuição de probabilidades

X	0	10	20	30
$p_i = P(X = x_i)$	0,2	0,3	0,3	0,2

$$\mu = E(X) = 0 . 0{,}2 + 10 . 0{,}3 + 20 . 0{,}3 + 30 . 0{,}2 = 15$$

$$E(X^2) = 0^2 . 0{,}2 + 10^2 . 0{,}3 + 20^2 . 0{,}3 + 30^2 . 0{,}2 = 330$$

$$Var(X) = 330 - 15^2 = 105 = \sigma^2$$

Supondo agora que μ seja desconhecido e será estimado com base em uma amostra de tamanho $n = 2$, com reposição.

Nestas condições, sejam

X_1 - valor de X observado na 1ª extração
X_2 - valor de X observado na 2ª extração

$n = 2$

A distribuição conjunta de (X_1, X_2) é

$X_1 \backslash X_2$	0	10	20	30	$P(X_1 = x)$
0	0,04	0,06	0,06	0,04	0,2
10	0,06	0,09	0,09	0,06	0,3
20	0,06	0,09	0,09	0,06	0,3
30	0,04	0,06	0,06	0,04	0,2
$P(X_2 = x)$	0,2	0,3	0,3	0,2	

distribuição de X_1

distribuição de X_2

$$P(X_1 = 0, X_2 = 0) = \underbrace{P(X_2 = 0 \mid X_1 = 0)}_{P(X_2 = 0)} P(X_1 = 0) = 0{,}2 \cdot 0{,}2 = 0{,}04$$

$$P(X_1 = 0, X_2 = 10) = \underbrace{P(X_2 = 10 \mid X_1 = 0)}_{P(X_2 = 10)} P(X_1 = 0) = 0{,}3 \cdot 0{,}2 = 0{,}06$$

Obs:

- X_1 e X_2 têm a mesma distribuição da variável X
- X_1 e X_2 são variáveis aleatórias independentes (Consequência da amostragem com reposição)

Para estimar μ, sejam os estimadores

estamos supondo que μ é desconhecido

$\hat{\mu}_1 = X_1$

$\hat{\mu}_2 = \dfrac{X_1 + X_2}{2}$

Como X_1 tem a mesma distribuição de X,

$E(\hat{\mu}_1) = \mu \ (=15)$ $\hat{\mu}_1$ é um estimador não viciado de μ

$Var(\hat{\mu}_1) = Var(X) = \sigma^2 = 105$

Para $\hat{\mu}_2 = \bar{X}$, Distribuição de Probabilidades de $\bar{X}$

$\bar{X}$	0	5	10	15	20	25	30
pi	0,04	0,12	0,21	0,26	0,21	0,12	0,04

$P(\bar{X} = 10) = P((0,20)) + P((20,0)) + P((10,10)) =$

$= 0,06 + 0,06 + 0,09 = 0,21$

$P(\bar{X} = 15) = P((0,30)) + P((30,0)) + P((10,20)) + P((20,10))$

$= 0,04 + 0,04 + 0,09 + 0,09 = 0,26$

$E(\bar{X}) = 0.0,04 + 5.0,12 + 10.0,21 + 15.0,26 + 20.0,21 +$

$+ 25.0,12 + 30.0,04 = 15 = \mu$

$\therefore \hat{\mu}_2 = \bar{X}$ é não viciado para μ.

Calculando $Var(\hat{\mu}_2) = Var(\bar{X})$ obtemos

$$Var(\hat{\mu}_2) = 52{,}5$$

Comparação entre os estimadores

$\hat{\mu}_1$ e $\hat{\mu}_2$ são não viciados para μ

$$Var(\hat{\mu}_2) = 52{,}5 < Var(\hat{\mu}_1) = 105$$

$\Rightarrow$ $\hat{\mu}_2$ é mais eficiente que $\hat{\mu}_1$.