Introdução à Probabilidade e à Estatística II

Estimação I

Lígia Henriques-Rodrigues

MAE0229 – 1º semestre 2018

1. Introdução

Objetivo: Obter, com base na informação disponível a partir da amostra, um valor ou um conjunto de valores para o parâmetro desconhecido.

Contexto: Suponhamos que pretendemos estudar uma determinada população, caracterizada por uma v.a. X, com distribuição $f(x;\theta)$ conhecida, mas cujo parâmetro caracterizador, θ, é desconhecido.

A solução passa por recolher informação via AAS $X = (X_1, \ldots, X_n)$ e respectiva concretização $x = (x_1, \ldots, x_n)$, para obter um valor para θ ou um conjunto de valores para θ que pertença(m) ao conjunto de valores que o parâmetro pode assumir (espaço paramétrico).

Tipos de Estimação	**Objetivo**
Pontual	obter um valor para θ
Intervalar	obter um conjunto de valores para θ

2. Estimação Pontual

Estimador Pontual de θ: Um estimador pontual T do parâmetro θ é qualquer função das observações da amostra aleatória, ou seja, $T = g(X_1, \ldots, X_n)$.

Estimativa Pontual de θ: Estimativa pontual é o valor assumido pelo estimador para a amostra observada $x_1, \ldots, x_n$, ou seja, $t = g(x_1, \ldots, x_n)$.

As estimativas são números reais fixos e portanto não são variáveis aleatórias.

Parâmetro Populacional	**Estimador Pontual**
Média populacional μ	Média amostral $\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$
Variância populacional σ^2	Variância amostral $S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X})^2$
Proporção populacional p	Proporção amostral $\hat{p} = \frac{\sum_{i=1}^{n} X_i}{n}$

2.1 Propriedades dos Estimadores

Definição: O estimador T é não viesado ou não viciado para θ se

$$E(T) = \theta, \quad \forall\ \theta.$$

Definição: Designamos de viés ou vício de um estimador T a diferença $V(T) = E(T) - \theta$.

Exercício 1: Mostre que $\hat{p}$ é um estimador não viesado de p.

Exercício 2: Considere uma AAS $(X_1, \ldots, X_n)$ da v.a. X e mostre que $\hat{\mu} = \overline{X}$ é um estimador não viesado de μ.

Exercício 3: Considere uma população com N elementos e a variância populacional

$$\sigma^2 = \frac{1}{N}\sum_{i=1}^{N}(X_i - \mu)^2,$$

em que $\mu = \frac{1}{N}\sum_{i=1}^{N} X_i$ é a média populacional. Um possível estimador para σ^2 é

$$\hat{\sigma}^2 = S_X^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X})^2.$$

1. Mostre que $E(S_X^2) = \frac{n-1}{n}\sigma^2$.
   Sugestão: relembre que $\frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X})^2 = \frac{1}{n}\sum_{i=1}^{n} X_i^2 - \overline{X}^2$.
2. Calcule o viés do estimador S_X^2.
3. Proponha, com base em S_X^2, um estimador não viesado para σ^2.

Definição: Uma sequência $\{T_n\}$ de estimadores de um parâmetro θ é consistente se, para todo o $\varepsilon > 0$,

$$P\{|T_n - \theta| > \varepsilon\} \to 0, \ n \to \infty.$$

Proposição: Uma sequência $\{T_n\}$ de estimadores de θ é consistente se

$$\lim_{n\to\infty} E(T_n) = \theta$$

e

$$\lim_{n\to\infty} Var(T_n) = 0.$$

Exercício 4: Mostre que $\hat{p}$ e $\overline{X}$ são estimadores consistentes dos parâmetros p e μ, respectivamente.

Exercício 5: Sabendo que $Var(S^2) = \frac{2\sigma^4}{n-1}$, mostre que S^2 é um estimador consistente para σ^2.

Definição: Se T e T' são dois estimadores não viesados de um mesmo parâmetro θ, e ainda

$$Var(T) < Var(T'),$$

então T diz-se mais eficiente do que T'.

Definição: Chama-se erro quadrático médio (EQM) do estimador T ao valor,

$$EQM(T;\theta) = E[(T-\theta)^2].$$

Nota: Mostra-se que $EQM(T;\theta) = Var(T) + (E(T)-\theta)^2$.

Definição: Se T e T' são dois estimadores de um mesmo parâmetro θ. Então, T diz-se mais eficiente do que T', se e só se

$$EQM(T;\theta) < EQM(T';\theta),\ \forall\theta \in \Theta.$$