Universidade Federal de Santa Maria
Departamento de Estatística

ESTIMAÇÃO PONTUAL

Estatística Inferencial

Professora: Laís Helen Loose

Introdução

ESTIMAÇÃO PONTUAL E INTERVALAR

Estimação Pontual

A estimação pontual utiliza os dados da amostra para chegar a um **único número** que representa um valor plausível para o parâmetro de interesse.

Estimação Intervalar

A estimação intervalar (intervalo de confiança) fornece um **intervalo de valores** em que se espera que o parâmetro desconhecido esteja.

DEFINIÇÕES

Amostra aleatória

Uma **amostra aleatória** de tamanho n é uma coleção de variáveis aleatórias $X_1, X_2, \ldots, X_n$ que são independentes e identicamente distribuídas.

Amostra observada

O conjunto formado pelos valores observados da amostra aleatória após a realização do experimento aleatório, denotado por $x_n = (x_1, \ldots, x_n)$, é chamado de **amostra observada.** A amostra observada é fixa, ou seja, não aleatória.

DEFINIÇÕES

Estatística

Uma **estatística** é uma variável aleatória $T = t(\underset{\sim}{X})$, definida como função da amostra aleatória e que não depende do parâmetro θ.

- **Exemplos:**
  - $\overline{X} = \dfrac{1}{n} \sum_{i=1}^{n} X_i$ (média amostral)
  - $S^2 = \dfrac{1}{n-1} \sum_{i=1}^{n} (X_i - \overline{X})^2$

- Um **estimador é uma estatística** que é usada para inferir o valor de um parâmetro desconhecido.
- **Como obter estimadores?** Métodos de estimação

DEFINIÇÕES

Métodos de estimação

Pode-se obter estimadores para a quantidade desconhecida θ aplicando algum método de estimação. Vamos estudar dois métodos:

- **Método dos momentos.**
- **Método de máxima verossimilhança.**

Existem outros métodos de estimação, a saber: mínimos quadrados, quasi-verossimilhança, bayesiano, etc.

Estimação Pontual

MÉTODO DOS MOMENTOS

Momentos populacional e amostral

Seja $X_1, \ldots, X_n$ uma amostra aleatória de uma variável aleatória X. Definimos o **k-ésimo momento populacional** por

$$\mu_k = \mathbb{E}(X^k)$$

O correspondente **k-ésimo momento amostral** é dado por

$$m_k = \frac{1}{n}\sum_{i=1}^{n} X_i^k$$

MÉTODO DOS MOMENTOS

Método dos Momentos

O método dos momentos consiste em igualar os momentos amostrais com os momentos populacionais correspondentes. O estimador de momentos é obtido pela solução da equação, ou do sistema de equações, quando possível.

Estimador e Estimativa

- **Um estimador pontual** é uma estatística usada para estimar o parâmetro desconhecido θ. Denotaremos o estimador por $\widehat{\theta}(\underset{\sim}{X}_n) = \widehat{\theta}(X_1, \ldots, X_n)$.
- **Estimativa** é o valor numérico do estimador após observarmos a amostra. Ou seja, é um número real fixo, portanto **não é uma variável aleatória.** Denotaremos a estimativa por $\widehat{\theta}(\underset{\sim}{x}_n) = \widehat{\theta}(x_1, \ldots, x_n)$.

MÉTODO DOS MOMENTOS

Exemplo

Seja $X_1, X_2, \ldots, X_n$ uma amostra aleatória de $X \sim$ Bernoulli(θ). Encontre o estimador de momentos para o parâmetro desconhecido.

MÉTODO DOS MOMENTOS

Exemplo

Seja $X_1, X_2, \ldots, X_n$ uma amostra aleatória de $X \sim \text{Bernoulli}(\theta)$. Encontre o estimador de momentos para o parâmetro desconhecido.

Solução

- Primeiro momento populacional $\mathbb{E}(X) = \theta$.
- Primeiro momento amostral: $m_1 = \dfrac{1}{n}\sum_{i=1}^{n} X_i = \overline{X}$.
- Dessa maneira, o estimador de momentos para o parâmetro é dado por:

$$\widehat{\theta}(\boldsymbol{X}) = \overline{X}$$

.

MÉTODO DOS MOMENTOS

Exemplo

Seja $X_1, X_2, \ldots, X_n$ uma amostra aleatória de $X \sim$ Uniforme$(-\theta, \theta)$. Encontre o estimador de momentos para o parâmetro desconhecido.

$$f_\theta(x) = \frac{1}{2\theta}, \quad -\theta < x < \theta$$

MÉTODO DOS MOMENTOS

Solução

- O primeiro momento populacional é nulo: $\mathbb{E}(X) = 0$ (não é útil para encontrar θ)
- Segundo momento populacional: $\mathbb{E}(X^2) = \text{Var}(X) + [\mathbb{E}(X)]^2 = \theta^2/3$
- Segundo momento amostral: $m_2 = \dfrac{1}{n}\sum_{i=1}^{n} X_i^2 = \overline{X}$
- Igualando o segundo momento populacional com o segundo amostral, encontramos que o estimador de momentos para θ é dado por:

$$\widehat{\theta}(\boldsymbol{X}) = \sqrt{\frac{3}{n}\sum_{i=1}^{n} X_i^2}$$

MÉTODO DOS MOMENTOS

Exemplo

Seja $X_1, X_2, \ldots, X_n$ uma amostra aleatória de $X \sim$ Normal(μ, σ^2). Encontre o estimador de momentos para os parâmetros desconhecidos $\theta = (\mu, \sigma^2)$.

$$f_\theta(x) = \frac{1}{2\pi\sigma^2} \exp\left\{-\frac{1}{2\sigma^2}(x-\mu)^2\right\}, \quad -\infty < x < \infty$$

MÉTODO DOS MOMENTOS

Solução

- Temos dois parâmetros, assim necessitamos de pelo menos duas equações.
- Primeiro momento populacional: $\mathbb{E}(X) = \mu$.
- Segundo momento populacional: $\mathbb{E}(X^2) = \text{Var}(X) + [\mathbb{E}(X)]^2 = \sigma^2 + \mu^2$.
- Igualando esses momentos com seus respectivos momentos amostrais, temos o seguinte sistema de equações:

MÉTODO DOS MOMENTOS

Solução

$$\widehat{\mu} = \frac{1}{n}\sum_{i=1}^{n} X_i = \overline{X}$$

$$\widehat{\sigma}^2 + \widehat{\mu} = \frac{1}{n}\sum_{i=1}^{n} X_i^2$$

- Resolvendo o sistema de equações:

$$\widehat{\mu} = \overline{X} \quad \text{e} \quad \widehat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X})^2 = \frac{n-1}{n}S^2$$

Método de Máxima Verossimilhança

MÁXIMA VEROSSIMILHANÇA

Distribuição conjunta

Seja $X_1, \ldots, X_n$ uma amostra aleatória de uma variável aleatória X cuja a distribuição depende de um parâmetro θ. **A função de densidade (função de probabilidade) conjunta** para a amostra aleatória será dada por:

- Caso contínuo:

$$f_\theta\left(x_1, x_2, \ldots, x_n\right) = \prod_{i=1}^{n} f_\theta\left(x_i\right)$$

- Caso discreto:

$$P_\theta\left(X_1 = x_1, X_2 = x_2, \ldots, X_n = x_n\right) = \prod^{n} P_\theta\left(X_i = x_i\right)$$

MÁXIMA VEROSSIMILHANÇA

Função de verossimilhança

Suponha que $x_1, x_2, \ldots, x_n$ são os valores observados da amostra aleatória. Uma vez que a amostra foi observada, a função de densidade de probabilidade conjunta pode ser reinterpretada como sendo uma função do parâmetro θ. Chamaremos esta função de **função de verossimilhança.**

$$L_{\underset{\sim}{x}_n}(\theta) = \begin{cases} \displaystyle\prod_{i=1}^{n} f_\theta\left(x_i\right) \\ \displaystyle\prod_{i=1}^{n} P_\theta\left(X_i = x_i\right) \end{cases}$$

MÁXIMA VEROSSIMILHANÇA

- **O método de estimação por máxima verossimilhança** consiste em encontrar o(s) valor(es) para o(s) parâmetro(s) θ que maximiza a função de verossimilhança $L_{\underset{\sim}{x}_n}(\theta)$, o que torna a função dos dados máxima.

- Os valores em que a função de verossimilhança atinge seu máximo são chamados de **estimativas de máxima verossimilhança.**

- Maximizar a função de verossimilhança equivale a maximizar o seu logaritmo $\ell(\theta) = \ln(L(\theta))$, uma vez que a função logarítmica é contínua e monótona crescente em seu domínio de definição.

MÁXIMA VEROSSIMILHANÇA

Exemplo

Seja $X_1, \ldots, X_n$ uma amostra aleatória de $X \sim f_\theta(x)$. Encontre o estimador de máxima verossimilhança para:

1. $X \sim$ Bernoulli(p)
2. $X \sim$ Exponencial(λ)
3. $X \sim$ Normal(μ, σ^2)

MÁXIMA VEROSSIMILHANÇA

Solução Distribuição Bernoulli

- Se $X \sim \text{Bernoulli}(p)$, então a função de probabilidade é dada por:

$$P(X = x) = p^x(1-p)^{1-x}$$

- Dessa forma, a função de verosssimilhança é dada por:

$$L(p) = \prod_{i=1}^{n} P(X = x) = \prod_{i=1}^{n} p^{x_i}(1-p)^{1-x_i} = p^{\sum_{i=1}^{n} x_i}(1-p)^{n-\sum_{i=1}^{n} x_i}$$

- E a função de log-verossimilhança:

$$\ell(p) = \ln(p) \cdot \sum_{i=1}^{n} x_i + \ln(1-p) \cdot \left(n - \sum_{i=1}^{n} x_i\right)$$

MÁXIMA VEROSSIMILHANÇA

Solução Distribuição Bernoulli

- Derivando $\ell(p)$ em relação a p e igualando a zero, temos:

$$\frac{d\ell(p)}{dp} = \frac{\sum_{i=1}^{n} x_i}{\widehat{p}} - \frac{(n - \sum_{i=1}^{n} x_i)}{1 - \widehat{p}} = 0$$

- Resolvendo em função de p, temos que a estimativa de máxima verossimilhança é dado por:

$$p(\boldsymbol{x}) = \frac{1}{n}\sum_{i=1}^{n} x_i = \overline{x}$$

- O estimador de máxima verossimilhança do parâmetro p da distribuição Bernoulli é a média amostral.

$$p(\boldsymbol{X}) = \frac{1}{n}\sum_{i=1}^{n} X_i = \overline{X}$$

MÁXIMA VEROSSIMILHANÇA

Solução Distribuição Exponencial

- Se $X \sim \text{Exponencial}(\lambda)$, então a função de densidade de probabilidade é dada por:

$$f(x) = \lambda \exp\{-\lambda x\}$$

- As funções de verosssimilhança e log-verossimilhança são dadas, respectivamente, por:

$$L(\lambda) = \prod_{i=1}^{n} f_\theta(x_i) = \prod_{i=1}^{n} \lambda \exp\{-\lambda x_i\} = \lambda^n \exp\left\{-\lambda \sum_{i=1}^{n} x_i\right\}$$

$$\ell(\lambda) = n \ln(\lambda) - \lambda \sum_{i=1}^{n} x_i$$

MÁXIMA VEROSSIMILHANÇA

Solução Distribuição Exponencial

- Derivando $\ell(\lambda)$ em relação a λ e igualando a zero, temos:

$$\frac{d\ell(\lambda)}{d\lambda} = \frac{n}{\lambda} - \sum_{i=1}^{n} x_i = 0$$

- Resolvendo a equação, a estimativa e o estimador de máxima verossimilhança são dados, respectivamente, por:

$$\lambda(\boldsymbol{x}) = \frac{n}{\sum\limits_{i=1}^{n} x_i} = \frac{1}{\overline{x}} \qquad \text{e} \qquad \lambda(\boldsymbol{X}) = \frac{n}{\sum\limits_{i=1}^{n} X_i} = \frac{1}{\overline{X}}$$

MÁXIMA VEROSSIMILHANÇA

Solução Distribuição Normal

- Se $X \sim$ Normal(μ, σ^2), então a função de densidade de probabilidade é dada por:

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{1}{2\sigma^2}(x-\mu)^2\right\}$$

- Seja $\theta = (\mu, \sigma^2)$, as funções de verosssimilhança e log-verossimilhança são dadas, respectivamente, por:

$$L(\theta) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{1}{2\sigma^2}(x_i-\mu)^2\right\} = (2\pi\sigma^2)^{-n/2} \exp\left\{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2\right\}$$

$$\ell(\theta) = -\frac{n}{2}\ln(2\pi\sigma^2) - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2$$

MÁXIMA VEROSSIMILHANÇA

Solução Distribuição Normal

- Derivando $\ell(\theta)$ em relação a μ:

$$\frac{d\ell(\theta)}{d\mu} = \frac{1}{\sigma^2}\sum_{i=1}^{n}(x_i - \mu)$$

- Derivando $\ell(\theta)$ em relação a σ^2 :

$$\frac{d\ell(\theta)}{d\sigma^2} = \frac{n}{2\sigma^2} + \frac{1}{2(\sigma^2)^2}\sum_{i=1}^{n}(x_i - \mu)^2 = \frac{n-1}{n}S^2$$

MÁXIMA VEROSSIMILHANÇA

Solução Distribuição Normal

- Resolvendo o sistema de equações formado pelas duas equações anteriores, temos que os estimadores de máxima verossimilhança para μ e σ^2 são, respectivamente:

$$\widehat{\mu}(\boldsymbol{X}) = \sum_{i=1}^{n} X_i = \overline{X}$$

$$\widehat{\sigma}^2(\boldsymbol{X}) = \frac{1}{n}\sum_{i=1}^{n}(X_i - \mu)^2$$

PROPRIEDADES

Propriedades dos estimadores

- Como podemos avaliar a qualidade de estimador?
  - Vício (viés).
  - Variância.
  - Erro quadrático médio.

PROPRIEDADES

Definição

Um estimador $\widehat{\theta}$ é dito dito não **viesado (viciado)** para o parâmetro θ se

$$\mathbb{E}(\widehat{\theta}) = \theta$$

- Uma vez que um estimador é uma variável aleatória e seu objetivo é estimar o valor do parâmetro desconhecido, então é interessante que seu valor médio seja exatamente o valor a ser estimado.
- Se o estimador for viesado (viciado), então definimos o seu **viés (vício)** como sendo a diferença:

$$B(\widehat{\theta}) = \mathbb{E}(\widehat{\theta}) - \theta$$

DISTRIBUIÇÃO AMOSTRAL DA MÉDIA

Exemplo: Viés de média amostral

Considere uma dada amostra aleatória de tamanho n de uma distribuição $\mathrm{N}(\mu, \sigma^2)$, em que a média μ é desconhecido e é o parâmetro que estamos interessados em estimar. Encontre o vício da média amostral.

DISTRIBUIÇÃO AMOSTRAL DA MÉDIA

Solução

- A esperança do estimador é dada por:

$$\begin{aligned}\mathbb{E}(\overline{X}) &= \mathbb{E}\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right)\\ &= \frac{1}{n}\mathbb{E}\left(\sum_{i=1}^{n} X_i\right)\\ &= \frac{1}{n}n\mu = \mu\end{aligned}$$

- Portanto, o viés do estimador é dado por:

$$\begin{aligned}B(\overline{X}) &= \mathbb{E}(\overline{X}) - \mu\\ &= \mu - \mu = 0\end{aligned}$$

DISTRIBUIÇÃO AMOSTRAL DA VARIÂNCIA AMOSTRAL

Exemplo: Viés da variância amostral

Considere uma dada amostra aleatória de tamanho n de uma distribuição N(μ, σ^2), em que a variância $\sigma^2 > 0$ é desconhecida. Considerando a variância amostra como um estimador para a variância populacional, encontre o seu viés.

DISTRIBUIÇÃO AMOSTRAL DA MÉDIA

Solução

$$
\begin{aligned}
\mathbb{E}(S^2) &= \mathbb{E}\left(\frac{n-1}{n}\sum_{i=1}^{n}(X_i-\overline{X})^2\right) = \frac{n-1}{n}\cdot\mathbb{E}\left(\sum_{i=1}^{n}(X_i-\overline{X})^2\right) \\
&= \frac{n-1}{n}\cdot\mathbb{E}\left(\sum_{i=1}^{n}(X_i^2-2X_i\overline{X}+\overline{X}^2)\right) = \frac{n-1}{n}\cdot\mathbb{E}\left(\sum_{i=1}^{n}X_i^2-n\overline{X}^2\right) \\
&= \frac{n-1}{n}\cdot\left(\sum_{i=1}^{n}\mathbb{E}(X_i^2)-n\mathbb{E}(\overline{X}^2)\right) \\
&= \frac{n-1}{n}\cdot\left(\sum_{i=1}^{n}(\mu^2+\sigma^2)-n(\mu+\sigma^2/n)\right) = \sigma^2
\end{aligned}
$$

PROPRIEDADES

Variância do estimador

- Considere as seguintes situações:
  - $\widehat{\theta}_1$ é um estimador não viesado para θ, mas apresenta uma variância grande.
  - $\widehat{\theta}_2$ é um estimador viesado para θ, mas apresenta uma variância pequena.
- Qual estimador escolher (melhor)?
- É razoável escolher o estimador que apresente a **menor variância.**

PROPRIEDADES

Erro quadrático médio

- Desejamos um estimador que seja não viesado (ou que tenha um viés pequeno) e que tenha uma variância pequena.
- O **erro quadrático médio é uma medida que combina o viés e a variância.**

PROPRIEDADES

Definição

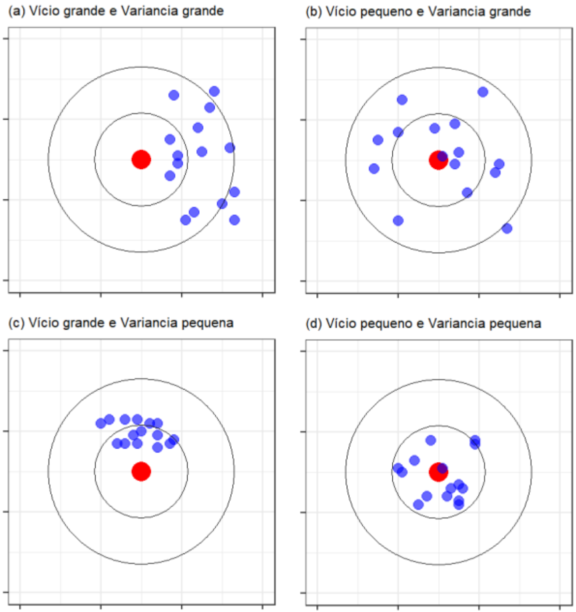
Seja $\widehat{\theta}$ um estimador para um parâmetro θ. **O erro quadrado médio** de $\widehat{\theta}$ é definido como:

$$\text{EQM}(\widehat{\theta}) = \mathbb{E}(\widehat{\theta} - \theta)^2$$

O erro quadrático médio pode ser reescrito como **função da variância e do viés** como:

$$\begin{aligned}\text{EQM}(\widehat{\theta}) &= \mathbb{E}\left(\widehat{\theta} - \mathbb{E}(\widehat{\theta})\right)^2 + \left(\mathbb{E}(\widehat{\theta}) - \theta\right)^2 \\ &= \text{Var}(\widehat{\theta}) + B(\widehat{\theta})^2\end{aligned}$$

Se $\lim_{n\to\infty} EQM(\widehat{\theta}) = 0$ então $\widehat{\theta}$ é dito consistente.


(a) Vício grande e Variancia grande
(b) Vício pequeno e Variancia grande
(c) Vício grande e Variancia pequena
(d) Vício pequeno e Variancia pequena