Propriedades dos Estimadores de Máxima Verossimilhança

ESTAT0078 – Inferência I

Prof. Dr. Sadraque E. F. Lucena

sadraquelucena@academico.ufs.br

http://sadraquelucena.github.io/inferencia1

UNIVERSIDADE FEDERAL DE SERGIPE

Introdução

Por que o Método de Máxima Verossimilhança (MV) é considerado o método padrão na inferência estatística moderna?

Diferente do Método dos Momentos, o MV possui propriedades teóricas ótimas para **grandes amostras**:

1. **Invariância:** Facilidade para estimar funções do parâmetro.
2. **Eficiência:** Garante a menor variância possível (assintoticamente).
3. **Normalidade:** Permite a construção fácil de intervalos de confiança.

Teorema 14.1: Princípio da Invariância

- Este teorema resolve um problema prático: muitas vezes não queremos estimar θ, mas uma função dele.
- Seja $\widehat{\theta}$ o estimador de máxima verossimilhança de θ. Se $g(\theta)$ é uma função bijetora (ou, sob certas condições, qualquer transformação), então o estimador de máxima verossimilhança de $g(\theta)$ é dado por:

$$\widehat{g(\theta)} = g(\widehat{\theta}).$$

- *Diferença Crucial:* O Método dos Momentos **não** garante isso (ex: $E[\overline{X}^2] \neq \mu^2$). Já o Método da Máxima Verossimilhança garante.

Exemplo 14.1

Seja $X \sim \text{Bernoulli}(p)$. O EMV é $\widehat{p} = \overline{X}$. Qual o EMV da variância populacional $g(p) = p(1-p)$?

Exemplo 14.2

Seja $X \sim \text{Exp}(\theta)$ com média θ. A densidade é $f(x) = \frac{1}{\theta} e^{-x/\theta}$. O EMV é $\widehat{\theta} = \overline{X}$. Um atuário precisa estimar a probabilidade de um sinistro ocorrer após o tempo $t = 1$:

$$S(1) = P(X > 1) = \int_1^\infty \frac{1}{\theta} e^{-x/\theta} dx = e^{-1/\theta}$$

Qual o EMV de S(1)?

Precisão do Estimador

- Para falarmos de eficiência (variância), precisamos quantificar quanta informação os dados carregam sobre θ.
- Definimos a **Informação de Fisher Esperada (unitária)**:

$$I_1(\theta) = -E\left[\frac{\partial^2 \ln f(X|\theta)}{\partial\theta^2}\right]$$

- Ela representa a "curvatura" média da log-verossimilhança.
  - Muita informação $\rightarrow$ Variância Baixa.
  - Pouca informação $\rightarrow$ Variância Alta.
- Para uma amostra aleatória de tamanho n, a informação total é $I_n(\theta) = nI_1(\theta)$.

Distribuição em Grandes Amostras

- Sob condições de regularidade, o EMV tem comportamento **Assintoticamente Normal**.
- Quando $n \to \infty$:

$$\sqrt{n}(\widehat{\theta} - \theta) \xrightarrow{d} N\left(0, \frac{1}{I_1(\theta)}\right)$$

- Ou, em termos de aproximação para amostras finitas:

$$\widehat{\theta} \approx N\left(\theta, \frac{1}{nI_1(\theta)}\right)$$

Distribuição em Grandes Amostras

Eficiência Assintótica

A variância $\frac{1}{nI_1(\theta)}$ é o **Limite Inferior de Cramér-Rao**. Isso significa que o EMV é o estimador mais preciso possível (entre os não-viesados) para grandes amostras.

O Método Delta

- Como obter a distribuição assintótica de uma função $g(\widehat{\theta})$?
- Usamos a expansão de Taylor de primeira ordem, técnica conhecida como **Método Delta**.
- Se $\sqrt{n}(\widehat{\theta} - \theta) \to N(0, \sigma^2)$, então:

$$\sqrt{n}(g(\widehat{\theta}) - g(\theta)) \xrightarrow{d} N\left(0, [g'(\theta)]^2 \cdot \sigma^2\right)$$

- Aplicando ao caso de MV:

$$g(\widehat{\theta}) \approx N\left(g(\theta), \frac{[g'(\theta)]^2}{nI_1(\theta)}\right)$$

Exemplo 14.4

Seja $X \sim \text{Bernoulli}(p)$. EMV $\hat{p} = \overline{X}$.

a. Determine a distribuição assintótica do EMV de p.

b. Determine a distribuição assintótica do EMV de $p(1-p)$.

Exemplo 14.5

Seja $X \sim \text{Poisson}(\theta)$. EMV $\widehat{\theta} = \overline{X}$.

a. Determine a distribuição assintótica do EMV de θ.

b. Determine a distribuição assintótica do EMV de $g(\theta) = P(X = 0) = e^{-\theta}$ (Probabilidade de zero ocorrências).

Fim