O Método da Máxima Verossimilhança

ESTAT0078 – Inferência I

Prof. Dr. Sadraque E. F. Lucena

sadraquelucena@academico.ufs.br

http://sadraquelucena.github.io/inferencia1

UNIVERSIDADE FEDERAL DE SERGIPE

Motivação

- Na aula passada, utilizamos a ideia de que momentos amostrais aproximam momentos populacionais para construir estimadores.
- Agora adotaremos uma abordagem diferente, baseada diretamente no modelo probabilístico: **O Método da Máxima Verossimilhança (MV)**.
- O foco passa a ser o seguinte: entre todos os valores possíveis de θ, qual explica melhor os dados que observamos?

Construção da Ideia de Verossimilhança

- Suponha que temos um conjunto de dados observado:

$$\mathbf{x} = (x_1, \ldots, x_n),$$

assumido como proveniente de ums distribuição $f(\mathbf{x}|\theta)$.

- **Pergunta central:** Entre todos os valores possíveis de θ, qual deles torna mais plausível que exatamente esses dados tenham sido observados?

$$\text{"Qual valor de } \theta \text{ explica melhor os dados observados?"}$$

Construção da Ideia de Verossimilhança

- Se o modelo é $f(x|\theta)$, então para cada θ podemos calcular quão provável é observar cada x_i.
- Para uma amostra independente:

$$\text{plausibilidade de } \mathbf{x} \text{ sob } \theta = f(x_1|\theta) \cdots f(x_n|\theta)$$

- Chamamos isso de função de verossimilhança.

Definição 13.1: Função de Verossimilhança

- Seja $X_1, \ldots, X_n$ uma amostra aleatória com densidade ou função de probabilidade $f(x|\theta)$.
- A **Função de Verossimilhança** é definida por:

$$L(\theta; \mathbf{x}) = \prod_{i=1}^{n} f(x_i|\theta).$$

- Ela mede quão compatível com os dados é cada valor de θ.
- O objetivo da Máxima Verossimilhança é escolher o valor de θ que maximiza essa compatibilidade.

Distinção Importante!

Embora a forma seja idêntica à da densidade conjunta, a interpretação é diferente:

- Em $f(\mathbf{x}|\theta)$, θ é fixo e os dados variam. (Integral = 1)
- Em $L(\theta; \mathbf{x})$, os dados são fixos e θ varia. (Não é densidade; integral $\neq$ 1)

O Método de Máxima Verossimilhança

- O Estimador de Máxima Verossimilhança (EMV) é o valor de θ que maximiza a verossimilhança:

$$\widehat{\theta}_{MV} = \arg\max_{\theta} L(\theta; \mathbf{x})$$

- Graficamente, buscamos o pico da curva $L(\theta)$.
- Intuitivamente: escolhemos o valor de θ que melhor reconstrói os dados que foram observados.

Log-Verossimilhança

- Para maximizar $L(\theta)$, precisamos trabalhar com o produtório

$$L(\theta) = \prod_{i=1}^{n} f(x_i|\theta),$$

  que geralmente é difícil de manipular e diferenciar.

- Para contornar isso, podemos usar uma propriedade útil:
  - A função logaritmo é **estritamente crescente**, portanto:

$$\arg\max_{\theta} L(\theta) = \arg\max_{\theta} \log L(\theta).$$

  - Ou seja, **o valor de θ que maximiza L é o mesmo que maximiza** $\log L$.

Definição 13.2: Função de Log-Verossimilhança

- A **Log-Verossimilhança** é dada por:

$$\ell(\theta; \mathbf{x}) = \log L(\theta; \mathbf{x}) = \sum_{i=1}^{n} \log f(x_i|\theta)$$

- Por que isso ajuda?
  - Produto → soma
  - Derivadas ficam mais simples
  - Evita problemas numéricos com produtos de valores pequenos

A Equação de Score

- Para encontrar o EMV, derivamos a log-verossimilhança:

$$\frac{\partial}{\partial\theta}\ell(\theta; \mathbf{x}) = 0.$$

- Esta é chamada de Equação de Score.
  - A solução dessa equação fornece candidatos ao EMV.
  - Para confirmar que é um máximo, verificamos a segunda derivada:

$$\frac{\partial^2}{\partial\theta^2}\ell(\theta; \mathbf{x}) < 0.$$

Algoritmo Geral do Método da Máxima Verossimilhança

Para encontrar o Estimador de Máxima Verossimilhança (EMV ou $\widehat{\theta}_{MV}$):

1. **Monte** a função $L(\theta) = \prod f(x_i)$.
2. **Aplique o log:** $\ell(\theta) = \ln L(\theta)$.
3. **Derive** em relação a θ e iguale a zero (Equação de Score):

$$\frac{\partial \ell(\theta)}{\partial \theta} = 0$$

Algoritmo Geral do Método da Máxima Verossimilhança

4. **Verifique** o sinal da segunda derivada (deve ser negativo):

$$\frac{\partial^2 \ell(\theta)}{\partial \theta^2} < 0$$

Exemplo 13.1

Seja $x_1, \ldots, x_n$ uma amostra aleatória obtida de $X \sim \text{Bernoulli}(p)$. Obtenha o EMV de p.

Exemplo 13.2

Seja $x_1, \ldots, x_n$ uma amostra aleatória obtida de $X \sim \text{Poisson}(\lambda)$. Obtenha o EMV de λ.

Exemplo 13.3

Considere uma amostra $x_1, \ldots, x_n$ com $X \sim \text{Exponencial}(\theta)$, parametrizada pela média $\theta > 0$. A densidade é $f(x) = \frac{1}{\theta} e^{-x/\theta}$, para $x > 0$. Obtenha o EMV de θ.

Algumas Observações

- Se a distribuição pertence à **Família Exponencial** (Normal, Poisson, Binomial, Gama, Exponencial), o EMV quase sempre será baseado na **Média Amostral** ($\overline{X}$).
- No caso da Distribuição Uniforme$(0, \theta)$.
  - A derivada não funciona (a função é constante e cai abruptamente).
  - O máximo ocorre na fronteira dos dados.
  - $\widehat{\theta}_{MV} = \max(X_1, \dots, X_n)$.
  - *Se você derivar, vai encontrar zero ou erro.*

Fim