Inferência Estatística I

Métodos de estimação pontual

Prof. Paulo Cerqueira Jr
Faculdade de Estatística - FAEST
Instituto de Ciências Exatas e Naturais - ICEN

https://github.com/paulocerqueirajr

Métodos de estimação

Introdução

- Vimos até então a propriedades dos estimadores.
- Contudo tais procedimentos não são métodos que possibilitam, em geral, a obtenção de estimadores em situações específicas.
- Vimos também que todo bom estimador deve ser função de uma estatística suficiente.
- Vamos considerar alguns métodos que possibilitam a obtenção de estimadores em situações específicas.
- São os seguintes:

1. Método de máxima verossimilhança;
2. Método dos momentos;
3. Método de mínimos quadrados.

Método da máxima verossimilhança

Definição 1 O estimador de máxima verossimilhança (EMV) de θ é o valor $\hat{\theta} \in \Theta$ que maximiza a função de verossimilhança $L(\theta \mid x)$.

> Tomamos como estimador aquele valor de θ que torna a amostra observada a mais provável de ocorrer.

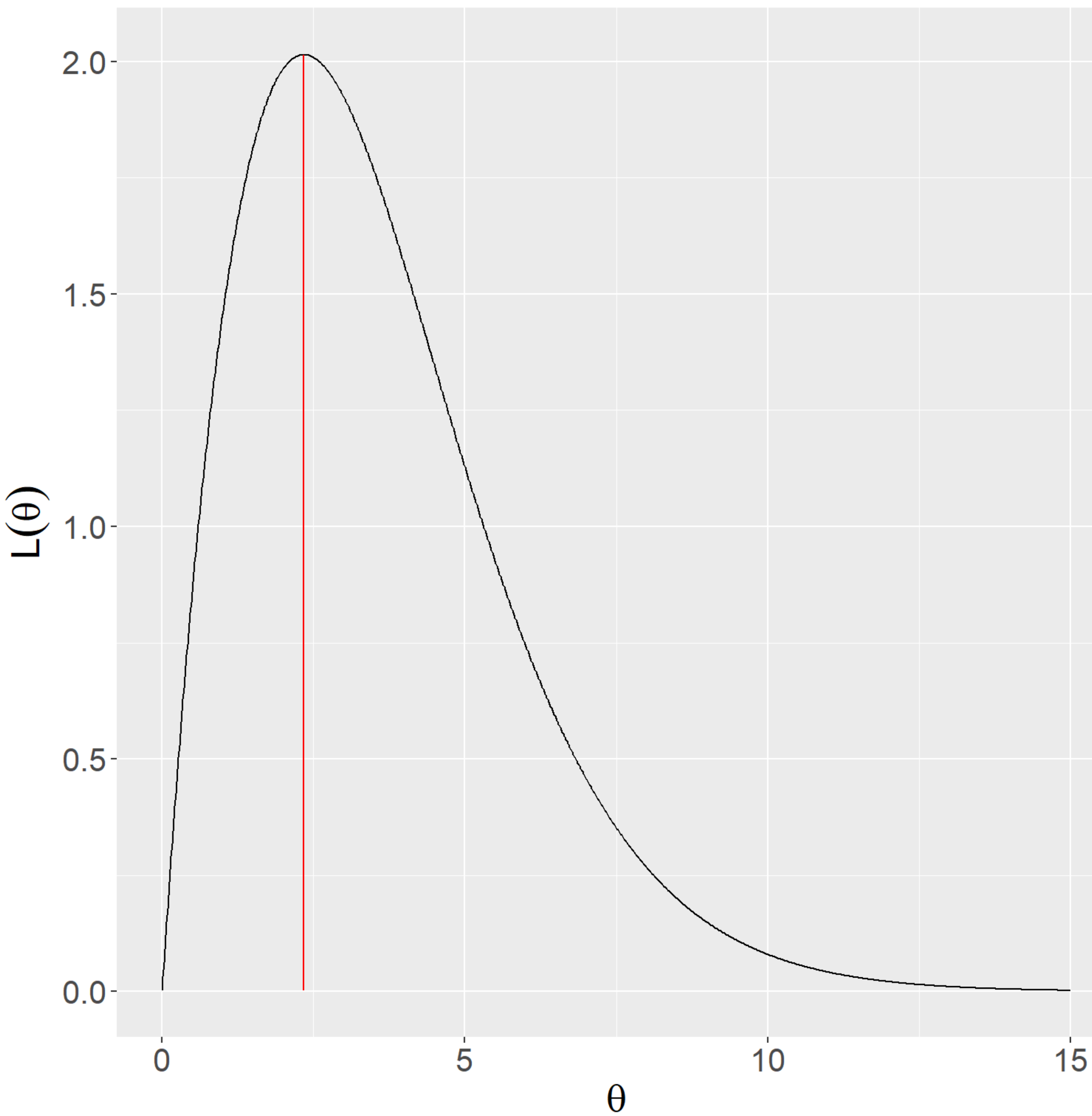
Exemplo:

Sejam $X_1, \ldots, X_{10}$ uma a.a. de $X \sim Pois(\theta)$. Obtenha a estimativa de máxima verossimilhança para θ se os valores amostrais foram:

3, 2, 2, 3, 4, 4, 1, 4, 1, 2

A função verossimilhança:

$$L(\theta) = \prod_{i=1}^{10} \frac{e^{-\theta}\theta^{x_i}}{x_i!}.$$



- O valor que maximiza $L(\theta)$ é 2.34

Propriedades do EMV:

1. Pode-se mostrar que o valor de θ que maximiza a função de verossimilhança $L(\theta \mid x)$, também maximiza $\ell(\theta; |) = \log L(\theta \mid x)$.

2. No caso em que θ é um intervalo da reta e $\ell(\theta \mid x)$ é derivável, o EMV é dado pelo valor de θ que satisfaz a equação de verossimilhança:

$$\ell^{'}(\theta \mid x) = \frac{d}{d\theta}\ell(\theta \mid x) = 0.$$

3. Para verificar se $L(\hat{\theta} \mid x)$ é um ponto de máximo, basta tomar:

$$\ell^{''}(\theta \mid x) = \frac{d}{d\theta}\ell^{'}(\theta \mid x) \mid_{\theta=\hat{\theta}} < 0.$$

4. Em situações mais complicadas, a solução da equação de verossimilhança não pode ser obtida explicitamente, tem-se a necessidade de utilizar procedimentos numéricos.

Exemplo

Exemplo 1 Sejam $X_1, \ldots, X_n$ uma a.a. de $X \sim Pois(\theta)$. Obtenha a estimador de máxima verossimilhança para θ.

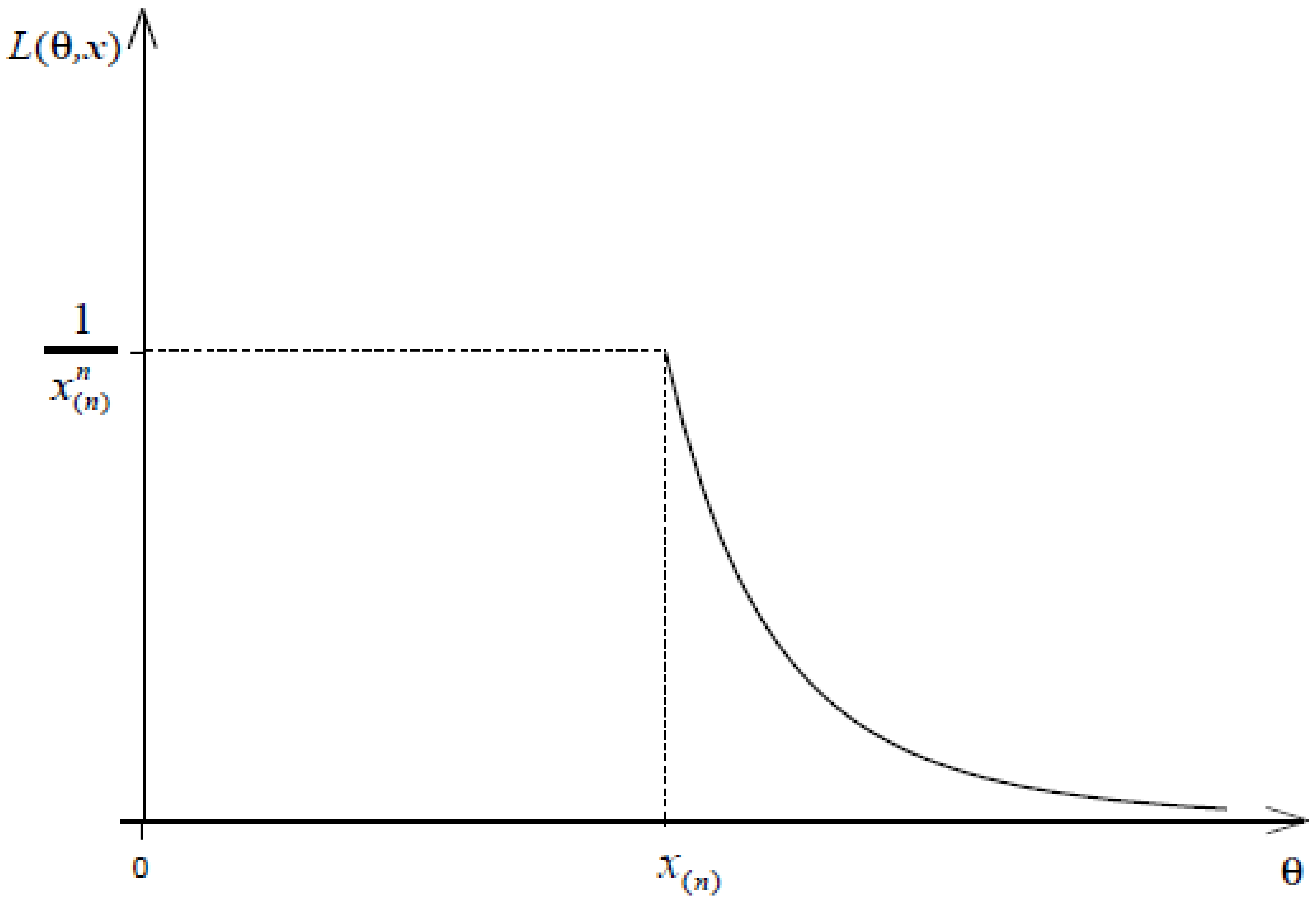
Exemplo

Exemplo 2 Sejam $X_1, \ldots, X_n$ uma a.a. de $X \sim N(\mu, 1)$. Obtenha a estimador de máxima verossimilhança para μ.

Exemplo (suporte depende do parâmetro)

Exemplo 3 Sejam $X_1, \ldots, X_n$ uma a.a. de $X \sim U(0, \theta)$. Obtenha a estimador de máxima verossimilhança para θ.

Exemplo



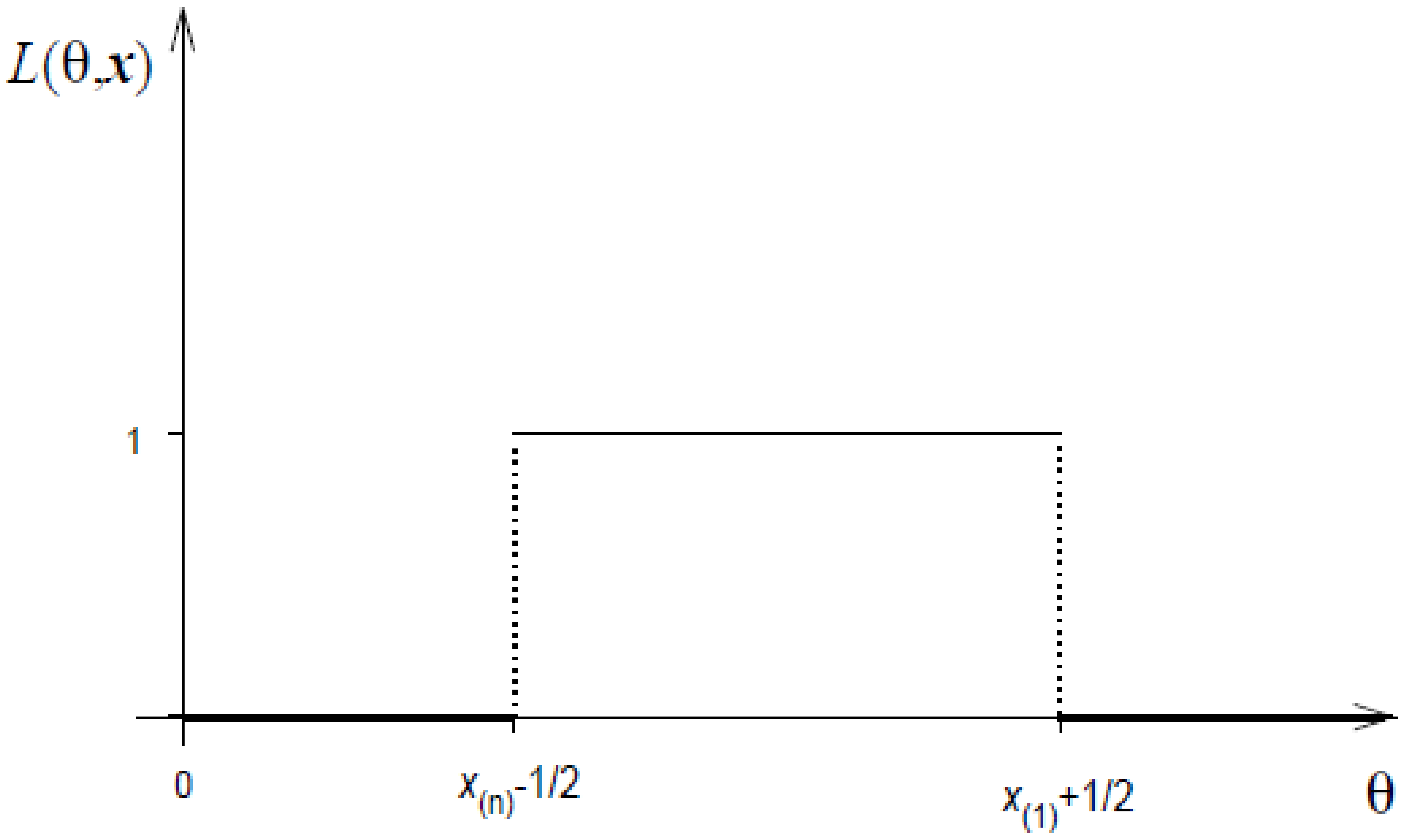
Exemplo (Caso discreto)

Exemplo 4 Temos uma caixa com bolas brancas e vermelhas. Sabe-se que a proporção θ de bolas vermelhas na caixa é $1/3$ ou $2/3$. Portanto, $\Theta = \{1/3, 2/3\}$. Uma amostra de 3 bolas, retiradas com reposição da caixa, apresentou bola vermelha na 1^a extração e branca na 2^a e 3^a extrações. Obtenha uma estimativa de máxima verossimilhança para θ.

Exemplo (suporte depende do parâmetro)

Exemplo 5 Sejam $X_1, \ldots, X_n$ uma a.a. de $X \sim U(\theta - 1/2, \theta + 1/2)$. Obtenha a estimador de máxima verossimilhança para θ.

Exemplo (suporte depende do parâmetro)



EMV via Método do escore

- Em alguns casos, principalmente quando a verossimilhança está associada a modelos mais complexos, a função de verossimilhança não apresenta solução analítica explícita.
- Em tais casos, os estimadores de máxima verossimilhança podem ser obtidos por meio de métodos numéricos.
- Vamos denotar por $U(\theta)$ a função escore, ou seja

$$U(\theta) = \frac{d}{dx} \log L(\theta \mid x),$$

- Determinamos o EMV através:

$$U(\hat{\theta}) = 0.$$

EMV via Método do escore

- Expandindo $U(\hat{\theta})$ em série de Taylor em torno de um ponto θ_0, obtemos

$$0 = U(\hat{\theta}) \approx U(\theta_0) + (\hat{\theta} - \theta_0)U^{'}(\theta_0).$$

- Dessa forma,

$$\hat{\theta} \approx \theta_0 - \frac{U(\theta_0)}{U^{'}(\theta_0)}.$$

- Da equação anterior, temos o processo iterativo de *Newton-Raphson*:

$$\theta_{j+1} \approx \theta_j - \frac{U(\theta_j)}{U^{'}(\theta_j)}$$

que é iniciado com o valor θ_0 e então o valor θ_1 é o obtido pela expressão anterior.

Exemplo

Exemplo 6 Sejam $X_1, \ldots, X_n$ uma a.a. de X, com função densidade dada por

$$f(x \mid \theta) = \frac{1}{2}(1 + \theta x), \quad -1 \leq x \leq 1, \quad -1 \leq \theta \leq 1.$$

Determine o EMV para θ.

Exemplo

- Amostra de tamanho 100.
- Gerado uma amostra com $\theta = 0.4$, usando o método da função de distribuição.

```
# derivada 1:
u.theta <- function(par=par, dat=d
  dv1 <- sum(dat/(1+(par*dat)))
  return(dv1)
}

# derivada 2:
u.theta2 <- function(par=par, dat=
  dv2 <- -1*sum(dat^2/(1+(par*dat)
  return(dv2)
}

set.seed(12345678)

# Dados:
```

```
Iteração= 1   theta inicial= 0.1578034
theta novo= 0.4880783

Iteração= 2   theta inicial= 0.4880783
theta novo= 0.4661251
```

Propriedades do EMV:

Importante:

1. O EMV é sempre função de uma estatística suficiente;
2. O EMV é assintoticamente não viciado;
3. O EMV é invariante a transformações, ou seja, se $\hat{\theta}$ é o EMV de θ, então $g(\hat{\theta})$ é o EMV de $g(\theta)$.

Exemplo 7 Sejam $X_1, \ldots, X_n$ uma a.a. de $X \sim Ber(\theta)$. Obtenha a estimador de máxima verossimilhança para $g(\theta) = \theta(1-\theta)$.

Caso multiparamétrico

- Até então, estamos considerando situações em que a função de verossimilhança depende de somente um parâmetro.
- Quando a verossimilhança apresenta mais de um parâmetro, ou seja, $\theta = (\theta_1, \ldots, \theta_r)$.
- Os EMV's $(\hat{\theta}_1, \ldots, \hat{\theta}_r)$ são obtidos como soluções das equações:

$$\frac{\partial}{\partial \theta_i} \log L(\theta \mid x), \; i = 1, 2, \ldots, r.$$

Exemplo

Exemplo 8 Sejam $X_1, \dots, X_n$ uma a.a. de $X \sim N(\mu, \sigma^2)$, ambos desconhecidos. Obtenha a estimador de máxima verossimilhança para $\theta = (\mu, \sigma^2)$.

Distribuição para grandes amostras

- Sob algumas condições temos que, para amostras grandes, o EMV de θ satisfaz:

[Math Processing Error]

e

[Math Processing Error]

- Assim, os EMV's de θ, $\hat{\theta}$, e $g(\theta)$, $g(\theta)$, são aproximadamente não-viciados e suas variâncias coincidem com os respectivos limites inferiores das variâncias dos estimadores não-viciados de θ e $g(\theta)$.

Exemplo

Exemplo 9 Seja $X_1, \ldots, X_n$ uma a.a. de $X \sim Poisson(\theta)$. Obtenha o EMV de $e^{-\theta}$ e sua distribuição assintótica.

Solução: Temos que $\hat{\theta} = \bar{X}$. Pela propriedade de invariância temos que

$$g(\hat{\theta}) = e^{-\bar{X}}.$$

que é o EMV para e^{θ}. A distribuição assintótica é dada por

[Math Processing Error]

Verossimilhanças para amostras independentes

- Quando temos duas amostras independentes de populações que dependem de um parâmetro θ, $X_1, \ldots, X_n, Y_1, \ldots, Y_m$, a função de verossimilhança pode ser escrita como

$$L(\theta \mid \mathbf{x}, \mathbf{y}) = L(\theta \mid \mathbf{x}) \times L(\theta \mid \mathbf{y})$$

- A função de log verossimilhança é dada por

$$\ell(\theta \mid \mathbf{x}, \mathbf{y}) = \log L(\theta \mid \mathbf{x}, \mathbf{y}) = \log L(\theta \mid \mathbf{x}) \times \log L(\theta \mid \mathbf{y}) = \ell(\theta \mid \mathbf{x}) \times \ell(\theta \mid \mathbf{y})$$

Exemplo

Exemplo 10 Sejam $X_1, \ldots, X_n$ uma a.a. de $X \sim N(\mu, 4)$ e $Y_1, \ldots, Y_m$ uma a.a. de $Y \sim N(\mu; 9)$, duas amostras independentes. Obtenha o EMV de μ.

A verossimilhança conjunta é dada por

$$
\begin{aligned}
L(\mu, \boldsymbol{x}, \boldsymbol{y}) &= L(\mu, \boldsymbol{x}) \times L(\mu, \boldsymbol{y}) \\
&= \left(\frac{1}{2\sqrt{2\pi}}\right)^n e^{-\frac{1}{8}\sum_{i=1}^n (x_i-\mu)^2} \times \left(\frac{1}{3\sqrt{2\pi}}\right)^m e^{-\frac{1}{18}\sum_{i=1}^m (y_i-\mu)^2} \\
&= \left(\frac{1}{2\sqrt{2\pi}}\right)^n \left(\frac{1}{3\sqrt{2\pi}}\right)^m e^{-\sum_{i=1}^n \frac{(x_i-\mu)^2}{8} - \sum_{i=1}^m \frac{(y_i-\mu)^2}{18}}
\end{aligned}
$$

Logo, o logaritmo da verossimilhança é dado por

$$
l(\mu, \boldsymbol{x}, \boldsymbol{y}) = n \log\left(\frac{1}{2\sqrt{2\pi}}\right) + m \log\left(\frac{1}{3\sqrt{2\pi}}\right) - \sum_{i=1}^n \frac{(x_i-\mu)^2}{8} - \sum_{i=1}^m \frac{(y_i-\mu)^2}{18}
$$

Exemplo

Tomando a derivada com relação a μ, temos:

$$\begin{aligned}
\frac{\partial l(\mu, \boldsymbol{x}, \boldsymbol{y})}{\partial \mu} &= -\sum_{i=1}^{n} \frac{2(x_i - \mu)(-1)}{8} - \sum_{i=1}^{m} \frac{2(y_i - \mu)(-1)}{18} \\
&= \sum_{i=1}^{n} \frac{(x_i - \mu)}{4} + \sum_{i=1}^{m} \frac{(y_i - \mu)}{9}
\end{aligned}$$

Igualando a zero, temos:

$$\sum_{i=1}^{n} \frac{(x_i - \hat{\mu})}{4} + \sum_{i=1}^{m} \frac{(y_i - \hat{\mu})}{9} = 0$$

$$\frac{1}{4}\sum_{i=1}^{n} X_i - \frac{n}{4}\,\hat{\mu} + \frac{1}{9}\sum_{i=1}^{m} Y_i - \frac{m}{9}\,\hat{\mu} = 0$$

$$\hat{\mu}\left(\frac{n}{4} + \frac{m}{9}\right) = \frac{1}{4}\sum_{i=1}^{n} X_i + \frac{1}{9}\sum_{i=1}^{m} Y_i$$

$$\hat{\mu} = \frac{\frac{1}{4}\sum_{i=1}^{n} X_i + \frac{1}{9}\sum_{i=1}^{m} Y_i}{\frac{n}{4} + \frac{m}{9}}$$

Método de momentos

- Métodos de estimação mais simples e antigo.
- Seja,

$$m_r = \frac{1}{n}\sum_{i=1}^{n} X_i^r,\ r \geq 1,$$

o $r-$ésimo momento amostral de uma a.a. $X_1, X_2, \ldots, X_n$.

- Seja,

$$\mu_r = E(X^r),\ r \geq 1,$$

o $r-$ésimo momento populacional.

Método de momentos

- O método dos momentos consiste na obtenção de estimadores para $\theta = (\theta_1, \theta_2, \ldots, \theta_k)$ `resolvendo o sistema de equações`, ou seja,

[Math Processing Error]

Exemplo

Exemplo 11 Suponha que o tempo de vida de um determinado aparelho eletrônico é uma v.a. com distribuição exponencial de parâmetro θ. Afim de estimar o valor de θ, uma amostra de 20 aparelhos foi submetida a testes e registrou-se seus tempos de vida (em anos):

$$2; 4; 4; 6; 4; 5; 3,5; 4,5; 3,5; 3; 5; 5; 4,5; 2; 5; 3,5; 3; 3,5; 4,5 \text{ e } 4.$$

Obtenha uma estimativa para θ pelo método dos momentos.

`Sol.` Temos que $(X_1, \ldots, X_n)$ é uma a.a. com $X \sim \mathrm{Exp}(\theta)$. Pelo método dos momentos, temos de resolver a equação:

$$\mu_1 = E(X) = \frac{1}{n}\sum_{i=1}^{n} X_i = m_1$$

sabemos que se $X \sim Exp(\theta)$, então $E(X) = 1/\theta$, logo,

$$1/\theta = \frac{1}{n}\sum_{i=1}^{n} X_i = \bar{X} \Rightarrow \hat{\theta} = \frac{1}{\bar{X}}.$$

O método de mínimos quadrados

- Quando dispomos uma uma relação de proporcionalidade entre duas variáveis X e Y usando a seguinte equação,

$$Y \approx \theta X$$

em que θ é o coeficiente de proporcionalidade.

- Uma estratégia para se obter estimativas para o parâmetro θ se dá através do método de mínimos quadrados.

Método de mínimos quadrados

- Assim, devemos procurar a estimativa que torne esta soma de quadrados mínima. Ou seja, queremos encontrar o valor de θ que minimiza a função

$$S(\theta) = \frac{d}{d\theta} \sum (Y - \theta X)^2 = \sum 2(Y - \hat{\theta} X)(-X) = 0$$

$$\sum(-2XY + \hat{\theta} 2X^2) = 0 \Rightarrow -2 \sum XY + 2\hat{\theta} \sum X^2 = 0$$

$$\hat{\theta} \sum X^2 = \sum XY \Rightarrow \hat{\theta} = \frac{\sum XY}{\sum X^2}.$$

Temos que $\hat{\theta}$ é o estimador pelo método de momentos.

Exemplo

Exemplo 12 Um engenheiro está estudando a resistência Y de uma fibra em função de seu diâmetro X, e descobriu que as variáveis são aproximadamente proporcionais, isto é, elas obedecem à relação

$$Y \approx \theta X.$$

Uma amostra de 5 fibras foi obtida e submetida a testes, que forneceram os resultados:

[Math Processing Error]

Exemplo

Usando os dados temos que

$$\sum XY = 51.05 \text{ e } \sum X^2 = 17.34$$

Assim, a estimativa do parâmetro de proporcionalidade pelo método de mínimos quadrados é dada por

$$\hat{\theta} \sum X^2 = \sum XY \Rightarrow \hat{\theta} = \frac{\sum XY}{\sum X^2} = \frac{51.05}{17.34} = 2.94.$$