Estatística Matemática

Estimação Pontual - Aula 1

Prof. Paulo Cerqueira Jr - cerqueirajr@ufpa.br
Faculdade de Estatística - FAEST
Programa de Pós-Graduação em Matemática e Estatística - PPGME
Instituto de Ciências Exatas e Naturais - ICEN

https://github.com/paulocerqueirajr

Introdução

Introdução

- Assuma que os valores $x_1, x_2, \ldots, x_n$ de uma amostra aleatória $X_1, X_2, \ldots, X_n$ com f.d ou f.p $f(\cdot;\theta)$ podem ser observados (a p.d.f. f é conhecida, Θ é desconhecido).
- Tomando como base as observações $x_1, x_2, \ldots, x_n$ deseja-se estimar o valor do parâmetro desconhecido Θ ou o valor de alguma função $\tau(\Theta)$ do parâmetro desconhecido. Esta estimação pode ser feita de duas maneiras:
  - **Estimação pontual**: O valor de alguma estatística, ex. $l(x_1, \ldots, x_n)$, representará o desconhecido $\tau(\Theta)$. Esta estatística é chamada de `estimador pontual`.
  - **Estimação intervalar**: Definimos duas estatísticas

$$l_1(x_1, \ldots, x_n) < l_2(x_1, \ldots, x_n)$$

tal que

$$[l_1(x_1, \ldots, x_n), l_2(x_1, \ldots, x_n)]$$

constitui um intervalo para o qual a probabilidade de conter $\tau(\Theta)$ pode ser determinada.

Métodos para encontrar estimadores pontuais

Métodos para encontrar estimadores pontuais

Um **estimador** é uma estatística que pode ser encarada tanto como:

- Variável aleatória;
- Função de valores observados.

Considere que $X_1, \ldots, X_n$ é uma amostra aleatória com p.d.f. $f(\cdot; \theta)$ e queremos estimar $T(\Theta)$, sendo: - $T(\cdot)$ é alguma função de θ - $t(X_1, \ldots, X_n)$ é um estimador de $T(\theta)$

> Representações do estimador:

1. Como variável aleatória:

$$T = t(X_1, \ldots, X_n)$$

2. Como função de valores observados:

$$t = t(x_1, \ldots, x_n)$$

`Obs:` O valor numérico obtido é chamado de **estimativa**.

Estimador para a média populacional (μ)

Como variável aleatória:

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

Como estimativa baseada em dados observados:

$$\bar{x}_n = \frac{1}{n}\sum_{i=1}^{n} x_i$$

Neste caso, é um valor numérico representando uma estimativa para μ baseada na amostra $x_1, \ldots, x_n$.

- Sobre a estimação de parâmetros:
  - **Estimação de θ**: Busca determinar o valor desconhecido do parâmetro θ.
  - **Estimação de $\tau(\theta)$**: Busca determinar o valor que a função conhecida $\tau(\cdot)$ assume para o parâmetro desconhecido θ.

Observações importantes

1. **Candidatos naturais**:
    - A média amostral é um candidato natural para estimar a média populacional.
2. **Limitações da intuição**:
    - Em casos mais complexos, a intuição pode falhar;
    - Técnicas formais são necessárias para sugerir estimadores adequados.
3. **Aviso sobre qualidade**:
    - As técnicas de estimação não garantem automaticamente bons estimadores;
    - Os estimadores obtidos devem sempre ser avaliados quanto à sua **qualidade**.

> **Nota**: Todo estimador pontual deve ser cuidadosamente avaliado antes de ser considerado adequado para inferência estatística.

Método dos Momentos

Método dos Momentos

- Este método é, talvez, o mais velho método para encontrar estimadores pontuais. Tem a virtude de ser bastante simples de usar e quase sempre fornece algum tipo de estimativa.
- Em muitos casos, infelizmente, este método fornece estimadores que podem ser melhorados. Entretanto, ele é uma opção interessante quando outros métodos se mostram problemáticos.
- Os estimadores do Método de Momentos são encontrados quando igualamos os primeiros k momentos amostrais aos correspondentes k momentos populacionais, e resolvemos os sistema resultante de equações simultâneas.
- Seja $X_1, X_2, \ldots, X_n$ uma amostra aleatória de uma população com p.m.f. ou p.d.f. $f(x|\theta_1, \theta_2, \ldots, \theta_k)$.

Método dos Momentos

- Seja M'_j o j-ésimo momento amostral.

$$M'_j = \frac{1}{n}\sum_{i=1}^{n} X_i^j$$

- Denote μ'_j como o j-ésimo momento populacional em relação à zero. $\mu'_j = E(X^j)$.
- O momento populacional μ'_j será tipicamente uma função de $\theta_1, \ldots, \theta_k$; denote $\mu'_j(\theta_1, \ldots, \theta_k)$.
- Sistema de equações a ser resolvido:

$$\begin{gathered} M'_1 = \mu'_1(\theta_1, \ldots, \theta_k) \\ M'_2 = \mu'_2(\theta_1, \ldots, \theta_k) \\ \vdots \\ M'_k = \mu'_k(\theta_1, \ldots, \theta_k) \end{gathered}$$

Método dos Momentos - Exemplos

Exemplo 1 Suponha $X_1, \ldots, X_n$ i.i.d. $N(\theta, \sigma^2)$. Determine o estimador pelo método de momentos.

`Solução:` Na notação anterior, temos $\theta_1 = \theta$ e $\theta_2 = \sigma^2$:

- Momentos amostrais:

$$M_1' = \frac{1}{n}\sum_{i=1}^{n} X_i^1 = \overline{X}$$

$$M_2' = \frac{1}{n}\sum_{i=1}^{n} X_i^2$$

- Momentos populacionais:

$$\mu_1' = E(X^1) = \theta$$

$$\begin{aligned} \mu_2' = E(X^2) &= \mathrm{Var}(X) + E^2(X) \\ &= \sigma^2 + \theta^2 \end{aligned}$$

Equações:

$$\begin{cases} \hat{\theta} & = & \overline{X} \\ \hat{\sigma}^2 + \hat{\theta}^2 & = & \frac{1}{n}\sum_{i=1}^{n} X_i^2 \end{cases}$$

Método dos Momentos - Exemplos

- Continuação, substituindo $\hat{\theta} = \bar{X}$, temos

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

$$\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n} X_i^2 - \overline{X}^2 = \frac{1}{n}\sum_{i=1}^{n} X_i^2 - 2\overline{X}^2 + \overline{X}^2$$

$$= \frac{1}{n}\sum_{i=1}^{n} X_i^2 - 2\overline{X}\frac{1}{n}\sum_{i=1}^{n} X_i + \overline{X}^2$$

$$= \frac{1}{n}\sum_{i=1}^{n}[X_i^2 - 2\overline{X}X_i + \overline{X}^2] = \frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X})^2$$

- Neste exemplo, a solução do método de momentos coincide com nossa intuição.
- O método será mais útil, entretanto, quando nenhum estimador óbvio existir.

Método dos Momentos - Exemplos

Exemplo 2 Seja $X_1, \ldots, X_n$ i.i.d. Binomial (k, p). Determine o estimador pelo método de momentos.

$$P(X_i = x|k,p) = \binom{k}{x} p^x (1-p)^{k-x}, \quad x = 0, 1, \ldots, k$$

Assumimos que ambos k e p são desconhecidos e nós desejamos estimadores para ambos os parâmetros.

- Momentos amostrais:

$$M_1' = \frac{1}{n}\sum_{i=1}^{n} X_i^1 = \overline{X}$$

$$M_2' = \frac{1}{n}\sum_{i=1}^{n} X_i^2$$

- Momentos populacionais:

$$\mu_1' = E(X^1) = kp$$

$$\mu_2' = E(X^2) = \mathrm{Var}(X) + E^2(X)$$

$$= kp(1-p) + k^2p^2$$

Método dos Momentos - Exemplos

Equações :

$$\begin{cases} \bar{X} = kp \\ \dfrac{1}{n}\displaystyle\sum_{i=1}^{n} X_i^2 = kp(1-p) + (kp)^2 \end{cases}$$

Substituindo $kp = \bar{X}$, temos

$$\frac{1}{n}\sum_{i=1}^{n} X_i^2 = \bar{X}(1-p) + \bar{X}^2$$

$$\frac{1}{n}\sum_{i=1}^{n} X_i^2 = \bar{X} - p\bar{X} + \bar{X}^2$$

$$\frac{1}{n}\left(\sum_{i=1}^{n} X_i^2\right) - \bar{X}^2 = \bar{X} - p\bar{X}$$

$$\frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2 = \bar{X} - \bar{X}p$$

$$\bar{X}p = \bar{X} - \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2$$

$$\hat{p} = \frac{\bar{X} - \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2}{\bar{X}}$$

Método dos Momentos - Exemplos

Logo,

$$kp = \bar{X}$$

$$\hat{k} = \frac{\bar{X}}{\hat{p}} = \frac{\bar{X}}{\frac{\bar{X} - \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2}{\bar{X}}} = \frac{\bar{X}^2}{\bar{X} - \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2}$$

- Estes não são os melhores estimadores para os parâmetros populacionais k e p.
- Veja que é possível obter **estimativas negativas** para estes parâmetros que são números positivos. Isso ocorre quando a média amostral é menor que a variância (**alto grau de variabilidade nos dados**).
- O método dos momentos, neste caso, pelo menos nos deu um conjunto de candidatos a estimadores pontuais para k e p.

Método dos Momentos - Exemplos

Exemplo 3 Seja $X_1, \ldots, X_n$ uma amostra aleatória da Poisson(λ). Determine o estimador pelo método de momentos.

$$M_1' = \frac{1}{n}\sum_{x=1}^{n} X_i = E(X) = \lambda$$

$$\hat{\lambda} = \overline{X}$$

- Então o estimador de λ, obtido pelo método de momentos, é $M_1' = \overline{X}$, ou seja, estimamos a média populacional λ com a média amostral $\overline{X}$.

Método da máxima verossimilhança

Método da máxima verossimilhança

- Para introduzir este método, considere o seguinte problema:
- Suponha que uma urna contenha bolas brancas e pretas. Sabemos também que a razão entre o nº de bolas brancas e pretas é $3/1$, entretanto, não se sabe qual das duas cores é mais numerosa dentro da urna. Isto é, a probabilidade de sortear uma bola preta pode ser tanto $1/4$ quanto $3/4$.
- Se n bolas são sorteadas (com reposição) desta urna, a distribuição de X (variável n° de bolas pretas) é a Binomial:

$$f(x|p) = \binom{n}{x} p^x (1-p)^{n-x}, \quad \text{para } x = 0, 1, 2, \ldots, n.$$

sendo p é a probabilidade de retirar uma bola preta. Temos que $p = \frac{1}{4}$ ou $p = \frac{3}{4}$.

Método da máxima verossimilhança

`Estratégia:` Sortear três bolas da urna $(n = 3)$, com reposição, e tentar estimar o parâmetro p.

- O problema de estimação é particularmente simples, neste caso, visto que devemos escolher entre dois números 0.25 e 0.75. Desta forma podemos avaliar todos os possíveis resultados desta amostragem:

Resultado: $x = n^o$ **de bolas pretas**	**0**	**1**	**2**	**3**
$f(x \mid p = 3/4)$	$1/64$	$9/64$	$27/64$	$27/64$
$f(x \mid p = 1/4)$	$27/64$	$27/64$	$9/64$	$1/64$

Método da máxima verossimilhança

- Neste exemplo, se encontraremos $x = 0$ em uma amostra $n = 3$, a estimativa $p = 0.25$ será preferida com relação a $p = 0.75$ visto que

$$f(x = 0|p = 0.25) = \frac{27}{64} > f(x = 0|p = 0.75) = \frac{1}{64}$$

- Isto é, a amostra $x = 0$ é mais provável (no sentido de ter maior probabilidade) de surgir em uma população com $p = \frac{1}{4}$ do que em uma população com $p = \frac{3}{4}$.
- Usando o raciocínio acima:
  - Devemos estimar $p = 0.25$, quando $x = 0$ ou 1.
  - Devemos estimar $p = 0.75$, quando $x = 2$ ou 3.
- O estimador pode ser definido como:

$$\hat{p} = \hat{p}(x) = \begin{cases} 0.25 & \text{para } x = 0 \text{ ou } 1. \\ 0.75 & \text{para } x = 2 \text{ ou } 3. \end{cases}$$

Método da máxima verossimilhança

- O estimador seleciona para cada possível resultado x, o valor de $p = \hat{p}$ tal que $f(x|\hat{p}) > f(x|p^*)$ sendo p^* o valor alternativo de p.
- De maneira geral, se diversos valores alternativos de p forem possíveis, podemos proceder da seguinte forma:
- Se encontraremos $x = 6$ em uma amostra com $n = 25$ de uma população Binomial, devemos substituir todos os possíveis valores de p na expressão

$$f(x = 6|p) = \binom{25}{6} p^6 (1-p)^{19}, \quad \text{para } 0 \leq p \leq 1$$

e escolher, como nossa estimativa, o valor de p que maximiza $f(x = 6|p)$.

- O valor de p, correspondente ao máximo, pode ser encontrado quando derivamos a função acima com respeito a p, e igualamos o resultado a zero. Ver a seguir…

Método da máxima verossimilhança

$$\frac{d}{dp} f(x=6|p) = \binom{25}{6} \left[6(p)^5(1-p)^{12} + p^6 19(1-p)^{18}(-1)\right]$$

$$= \binom{25}{6} p^5(1-p)^{18}\left[6(1-p) - 19p\right]$$

- Igualando a zero

$$\binom{25}{6} p^5(1-p)^{18}\left[6 - 25p\right] = 0$$

$$p^5(1-p)^{18}\left[6 - 25p\right] = 0$$

Método da máxima verossimilhança

- Antes de definir os estimadores de máxima verossimilhança, considere a seguinte revisão:
- A função de verossimilhança de n variáveis aleatórias $X_1, X_2, \ldots, X_n$ é definida pela p.m.f. ou p.d.f. conjunta $f(x_1, \ldots, x_n|\theta)$, que é considerada uma função de θ.
- Em particular, se $X_1, X_2, \ldots, X_n$ é uma amostra aleatória com f.d. ou f.p. $f(x|\theta)$, então temos:

$$L(\theta; X_1, \ldots, X_n) = \prod_{i=1}^{n} f(X_i|\theta).$$

- Notação $L(\theta; X_1, \ldots, X_n)$ é usada para nos lembrar que a verossimilhança é encarada como uma função de θ.

Método da máxima verossimilhança

- Suponha que os valores observados na amostra sejam $x'_1, x'_2, \cdots, x'_n$.
  Queremos saber de qual f.d. ou f.p. este conjunto de valores é proveniente com maior chance. Em outras palavras:
  - Queremos saber de qual f.d. ou f.p. (qual valor de θ) a verossimilhança apresentará seu maior valor tendo em vista que observamos $x'_1, x'_2, \cdots, x'_n$.
  - Desejamos encontrar o valor de $\theta \in \Theta$, denotado por $\hat{\Theta}$, que irá maximizar a função de verossimilhança $L(\theta; x'_1, \cdots, x'_n)$.
- $\hat{\theta}$ é em geral uma função da amostra aleatória $X_1, \cdots, X_n$.
- $\hat{\theta} = h(X_1, \cdots, X_n)$.

> $\hat{\theta}$ é chamado de Estimador de Máxima Verossimilhança (EMV) de θ.

- A definição formal é dada a seguir…

Método da máxima verossimilhança

Definição 1 (Estimador de Máxima Verossimilhança (EMV).) Seja $L(\theta; x_1, \ldots, x_n)$ a função de verossimilhança associada às variáveis aleatórias $X_1, \ldots, X_n$. Se $\hat{\theta} = h(X_1, \ldots, X_n)$ é o valor de $\theta \in \Theta$ que maximiza $L(\theta; X_1, \ldots, X_n)$, então $\hat{\theta}$ é o EMV de θ.

- Para a amostra observada $x_1, \ldots, x_n$, o número calculado através da formulação $\hat{\theta} = h(x_1, \ldots, x_n)$ será a estimativa de máxima verossimilhança de θ.
- Muitas funções de verossimilhança satisfazem condições de regularidade, tal que o EMV é obtido pela solução da equação: $\frac{dL(\theta)}{d\theta} = 0$.

Método da máxima verossimilhança

- Outro ponto importante é que $L(\theta; X_1, \ldots, X_n)$ e $\ln L(\theta; X_1, \ldots, X_n)$ apresentam ponto de máximo para o mesmo valor de θ.
- Muitas vezes é mais fácil encontrar o máximo usando o logaritmo da verossimilhança.
- Se a função de verossimilhança envolve k parâmetros, isto é:

$$L(\theta_1, \ldots, \theta_k; X_1, \ldots, X_n) = \prod_{i=1}^{n} f(X_i | \theta_1, \ldots, \theta_k),$$

então os EMVs de $\theta_1, \ldots, \theta_k$ serão as variáveis aleatórias

$$\hat{\theta}_1 = h_1(X_1, \ldots, X_n), \quad \hat{\theta}_2 = h_2(X_1, \ldots, X_n), \quad \hat{\theta}_k = h_k(X_1, \ldots, X_n)$$

onde $\hat{\theta}_1, \hat{\theta}_2, \ldots, \hat{\theta}_k$ são os valores em θ que maximizam

$$L(\theta_1, \ldots, \theta_k; X_1, \ldots, X_n).$$

Método da máxima verossimilhança

- Se certas condições de regularidade são satisfeitas, o ponto onde a verossimilhança atinge seu máximo é solução do sistema de equações ao lado.

$$
\begin{gathered}
\frac{\partial}{\partial \theta_1} \ln L(\theta_1, \ldots, \theta_k; X_1, \ldots, X_n) = 0 \\
\frac{\partial}{\partial \theta_2} \ln L(\theta_1, \ldots, \theta_k; X_1, \ldots, X_n) = 0 \\
\vdots \\
\frac{\partial}{\partial \theta_k} \ln L(\theta_1, \ldots, \theta_k; X_1, \ldots, X_n) = 0
\end{gathered}
$$

- Neste caso, também pode ser mais fácil trabalhar com `o log da verossimilhança`.

Método da máxima verossimilhança - Exemplos

Exemplo 4 Suponha que uma amostra aleatória $X_1, X_2, \ldots, X_n$ é obtida a partir da distribuição Bernoulli(p).

$$f(x|p) = p^x(1-p)^{1-x}, \quad 0 \leq p \leq 1$$

- A amostra $x_1, \ldots, x_n$ será uma sequência de 0's e 1's.
- Função de verossimilhança:

$$L(p; x_1, \ldots, x_n) = \prod_{i=1}^{n} p^{x_i}(1-p)^{1-x_i} = p^{\sum_{i=1}^{n} x_i}(1-p)^{n-\sum_{i=1}^{n} x_i}$$

- Para simplificar a notação considere

$$y = \sum_{i=1}^{n} x_i.$$

Método da máxima verossimilhança - Exemplos

- Aplicando o logaritmo:

$$\ln L(p; x_1, \ldots, x_n) = y \ln p + (n - y) \ln(1 - p)$$

- Derivando em relação a p:

$$\frac{d}{dp} \ln L(p; x_1, \ldots, x_n) = \frac{y}{p} - \frac{n - y}{1 - p}$$

- Igualando a zero:

$$\frac{y}{p} - \frac{n - y}{1 - p} = 0$$

Método da máxima verossimilhança - Exemplos

- Resolvendo:

$$y(1-p) = p(n-y)$$

$$y - yp = pn - yp$$

$$y = pn$$

$$\hat{p} = \frac{y}{n} = \frac{1}{n}\sum_{i=1}^{n} x_i = \overline{x}$$

- Neste caso o EMV coincide com o estimador do método dos momentos.

Método da máxima verossimilhança - Exemplos

Exemplo 5 Considere $X_1, \ldots, X_n$ uma amostra aleatória da distribuição $N(\mu, \sigma^2)$.

- Função de verossimilhança:

$$L(\mu, \sigma^2; x_1, \ldots, x_n) = \prod_{i=1}^{n} (2\pi\sigma^2)^{-1/2} \exp\left\{ -\frac{1}{2\sigma^2}(x_i - \mu)^2 \right\}$$

$$= (2\pi\sigma^2)^{-n/2} \exp\left\{ -\frac{1}{2\sigma^2} \sum_{i=1}^{n} (x_i - \mu)^2 \right\} \quad \sigma > 0, \ -\infty < \mu < \infty$$

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

Método da máxima verossimilhança - Exemplos

- Derivando:

$$\frac{\partial \ln L}{\partial \mu} = \frac{1}{\sigma^2} \sum_{i=1}^{n} (x_i - \mu) = \frac{n}{\sigma^2} (\overline{x} - \mu)$$

$$\frac{\partial \ln L}{\partial \sigma^2} = -\frac{n}{2\sigma^2} + \frac{1}{2\sigma^4} \sum_{i=1}^{n} (x_i - \mu)^2$$

Método da máxima verossimilhança - Exemplos

- Igualando a zero os resultados anteriores iremos obter:

> Equação 1:

$$\frac{n}{\sigma^2}(\overline{x}-\mu)=0$$

$$\overline{x}-\mu=0$$

$$\hat{\mu}=\overline{x}$$

Método da máxima verossimilhança - Exemplos

> Equação 2:

$$-\frac{n}{2\sigma^2}+\frac{1}{2\sigma^4}\sum_{i=1}^{n}(x_i-\overline{x})^2=0,\ \text{usando}\ \mu=\overline{x}$$

$$-\frac{n}{2\sigma^2}\left[1-\frac{1}{n\sigma^2}\sum_{i=1}^{n}(x_i-\overline{x})^2\right]=0$$

$$\frac{1}{n\sigma^2}\sum_{i=1}^{n}(x_i-\overline{x})^2=1$$

$$\hat{\sigma}^2=\frac{1}{n}\sum_{i=1}^{n}(x_i-\overline{x})^2.$$

Método da máxima verossimilhança - Exemplos

Exemplo 6 Suponha que a variável aleatória X tenha distribuição Uniforme.

$$f(x|\theta) = \begin{cases} \frac{1}{\theta} & \text{se } \theta - \frac{1}{2} \leq x \leq \theta + \frac{1}{2} \\ 0 & \text{caso contrário} \end{cases}$$

$$f(x|\theta) = \mathbf{1}_{\left[\theta - \frac{1}{2}, \theta + \frac{1}{2}\right]}(x)$$

sendo $-\infty < \theta < \infty$.

- A função de verossimilhança para uma amostra de tamanho n será:

$$L(\theta; x_1, \ldots, x_n) = \prod_{i=1}^{n} f(x_i|\theta) = \prod_{i=1}^{n} \mathbf{1}_{\left[\theta - \frac{1}{2}, \theta + \frac{1}{2}\right]}(x_i)$$

- O produtório resultará em 1 se e somente se todos os x_i pertencerem ao intervalo $[\theta - \frac{1}{2}, \theta + \frac{1}{2}]$.
- Isso será verdade somente se $\theta - 1/2 \leq x_{(1)}$ e $x_{(n)} \leq \theta + 1/2$ onde $x_{(1)} = \min\{x_1, \ldots, x_n\}$ e $x_{(n)} = \max\{x_1, \ldots, x_n\}$.

Método da máxima verossimilhança - Exemplos

- Veja que:

$$\theta - \frac{1}{2} \leq X_{(1)}, \quad \theta \leq X_{(1)} + \frac{1}{2}$$

- Então temos

$$X_{(n)} - \frac{1}{2} \leq \theta \leq X_{(1)} + \frac{1}{2}$$

- A função de verossimilhança será reescrita como segue:

$$L = \prod_{i=1}^{n} I_{[\theta - \frac{1}{2}, \theta + \frac{1}{2}]}(x_i) = I_{[X_{(n)} - \frac{1}{2}, X_{(1)} + \frac{1}{2}]}(\theta)$$

Método da máxima verossimilhança - Exemplos

Exemplo 7 $X_1, \ldots, X_n$ uma amostra aleatória com p.d.f. $f(x|\theta) = \theta x^{-2}, \quad 0 < \theta \leq x < \infty$.

- A função de verossimilhança:

$$L(\theta; X_1, \ldots, X_n) = \prod_{i=1}^{n} \left(\theta x_i^{-2} I_{[\theta,\infty)}(x_i)\right) = \theta^n \prod_{i=1}^{n} x_i^{-2} I_{[\theta,\infty)}(x_i)$$

- Este produtório não resultará em zero, somente se todos os x_i satisfizerem $\theta \leq x_i < \infty$.
- Se $\theta \leq x_{(1)} = \min\{x_1, \ldots, x_n\}$ então todos os x_i são maiores que θ.
- Conclusão: Podemos escrever

$$\prod_{i=1}^{n} x_i^{-2} \mathbf{I}_{[\theta,\infty)}(x_i) = \left(\prod_{i=1}^{n} x_i^{-2}\right) \mathbf{I}_{(-\infty, x_{(1)})}(\theta)$$

Método da máxima verossimilhança - Exemplos

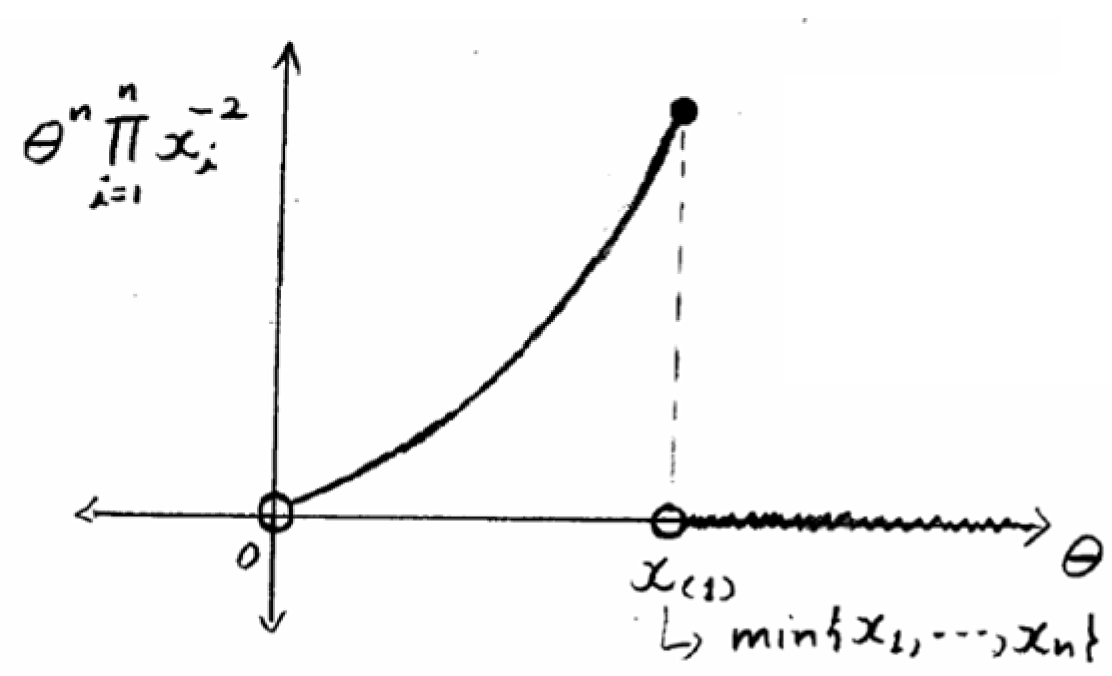
- A verossimilhança será:

$$L = \theta^n \left(\prod_{i=1}^{n} x_i^{-2} \right) I_{[0, x_{(1)}]}(\theta)$$

- Visto que esta expressão envolve uma função indicadora dependendo de θ, `será inapropriado aplicar o log e calcular derivadas.`

Método da máxima verossimilhança - Exemplos

- Veja que θ^n é uma função crescente de θ e $\prod_{i=1}^n x_i^{-2}$ não depende de θ.



- Se $0 < \theta \leq x_{(1)} \Rightarrow L = \theta^n \prod_{i=1}^n x_i^{-2}$
- Se $\theta > x_{(1)} \Rightarrow L = 0$

- Podemos ver que o valor máximo de $L(\theta; X_1, \cdots, X_n)$ é obtido quando $\theta = x_{(1)}$. Portanto $\hat{\theta} = X_{(1)} = \min\{X_1, \cdots, X_n\}$ é o EMV de θ.

Método da máxima verossimilhança - Exemplos

Observação:

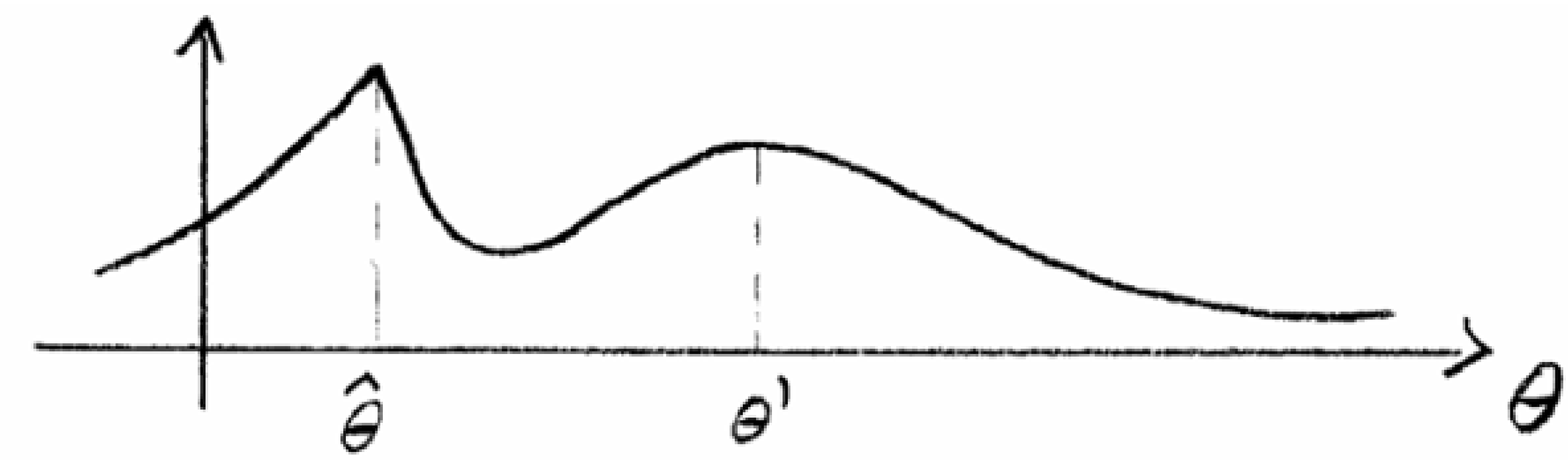
$$E(X|\theta) = \int_{\theta}^{\infty} x\theta x^{-2} dx = \theta \int_{\theta}^{\infty} x^{-1} dx = \theta \ln x \Big|_{\theta}^{\infty} = \infty$$

- O primeiro momento populacional não existe.
- Então, o estimador do método dos momentos para θ `não existe` neste caso.

Método da máxima verossimilhança

Os quatro exemplos que acabamos de estudar ilustram a aplicação do método da máxima verossimilhança.

- Os dois últimos exemplos mostram que não devemos sempre confiar no procedimento de derivação para identificar o ponto de máximo.
- A verossimilhança pode ser representada, por exemplo, pelo gráfico:



- Aqui, o verdadeiro máximo está em $\hat{\theta}$, mas a derivada de L igualada a zero iria localizar θ' como máximo.

Método da máxima verossimilhança

- Devemos lembrar também que a equação $\partial L/\partial\theta = 0$ localiza pontos de mínimo ou de máximo.

> Teste da 2ª derivada:

- Se $\frac{\partial^2 L}{\partial\theta^2} < 0$ então $\hat{\theta}$ é um ponto de máximo.
- Se $\frac{\partial^2 L}{\partial\theta^2} > 0$ então $\hat{\theta}$ é um ponto de mínimo.
- Se $\frac{\partial^2 L}{\partial\theta^2} = 0$ então $\hat{\theta}$ pode ser um ponto de inflexão.

Método da máxima verossimilhança

Teorema 1 (Invariância dos EMVs.) Se $\hat{\theta}$ é o EMV de θ, então para qualquer função $T(\theta)$, o EMV de $T(\theta)$ será $T(\hat{\theta})$.

Prova: [ver Casella e Berger (2002), pag. 320].

Exemplo 8 Para o caso $N(\mu_0, \sigma^2)$ com média μ_0 conhecida, o EMV de

$$\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \mu_0)^2.$$

- Pela propriedade da invariância, o EMV de $\hat{\sigma} = \sqrt{\hat{\sigma}^2}$
- Ou seja,

$$\hat{\sigma} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_i - \mu_0)^2}$$

Método da máxima verossimilhança - Invariância

Exemplo 9

- Para o caso $N(\theta, \sigma^2)$ temos que o EMV de θ é $\overline{X}$.
- Usando o Teorema 1, podemos ver que o EMV de θ^2 será $\overline{X}^2$.

Exemplo 10 Seja $\hat{p}$ o EMV de $p =$ probabilidade de sucesso da Binomial.
Temos que o EMV de $\sqrt{p(1-p)}$ é dado por $\sqrt{\hat{p}(1-\hat{p})}$.

- A propriedade de invariância dos EMVs também é válida para o caso multivariado (θ pode ser um vetor no Teorema 1).
- Se o EMV de $(\theta_1, \ldots, \theta_k)$ é $(\hat{\theta}_1, \ldots, \hat{\theta}_k)$ e se $t(\theta_1, \ldots, \theta_k)$ é qualquer função dos parâmetros, o EMV de $t(\theta_1, \ldots, \theta_k)$ será $t(\hat{\theta}_1, \ldots, \hat{\theta}_k)$.