3. Estimação pontual

USP-ICMC-SME

2013

Roteiro

Formulação do problema.
O problema envolve um fenômeno aleatório.
Interesse em alguma característica da população. Pode ser um parâmetro θ (verdadeiro valor desconhecido) ou uma função deste.
Identificar uma v.a. X cuja distribuição depende de θ, com função densidade dada por $f(x;\theta)$.
Dizemos que a v.a. X fornece informação sobre θ.

Problemas de inferência

Com base em uma amostra aleatória $X_1, X_2, \ldots, X_n$,

1. apresentar um valor para θ (estimação pontual),
2. apresentar um conjunto de possíveis valores para θ, sendo que muitas vezes este conjunto é um intervalo (estimação intervalar) e
3. testar hipóteses sobre θ ($\theta = 0$?, $\theta > 2$?, $\theta \neq 0,5$?, $\ldots$).

Estimador

Um estimador $\widehat{\theta}$ (teta chapéu) é uma estatística T utilizada para obter um valor para θ.
$\widehat{\theta} = T(X_1, \ldots, X_n)$. Com base nos valores observados $x_1, \ldots, x_n$ obtemos uma estimativa $T(x_1, \ldots, x_n)$, também denotada por $\widehat{\theta}$.

Exemplo

Método de substituição.
A média amostral é um estimador da média populacional.
A variância amostral é um estimador da variância populacional.

Propriedades

Um estimador é uma v.a. e tem uma distribuição amostral.

Viés: relacionado ao erro sistemático.

Precisão: proximidade das estimativas entre si. Inversamente relacionada à variância.

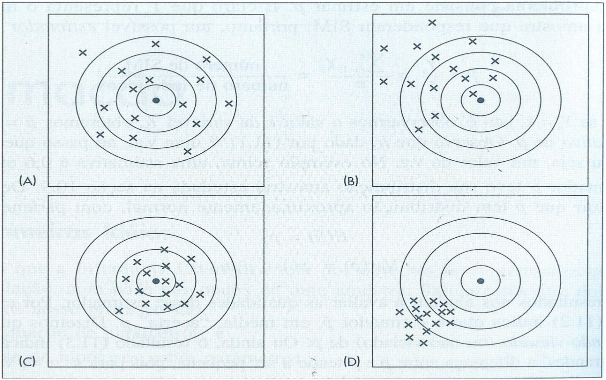
Acurácia: proximidade entre as estimativas e o verdadeiro valor.

Consistência: proximidade entre as estimativas e o verdadeiro valor quando $n \to \infty$.

Existe um método geral que garanta obter estimadores com estas propriedades?

Exemplo

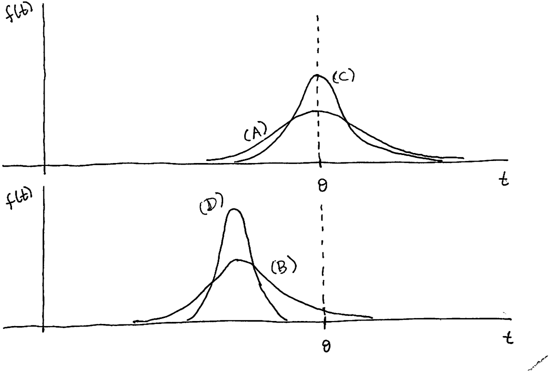
Quatro atiradores



(B) viesado, pouco acurado e pouca precisão. (D) viesado, pouco acurado e alta precisão. (A) sem viés, pouco acurado e baixa precisão. (C) sem viés, acurado e boa precisão.

Exemplo

Quatro estimadores. Represente graficamente quatro estimadores para θ, todos com distribuição normal e que sejam semelhantes ao exemplo dos quatro atiradores.



Propriedades

O estimador $\widehat{\theta}$ é não viesado para θ se

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

para todo θ.
Se o estimador é viesado, a diferença

$$E(\widehat{\theta}) - \theta$$

é chamada de viés do estimador.

Exemplo

Vimos que se $X_1, X_2, \ldots, X_n$ é uma amostra aleatória de uma população com média μ e variância σ^2, então $\overline{X}$ tem média μ, ou seja, a média amostral é um estimador não viesado para a média populacional.

Pode ser provado que a variância amostral $S^2 = \sum_{i=1}^{n}(X_i - \bar{X})^2/(n-1)$ é um estimador não viesado para a variância populacional σ^2.

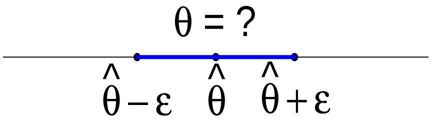
Propriedades

Se $\widehat{\theta}$ é um estimador de θ com base em uma a.a. de tamanho n, dizemos que $\widehat{\theta}$ é consistente para θ se, para todo $\epsilon > 0$,

$$\lim_{n\to\infty} \mathsf{P}\left(|\widehat{\theta} - \theta| \leq \epsilon\right) = 1.$$

$\widehat{\theta} - \theta$ representa o erro de estimação.

Se $\widehat{\theta}$ é consistente para θ, então em termos de probabilidade, $\widehat{\theta}$ pode estar tão próximo de θ quanto queiramos (ϵ pequeno), basta tomarmos n suficientemente grande.



Propriedades

Resultado

$\widehat{\theta}$ é um estimador consistente para θ se

$$\lim_{n\to\infty} E(\widehat{\theta}) = \theta \quad \textit{e} \quad \lim_{n\to\infty} Var(\widehat{\theta}) = 0.$$

Exemplo

Vimos que se $X_1, X_2, \ldots, X_n$ é uma amostra aleatória de uma população com média μ e variância σ^2, então $\overline{X}$ tem média μ e variância σ^2/n, ou seja, $Var(\overline{X}) \to 0$, quando $n \to \infty$. Já vimos que $\overline{X}$ é não viesado, de modo $\lim_{n\to\infty} E(\overline{X}) = \mu$. Logo, $\overline{X}$ é consistente para μ.

$\widehat{\mu} = \overline{X} + 27/n$ é não viesado? É consistente?

Propriedades

$\widehat{\theta}_1$ e $\widehat{\theta}_2$ são dois estimadores não viesados de θ. Se

$$Var(\widehat{\theta}_1) < Var(\widehat{\theta}_2),$$

então $\widehat{\theta}_1$ é mais eficiente do que $\widehat{\theta}_2$.

Método de máxima verossimilhança

Exemplo

$X \sim$ binomial$(16, p)$, sendo que $p \in \{1/2, 6/7\}$. Em uma amostra observamos $x = 13$ sucessos. Com base nesta amostra, apresente uma estimativa para p.

Temos que P$(X = x) = \binom{16}{x} p^x (1-p)^{16-x}$. Observamos $x = 13$ e perguntamos de qual população este valor foi amostrado: $p = 1/2$ ou $p = 6/7$?

Se $p = 1/2$, P$(X = 13; p = 1/2) = \binom{16}{13}(1/2)^{13} \times (1/2)^3 = 0,0085$.

Se $p = 6/7$, P$(X = 13; p = 6/7) = \binom{16}{13}(6/7)^{13} \times (1/7)^3 = 0,2201$.

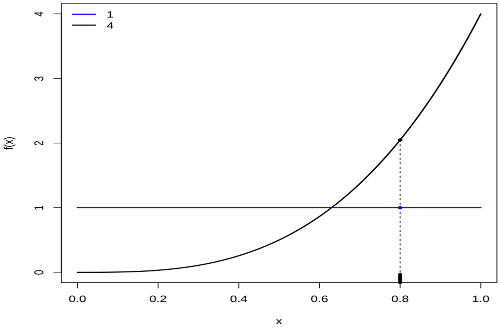
Logo, P$(X = 13; p = 6/7)/$P$(X = 13; p = 1/2) = 25$. Se $p = 6/7$, a probabilidade de 13 sucessos em 16 realizações é 25 vezes maior comparando com $p = 1/2$.

Escolhemos $\widehat{p} = 6/7$ como estimativa de p.

Método de máxima verossimilhança

Exemplo

X tem função densidade $f(x;\theta) = \theta x^{\theta-1}$, $x \in (0,1)$ e $\theta \in \{1,4\}$. Uma única observação foi coletada obtendo-se $x = 0,8$. Com base nesta observação, apresente uma estimativa para θ.



Método de máxima verossimilhança

Se $\theta = 1$, $f(0,8;1) = 1$.
Se $\theta = 4$, $f(0,8;4) = 2,048$.
Logo, $f(0,8;4)/f(0,8;1) = 2$. Se $\theta = 4$, a densidade de probabilidade de $x = 0,8$ é duas vezes maior comparando com $\theta = 1$.
Escolhemos $\widehat{\theta} = 4$ como estimativa de θ.
Problema. E se tivermos $n > 1$ e $\theta \in (0, \infty)$?

Método de máxima verossimilhança

$X_1, X_2, \ldots, X_n$ é uma amostra aleatória de uma população com função densidade $f(x;\theta)$.
Para uma amostra observada $x_1, x_2, \ldots, x_n$, a função verossimilhança de θ, denotada por $L(\theta)$, é dada por

$$L(\theta) = f(x_1;\theta)f(x_2;\theta) \times \cdots \times f(x_n;\theta) = \prod_{i=1}^{n} f(x_i;\theta). \tag{1}$$

Um estimador de máxima verossimilhança (EMV) de θ é um valor $\widehat{\theta}$ que maximiza $L(\theta)$.
$\widehat{\theta}$ que maximiza $L(\theta)$ também maximiza $\ell(\theta) = \ln(L(\theta))$, pois a função logaritmo é monótona crescente.

Método de máxima verossimilhança

Observação

Se X é uma v.a. discreta, a função verossimilhança de θ é

$$L(\theta) = P(X_1 = x_1, X_2 = x_2, \ldots, X_n = x_n; \theta), \quad (2)$$

ou seja, $L(\theta)$ é a probabilidade de obter $x_1, \ldots, x_n$ interpretada como função de θ. Um EMV de θ maximiza a probabilidade de ocorência dos valores da amostra.

Se X é uma v.a. contínua, $L(\theta)$ é a função densidade calculada em $x_1, \ldots, x_n$, interpretada como função de θ. Um EMV de θ maximiza a função densidade dos valores da amostra.

Exemplo

$X_1, \ldots, X_n$ é uma amostra aleatória de uma população Bernoulli com parâmetro p, $p \in (0,1)$. Obtenha o EMV de p.

Função densidade:

$$f(x) = \begin{cases} p^x(1-p)^{1-x}, & \text{se } x \in \{0,1\}, \\ 0, & c.c. \end{cases}.$$

Função verossimilhança de p:

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

Exemplo

Logaritmo da função verossimilhança:

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

Derivada do logaritmo da função verossimilhança em relação a p:

$$\frac{\partial \ell(p)}{\partial p} = \frac{\sum_{i=1}^{n} x_i}{p} - \frac{n - \sum_{i=1}^{n} x_i}{1-p}.$$

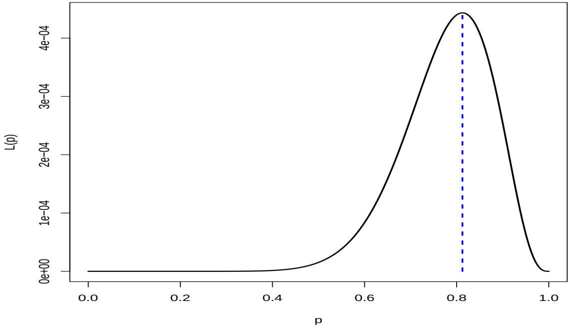
Igualando a derivada a 0 e resolvendo para p:

$$\widehat{p} = \frac{\sum_{i=1}^{n} x_i}{n}.$$

Uma vez que $\frac{\partial^2 \ell(p)}{\partial p^2}|_{\widehat{p}} < 0$, $\widehat{p} = \overline{X}$ é o EMV de p.

Exemplo

Em $n = 16$ *repetições de um experimento de Bernoulli foram obtidos 13 sucessos. Com base nesta amostra, apresente a EMV para* p.
Pelo exemplo anterior, $\widehat{p} = 13/16 = 0,8125$.



Exemplo

O tempo de falha, em horas, de um componente eletrônico tem distribuição exponencial com parâmetro θ desconhecido. Oito unidades são selecionadas ao acaso e testadas com os seguintes de tempos de falha:

$$11,96 \quad 5,03 \quad 67,40 \quad 16,07 \quad 31,50 \quad 7,73 \quad 11,10 \quad 22,38.$$

Obtenha a EMV de θ.

Função verossimilhança de θ:

$$L(\theta) = \prod_{i=1}^{n} \theta e^{-\theta x_i} = \theta^n e^{-\theta \sum\limits_{i=1}^{n} x_i}.$$

Logaritmo da função verossimilhança:

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

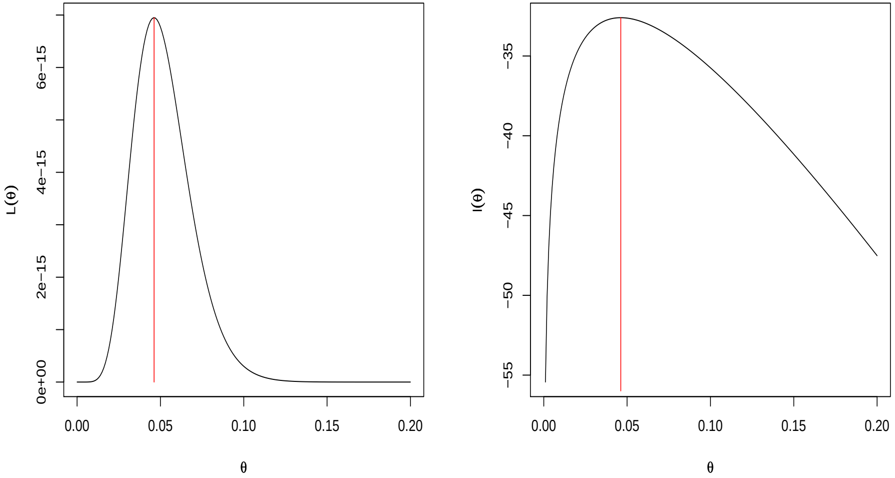
Derivada do logaritmo da função verossimilhança em relação a θ:

$$\frac{\partial \ell(\theta)}{\partial \theta} = \frac{n}{\theta} - \sum_{i=1}^{n} x_i.$$

Igualando a derivada a 0 e resolvendo para θ:

$$\widehat{\theta} = \frac{n}{\sum_{i=1}^{n} x_i} = \frac{1}{\bar{x}}.$$

Uma vez que $\frac{\partial^2 \ell(\theta)}{\partial \theta^2}|_{\widehat{\theta}} < 0$, $\widehat{\theta} = 1/\overline{X}$ é o EMV de θ.
Dos dados obtemos $\overline{x} = 21,65$ e $\widehat{\theta} = 0,0462$.


L(θ)
0e+00
2e−15
4e−15
6e−15
0.00
0.05
0.10
0.15
0.20
θ
l(θ)
−35
−40
−45
−50
−55
0.00
0.05
0.10
0.15
0.20
θ

Exemplo

Seja $X_1, X_2, \ldots, X_n$ *é uma amostra aleatória de uma população normal com média* μ *e variância* σ^2*. Obtenha os EMV de* μ *e* σ^2*.*

Pode ser provado que

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

são os EMV de μ e σ^2.
Notar que $\widehat{\sigma^2} = (n-1)S^2/n < S^2$; logo, $\widehat{\sigma^2}$ é um estimador viesado de σ^2.

Propriedades dos EMV

Invariância. Se $\widehat{\theta}_1, \ldots, \widehat{\theta}_k$ são os EMV de $\theta_1, \ldots, \theta_k$, então o EMV de $h(\theta_1, \ldots, \theta_k)$ é $h(\widehat{\theta}_1, \ldots, \widehat{\theta}_k)$.

Exemplo

Para a distribuição normal, o EMV do desvio padrão σ é $\widehat{\sigma} = \sqrt{\widehat{\sigma^2}}$.
O EMV do coeficiente de variação σ/μ é $\sqrt{\widehat{\sigma^2}}/\overline{X}$.

Exemplo

Para a distribuição Bernoulli, o EMV da chance de sucesso $p/(1-p)$ é $\widehat{p}/(1-\widehat{p})$, em que $\widehat{p} = \overline{X}$ é a proporção amostral de sucessos.

Propriedades dos EMV

Sob certas condições que são geralmente válidas, o EMV $\widehat{\theta}$ de θ é consistente.
Se o tamanho da amostra n for suficientemente grande,

Propriedades dos EMV

Sob certas condições que são geralmente válidas, o EMV $\widehat{\theta}$ de θ é consistente.

Se o tamanho da amostra n for suficientemente grande,

- $\widehat{\theta}$ é aproximadamente não viesado para θ e

Propriedades dos EMV

Sob certas condições que são geralmente válidas, o EMV $\widehat{\theta}$ de θ é consistente.

Se o tamanho da amostra n for suficientemente grande,

- $\widehat{\theta}$ é aproximadamente não viesado para θ e
- $\widehat{\theta}$ tem distribuição aproximadamente normal.