

Viés

Em Estatística, a palavra viés tem um significado preciso e tem a ver com a esperança da distribuição de um estimador.

Definição 32 (Estimador não-viesado)

*Um estimador $\delta(\mathbf{X})$ de uma função $g(\theta)$ é dito **não-viesado** se $E_{\theta}[\delta(\mathbf{X})] = g(\theta)$ para todo $\theta \in \Omega$. Um estimador que não atende a essa condição é dito viesado. O **viés** de δ é definido como $B_{\delta}(\theta) := E_{\theta}[\delta(\mathbf{X})] - g(\theta)$.*

Exemplo 12 (Tempos de falha de lâmpadas)

Lembremos do exemplo das lâmpadas da fábrica de Astolfo. Neste caso, não é difícil mostrar que $E[\hat{\theta}_{EMV}] = \frac{n}{n-1}\theta = 3\theta/2$. Desta forma, o viés do EMV é $B_{\hat{\theta}_{EMV}}(\theta) = 3\theta/2 - \theta = \theta/2$. É possível encontrar $\delta(\mathbf{X})$ não-viesado? Esse estimador é bom?

Estimadores não-viesados sempre?

Quando avaliamos estimadores, o erro quadrático médio e o viés são alguns *aspectos* a serem considerados, mas há um compromisso (*trade-off*) entre eles, de certa forma.

Observação 13 (Erro quadrático, variância e viés)

$$R(\theta, \delta) = \text{Var}_{\theta}(\delta) + [B_{\delta}(\theta)]^2.$$

No exemplo das lâmpadas, é possível mostrar que $\delta_2(\mathbf{X}) = 1/S$ tem o menor EQM, mas tem viés $B_{\delta_2}(\theta) = \frac{n-2}{n-1}\theta = \theta/2$, assim como o EMV.

Estimador não-viesado da variância

A variância amostral como a temos definido até aqui é viesada. Uma pequena modificação leva a um estimador não viesado da variância.

Teorema 18 (Estimador não-viesado da variância)

Seja $\mathbf{X} = \{X_1, X_2, \dots, X_n\}$ uma amostra aleatória, com $E[X_1] = m$ e $\text{Var}(X_1) = v < \infty$. Então

$$\delta_1(\mathbf{X}) = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X}_n)^2$$

é um estimador não-viesado de v .

Prova: usar a igualdade

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

e usar a linearidade da esperança e o fato de que temos uma amostra aleatória.

Nem tudo são flores

Não-viesamento é uma característica desejável, mas nem sempre um estimador não-viesado (i) existe ou (ii) é um bom estimador.

- Não existência. Exemplo: $X_1, X_2, \dots, X_n \sim \text{Bernoulli}(p)$, estimador para \sqrt{p} ?
- Estimador não-viesado ruim: $X \sim \text{Geometrica}(p)$. Quais as propriedades do estimador não viesado, $\delta(X)$?

Escolhendo entre desenhos amostrais

Exemplo 13 (Estudando chegada de clientes)

Exemplo 8.8.1 em DeGroot. Suponha que Palmirinha esteja interessada em estudar quantos clientes chegam à sua loja de pamonha num determinado intervalo. Para isso, ela vai modelar o fenômeno como um processo de Poisson:

$$Y(\Delta_t) \sim \text{Poisson}(\theta\Delta_t),$$

isto é, o número Y de clientes num intervalo de tempo Δ_t tem distribuição Poisson com média $\theta\Delta_t$. Palmirinha pode

- *Fixar um número n de clientes a serem observados e marcar o tempo, X que leva para chegarem n clientes ou;*
- *Fixar um determinado intervalo de tempo, t , e contar o número Y de clientes que chegam neste intervalo.*

Pergunta: *qual desenho é melhor para estimar θ ?*

Informação de Fisher

Como medir a quantidade de informação (sobre um parâmetro θ) contida em uma amostra aleatória? A (matriz de) informação de Fisher oferece a resposta.

Definição 33

Seja X uma variável aleatória com f.d.p/f.m.p. $f(x | \theta)$, $\theta \in \Omega \subseteq \mathbb{R}$. Suponha que $f(x | \theta)$ é duas vezes diferenciável com respeito a θ . Defina $\lambda(x | \theta) = \log f(x | \theta)$ e

$$\lambda'(x | \theta) = \frac{\partial \lambda(x | \theta)}{\partial \theta} \quad \text{e} \quad \lambda''(x | \theta) = \frac{\partial^2 \lambda(x | \theta)}{\partial \theta^2}. \quad (20)$$

Definimos a **informação de Fisher** como

$$I(\theta) = E_{\theta} [\{\lambda''(x | \theta)\}^2] \stackrel{(1)}{=} -E_{\theta} [\lambda''(x | \theta)] = \text{Var}_{\theta} (\lambda'(x | \theta)). \quad (21)$$

Prova de ⁽¹⁾: diferenciar sob o sinal da integral e usar a regra da cadeia.

Informação de Fisher: exemplos

- Bernoulli;
- Normal;

Informação de Fisher: exemplos

- Bernoulli;

$$I(p) = \frac{1}{p(1-p)}.$$

- Normal;

$$I(\mu) = \frac{1}{\sigma^2}.$$

Informação de Fisher de uma amostra aleatória

Teorema 19

Seja $\mathbf{X} = \{X_1, X_2, \dots, X_n\}$ uma amostra aleatória e seja $I_n(\theta) = E_\theta [-\lambda''_n(\mathbf{X} | \theta)]$ a informação de Fisher da amostra. Então

$$I_n(\theta) = nI(\theta).$$

Prova: Usar as propriedades do log, da derivada e a lei de esperanças. Ver De Groot, Teorema 8.8.2.

Voltando ao dilema de Palmirinha

Podemos usar a informação de Fisher para analisar os desenhos propostos por Palmirinha. Não é difícil derivar

$$I_X(\theta) = \frac{n}{\theta^2} \quad \text{e} \quad I_Y(\theta) = \frac{t}{\theta}.$$

Portanto, os desenhos são equivalentes se $n = t\theta$, o que não ajuda muito, já que θ é desconhecido. Por outro lado, vemos que neste caso não é possível decidir entre os desenhos baseado apenas na informação de Fisher.

Extra: faça uma análise Bayesiana deste problema, derivando a esperança *a priori* da informação de Fisher sob os dois desenhos.

O Teorema de Crámer-Rao

Outro uso importante da informação de Fisher é encontrar uma cota inferior para a variância de um estimador. Para isso, empregamos um dos resultados mais importantes da Estatística:

Teorema 20 (Teorema de Crámer-Rao¹²)

Seja $\mathbf{X} = \{X_1, X_2, \dots, X_n\}$ uma amostra aleatória com f.d.p./f.m.p $f(x | \theta)$, com as mesmas premissas da definição 33. Suponha que $T = r(\mathbf{X})$ é uma estatística com variância finita. Seja $m(\theta) = E_\theta(T)$ uma função diferenciável de θ . Então,

$$\text{Var}_\theta(T) \geq \frac{[m'(\theta)]^2}{nI(\theta)}, \quad (22)$$

com igualdade apenas se existem u e v tal que

$$T = u(\theta)\lambda'_n(\mathbf{X} | \theta) + v(\theta).$$

Prova: Usar Cauchy-Schwarz e diferenciar sob o sinal da integral.

¹²Em homenagem ao estatístico indo-estadunidense Calyampudi Radhakrishna Rao (1920-) e ao matemático sueco Harald Cramér (1893–1985).

Um corolário útil

Se T é um estimador não-viesado, temos uma expressão útil para a cota de Crámer-Rao.

Observação 14 (Variância de um estimador não-viesado)

Se T é um estimador não-viesado de θ , temos

$$\text{Var}_{\theta}(T) \geq \frac{1}{nI(\theta)}$$

Prova: T é não viesado $\implies m(\theta) = \theta \implies m'(\theta) = 1 \forall \theta \in \Omega$ \square

Eficiência

Com esse Teorema de Crámer-Rao em mãos, estamos em posição de definir um critério de otimalidade para estimadores.

Definição 34 (Estimador eficiente)

Um estimador $\delta(\mathbf{X})$ é dito **eficiente** de (sua esperança) $m(\theta)$ se

$$\text{Var}_{\theta}(\delta) = \frac{[m'(\theta)]^2}{nI(\theta)}.$$

Exemplo 14

Seja X_1, X_2, \dots, X_n uma amostra aleatória de uma distribuição Poisson com parâmetro θ . Podemos mostrar que \bar{X}_n é um estimador eficiente de θ .

Distribuição assintótica de um estimador eficiente

Podemos usar o TCL para estudar a distribuição assintótica de um estimador eficiente.

Teorema 21 (Distribuição assintótica de um estimador eficiente)

Assumindo as condições de regularidade usuais, considere δ um estimador eficiente de $m(\theta)$. Assuma também que $m'(\theta) \neq 0 \forall \theta \in \Omega$. Então a distribuição assintótica de

$$\frac{\sqrt{nl(\theta)}}{m'(\theta)} [\delta - m(\theta)]$$

é normal padrão.

Prova: Ver De Groot, Teorema 8.8.4. Escrever $E_\theta[\delta]$ e $\text{Var}_\theta(\delta)$ explicitamente, usar a condição $\delta = u(\theta)\lambda'_n(\mathbf{X} | \theta) + v(\theta)$, e aplicar as leis de esperanças e variâncias.




Observação 15 (Normalidade Assintótica do EMV)

Supondo que o EMV possa ser derivado ao resolver a equação $\lambda'_n(\mathbf{X} | \theta) = 0$ e que $\lambda''_n(\mathbf{X} | \theta)$ e $\lambda'''_n(\mathbf{X} | \theta)$ satisfazem certas condições técnicas, $\sqrt{nl(\theta)} (\hat{\theta}_{EMV} - \theta)^2$ tem distribuição aproximadamente normal padrão.

O que aprendemos?

- 💡 Viés;
“Um estimador viesado é aquele cuja esperança não coincide com a função estimada”
- 💡 Informação de Fisher;
“A informação de Fisher é uma quantidade derivada de uma distribuição que mede a quantidade de informação contida em uma amostra aleatória advinda desta distribuição”
- 💡 Crámer-Rao;
“A desigualdade de Crámer-Rao dá uma cota inferior para a variância de um estimador”
- 💡 Distribuição assintótica de estimadores eficientes (e EMV);
“Sob condições de regularidade, vale um TCL para estimadores eficientes e para o EMV”

Leitura recomendada

-  De Groot seções 8.7 e 8.8;
-  * Casella & Berger (2002), seção 7.3.
-  * Schervish (1995), Teorema 5.13.

- **Exercícios recomendados**

-  De Groot.

- Seção 8.7: exercícios 4, 6, 11 e 13;

- Seção 8.8: exercícios 5, 7 e 10.