8.2 Fisher Information CR Inequality

- $X \sim f(x,\theta), \;\; \theta \in \Theta = (a,b).$
- Regular cases:

$$f(x,\theta) > 0, \;\; \frac{\partial}{\partial\theta}\int f = \int \frac{\partial}{\partial\theta} f,$$

Since $\int f(x,\theta) = 1$, then

$$\int \frac{\partial f(x,\theta)}{\partial \theta} dx = 0,$$

i.e.

$$\int \frac{\partial \ln f(x,\theta)}{\partial \theta} f(x,\theta) dx = 0.$$

$$\Rightarrow \quad \int [\frac{\partial^2 \ln f(x,\theta)}{\partial \theta^2} f(x,\theta)$$

$$+ \frac{\partial \ln f(x,\theta)}{\partial \theta} \frac{\partial f(x,\theta)}{\partial \theta}] dx = 0,$$

$$\Rightarrow \quad \int (\frac{\partial \ln f(x,\theta)}{\partial \theta})^2 f(x,\theta) dx$$

$$+ \int \frac{\partial^2 \ln f(x,\theta)}{\partial \theta^2} f(x,\theta) dx = 0.$$

Therefore,

$$I(\theta) =: \int \left(\frac{\partial \ln f(x,\theta)}{\partial \theta} \right)^2 f(x,\theta)dx$$

$$= E\left[\frac{\partial \ln f(x,\theta)}{\partial \theta} \right]^2$$

$$= -\int \frac{\partial^2 \ln f(x,\theta)}{\partial \theta^2} f(x,\theta)dx$$

$$= -E\frac{\partial^2 \ln f(x,\theta)}{\partial \theta^2}.$$

首都师范大学
数学科学学院
统计学系
Department of Statistics

- The larger of $I(\theta)$, the more information of θ.

- If $I(\theta) = 0$, then $\frac{\partial \ln f}{\partial \theta} = 0$. No information in $\ln f(x, \theta)$.

- $\frac{\partial \ln f(x,\theta)}{\partial \theta}$ is an important function.

MLE $\hat{\theta}$ satisfies:

$$\sum_{i=1}^{n} \frac{\partial \ln f(x_i, \hat{\theta})}{\partial \theta} = 0.$$

Example 1: $X \sim N(\theta, \sigma_0^2)$,

$$f(x,\theta) = \frac{1}{\sqrt{2\pi}\sigma_0} \exp\{-\frac{(x-\theta)^2}{2\sigma_0^2}\}.$$

Then

$$\ln f(x,\theta) = -\ln(\sqrt{2\pi}\sigma_0) - \frac{(x-\theta)^2}{2\sigma_0^2}.$$

$\Rightarrow$

$$\frac{\partial \ln f(x,\theta)}{\partial \theta} = \frac{x-\theta}{\sigma_0^2},$$

$\Rightarrow$

$$\frac{\partial^2 \ln f(x,\theta)}{\partial \theta^2} = -\frac{1}{\sigma_0^2}.$$

Then

$$I(\theta) = \frac{1}{\sigma_0^2}.$$

Example 2: $X \sim B(1, \theta)$, then

$$\begin{aligned}\ln f(x,\theta) &= x \ln\theta \\ &\quad + (1-x)\ln(1-\theta),\end{aligned}$$

$$\frac{\partial^2 \ln f(x,\theta)}{\partial\theta^2} = -\frac{x}{\theta^2} - \frac{(1-x)}{(1-\theta)^2}.$$

Hence,

$$I(\theta) = E\left[\frac{x}{\theta^2} + \frac{1-x}{(1-\theta)^2}\right]$$

$$= \frac{1}{\theta} + \frac{1}{1-\theta} = \frac{1}{\theta(1-\theta)}.$$

- Suppose $X_1, \cdots, X_n$ iid. $\sim f(x, \theta)$, then

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

and

$$\ln L(\theta) = \sum_{i=1}^{n} \ln f(x_i, \theta).$$

$\Rightarrow$

$$\frac{\partial \ln L(\theta)}{\partial \theta} = \sum_{i=1}^{n} \frac{\partial \ln f(x_i, \theta)}{\partial \theta}.$$

$$\begin{aligned} I_n(\theta) &= E[\frac{\partial \ln L(\theta)}{\partial \theta}]^2 \\ &= \sum_{i=1}^{n} E(\frac{\partial \ln f(x_i, \theta)}{\partial \theta})^2 \\ &= nI(\theta). \end{aligned}$$

- **Rao-Cramer inequality**:

$$Y = u(X_1, \cdots, X_n) \sim \theta,$$

$$E(Y) = k(\theta).$$

$$\Rightarrow k(\theta) = \int u(x_1, \cdots, x_n) \prod f(x_i, \theta) dx_i$$

$$k'(\theta) = \int u(x_1, \cdots, x_n)$$

$$(\sum_{i=1}^{n} \frac{\partial \ln f(x_i, \theta)}{\partial \theta}) \prod f(x_i, \theta) dx_i.$$

Let $Z = \sum_{i=1}^{n} \frac{\partial \ln f(x_i,\theta)}{\partial \theta}$, then

$$EZ = 0,$$

$$Var(Z) = nE\left[\frac{\partial \ln f(X,\theta)}{\partial \theta}\right]^2$$

$$= nI(\theta).$$

首都师范大学 数学科学学院 统计学系 Department of Statistics

Since $k'(\theta) = E(YZ)$, then

$$k'(\theta) \leq \sqrt{Var(Y)Var(Z)}$$

$\Rightarrow$

$$Var(Y) \geq \frac{[k'(\theta)]^2}{Var(Z)}$$

$$= \frac{[k'(\theta)]^2}{nI(\theta)}.$$

- If $k(\theta) = \theta$, then Y is unbiased and

$$Var(Y) \geq \frac{1}{nI(\theta)}$$

CR-lower bound.

Example 1: $X_1, \cdots, X_n$ iid. $\sim N(\theta, \sigma^2)$ with known σ^2. Then

$$Var(Y) \geq \frac{\sigma^2}{n}.$$

Example 2: $X_1, \cdots, X_n$ iid. $\sim B(1, \theta)$. Then

$$Var(Y) \geq \frac{\theta(1-\theta)}{n}.$$

Def 1: Y is a efficient unbiased estimator of $\theta \Longleftrightarrow$ the variance of Y attains the CR lower bound.

Def 2: Efficiency

$$\frac{\text{The CR lower bound}}{Var(Y)}$$

is called the efficiency of unbiased estimator Y.

Example 3: $X_1, \cdots, X_n$ iid. $\sim N(\mu, \theta)$, $0 < \theta < +\infty$ with known μ.

$$E(S^2) = \theta.$$

and

$$\ln f(x, \theta) = -\frac{(x-\mu)^2}{2\theta} - \frac{\ln \theta}{2} + c.$$

首都师范大学 数学科学学院 统计学系 Department of Statistics

$$I(\theta) = -E\left[\frac{\partial \ln^2 f(X,\theta)}{\partial\theta^2}\right]$$

$$= E\left[\frac{(X-\mu)^2}{\theta^3} - \frac{1}{2\theta^2}\right] = \frac{1}{2\theta^2}.$$

and

$$\frac{1}{nI(\theta)} = \frac{2\theta^2}{n}.$$

Since

$$Var(S^2) = \frac{2(n-1)}{(n-1)^2}\theta^2 = \frac{2\theta^2}{n-1}$$

$$> \frac{2\theta^2}{n},$$

then

$$\text{eff}(S^2) = \frac{n-1}{n}.$$

Take $S_0^2 = \frac{1}{n}\sum_{i=1}^n (X_i - \mu)^2$, then

$$\mathrm{eff}(S_0^2) = 1.$$

- Asymptotically efficient:

$$\mathrm{eff}(Y_n) \to 1 \ (n \to \infty).$$