Cramer-Rao Lower Bound

Definitions

Let $X_1, \dots, X_n$ r.s. with PDF $f(x;\theta)$. The quantity $I_X(\theta) = E\left[\left(\frac{\partial(\log f(x;\theta))}{\partial\theta}\right)^2\right]$ is called the Fischer information included in the r.v. X about θ.

The quantity $\frac{\partial(\log f(x;\theta))}{\partial\theta}$ is called the score function

Notes

- The score function has expected value equal to 0.

  Proof

  $$\int f(x;\theta)\,dx = 1 \xRightarrow{\partial/\partial\theta} \int \frac{\partial}{\partial\theta} f(x;\theta)\,dx = 0 \Rightarrow$$

  $$\Rightarrow \int \frac{\frac{\partial f(x;\theta)}{\partial\theta}}{f(x;\theta)} \cdot f(x;\theta)\,dx = 0 \Rightarrow \int \frac{\partial \log f(x;\theta)}{\partial\theta} \cdot f(x;\theta)\,dx = 0 \quad (*)$$

  $$\Rightarrow E\left[\frac{\partial \log f(x;\theta)}{\partial\theta}\right] = 0$$

- The Fischer information is $I_X(\theta) = -E\left[\frac{\partial^2}{\partial\theta^2} \log f(x;\theta)\right]$

  Proof

  We derive a second time:

  $$(*) \xRightarrow{\partial/\partial\theta} \int\left[\frac{\partial^2}{\partial\theta^2} \log f(x;\theta) \cdot f(x;\theta) + \frac{\partial}{\partial\theta} \log f(x;\theta) \cdot \frac{\partial}{\partial\theta} f(x;\theta)\right]dx = 0$$

  $$\Rightarrow \int \frac{\partial^2}{\partial\theta^2} \log f(x;\theta) \cdot f(x;\theta)\,dx + \int \frac{\partial}{\partial\theta} \log f(x;\theta) \cdot \underbrace{\frac{\partial/\partial\theta\, f(x;\theta)}{f(x;\theta)}}_{\frac{\partial}{\partial\theta} \log f(x;\theta)} \cdot f(x;\theta)\,dx = 0$$

  $$\Rightarrow E\left[\frac{\partial}{\partial\theta^2} \log f(x;\theta)\right] + E\left[\left(\frac{\partial}{\partial\theta} \log f(x;\theta)\right)^2\right] = 0$$

  $$\Rightarrow I_X(\theta) = -E\left[\frac{\partial}{\partial\theta^2} \log f(x;\theta)\right)$$

Theorem (Cramer-Rao inequality)

Let $X_1, \dots, X_n$ with PDF $f(x;\theta)$ and $T = T(\underline{X})$ estimator of θ with bias $b_T(\theta) = E[T] - \theta$. Then:

$$V(T) \geq \frac{[1+b_\theta'(T)]^2}{\nu I_X(\theta)}$$

Generally, if $T=T(\underline{X})$ is an estimator of $g(\theta)$ and $b_{g(\theta)}(T)=g(\theta)-E[T]$ then:

$$V(T) \geq \frac{[g'(\theta)+b_{g(\theta)}'(T)]^2}{\nu I_X(\theta)}$$

Proof

X has PDF $f(x;\theta)$ so $X_1,\dots,X_\nu$ r.s of PDF $f(x;\theta)$

We have: $E[T]=\int\int\cdots\int T(\underline{x}) f(x_1;\theta)\cdots f(x_\nu;\theta)\,dx_1\cdots dx_\nu = \theta + b_\theta(T)$

$$\overset{\partial/\partial\theta}{\Longrightarrow} 1+b_\theta'(T) = \sum_{i=1}^{\nu}\int\int\cdots\int T(\underline{x}) f(x_1;\theta)\cdots\frac{\partial f(x_i;\theta)}{\partial\theta}\,dx_1\cdots dx_\nu =$$

$$= \int\int\cdots\int T(\underline{x}) f(x_1;\theta)\cdot\ldots\cdot f(x_\nu;\theta)\cdot\sum_{i=1}^{\nu}\left(\frac{\frac{\partial f(x_i;\theta)}{\partial\theta}}{f(x_i;\theta)}\right)dx_1\cdots dx_\nu =$$

$$= \int\int\cdots\int T(\underline{X})\cdot\prod_{i=1}^{\nu} f(x_i;\theta)\cdot\sum_{i=1}^{\nu}\frac{\partial(\log f(x_i;\theta))}{\partial\theta}\,dx_1\cdots dx_\nu =$$

$$= E\left[T(\underline{X})\sum_{i=1}^{\nu}\frac{\partial(\log f(x_i;\theta)}{\partial\theta}\right]$$

Also: $\int\int\cdots\int f(x_1;\theta)\cdots f(x_\nu;\theta) = 1$ so by deriving like before, we get $E\left[\sum_{i=1}^{\nu}\frac{\partial\log f(x_i;\theta)}{\partial\theta}\right]=0$.

We set $W=\sum_{i=1}^{\nu}\frac{\partial\log f(x_i;\theta)}{\partial\theta}$ and it is $\begin{cases} E[TW]=1+b_\theta'(T) \\ E[W]=0\end{cases}$

and $Cov(T,W)=E[TW]-E[T]\cdot E[W]=1+b_\theta'(T)$

Then we have: $V(W)=\sum_{i=1}^{\nu}V\left(\frac{\partial\log f(x_i;\theta)}{\partial\theta}\right)=\sum_{i=1}^{\nu}E\left[\left(\frac{\partial\log f(x_i;\theta)}{\partial\theta}\right)^2\right]=$

$=\nu I_X(\theta)=I_{\underline{X}}(\theta)$

So from Cauchy-Schwarz inequality we get:

$$(Cov(T,W))^2 \leq V(T)\cdot V(W) \Rightarrow$$

$$\Rightarrow V(T)\geq\frac{(1+b_\theta'(T))^2}{\nu I_X(\theta)}$$

$\square$

Note

a) If $E[T(\underline{x})] = g(\theta) + b_\theta'(T)$, then $V(T) \geq \frac{(g'(\theta) + b_\theta'(T))^2}{\nu I_X(\theta)}$

b) If $b_{g(\theta)}(T) = 0$, then $V(T) \geq \frac{(g'(\theta))^2}{\nu I_X(\theta)}$

c) If $g(\theta) = \theta$ and $b_\theta(T) = 0$ then $V(T) \geq \frac{1}{\nu I_X(\theta)} = \frac{1}{I_{\underline{X}}(\theta)} = I_{\underline{X}}(\theta)^{-1}$

Example

Let $X_1, \dots, X_\nu$ r.s of Poisson(θ). Find m.v.u.e using the C-R lower bound.

Solution

We have: $I_X(\theta) = E\left[\left(\frac{\partial \log f(x;\theta)}{\partial \theta}\right)^2\right] = E\left[\left(\frac{\partial}{\partial\theta}\left(\log\left(e^{-\theta}\frac{\theta^x}{x!}\right)\right)\right)^2\right] =$

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

$= \frac{1}{\theta^2}E[(x-\theta)^2] = \frac{1}{\theta^2}V(X) = \frac{1}{\theta^2}\cdot\theta = \frac{1}{\theta}$, so $I_X(\theta) = \frac{1}{\theta}$.

So the C-R lower bound is: $V(T) \geq \frac{1}{\nu I_X(\theta)} = \frac{\theta}{\nu}$

We know that $T = T(\underline{x}) = \sum_{i=1}^{\nu}\frac{x_i}{\nu} = \bar{X}$ u.e. of θ.

We have $E[\bar{X}] = \theta$ and $V(\bar{X}) = V\left(\frac{\sum x_i}{\nu}\right) = \frac{1}{\nu^2}\sum_{i=1}^{\nu}V(x_i) = \frac{1}{\nu^2}\nu\theta =$

$= \frac{\theta}{\nu}$. Hence, T is m.v.u.e of θ.