Lecture 20

Maximum Likelihood Estimation (III)

Cramer-Rao Lower Bound

1. CRLB Theorem

> Under Assumption (1)-(2),
> $$\mathrm{Var}(\hat{\theta}(Y)) \geq \frac{\left(\frac{\partial}{\partial\theta}\mathbb{E}[\hat{\theta}(Y)]\right)^2}{I(\theta)},$$
> for any ~~unbiased~~ estimator $\hat{\theta}$.

Pf: $\mathrm{Var}(\hat{\theta}(Y))\, I(\theta)$

$$= \int \left(\hat{\theta}(y) - \mathbb{E}[\hat{\theta}(y)]\right)^2 f_\theta(y)\,dy \cdot \int \left(\frac{\partial}{\partial\theta}\ln f_\theta(y)\right)^2 f_\theta(y)\,dy$$

$$= \mathbb{E}[A^2]\,\mathbb{E}[B^2]$$

$$\geq \mathbb{E}[AB]^2 \qquad \text{(Cauchy inequality)}$$

$$= \left[\int \left(\hat{\theta}(y) - \mathbb{E}[\hat{\theta}(y)]\right)\left(\frac{\partial}{\partial\theta}\ln f_\theta(y)\right) f_\theta(y)\,dy\right]^2$$

$$= \left[\int \left(\hat{\theta}(y) - \mathbb{E}[\hat{\theta}(Y)]\right)\frac{\partial}{\partial\theta} f_\theta(y)\,dy\right]^2$$

$$= \left[\int \hat{\theta}(y)\frac{\partial}{\partial\theta} f_\theta(y)\,dy - \mathbb{E}[\hat{\theta}(Y)]\int \frac{\partial}{\partial\theta} f_\theta(y)\,dy\right]^2$$

$$= \left[\frac{\partial}{\partial\theta}\mathbb{E}[\hat{\theta}(Y)] - 0\right]^2$$

$$= \left(\frac{\partial}{\partial\theta}\mathbb{E}[\hat{\theta}(Y)]\right)^2.$$

(2) if $\hat{\theta}$ is ~~an~~ a general estimator, then

$$\frac{\partial}{\partial\theta}\ln f_\theta(y) = k(\theta)\left(\hat{\theta}(y) - \mathbb{E}[\hat{\theta}(Y)]\right)$$

$$\Rightarrow \ln f_\theta(y) = \int_a^\theta k(\theta')\left(\hat{\theta}(y) - \mathbb{E}[\hat{\theta}(Y)]\right)d\theta' + C(y).$$

Fundamental theorem Calculus

$$\Rightarrow \quad = \underbrace{-\int_a^\theta k(\theta')\,\mathbb{E}[\hat{\theta}(Y)]\,d\theta'}_{\ln A(\theta)} + \underbrace{C(y)}_{\ln h(y)} + \underbrace{\hat{\theta}(y)\int_a^\theta k(\theta')\,d\theta'}_{Q(\theta)}$$

$$\Rightarrow \quad f_\theta(y) = A(\theta)\exp\left(Q(\theta)\,\hat{\theta}(y)\right)\cdot h(y).$$

$\Rightarrow$ $f_\theta(y)$ is a one-parameter exponential family.

Conversely, if $f_\theta(y) = A(\theta)\exp\left(Q(\theta)\hat{\theta}(y)\right)\cdot h(y)$, then CRLB equality is satified, when

$\hat{\theta}(y)$ is a suff. statistic.

Conditions for Equality

(1). $\frac{\partial}{\partial\theta} \ln f_\theta(y) = I(\theta)\left[\hat{\theta}(y) - \theta\right]$,

if $\hat{\theta}(Y)$ is an unbiased estimator.

Pf: $\cancel{\frac{\partial^2}{\partial\theta^2} \ln f_\theta(y) = I'(\theta)(\hat{\theta}-\theta) - I(\theta)}$

$\cancel{\text{So } \mathbb{E}\left[-\frac{\partial^2}{\partial\theta^2} \ln f_\theta(y)\right] = -I'(\theta)\left(\mathbb{E}[\hat{\theta}-\theta]\right) + I(\theta)}$

$= I(\theta).$

Pf: Cauchy holds when

$$A = \alpha \cdot B.$$

$\mathbb{E}[\hat{\theta}] = \theta$

So

$$\frac{\partial}{\partial\theta} \ln f_\theta(y) = k(\theta)\cdot\left(\hat{\theta}(y) - \theta\right).$$

$$\Rightarrow \frac{\partial^2}{\partial\theta^2} \ln f_\theta(y) = k'(\theta)(\hat{\theta}(y)-\theta) - k(\theta)$$

$$\Rightarrow \underbrace{-\mathbb{E}\left[\frac{\partial^2}{\partial\theta^2} \ln f_\theta(y)\right]}_{= I(\theta)} = k(\theta)$$

So $\frac{\partial}{\partial\theta} \ln f_\theta(y) = I(\theta)\left(\hat{\theta}(y) - \theta\right)$.

2. Examples

Example 1

$$Y_k \overset{iid}{\sim} \mathcal{N}(\theta, \sigma^2), \quad k = 1, \dots, n$$

Find CRLB for $\hat{\theta}(\underline{Y}) = \frac{1}{n}\sum_{i=1}^{n} Y_i$. Is $\hat{\theta}(\underline{Y})$ MVUE?

Solution $I(\theta) = \frac{n}{\sigma^2}$.

$$\text{So } Var(\hat{\theta}) \geq \frac{\left(\frac{\partial}{\partial\theta}\mathbb{E}[\hat{\theta}(Y)]\right)^2}{I(\theta)}.$$

if $\hat{\theta}(\underline{Y}) = \frac{1}{n}\sum_{i=1}^{n} Y_i$,

then $\mathbb{E}[\hat{\theta}(Y)] = \frac{1}{n}\sum_{i=1}^{n}\mathbb{E}[Y_i] = \theta$.

So $\frac{\partial}{\partial\theta}\mathbb{E}[\hat{\theta}(Y)] = 1$.

Hence $Var(\hat{\theta}) \geq \frac{1}{I(\theta)} = \frac{\sigma^2}{n}$.

$$Var(\hat{\theta}) = Var\left(\frac{1}{n}\sum_{i=1}^{n} Y_i\right)$$
$$= \frac{1}{n^2}\sum_{i=1}^{n} Var(Y_i) = \frac{\sigma^2}{n}.$$

So lower bound is attained at $\hat{\theta}(\underline{Y}) = \frac{1}{n}\sum_{i=1}^{n} Y_i$.

Therefore,

$\hat{\theta}(\underline{Y}) = \frac{1}{n}\sum_{i=1}^{n} Y_i$ is MVUE.

Example 2

$$Y_k \overset{iid}{\sim} s_\theta(k) + N_k \qquad k = 1, \dots, n$$

where $S_k(\theta)$ is a function of k.

$N_k \sim N(0, \sigma^2)$.

Find CRLB for any unbiased estimator $\hat{\theta}$.

Solution:

$$\ln f_\theta(\underline{y}) = \cancel{N} \frac{n}{2} \ln(2\pi\sigma^2) - \frac{1}{2\sigma^2} \sum_{i=1}^{n} (y_i - s_k(\theta))^2$$

$$\frac{\partial}{\partial\theta} \ln f_\theta(\underline{y}) = \cancel{n} \frac{1}{\cancel{2}\sigma^2} \sum_{k=1}^{n} (y_k - s_k(\theta)) \frac{\partial}{\partial\theta} s_k(\theta)$$

$$\frac{\partial^2}{\partial\theta^2} \ln f_\theta(\underline{y}) = \frac{1}{\sigma^2} \sum_{k=1}^{n} \Big[(y_k - s_k(\theta)) \frac{\partial^2}{\partial\theta^2} s_k(\theta) - \left(\frac{\partial}{\partial\theta} s_k(\theta)\right)^2 \cancel{(y_k - s_k(\theta))} \Big]$$

$$= \frac{1}{\sigma^2} \sum_{k=1}^{n} (y_k - s_k(\theta)) \left[\frac{\partial^2}{\partial\theta^2} s_k(\theta)\right] - \left[\left(\frac{\partial}{\partial\theta} s_k(\theta)\right)^2\right].$$

$$\mathbb{E}\left[-\frac{\partial^2}{\partial\theta^2}\right] = \frac{1}{\sigma^2} \sum_{k=1}^{n} \left(\frac{\partial}{\partial\theta} s_k(\theta)\right)^2.$$

So
$$Var(\hat{\theta}) \geq \frac{1}{I(\theta)} = \frac{\sigma^2}{\sum_{k=1}^{n} \left(\frac{\partial}{\partial\theta} s_k(\theta)\right)^2}.$$

E.g. If $s_k(\theta) = \theta$, then $Var(\hat{\theta}) \geq \frac{\sigma^2}{n}$

If $s_k(\theta) = A\cos(\omega_0 k + \theta)$, then $Var(\hat{\theta}) \geq \frac{2\sigma^2}{nA^2}$.

3. Extension to Vector

CRLB (vector)

$$\boxed{\begin{array}{l}\text{Under assumption (1)–(2),} \\ \qquad \underline{\mathrm{Cov}}(\hat{\underline{\theta}}(\underline{Y})) - \underline{I}(\underline{\theta})^{-1} \succeq 0, \\ \text{for any unbiased estimator } \hat{\underline{\theta}}(\underline{Y}).\end{array}}$$

$$\underline{I}(\underline{\theta})_{ij} = -\mathbb{E}\left[\frac{\partial^2 \ln f_{\underline{\theta}}(\underline{y})}{\partial \theta_i \, \partial \theta_j} \right].$$

Example

$$Y_k = A + Bk + N_k, \quad N_k \sim N(0, \sigma^2).$$

Find CRLB for (A, B), for any unbiased estimator $\hat{\underline{\theta}} = (\hat{A}, \hat{B})$.

Solution:

$$\underline{I}(\underline{\theta}) = \begin{bmatrix} -\mathbb{E}\left[\frac{\partial^2 \ln f_{\underline{\theta}}(\underline{y})}{\partial A^2}\right] & -\mathbb{E}\left[\frac{\partial^2 \ln f_{\underline{\theta}}(\underline{y})}{\partial A \, \partial B}\right] \\ -\mathbb{E}\left[\frac{\partial^2 \ln f_{\underline{\theta}}(\underline{y})}{\partial B \, \partial A}\right] & -\mathbb{E}\left[\frac{\partial^2 \ln f_{\underline{\theta}}(\underline{y})}{\partial B^2}\right] \end{bmatrix}$$

$$\ln f_{\underline{\theta}}(\underline{y}) = -\frac{n}{2}\ln(2\pi\sigma^2) - \frac{1}{2\sigma^2}\sum_{k=1}^{n}(y_k - A - Bk)^2$$

$$\frac{\partial}{\partial A}\ln f_{\underline{\theta}}(\underline{y}) = \frac{1}{\sigma^2}\sum_{k=1}^{n}(y_k - A - Bk)$$

$$\frac{\partial}{\partial B}\ln f_{\underline{\theta}}(\underline{y}) = \frac{1}{\sigma^2}\sum_{k=1}^{n}k(y_k - A - Bk).$$

Therefore,

$$\frac{\partial^2}{\partial A^2} \ln f_{\underline{\theta}}(\underline{y}) = -\frac{n}{\sigma^2},$$

$$\frac{\partial^2}{\partial B^2} \ln f_{\underline{\theta}}(\underline{y}) = \frac{-1}{\sigma^2} \sum_{k=1}^{n} k^2$$

$$\frac{\partial^2}{\partial A \partial B} \ln f_{\underline{\theta}}(\underline{y}) = \frac{-1}{\sigma^2} \sum_{k=1}^{n} k .$$

So $$\underline{I}(\underline{\theta}) = \frac{1}{\sigma^2}\begin{bmatrix} n & \sum k \\ \sum k & \sum k^2 \end{bmatrix} = \frac{1}{\sigma^2}\begin{bmatrix} n & \frac{n(n-1)}{2} \\ \frac{n(n-1)}{2} & \frac{n(n-1)(2n-1)}{6} \end{bmatrix}$$

For any unbiased estimator $\hat{\theta} = (\hat{A}, \hat{B})$,

$$\underline{Cov}(\hat{\underline{\theta}}(\underline{Y})) \geq \underline{I}(\underline{\theta})^{-1}$$

$$= \sigma^2 \begin{bmatrix} \frac{2(n-1)}{n(n+1)} & -\frac{6}{n(n+1)} \\ -\frac{6}{n(n+1)} & \frac{12}{n(n^2-1)} \end{bmatrix}$$

Therefore, $$Var(\hat{A}) \geq \frac{2(2n-1)\sigma^2}{n(n+1)}$$

$$Var(\hat{B}) \geq \frac{12\sigma^2}{n(n^2-1)} .$$