MA 3140: Statistical Inference

Dr. Sameen Naqvi
Department of Mathematics, IIT Hyderabad
Email id: sameen@math.iith.ac.in

If we cannot find an estimator that attains the lower bound, we have to decide whether no estimator can attain it or whether we must look at more estimators?

Remark 4

- Recall that Cauchy-Schwartz Inequality was used to prove CR Inequality:

$$Cov^2\left[T(\boldsymbol{X}), \frac{\partial}{\partial\theta}\log f_\theta(\boldsymbol{X})\right] \leq Var(T(\boldsymbol{X}))\ Var\left[\frac{\partial}{\partial\theta}\log f_\theta(\boldsymbol{X})\right].$$

Now, note that Cauchy-Schwartz Inequality has a condition for equality also.

This holds true when $T(\boldsymbol{X})$ and $\frac{\partial}{\partial\theta}\log f_\theta(\boldsymbol{X})$ are linearly related.
Here, $S_\theta(\boldsymbol{x}) = \frac{\partial}{\partial\theta}\log f_\theta(\boldsymbol{x})$ is also referred to as the Score Function. For iid r.v.s, $S_\theta(\boldsymbol{x}) = \frac{\partial}{\partial\theta}\log f_\theta(\boldsymbol{x}) = \sum_{i=1}^{n}\frac{\partial}{\partial\theta}\log f_\theta(x_i)$.

Corollary (CRLB Attainment)

Let $X_1, \ldots, X_n$ be iid with common pdf $f_\theta(x)$. Suppose that the family $\{f_\theta : \theta \in \Theta\}$ satisfies the conditions of CFR Inequality. Then, equality holds if and only if for all $\theta \in \Theta$,

$$T(\boldsymbol{x}) - \psi(\theta) = k(\theta) \frac{\partial}{\partial \theta} \log f_\theta(\boldsymbol{x})$$

for some function $k(\theta)$; here, $\psi(\theta) = E[T(\boldsymbol{X})]$.

Example 4: Normal Distribution

Let $X_1, \ldots, X_n$ be a random sample from $N(\mu, \sigma^2)$.

Case 1: $\sigma^2 = \sigma_0^2$ (say) is known. Find the best UE for μ using the method of CRLB.

Solution:

$$
\begin{aligned}
f_\mu(x) &= \frac{1}{\sigma_0\sqrt{2\pi}} e^{-\frac{1}{2\sigma_0^2}(x-\mu)^2} \\
\log f_\mu(x) &= -\frac{1}{2}\log\sigma_0^2 - \frac{1}{2}\log 2\pi - \frac{(x-\mu)^2}{2\sigma_0^2} \\
\frac{\partial}{\partial\mu}\log f_\mu(x) &= \frac{x-\mu}{\sigma_0^2}
\end{aligned}
$$

Example 4: Normal Distribution cont'd

Now,

$$E\left[\frac{\partial}{\partial\mu}\log f_\mu(X)\right]^2 = E\left[\frac{(X-\mu)^2}{\sigma_0^4}\right] = \frac{\sigma_0^2}{\sigma_0^4} = \frac{1}{\sigma_0^2}.$$

So, $I(\mu) = \frac{n}{\sigma_0^2}$, and CRLB for variance of an UE of μ is $\frac{\sigma_0^2}{n}$.

We also know that $E(\overline{X}) = \mu$ and $Var(\overline{X}) = \frac{\sigma_0^2}{n}$.

Thus, $\overline{X}$ is the is the best UE of μ.

Example 4: Normal Distribution cont'd

Case 2: $\mu = \mu_0$ (say) is known. Find the best UE for σ^2 using the method of CRLB.

Solution:

$$
\begin{aligned}
f_{\sigma^2}(x) &= \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2\sigma^2}(x-\mu_0)^2} \\
\log f_{\sigma^2}(x) &= -\frac{1}{2}\log\sigma^2 - \frac{1}{2}\log 2\pi - \frac{(x-\mu_0)^2}{2\sigma^2} \\
\frac{\partial}{\partial\sigma^2}\log f_{\sigma^2}(x) &= -\frac{1}{2\sigma^2} + \frac{(x-\mu_0)^2}{2\sigma^4} \\
&= \frac{1}{2\sigma^2}\left[\frac{(x-\mu_0)^2}{\sigma^2} - 1\right]
\end{aligned}
$$

Example 4: Normal Distribution cont'd

Now,

$$\begin{aligned} E\left[\frac{\partial}{\partial\sigma^2}\log f_{\sigma^2}(X)\right]^2 &= \frac{1}{4\sigma^4}E\left[\left(\frac{X-\mu_0}{\sigma}\right)^2-1\right]^2 \\ &= \frac{1}{4\sigma^4}Var\left(\frac{X-\mu_0}{\sigma}\right)^2 \\ &= \frac{1}{4\sigma^4}\times 2 = \frac{1}{2\sigma^4} \end{aligned}$$

The above result holds because $(\frac{X-\mu_0}{\sigma})^2 \sim \chi^2_1$.

So, $I(\sigma^2) = \frac{n}{2\sigma^4}$, and CRLB for variance of an UE of σ^2 is $\frac{2\sigma^4}{n}$.

Example 4: Normal Distribution cont'd

- Note that S^2 does not attain CRLB because $Var(S^2) = \frac{2\sigma^4}{n-1}$.
- Is there a better UE of σ^2 than S^2? Is the CRLB attainable?
- In order to check this, let us consider

$$\frac{\partial}{\partial \sigma^2} \log f_{\sigma^2}(\boldsymbol{x}) = \frac{n}{2\sigma^4}\left[\sum_{i=1}^{n} \frac{(x_i - \mu_0)^2}{n} - \sigma^2\right].$$

Thus taking $k(\sigma^2) = \frac{n}{2\sigma^4}$ shows that the best UE of σ^2 is $\frac{1}{n}\sum_{i=1}^{n}(X_i - \mu_0)^2$.

Example 4: Normal Distribution cont'd

- **Note:** The above conclusion can also be obtained from the following facts:

  Consider $T = \frac{1}{n}\sum_{i=1}^{n}(X_i - \mu_0)^2$ which is unbiased for σ^2.

  This holds true as $\sum_{i=1}^{n}\left(\frac{X_i-\mu_0}{\sigma}\right)^2 \sim \chi^2_n$, thereby giving $E(\frac{nT}{\sigma^2}) = n$ and $Var(\frac{nT}{\sigma^2}) = 2n$.

  Thus, $Var(T) = \frac{2\sigma^4}{n}$, and $T = \frac{1}{n}\sum_{i=1}^{n}(X_i - \mu_0)^2$ is the best UE of σ^2.

What are the parametric functions for which this lower bound is attained?

If the CRLB for the variance of an UE of $g(\theta)$ is attained, then the class of parametric functions, for which the UEs attain CRLB, is the class of linear functions of $g(\theta)$.

Example: Poisson

- Let $X_1, \ldots, X_n \sim P(\lambda)$, $\lambda > 0$. Suppose you are interested to estimate $g(\lambda) = \lambda^2$.
- **Solution:**

$$\text{CRLB for variance of UE of } g(\lambda) = [g^{'}(\lambda)]^2 \times \text{ CRLB for } \lambda$$
$$= 4\lambda^2 \times \frac{\lambda}{n} = \frac{4\lambda^3}{n}.$$

Let $Y = \sum X_i \sim P(n\lambda)$. Define $U = \frac{1}{n^2} Y(Y-1)$.
Then,

$$E(U) = \frac{1}{n^2}(EY^2 - EY) = \frac{1}{n^2}(n\lambda + n^2\lambda^2 - n\lambda) = \lambda^2.$$

and

$$Var(U) = \frac{4\lambda^3}{n} + \frac{2\lambda^2}{n^2} > \frac{4\lambda^3}{n}.$$

Hence, CRLB is not attained.

Efficiency of Estimators

- Let T_1 and T_2 be two UEs of $g(\theta)$ such that $E(T_1^2) < \infty$ and $E(T_2^2) < \infty$. The efficiency of T_1 relative to T_2 is defined as:

$$Ef(T_2|T_1) = \frac{Var(T_2)}{Var(T_1)}.$$

  We say that T_2 is more efficient than T_1 if $Ef(T_2|T_1) < 1$.

- Efficiency of an UE can also be defined w.r.t. CRLB, i.e.,

$$Ef(T) = \frac{Var(T)}{CRLB}.$$

  We say that T is most efficient if $Ef(T) = 1$.

Example 1: Poisson Distribution

- Let $X \sim P(\lambda)$. Suppose you are interested to estimate $P(X=0) = e^{-\lambda} = g(\lambda)$.

  Solution:

  CRLB for $e^{-\lambda} = [g^{'}(\lambda)]^2 \times$ CRLB for $\lambda = e^{-2\lambda}\lambda$.

  Consider an estimator

$$\beta(X) = \begin{cases} 1 & \text{if } X = 0 \\ 0 & \text{if } X = 1, 2, \ldots \end{cases}$$

  Here,

$$E[\beta(X)] = 1.P(X=0) + 0\sum_{i=1}^{\infty} P(X=i) = e^{-\lambda}.$$

  Thus, $\beta(X)$ is unbiased for $e^{-\lambda}$.

Example 1: Poisson Distribution contd.

It can be easily seen that

$$E[\beta^2(X)] = e^{-\lambda}, \qquad Var[\beta(X)] = e^{-\lambda} - e^{-2\lambda}.$$

Note that

$$Var[\beta(X)] = e^{-\lambda} - e^{-2\lambda} > \lambda e^{-2\lambda} = \text{CRLB for } e^{-\lambda}$$

because $e^{\lambda} > 1 + \lambda$, $\lambda > 0$ is always true.

Although CRLB is not attained, it can be shown that β is the only UE of $e^{-\lambda}$, and hence, it is the most efficient estimator.

Example 2

Let $X_1, \ldots, X_n$ be iid r.v. with mean μ and variance σ^2 $(< \infty)$. Consider the following two estimators

$$T_1 = \overline{X}, \qquad T_2 = \frac{2}{n(n+1)} \sum_{i=1}^{n} iX_i$$

and compare their efficiencies.

Solution:

Consider T_1. Here,

$$E(T_1) = \mu, \quad Var(T_1) = \frac{\sigma^2}{n}.$$

So, T_1 is unbiased for μ.

Example 2 contd.

Consider T_2. Here,

$$E(T_2) = \frac{2}{n(n+1)} \sum_{i=1}^{n} i\mu = \mu$$

and

$$Var(T_2) = \frac{4}{n^2(n+1)^2} \sum_{i=1}^{n} i^2\sigma^2 = \frac{2}{3}\frac{2n+1}{n(n+1)}\sigma^2.$$

So, T_2 is unbiased for μ.
Now,

$$Ef(T_2|T_1) = \frac{Var(T_2)}{Var(T_1)} = \frac{2}{3}\frac{2n+1}{n+1} > 1 \text{ for } \; n > 1.$$

In general, T_1 is more efficient than T_2.

Thanks for your patience!