UNIT - III

Cramer Rao Inequality

Statement:

If t is an unbiased estimator for $\gamma(\theta)$ a function of parameter θ then

$$\text{var}(t) \geq \frac{\left[\frac{d}{d\theta}\gamma(\theta)\right]^2}{E\left(\frac{\partial}{\partial\theta}\log L\right)^2}$$

$$= \frac{[\gamma'(\theta)]^2}{I(\theta)}$$

where $I(\theta)$ is the information on θ supplied by the sample.

In otherwords Cramer-Rao inequality provides a lower bound $(\gamma'(\theta))^2 / I(\theta)$ to the variance of an unbiased estimator of $\gamma(\theta)$

Proof:

The assumptions made are known as Regularity conditions for Cramer-Rao inequality

1. The parameter space Θ is a non-degenerate open interval on the real line R $(-\infty, \infty)$

2. For almost all $x = (x_1, x_2 \ldots x_n)$ and $\forall\ \theta \in \Theta$ $\frac{\partial}{\partial\theta}L(x,\theta)$ exist. The exceptional set if any, is independent of θ

3. The range of integeration is independent of the parameter θ, so that $f(x,\theta)$ is differentiable under integeral sign.

4. The conditions of uniform convergence of integerals are satisfied so that differentiation under the integeral sign is valid

5. $I(\theta) = E\left[\left\{\frac{\partial}{\partial\theta} \log L(x,\theta)\right\}^2\right]$ exists and +ve $\forall$ $\theta \in \Theta$

Let X be a r.v. following the p.d.f $f(x,\theta)$. L be the likelihood function of the random sample $(x_1, x_2 \ldots x_n)$ from this popn. Then

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

L is the joint p.d.f of $x_1, x_2 \ldots x_n$

$$\int L(x,\theta)\, dx = 1$$

where $\int dx = \int\int \cdots \int dx_1\, dx_2 \ldots dx_n$

differentiating w.r.t θ and using the regularity conditions we get

$$\int \frac{\partial}{\partial\theta} L\, dx \Rightarrow \int \left(\frac{\partial}{\partial\theta} \log L\right) L\, dx = 0$$

$$= E\left(\frac{\partial}{\partial\theta} \log L\right) = 0 \quad \text{——— (1)}$$

Let $t = t(x_1, x_2 \ldots x_n)$ be an UBE of $\gamma(\theta)$ such that

$$E(t) = \gamma(\theta)$$

$$\Rightarrow \int t\, L\, dx = \gamma(\theta) \quad \text{——— (2)}$$

differenting w.r.t θ we get

$$\int t \frac{\partial L}{\partial \theta} dx = \gamma'(\theta)$$

$$\Rightarrow \int t\left(\frac{\partial}{\partial \theta} \operatorname{Log} L\right) L\, dx = \gamma'(\theta)$$

$$\Rightarrow E\left(t \frac{\partial}{\partial \theta} \operatorname{Log} L\right) = \gamma'(\theta) \quad \text{--- (3)}$$

$$\operatorname{cov}\left(t, \frac{\partial}{\partial \theta} \operatorname{Log} L\right) = E\left(t \cdot \frac{\partial}{\partial \theta} \operatorname{Log} L\right) - E(t) \cdot E\left(\frac{\partial}{\partial \theta} \operatorname{Log} L\right)$$

$$= \gamma'(\theta) \quad \text{--- (4)}$$

as $E\left(\frac{\partial}{\partial \theta} \operatorname{Log} L\right) = 0$ and using (3)

We have

$$\{r(x, y)\}^2 \leq 1$$

$$[\operatorname{cov}(x, y)]^2 \leq \operatorname{Var}(x) \operatorname{Var}(y)$$

$$\therefore \operatorname{cov}\left(t, \frac{\partial}{\partial \theta} \operatorname{Log} L\right)^2 \leq \operatorname{Var} t \ \operatorname{var}\left(\frac{\partial}{\partial \theta} \operatorname{Log} L\right)$$

$$\gamma'(\theta)^2 \leq \operatorname{Var} t \left[E\left(\frac{\partial}{\partial \theta} \operatorname{Log} L\right)\right]^2 - \left(E\left(\frac{\partial}{\partial \theta} \operatorname{Log} L\right)\right)^2$$

$$\Rightarrow \gamma'(\theta)^2 \leq \operatorname{Var} t \ E\left[\left(\frac{\partial}{\partial \theta} \operatorname{Log} L\right)^2\right]$$

$$\Rightarrow \operatorname{Var} t \geq \frac{[\gamma(\theta)]^2}{E\left[\left(\frac{\partial}{\partial \theta} \operatorname{Log} L\right)^2\right]}$$

which is the cramer-Rao inequality

Remarks:

If t is an UBE of parameter θ then

$E(t) = \theta \Rightarrow \gamma(\theta) = \theta$ or $\gamma'(\theta) = 1$

then

$$v(t) \geq \frac{1}{E\left[\left(\frac{\partial}{\partial\theta}\log L\right)^2\right]}$$

$$= \frac{1}{I(\theta)}$$

$I(\theta)$ is called as amount of information on θ supplied by the sample $x_1, x_2 \ldots x_n$

and

$\frac{1}{I(\theta)}$ is the information limit to the variance of estimator $t = t(x_1, x_2 \ldots x_n)$.

Then an UBE t of $\gamma(\theta)$ from which Cramer-Rao lower bound obtained is called minimum variance bound estimator (MVB) estimator

Condition for equality sign in Cramer-Rao inequality

In proving C.R inequality

$$\mathrm{var}(t) \geq \frac{\gamma'(\theta)^2}{I(\theta)} \quad \text{—— (1)}$$

where

$$I(\theta) = E\left(\frac{\partial}{\partial\theta}\log L\right)^2$$

We used

$$\gamma'(\theta)^2 \leq E[t - \gamma(\theta)]^2 \cdot E\left(\frac{\partial}{\partial\theta}\log L\right)^2 \quad \text{—— (2)}$$

The sign of inequality (1) will holds good only when sign of inequality holds good is (2)

ie if and only if the variables $(t-\gamma(\theta))$ and $\left(\frac{\partial}{\partial\theta}\log L\right)$ are proportional to each other

ie $$\frac{t-\gamma(\theta)}{\frac{\partial}{\partial\theta}\log L} = \lambda = \lambda(\theta)$$

where λ is constant and independent of $x_1, x_2 \ldots x_n$ but may depend on θ

$$\frac{\partial}{\partial\theta}\log L = \frac{t-\gamma(\theta)}{\lambda(\theta)} = [t-\gamma(\theta)]\,A(\theta) \quad \text{--- (3)}$$

where $A = A(\theta) = 1/\lambda(\theta)$

Is the necessary and sufficient condition for an MBE t to attain the lower bound of its variance. its given in (3).

Now CR minvariance bound is

$$\text{Var}(t) = [\gamma'(\theta)]^2 \Big/ E\left(\frac{\partial}{\partial\theta}\log L\right)^2 \quad \text{--- (4)}$$

But we know

$$E\left(\frac{\partial}{\partial\theta}\log L\right)^2 = E\left[A(\theta)\,(t-\gamma(\theta))\right]^2$$

$$= [A(\theta)]^2 \, E(t - \gamma(\theta))^2$$

$$E\left(\frac{\partial}{\partial\theta}\log L\right)^2 = [A(\theta)]^2 \, \mathrm{Var}\, t \quad \text{—— (5)}$$

substituting (5) in (4).

$$\mathrm{Var}(t) = \frac{[\gamma'(\theta)]^2}{A(\theta)^2 \, \mathrm{Var}(t)} \Rightarrow (\mathrm{Var}(t))^2 = \frac{[\gamma(\theta)]^2}{[(A(\theta))]^2}$$

$$\Rightarrow \mathrm{Var}\, t = \left|\frac{\gamma'(\theta)}{A(\theta)}\right|$$

$$= |\gamma'(\theta) \, \lambda(\theta)|$$

If the likelihood function L is expressible in the form of eqn (3) then

i) t is UBE of $\gamma(\theta)$

ii) MVB estimator t for $\gamma(\theta)$ exist and

iii) $\mathrm{Var}(t) = \dfrac{\gamma'(\theta)}{A(\theta)} = |\gamma'(\theta)\lambda(\theta)|$

1. Obtain the MVB estimator for μ in Normal population $N(\mu, \sigma^2)$ where σ^2 is known

If $x_1, x_2 \cdots x_n$ is a random sample of size n from the normal population

then

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

$$= \left(\frac{1}{\sqrt{2\pi}\sigma}\right)^n e^{-\left[\sum_{i=1}^{n} (x_i - \mu)^2 / 2\sigma^2\right]}$$

$$\log L = -n \log \sqrt{2\pi}\sigma - \frac{1}{2\sigma^2} \sum_{i=1}^{n} (x_i - \mu)^2$$

$$= k - \frac{1}{2\sigma^2} \sum_{i=1}^{n} (x_i - \mu)^2$$

$$\frac{\partial \log L}{\partial \mu} = -\frac{1}{2\sigma^2} \sum_{i=1}^{n} 2(x_i - \mu)(-1)$$

$$= \frac{\sum_{i=1}^{n} (x_i - \mu)}{\sigma^2}$$

$$= \frac{\sum x_i - n\mu}{\sigma^2}$$

$$= \frac{n\bar{x} - n\mu}{\sigma^2}$$

$$= \frac{(\bar{x} - \mu)}{\sigma^2/n}$$

which is in the form of

$$\frac{\partial}{\partial \theta} \log L = [t - \gamma(\theta)] \frac{1}{\lambda(\theta)}$$

Hence $\bar{x}$ is MVB unbiased estimator of μ and

$$V(\hat{\mu}) = V(\bar{x}) = \sigma^2/n$$

2. Find if MVB estimator exists for θ in the Cauchy's population.

$$dF(x,\theta) = \frac{1}{\pi} \frac{1}{1+(x-\theta)^2}, \quad -\infty < x < \infty$$

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

$$= \left(\frac{1}{\pi}\right)^n \prod_{i=1}^{n} \left(\frac{1}{1+(x_i-\theta)^2}$$

$$\log L = -n \log \pi - \sum_{i=1}^{n} \log\left(1+(x_i-\theta)^2\right)$$

$$\frac{\partial \log L}{\partial \theta} = 0 + \sum_{i=} \left(\frac{1 \sum_{i=1}^{n} 2(x_i-\theta)}{1+(x_i-\theta)^2}\right)$$

$$= 2 \sum_{i=1}^{n} \frac{(x_i-\theta)}{1+(x_i-\theta)^2}$$

This cannot be expressed in the form of

$$\frac{\partial \log L}{\partial \theta} = [t - \gamma(\theta)] \cdot \frac{1}{\lambda(\theta)}$$

MVB estimator does not exist for Cauchy's population.

So Cramer Rao bound is not attainable by the variance of any unbiased estimator θ.

3. A random sample $x_1, x_2 \dots x_n$ is taken from Normal population with mean 0 and variance σ^2. Examine if $\sum_{i=1}^{n} x_i^2/n$ is MVB estimator for σ^2

Solution

$$X \sim N(0, \sigma^2)$$

$$f(x, \sigma^2) = \frac{1}{\sqrt{2\pi}\,\sigma} e^{-\left(\frac{x^2}{2\sigma^2}\right)} \qquad -\infty < x < \infty$$

$$L = \prod_{i=1}^{n} f(x_i, \sigma^2)$$

$$L = \left(\frac{1}{\sqrt{2\pi}\,\sigma}\right)^n e^{-\sum_{i=1}^{n} x_i^2/2\sigma^2}$$

$$\cancel{\log L = -\frac{n}{2}\log(2\pi) + \frac{n}{2}\log\sigma^2}$$

$$\Rightarrow \cancel{\log} L = \left(\frac{1}{2\pi\sigma^2}\right)^{n/2} e^{\left[-\sum_{i=1}^{n}\left(x_i^2/2\sigma^2\right)\right]}$$

$$\log L = -\frac{n}{2}\log 2\pi - \frac{n}{2}\log\sigma^2 - \frac{n}{2}\log\sigma^2 - \frac{1}{2\sigma^2}\sum_{i=1}^{n} x_i^2$$

$$\frac{\partial \log L}{\partial \sigma^2} = -\frac{n}{2\sigma^2} + \frac{1}{2\sigma^4}\sum_{i=1}^{n} x_i^2$$

$$= \frac{\sum_{i=1}^{n} x_i^2 - n\sigma^2}{2\sigma^4}$$

$$= \frac{\left(\sum_{i=1}^{n} x_i^2/n\right) - \sigma^2}{2\sigma^4/n}$$

which is ~~the~~ in the form of

$$\frac{\partial}{\partial\theta} \log L = [t - \gamma(\theta)] \frac{1}{\lambda(\theta)}$$

hence $\hat{\sigma}^2 = \sum_{i=1}^{n} x_i^2/n$ is MVB estimator

$$V(\hat{\sigma}^2) = \frac{2\sigma^4}{n}$$

4. Show that $\bar{X} = \sum_{i=1}^{n} X_i/n$ is random sampling from

$$f(x,\theta) = \frac{1}{\theta} e^{-x/\theta}, \quad 0 < x < \infty$$

$$0 \quad \text{otherwise}$$

where $0 < \theta < \infty$ is MVB estimator of θ and variance θ^2/n

Solution:

Let $x_1, x_2 \dots x_n$ be a random sample of ~~size~~ n from population with p.d.f

$$f(x,\theta) = \frac{1}{\theta} e^{-x/\theta} \quad 0 < x < \infty$$

$$L = \prod_{i=1}^{n} f(x_i,\theta) = \left(\frac{1}{\theta}\right)^n e^{-\left(\sum_{i=1}^{n} x_i/\theta\right)}$$

$$\log L = -n\log\theta - \frac{1}{\theta}\sum_{i=1}^{n} x_i$$

$$\frac{\partial \log L}{\partial \theta} = -\frac{n}{\theta} + \frac{1}{\theta^2}\sum_{i=1}^{n} x_i$$

$$= \frac{\sum_{i=1}^{n} x_i - n\theta}{\theta^2}$$

$$= \frac{\bar{x} - \theta}{(\theta^2/n)}$$

$$= (\bar{x} - \theta)\frac{1}{\lambda(\theta)}$$

which is in the form of

$$t - \gamma(\theta)\frac{1}{\lambda(\theta)}$$

hence $\bar{x}$ is MVB estimator of θ and

$$\text{var}(\bar{x}) = \lambda(\theta) = \frac{\theta^2}{n}$$

5. If $x_1, x_2 \dots x_n$ is a r.s from a popn with density $f(x, \theta) = \theta e^{-\theta x}$ $0 < x < \infty$ Examine whether MVBE exist for parameter θ

Solution

given $f(x, \theta) = \theta e^{-\theta x}$ $0 < x < \infty$

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

$$= \theta^n e^{-\theta \sum_{i=1}^{n} x_i}$$

$$\log L = \log\left(\theta^n e^{-\theta \sum x_i}\right)$$

$= n \log \theta - \theta \sum x_i$

$$\frac{\partial \log L}{\partial \theta} = \frac{n}{\theta} - \sum x_i$$

$$\hat{\theta} = \frac{\sum x_i}{n} = \bar{x}$$

$$\frac{\partial \log L}{\partial \theta} = (t - \gamma(\theta)) A(\theta)$$

$$= \left(\frac{n}{\theta} - n\bar{x}\right)$$

$$= -n\left(\bar{x} - 1/\theta\right)$$

$\therefore$ $\bar{x}$ is MVBE of $1/\theta = \gamma(\theta)$

Rao Blackwell theorem

Let X and Y be random variables such that

i) $E(Y) = \mu$

ii) $V(Y) = \sigma^2_Y > 0$

Let $E(Y/X = x) = \phi(x)$

Then $E(\phi(X)) = \mu$ and

$V[\phi(X)] \leq Var(Y)$.

Proof: Let

$f_{XY}(x, y)$ be the joint p.d.f of r.v's X and Y

$f_1(\cdot)$ and $f_2(\cdot)$ are marginal p.d.f of X and Y respectively.

$h(y/x)$ be the conditional p.d.f of Y given $X = x$ such that

$$h(y/x) = f(x,y) / f_1(x).$$

$$E(Y/X=x) = \int_{-\infty}^{\infty} y \cdot h(y/x)\, dy$$

$$= \int_{-\infty}^{\infty} y \frac{f(x,y)}{f_1(x)}\, dy$$

$$= \frac{1}{f_1(x)} \int_{-\infty}^{\infty} y\, f(x,y)\, dy$$

$$\frac{1}{f_1(x)} \int_{-\infty}^{\infty} y\, f(x,y)\, dy = \phi(x)$$

$$= \phi(x)\, \cancel{f_1(x)} \longrightarrow (A)$$

$$\Rightarrow \int_{-\infty}^{\infty} y\, f(x,y)\, dy = \phi(x)\, f_1(x) \quad \text{—} (1).$$

Hence the conditional distn of Y given $X = x$ ~~depe~~ does not depend of parameter μ.

Hence X is sufficient statistic for μ.

Also

$$E[\phi(X)] = E[E(Y/X=x)]$$

$$= E(Y) = \mu.$$

Which proves part i) of the theorem

$$Var(Y) = E[Y - E(Y)]^2$$

$$= E[Y - \mu]^2$$

$$= E[Y - \phi(X) + \phi(X) - \mu]^2$$

$$= E[Y - \phi(X)]^2 + E[\phi(X) - \mu]^2 + 2E(Y - \phi(x))(\phi(x) - \mu) \quad \text{(2)}$$

Now consider the product term

$$E(Y - \phi(x))(\phi(x) - \mu) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (y - \phi(x))(\phi(x) - \mu) f(x, y)\, dx\, dy$$

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (y - \phi(x))(\phi(x) - \mu) f_1(x)\, h(y/x)\, dx\, dy$$

$$= \int_{-\infty}^{\infty} (\phi(x) - \mu)\left[\int_{-\infty}^{\infty} (y - \phi(x))\, h(y/x)\, dy\right] dx$$

since $E(Y/X = x) = \phi(x)$

but $\int_{-\infty}^{\infty} (y - \phi(x))\, h(y/x)\, dx = 0$

$$\therefore Var(y) = E[Y - \phi(x)]^2 + E[\phi(X) - \mu]^2 + 0$$

$$Var(y) = E[Y - \phi(X)]^2 + Var[\phi(X)]$$

$Var(Y) \geq Var[\phi(X)]$ as

$E(Y - \phi(X))^2 \geq 0$

$$\therefore Var\, \phi(X) \leq Var\, Y$$

Hence the theorem

Rao Blackwell theorem enables us to obtain MVU estimators:

* If a sufficient estimator exists for a parameter then in our search of MVUB we may restrict to the function of sufficient statistic

It is slightly different.

Let $U = U(x_1, x_2 \ldots x_n)$ be UBE of $\gamma(\theta)$ and $T = T(x_1, x_2 \ldots x_n)$ be suff stat for $\gamma(\theta)$ consider the function $\phi(T)$ of the sufficient statistic as

$$\phi(t) = E[U/T=t]$$

which is independent of θ since T is sufficient for θ

then

$$E[\phi(T)] = \gamma(\theta) \quad \text{and}$$

$$var(\phi(T)) \leq var(U)$$

This result implies

1. Starting with UBE U we can improve upon by defining $\phi(T)$ of the sufficient statistics.

2. This technique of obtaining improved estimators is called Blackwellisation