CRAMER-RAO INEQUALITY

CRAMER-RAO INEQUALITY
Syllabus:

- Statement and proof of Cramer-Rao inequality
- Definition of minimum variance bound unbiased estimator (MVBUE) of $f(\theta)$.
- Proof of the following results .
- 1) If MVBUE exists for θ , then MVBUE exists for $f(\theta)$, provided f is linear function.
- 2) If T is MVBUE for $f(\theta)$, then T is sufficient for $f(\theta)$.

CRAMER-RAO INEQUALITY
Syllabus ....Continue

- If T is an unbiased estimator of θ satisfying regularity condition, then T satisfies the relation

$$Var(T) \geq \frac{1}{nI(\theta)} \; ; \quad where \quad I(\theta) \quad is \quad the \quad fisher \quad informatio \quad n$$

- Examples and problems

Comparing estimators :

- Unbiased estimators can be compared via their variances and biased estimators by comparing mean square errors. (Generally unbiased estimators are preferable.)
- We can compare several unbiased estimators and find which one has smallest variance, but this does not allow us to tell if an estimator has the smallest variance amongst all unbiased estimators
- The Cramer-Rao lower bound provides a uniform lower bound on the variance of all unbiased estimators of $f = g(\theta)$.
- So if the variance of an unbiased estimator is equal to the Cramer-Rao lower bound it must have minimum variance amongst all unbiased estimators (so is said to be a minimum variance unbiased estimator of f).

CRAMER-RAO INEQUALITY : Statement

- Let $X_1, X_2, \ldots.. X_n$ be a random sample from $f(x;\theta)$, $\theta \in \Theta$. Assume Θ *is a subset of the real line. Let* $T = u(X_1, X_2, \ldots.. X_n)$ be an unbiased estimator of $k(\theta)$. Assume $f(x;\theta)$ satisfies the regularity conditions, then

$$Var(T) \geq \frac{\left(k'(\theta)\right)^2}{nE\left[\frac{\partial}{\partial\theta}\ln f(x;\theta)\right]^2}$$

Regularity Conditions:

1) $$\frac{\partial}{\partial\theta} \ln f(x;\theta) \text{ exists} \quad \text{for all } x \text{ and all } \theta$$

2) $$\frac{\partial}{\partial\theta}\int \cdots\cdots \int \prod_{1}^{n} f(x_i;\theta)\,dx_1\,dx_2 \cdots\cdots dx_n .$$

$$= \int \cdots\cdots \int \frac{\partial}{\partial\theta} \prod_{1}^{n} f(x_i;\theta)\,dx_1\,dx_2 \cdots\cdots dx_n .$$

3) $$\frac{\partial}{\partial\theta}\int \cdots\cdots \int t(x_1, x_2, \ldots\ldots x_n) \prod_{1}^{n} f(x_i;\theta)\,dx_1\,dx_2 \cdots\cdots dx_n .$$

$$= \int \cdots\cdots \int t(x_1, x_2, \ldots\ldots x_n) \frac{\partial}{\partial\theta} \prod_{1}^{n} f(x_i;\theta)\,dx_1\,dx_2 \cdots\cdots dx_n .$$

4) $$0 < E\left[\frac{\partial}{\partial\theta} \ln f(x;\theta)\right]^2 < \infty \quad \text{for all } \theta \quad \text{in } \Theta$$

Proof : We have $\frac{\partial}{\partial\theta}\left[f(x_1;\theta).......f(x_n;\theta)\right]$

$$=\left[\frac{\partial}{\partial\theta}f(x_1;\theta)\right]\prod_{j\neq 1}f(x_j;\theta)+\left[\frac{\partial}{\partial\theta}f(x_2;\theta)\right]\prod_{j\neq 2}f(x_j;\theta)$$

$$+..........+\left[\frac{\partial}{\partial\theta}f(x_n;\theta)\right]\prod_{j\neq n}f(x_j;\theta)$$

$$=\left[\frac{1}{f(x_1;\theta)}\frac{\partial}{\partial\theta}f(x_1;\theta)\right]f(x_1;\theta)\prod_{j\neq 1}f(x_j;\theta)+\left[\frac{1}{f(x_2;\theta)}\frac{\partial}{\partial\theta}f(x_2;\theta)\right]f(x_2;\theta)\prod_{j\neq 2}f(x_j;\theta)$$

$$+..........+\left[\frac{1}{f(x_n;\theta)}\frac{\partial}{\partial\theta}f(x_n;\theta)\right]f(x_n;\theta)\prod_{j\neq n}f(x_j;\theta)$$

$$=\left[\frac{\partial}{\partial\theta}\ln f(x_1;\theta)\right]\prod_{j=1}f(x_j;\theta)+\left[\frac{\partial}{\partial\theta}\ln f(x_2;\theta)\right]\prod_{j=1}f(x_j;\theta)$$

$$+..........+\left[\frac{\partial}{\partial\theta}\ln f(x_n;\theta)\right]\prod_{j\neq 1}f(x_j;\theta)$$

$$=\left[\sum_{j=1}^{n}\frac{\partial}{\partial\theta}\ln f(x_j;\theta)\right]\prod_{j=1}^{n}f(x_j;\theta)$$

Let T=u($X_1, X_2, \ldots X_n$) be an unbiased estimator of k(θ) so

$$E(T) = E\left[u\left(X_1, X_2, \ldots X_n\right)\right] = k(\theta)$$

That is in the continuous case

$$k(\theta) = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} u(x_1, x_2, \ldots, x_n) f(x_1;\theta) \ldots f(x_n;\theta)\, dx_1 dx_2 \ldots dx_n .$$

$$k'(\theta) = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} u(x_1, x_2, \ldots, x_n) \frac{\partial}{\partial \theta} f(x_1;\theta) \ldots f(x_n;\theta)\, dx_1 dx_2 \ldots dx_n .$$

$$k'(\theta) = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} u(x_1, x_2, \ldots, x_n) \left[\frac{\partial}{\partial \theta} f(x_1;\theta) \ldots f(x_n;\theta)\right] dx_1 dx_2 \ldots dx_n .$$

$$k'(\theta) = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} u(x_1, x_2, \ldots, x_n) \left[\sum_{j=1}^{n} \frac{\partial}{\partial \theta} \ln f(x_j;\theta)\right]\left[\prod_{j=1}^{n} f(x_j;\theta)\right] dx_1 dx_2 \ldots dx_n .$$

Define the random variable Z by $\left[\sum_{j=1}^{n} \frac{\partial}{\partial \theta} \ln f(x_j;\theta) \right]$,

so $k'(\theta) = E(TZ)$

Now $\int_{-\infty}^{\infty} f(x;\theta)\,dx = 1,$ *taking derivative w.r.t.* θ *we get* $\int_{-\infty}^{\infty} \frac{\partial}{\partial \theta} f(x;\theta) = 0$

The latter expression can be written as $\int_{-\infty}^{\infty} \frac{\frac{\partial f(x;\theta)}{\partial \theta}}{f(x;\theta)} f(x;\theta)\,dx = 0$

$$\Rightarrow \int_{-\infty}^{\infty} \frac{\partial \ln f(x;\theta)}{\partial \theta} f(x;\theta)\,dx = 0$$

According to above we have

$$E(Z) = E\left[\sum_{1}^{n} \frac{\partial \ln f(x_j;\theta)}{\partial \theta} \right] = \sum_{1}^{n} E\left[\frac{\partial \ln f(x_j;\theta)}{\partial \theta} \right] = 0$$

Moreover, Z is the sum of n indep. r.v. each with mean zero and consequently with variance $E\left\{\left[\frac{\partial \ln f(x;\theta)}{\partial \theta}\right]^2\right\}$

Hence variance of Z is the sum of the n variances,

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

Recall that $E(TZ) = E(T)E(Z) + \rho\sigma_T\sigma_Z$

Since $E(TZ) = k'(\theta), E(T) = k(\theta), \; E(Z) = 0$

We have $k'(\theta) = k(\theta).0 + \rho\sigma_T\sigma_Z$ or $\rho = \dfrac{k'(\theta)}{\sigma_T\sigma_Z}$

Now $\rho^2 \le 1$, because $-1 \le \rho \le 1$, Hence $\dfrac{[k'(\theta)]^2}{Var(T)Var(Z)} \le 1$

or $\dfrac{[k'(\theta)]^2}{Var(Z)} \le Var(T)$, i.e. we have $Var(T) \ge \dfrac{[k'(\theta)]^2}{nE\left[\left(\frac{\partial \ln f(x;\theta}{\partial \theta}\right)^2\right]}$

Definition : *The expression* $E\left[\left(\frac{\partial \ln f(x;\theta)}{\partial \theta}\right)^2\right]$, *denoted by* $I(\theta)$ *is called fisher's information (about* θ*). That is* $I(\theta) = \int_{-\infty}^{\infty}\left(\frac{\partial \ln f(x;\theta)}{\partial \theta}\right)^2 f(x;\theta)dx$.

$I(\theta)$ *can be computed from* $I(\theta) = -\int_{-\infty}^{\infty}\frac{\partial^2 \ln f(x;\theta)}{\partial \theta^2} f(x;\theta)dx$.

Sometimes one expression is easier to compute than the other, but often we prefer the second expression.

Fisher information in a random sample of size n is n times the fisher information in one observation. i.e. $I_n(\theta) = nI(\theta)$

So Cramer – Rao Inequality becomes $Var(T) \geq \frac{[k'(\theta)]^2}{nI(\theta)}$

If T is an unbised estimator of θ*, so that* $k(\theta) = \theta$*, then the Cramer – Rao inequality becomes, since* $k'(\theta) = 1$, $Var(T) \geq \frac{1}{nI(\theta)}$

Minimum Variance Bound Unbiased Estimator (MVBUE):

An unbiased estimator T of a parameter k(θ) is called an MVBUE if Var(T) attains Cramer-Rao lower bound.

In other words :

1) *An estimator* T *of* $k(\theta)$ *is said to be MVBUE if* $E(T) = k(\theta)$ *and*

$$Var(T) = \frac{[k'(\theta)]^2}{nI(\theta)}$$

2) *An estimator* T *of* θ *is said to be MVBUE if* $E(T) = \theta$ *and*

$$Var(T) = \frac{1}{nI(\theta)}$$

Criterion for the MVBUE .

If T *is MVBUE of the parameter* $k(\theta)$ *, then it must satisfy the relation*

$$\frac{\partial}{\partial\theta}\ln L = A(\theta)[T - k(\theta),\ ], \qquad \left[Var(T) = \frac{k'(\theta)}{A(\theta)}\right]$$

This is therefore the necessary and sufficient condition for an estimator T *of parameter* $k(\theta)$ *to be MVBUE* .

Result: If MVBUE exists for θ , then MVBUE exists for $k(\theta)$, provided k is linear function.

Proof :

$$Let \quad T \text{ is } MVBUE, \quad Hence \quad E(T) = \theta \quad and \quad Var(T) = \frac{1}{nI(\theta)}$$

$$Let \quad k(T) = aT + b \quad and \quad k(\theta) = a\theta + b$$

$$We \ have \ to \ show \quad E(k(T)) = a\theta + b \quad and \quad Var((k(T)) = \frac{[k'(\theta)]^2}{nI(\theta)} = \frac{a^2}{nI(\theta)}$$

$$1) \ E(aT + b) = a\theta + b \quad Since \quad E(T) = \theta$$

$$2) \ Var(aT + b) = a^2 Var(T) = a^2 \frac{1}{nI(\theta)} \quad since \quad Var(T) = \frac{1}{nI(\theta)}$$

$$Hence \quad MVBUE \quad exist \quad for \quad k(\theta).$$

Result : An MVBUE of a parameter *k(θ)* must be a sufficient statistic for *k(θ)*

Proof:

Let T is MVBUE of the parameter $k(\theta)$ *, then it must satisfy the relation* $\frac{\partial}{\partial\theta}\ln L = A(\theta)[T - k(\theta),]$, $\left[Var(T) = \frac{k'(\theta)}{A(\theta)} \right]$

Integrating it w.r.t. θ *(by parts) and writing* $\int A(\theta)d\theta = C(\theta)$ *we get*

$$\ln L = [T(\underline{x}) - k(\theta)]C(\theta) + \int k'(\theta)C(\theta)d\theta + B(\underline{x})$$

$$\ln L = [T(\underline{x}) - k(\theta)]C(\theta) + D(\theta) + B(\underline{x})$$

$$L = \exp\{ [T(\underline{x}) - k(\theta)]C(\theta) + D(\theta) + B(\underline{x}) \}$$

$$L = \exp\{ [T(\underline{x}) - k(\theta)]C(\theta) + D(\theta) \}\exp\{B(\underline{x})\}$$

Compare it with $L(x;\theta) = g_\theta(t(\underline{x}))h(\underline{x})$, *we find by Neyman factorization* th^m $T(\underline{X})$ *is a sufficient statistic for parameter* $k(\theta)$

Example 1:

For $N(\theta, \sigma^2)$, *where* $-\infty < \theta < \infty$, *and* σ^2 *is known*, *show that the sample mean is the MVBUE of* θ.

Solution : $f(x;\theta) = \dfrac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\dfrac{(x-\theta)^2}{2\sigma^2}\right] \quad -\infty < x < \infty$

$$\ln f(x;\theta) = \ln\left(2\pi\sigma^2\right)^{(-1/2)} - \frac{(x-\theta)^2}{2\sigma^2}$$

$$\ln f(x;\theta) = \left(-\frac{1}{2}\right)\ln\left(2\pi\sigma^2\right) - \frac{(x-\theta)^2}{2\sigma^2}$$

$$\text{Thus} \quad \frac{\partial}{\partial\theta}\ln f(x;\theta) = -\frac{2(x-\theta)(-1)}{2\sigma^2} = \frac{x-\theta}{\sigma^2}$$

$$\text{and} \quad \frac{\partial^2}{\partial\theta^2}\ln f(x;\theta) = -\frac{1}{\sigma^2}\ ;\ I(\theta) = -E\left[\frac{\partial^2}{\partial\theta^2}\ln f(x;\theta)\right] = \frac{1}{\sigma^2}$$

An estimator T of θ is said to be MVBUE if $E(T) = \theta$ and

$$Var(T) = \frac{1}{nI(\theta)}$$

$$E(\overline{X}) = \theta \quad \text{since} \quad \overline{X} \sim N\left(\theta, \frac{\sigma^2}{n}\right)$$

$$Var(\overline{X}) = \frac{\sigma^2}{n}$$

$$= \frac{1}{\frac{n}{\sigma^2}}$$

$$= \frac{1}{nI(\theta)} \qquad \text{because} \quad I(\theta) = \frac{1}{\sigma^2}$$

Hence $\overline{X}$ is MVBUE for θ

Example 2:

Let $X \sim B(1,\theta,)$ *show that* $\overline{X}$ *is MVBUE for* θ.

Solution : $f(x;\theta) = \theta^{x}(1-\theta)^{1-x} \qquad x = 0,1$

$$\ln f(x;\theta) = x\ln\theta + (1-x)\ln(1-\theta)$$

Thus $\dfrac{\partial}{\partial\theta}\ln f(x;\theta) = \dfrac{x}{\theta} - \dfrac{1-x}{1-\theta}$

and $\dfrac{\partial^2}{\partial\theta^2}\ln f(x;\theta) = -\dfrac{x}{\theta^2} - \dfrac{1-x}{(1-\theta)^2}$

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

$$i.e.\ I(\theta) = \frac{\theta}{\theta^2} + \frac{1-\theta}{(1-\theta)^2} = \frac{1}{\theta} + \frac{1}{1-\theta} = \frac{(1-\theta)+\theta}{\theta(1-\theta)} = \frac{1}{\theta(1-\theta)}$$

An estimator T of θ is said to be MVBUE if $E(T) = \theta$ and

$$Var(T) = \frac{1}{nI(\theta)}$$

$$E(\overline{X}) = E\left(\frac{\sum_{i=1}^{n} X_i}{n}\right) = \frac{1}{n} n\theta = \theta \qquad \text{since} \quad \sum_{i=1}^{n} X_i \sim B(n,\theta)$$

$$Var(\overline{X}) = Var\left(\frac{\sum_{i=1}^{n} X_i}{n}\right) = \frac{1}{n^2} n\theta(1-\theta) = \frac{\theta(1-\theta)}{n} = \frac{1}{\frac{n}{\theta(1-\theta)}}$$

$$= \frac{1}{nI(\theta)} \qquad \text{because} \quad I(\theta) = \frac{1}{\theta(1-\theta)}$$

Hence $\overline{X}$ is MVBUE for θ

Example 3:

Let $X \sim P(\theta)$ *show that* $\overline{X}$ *is MVBUE for* θ.

Solution : $f(x;\theta) = \dfrac{e^{-\theta}\theta^{x}}{x!} \qquad x = 0,1,2,3,4,.........$

$$\ln f(x;\theta) = -\theta + x \ln \theta - \ln x!$$

Thus $\dfrac{\partial}{\partial\theta} \ln f(x;\theta) = \dfrac{x}{\theta} - 1$

and $\dfrac{\partial^2}{\partial\theta^2} \ln f(x;\theta) = -\dfrac{x}{\theta^2}$

$$I(\theta) = -E\left\{\left[\frac{\partial^2 \ln f(x;\theta)}{\partial\theta^2}\right]\right\} = -E\left\{-\frac{X}{\theta^2}\right\} = \frac{E(X)}{\theta^2} = \frac{\theta}{\theta^2} = \frac{1}{\theta}$$

An estimator T *of* θ *is said to be MVBUE if* $E(T) = \theta$ *and*

$$Var(T) = \frac{1}{nI(\theta)}$$

$$E(\overline{X}) = E\left(\frac{\sum_{i=1}^{n} X_i}{n}\right) = \frac{1}{n} n\theta = \theta \qquad \text{since} \quad \sum_{i=1}^{n} X_i \sim P(n\theta)$$

$$Var(\overline{X}) = Var\left(\frac{\sum_{i=1}^{n} X_i}{n}\right) = \frac{1}{n^2} n\theta = \frac{\theta}{n} = \frac{1}{\frac{n}{\theta}}$$

$$= \frac{1}{nI(\theta)} \qquad \text{because} \quad I(\theta) = \frac{1}{\theta}$$

Hence $\overline{X}$ *is MVBUE for* θ

Example 3: (Alternative approach)

Let $X \sim P(\theta)$ *show that* $\overline{X}$ *is MVBUE for* θ.

Criterion for the MVBUE .

If T *is MVBUE of the parameter* $k(\theta)$ *, then it must satisfy the relation*

$$\frac{\partial}{\partial \theta} \ln L = A(\theta)[T - k(\theta),\ ], \qquad \left[Var(T) = \frac{k'(\theta)}{A(\theta)} \right]$$

This is therefore the necessary and sufficient condition for an estimator T *of parameter* $k(\theta)$ *to be MVBUE* .

$$f(x;\theta) = \frac{e^{-\theta}\theta^{x}}{x!} \qquad x = 0,1,2,3,4,.........$$

$$\ln f(x;\theta) = -\theta + x \ln \theta - \ln x!$$

$$L = f(x_1;\theta) f(x_2;\theta)......... \quad ... f(x_n;\theta)$$

$$\ln L = \ln f(x_1;\theta) + \ln f(x_2;\theta) + ... \quad .. + f(x_n;\theta)$$

$$\ln L = \sum \ln f(x_i;\theta)$$

$$\ln L = -n\theta + (\Sigma x_i)\ln\theta - \Sigma \ln(x_i!) \quad \textit{since} \quad \ln f(x) = -\theta + x\ln\theta - \ln(x!)$$

$$\frac{\partial}{\partial\theta}\ln L = -n + \frac{\Sigma x_i}{\theta}$$

$$\frac{\partial}{\partial\theta}\ln L = -\frac{n\theta}{\theta} + \frac{n\overline{x}}{\theta}$$

$$\frac{\partial}{\partial\theta}\ln L = \frac{n}{\theta}\left(\overline{x} - \theta\right)$$

Compare with $\frac{\partial}{\partial\theta}\ln L = A(\theta)\left[T - k(\theta)\right]$

We find that $T = \overline{X}$ *is an MVBUE of* θ. *Since* $Var(T) = \frac{k'(\theta)}{A(\theta)} = \frac{1}{A(\theta)}$

$$= \frac{\theta}{n}$$

Example 4

Determine whether the MVBUE of the parameter σ^2 *of* $N(\mu, \sigma^2)$ *distribution exists,* μ *being known.*

Criterion for the MVBUE.

If T *is MVBUE of the parameter* $k(\theta)$*, then it must satisfy the relation*

$$\frac{\partial}{\partial \theta} \ln L = A(\theta)\left[T - k(\theta),\right], \qquad \left[Var(T) = \frac{k'(\theta)}{A(\theta)}\right]$$

This is therefore the necessary and sufficient condition for an estimator T *of parameter* $k(\theta)$ *to be MVBUE.*

Let $\theta = \sigma^2$

$$L = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\theta}} \exp\left[-\frac{(x_i - \mu)^2}{2\theta}\right]$$

$$L = \left(\frac{1}{\sqrt{2\pi\theta}}\right)^n \exp\left[\frac{-\sum (x_i - \mu)^2}{2\theta}\right]$$

$$\ln L = -\frac{n}{2}\ln\left(2\pi\theta\right) - \frac{\sum (x_i - \mu)^2}{2\theta}$$

$$\frac{\partial}{\partial \theta}\ln L = -\frac{n}{2\theta} + \frac{1}{2\theta^2}\sum (x_i - \mu)^2$$

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

$$= \frac{n}{2\theta^2}\left(\frac{\sum (x_i - \mu)^2}{n} - \theta\right)$$

Compare with $\frac{\partial}{\partial \theta}\ln L = A(\theta)\left[T - k(\theta)\right]$

We find that $T = \frac{\sum (x_i - \mu)^2}{n}$ *is an MVBUE of* θ.

Since $Var(T) = \frac{k'(\theta)}{A(\theta)} = \frac{1}{A(\theta)} = \frac{2\theta^2}{n}$

Definition 1 : Let T be an unbiased estimator of a parameter θ in such a case of point estimation. The statistic T is called an efficient estimator of θ if and only if the variance of T attains the Cramer-Rao lower bound.

Definition 2: In cases in which we can differentiate with respect to a parameter under an integral or summation symbol, the ratio of the Cramer-Rao lower bound to the actual variance of any unbiased estimator of a parameter is called the efficiency of that statistic.

Let S^2 *denote the varince of a random sample of size* $n > 1$ *from a distribution that is* $N(\mu, \theta)$, *where* μ *is known. What is the efficiency of the estimator* $\frac{nS^2}{n-1}$ *of* θ.

Solution : *Cramer lower bound is* $\frac{2\theta^2}{n}$. (*Refer example* 4)

We know that $E\left[\frac{nS^2}{n-1}\right] = \theta$

Now $\frac{nS^2}{\theta} \sim \chi^2_{n-1}$ *so the variance of* $\frac{nS^2}{\theta}$ *is* $2(n-1)$

$$\textit{Accordingly, } Var\left(\frac{(n-1)nS^2}{(n-1)\theta}\right) = 2(n-1)$$

i.e. $\frac{(n-1)^2}{\theta^2} Var\left(\frac{nS^2}{(n-1)}\right) = 2(n-1) \Rightarrow Var\left(\frac{nS^2}{(n-1)}\right) = \frac{2\theta^2}{n-1}$

Thus the efficiency of the estimator $\frac{nS^2}{n-1}$ *is* $\dfrac{\frac{2\theta^2}{n}}{\frac{2\theta^2}{n-1}} = \frac{n-1}{n}$