Remark: $E\left(\frac{\partial}{\partial\theta}\log f_\theta(x)\right)^2 = I(\theta)$ is called the Fisher information

$\nwarrow V\left(\frac{\partial}{\partial\theta}\log f_\theta(x)\right)$ as $E\left(\frac{\partial}{\partial\theta}\log f_\theta(x)\right) = 0$

Alt form: $I(\theta) = -E\left(\frac{\partial^2 \log f_\theta(x)}{\partial\theta^2}\right)$

Remark: If $\exists$ an unbiased estimator whose variance equals CRLB, then it is UMVUE.

Remark: There can be situations wherein UMVUE has variance higher than CRLB. In such cases, CRLB is not achievable.

Examples:

(i) $X_1, \ldots, X_n$ i.i.d. $B(1,\theta)$

$$f_\theta(x) = \theta^x (1-\theta)^{1-x}$$

$$\log f_\theta(x) = x\log\theta + (1-x)\log(1-\theta)$$

$$\frac{\partial \log f_\theta(x)}{\partial\theta} = \frac{x}{\theta} + (1-x)\frac{1}{1-\theta}(-1)$$

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

$$\frac{\partial^2 \log f_\theta(x)}{\partial\theta^2} = -\frac{x}{\theta^2} - (1-x)\frac{1}{(1-\theta)^2}$$

$$E\left(\frac{\partial^2 \log f_\theta(x)}{\partial\theta^2}\right) = -\frac{\theta}{\theta^2} - \frac{1-\theta}{(1-\theta)^2} = -\frac{1}{\theta(1-\theta)}$$

$$I(\theta) = -E\left(\frac{\partial^2 \log f_\theta(x)}{\partial\theta^2}\right) = \frac{1}{\theta(1-\theta)}$$

Estimand: $g(\theta) = \theta$

$$CRLB = \frac{(g'(\theta))^2}{n\,I(\theta)} = \frac{\theta(1-\theta)}{n}$$

$\delta(\underline{x}) = \frac{\sum X_i}{n}$ u.e. for $g(\theta)$

$$V\left(\frac{\sum X_i}{n}\right) = \frac{1}{n^2}\sum V(X_i)$$

$$= \frac{1}{n^2}\, n\, \theta(1-\theta) = \frac{\theta(1-\theta)}{n} = CRLB$$

$\Rightarrow \bar{X}$ is UMVUE for θ.

Example (ii)

$X_1, \ldots X_n$ r.s $N(\theta, 1)$

$$f_\theta(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(x-\theta)^2}$$

$$\log f_\theta(x) = c - \frac{1}{2}(x-\theta)^2$$

$$\frac{\partial \log f_\theta(x)}{\partial \theta} = -\frac{1}{\cancel{2}}\,\cancel{2}(x-\theta)(-1) = x-\theta$$

$$E\left(\frac{\partial \log f_\theta(X)}{\partial \theta}\right)^2 = E(X-\theta)^2 = 1 = I(\theta)$$

or $$\frac{\partial^2 \log f_\theta(x)}{\partial \theta^2} = -1 \Rightarrow -E\left(\frac{\partial^2 \log f_\theta(X)}{\partial \theta^2}\right) = I(\theta) = 1$$

$g(\theta) = \theta$ say

$$CRLB = \frac{(g'(\theta))^2}{n\, I(\theta)} = \frac{1}{n}$$

$$V(\bar{X}) = \frac{1}{n} = CRLB$$

$\Rightarrow \bar{X}$ is UMVUE for θ

If $g(\theta) = \theta^2$

$$CRLB = \frac{(g'(\theta))^2}{n\, I(\theta)} = \frac{4\theta^2}{n}$$

Maximum Likelihood Estimator (MLE)

Let $X_1, \ldots, X_n$ be an i.i.d. random sample from $f_\theta(x)$ (p.d.f. or p.m.f), $\theta \in \Theta$

jt p.d.f. (or p.m.f)

$$f_{X_1,\ldots,X_n} = \prod_{i=1}^{n} f_\theta(x_i)$$

Likelihood f^n:

$$L(\theta) = \prod_{i=1}^{n} f_\theta(x_i)$$ viewed as a function of θ given the observations $x_1, \ldots, x_n$

Note that for a discrete dist^n setup $L(\theta)$ is probability of observing $(x_1, \ldots, x_n)$ and for a continuous dist^n setup $L(\theta)$ is proportional to a probability statement.

MLE approach: Find θ which maximises the likelihood (linked with the above prob statement interpretation)

Def^n: $\hat{\theta}$ is an MLE of θ if

$$\hat{\theta} = \underset{\theta \in \Theta}{\arg\max}\, L(\theta)$$

Remark: MLE is a function of sufficient statistic

T is suff $\Rightarrow$ by NFFT

$$f_\theta(\underset{\sim}{x}) = h(\underset{\sim}{x})\, g_\theta(t(\underset{\sim}{x})) = L(\theta)$$

Thus, maximization of L w.r.t θ

$\Leftrightarrow$ maximization of $g_\theta(t(\underset{\sim}{x}))$ w.r.t θ

Hence, MLE is a f^n of suff stat $T(\underline{X})$

<u>Remark</u>: Note that (as log is a monotone f^n),

$$\hat{\theta} = \underset{\theta \in \Theta}{\arg\max} \log L(\theta)$$

It is often convenient to work with $\log L(\theta)$ to find MLE

$l(\theta) = \log L(\theta)$ – log likelihood f^n.

<u>Remark</u>: Invariance property of MLE

$\theta \in \Theta \subseteq \mathbb{R}^k$, say

Let $\hat{\theta}$ be MLE of θ and $g(\cdot)$ be a f^n from Θ to a subset of $\mathbb{R}^m$ (say). Then $g(\hat{\theta})$ is MLE of $g(\theta)$.

<u>Remark</u>: Suppose $L(\theta)$ (or $l(\theta)$) is differentiable w.r.t. θ and the maximum of $L(\theta)$ ($l(\theta)$) is an interior point of Θ and not a point on the boundary then $\hat{\theta}$ satisfies

$$\left.\frac{\partial\, l(\theta)}{\partial \theta}\right|_{\theta=\hat{\theta}} = 0$$

and $$\left.\frac{\partial^2 l(\theta)}{\partial \theta}\right|_{\theta=\hat{\theta}} < 0$$

Similar conditions for multiparameter setup.

In such a situation, MLE can be obtained by solving

$$\frac{\partial \ell(\theta)}{\partial \theta} = 0$$

and verifying that $\left.\frac{\partial^2 \ell(\theta)}{\partial \theta^2}\right|_{\theta=\hat{\theta}} < 0$.