A result related to CSS :-

Suppose T_n : CSS

$g(\theta)$: A f$^{\underline{n}}$ of θ needs to be estimated.

Suppose, $\exists$ ~~an~~ a u.e. of $g(\theta)$. Then $g(\theta)$ has one and only one u.e., which depends on T_n.

Proof :- Let U_n be ~~an~~ a u.e of $g(\theta)$.

$$\eta(T_n) = E(U_n \mid T_n) \text{ is also u.e. of } g(\theta),$$

& it depends on T_n.

Let $\eta^*(T_n)$ be another u.e. of $g(\theta)$.

Observe that $E[\eta^*(T_n) - \eta(T_n)] = g(\theta) - g(\theta) = 0 \xRightarrow{\text{by the def}^{\underline{n}}\text{ of CSS}} \eta^*(T_n) = \eta(T_n)$ a.e.

Hence, proved.

Lehmann- Scheffe theorem:-

Statement:- T_n is CSS, and let U_n be a u.e. of $g(\theta)$. Then $\exists$ an unique UMVUE (Uniformly minimum variance unbiased estimator) of $g(\theta)$, which $\eta(T_n) = E(U_n | T_n)$.

Method of finding UMVUE (it follows from the Lehmann - Scheffe theorem).

Steps:-

1. Find a sufficient statistic T_n.
2. Check that T_n is complete or not.
3. If T_n is complete, find a $f^{\underline{n}}$ of T_n, which is u.e. of $g(\theta)$.

Examples :-   i) Suppose $X_1, \dots, X_n \overset{i.i.d}{\sim} N(\theta, 1)$.

Want to find UMVUE of θ.

Already, we know $\frac{1}{n}\sum_{i=1}^{n} X_i$ is a u.e. of θ.

And, moreover T_n is CSS of θ.

Hence, $\frac{1}{n}\sum_{i=1}^{n} X_i$ is the UMVUE of θ.

ii) Suppose $X_1, \dots, X_n \overset{i.i.d.}{\sim} N(\theta, 1)$

Let $g(\theta) = \theta^2$.

We know $\underbrace{\frac{1}{n}\sum_{i=1}^{n} X_i}_{=\bar{X}_n}$ is CSS for θ.

$E(\bar{X}_n) = \theta$ & $Var(\bar{X}_n) = E(\bar{X}_n^2) - \{E(\bar{X}_n)\}^2 \Leftrightarrow \frac{1}{n} = E(\bar{X}_n^2) - \theta^2$

$\Leftrightarrow E(\bar{X}_n^2 - \frac{1}{n}) = \theta^2$ Hence, $\bar{X}_n^2 - \frac{1}{n}$ is the UMVUE of θ^2

iii) $X_1, \ldots, X_n \overset{i.i.d.}{\sim} f(x,\theta) = e^{-(x-\theta)}$; $x \geq \theta$.

Check that $X_{(1)}$ is the CSS of θ.

The problem boils down to find a u.e. of θ depending on $X_{(1)}$.

$$f_{X_{(1)}}(x) \overset{\text{check}}{=} n\left[1 - F_X(x)\right]^{n-1} f_X(x)$$

$$\overset{\text{check}}{=} n\left[1 - \left(1 - e^{-(x-\theta)}\right)\right]^{n-1} e^{-(x-\theta)}$$

$$f_{X_{(1)}}(x) = n e^{-n(x-\theta)}, \quad x \geq \theta.$$

$$E[X_{(1)}] = \int_\theta^\infty y \times n e^{-n(y-\theta)}\, dy \overset{\text{check}}{=} \theta + \frac{1}{n}.$$

$\Leftrightarrow E\left[X_{(1)} - \frac{1}{n}\right] = \theta$. Hence, $\left[X_{(1)} - \frac{1}{n}\right]$ is the UMVUE of θ.

Cramer-Rao Lower bound (CRLB) :-

(Essentially it provides the lower bound of the variances of the unbiased estimators).

Set up :- $X_1, \ldots, X_n \overset{i.i.d}{\sim} f(x; \theta)$.

Let $U_n(X_1, \ldots, X_n)$: An unbiased estimator of $g(\theta)$.

A few assumption :-

i) The range of the r.v. does NOT depend θ.

ii) $g(\theta)$ is differentiable w.r.t. θ.

iii) $\frac{\partial}{\partial \theta} f(x; \theta)$ exists (is finite).

Under $\langle i \rangle - \langle iii \rangle$, $\quad Var[U_n(X_1, \ldots, X_n)] \geq \dfrac{\{g'(\theta)\}^2}{nE\left[\frac{\partial}{\partial \theta} \log f(x, \theta)\right]^2}$ — (*)

(*) is called the CRLB.

A few remarks :-

i) C.R.L.B. $= \dfrac{\{g'(\theta)\}^2}{-n\, E\left[\frac{\partial^2}{\partial\theta^2} \log f(x;\theta)\right]}$

$\longrightarrow$ (1)

ii) $E\left[\frac{\partial}{\partial\theta} \log f(x;\theta)\right]^2 = I(\theta)$ is called the fisher information or $I(\theta) = -E\left[\frac{\partial^2}{\partial\theta^2} \log f(x;\theta)\right]$.

iii) Note that $I(\theta) = Var\left(\frac{\partial}{\partial\theta} \log f(x;\theta)\right)$ $\longrightarrow$ (2)

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

To prove (2), (1) indicates that $E\left[\frac{\partial}{\partial\theta} \log f(x;\theta)\right] = 0$ $\longrightarrow$ (3)

Proof ③ :-

$$E\left[\frac{\partial}{\partial\theta}\log f(x;\theta)\right]$$

$$= \int \frac{\partial}{\partial\theta}\log f(x;\theta)\, f(x;\theta)\, dx.$$

$$= \int \frac{f'(x;\theta)}{f(x;\theta)}\, f(x;\theta)\, dx.$$

$$= \int f'(x;\theta)\, dx \quad \left[\text{Assume } f(x) > 0\ \forall x\right].$$

$$= \int \frac{\partial}{\partial\theta} f(x;\theta)\, dx \;\underset{\text{Argument of DCT needed}}{\overset{\text{Very careful}}{=\!=}}\; \frac{\partial}{\partial\theta}\int f(x;\theta)\, dx.$$

$$= \frac{\partial}{\partial\theta}[1] \text{ since } f(x;\theta) \text{ is a p.d.f.}$$
$$= 0.$$

iv) Higher of $I(\theta)$, lower the accuracy.

v) If we can find an u.e, whose variance is same as CRLB, then it will be UMVUE.

vi) Many cases, Variance (UMVUE) > CRLB.
In other words, for those cases, CRLB is NOT achievable.

Examples :- Let $X_1, \dots, X_n \overset{i.i.d.}{\sim} \text{Ber}(1, \theta)$, $\theta \in (0,1)$.

$$f(x; \theta) = \theta^{x} (1-\theta)^{1-x}; \quad x = 0 \text{ or } 1.$$

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

$$\Rightarrow \frac{\partial}{\partial\theta} \log f(x;\theta) = \frac{x}{\theta} - \frac{1-x}{1-\theta}$$

$$= \frac{x}{\theta} + \frac{x}{1-\theta} - \frac{1}{1-\theta}$$

$$= \frac{x}{\theta(1-\theta)} - \frac{1}{1-\theta} \quad \text{———} \quad (*)$$

$$I(\theta) = \mathrm{Var}\left[\frac{\partial}{\partial\theta} \log f(x;\theta)\right]$$

$$= \mathrm{Var}\left[\frac{X}{\theta(1-\theta)} - \frac{1}{1-\theta}\right] \quad \text{Using } (*)$$

$$= \mathrm{Var}\left[\frac{X}{\theta(1-\theta)}\right] = \frac{1}{\theta^2(1-\theta)^2} \times \theta(1-\theta) = \frac{1}{\theta(1-\theta)}.$$

$\mathrm{CRLB} = \dfrac{1}{n \times \frac{1}{\theta(1-\theta)}} = \dfrac{\theta(1-\theta)}{n} = \mathrm{Var}\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right)$. Hence, $\frac{1}{n}\sum_{i=1}^{n} X_i$ is the UMVUE!