THEORY OF POINT ESTIMATION

Consider the following set up —
X = A random variable, x = Observation on X.
$\mathscr{X}$ = Sample space.
$p(x) \in \{p_\theta(x) : \theta \in \Omega\} \to$ a family of pdfs or pmfs.
Both X and θ may be multidimensional.
To estimate $g(\theta)$ = a real-valued function of θ.
Let $T = t(X)$ be a real-valued function of x.

Definition: Any statistic $T = t(X)$ is called an estimator of $g(\theta)$ if we estimate $g(\theta)$ by $t(x)$ for $X = x$. $t(x)$ is called the estimate of $g(\theta)$ corresponding to $X = x$.

The probability distribution of a good estimator T should have a good degree of concentration around the true value of $g(\theta)$.

A measure of this is given by
Mean square error (MSE) $= MSE_\theta(T) = E_\theta(T - g(\theta))^2$. Clearly, MSE depends on θ.

Taking MSE as a measure of goodness of an estimator we can make the following definition.

Definition:- An estimator T of $g(\theta)$ is called the best estimator if $MSE_\theta(T) \le MSE_\theta(T') \ \forall \theta \in \Omega$, whatever be the other estimator T' of $g(\theta)$.

Proposition:- No best estimator, satisfying the above definition, exists.

Proof: If possible, let T be the best estimator of $g(\theta)$.
Then $MSE_\theta(T) \le MSE_\theta(T') \ \forall \theta \in \Omega$, whatever be the other estimator T' of $g(\theta)$.
Consider the particular value θ_0 of θ, and let us define the estimator $T' = g(\theta_0)$.
Then, $MSE_\theta(T') = 0$
$\Rightarrow MSE_{\theta_0}(T) \le 0 \Rightarrow MSE_{\theta_0}(T) = 0$.
$\Rightarrow T = g(\theta_0)$ with probability 1.
But θ_0 is any arbitrary value of θ.
Hence we must have $T = g(\theta)$ with probability 1.

But such a choice of T is impossible since θ is unknown to us so that we can not chose $T = g(\theta)$. Hence the proposition.

Since no ~~such~~ best estimator exists within the class of all estimators for $g(\theta)$, we may consider a reasonable sub-class of estimators and proceed to find the best estimator within this sub-class. One such reasonable sub-class is the class of "unbiased estimators".

Unbiasedness :— An estimator T of $g(\theta)$ is said to be unbiased if $E_\theta(T) = g(\theta) \ \forall \theta \in \Omega$.

For an unbiased estimator T of $g(\theta)$ $MSE_\theta(T) = Var_\theta(T)$.

Definition: An unbiased estimator T of $g(\theta)$ is said to be best within the class of u.e. of $g(\theta)$ if

$$Var_\theta(T) \le Var_\theta(T') \ \forall \theta \in \Omega, \text{ whatever be the other u.e. } T' \text{ of } g(\theta).$$

This best estimator is called the uniformly ~~or UMVUE~~ minimum variance unbiased estimator or UMVUE (commonly it is known as the MVUE).

Note 1: From law of large numbers it follows that for a large number of repetitions of an experiment the average of the values assummed by an u.e. T of $g(\theta)$ will tend to $g(\theta)$ with probability 1.

This justifies the restriction "The class of unbiased estimators".

Note 2: In some situations, no u.e. of $g(\theta)$ may exist, i.e. the class of unbiased estimators is empty.

Example 1: $P_N(X=x) = \dfrac{\binom{D}{x}\binom{N-D}{n-x}}{\binom{N}{n}}$; $x = 0,1,2,\dots,\min(D,n)$; $N = D, D+1, \dots$

Let $T = t(X)$ be an u.e. of N. $t(x)$ is then defined for $x = 0,1,2,\dots,\min(D,n)$.

Let $M = \max\limits_{0 \le x \le \min(D,n)} t(x)$; $m = \min\limits_{0 \le x \le \min(D,n)} t(x)$

Then $m \le t(x) \le M \quad \forall x$.

$\Rightarrow m \le E_N[t(X)] \le M \ \forall N$ ---- (1)

But T is an u.e. of N so that

$E_N[t(X)] = N \quad \forall N$ ---- (2)

Then (1) and (2) contradict one another for $N > M$.

$\Rightarrow \nexists$ any u.e. of N.

Example 2. $X \sim Bin(n, \theta)$

To estimate $g(\theta) = \frac{1}{\theta}$

Let $T = t(X)$ be an u.e. of $g(\theta)$

Then $E_\theta[t(X)] = \frac{1}{\theta} \quad \forall \theta \in (0,1)$

or, $\sum_{x=0}^{n} t(x)\binom{n}{x}\theta^x(1-\theta)^{n-x} = \frac{1}{\theta} \quad \forall \theta \in (0,1) \quad \cdots (1)$

Now RHS of (1) can be made arbitrarily large by taking θ sufficiently close to 0. But LHS is bounded since

$$\left|\sum t(x)\binom{n}{x}\theta^x(1-\theta)^{n-x}\right| \leq \sum_{x=0}^{n} |t(x)|\binom{n}{x}\theta^x(1-\theta)^{n-x} \leq \sum_{x=0}^{n} |t(x)|\binom{n}{x},$$

since $\theta^x(1-\theta)^{n-x} < 1$ as $0<\theta<1$.

Hence (1) cannot hold for any T.

$\Rightarrow \nexists$ an u.e. of $g(\theta) = \frac{1}{\theta}$

(H.T.)

Example 3. $X \sim B(n, \theta)$

To estimate $g(\theta) = \frac{\theta}{1-\theta}$. Show that $\nexists$ u.e. of $g(\theta)$.

Solution: $E_\theta[T(X)] = \frac{\theta}{1-\theta} \Rightarrow \sum_{x=0}^{n} t(x)\binom{n}{x}\theta^x(1-\theta)^{n-x} = \frac{1}{1-\frac{1}{\theta}}$

RHS $\to \infty$ as $\theta \to 1$. But LHS $\leq \sum_{x=0}^{n} |t(x)|\binom{n}{x}$.

Definition: A function $g(\theta)$ is said to be estimable if $\exists$ at least one u.e. of $g(\theta)$.

So, whenever, we shall speak of an u.e. of $g(\theta)$ we shall assume that $g(\theta)$ is estimable.

Note 3: The UMVUE of $g(\theta)$ may be inadmissible within the class of all estimators of $g(\theta)$ in the sense that there may exist a biased estimator of $g(\theta)$ which is better than the UMVUE.

Example: $N(\mu, \sigma^2) \longrightarrow \mu, \sigma^2$ both unknown.

Let sample observations $x_1, x_2, \cdots, x_n$.

To estimate σ^2

Let $S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2$

Then S^2 is MVUE of σ^2.

Let $T_c = cS^2$ (for some constant c)

Then $E(T_c) = c\,E(S^2) = c\sigma^2$

$\Rightarrow T_c$ is u.e. iff $c = 1$,

$$\begin{aligned}
MSE(T_c) &= E(T_c-\sigma^2)^2 \\
&= E(cs^2-\sigma^2)^2 \\
&= c^2 Var(s^2) + [E(cs^2)-\sigma^2]^2 \\
&= c^2 Var(s^2) + \sigma^4(c-1)^2
\end{aligned}$$

Now $\frac{(n-1)s^2}{\sigma^2} \sim \chi^2_{n-1} \Rightarrow Var\left(\frac{(n-1)s^2}{\sigma^2}\right) = 2(n-1)$

$$\Rightarrow Var(s^2) = \frac{2\sigma^4}{n-1}$$

$\therefore MSE(T_c) = \sigma^4\left[\frac{2c^2}{n-1} + (c-1)^2\right]$

$\frac{d\, MSE(T_c)}{dc} = 0$

$\Rightarrow \frac{4c}{n-1} + 2(c-1) = 0$

$\Rightarrow c = \frac{n-1}{n+1}$, whatever be (μ, σ^2)

$\Rightarrow MSE\left(T_{\frac{n-1}{n+1}}\right) < MSE(T_c)$, whatever (μ, σ^2).

$\Rightarrow MSE\left(T_{\frac{n-1}{n+1}}\right) < MSE(T_1) = Var(s^2)$, whatever (μ, σ^2)

$\Rightarrow s^2$, though MVUE, is not admissible.

Some lower bounds to the variance of an u.e. of $g(\theta)$.

Case of Single parameter:

Consider $\mathcal{P} = \{p_\theta(x), \theta \in \Omega\}$, a family of pdf's.
We take θ as a real-valued parameter. To obtain some lower bound to the variance of any u.e. of the real-valued estimable function $g(\theta)$.

Crammer-Rao lower bound: $\mathcal{P}$ is said to satisfy Crammer-Rao regularity conditions if

(i) Ω is non-degenerate open subset of real line.

(ii) $\frac{\partial p_\theta(x)}{\partial \theta}$ exists $\forall \theta \in \Omega$.

(iii) $\frac{\partial}{\partial \theta}\int p_\theta(x)\,dx = \int \frac{\partial}{\partial \theta} p_\theta(x)\,dx$

(iv) $I(\theta) = E_\theta\left[\frac{\partial}{\partial \theta}\ln p_\theta(x)\right]^2 = E_\theta\left[\frac{1}{p_\theta(x)}\frac{\partial}{\partial \theta}p_\theta(x)\right]^2$

$I(\theta)$ is called the Fisher's Information function for the information contained in x about θ.

$I(\theta)$ exists and is positive.

Theorem: Let $\mathcal{P}$ be any family of pdf's satisfying the C-R regularity conditions, and $T = t(x)$ is any u.e. of a differentiable parametric function $g(\theta)$ satisfying

(v) $\frac{\partial}{\partial\theta}\int t(x)\, p_\theta(x)\,dx = \int t(x) \cdot \frac{\partial}{\partial\theta} p_\theta(x)\,dx$

Then, $Var_\theta(T) \geq \frac{[g'(\theta)]^2}{I(\theta)} \quad \forall \theta$

Proof: Let $S(x,\theta) = \frac{\partial}{\partial\theta} \ln p_\theta(x)$

Then, $E_\theta(S) = \int [\frac{\partial}{\partial\theta} \ln p_\theta(x)]\, p_\theta(x)\,dx$

$$= \int \frac{1}{p_\theta(x)} [\frac{\partial}{\partial\theta} p_\theta(x)] \cdot p_\theta(x)\,dx.$$

$$= \int \frac{\partial}{\partial\theta} p_\theta(x)\,dx$$

$$= \frac{\partial}{\partial\theta}\int p_\theta(x)\,dx = 0 \quad \forall\theta$$

$\therefore\ Var_\theta(S) = E(S^2) = I(\theta) \quad \forall\theta$ (regarding condition (iv))

$Cov_\theta(S,T) = E_\theta(S.T)$

$$= \int \frac{\partial}{\partial\theta} \ln p_\theta(x) \cdot t(x)\, p_\theta(x)\,dx$$

$$= \int \frac{1}{\cancel{p_\theta(x)}} \cdot \frac{\partial}{\partial\theta} p_\theta(x) \cdot t(x)\, \cancel{p_\theta(x)}\,dx$$

$$= \frac{\partial}{\partial\theta}\int p_\theta(x)\, t(x)\,dx. \quad [\text{condition (v)}]$$

$$= \frac{\partial}{\partial\theta} g(\theta)$$

$$= g'(\theta) \quad \forall\theta.$$

Now, $\rho^2(S,T) \leq 1 \quad \forall\theta$

or, $Cov_\theta^2(T,S) \leq Var_\theta(T) \cdot Var_\theta(S)$ for all θ.

or, $[g'(\theta)]^2 \leq Var_\theta(T) \cdot I(\theta) \quad \forall\theta$

or, $Var_\theta(T) \geq \frac{[g'(\theta)]^2}{I(\theta)} \quad \forall\theta$.

Case of equality

'=' holds iff $T \propto S(x,\theta)$ with probability 1.

or, $T - g(\theta) = \lambda(\theta)\, S(x,\theta)$ with probability 1.

Since equality holds in this case, we have,

$$Var_\theta(T) = \frac{[g'(\theta)]^2}{I(\theta)}$$

or, $\lambda^2(\theta)\, I(\theta) = \dfrac{[g'(\theta)]^2}{I(\theta)}$

or, $\lambda^2(\theta) = \left[\dfrac{g'(\theta)}{I(\theta)}\right]^2$

$\Rightarrow \lambda(\theta) = \pm \dfrac{g'(\theta)}{I(\theta)}$.

But, $\lambda(\theta) = -\dfrac{g'(\theta)}{I(\theta)}$ is impossible (check)

$\Rightarrow \lambda(\theta) = \dfrac{g'(\theta)}{I(\theta)}$.

i.e. $T - g(\theta) = \dfrac{g'(\theta)}{I(\theta)} \cdot S(x,\theta)$ with probability 1.

Distribution admitting u.e.'s with variance attaining C-R lower bound:-

'=' holds iff

$t(x) - g(\theta) = \lambda(\theta) \frac{\partial}{\partial\theta} \ln p_\theta(x)$ with probability 1, where $\lambda(\theta) = \dfrac{g'(\theta)}{I(\theta)}$.

or, $\frac{\partial}{\partial\theta} \ln p_\theta(x) = \dfrac{t(x)}{\lambda(\theta)} - \dfrac{g(\theta)}{\lambda(\theta)}$ with probability 1.

or, $\ln p_\theta(x) = Q(\theta)\cdot t(x) + c(\theta) + h(x)$

i.e. $p_\theta(x) = e^{Q(\theta)\cdot t(x) + c(\theta) + h(x)}$

$= K(\theta)\, e^{Q(\theta)\, t(x)}\, H(x)$, where $K(\theta) = e^{c(\theta)}$ and $H(x) = e^{h(x)}$.

which is of the exponential form.

Again, if $p_\theta(x)$ be of the above form, then

$\ln p_\theta(x) = Q(\theta)\, t(x) + c(\theta) + h(x)$

$$\begin{aligned} S(x,\theta) = \frac{\partial}{\partial\theta} \ln p_\theta(x) &= Q'(\theta)\, t(x) + c'(\theta) \\ &= Q'(\theta)\left[t(x) - \left\{ -\frac{c'(\theta)}{Q'(\theta)} \right\} \right] \\ &= \frac{1}{\lambda(\theta)} [t(x) - g(\theta)] \end{aligned}$$

where, $\lambda(\theta) = \dfrac{1}{Q'(\theta)}$, $g(\theta) = -\dfrac{c'(\theta)}{Q'(\theta)}$.

Theorem: The necessary and sufficient condition for $\mathcal{P}$ to admit an u.e. $T(x) = t(x)$ of some $g(\theta)$ with variance attaining C-R lower bound is that $p_\theta(x)$ is of the exponential form, viz, $p_\theta(x) = e^{Q(\theta)t(x) + c(\theta) + h(x)}$ and in this case $g(\theta) = -\frac{c'(\theta)}{Q'(\theta)}$.

Corollary: If $Var_\theta(T)$ attains C-R lower bound then T is a sufficient statistic for $\mathcal{P}$. [Since pdf of the exponential form has $T = t(x)$ as a sufficient statistic]

Example 1. $X = (x_1, x_2, \ldots, x_n) \to$ outcome of n independent bernoullian trials with probability of success θ, $0<\theta<1$,

$$p_\theta(x) = \theta^{\Sigma x_i}(1-\theta)^{n-\Sigma x_i} = \left(\frac{\theta}{1-\theta}\right)^{\Sigma x_i}(1-\theta)^n$$
$$= e^{Q(\theta)t(x) + c(\theta) + h(x)},$$

where, $Q(\theta) = \ln\frac{\theta}{1-\theta}$, $t(x) = \Sigma x_i$, $c(\theta) = n\ln(1-\theta)$, $e^{h(x)} = 1$, and this is of the exponential form.

$$Q'(\theta) = \frac{1}{\theta(1-\theta)}$$

$$c'(\theta) = -\frac{n}{1-\theta}$$

$$g(\theta) = -\frac{c'(\theta)}{Q'(\theta)} = n\theta$$

$\Rightarrow$ For this parametric function $g(\theta) = n\theta$, $\exists$ an u.e. $T = \Sigma x_i$, whose variance attains C-R lower bound. [As. $I(\theta) = g'(\theta) \cdot Q'(\theta)$]

$\therefore$ C-R lower bound $= Var_\theta(\Sigma x_i) = n\theta(1-\theta)$ $\forall\theta$.

For any other u.e. T' of $g(\theta)$,

$$Var_\theta(T') \geq n\theta(1-\theta) \quad \forall\theta$$

Now, $E_\theta(\bar{x}) = E_\theta\left(\frac{\Sigma x_i}{n}\right) = \theta$ $\forall\theta$

$\Rightarrow \bar{x}$ is an u.e. of θ.

Since $\exists$ a 1:1 relation between an u.e. of $n\theta$ and that of θ, it follows that $T = \bar{x}$ gives an u.e. of θ with variance attaining C-R lower bound i.e. C-R lower bound $= Var_\theta(\bar{x}) = \frac{\theta(1-\theta)}{n}$ $\forall\theta$.

Examples (H.T.): $x_1, x_2, \ldots, x_n$ are iid. (i) Poisson(θ), (ii) $N(\theta,1)$ & (iii) $N(0,\theta)$.
In each of the above three cases identify parametric functions for which u.e. exists with variance attaining the C-R lower bound. Also find the C-R lower bound.

i) $\frac{e^{-n\theta}\theta^{\Sigma x_i}}{\prod_{i=1} x_i!} = e^{\ln\theta\Sigma x_i - n\theta - \Sigma\ln x_i}$, $g(\theta) = n\theta$, $T(x) = \Sigma x_i$

ii) $\frac{1}{(2\pi)^{\frac{n}{2}}} e^{-\frac{1}{2}\Sigma x_i^2 + \theta\Sigma x_i - \frac{1}{2}n\theta^2}$, $g(\theta) = n\theta$, $T(x) = \Sigma x_i$

iii) $\frac{1}{(2\pi)^{\frac{n}{2}}} e^{-\frac{n}{2}\ln\theta - \frac{1}{2\theta}\Sigma x_i^2}$, $g(\theta) = n\theta$, $T(x) = \Sigma x_i^2$.

Notes:

1. The C-R inequality can also be applied to get a lower bound to the MSE of a biased estimator T of $g(\theta)$.

Let T be a biased estimator of $g(\theta)$.

$$
\begin{aligned}
MSE_\theta(T) &= E_\theta[T-g(\theta)]^2 \\
&= E_\theta[\{T-E_\theta(T)\} + \{E_\theta(T)-g(\theta)\}]^2 \\
&= Var_\theta(T) + b^2(\theta) \\
&\geq \frac{[\frac{d}{d\theta}E_\theta(T)]^2}{I(\theta)} + b^2(\theta) \\
&= b^2(\theta) + \frac{[g'(\theta)+b'(\theta)]^2}{I(\theta)} \quad [\text{Since, } b(\theta) = E_\theta(T) - g(\theta)]
\end{aligned}
$$

2. If $g(\theta)=\theta$, Then for any u.e. of $g(\theta)$

$$Var_\theta(T) \geq \frac{1}{I(\theta)} \quad \forall \theta.$$

3. If $p_\theta(x)$, beside the regularity conditions already stated, also satisfies

$$\frac{\partial^2}{\partial\theta^2}\int p_\theta(x)\,dx = \int \frac{\partial^2}{\partial\theta^2} p_\theta(x)\,dx,$$

then $$Var_\theta(T) \geq -\frac{[g'(\theta)]^2}{E_\theta[\frac{\partial^2}{\partial\theta^2}\ln p_\theta(x)]} \quad \forall\theta$$

Proof: It is sufficient to show that

$$I(\theta) = -E\left[\frac{\partial^2}{\partial\theta^2}\ln p_\theta(x)\right] \quad \forall\theta$$

We have $\int \frac{\partial^2}{\partial\theta^2} p_\theta(x)\,dx = 0$

$$\Rightarrow \int \frac{\partial}{\partial\theta}\left[p_\theta(x)\cdot\frac{\partial}{\partial\theta}\ln p_\theta(x)\right]dx = 0, \text{ Since } \frac{\partial}{\partial\theta}\ln p_\theta(x) = \frac{1}{p_\theta(x)}\frac{\partial}{\partial\theta}p_\theta(x).$$

$$\Rightarrow \int p_\theta(x)\left[\left\{\frac{\partial}{\partial\theta}\ln p_\theta(x)\right\}^2 + \frac{\partial^2}{\partial\theta^2}\ln p_\theta(x)\right]dx = 0,$$

Since $$\left(\frac{\partial \ln p_\theta(x)}{\partial\theta}\right)^2 = \frac{1}{p_\theta^2(x)}\left[\frac{\partial}{\partial\theta}p_\theta(x)\right]^2$$

and $$\frac{\partial^2}{\partial\theta^2}\ln p_\theta(x) = \frac{\partial}{\partial\theta}\left\{\frac{1}{p_\theta(x)}\frac{\partial}{\partial\theta}p_\theta(x)\right\} = -\frac{1}{p_\theta^2(x)}\left[\frac{\partial}{\partial\theta}p_\theta(x)\right]^2 + \frac{\frac{\partial^2}{\partial\theta^2}p_\theta(x)}{p_\theta(x)}$$

$$\Leftrightarrow E_\theta\left[\frac{\partial}{\partial\theta}\ln p_\theta(x)\right]^2 + E_\theta\left[\frac{\partial^2}{\partial\theta^2}\ln p_\theta(x)\right] = 0$$

$$\Leftrightarrow I(\theta) = -E_\theta\left[\frac{\partial^2}{\partial\theta^2}\ln p_\theta(x)\right].$$

4. If $p_\theta(x) = \prod_{i=1}^{n} f_\theta(x_i)$, where $x = (x_1, x_2, \cdots, x_n)$, x_i's being iid with common pdf $f_\theta(x)$.

Then, $Var_\theta(T) \geq \dfrac{[g'(\theta)]^2}{n E_\theta [\frac{\partial}{\partial\theta} \ln f_\theta(x)]^2} \quad \forall \theta$

Proof: It is sufficient to show that

$I(\theta) = n E_\theta [\frac{\partial}{\partial\theta} \ln p_\theta(x)]^2$

$$[\tfrac{\partial}{\partial\theta} \ln p_\theta(x)]^2 = [\tfrac{\partial}{\partial\theta} \sum_{i=1}^{n} \ln f_\theta(x_i)]^2$$

$$= [\sum_{i=1}^{n} \tfrac{\partial}{\partial\theta} \ln f_\theta(x_i)]^2$$

$$= \sum_{i=1}^{n} [\tfrac{\partial}{\partial\theta} \ln f_\theta(x_i)]^2 + \sum_{\substack{i=1 \\ i\neq j}}^{n} \sum_{j=1}^{n} \tfrac{\partial}{\partial\theta} \ln f_\theta(x_i) \cdot \tfrac{\partial}{\partial\theta} \ln f_\theta(x_j)$$

$$\Rightarrow I(\theta) = \sum_{i=1}^{n} E_\theta [\tfrac{\partial}{\partial\theta} \ln f_\theta(x_i)]^2 + \sum_{\substack{i=1 \\ (i\neq j)}}^{n} \sum_{j=1}^{n} E_\theta \left(\tfrac{\partial}{\partial\theta} \ln f_\theta(x_i)\right) \cdot E_\theta \left(\tfrac{\partial}{\partial\theta} \ln f_\theta(x_j)\right)$$

$$= n E_\theta [\tfrac{\partial}{\partial\theta} \ln f_\theta(x)]^2 \quad [\text{since } x_i\text{'s are iid}].$$

5. For some $p_\theta(x)$, there may not exist any $g(\theta)$ for which $\exists$ an u.e. whose variance attains the C-R lower bound. Eg. $x = (x_1, x_2, \cdots, x_n)$; x_i's iid $\sim f_\theta(x)$, where

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

Since $p_\theta(x)$ is not of the exponential form $\nexists$ any $g(\theta)$ whose u.e. has variance equal to C-R lower bound.

6. C-R lower bound will not be applicable if the regularity conditions be not satisfied.

Example: $x_1, x_2, \cdots, x_n$ iid $\sim f_\theta(x)$, where
$$f_\theta(x) = e^{\theta - x}; \quad x \geq \theta,\ \theta > 0$$
$$= 0 \quad , \text{otherwise.}$$

In this case the regularity conditions are not satisfied since $\frac{\partial}{\partial\theta} f_\theta(x)$ does not exist at $x = \theta$.

Suppose we want to estimate θ.

Then, C-R lower bound $= \dfrac{1}{n E[\frac{\partial}{\partial\theta} \ln f_\theta(x)]^2} = \dfrac{1}{n}$

Consider $T = x_{(1)} - \frac{1}{n}$. Then $E_\theta(T) = \theta$ (check)

But $Var(T) = \frac{1}{n^2} < \frac{1}{n} =$ C-R lower bound

Bhattacharya System of lower bounds: A generalisation of C-R lower bound :—

For some estimable function $g(\theta)$, there may not exist any unbiased estimator whose variance attains the C-R lower bound. e.g. $(x_1, x_2, \cdots, x_n) \overset{iid}{\sim} N(\theta, 1)$.

To estimate $g(\theta) = \theta^2$

Here $\nexists$ any u.e. of θ^2 whose variance attains the C-R lower bound. (Check)

In such a situation, there may exist an u.e. of $g(\theta)$ whose variance attains some sharper (or larger) lower bound. One such system of lower bounds is the Bhattacharya lower bounds.

A family $\mathcal{P} = \{p_\theta(x); \theta \in \Omega\}$ is said to satisfy Bhattacharya regularity conditions if

(i) Ω is an open interval of the real line

(ii) $\frac{\partial^i}{\partial\theta^i} p_\theta(x)$ exists $\forall\, \theta$, $i = 1(1)k$.

(iii) $0 = \frac{\partial^i}{\partial\theta^i} \int p_\theta(x)\,dx = \int \frac{\partial^i}{\partial\theta^i} p_\theta(x)\,dx \quad \forall\theta;\ i = 1(1)k$.

(iv) Let $V_{ij}(\theta) = E_\theta\left[\frac{1}{p_\theta(x)} \cdot \frac{\partial^i}{\partial\theta^i} p_\theta(x) \cdot \frac{1}{p_\theta(x)} \frac{\partial^j}{\partial\theta^j} p_\theta(x)\right];\ i, j = 1(1)k$.

All $V_{ij}(\theta)$'s are finite and $V^{k\times k} = ((V_{ij}))$ is non-singular.

Here k is some positive integer.

For $k = 1$, the above regularity conditions reduce to C-R regularity conditions.

For $k > 1$, the conditions are more stringent than the C-R regularity conditions.

Theorem: Let $\mathcal{P}$ be a family of pdf's satisfying Bhattacharya regularity conditions, and let $g(\theta)$ be a real valued estimable function of θ and is k times differentiable. Then, for any unbiased estimator T of $g(\theta)$ satisfying

(v) $\frac{\partial^i}{\partial\theta^i} \int t(x)\, p_\theta(x)\,dx = \int t(x) \frac{\partial^i}{\partial\theta^i} p_\theta(x)\,dx \quad \forall\, i = 1(1)k$.

Then $\quad Var_\theta(T) \geqslant g' V^{-1} g$, where $g'(\theta) = \left(g^{(1)}(\theta), g^{(2)}(\theta), \cdots, g^{(k)}(\theta)\right)$,

$$g^{i}(\theta) = \frac{\partial^i}{\partial\theta^i} g(\theta);\ i = 1(1)k.$$

Proof: Let $s_i(x,\theta) = \frac{1}{p_\theta(x)} \frac{\partial^i}{\partial\theta^i} p_\theta(x)$; $i = 1(1)k$.

Then, $E_\theta[S_i] = \int \frac{1}{p_\theta(x)} \frac{\partial^i}{\partial\theta^i} p_\theta(x) \cdot p_\theta(x)\, dx$

$= \frac{\partial^i}{\partial\theta^i} \int p_\theta(x)\, dx$

$= 0$; $i = 1(1)k$ (by condition (iii))

$Cov(s_i, s_j) = E_\theta(s_i \cdot s_j) = v_{ij}$; $i, j = 1(1)k$.

$E_\theta(T) = g(\theta)$

$Cov(s_i, T) = \int t(x) \cdot \frac{1}{p_\theta(x)} \frac{\partial^i}{\partial\theta^i} p_\theta(x) \cdot p_\theta(x)\, dx$.

$= \frac{\partial^i}{\partial\theta^i} \int t(x)\, p_\theta(x)\, dx$ [by condition (ii)]

$= \frac{\partial^i}{\partial\theta^i} g(\theta) = g^{(i)}(\theta)$; $i = 1(1)k$.

Let $\Sigma^{\overline{k+1} \times \overline{k+1}}$ = var-cov matrix of $\begin{pmatrix} T \\ s_1 \\ \vdots \\ s_k \end{pmatrix}$

$$= \left(\begin{array}{c|cccc} var_\theta(T) & g^{(1)}(\theta) & g^{(2)}(\theta) & \cdots & g^{(k)}(\theta) \\ \hline & v_{11} & v_{12} & \cdots & v_{1k} \\ \underset{\sim}{g} & & v_{22} & \cdots & v_{2k} \\ & & & \ddots & \vdots \\ & & & & v_{kk} \end{array}\right)$$

$$= \left(\begin{array}{c|c} var_\theta(T) & \underline{g}' \\ \hline \underline{g} & V \end{array}\right)$$

Since, Σ is a var-cov matrix, it must be non-negative definite.

$\Rightarrow |\Sigma| \geq 0$.

i.e. $|V| \cdot |var_\theta(T) - \underline{g}' V^{-1} \underline{g}| \geq 0$.

i.e. $(var_\theta(T) - \underline{g}' V^{-1} \underline{g}) \cdot |V| \geq 0$.

i.e. $Var_\theta(T) - \underline{g}' V^{-1} \underline{g} \geq 0$, since by condition (iv), $|V| > 0$.

$\Leftrightarrow Var_\theta(T) \geq \underline{g}' V^{-1} \underline{g}$ (Proved)

Equality case:

'=' holds iff $|\Sigma| = 0$

i.e. $rank(\Sigma) < k+1$

But $rank(V) = k$, since V is non-singular.

Hence, $r(\Sigma) = k$.

Lemma:- Let $\underline{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_p \end{pmatrix}$, $\Sigma = \text{Disp}(x)$. Then, Σ is of rank r iff with probability 1, $x_1, x_2, \cdots, x_p$ satisfy $(p-r)$ independent relations of the form

$$a_{11}(x_1-\mu_1) + a_{12}(x_2-\mu_2) + \cdots + a_{1p}(x_p-\mu_p) = 0$$
$$a_{21}(x_1-\mu_1) + a_{22}(x_2-\mu_2) + \cdots + a_{2p}(x_p-\mu_p) = 0$$
$$\cdots\cdots\cdots$$
$$a_{p-r,1}(x_1-\mu_1) + a_{p-r,2}(x_2-\mu_2) + \cdots + a_{p-r,p}(x_p-\mu_p) = 0$$

In our case, $p = k+1$, $\underline{x} = \begin{pmatrix} T \\ S_1 \\ \vdots \\ S_k \end{pmatrix}$, $r = k$.

Hence, by the above lemma, for "=" to hold with probability 1 $T, S_1, S_2, \cdots, S_k$ should satisfy one linear relation of the type

$$a_0(T - g(\theta)) + a_1 S_1 + a_2 S_2 + \cdots + a_k S_k = 0$$
or, $T - g(\theta) = l_1 S_1 + l_2 S_2 + \cdots + l_k S_k$

or, $T - g(\theta) = \underline{l}'\underline{S}$ where $S = \begin{pmatrix} S_1 \\ \vdots \\ S_k \end{pmatrix}$, $\underline{l} = \begin{pmatrix} l_1 \\ \vdots \\ l_k \end{pmatrix}$.

In the '=' case,

$Var_\theta(T) = g'V^{-1}g = Var_\theta(g'V^{-1}S)$.

Hence $Var_\theta(\underline{l}'S - g'V^{-1}S) = Var_\theta(T - g(\theta) - g'V^{-1}S)$

$$= Var_\theta(T) + Var(g'V^{-1}S) - 2.\,Cov(T - g(\theta), g'V^{-1}S)$$
$$= \underline{g}'V^{-1}\underline{g} + \underline{g}'V^{-1}VV^{-1}\underline{g} - 2\underline{g}'V^{-1}\underline{g}$$
$$= 0$$

$\Rightarrow \underline{l}'S - \underline{g}'V^{-1}S = 0$ with probability 1.

or, $\underline{l}'S = \underline{g}'V^{-1}S$ with probability 1.

Hence equality holds iff

$T - g(\theta) = \underline{g}'V^{-1}S$ with probability 1.

Notes

1. For $k=1$, Then $\underline{g}' = g^{(1)}(\theta)$ [i.e. $g^{(1)}(\theta) = \frac{\partial}{\partial\theta} g(\theta)$].

$$V = v_{11} = E_\theta\left[\frac{1}{p_\theta(x)}\frac{\partial}{\partial\theta}p_\theta(x)\right]^2$$
$$= E_\theta\left[\frac{\partial}{\partial\theta}\ln p_\theta(x)\right]^2$$
$$= I(\theta)$$

$$\therefore g'V^{-1}g = \frac{g^{(1)}(\theta)}{I(\theta)} \cdot g^{(1)}(\theta) = \frac{[g^{(1)}(\theta)]^2}{I(\theta)}$$
$$= \text{C-R lower bound.}$$

i.e. C-R lower bound is a particular case of Bhattacharya lower bound.

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

$g^{(1)}(\theta) = 1$, $g^{(i)}(\theta) = 0$; $i = 2(1)k$.

$\therefore \underline{g}'^{\,1\times k} = (1, 0, 0, \cdots, 0)$

$\therefore g'V^{-1}g = (1, 0, 0, \cdots, 0)\begin{pmatrix} v^{11} & v^{12} & \cdots & v^{1k} \\ \cdot & \cdot & \cdots & v^{kk} \end{pmatrix}\begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$, where $V^{-1} = ((v^{ij}))$

$$= v^{11} = \frac{1}{v_{11} - \underset{\sim}{v}' V_2^{-1} \underset{\sim}{v}}, \text{ where } \underset{\sim}{v}' = (v_{12}\, v_{13} \cdots v_{1k})$$

$$V_2 = \begin{pmatrix} v_{22} & v_{23} & \cdots & v_{2k} \\ & \ddots & & \vdots \\ & & & v_{kk} \end{pmatrix}$$

$$= \frac{|V_2|}{|V|}, \text{ for } k \geq 2$$

$$= \frac{1}{I(\theta)}, \text{ for } k = 1.$$

For different k's we obtain different lower bounds, i.e. we have a sequence of lower bounds $\{\Delta_k\}$ for $k = 1, 2, \ldots$, where $\Delta_k = k^{th}$ Bhattacharya lower bound $= \underset{\sim}{g}_k' V_k^{-1} \underset{\sim}{g}_k$

$$\underset{\sim}{g}_k' = \left(g^{(1)}(\theta), g^{(2)}(\theta), \ldots, g^{(k)}(\theta)\right), \quad V_k = ((v_{ij}))_{\substack{i=1(1)k \\ j=1(1)k}}$$

Theorem: $\{\Delta_k\}$ is a non-decreasing sequence i.e. $\Delta_{k+1} \geq \Delta_k \;\forall k$.

Proof: $\Delta_{k+1} = \underset{\sim}{g}_{k+1}' V_{k+1}^{-1} \underset{\sim}{g}_{k+1}$, where $\underset{\sim}{g}_{k+1}' = \left(g^{(1)}(\theta), g^{(2)}(\theta), \ldots, g^{(k)}(\theta), g^{(k+1)}(\theta)\right)$

$$V_{k+1} = \left(\begin{array}{cccc|c} v_{11} & v_{12} & \cdots & v_{1k} & v_{1,k+1} \\ v_{21} & v_{22} & \cdots & v_{2k} & v_{2,k+1} \\ \cdots & & & & \cdots \\ v_{k1} & v_{k2} & \cdots & v_{kk} & \\ \hline v_{k+1,1} & v_{k+1,2} & \cdots & v_{k+1,k} & v_{k+1,k+1} \end{array}\right) = \left(\begin{array}{c|c} V_k & \underset{\sim}{v}_k \\ \hline \underset{\sim}{v}_k' & v_{k+1,k+1} \end{array}\right).$$

$$\underset{\sim}{v}_k' = (v_{k+1,1}, v_{k+1,2}, \ldots, v_{k+1,k})$$

Let $C^{\overline{k+1} \times \overline{k+1}}$ be any non-singular matrix defined as $C = \begin{pmatrix} I_k & 0 \\ -\underset{\sim}{v}_k' V_k^{-1} & 1 \end{pmatrix}$.

$$\text{Then, } C V_{k+1} C' = \begin{pmatrix} I_k & 0 \\ -\underset{\sim}{v}_k' V_k^{-1} & 1 \end{pmatrix} \begin{pmatrix} V_k & \underset{\sim}{v}_k \\ \underset{\sim}{v}_k' & v_{k+1,k+1} \end{pmatrix} \begin{pmatrix} I_k & -V_k^{-1} \underset{\sim}{v}_k \\ 0 & 1 \end{pmatrix}$$

$$= \begin{pmatrix} V_k & 0 \\ 0 & \underbrace{v_{k+1,k+1} - \underset{\sim}{v}_k' V_k^{-1} \underset{\sim}{v}_k}_{= v_{k+1.12\cdots k}} \end{pmatrix}$$

Now $C V_{k+1} C'$ is p.d., since V_{k+1} is p.d. C is non-singular
$\Rightarrow v_{k+1.12\cdots k} > 0$.

$$\text{Now } \Delta_{k+1} = (C \underset{\sim}{g}_{k+1})' [C V_{k+1} C']^{-1} (C \underset{\sim}{g}_{k+1})$$

$$= \begin{pmatrix} \underset{\sim}{g}_k \\ g_{k+1} - \underset{\sim}{v}_k' V_k^{-1} \underset{\sim}{g}_k \end{pmatrix}' \begin{bmatrix} V_k & 0 \\ 0 & v_{k+1.12\cdots k} \end{bmatrix}^{-1} \begin{pmatrix} \underset{\sim}{g}_k \\ g_{k+1} - \underset{\sim}{v}_k' V_k^{-1} \underset{\sim}{g}_k \end{pmatrix}$$

$$= \begin{pmatrix} \underset{\sim}{g}_k \\ g_{k+1} - \underset{\sim}{v}_k' V_k^{-1} \underset{\sim}{g}_k \end{pmatrix}' \begin{bmatrix} V_k^{-1} & 0 \\ 0 & v_{k+1.12\cdots k}^{-1} \end{bmatrix} \begin{pmatrix} \underset{\sim}{g}_k \\ g_{k+1} - \underset{\sim}{v}_k' V_k^{-1} \underset{\sim}{g}_k \end{pmatrix}$$

$$= \underset{\sim}{g}_k' V_k^{-1} \underset{\sim}{g}_k + \frac{[g_{k+1} - \underset{\sim}{v}_k' V_k^{-1} \underset{\sim}{g}_k]^2}{v_{k+1.123\cdots k}}$$

$$\geq \underset{\sim}{g}_k' V_k^{-1} \underset{\sim}{g}_k = \Delta_k.$$

Note: If for some $g(\theta)$, $\nexists$ any u.estimator whose variance attains the kth bound we may try to obtain some sharper bound by considering the $(k+1)$th bound. If the kth bound is already attained by some unbiased estimator of $g(\theta)$, there is nothing to be available by considering the $(k+1)$th bound and in this case $\Delta_k = \Delta_{k+1}$. However, $\Delta_k = \Delta_{k+1}$ does not necessarily imply that the kth bound is obtained.

Examples:

1. $N(\theta, 1)$
$\downarrow$
$x_1, x_2, \ldots, x_n$ be a random sample.

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

$\nexists$ any u.e. of $g(\theta)$ whose variance attains the lower bound Δ_1.

Here $p_\theta(x) = \frac{1}{(2\pi)^{\frac{n}{2}}} e^{-\frac{1}{2}\sum_{i=1}^{n}(x_i-\theta)^2}$

$$\frac{\partial}{\partial\theta} p_\theta(x) = \frac{1}{(2\pi)^{\frac{n}{2}}} e^{-\frac{1}{2}\sum_{i=1}^{n}(x_i-\theta)^2} \cdot \sum_{i=1}^{n}(x_i-\theta)$$

$$\frac{\partial^2}{\partial\theta^2} p_\theta(x) = p_\theta(x)\left[\left\{\Sigma(x_i-\theta)\right\}^2 - n\right] \quad \text{(Check)}$$

Then, $S_1 = \frac{1}{p_\theta(x)}\frac{\partial}{\partial\theta}p_\theta(x) = \Sigma(x_i-\theta)$

$$S_2 = \frac{1}{p_\theta(x)}\frac{\partial^2}{\partial\theta^2}p_\theta(x) = \left\{\Sigma(x_i-\theta)\right\}^2 - n.$$

$E(S_1) = 0.$

$$\begin{aligned} E(S_2) &= E\left[\Sigma(x_i-\theta)\right]^2 - n \\ &= n + 0 - n \\ &= 0 \end{aligned}$$

$V_{11} = E_\theta(S_1^2) = E_\theta\left[\Sigma(x_i-\theta)\right]^2 = n$

$$\begin{aligned} V_{12} = Cov_\theta(S_1, S_2) &= E_\theta[S_1 S_2] \\ &= E\left[\left\{\Sigma(x_i-\theta)\right\}\left\{\left(\Sigma(x_i-\theta)\right)^2 - n\right\}\right] \\ &= E\left[\left\{\Sigma(x_i-\theta)\right\}^3 - n\Sigma(x_i-\theta)\right] \\ &= E\left[\Sigma(x_i-\theta)^3 + 3\sum_{i\neq j}\Sigma(x_i-\theta)^2(x_j-\theta) + \sum_{i\neq j\neq k}\Sigma\Sigma(x_i-\theta)(x_j-\theta)(x_k-\theta)\right] \\ &= 0+0+0 \\ &= V_{21} \end{aligned}$$

$$\begin{aligned} V_{22} = E(S_2^2) &= E\left[\left(\Sigma(x_i-\theta)\right)^2 - n\right]^2 \\ &= n^2 + E\left\{\Sigma(x_i-\theta)\right\}^4 - 2n\,E\left\{\Sigma(x_i-\theta)\right\}^2 \\ &= n^2 + E\Big[\Sigma(x_i-\theta)^4 + 3\sum_{i\neq j}\Sigma(x_i-\theta)^2(x_j-\theta)^2 + 4\sum_{i\neq j}\Sigma(x_i-\theta)^3(x_j-\theta) \\ &\quad + 6\sum_{i\neq j\neq k}\Sigma\Sigma(x_i-\theta)(x_j-\theta)(x_k-\theta) + \sum_{i\neq j\neq k\neq l}\Sigma\Sigma\Sigma(x_i-\theta)(x_j-\theta)(x_k-\theta)(x_l-\theta)\Big] - 2n^2 \\ &= n^2 + 3n + 3n(n-1) + 0 + 0 + 0 - 2n^2 \\ &= 2n^2. \end{aligned}$$

$\therefore V_2 = \begin{pmatrix} V_{11} & V_{12} \\ V_{21} & V_{22} \end{pmatrix} = \begin{pmatrix} n & 0 \\ 0 & 2n^2 \end{pmatrix} = n \begin{pmatrix} 1 & 0 \\ 0 & 2n \end{pmatrix}$

$V_2^{-1} = \frac{1}{n} \begin{pmatrix} 1 & 0 \\ 0 & \frac{1}{2n} \end{pmatrix}$

$g^{(1)}(\theta) = 2\theta, \quad g^{(2)}(\theta) = 2.$

$\therefore \underset{\sim}{g}_2' = (2\theta \;\; 2) = 2(\theta \;\; 1)$

Bhattacharya 2nd lower bound

$= \underset{\sim}{g}_2' V_2^{-1} \underset{\sim}{g}_2$

$= \frac{4}{n} (\theta \;\; 1) \begin{pmatrix} 1 & 0 \\ 0 & \frac{1}{2n} \end{pmatrix} \begin{pmatrix} \theta \\ 1 \end{pmatrix}$

$= \frac{4}{n} \left(\theta^2 + \frac{1}{2n}\right).$

2nd lower bound is attained by an u.e. T iff

$T - g(\theta) = \underset{\sim}{g}_2' V_2^{-1} S = \frac{2}{n} (\theta \;\; 1) \begin{pmatrix} 1 & 0 \\ 0 & \frac{1}{2n} \end{pmatrix} \begin{pmatrix} \Sigma(x_i - \theta) \\ \{\Sigma(x_i-\theta)\}^2 - n \end{pmatrix}$

$= \frac{2}{n} \cdot \left[\theta \Sigma(x_i - \theta) + \frac{1}{2n}\{\Sigma(x_i-\theta)\}^2 - \frac{1}{2}\right]$

$= \frac{2}{n} \left[n\theta(\bar{x} - \theta) + \frac{n^2}{2n} (\bar{x}-\theta)^2 - \frac{1}{2}\right]$

$= \bar{x}^2 - \frac{1}{n} - \theta^2$

i.e. $T - \theta^2 = \bar{x}^2 - \frac{1}{n} - \theta^2$

$\Rightarrow T = \bar{x}^2 - \frac{1}{n}$

$\Rightarrow E(T) = \theta^2$

$\Rightarrow T = \bar{x}^2 - \frac{1}{n}$ is an u.e. of θ^2 with its variance attaining Bhattacharya 2nd lower bound.

(H.T.) 2. $x_1, x_2, \ldots, x_n \overset{iid}{\sim} \text{Poisson}(\theta)$
Find 2nd B-l.b. to an u.e. of θ^2 and an u.e. of θ^2 attaining this lower bound.

A theorem on the attainment of Bhattacharya lower bound for an exponential family.

Consider the exponential family $\mathcal{P} = \{p_\theta(x) : \theta \in \Omega\}$, where

$$p_\theta(x) = h(x)\, e^{\psi_1(\theta) t(x) + \psi_2(\theta)}; \quad \psi_1'(\theta) \neq 0.$$

Suppose, we want to estimate a real-valued estimable function $g(\theta)$ of θ, where $g(\theta)$ is k times differentiable w.r.t. θ

Let $\hat{g}(x)$ be any unbiased estimator of $g(\theta)$, satisfying regularity condition (iv).

<u>Theorem</u>: (i) If $Var_\theta(\hat{g}(x))$ attains k^{th} Bhattacharya-l.b. but not $(k-1)^{th}$ lower bound, then $\hat{g}(x)$ is a polynomial of degree k in t.

(ii) The variance in any polynomial in t of degree k, which is an u.e. of $g(\theta)$, attains Bhattacharya k^{th} l.b.

(ii) $\Rightarrow$ If $\exists$ a k^{th} degree polynomial $\hat{g}$ in $t \ni E(\hat{g}) = g(\theta)$, then $Var_\theta(\hat{g}) = \Delta_k$ and if $\nexists$ any k th polynomial $\hat{g}$ in t, which is an unbiased estimator of $g(\theta)$, then Bhattacharya k^{th} lower bound is not attained.

<u>Example 1</u>: Let $x_1, x_2, \ldots, x_n$ are iid $\sim N(\theta, 1)$. Let $g(\theta) = \theta^2$.

$\hat{g} = \bar{x}^2 - \frac{1}{n}$ is a polynomial of degree 2 in $t = \bar{x}$.

Theorem $\Rightarrow$ $\hat{g}$ attained B. 2nd l.b.

<u>Example 2</u>: Let $x_1, x_2, \ldots, x_n$ are iid $\sim P(\theta)$.

Then, $f(\underset{\sim}{x}) = \dfrac{e^{-n\theta}\,\theta^{\Sigma x_i}}{\Pi x_i!}$; $x_i = 0, 1, \ldots, \infty$; $i = 1(1)n$.

This can be written in the exponential form with $t = \sum_{i=1}^{n} x_i \sim$ Poisson $(n\theta)$

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

$E(T^2) = n\theta + (n\theta)^2 = n\theta + n^2\theta^2 = E(T) + n^2\theta^2$.

$E(T) = n\theta$

$\Rightarrow \theta^2 = E\left(\frac{T^2 - T}{n^2}\right)$.

Thus $\frac{T^2 - T}{n^2}$ is an u.e. of θ^2 and ~~therefore~~, further, it is a polynomial of degree 2 in T.

$\Rightarrow Var_\theta\left(\frac{T^2-T}{n^2}\right)$ = Bhattacharya 2nd lower bound $= \Delta_2$.

<u>Example 3</u>: $x_1, x_2, \ldots, x_n$ iid $\sim N(\theta, 1)$.

To estimate $g(\theta) = e^{-\theta}$.

$\nexists$ any k^{th} degree polynomial in $T = \bar{x}$ which is an u.e. of $g(\theta)$.

$\Rightarrow \nexists$ any u.e. of $g(\theta)$ with variance attaining the k^{th} Bhattacharya lower bound.

<u>Proof of the Theorem</u>: We have $Var_\theta(\hat{g}(x))$ attains the k^{th} l.b. but not the $(k-1)^{th}$ lower bound iff $\hat{g}(x)$ can be written as a linear combination of $S_1, S_2, \ldots, S_k$, but not of $S_1, S_2, \ldots, S_{k-1}$ with probability 1.

i.e. $\hat{g}(x) = a_0(\theta) + \sum_{i=1}^{k} a_i(\theta) S_i$, where $a_k(\theta) \neq 0$

$$S_1 = \frac{1}{p_\theta(x)} \frac{\partial}{\partial\theta} p_\theta(x)$$

$$= \frac{1}{p_\theta(x)} \cdot p_\theta(x) \left\{ t(x) \cdot \psi_1'(\theta) + \psi_2'(\theta) \right\}$$

$$= \psi_1'(\theta)\, t(x) + \psi_2'(\theta) \longrightarrow \text{polynomial of degree 1 in } t(x).$$

$$S_2 = \frac{1}{p_\theta(x)} \cdot \frac{\partial^2}{\partial\theta^2} p_\theta(x)$$

$$= \frac{1}{p_\theta(x)} \cdot \frac{\partial}{\partial\theta} \left[p_\theta(x) \{ \psi_1'(\theta) t(x) + \psi_2'(\theta) \} \right]$$

$$= \frac{1}{p_\theta(x)} \cdot \left[p_\theta(x) \{ \psi_1'(\theta) t(x) + \psi_2'(\theta) \}^2 + p_\theta(x) \{ \psi_1''(\theta) \cdot t(x) + \psi_2''(\theta) \} \right]$$

$$= \left[\psi_1'(\theta) t(x) + \psi_2'(\theta) \right]^2 + \left[\psi_1''(\theta) \cdot t(x) + \psi_2''(\theta) \right].$$

In general,

$$S_i = \left[\psi_1'(\theta) \cdot t(x) + \psi_2'(\theta) \right]^i + P_{i-1}(t(x), \theta),$$

where $P_{i-1}(t(x), \theta)$ = polynomial in $t(x)$ of degree at most $(i-1)$.

$$= \sum_{j=0}^{i-1} \beta_{ij}(\theta) \cdot t^j, \text{ (say)}$$

$$\therefore S_i = \left[\psi_1'(\theta) \cdot t + \psi_2'(\theta) \right]^i + \sum_{j=0}^{i-1} \beta_{ij}(\theta) t^j$$

$$= \sum_{j=0}^{i} \binom{i}{j} t^j \psi_1'^{\,j} \psi_2'^{\,i-j} + \sum_{j=0}^{i-1} \beta_{ij}(\theta) t^j$$

So, we get

$$\hat{g}(x) = a_0(\theta) + \sum_{i=1}^{k} a_i(\theta) \left\{ \sum_{j=0}^{i} \binom{i}{j} t^j \psi_1'^{\,j} \psi_2'^{\,i-j} + \sum_{j=0}^{i-1} \beta_{ij}(\theta) t^j \right\}$$

$$= \sum_{j=0}^{k} \left\{ \sum_{i=j}^{k} a_i(\theta) \binom{i}{j} t^j \psi_1'^{\,j} \psi_2'^{\,i-j} \right\} + \sum_{j=0}^{k-1} t^j \sum_{i=j+1}^{k} \beta_{ij}(\theta) \quad \text{for } 0 \le j \le i \le k$$

[Interchanging the order]

$$= \sum_{j=0}^{k} t^j \left\{ \sum_{i=j}^{k} a_i(\theta) \binom{i}{j} \psi_1'^{\,j} \psi_2'^{\,i-j} \right\} + \sum_{j=0}^{k-1} t^j \sum_{i=j+1}^{k} \beta_{ij}(\theta)$$

$$= a_k(\theta) \psi_1'^{\,k} t^k + \sum_{j=0}^{k-1} t^j \left\{ \sum_{i=j}^{k} a_i(\theta) \binom{i}{j} \psi_1'^{\,j} \psi_2'^{\,i-j} + \sum_{i=j+1}^{k} \beta_{ij}(\theta) \right\}$$

$$= \sum_{j=0}^{k} c_j t^j$$

Here, $C_k = a_k(\theta) \psi_1' \neq 0$, since $\psi_1' \neq 0$, $a_k(\theta) \neq 0$.

Hence, $\hat{g}(x)$ is a polynomial of degree k in $t(x)$.

Uses of these Lower bounds

1. These lower bounds give idea about the maximum precision i.e. inverse of variance which we can expect in estimating $g(\theta)$ unbiasedly.
2. They give a characterization of MVUE (or UMVUE). Suppose $\exists$ an u.e. T_0 of $g(\theta)$ such that $\text{Var}_\theta(T_0) = \Delta_k$ for some k, then $\text{Var}_\theta(T) \geq \text{Var}_\theta(T_0)$ for any u.e. T of $g(\theta)$ and this imply T_0 is UMVUE of $g(\theta)$.

Limitations of the lower bounds

1. The lower bounds are valid only under regularity conditions both on $p_\theta(x)$ and $t(x)$.
2. For some $g(\theta)$, there may not exist any u.e. with variance attaining kth lower bound for some k though MVUE exists. Then the method is fixed. e.g. say $x_1, x_2, \cdots, x_n \overset{iid}{\sim} N(\theta,1)$ and $g(\theta) = e^{-\theta}$. $\nexists$ any u.e. of $g(\theta)$ whose variance attains kth lower bound for some k but MVUE of $g(\theta)$ exists.

Limitations of lower bounds in finding UMVUE leads to other methods for finding UMVUE of an estimable function $g(\theta)$.

Method of Covariance

Consider the family $\mathcal{P} = \{p_\theta(x);\ \theta \in \Omega\}$ of pdfs. [Here θ may be vector valued, unlike before, where θ was a scalar quantity].

Suppose we want to estimate $g(\theta)$, a real-valued estimable function of θ.

Let U_g = Class of all u.e.'s of $g(\theta)$ with finite variance

$$= \{t(x) / E_\theta\{t(x)\} = g(\theta),\ Var_\theta(t(x)) < \infty;\ \forall \theta \in \Omega\}$$

So our problem is to find the best estimator (i.e. with least variance) of $g(\theta)$ in U_g.

Let U_0 = class of all u.e.'s of zero with finite variance

$$= \{h(x) \mid E_\theta\{h(x)\} = 0,\ Var_\theta(h(x)) < \infty,\ \forall \theta \in \Omega\}$$

Theorem: An estimator $T \in U_g$ will be UMVUE iff

$$Cov(T, h) = 0 \text{ for all } h \in U_0 \text{ and } \forall \theta \in \Omega.$$

Proof: "If part"

Let $T \in U_g$ be $\ni Cov_\theta(T,h) = 0\ \forall h \in U_0$ and $\forall \theta \in \Omega$.

Consider any other estimator $T^* \in U_g$

Then, $E_\theta(T^* - T) = g(\theta) - g(\theta) = 0.$

$\Rightarrow T^* - T \in U_0$

$\Rightarrow Cov_\theta(T, T^* - T) = 0\ \forall \theta \in \Omega$

$$\therefore Var_\theta(T^*) = Var_\theta(T + T^* - T)$$
$$= Var_\theta(T) + Var_\theta(T^* - T) + 2\cdot Cov(T, T^* - T)$$
$$= Var_\theta(T) + Var_\theta(T^* - T)$$
$$\geq Var_\theta(T).$$

$\Rightarrow T$ is UMVUE of $g(\theta)$.

"Only if part"

Let $T \in U_g$ be UMVUE of $g(\theta)$.

Then $Var_\theta(T) \leq Var_\theta(T^*)\ \forall \theta,\ \forall T^* \in U_g$.

Let us take $T^* = T + \epsilon . h$, for any $h \in U_0$, and ϵ is any given non-zero constant,

Then, $E_\theta(T^*) = E_\theta(T) = g(\theta), \forall \theta$.

$\Rightarrow T^* \in U_g$

Now, $Var_\theta(T) \le Var_\theta(T^*) = Var_\theta(T + \epsilon . h) = Var_\theta(T) + \epsilon^2 . Var_\theta(h) + 2.\epsilon\, Cov_\theta(T,h)$

$\Rightarrow \epsilon [\epsilon . Var_\theta(h) + 2.Cov(T,h)] \geq 0$

$\Rightarrow \epsilon\, Var_\theta(h) + 2.Cov(T,h) \geq 0$ if $\epsilon > 0$ ---- (*)

and $\epsilon . Var_\theta(h) + 2.Cov(T,h) \le 0$ if $\epsilon < 0$ ---- (**)

Let $\epsilon \to 0$ through positive values

Then (*) $\Rightarrow 2\, Cov(T,h) \geq 0$

i.e. $Cov(T,h) \geq 0$ --- (1)

Similarly, let $\epsilon \to 0^-$

Then (**) $\Rightarrow 2\, Cov.(T,h) \le 0$

or, $Cov.(T,h) \le 0$ --- (2)

(1) & (2) implies $Cov(T,h) = 0$ (proved).

Corollary 1: Let T be UMVUE of $g(\theta)$ and T' be any other u.e. of $g(\theta)$. Then, $Cov_\theta(T,T') > 0$ i.e. $\rho_\theta(T,T') > 0$.

Proof: Since T and T' are unbiased estimators of $g(\theta)$,

$E_\theta(T) = E_\theta(T') = g(\theta)\ \forall \theta$.

$\Rightarrow T - T' \in U_0$

Since T is MVUE, $Cov_\theta(T, T-T') = 0$.

i.e. $Var_\theta(T) = Cov(T,T')$

$$\therefore \rho_\theta(T,T') = \frac{Cov_\theta(T,T')}{\sqrt{Var_\theta(T)}\sqrt{Var_\theta(T')}} = \sqrt{\frac{Var_\theta(T)}{Var_\theta(T')}} > 0$$

$$= \sqrt{e_\theta(T',T)},$$

where $e_\theta(T',T)$ = efficiency of T' w.r.t. T.

Corollary 2: MVUE of $g(\theta)$ is unique.

Proof: If possible, let T and T' be two MVUE of $g(\theta)$.

Then, $Var_\theta(T) = Var_\theta(T')\ \forall \theta$

$\Rightarrow \rho_\theta(T,T') = 1\ \forall \theta$

$\Rightarrow T = A(\theta) + B(\theta).T'$ with probability 1, $B(\theta) > 0$.

$\Rightarrow Var_\theta(T) = B^2(\theta)\, Var_\theta(T')$

$\Rightarrow B^2(\theta) = 1$

$\Rightarrow B(\theta) = 1$ since $\rho_\theta(T,T') = 1$

$\therefore T = A(\theta) + T'$ with probability 1

$\Rightarrow E_\theta(T) = A(\theta) + E(T')$

or, $g(\theta) = A(\theta) + g(\theta)$

$\Rightarrow A(\theta) = 0$

$\therefore T = T'$ with probability 1.

i.e. MVUE of $g(\theta)$ is unique.

Corollary 3: If T is the MVUE of $g(\theta)$, then $a+bT$ is the MVUE of $a+bg(\theta)$, where a and b are given constants.

Proof: The corollary is trivial since $\exists$ a 1:1 correspondence between $g(\theta)$ and $a+bg(\theta)$.

Corollary 4: T is MVUE of $E_\theta(T)$

$\Rightarrow T^2$ is MVUE of $E_\theta(T^2)$.

Proof: Since T is MVUE of $E_\theta(T)$,

$Cov_\theta(T,h) = 0 \ \forall h \in U_0$.

i.e. $E_\theta(T,h) = 0 \ \forall h \in U_0$.

$\Rightarrow T.h \in U_0$

$\Rightarrow Cov_\theta(T, T.h) = 0 \ \forall h \in U_0$

$\Rightarrow E_\theta(TTh) = 0 \ \forall h \in U_0$

or, $E_\theta(T^2 h) = 0 \ \forall h \in U_0$

$\Rightarrow T^2$ is MVUE of its expectation.

Generalization

If T be the MVUE of $E_\theta(T)$, then T^K is MVUE of $E(T^K)$, where K is a positive integer $(K \geq 1)$.

Corollary 5: Let $T_1, T_2, \cdots, T_K$ be the MVUEs of $g_1(\theta), g_2(\theta), \cdots, g_K(\theta)$. Then $a_1T_1 + a_2T_2 + \cdots + a_KT_K$ is the MVUE of $a_1g_1(\theta) + a_2g_2(\theta) + \cdots + a_Kg_K(\theta)$.

Proof: Since T_i is MVUE of $g_i(\theta)$,

$E_\theta(T_i, h) = 0 \ \forall h \in U_0$.

$\Rightarrow E_\theta(a_iT_i, h) = 0 \ \forall h \in U_0$

$\Rightarrow E_\theta\left(\sum_{i=1}^{K} a_i T_i, h\right) = 0 \ \forall h \in U_0$

$\Rightarrow \sum_{i=1}^{K} a_i T_i$ is the MVUE of $\sum_{i=1}^{K} a_i g_i(\theta)$.

Corollary 6: T_1 is MVUE of $E_\theta(T_1)$, T_2 is MVUE of $E_\theta(T_2)$

$\Rightarrow T_1T_2$ is MVUE of $E_\theta(T_1T_2)$.

Proof: Since T_1 is MVUE of $E_\theta(T_1)$

$Cov_\theta(T_1, h) = 0 \quad \forall h \in U_0$

$\Rightarrow T_1 h \in U_0$

$\Rightarrow E_\theta(T_2, T_1 h) = 0 \quad \forall h \in U_0$, since T_2 is MVUE of $E_\theta(T_2)$.

i.e. $E_\theta(T_1 T_2 \cdot h) = 0 \quad \forall h \in U_0$

$\Rightarrow T_1 T_2$ is the MVUE of $E_\theta(T_1T_2)$.

Example: T_i is MVUE of $E_\theta(T_i)$; $i = 1(1)k$.

$\Rightarrow \sum\sum_{i,j=1}^{k} b_{ij} T_i T_j$ is MVUE of its expectation.

Proof: From corollary 6, $\left.\begin{aligned} &E(T_i T_j \cdot h) = 0 \quad \forall h \in U_0 \\ &\Rightarrow E_\theta(b_{ij} T_i T_j \cdot h) = 0 \quad \forall h \in U_0 \end{aligned}\right\} i, j = 1(1)k.$

$\Rightarrow E_\theta\left(\sum\sum_{i,j=1}^{k} b_{ij} T_i T_j \cdot h\right) = 0 \quad \forall h \in U_0.$

$\Rightarrow \sum\sum_{i,j=1}^{k} b_{ij} T_i T_j$ is MVUE of its expectation.

Corollary 7: Any polynomial in T_i is the MVUE of its expectation.

Proof: Follows from generalisation of corollary 4 and corollary 5.

Example: $P_\theta[X = -1] = \theta$, $P_\theta[X = x] = (1-\theta)^2 \theta^x$; $x = 0, 1, 2, \ldots$

Now, $h(x) \in U_0$ iff $E_\theta(h) = 0 \quad \forall \theta$

$$\text{i.e. } h(-1)\cdot\theta + (1-\theta)^2 \sum_{x=0}^{\infty} h(x)\, \theta^x = 0 \quad \forall \theta$$

$$\text{or, } \sum_{x=0}^{\infty} h(x)\, \theta^x = -\frac{\theta}{(1-\theta)^2} h(-1) \quad \forall \theta$$

$$= -\sum_{x=0}^{\infty} x \cdot \theta^x \cdot h(-1) \quad \forall \theta$$

$$\Rightarrow h(x) = -x \cdot h(-1); \; x = 0, 1, 2, \ldots \quad \ldots (*)$$

An estimator T is MVUE of its expectation iff

$$E_\theta(T \cdot h) = 0 \quad \forall h \in U_0$$

i.e. iff $t(x) h(x) = -x \cdot h(-1)\, t(-1)$; $x = 0, 1, 2, \ldots \quad (**)$.

Dividing $(**)$ by $(*)$, we get

$t(x) = t(-1)$; $x = 1, 2, \ldots$

and $t(0)$ is arbitrary.

In this case, $E_\theta(T) = t(0)\, P_\theta(x = 0) + t(-1)\, P_\theta(x \neq 0)$

$$= (1-\theta)^2 t(0) + t(-1)\left[1 - (1-\theta)^2\right]$$

$$= t(-1) + \{t(0) - t(-1)\}(1-\theta)^2 = c_1 + c_2 (1-\theta)^2, \text{ say}$$

Thus, any estimable function $g(\theta)$ admits a MVUE iff $g(\theta) = c_1 + c_2(1-\theta)^2$ for some constants c_1, c_2, and for such a $g(\theta)$ the MVUE is of the form

$$t(x) = c_1 \quad \text{for } x \neq 0$$
$$= c_1 + c_2 \quad \text{for } x = 0.$$

Consider $g(\theta) = (1-\theta)^2$. Here $c_1 = 0$, $c_2 = 1$.

$\therefore$ MVUE of $g(\theta)$ has the form

$$t(x) = 0 \quad \text{for } x \neq 0$$
$$= 1 \quad \text{" } x = 0$$

Now consider $g(\theta) = \theta$

This $g(\theta)$ cannot be put in the form $c_1 + c_2(1-\theta)^2$.

$\Rightarrow g(\theta) = \theta$ does not have any MVUE.

However, we can find an u.e. of $g(\theta)$ as

$$t(x) = 1 \quad \text{for } x = -1$$
$$= 0 \quad \text{for } x \neq -1$$

$[\text{Then } E_\theta(T) = \theta \ \forall \theta]$.

Example: suppose $y_1, y_2, \ldots, y_n$ are n uncorrelated random vectors with

$E(\underset{\sim}{y}) = A'\underset{\sim}{\beta}$, $\underset{\sim}{\beta} = \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_m \end{pmatrix}$, $\underset{\sim}{y} = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$

$\text{Disp}(\underset{\sim}{y}) = \sigma^2 I_n$

To estimate $\underset{\sim}{l}'\underset{\sim}{\beta}$.

The least squares estimate is

$\underset{\sim}{l}'\hat{\underset{\sim}{\beta}} = \lambda' A \underset{\sim}{y}$, where λ is a solution to

$AA'\underset{\sim}{\lambda} = \underset{\sim}{l}$ and $\hat{\underset{\sim}{\beta}}$ satisfies

$AA'\underset{\sim}{\beta} = A\underset{\sim}{y}$ i.e. $\hat{\beta} = (AA')^{-} A\underset{\sim}{y}$, if AA' is not of full rank

$= (AA')^{-1} A\underset{\sim}{y}$, if AA' is of full rank.

Estimate of σ^2 is $\hat{\sigma}^2 = \dfrac{S_e^2}{n-r}$, where $r = \text{Rank}(A)$

$S_e^2 = \underset{\sim}{y}'\underset{\sim}{y} - \hat{\beta}' A \underset{\sim}{y}$

Suppose $\underset{\sim}{y} \sim N_n(A'\underset{\sim}{\beta}, \sigma^2 I_n)$

Then, $E(h(\underset{\sim}{y})) = 0$

$\Rightarrow \text{const.} \int h(\underset{\sim}{y})\, e^{-\frac{1}{2\sigma^2}(\underset{\sim}{y} - A'\underset{\sim}{\beta})'(\underset{\sim}{y} - A'\underset{\sim}{\beta})} dy = 0$

$\Leftrightarrow \int h(\underset{\sim}{y})\, e^{-\frac{1}{2\sigma^2}(\underset{\sim}{y}'\underset{\sim}{y} - 2\underset{\sim}{\beta}' A\underset{\sim}{y})} dy = 0 \quad \text{---} (1)$

$\Leftrightarrow \int h(\underset{\sim}{y})\, e^{-\frac{1}{2\sigma^2}(\underset{\sim}{y}'\underset{\sim}{y} - 2\underset{\sim}{\beta}' \underset{\sim}{\theta}(\underset{\sim}{y}))} d\underset{\sim}{y} = 0$, where $\underset{\sim}{\theta}(\underset{\sim}{y}) = A\underset{\sim}{y}$.

Differentiating the above w.r.t. β_i we have

$\int h(\underset{\sim}{y})\, e^{-\frac{1}{2\sigma^2}(\underset{\sim}{y}'\underset{\sim}{y} - 2\underset{\sim}{\beta}'\theta(\underset{\sim}{y}))} \theta_i(\underset{\sim}{y})\, d\underset{\sim}{y} = 0$

$\Rightarrow E[h(\underset{\sim}{y})\, \theta_i(\underset{\sim}{y})] = 0 \quad \forall h \in U_0$

$\Rightarrow \theta_i(\underline{y})$ is MVUE of $E(\theta_i(\underline{y}))$.

$\Rightarrow \underline{\ell}'\hat{\underline{\beta}} = \lambda'\theta(\underline{y}) = \sum_{i=1}^{m} \lambda_i \theta_i(\underline{y})$ is MVUE of its expectation,

$$E(\underline{\lambda}'\underline{\theta}(\underline{y})) = \lambda' AA'\underline{\beta} = \underline{\ell}'\underline{\beta}$$

i.e. $\underline{\ell}'\hat{\underline{\beta}}$ is MVUE of $\underline{\ell}'\underline{\beta}$, under the normality assumption.

$$\begin{aligned} S_e^2 &= \underline{y}'\underline{y} - \hat{\underline{\beta}}'A\underline{y} \\ &= \underline{y}'\underline{y} - \underline{y}'A'(AA')^- A\underline{y} \\ &= \underline{y}'\underline{y} - \underline{y}'A'BA\underline{y}, \quad B = (AA')^- \text{ if } AA' \text{ is not of full rank} \\ &\qquad\qquad\qquad\qquad\quad = (AA')^{-1} \text{ if } AA' \text{ is of full rank.} \\ &= \sum y_i^2 - \sum_i\sum_j b_{ij}\theta_i(\underline{y})\theta_j(\underline{y}), \quad B = ((b_{ij})) \end{aligned}$$

Now $\theta_i(\underline{y})$ is MVUE of its expectation.

$\Rightarrow \sum_i\sum_j b_{ij}\theta_i(\underline{y})\theta_j(\underline{y})$ is MVUE of its expectation

Now differentiating (1) w.r.t. σ^2, we have,

$$\text{Const} \int h(y)\, e^{-\frac{1}{2\sigma^2}(\underline{y}'\underline{y} - 2\underline{\beta}'A\underline{y})} (\underline{y}'\underline{y} - 2\underline{\beta}'A\underline{y})\, d\underline{y} = 0$$

$\Rightarrow E[h(\underline{y})\{\underline{y}'\underline{y} - 2\underline{\beta}'A\underline{y}\}] = 0$

$\Rightarrow E[h(\underline{y}).\underline{y}'\underline{y}] = 0 \quad \forall h \in U_0$

$$\text{Since } E[h(\underline{y})\underline{\beta}A\underline{y}] = E[h(\underline{y}).\sum\beta_i\theta_i(\underline{y})] = \sum\beta_i E(h(\underline{y}),\theta_i(\underline{y})) = 0.$$

$\Rightarrow \underline{y}'\underline{y}$ is MVUE of its expectation.

$\therefore S_e^2 = \underline{y}'\underline{y} - \sum_i\sum_j b_{ij}\theta_i(\underline{y})\theta_j(\underline{y})$ is MVUE of its expectation i.e. $(n-r)\sigma^2$,

$\Rightarrow s_e^2 = \frac{S_e^2}{n-r}$ is the MVUE of σ^2.

USE OF COMPLETE SUFFICIENT STATISTICS

$\mathcal{P} = \{p_\theta(x);\ \theta \in \Omega\}$, θ is real valued or vector valued.

To estimate $g(\theta)$ = an estimable real valued function of θ.

Theorem 1 (Rao-Blackwell Theorem)

Let T be a sufficient statistic of $\mathcal{P}$ and U be an u.e. of $g(\theta)$.
Define $h(T) = E(U/T)$.
Then (i) $E_\theta h(T) = g(\theta) \quad \forall \theta$

(ii) $Var_\theta h(T) \leq Var_\theta(U) \quad \forall \theta$

'=' holds iff $U = h(T)$ a.e.

Implication: Given any u.e U of $g(\theta)$, not based on T, we can always find an u.e. based on T which is uniformly better. ~~than~~ Thus to find MVUE of $g(\theta)$, we restrict ourselves to the class of u.e.'s based on T only.

Proof: T is sufficient for $\mathcal{P}$.

$\Rightarrow h(T) = E(U/T)$ is independent of θ.

$\Rightarrow h(T)$ is a sufficient statistic

(i) $g(\theta) = E_\theta(U) = E_\theta E_T(U/T) = E_\theta h(T) \quad \forall \theta$

(ii)
$$\begin{aligned} Var_\theta(U) &= E_\theta E[U - E(U|T) + E(U|T) - g(\theta) \,/ T]^2 \\ &= E_\theta Var(U|T) + Var_\theta[E(U|T)] \\ &\geq Var_\theta h(T), \text{ since } Var(U|T) \geq 0. \end{aligned}$$

'=' holds iff $E_\theta Var(U/T) = 0$

i.e. $E_\theta E[\{U - E(U|T)\}^2 / T] = 0$

i.e. $E_\theta E[\{U - h(T)\}^2 / T] = 0$

i.e. $E_\theta[U - h(T)]^2 = 0$

i.e. $U = h(T)$ with probability 1.

Hence the theorem.

Theorem 2: [Lehmann-Scheffe Theorem]

Let there exists a complete sufficient statistic T for $\mathcal{P}$. Then every estimable function $g(\theta)$ has unique MVUE and it is given by the unique u.e. of $g(\theta)$ based on T.

Implication: To find the MVUE of $g(\theta)$ $\Leftrightarrow$ to find an u.e. of $g(\theta)$ based on the complete sufficient statistic T. Also to find such an estimator, we may start any u.e. U of $g(\theta)$ and find $E(U|T)$.

Proof: $g(\theta)$ is estimable $\Rightarrow \exists$ at least one u.e. of $g(\theta)$.

$\Rightarrow \exists$ at least one u.e. of $g(\theta)$ based on T [from Theorem 1]

T is complete

$\Rightarrow \exists$ at most one u.e. of $g(\theta)$ based on T.

[If $h_1(T)$ and $h_2(T)$ be two u.e.'s of $g(\theta)$ based on T Then

$E_\theta[h_1(T) - h_2(T)] = 0 \quad \forall \theta$

$\Rightarrow h_1(T) - h_2(T) = 0$ a.e.

$\Leftrightarrow h_1(T) = h_2(T)$ a.e.]

Hence, combining the two we get $\exists$ an unique u.e. of $g(\theta)$ based on T and by Theorem 1, it is the MVUE of $g(\theta)$.

Examples:

1. $x_1, x_2, \cdots, x_n$ are results of n independent Bernoullian trials with probability of success θ.
Then, $T = \sum x_i$ is a complete sufficient statistic $\sim Bin(n, \theta)$.

(i) $g(\theta) = \theta$

$E_\theta(T) = n\theta \quad \forall \theta$

$\Rightarrow E_\theta(\frac{T}{n}) = \theta \quad \forall \theta$

$\Rightarrow \frac{T}{n}$ is MVUE of θ.

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

$$\begin{aligned} E_\theta(T^2) &= var_\theta(T) + [E_\theta(T)]^2 && \forall \theta \\ &= n\theta(1-\theta) + n^2\theta^2 && \forall \theta \\ &= n\theta - n\theta^2 + n^2\theta^2 && \forall \theta \\ &= E(T) + n\theta^2(n-1) && \forall \theta \end{aligned}$$

$\Rightarrow E_\theta\left\{\frac{T(T-1)}{n(n-1)}\right\} = \theta^2 \quad \forall \theta.$

$\Rightarrow \frac{T(T-1)}{n(n-1)}$ is the MVUE of θ^2.

(iii) $g(\theta) = Var_\theta(\text{MVUE of } \theta)$

$= Var_\theta(\frac{T}{n})$

$= \frac{\theta(1-\theta)}{n}.$

Now, the MVUE of θ and θ^2 are $\frac{T}{n}$ and $\frac{T(T-1)}{n(n-1)}$ respectively.

$\Rightarrow \frac{T}{n^2}\left[1 - \frac{T-1}{n-1}\right] = \frac{T(n-T)}{n^2(n-1)}$ is the MVUE of $g(\theta)$.

2. $X_1, X_2, \cdots, X_n$ are iid $\sim P(\theta)$, $0<\theta<\infty$

$T = \sum X_i$ is a complete sufficient statistic
$\sim P(n\theta)$

i) $g(\theta) = \theta$, $E_\theta(T) = n\theta$

$$\Rightarrow \frac{T}{n} \text{ is the MVUE of } \theta.$$

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

Now
$$E_\theta(T^2) = Var_\theta(T) + \{E_\theta(T)\}^2$$
$$= n\theta + n^2\theta^2$$
$$= E(T) + n^2\theta^2$$

$$\Rightarrow \frac{T(T-1)}{n^2} \text{ is the MVUE of } \theta^2.$$

3. Let $x_1, x_2, \cdots, x_n$ are independent observations on x having pmf

$$\pi_x(\theta) = P_\theta[X=x]\,;\ x = 0, 1, \cdots$$

To estimate $g(\theta) = \pi_\theta(r) = P_\theta[X=r]$

Suppose $\exists$ a complete sufficient statistic T.

Define

$$U = 1 \quad \text{if } x_1 = r.$$
$$= 0 \quad \text{if } x_1 \neq r.$$

$$E_\theta(U) = P_\theta[X_1 = r] = \pi_\theta(r) \quad \forall\,\theta$$

$\therefore$ The MVUE of $\pi_\theta(r)$ is

$$h(T) = E(U|T) = P[X_1 = r/T].$$

4. Let $x_1, x_2, \cdots, x_n$ be iid with common pmf

$$P_\theta[X_i = x] = a(x)\,\theta^x / f(\theta)\,;\ x = 0(1)\infty\,;\ \theta>0,$$
$$f(\theta) = \sum_{x=0}^{\infty} a(x)\,\theta^x.$$

Let $T = \sum_{i=1}^{n} x_i = t(x)$. The pmf of T is

$$P_\theta(T=t) = \sum_{\substack{(x_1,\cdots,x_n)\ni \\ t(\underset{\sim}{x})=t}} \prod_{i=1}^{n} a(x_i)\,\theta^{\sum x_i} / \{f(\theta)\}^n\,;\ t = 0, 1, 2, \cdots$$

$$= c(t,n)\,\theta^t / \{f(\theta)\}^n\,;\ t = 0, 1, 2, \cdots$$

where
$$c(t,n) = \sum_{\substack{(x_1,\cdots,x_n)\ni \\ t(x)=t}} \prod_{i=1}^{n} a(x_i)$$

~~Exer~~

<u>Exercise</u> : T is a complete sufficient statistic.

To estimate θ^r

Define $u_r(t) = 0$ if $t < r$

$$= \frac{c(t-r, n)}{c(t, n)} \quad \text{if } t \geq r$$

Then, $$E_\theta[U_r(T)] = \sum_{t=r}^{\infty} \frac{c(t-r,n)}{c(t,n)} \cdot \frac{c(t,n)\,\theta^t}{\{f(\theta)\}^n}$$

$$= \theta^r \sum_{t-r=0}^{\infty} c(t-r, n) \frac{\theta^{t-r}}{\{f(\theta)\}^n}$$

$$= \theta^r$$

$\Rightarrow U_r(T)$ is an u.e. and hence MVUE of θ^r

To estimate the variance of the MVUE of θ^r

$$Var_\theta(U_r(T)) = E_\theta[U_r(T)]^2 - \theta^{2r}$$
$$= E_\theta[\{U_r(T)\}^2 - U_{2r}(T)]$$

$\Rightarrow \{U_r(T)\}^2 - U_{2r}(T)$ is u.e. and hence MVUE of $Var_\theta(U_r(T))$.

Examples:

a. $X_1, X_2, \ldots, X_n$ iid $\sim$ Poisson (θ); $0<\theta<\infty$

$$P_\theta[X_i = x] = \frac{e^{-\theta}\theta^x}{x!}; \; x = 0,1,2,\ldots$$

$$= \frac{a(x).\,\theta^x}{f(\theta)}, \text{ where } a(x) = \frac{1}{x!}, \; f(\theta) = e^\theta.$$

$T = \sum_{i=1}^{n} X_i$ is a complete sufficient statistic.

$T \sim$ Poisson $(n\theta)$

$$P_\theta[T=t] = \frac{e^{-n\theta}(n\theta)^t}{t!}; \; t = 0,1,2,\ldots$$

$$\Rightarrow c(t,n) = \frac{n^t}{t!}$$

$$U_r(t) = \frac{c(t-r,n)}{c(t,n)} = \frac{n^{t-r}/(t-r)!}{n^t/t!} = \frac{1}{n^r}\, t(t-1)(t-2)\cdots(t-r+1)$$

and this is the MVUE of θ^r.

In particular for $r=2$, the MVUE of θ^2 is $\frac{T(T-1)}{n^2}$ and the MVUE of the variance of MVUE of θ^2 is

$$\{U_2(T)\}^2 - U_4(T) = \frac{T^2(T-1)^2}{n^4} - \frac{T(T-1)(T-2)(T-3)}{n^4} = \frac{T(T-1)}{n^4}[T^2 - T - T^2 + 5T - 6] = \frac{2T(T-1)(2T-3)}{n^4}.$$

b. $X_1, X_2, \ldots, X_n$ are iid Negative binomial with pmf

$$P_\theta[X_i = x] = \binom{k+x-1}{x}\theta^x (1-\theta)^k; \; x = 0,1,2,\ldots, \; 0<\theta<1$$

$$= \frac{a(x)\,\theta^x}{f(\theta)}, \text{ where } a(x) = \binom{k+x-1}{x}, \; f(\theta) = (1-\theta)^{-k}.$$

(H.T.) Exercise: Find the MVUE of θ^r and also the MVUE of the variance of the MVUE of θ^r.

5. $x_1, x_2, \cdots, x_n$ iid $\sim N(\theta, 1)$.

$T = \bar{x}$ is a complete sufficient statistic.

(i) To estimate $g(\theta) = \theta$

$E_\theta(\bar{x}) = \theta$

$\Rightarrow \bar{x}$ is MVUE of θ.

(ii) To estimate $g(\theta) = \theta^2$

$E(\bar{x}^2) = Var_\theta(\bar{x}) + \{E(\bar{x})\}^2$

$= \frac{1}{n} + \theta^2$

$\Rightarrow \bar{x}^2 - \frac{1}{n}$ is an u.e. and hence, the MVUE of θ^2.

(iii) To estimate $g(\theta) = e^\theta$

$\bar{x} \sim N(\theta, \frac{1}{n})$

$E(e^{t\bar{x}}) = e^{t\theta + \frac{1}{2n}t^2}$

$E(e^{\bar{x} - \frac{1}{2n}}) = e^\theta$

$\Rightarrow e^{\bar{x} - \frac{1}{2n}}$ is u.e. and hence MVUE of e^θ.

6. $x_1, x_2, \cdots, x_n$ iid $\sim N(\mu, \sigma^2)$; μ, σ^2 unknown

Let $\theta = (\mu, \sigma^2)$.

$T = (\bar{x}, \Sigma(x_i - \bar{x})^2)$ is complete sufficient.

(i) To estimate μ^2

$$E_\theta(\bar{x}^2) = Var_\theta(\bar{x}) + \{E_\theta(\bar{x})\}^2$$

$$= \frac{\sigma^2}{n} + \mu^2 = E_\theta\left[\frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n(n-1)}\right] + \mu^2$$

$$\Rightarrow E_\theta\left[\bar{x}^2 - \frac{\Sigma(x_i - \bar{x})^2}{n(n-1)}\right] = \mu^2$$

$\Rightarrow \bar{x}^2 - \frac{\Sigma(x_i - \bar{x})^2}{n(n-1)}$ is u.e. and hence MVUE of μ^2.

(ii) To estimate

$g(\theta) = \Phi\left(-\frac{\mu}{\sigma}\right)$

Let $U = 1$ if $x_1 \le 0$

$= 0$ if $x_1 > 0$.

Then, $E_\theta(U) = P_\theta[x_1 \le 0] = P_\theta\left[\frac{x_1 - \mu}{\sigma} \le -\frac{\mu}{\sigma}\right] = \Phi\left(-\frac{\mu}{\sigma}\right)$

$\Rightarrow U$ is an u.e. of $\Phi\left(-\frac{\mu}{\sigma}\right)$

$\Rightarrow$ The MVUE of $\Phi\left(-\frac{\mu}{\sigma}\right)$ is $E[U / (\bar{x}, \Sigma(x_i - \bar{x})^2)]$.

Now $E(U / \bar{x}, \Sigma(x_i - \bar{x})^2) = P[x_1 \le 0 / \bar{x}, \Sigma(x_i - \bar{x})^2]$

$$= P\left[\frac{\sqrt{n}(x_1 - \bar{x})}{\sqrt{n-1}\sqrt{\Sigma(x_i - \bar{x})^2}} \le -\frac{\sqrt{n}\,\bar{x}}{\sqrt{n-1}\sqrt{\Sigma(x_i - \bar{x})^2}} \Big/ \bar{x}, \Sigma(x_i - \bar{x})^2\right] \quad \text{----- (1)}$$

Evaluation of (1)

Lemma: Let $x_1, x_2, \cdots, x_n$ be iid $\sim N(\mu, \sigma^2)$ and $m_1, m_2, \cdots, m_n$ be n given numbers.

Define

$$Z = \frac{\sum x_i (m_i - \bar{m})}{\left\{ \sum (x_i - \bar{x})^2 \sum (m_i - \bar{m})^2 \right\}^{\frac{1}{2}}}, \text{ where, } \bar{m} = \frac{1}{n} \sum_{i=1}^{n} m_i$$

Then, Z is distributed independently of $\bar{x}$ and $\sum (x_i - \bar{x})^2$ and

$Z^2 \sim \text{Beta}\left(\frac{1}{2}, \frac{n-2}{2}\right)$.

Proof: Let $\underset{\sim}{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$. The pdf of $\underset{\sim}{x}$ is

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

Let $C^{n \times n}$ be an $\perp$ matrix defined as

$$C = \begin{pmatrix} \frac{1}{\sqrt{n}} & \frac{1}{\sqrt{n}} & \cdots & \frac{1}{\sqrt{n}} \\ \frac{m_1 - \bar{m}}{\sqrt{\sum (m_i - \bar{m})^2}} & \frac{m_2 - \bar{m}}{\sqrt{\sum (m_i - \bar{m})^2}} & \cdots & \frac{m_n - \bar{m}}{\sqrt{\sum (m_i - \bar{m})^2}} \\ - & - & - & - \\ - & - & - & - \end{pmatrix}$$

Other rows of C are so defined that C is an $\perp$ matrix and sum of elements of each row is zero.

Let $\underset{\sim}{y} = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = C \underset{\sim}{x}$.

Then $y_1 = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} x_i = \sqrt{n}\, \bar{x}$

$$y_2 = \frac{\sum x_i (m_i - \bar{m})}{\sqrt{\sum (m_i - \bar{m})^2}}$$

$|J| = 1$

Hence the pdf of $\underset{\sim}{y}$ is

$$\text{const. } e^{-\frac{1}{2\sigma^2} \left[(y_1 - \sqrt{n}\mu)^2 + \sum_{i=2}^{n} y_i^2 \right]}$$

$\Rightarrow$ y_i's are independently distributed and $y_1 \sim N(\sqrt{n}\mu, \sigma^2)$ and $y_i \sim N(0, \sigma^2)$; $i = 2(1)n$.

Now, $Z = \frac{y_2}{\sqrt{\sum (x_i - \bar{x})^2}} = \frac{y_2}{\sqrt{\sum_{i=2}^{n} y_i^2}}$; which is independent of y_1 and hence $\bar{x}$.

Now, $Z^2 = \frac{y_2^2 / \sigma^2}{y_2^2/\sigma^2 + \sum_{i=3}^{n} y_i^2 / \sigma^2}$, where $y_2^2/\sigma^2 \sim \chi_1^2$ and $\sum_{i=3}^{n} y_i^2 / \sigma^2 \sim \chi^2_{n-2}$ are distributed independently.

$\Rightarrow$ $Z^2 \sim \text{Beta}\left(\frac{1}{2}, \frac{n-2}{2}\right)$ and it is independent of $\sum_{i=2}^{n} y_i^2 = \sum_{i=1}^{n} (x_i - \bar{x})^2$

Hence the lemma.

Particular case: let $m_1 = 1-\frac{1}{n}$, $m_2 = m_3 = \cdots = m_n = -\frac{1}{n}$.

$$\sum_{i=1}^{n} m_i = 0 \Rightarrow \bar{m} = 0$$

$$\sum_{i=1}^{n}(m_i-\bar{m})^2 = \sum_{i=1}^{n} m_i^2 = 1-\frac{2}{n}+\frac{1}{n} = \frac{n-1}{n}$$

$$\sum_{i=1}^{n} x_i(m_i-\bar{m}) = \sum_{i=1}^{n} x_i m_i = x_1 - \frac{1}{n}\sum x_i = x_1-\bar{x}.$$

So, $Z = \dfrac{(x_1-\bar{x})\sqrt{n}}{\sqrt{n-1}\sqrt{\sum(x_i-\bar{x})^2}}$ and it is distributed independently of $\bar{x}$ and $\sum(x_i-\bar{x})^2$.

Further, $Z^2 \sim \text{Beta}\left(\frac{1}{2}, \frac{n-2}{2}\right)$.

Evaluation of (1)

$$(1) = P[Z \le z_0];\quad z_0 = -\frac{\sqrt{n}\,\bar{x}}{\sqrt{n-1}\sqrt{\sum(x_i-\bar{x})^2}}.$$

When $z_0 \geq 0$

$$\begin{aligned} P[Z\le z_0] &= P[|Z|\le z_0] + P[Z\le -z_0] \\ &= P[Z^2\le z_0^2] + 1 - P[Z \geq -z_0] \\ &= P[Z^2 \le z_0^2] + 1 - P[Z\le z_0] \quad [\text{Since } Z \text{ has a symmetric distn. about zero}] \end{aligned}$$

$$\begin{aligned} \Rightarrow P[Z\le z_0] &= \frac{1}{2}\left[1+P[Z^2\le z_0^2]\right] \\ &= \frac{1}{2}\left[1+I_{z_0^2}\left(\tfrac{1}{2},\tfrac{n-2}{2}\right)\right], \text{ where } I_x(m,n) = \text{const.}\int_0^x y^{m-1}(1-y)^{n-1}dy \end{aligned}$$

When $z_0 \le 0$

$$\begin{aligned} P[Z\le z_0] &= P[-Z \geq z_0^*], \text{ writing } z_0^* = -z_0 \geq 0. \\ &= P[Z\geq z_0^*], \text{ since distn. of } Z \text{ is symmetric.} \\ &= 1-\frac{1}{2}\left[1+I_{z_0^2}\left(\tfrac{1}{2},\tfrac{n-2}{2}\right)\right] \\ &= \frac{1}{2}\left[1-I_{z_0^2}\left(\tfrac{1}{2},\tfrac{n-2}{2}\right)\right] \end{aligned}$$

7. Let $x_1, x_2, \dots, x_n$ are iid $\sim R(\theta_1,\theta_2)$; $\theta = (\theta_1,\theta_2)$.

Here, $T = (x_{(1)}, x_{(n)})$ is complete sufficient statistic.

The pdf of $x_{(n)}$ is

$$\frac{n}{(\theta_2-\theta_1)^n}\{x_{(n)}-\theta_1\}^{n-1},\quad \theta_1\le x_{(n)}\le\theta_2$$

$$\Rightarrow E_\theta[x_{(n)}-\theta_1] = \frac{n}{n+1}(\theta_2-\theta_1) \text{ (check)} \quad \text{---- (1)}$$

The pdf of $x_{(1)}$ is

$$\frac{n}{(\theta_2-\theta_1)^n}\{\theta_2-x_{(1)}\}^{n-1},\quad \theta_1\le x_{(1)}\le\theta_2$$

$$\Rightarrow E_\theta[\theta_2-x_{(1)}] = \frac{n}{n+1}(\theta_2-\theta_1) \quad \text{-------- (2)}$$

i) $g(\theta) = \theta_2 - \theta_1$

① + (2) $\Rightarrow E_\theta[X_{(n)} - X_{(1)}] + (\theta_2 - \theta_1) = \frac{2n}{n+1}(\theta_2 - \theta_1) \quad \forall \theta$

$$\Rightarrow E_\theta[X_{(n)} - X_{(1)}] = \frac{n-1}{n+1}(\theta_2 - \theta_1) \quad \forall \theta.$$

$\Rightarrow \frac{n+1}{n-1}\{X_{(n)} - X_{(1)}\}$ is u.e. and hence MVUE of $\theta_2 - \theta_1$.

ii) $g(\theta) = \frac{\theta_1 + \theta_2}{2}$.

MVUE of $g(\theta)$ is $\frac{X_{(1)} + X_{(n)}}{2}$. (check)

iii) MVUE of θ_1 is $\frac{nX_{(1)} - X_{(n)}}{n-1}$ and that of θ_2 is $\frac{nX_{(n)} - X_{(1)}}{n-1}$ respectively. (check).

8. $\underset{\sim}{Y}^{n\times 1} \sim N_n(A'\underset{\sim}{\beta}^{m\times 1}, \sigma^2 I)$; $\underset{\sim}{\beta}^{m\times 1}$ and σ^2 are unknown.

$\underset{\sim}{\theta} = (\beta_1, \beta_2, \dots, \beta_m, \sigma^2)$

The pdf of $\underset{\sim}{y}$ is

$$\text{const. } e^{-\frac{1}{2\sigma^2}(\underset{\sim}{y} - A'\underset{\sim}{\beta})'(\underset{\sim}{y} - A'\underset{\sim}{\beta})}$$

Now, $(\underset{\sim}{y} - A'\underset{\sim}{\beta})'(\underset{\sim}{y} - A'\underset{\sim}{\beta}) = (\underset{\sim}{y} - A'\hat{\underset{\sim}{\beta}})'(\underset{\sim}{y} - A'\hat{\underset{\sim}{\beta}}) + (\underset{\sim}{y} - A'\underset{\sim}{\beta})'(\underset{\sim}{y} - A'\underset{\sim}{\beta}) - (\underset{\sim}{y} - A'\hat{\underset{\sim}{\beta}})'(\underset{\sim}{y} - A'\hat{\underset{\sim}{\beta}})$

where, $\hat{\beta}$ is the solution of $A'A\underset{\sim}{\beta} = A\underset{\sim}{y}$.

$$= S_e^2 + \underset{\sim}{\beta}'AA'\underset{\sim}{\beta} - 2\underset{\sim}{\beta}'A\underset{\sim}{y} - \hat{\beta}'AA'\hat{\beta} + 2\hat{\beta}'A\underset{\sim}{y}$$

$$= S_e^2 + \underset{\sim}{\beta}'AA'\underset{\sim}{\beta} - 2\underset{\sim}{\beta}'A\underset{\sim}{y} + \hat{\underset{\sim}{\beta}}'AA'\hat{\underset{\sim}{\beta}}$$

$$= S_e^2 + \underset{\sim}{\beta}'AA'\underset{\sim}{\beta} - 2\underset{\sim}{\beta}'AA'\hat{\beta} + \hat{\underset{\sim}{\beta}}'AA'\hat{\underset{\sim}{\beta}}$$

$\therefore$ The pdf of $\underset{\sim}{y}$ is

$$\text{constant } e^{-\frac{1}{2\sigma^2}[S_e^2 - 2\underset{\sim}{\beta}'AA'\hat{\underset{\sim}{\beta}} + \hat{\underset{\sim}{\beta}}'AA'\hat{\underset{\sim}{\beta}} + \underset{\sim}{\beta}'AA'\underset{\sim}{\beta}]}$$

$$= \text{constant } e^{-\frac{1}{2\sigma^2}\cdot\underset{\sim}{\beta}'AA'\underset{\sim}{\beta}} \cdot e^{-\frac{1}{2\sigma^2}[S_e^2 + \hat{\underset{\sim}{\beta}}AA'\hat{\underset{\sim}{\beta}}] + \frac{1}{\sigma^2}\sum \hat{\beta}_i \sum b_{ij}\beta_j},$$

where $AA' = ((b_{ij}))$.

$\Rightarrow (S_e^2 + \hat{\beta}'AA'\hat{\beta}, \hat{\beta}_i ; i = 1(1)m)$ is complete sufficient statistic.

(This follows from the fact that the distn. belongs to multiparameter exponential family)

$\Rightarrow (S_e^2, \hat{\beta}_i ; i = 1(1)m)$ is also complete sufficient statistic.

(i) To estimate $\underset{\sim}{\ell}'\underset{\sim}{\beta}$, a linear estimable function of $\underset{\sim}{\beta}$.

$$E(\underset{\sim}{\ell}'\hat{\beta}) = \underset{\sim}{\ell}'\underset{\sim}{\beta}$$

$\Rightarrow \underset{\sim}{\ell}'\hat{\underset{\sim}{\beta}}$ is an u.e. and hence MVUE of $\underset{\sim}{\ell}'\underset{\sim}{\beta}$.

(ii) To estimate σ^2.

$E\left(\frac{S_e^2}{n-r}\right) = \sigma^2$, where $r = \text{rank}(A)$

$\Rightarrow \frac{S_e^2}{n-r}$ is u.e. and hence MVUE of σ^2.