

$I(\theta) = -E_{\theta} \left[\frac{\partial^2}{\partial \theta^2} \log f_{\theta}(x_1) \right]$ is the Fisher information.

We can also start with

$$\frac{\partial}{\partial \theta} \log p_{\theta}(x) = \sum_{i=1}^n \frac{\partial}{\partial \theta} \log f_{\theta}(x_i) \text{ and show that}$$

$$E_{\theta} \left[\left\{ \frac{\partial}{\partial \theta} \log p_{\theta}(x) \right\}^2 \right] = n E_{\theta} \left[\left\{ \frac{\partial}{\partial \theta} \log f_{\theta}(x_1) \right\}^2 \right] \\ = n I(\theta).$$

$$\therefore \text{var}_{\theta} (T(x)) \geq \underbrace{\frac{d}{d\theta} E_{\theta} [T(x)]^2}_{n I(\theta)}$$

If T is unbiased for θ , Then $\frac{d}{d\theta} E_{\theta} [T(x)] = 1$

If T is unbiased for $a(\theta)$, Then $\frac{d}{d\theta} E_{\theta} [T(x)] = a'(\theta)$.

If $w = w(\underline{x})$ is an estimator. Then

$$\text{Var}_{\theta}(w) \geq \frac{\left\{ \frac{d}{d\theta} \mathbb{E}_{\theta}(w) \right\}^2}{I_n(\theta)},$$

where $I_n(\theta) = \mathbb{E}_{\theta} \left[\left\{ \frac{\partial}{\partial \theta} \log p_{\theta}(\underline{x}) \right\}^2 \right]$ is

The Fisher information contained in the sample $\underline{X} \sim$
 $= (X_1, \dots, X_n)$

$$\mathbb{E}_{\theta}(w) = \tau(\theta) \quad \forall \theta \in \mathbb{H}$$

So, another way of stating the Cramer-Rao lower bound:

The variance of any unbiased estimator of $\tau(\theta)$
must be greater than or equal to $\frac{\{\tau'(\theta)\}^2}{I_n(\theta)}$.

If $Z(\theta) = \theta$. Then, for any w with $I\!E_\theta(w) = \theta$ ∇_θ
 $\text{Var}_\theta(w) > \frac{1}{I_n(\theta)}$.

it must hold that

$$\text{Var}_\theta(w) > \frac{1}{I_n(\theta)}.$$

$$I_n(\theta) = I\!E_\theta \left[\left\{ \frac{\partial}{\partial \theta} \log p_\theta(\underline{x}) \right\}^2 \right]$$

$$= \text{Var}_\theta \left[\frac{\partial}{\partial \theta} \log p_\theta(\underline{x}) \right]$$

$$x_1, \dots, x_n \stackrel{i.i.d.}{\sim} f_\theta \Rightarrow p_\theta(\underline{x}) = \prod_{i=1}^n f_\theta(x_i)$$

$$\therefore \frac{\partial}{\partial \theta} \log p_\theta(\underline{x}) = \frac{\partial}{\partial \theta} \sum_{i=1}^n \log f_\theta(x_i)$$

$$= \sum_{i=1}^n \frac{\partial}{\partial \theta} \log f_\theta(x_i)$$

$$I_h(\theta) = \text{var}_\theta \left[\frac{\partial}{\partial \theta} \log b_\theta(x) \right]$$

$$= \text{var}_\theta \left[\sum_{i=1}^n \frac{\partial}{\partial \theta} \log f_\theta(x_i) \right]$$

$$= \sum_{i=1}^n \text{var}_\theta \left[\frac{\partial}{\partial \theta} \log f_\theta(x_i) \right]$$

$$= n \text{var}_\theta \left[\frac{\partial}{\partial \theta} \log f_\theta(x_1) \right]$$

$$= n \underbrace{\mathbb{E}_\theta \left[\left(\frac{\partial}{\partial \theta} \log f_\theta(x_1) \right)^2 \right]}_{\text{Fisher information in a single observation}} \quad \left[\because \mathbb{E}_\theta \left[\frac{\partial}{\partial \theta} \log f_\theta(x_1) \right] \approx 0 \text{ H.O.E.} \right]$$

$$= n I(\theta).$$

$$I_n(\theta) = n I(\theta)$$

C.R. lower bound

for one obs.

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

for n obs.

$$\frac{[z'(\theta)]^2}{n I(\theta)}$$

With addition conditions

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

In this case,

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

Ef. Let X_1, \dots, X_n iid Bernoulli (θ), $\theta \in [0, 1]$.

$$f_\theta(x) = \begin{cases} \theta & \text{if } x=1 \\ 1-\theta & \text{if } x=0 \\ 0 & \text{o.w.} \end{cases} \quad \leftarrow \text{p.m.f. of } X$$

$$= \begin{cases} \theta^x (1-\theta)^{1-x} & \text{if } x \in \{0, 1\} \\ 0 & \text{o.w.} \end{cases}$$

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

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

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

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

$$= \frac{1}{\theta^2(1-\theta)^2} I E_{\theta} [(x-\theta)^2]$$

$$= \frac{1}{\theta^2(1-\theta)^2} \text{var}_{\theta}(x) = \frac{\theta(1-\theta)}{\theta^2(1-\theta)^2}$$

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

C.R. lower bound :

$$\text{var}_{\theta}(w) \geq \frac{1}{n I(\theta)} = \frac{\theta(1-\theta)}{n} \quad \begin{matrix} \text{for any unbiased} \\ \text{estimator } \hat{\theta} \end{matrix}$$

$$\text{Check That } \text{var}_{\theta}(\bar{X}) = \frac{\theta(1-\theta)}{n}$$

So, \bar{X} achieves the min. possible variance.

Suppose that with a single observation x , we can find $g(x)$ such that

$$\text{var}_{\theta}(g(x)) = \frac{1}{I(\theta)} \quad \leftarrow \text{The C.R. lower bound with a single obs.}$$

Then, $\text{var}_{\theta} \left(\frac{1}{n} \sum_{i=1}^n g(x_i) \right) = \frac{1}{n I(\theta)} \quad \leftarrow \text{The C.R. lower bound for the entire sample.}$

Thus, $T = T(\underline{x}) = \frac{1}{n} \sum_{i=1}^n g(x_i)$ is the UMVUE of θ .

$$\sqrt{n} \left(\hat{\theta}_{MLE} - \theta \right) \xrightarrow{d} N \left(0, \frac{1}{I(\theta)} \right) \quad \left[\begin{array}{l} \text{informal} \\ \text{statement} \end{array} \right]$$

$$\hat{\theta}_{MLE} \stackrel{a}{\sim} N \left(\theta, \frac{1}{n I(\theta)} \right) \quad \left[\begin{array}{l} \uparrow \\ \text{C.R. lower bound.} \end{array} \right]$$

$$\sqrt{n} \left(g(\hat{\theta}_{MLE}) - g(\theta) \right) \xrightarrow{D} N\left(0, \frac{\{g'(\theta)\}^2}{I(\theta)}\right)$$

$$g(\hat{\theta}_{MLE}) \stackrel{a}{\sim} N\left(g(\theta), \frac{\{g'(\theta)\}^2}{n I(\theta)}\right)$$

↑ C.L lower bound for
v.E.s of $g(\theta)$

$$T_n \xrightarrow{D} T \quad T \text{ is C.W.P.I.}$$

$$T_n \xrightarrow{P} T$$

$$\text{If } T_n \xrightarrow{D} c, \text{ then } T_n \xrightarrow{P} c$$

$$\sqrt{n} (T_n - \theta) \xrightarrow{D} N(0, \sigma^2 \gamma(\theta))$$

$$T_n - \theta \xrightarrow{D} 0$$

$$T_n \xrightarrow{P} \theta$$

informal discussions.

The multiparameter case

Let $\underline{\theta} = (\theta_1, \dots, \theta_K)$ be a vector of parameters.

We are interested in $\underline{T}(\underline{\theta})$, which is a m -dimensional vector

We want to find a lower bound on the variance of an unbiased estimator of $\underline{T}(\underline{\theta})$.

Thus, the estimator $\underline{W} = \underline{W}(\underline{x})$ has to be an m -dim. vector.

$\text{Var}_{\underline{\theta}}(\underline{W}) = \underline{V}$ is an $m \times m$ matrix.

For random variables, $\text{Var}_\theta(T) \geq \text{Var}_\theta(S)$ H & E(H)

$$\Leftrightarrow \text{Var}_\theta(T) - \text{Var}_\theta(S) \geq 0 \quad \text{H & E(H)}.$$

So, we need to define " \geq " for matrices.

We are looking for a lower bound on a variance matrix, which is p.s.d.

Def: Let A & B be p.s.d. matrices. We say that

A is greater than or equal to B if $A - B$ is p.s.d.

We denote it by writing $A \succeq B$.

A is said to be strictly greater than B if $A - B$ is p.d.

We denote this by writing $A \succ B$.

Lower order between p.s.d. matrices.

$$A \succcurlyeq B \quad (\Leftrightarrow) \quad \underline{x}^T A \underline{x} \geq \underline{x}^T B \underline{x} \quad \forall \underline{x} \in \mathbb{R}^m.$$

$$V = \text{var}_\theta(\underline{w})$$

$$\underline{a}^T V \underline{a} = \underline{a}^T \text{var}_\theta(\underline{w}) \underline{a} = \text{var}_\theta(\underline{a}^T \underline{w})$$

A lower bound for V would be such that

$$\underline{a}^T V \underline{a} \geq \underline{a}^T L \underline{a} \quad \forall \underline{a}$$

\uparrow

$$\text{var}_\theta(\underline{a}^T \underline{w}) \geq \underline{a}^T L \underline{a} \quad \forall \underline{a}$$

Cramer - Rao lower bound : For any unbiased estimator

$\underline{W} = \underline{W}(X)$ & $\underline{\tau}(\underline{\theta}) \in \mathbb{R}^m$, we have

$$\text{Var}_{\underline{\theta}}(\underline{W}) \geq (\nabla_{\underline{\theta}} \underline{\tau}(\underline{\theta}))^T \{I_n(\underline{\theta})\}^{-1} (\nabla_{\underline{\theta}} \underline{\tau}(\underline{\theta}))$$

$$\nabla_{\underline{\theta}} \underline{\tau}(\underline{\theta}) = \left(\left(\frac{\partial}{\partial \theta_i} \tau_j(\underline{\theta}) \right) \right)_{i,j}^{m \times m}$$

$$I_n(\underline{\theta}) = \mathbb{E}_{\underline{\theta}} \left[\begin{array}{cc} \nabla_{\underline{\theta}} \log p_{\underline{\theta}}(X) & \nabla_{\underline{\theta}} \log p_{\underline{\theta}}(X)^T \end{array} \right]_{K \times K}$$

$$= \left(\left(\mathbb{E}_{\underline{\theta}} \left[\frac{\partial}{\partial \theta_i} \log p_{\underline{\theta}}(X) \frac{\partial}{\partial \theta_j} \log p_{\underline{\theta}}(X) \right] \right) \right)_{i,j}$$

Under additional conditions

$$\begin{aligned} I_n(\theta) &= - \mathbb{E}_\theta \left[\bar{V}_\theta^2 \log f_\theta(x) \right] \\ &= - \left(\left(\mathbb{E}_\theta \left[\frac{\partial^2}{\partial \theta_i \partial \theta_j} \log f_\theta(x) \right] \right)_{i,j} \right) \end{aligned}$$

For iid sample:

$$I_n(\theta) = n I(\theta), \text{ where}$$

$$\begin{aligned} I(\theta) &= \left(\left(\mathbb{E}_\theta \left[\frac{\partial}{\partial \theta_i} \log f_\theta(x) \frac{\partial}{\partial \theta_j} \log f_\theta(x) \right] \right)_{i,j} \right) \\ &= - \left(\left(\mathbb{E}_\theta \left[\frac{\partial^2}{\partial \theta_i \partial \theta_j} \log f_\theta(x) \right] \right)_{i,j} \right) \end{aligned}$$

Under additional
conditions.

- If T is an unbiased estimator of θ and it attains the Cramer-Rao lower bound, then it is the UMVUE of θ .

Eg. Let $X_1, \dots, X_n \stackrel{iid}{\sim} \text{Bernoulli}(\theta)$, $\theta \in [0, 1]$.

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

C.R. lower bound

$$\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i \quad \text{has} \quad \text{E}_{\theta}(\bar{X}) = \theta \quad \forall \theta \leftarrow \text{U.E.}$$

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

$$\text{Var}_{\theta}(\bar{X}) = \frac{\theta(1-\theta)}{n} \quad \forall \theta \leftarrow \begin{array}{l} \text{attains} \\ \text{the C.R.} \end{array}$$

$\therefore \bar{X}$ is the UMVUE of θ .

Eg. $X_1, \dots, X_n \stackrel{iid}{\sim} N(\mu, 1)$, $\mu \in \mathbb{R}$.

$$I(\mu) = 1$$

C.R. lower bound $\frac{1}{n}$.

$$\mathbb{E}_\mu(\bar{X}) = \mu \quad \forall \mu \in \mathbb{R} \quad \leftarrow \text{U.E.}$$

$$\text{Var}_\mu(\bar{X}) = \frac{1}{n} \quad \forall \mu \in \mathbb{R} \quad \leftarrow \text{attains C.R. lower bound.}$$

$\therefore \bar{X}$ is the UMVUE of μ .

Eg. $X_1, \dots, X_n \stackrel{iid}{\sim} N(\mu, 1)$, $\mu \in \mathbb{R}$.

We want to estimate μ^2 .

$$I(\mu) \approx 1$$

C.R. lower bound

$$\frac{\left\{ \frac{d}{d\mu} \mu^2 \right\}^2}{n} = \frac{4\mu^2}{n}$$

$$E_{\mu}(\bar{x}^2) = \underbrace{\text{var}_{\mu}(\bar{x})}_{\frac{1}{n}} + \underbrace{\{E_{\mu}(\bar{x})\}^2}_{\mu}$$

$$V(x) = E(x^2) - \{E(x)\}^2$$

$$E(x^2) = V(x) + \{E(x)\}^2$$

$$\therefore E_{\mu}(\bar{x}^2 - \frac{1}{n}) = \mu^2$$

$T = \bar{x}^2 - \frac{1}{n}$ is an U.E. of μ^2 .

$$\begin{aligned} \text{var}_{\mu}(T) &= \text{var}_{\mu}(\bar{x}^2) \\ &= E_{\mu}(\bar{x}^4) - \{E_{\mu}(\bar{x}^2)\}^2 \end{aligned}$$

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

$$z - \mu \sim N(0, \sigma^2)$$

$$\text{For } z \sim N(\mu, \sigma^2),$$

$$\begin{aligned} E(z^4) &= E((z - \mu + \mu)^4) \\ &= E((z - \mu)^4) + 4 E(z - \mu)^3 \mu + 6 E(z - \mu)^2 \mu^2 \\ &\quad + 4 E(z - \mu) \mu^3 + \mu^4 \\ &= 3\sigma^4 + 6\sigma^2\mu^2 + \mu^4 \quad [\text{if } x \sim N(0, \sigma^2), \text{ then } E(x^4) = 3\sigma^4] \end{aligned}$$

$$E_{\mu}(\bar{x}^4) = 3\sigma^4 + 6\mu^2\sigma^2 + \mu^4$$

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

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

$$\text{Var}_{\mu}(\bar{x}^2) = \frac{3}{n^2} + \frac{6\mu^2}{n} + \mu^4 - \left(\mu^2 + \frac{1}{n}\right)^2$$

$$= \frac{3}{n^2} + \frac{6\mu^2}{n} + \mu^4 - \cancel{\mu^4} - \frac{2\mu^2}{n} - \frac{1}{n^2}$$

$$= \frac{2}{n^2} + \frac{4\mu^2}{n}$$

C.R. lower bound

> 0

$\bar{x}^2 > \text{C.R. lower bound.}$

$T = \bar{x}^2 - \frac{1}{n}$ is the UMVUE of μ^2 .

But it does not attain the C.R. lower bound.

Eg. X_1, \dots, X_n iid Unif $(0, \theta)$, $\theta > 0$

C.R. lower bound

$$\frac{\theta^2}{n}.$$

$\hat{\theta}_{MLE} = X_{(n)}$ has b.d.f $f_{(n)}(y) = \frac{n y^{n-1}}{\theta^n}, 0 < y < \theta$

$$E_\theta(X_{(n)}) = \frac{n}{n+1} \theta \leftarrow \text{not unbiased.}$$

$$\text{Var}_\theta(X_{(n)}) = \frac{n}{(n+1)^2(n+2)} \theta^2 = O\left(\frac{n}{n^3}\right) = O\left(\frac{1}{n^2}\right)$$

$\ll \frac{1}{n}$

$$\text{If we use } T = \frac{n+1}{n} X_{(n)}$$

$$E_\theta(T) = \frac{n+1}{n} E_\theta(X_{(n)}) = \frac{n+1}{n} \cdot \frac{n}{n+1} \theta = \theta \quad \forall \theta$$

$$\text{Var}_\theta(T) = \left(\frac{n+1}{n}\right)^2 \text{Var}_\theta(X_{(n)}) = \frac{\theta^2}{n(n+2)} < \frac{\theta^2}{n}.$$

C.R. lower bound does not apply in this case.

Characterization of UMVUE.

Let T be an u.E. of $\theta \Rightarrow E_\theta(T) = \theta \quad \forall \theta \in \Theta$

Let Z be an estimator s.t. $E_\theta(Z) = 0 \quad \forall \theta \in \Theta$

Z is an unbiased estimator of 0 .

$\Rightarrow T + Z$ is an unbiased estimator of θ .

For any $\lambda \in \mathbb{R}$, $T + \lambda Z$ is an unbiased estimator of θ .

$$\text{var}_\theta(T + \lambda Z) = \text{var}_\theta(T) + \lambda^2 \text{var}_\theta(Z) + 2\lambda \text{cov}_\theta(T, Z)$$

Result: Let \mathcal{Z} denote the class of all unbiased estimators of 0 . Then, T is the UMVUE of $E_\theta(T)$ if and only if $\text{cov}_\theta(T, Z) = 0 \quad \forall \theta \in \Theta \quad \forall Z \in \mathcal{Z}$.

Pf: If part: we want to show that if $\text{Cov}_\theta(T, z) = 0 \quad \forall z \in \mathbb{Z}$

Then T is the UMVUE of $E_\theta(T)$.

For simplicity, write $E_\theta(T) = Z(\theta)$.

Let S be another U.E. of $Z(\theta)$. $E_\theta(S) = Z(\theta) \quad \forall \theta$

$$\therefore E_\theta(T-S) = Z(\theta) - Z(\theta) = 0 \quad \forall \theta$$

$$\therefore T-S \in \mathcal{Z}.$$

$$\therefore \text{Cov}_\theta(T, T-S) = 0 \quad \forall \theta$$

$$(=) \text{Var}_\theta(T) - \text{Cov}_\theta(T, S) = 0$$

$$(=) \text{Cov}_\theta(T, S) = \text{Var}_\theta(T) \quad \forall \theta$$

$$\text{Var}_\theta(S) = \text{Var}_\theta(T + S - T)$$

$$= \text{Var}_\theta(T) + \underbrace{\text{Var}_\theta(S-T)}_{\geq 0} + 2 \underbrace{\text{Cov}_\theta(T, S-T)}_0$$

$$\therefore \text{Var}_\theta(S) \geq \text{Var}_\theta(T). \quad \therefore \text{Cov}_\theta(T, T-S) = 0$$

Only if part we need to show that if T is the

UMVUE of $Z(\theta)$, then $\text{Cov}_\theta(T, Z) = 0 \quad \forall \theta \in \Theta$
 $\forall z \in \mathcal{Z}$

Take an arbitrary $z \in \mathcal{Z}$

For any $\lambda \in \mathbb{R}$, $E_\theta(T + \lambda z) = E_\theta(T) = Z(\theta) \quad \forall \theta \in \Theta$.

$\therefore T + \lambda z$ is an U.E. of $Z(\theta)$

Since T is the UMVUE, we must have

$$\text{Var}_\theta(T + \lambda z) \geq \text{Var}_\theta(T)$$

$$(\Rightarrow) \text{Var}_\theta(T) + \lambda^2 \text{Var}_\theta(z) + 2\lambda \text{Cov}_\theta(T, z) \geq \text{Var}_\theta(T)$$

$$(\Rightarrow) \lambda^2 \text{Var}_\theta(z) + 2\lambda \text{Cov}_\theta(T, z) \geq 0$$

$$(\Rightarrow) \lambda \text{Cov}_\theta(T, z) \geq -\frac{\lambda^2}{2} \text{Var}_\theta(z)$$

For $\lambda > 0$, $\text{Cov}_\theta(T, z) \geq -\frac{\lambda}{2} \text{Var}_\theta(z)$ has to hold for every $z \in \mathcal{Z}$.

Take $\lambda \rightarrow 0$, to set $\text{Cov}_\theta(T, Z) \geq 0$.

For $\lambda < 0$, $\text{Cov}_\theta(T, Z) \leq -\frac{\lambda}{2} \text{var}_\theta(Z)$. true for any $\lambda < 0$

Take $\lambda \rightarrow 0$, to set $\text{Cov}_\theta(T, Z) \leq 0$

$\therefore \text{Cov}_\theta(T, Z) = 0 \quad \forall \theta \in \Theta$

Let T be the UMVUE of $Z(\theta)$

S be an U.E. of $Z(\theta)$.

Then, $\text{Cov}_\theta(T, S-T) = 0 \quad \forall \theta$

$$\Rightarrow \text{Cov}_\theta(T, S) = \text{var}_\theta(T) \quad \forall \theta$$

$$\begin{aligned} \text{Corr}_\theta(T, S) &= \frac{\text{Cov}_\theta(T, S)}{\sqrt{\text{var}_\theta(T) \text{var}_\theta(S)}} = \frac{\text{var}_\theta(T)}{\sqrt{\text{var}_\theta(T) \text{var}_\theta(S)}} \\ &= \sqrt{\frac{\text{var}_\theta(T)}{\text{var}_\theta(S)}} \end{aligned}$$

$$\therefore \frac{\text{Var}_\theta(T)}{\text{Var}_\theta(S)} = \text{Corr}_\theta^2(T, S) = e(S|T) = e(S)$$

\uparrow efficiency of the estimator
 S (w.r.t. the UMVUE
 T)

$0 \leq e(S) \leq 1.$

Efficiency is inversely proportional to the variance.

The more the variance, the less efficient the estimator.

\textcircled{X} MLE's are superefficient in the sense that] info. incl.
 Their asymptotic variance matches the variance of
 the UMVUE.

Our result: A statistic T is UMVUE for its expectation, if it is uncorrelated with all unbiased estimators of 0 .

Cov. Let T_1 be the UMVUE of $\tau_1(\theta)$ and T_2 be the UMVUE of $\tau_2(\theta)$. Then, $T_1 T_2$ is the UMVUE of $I\bar{E}_\theta(T_1 T_2)$.

Pf: It is enough to show that $\text{Cov}_\theta(T_1 T_2, z) = 0 \quad \forall z \in \mathbb{Z}$
 $\vdash \quad (\Leftarrow) \quad I\bar{E}_\theta(T_1 T_2 z) = 0 \quad \forall \theta \quad \forall z \in \mathbb{Z}$
 $\qquad \qquad \qquad (\because I\bar{E}_\theta(z) = 0 \quad \forall z)$

Since T_2 is the UMVUE of $\tau_2(\theta)$, so, $I\bar{E}_\theta(T_2 z) = 0 \quad \forall \theta \in \mathbb{H}$

$\therefore T_2 z \in \mathbb{Z}$. So, $\text{Cov}_\theta(T_1, T_2 z) = 0 \quad \forall \theta \in \mathbb{H}$
 $\quad \quad \quad (\Leftarrow) \quad I\bar{E}_\theta(T_1 T_2 z) = 0 \quad \forall \theta \in \mathbb{H}$

COR. If T is the UMVUE of $Z(\theta)$, Then

(i) T^2 is the UMVUE of $E_\theta(T^2)$

(ii) T^3 is the UMVUE of $E_\theta(T^3)$

.

.

(k) T^K is the UMVUE of $E_\theta(T^K)$.

* If T_1 is the UMVUE of $Z_1(\theta)$ & T_2 is the UMVUE of $Z_2(\theta)$. Then $T_1 + T_2$ is the UMVUE of $Z_1(\theta) + Z_2(\theta)$.

Pf.: For any $z \in \mathbb{Z}$. $\text{Cov}_\theta(T_1 + T_2, z) = \text{Cov}_\theta(T_1, z) + \text{Cov}_\theta(T_2, z)$
 $= 0 \quad \forall \theta \quad (\because T_1 \text{ & } T_2 \text{ are UMVUE's})$

$\therefore T_1 + T_2$ is the UMVUE of $E_\theta(T_1 + T_2)$
 $= Z_1(\theta) + Z_2(\theta)$.

* If T is the UMVUE of $E_\theta(T)$, Then $a_0 + a_1 T + \dots + a_K T^K$ is the UMVUE of $E_\theta(a_0 + a_1 T + \dots + a_K T^K)$.

Let T be an UMVUE of $\bar{z}(\theta)$

For any unbiased estimator S of $\bar{z}(\theta)$

$$\text{Cov}_\theta(T, S) = \text{Var}_\theta(T).$$

$$\begin{aligned}\text{Var}_\theta(S - T) &= \text{Var}_\theta(S) + \text{Var}_\theta(T) - 2 \underbrace{\text{Cov}_\theta(T, S)}_{\text{Var}_\theta(T)} \\ &= \text{Var}_\theta(S) - \text{Var}_\theta(T)\end{aligned}$$

If $\text{Var}_\theta(S) = \text{Var}_\theta(T)$ [That is S is also an UMVUE]

Then, $\text{Var}_\theta(S - T) = 0$

$$\Rightarrow S = T \text{ with prob. 1.}$$

So, an UMVUE, if it exists, must be unique with prob 1.

- An estimator T is UMVUE if and only if T is uncorrelated with every unbiased estimator of θ .
- If T_1, T_2 are UMVUEs, then $T_1 T_2$ is a UMVUE.
- If T is an UMVUE, then $a_0 + a_1 T + \dots + a_k T^k$ is the UMVUE of $a_0 + a_1 E_\theta(T) + \dots + a_k E_\theta(T^k)$.
Here, a_0, a_1, \dots, a_k are known constants.

A different way of getting to the UMVUE

use sufficient statistics

Let T be an unbiased estimator of $\tau(\theta)$.

$$E_\theta(T) = \tau(\theta) \quad \forall \theta \in \mathbb{R}$$

Theorem (Rao - Blackwell) Let T be an unbiased estimator of $\tau(\theta)$ and let S be a sufficient statistic for θ . Define $T^* = \text{IE}_{\theta}(T|S)$. Then,

(a) T^* is a proper estimator.

(b) $\text{IE}_{\theta}(T^*) = \tau(\theta) \quad \forall \theta \in \mathbb{H} \leftarrow T^* \text{ is an } \underline{\text{U.E.}}$
unbiased estimator

(c) $\text{Var}_{\theta}(T^*) \leq \text{Var}_{\theta}(T) \quad \forall \theta \in \mathbb{H} \leftarrow T^* \text{ is uniformly better than } T$

Pf: S is sufficient for θ

\Rightarrow The dist of X_1, \dots, X_n given S is free of θ .

\Rightarrow The dist. of $T(X_1, \dots, X_n)$ given S is free of θ

$\Rightarrow \text{IE}_{\theta}(T|S)$ is free of θ .

$\therefore T^*$ does not involve θ . $\Rightarrow T^*$ can be used as an estimator.

$$\mathbb{E}_\theta(T^*) = \mathbb{E}_\theta(\mathbb{E}_\theta(T|S)) \quad [\mathbb{E}(X) = \mathbb{E}\mathbb{E}(X|Y)]$$

$$= \mathbb{E}_\theta(T) = z(\theta) \quad \forall \theta \in \mathbb{H}$$

$\therefore T^*$ is an unbiased estimator of θ .

$$\text{var}_\theta(T) = \mathbb{E}_\theta \text{var}_\theta(T|S) + \text{var}_\theta \underbrace{\mathbb{E}_\theta(T|S)}_{T^*}$$

$$= \mathbb{E}_\theta \underbrace{\text{var}_\theta(T|S)}_{\geq 0 \text{ w.p.}} + \text{var}_\theta(T^*)$$

$$\geq 0$$

$$\Rightarrow \text{var}_\theta(T) \geq \text{var}_\theta(T^*) \quad \forall \theta \in \mathbb{H}$$

So, T^* has uniformly smaller variance than T .

" $=$ " holds, i.e., $\text{var}_\theta(T) = \text{var}_\theta(T^*)$, iff

$$\mathbb{E}_\theta \text{var}_\theta(T|S) = 0 \iff \text{var}_\theta(T|S) = 0 \text{ w.p. 1}$$

[$\because \text{var}_\theta(T|S)$ is a non-negative r.v.

and for a non-negative r.v. Z , $\mathbb{E}(Z) = 0$ iff $Z = 0$ w.p. 1]

$$\text{var}_\theta(T|S) = 0 \text{ w.p. 1}.$$

$$\text{var}_\theta(T|S) = \mathbb{E}_\theta \left[\{ T - \mathbb{E}_\theta(T|S) \}^2 \mid S \right]$$

$$\therefore \text{var}_\theta(T|S) = 0 \text{ w.p. 1} \iff T - \mathbb{E}_\theta(T|S) = 0 \text{ w.p. 1}$$

$$\iff T = \mathbb{E}_\theta(T|S) \text{ w.p. 1}$$

$$\iff T = T^* \text{ w.p. 1}.$$

The procedure where we start with an uns. estimator T and find $T^* = \mathbb{E}_\theta(T|S)$ for a sufficient statistic S is sometimes called Rao-Blackwellization.

. Rao-Blackwellization does not give us an improvement if and only if after taking the conditional expectation, we get back the same statistic.

Eg. Let X_1, X_2 be iid $N(\theta, 1)$, $\theta \in \mathbb{R}$.

$$\bar{X} = \frac{X_1 + X_2}{2}, \quad \mathbb{E}_\theta(\bar{X}) = \theta, \quad \text{var}_\theta(\bar{X}) = \frac{1}{2}$$

$$S = X_1, \quad T^* = \mathbb{E}_\theta(\bar{X} | X_1)$$

$$\text{Then, } \mathbb{E}_\theta(T^*) = \theta \neq \theta$$

$$\text{var}_\theta(T^*) \leq \text{var}_\theta(T) \quad \forall \theta$$

$$\begin{aligned}
 T^* &= E_{\theta}(\bar{X} | X_1) = E_{\theta}\left(\frac{X_1 + X_2}{2} | X_1\right) \\
 &= E_{\theta}\left(\frac{X_1}{2} | X_1\right) + E_{\theta}\left(\frac{X_2}{2} | X_1\right) \\
 &= \frac{X_1}{2} + E_{\theta}\left(\frac{X_2}{2}\right) \quad [\because X_2 \text{ is indep of } X_1] \\
 &= \frac{X_1}{2} + \frac{\theta}{2}
 \end{aligned}$$

T^* involves The unknown parameter θ .

$\therefore T^*$ is not a statistic. X_1 is not sufficient.

• Conditioning w.r.t. a sufficient statistic is important.

Theorem. (Lehmann - Scheffe). Let T be an unbiased estimator of $\tau(\theta)$ and let S be a complete sufficient statistic for θ . Then, $E_\theta(T|S)$ is the UMVUE of $\tau(\theta)$.

Suppose that $T_1 = f_1(S)$ and $T_2 = f_2(S)$ are unbiased estimators of $\tau(\theta)$.

$$E_\theta(T_1) = \tau(\theta) = E_\theta(T_2) \quad \forall \theta \in \mathbb{H}$$

$$\therefore E_\theta(T_1 - T_2) = 0 \quad \forall \theta \in \mathbb{H}$$

$$\Rightarrow E_\theta(f_1(S) - f_2(S)) = 0 \quad \forall \theta \in \mathbb{H}$$

$$\Rightarrow P_\theta(f_1(S) - f_2(S) = 0) = 1 \quad \forall \theta \in \mathbb{H}$$

$$\Leftrightarrow P_\theta(f_1(S) = f_2(S)) = 1 \quad \forall \theta \in \mathbb{H}, \therefore T_1 = T_2 \text{ w.p. 1.}$$

Therefore, we cannot have two unbiased estimators of $\tau(\theta)$ which are functions of the complete sufficient statistic S .

Lehmann-Scheffe Theorem tells us that if we start with a function of a complete sufficient statistic, which is unbiased for $\tau(\theta)$, then it is the UMVUE of $\tau(\theta)$.

Eg. Let X_1, \dots, X_n $\stackrel{\text{iid}}{\sim}$ Bernoulli (p), $0 \leq p \leq 1$.

Then, $\sum_{i=1}^n X_i$ is complete sufficient for p

Also, $\bar{X} = \frac{\sum_{i=1}^n X_i}{n}$ is unbiased for p .

$\therefore \bar{X}$ is the UMVUE of p .

Similarly, for $X_1, \dots, X_n \stackrel{iid}{\sim} \text{Poisson}(\theta)$, $\theta > 0$,

\bar{X} is the UMVUE of θ .

for $X_1, \dots, X_n \stackrel{iid}{\sim} \text{Normal}(\theta, 1)$, $\theta \in \mathbb{R}$

\bar{X} is the UMVUE of θ .

Eg. $X_1, \dots, X_n \stackrel{iid}{\sim} \text{Uniform}(0, \theta)$, $\theta > 0$.

$$\hat{\theta}_{MLE} = X_{(n)} = \max \{X_1, \dots, X_n\}$$

$X_{(n)}$ is complete sufficient.

$$E_\theta(X_{(n)}) = \frac{n}{n+1} \theta \quad \forall \theta \quad \leftarrow \text{not unbiased.}$$

$$E_\theta \left(\underbrace{\frac{n+1}{n} X_{(n)}}_T \right) = \theta \quad \forall \theta$$

$T = \frac{n+1}{n} X_{(n)}$ is a fn of
a complete sufficient stat.
 T is unbiased.
 $\therefore T$ is the UMVUE of θ .

Tests of hypotheses

Let X_1, \dots, X_n be a sample with j.t. dist.
(pmf / pdf)

$$p_\theta(\underline{x}), \quad \theta \in \mathbb{H}$$

In tests of hypotheses, we are interested to check whether $\theta \in \mathbb{H}_0$ or $\theta \in \mathbb{H}_1$, where $\mathbb{H}_0, \mathbb{H}_1$ are non-overlapping subsets of \mathbb{H} .

The assertions that $\theta \in \mathbb{H}_0$ or $\theta \in \mathbb{H}_1$ are known as hypotheses.

A hypothesis is any statement regarding the underlying model.

In our parametric setup, the underlying model is completely specified w.r.t. the unknown parameter.

So, any statement about the underlying model will be a statement about the unknown parameter.

Eg. $X_1 - X_n \stackrel{i.i.d}{\sim} \text{Bernoulli}(p)$. $p \in [0, 1]$.

We may want to check whether $p \leq 0.05$
or $p > 0.05$