CHAPTER - III

ASYMPTOTIC PROPERTIES AND EFFICIENCY OF CONSISTENT ESTIMATORS

3.1 INTRODUCTION :

The first section of this chapter deals with definitions of some terms which are useful for further discussion.

The second section refers to the discussion of some asymptotic properties of consistent estimators, mainly the consistent asymptotical normal estimators.

The third section deals with the discussion on the contribution by Bahadur (1960) in the area of efficiency of the consistent estimators. It also includes the discussion on the concepts viz. Concentration probability and Asymptotic Relative Efficiency considered by Bahadur for comparision of efficiency of two consistent estimators.

The last section deals with the discussion on the consistency of testing procedures and related theorem.

For ready references we state the following definitions:

3.1.1 : Asymptotic Normal estimator :

A sequence T_n, $n \geq 1$ of estimators is said to be asymptotically normal for θ if there exist a sequence b_n, $n \geq 0$ of non-negative numbers as $n \longrightarrow \infty$ such that

$$b_n(T_n - \theta) \xrightarrow{D} Z^* \quad \text{where } Z^* \rightsquigarrow N(0, \nu(\theta))$$

3.1.2 : Consistent Asymptotically Normal (CAN):

Let T_n be a sequence of estimators for $g(\theta)$. Then T_n is said to be CAN for $g(\theta)$ if

$$b_n\{T_n - g(\theta)\} \xrightarrow{D} Z^* \text{ for every } \theta \in \Theta.$$

3.1.3 : Rate of convergence : $o_p(.)$ and $O_p(.)$ are called rate of convergence in probability where $X_n = o_p(Y_n)$ as $n \longrightarrow \infty$ if $X_n/Y_n \xrightarrow{P} 0$ and $X_n = O_p(Y_n)$ if there exist a constant K, 0<K<∞ such that

$$\lim_{n\to\infty} P\left[|X_n/Y_n| \leq K\right] = 1.$$

3.2 Asymptotic properties of consistent estimators :

In this section we discuss some asymptotic properties of consistent estimators.

3.2.1 Lemma: Let $b_n \uparrow \infty$ and $b_n(T_n - \theta) \longrightarrow Z^*$ and Z^* is a.s. (Almost sure) bounded. Then $T_n \xrightarrow{P} \theta$.

Proof: Consider for ε given

$$P\left\{ | T_n - \theta | > \varepsilon \right\} = P\left\{ |b_n(T_n - \theta)| > b_n\varepsilon \right\}$$

Let $2\varepsilon' \leq \varepsilon$ and $P(|Z^*| > C) < \varepsilon'$

and n_1 is such that $b_n\varepsilon > C$ for every $n > n_1$. So

$$P\left[| b_n(T_n - \theta) | > b_n \varepsilon \right] < P\left[|b_n(T_n - \theta)| > C \right]$$

BARR. BALASAHEB KHARDEKAR LIBRARY
SHIVAJI UNIVERSITY, KOLHAPUR

$< P\left[\ |Z^*| > C\ \right] + \varepsilon'$ for every $n \geq n_2$

$< 2\varepsilon'$

$< \varepsilon$ for every $n > \max(n_1, n_2)$.

This implies that T_n is consistent for θ.

Example : 3.2.1 : Let $X_1, X_2, \ldots, X_n$ be i.i.d $b(1, \theta)$ $0<\theta<1$ and $T_n = \bar{X}$.

Thus
$$\sqrt{n}\ (T_n - \theta) = \sqrt{n}\ (\bar{X} - \theta)$$
$$= \frac{1}{\sqrt{n}} \sum_{i=1}^{n} (X_i - \theta)$$
$$= \sqrt{\theta(1-\theta)}.\ \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \left[\frac{X_i - \theta}{\sqrt{\theta(1-\theta)}}\right]$$
$$\xrightarrow{D} \sqrt{\theta(1-\theta)}\ .\ Z_i$$

By using CLT where $Z \rightsquigarrow N(0, 1)$.

That is $\sqrt{n}\ (T_n - \theta) \xrightarrow{D} Z^*$, where $Z^* \rightsquigarrow N(0, \theta(1-\theta))$

and we know that $\bar{X}$ is consistent estimator of θ.

3.2.2 If $\{T_n\}$ is asymptotically normal for $g(\theta)$ and $b_n \uparrow b < \infty$ then T_n is not consistent. This can be verified as below, suppose that T_n is consistent for θ then $(T_n - \theta) \xrightarrow{P_\theta} 0$ implies that $(T_n - \theta) \xrightarrow{D} \delta_0$ that is degenerate distribution.

So $(T_n - \theta) = \frac{1}{b_n} b_n(T_n - \theta) \xrightarrow{D} \frac{1}{b} Z$

but $\frac{1}{b} Z \rightsquigarrow N(0, 1/\sigma^2)$ which is not degenerate distribution. Hence controdiction to assumption, so T_n is not consistent.

Example 3.2.2: Let $X_1, X_2, \ldots, X_n$ be i.i.d. $N(\theta, 1)$, $g(\theta) = \theta$.

We define $T_n = \sum_{i=1}^{n} W_i X_i \,/\, \sum_{i=1}^{n} W_i$ is weighted mean and $W_i > 0$.

Consider $b_n = \sqrt{\sum_{i=1}^{n} W_i}$

$$\text{So} \quad b_n(T_n - \theta) = b_n \left\{ \frac{\sum_{i=1}^{n} W_i X_i}{\sum_{i=1}^{n} W_i} - \theta \right\}$$

$$= \frac{\sqrt{\sum_{i=1}^{n} W_i}}{\sum_{i=1}^{n} W_i} \left[\sum_{i=1}^{n} W_i (X_i - \theta)\right]$$

$$= \frac{1}{\sqrt{\sum_{i=1}^{n} W_i}} \sum_{i=1}^{n} W_i (X_i - \theta) \rightsquigarrow \frac{1}{\sqrt{\sum_{i=1}^{n} W_i}} Z^*$$

By using CLT where $Z^* \rightsquigarrow N(0, \sum_{i=1}^{n} W_i^2)$. So $b_n(T_n - \theta)$ has normal distribution. This implies that T_n is asymptotically normal. Consistency of T_n depends on nature of W_i. Here $b_n^2 = \sum_{i=1}^{n} W_i$. If $W_i = 1/i^2$ for every i then $b_n^2 = \sum_{i=1}^{n} 1/i^2$ is convergent. This implies that $b_n^2 \uparrow$ constant that is $b_n \uparrow$ constant so $b_n \not\longrightarrow \infty$. Hence estimator T_n is not consistent.

While $W_i = \frac{1}{i}$ for every i then b_n^2 is divergent and so b_n is divergent. Hence T_n is CAN.

3.2.3 : T_n is consistent for $g(\theta)$ but need not be asymptotically normal. This can be varified by the following example.

Example : 3.2.3 : Let $X_1, X_2, \ldots, X_n$ be i.i.d. $U(0, \theta)$ and $T_n = X_{(n)}$ is consistent for $g(\theta) = \theta$, but T_n is not asymptotically normal.

Suppose T_n be the asymptotically normal estimator for $g(\theta) = \theta$. There exist a sequence b_n of positive real numbers. Such that $b_n(X_{(n)} - \theta) \longrightarrow Z$ where $Z \rightsquigarrow N(0, \nu(\theta))$. Note that $0 < X_{(n)} < \theta$

so $$P_\theta \left[X_{(n)} - \theta > 0\right] = 0$$

or

$$P_\theta \left[b_n(X_{(n)} - \theta) < 0\right] = 1.$$

This implies that $P_\theta \left[Z_n < 0\right] = 1$.

This does not implies that Z_n converges to 0 normal random variable in distribution. Again consider $S_n = 2\overline{X}$ is asymptotically normal for θ when $X_1, X_2, \ldots, X_n$ be i.i.d.

U(0, θ) because

$$\sqrt{n}(S_n - \theta) = \sqrt{n}\,(2\bar{X} - \theta)$$

$$= \sqrt{n}\,\frac{2}{n}\sum_{i=1}^{n}(X_i - \frac{\theta}{2})$$

$$= \frac{2}{\sqrt{n}}\sum_{i=1}^{n}(X_i - \frac{\theta}{2})$$

$$= 2.\frac{\theta}{\sqrt{12}}\left\{\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\left[\frac{X_i - \theta/2}{\theta/\sqrt{12}}\right]\right\} \xrightarrow{D} 2.\frac{\theta}{\sqrt{12}}\, Z$$

where $Z \leadsto N(0, 1)$.

So $\sqrt{n}\,(S_n - \theta) \xrightarrow{D} Z \leadsto N(0, \theta^2/3)$.

This implies that S_n is consistent for θ.

3.2.4 : A consistent estimator need not be even asymptotically unbiased estimator.

Example : 3.2.4 : Consider here $X_1, X_2, \ldots, X_n$ be i.i.d. N(θ,1).

Let $\{a_n\}$ be a positive sequence such that

$$\int_{-a_n}^{a_n}\frac{1}{\sqrt{2\pi}}\, e^{-t^2/2}\, dt = \frac{1}{n+1}$$

Note that $a_n \longrightarrow 0$ as $n \longrightarrow \infty$.

We define an estimator T_n by

$$T_n = \begin{cases} \frac{1}{n-2} \sum_{i=3}^{n} X_i + n & \text{if } \frac{X_1 - X_2}{\sqrt{2}} \in [-a_n, a_n] \\ \frac{1}{n} \sum_{i=1}^{n} X_i & \text{otherwise.} \end{cases}$$

$\{T_n\}$ is consistent but not unbiased. We shall see first consistency of T_n. Consider

$$P_\theta [\mid T_n - \theta \mid \geq \varepsilon]$$

$$= P [\mid \tfrac{1}{n-2} \sum_{i=3}^{n} X_i + n - \theta \mid \geq \varepsilon, \frac{X_1 - X_2}{\sqrt{2}} \in (-a_n, a_n)] +$$

$$P [\mid \bar{X}_n - \theta \mid \geq \varepsilon, \frac{X_1 - X_2}{\sqrt{2}} \notin (-a_n, a_n)].$$

$$\leq P [\mid \tfrac{1}{n-2} \sum_{i=3}^{n} X_i + n - \theta \mid \geq \varepsilon, \frac{X_1 - X_2}{\sqrt{2}} \in (-a_n, a_n)] +$$

$$P [\mid \bar{X}_n - \theta \mid \geq \varepsilon].$$

$$\leq P [\frac{X_1 - X_2}{\sqrt{2}} \in (-a_n, a_n)] + P [\mid \bar{X}_n - \theta \mid \geq \varepsilon]$$

$$= \frac{1}{n+1} + P [\mid \bar{X}_n - \theta \mid \geq \varepsilon]$$

But we know that as $n \longrightarrow \infty$, $\bar{X}_n \longrightarrow \theta$ so

$$P [\mid \bar{X}_n - \theta \mid \geq \varepsilon] \longrightarrow 0, \quad \text{as } n \longrightarrow \infty$$

This implies that T_n is consistent estimator of θ.

Now we see

$$E(T_n) = P[\frac{X_1 - X_2}{\sqrt{2}} \in (-a_n, a_n)] . E[\tfrac{1}{n-2} \sum_{i=3}^{n} X_i + n]$$

$$+ P[\frac{X_1 - X_2}{\sqrt{2}} \notin (-a_n, a_n)] . E[\tfrac{1}{n} \sum_{i=1}^{n} X_i]$$

$$= \frac{1}{n+1} \frac{1}{n-2} (n-2)\theta + \frac{n}{n+1} + (1 - \frac{1}{n+1}) \frac{n\theta}{n}$$

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

$$= \theta + \frac{n}{n+1} \longrightarrow \theta + 1 \quad \text{as } n \longrightarrow \infty.$$

Hence T_n is consistent estimator but not unbiased.

<u>3.2.5</u> : Let $\{T_n\}$ be CAN for θ with rate b_n then $g(T_n)$ is CAN for $g(\theta)$ with rate b_n if g is differentiable and $g'(\theta) \neq 0$ for every $\theta \in \Theta$.

<u>Proof</u> : Since T_n is given to be CAN for θ we have [I] T_n is consistent for θ that is $T_n \xrightarrow{P_\theta} \theta, \quad \forall \theta \in \Theta.$

[II] $b_n(T_n - \theta) \xrightarrow{D} N(0, \nu(\theta)), \quad \forall \theta \in \Theta.$

Let $\theta_0 \in \Theta$ in particular I & II holds for θ_0. Define a function

$$h(\theta) = \begin{cases} \dfrac{g(\theta) - g(\theta_0)}{\theta - \theta_0} - g'(\theta_0) & \text{if } \theta \neq \theta_0 \\ 0 & \text{otherwise} \end{cases} \qquad (3.2.5)$$

Note that h is continuous function at θ_0 and $h(\theta) \longrightarrow 0$ if $\theta \longrightarrow \theta_0$.

Since $h(T_n) \xrightarrow{D} 0$ and $b_n(T_n - \theta) \longrightarrow N(0, \nu(\theta))$

BY slutsky theorem (Bhat (1986)) it is clear that

$$b_n(T_n - \theta_0).\ h(T_n) \xrightarrow{P_{\theta_0}} 0$$

Thus from (3.2.5) we have

$$b_n(T_n - \theta_o).\ \ h(T_n) = b_n(g(T_n) - g(\theta_o)) - b_n(T_n - \theta_o)\, g'(\theta_o) \xrightarrow{P_{\theta_o}} 0 \qquad (3.2.6)$$

Note that II also implies that

$$b_n(T_n - \theta_o)\ \ g'(\theta_o) \xrightarrow{D} N\{0,\ (g'(\theta_o))^2\ \nu(\theta_o)]\ \text{ if } g'(\theta_o) \neq 0$$

consider from (3.2.6)

$$b_n\Big[\frac{g(T_n) - g(\theta_o)}{g'(\theta_o)}\Big] - b_n\big[T_n - \theta_o\big] \xrightarrow{P_{\theta_o}} 0$$

since $g'(\theta_o) \neq 0$.

This implies that $b_n\Big[\frac{g(T_n) - g(\theta_o)}{g'(\theta_o)}\Big] \xrightarrow{D} N(0,\ \nu(\theta_o)$

so $g'(\theta_o).\ b_n\Big[\frac{g(T_n) - g(\theta_o)}{g'(\theta_o)}\Big] \xrightarrow{D} N\big[0,\ (g'(\theta_o))^2\ \nu(\theta_o)\big]$

This implies that

$$b_n\big[g(T_n) - g(\theta_o)\big] \xrightarrow{D} N\big[0,\ (g'(\theta_o))^2\ \nu(\theta_o)\big]$$

Hence $g(T_n)$ is CAN for $g(\theta)$ with rate b_n.

<u>Example</u> : <u>3.2.5</u> : Suppose $X_1, X_2, \ldots, X_n$ be i.i.d. $N(\theta,\ 1)$.

We know here $\bar{X}$ is CAN for θ.

[I] Let $g(\theta) = e^{\theta}$.

since $g'(\theta) = e^{\theta} \neq 0,\ \forall\ \theta \in \Theta$. This implies that $e^{\bar{X}}$ is CAN for e^{θ}.

[II] Let $g(\theta) = \theta^2$

Note that $g'(\theta) = \begin{cases} 2\theta & \text{if } \theta \neq 0 \\ 0 & \text{if } \theta = 0 \end{cases}$

Hence $\sqrt{n}\,(g(\bar{X}) - g(\theta_0)) \xrightarrow{D} N(0,\ 4\theta_0^2)$ for $\theta_0 \neq 0$

and $\sqrt{n}\,[g(\bar{X}) - g(0)] = \sqrt{n}\,\bar{X}^2$ which χ^2 distribution.

That is $\bar{X}^2$ is not CAN for θ^2.

Example : 3.2.6 : Let $X_1, X_2, \ldots, X_n$ be i.i.d. $N(\theta,\ 1)$,

$g(\theta) = \text{maximum}\ (\theta,\ 0) = \frac{\theta + |\theta - 0|}{2} = \frac{\theta + |\theta|}{2}$

Maximum function is continuous and $g(\bar{X})$ is consistent for $g(\theta)$. But $|\theta|$ is not differentiable at 0. This implies that g is not differentiable at $\theta = 0$ and if for $\theta < 0$ all derivative vanishes and theorem is not applicable for $\theta = 0$. If $\theta > 0$ then in this case

$$\sqrt{n}\,[g(\bar{X}) - g(\theta)] \xrightarrow{D} N(0,\ 1)$$

$$\sqrt{n}\,[\max\ (\bar{X},\ 0) - \max_{\theta > 0}\ (0,\ \theta)] \xrightarrow{D} N(0,\ 1).$$

3.2.6 : Lemma : If $X_1, X_2, \ldots, X_n$ be i.i.d. with

$$f(x,\ \theta) = \frac{a(x) \cdot \theta^x}{g(\theta)}, \quad x = 0, 1, 2, \ldots.$$

Then $T_n = \bar{X}$ is CAN for $\frac{g'(\theta)}{g(\theta)}\,\theta$.

<u>Proof</u> : Here $E_\theta(X) = \sum_{i=0}^{\infty} \frac{a(x).\theta^x}{g(\theta)} = \frac{\theta}{g(\theta)} \sum_{x=0}^{\infty} x.a_{(x)}.\theta^{x-1}$

$$= \frac{g'(\theta)}{g(\theta)}.\theta.$$

$$E_\theta(X^2) = \frac{\sum_{i=0}^{\infty}\left[x(x-1) + x\right].a(x).\theta^x.}{g(\theta)}$$

$$= \frac{g''(\theta)}{g(\theta)}.\theta^2 + \frac{g'(\theta)}{g(\theta)}.\theta$$

so $\quad Var_\theta(X) = \frac{g'(\theta)}{g(\theta)}.\theta + \frac{g''(\theta)}{g(\theta)}.\theta^2 - \left[\frac{g'(\theta)}{g(\theta)}.\theta\right]^2$

$$= \frac{g'(\theta)}{g(\theta)}\,\theta\left[1 - \frac{g'(\theta)}{g(\theta)}.\theta\right] + \frac{g''(\theta)}{g(\theta)}.\theta^2$$

$$= K_\theta \quad \text{(say)}.$$

But we know that

$\sqrt{n}\,(T_n - \theta) \rightsquigarrow z^* = N\left[0,\ \nu(\theta)\right]$. Implies that

$$\sqrt{n}\left[\sum_{i=1}^{n}\frac{X_i}{n} - \theta\right] = \frac{1}{\sqrt{n}}\left[\sum_{i=1}^{n}(X_i - \theta)\right]$$

$$= \sqrt{K_\theta}\,\frac{1}{\sqrt{n}}\left\{\frac{\sum_{i=1}^{n}(X_i - \theta)}{\sqrt{K_\theta}}\right\}$$

$$\rightsquigarrow \sqrt{K_\theta}\,.z^*$$

so $\quad T_n = \bar{X}$ is CAN for $\frac{g'(\theta)}{g(\theta)}.\theta.$

<u>3.2.7</u> : <u>Lemma</u> : If $X_1, X_2, \ldots, X_n$ be i.i.d. with

$f(x,\theta) = C(\theta).h(x).e^{\theta x}$. Then $T_n = \bar{X}$ is CAN for $\frac{C'(\theta)}{C(\theta)}$.

Proof : Here $E_\theta(X) = \frac{C'(\theta)}{C(\theta)}$

and

$$Var_\theta(X) = \frac{\left\{ \frac{C(\theta).C''(\theta) - [C'(\theta)]^2}{[C(\theta)]^2} \right\}^2}{Var\left[\frac{\partial}{\partial\theta} \log f_\theta(x) \right]}$$

$$= \frac{[C'(\theta)]^2 - C(\theta).C''(\theta)}{[C(\theta)]^2}$$

$$= K_\theta \quad \text{(say)}.$$

But we know that

$$\sqrt{n}\,(T_n-\theta) = \sqrt{n}\left[\frac{\Sigma X_i}{n} - \frac{C'(\theta)}{C(\theta)} \right]$$

$$= \frac{1}{\sqrt{n}}\left[\sum_{i=1}^{n}(X_i - \frac{C'(\theta)}{C(\theta)} \right]$$

$$= \sqrt{K_\theta} \cdot \frac{1}{\sqrt{n}}\left[\frac{\Sigma(X_i) - \frac{C'(\theta)}{C(\theta)}}{\sqrt{K_\theta}}\right] \rightsquigarrow K_\theta \,.\, z^*$$

Hence $\bar{X}$ is CAN for $\frac{C'(\theta)}{C(\theta)}$.

3.3 EFFICIENCY OF CONSISTENT ESTIMATORS :

3.3.1 : In this section we discuss the efficiency of consistent estimators.

The definition of efficiency by Bahadur(1960) is based on the coverage probability $\varphi_n(\theta)$ of consistent estimators. It provides an estimate of the rate of approaching coverage probabilities to one.

If T_n and U_n are consistent estimators the problem is to choose one of these two. To choose the good consistent estimator with the help of variance and concentration of probabilities. Consider here

$$P_\theta \left\{ \mid T_n - g(\theta) \mid \le \varepsilon \right\} = P_\theta \left\{ \left| \frac{T_n - g(\theta)}{\lambda} \right| \le \frac{\varepsilon}{\lambda} \right\}$$

$$= P_\theta \left\{ \mid Z \mid \le \frac{\varepsilon}{\lambda} \right\}$$

$$= \Phi(\varepsilon/\lambda) - \Phi(-\varepsilon/\lambda)$$

$$= 2\,\Phi(\varepsilon/\lambda) - 1$$

$$= \varphi_n(\theta, T_n, \varepsilon)$$

This implies that $\varphi_n(\theta, T_n, \varepsilon) = 2\,\Phi(\varepsilon/\lambda) - 1$ (3.3.1)

and $\varphi_n(\theta, T_n, \varepsilon)$ is called coverage probability of θ.

(3.3.1) implies that

$$\frac{1}{2} + \frac{\varphi_n(\theta, T_n, \varepsilon)}{2} = \Phi(\varepsilon/\lambda)$$

that is $\varepsilon^{-1}\,\Phi^{-1}\left\{ \frac{1}{2} + \frac{\varphi_n(\theta, T_n, \varepsilon)}{2} \right\} = \lambda^{-1}$

so $\lambda = \varepsilon \left\{ \Phi^{-1}\left[\frac{1}{2} + \frac{\varphi_n(\theta, T_n, \varepsilon)}{2} \right] \right\}^{-1}$ (3.3.2)

Here $\lambda = \lambda_g(T_n, \varepsilon, \theta)$ is effective standard deviation and it is decreasing function of φ_n the concentration probability.

If $T_n - g(\theta) = Op(1)$ as $n \longrightarrow \infty$ then

$\lambda_g(T_n, \varepsilon, \theta) \longrightarrow 0$ as $n \longrightarrow \infty$.

For some n , ε and $g(\theta)$ if T_n and U_n are two consistent estimators and if $\varphi_n(g(\theta),T_n,\varepsilon)<\varphi_n(g(\theta),U_n,\varepsilon)$ then $\lambda_g(T_n, \varepsilon, \theta) > \lambda_g(U_n, \varepsilon, \theta)$. So U_n is efficient consistent estimator than T_n.

Example : 3.3.1 : Let $X_1, X_2, \ldots, X_n$ be i.i.d. $N(\theta, 1)$

$$T_n = \bar{X}_n \text{ and } S_n = \frac{\bar{X}_n}{2} .$$

So $$\varphi_n = P_\theta \left\{ | \bar{X}_n - \theta | < \varepsilon \right\} = P_\theta \left\{ \left| \frac{\bar{X}_n - \theta}{1/\sqrt{n}} \right| < \frac{\varepsilon}{1/\sqrt{n}} \right\}$$

$$= 2 \Phi (\sqrt{n} . \varepsilon) - 1$$

Here $\lambda = \frac{1}{\sqrt{n}}$

and $$\varphi_n = P_\theta \left\{ \left| \frac{\bar{X}_n}{2} - \theta \right| < \varepsilon \right\}$$

$$= P_\theta \left\{ \left| \frac{\frac{\bar{X}n}{2} - \theta}{\sqrt{2/n}} \right| < \frac{\varepsilon}{\sqrt{2/n}} \right\}$$

$$= 2 \Phi (\sqrt{n/2} \; \varepsilon)$$

Here $\lambda = \sqrt{2/n}$

Hence $\frac{1}{\sqrt{n}} < \sqrt{2/n}$

So $\bar{X}_n$ is more efficient consistent estimator than $\frac{n}{2}$.

Example : 3.3.2 : Let $X_1, X_2, \ldots, X_n$ be i.i.d. $U(0, \theta)$,

$T_n = X_{(n)}$ and $U_n = 2\,\overline{X}_n$.

We know that $P_\theta \left\{ \mid X_{(n)} - \theta \mid < \varepsilon \right\} = 2\,\Phi(\varepsilon/\lambda) - 1$

that is $(1 - \varepsilon/\theta)^n = 2\,\Phi(\varepsilon/\lambda) - 1$

This implies that $\lambda = \varepsilon \left[\Phi^{-1} \left\{ \frac{1}{2} + \frac{1}{2} \left(\frac{\theta-\varepsilon}{\theta}\right)^n \right\} \right]^{-1}$

similarly for $U_n = 2\,\overline{X}_n$

$$\lambda = \varepsilon \left[\Phi^{-1} \left\{ \frac{1}{2} + \frac{1}{2} \left(\frac{\theta-\varepsilon}{2\theta}\right) \right\} \right]^{-1}$$

Hence by comparing λ for $X_{(n)}$ and $2\,\overline{X}_n$, $X_{(n)}$ is more efficient consistent estimator than $2\,\overline{X}_n$.

3.3.2 : Comparison of different two consistent estimators can be done by following method also. Let $\{U_n\}$ and $\{T_n\}$ be consistent estimators for $g(\theta)$.

The upper asymptotic efficiency of T_n relative to U_n is

$$\overline{e}_g(T_n, U_n, \theta) = \overline{\lim_{\varepsilon \to 0}}\ \overline{\lim_{n \to \infty}}\ \frac{\lambda_g^2(U_n, \varepsilon, \theta)}{\lambda_g^2(T_n, \varepsilon, \theta)}$$

where $T_g(U_n, \varepsilon, \theta)$ is effective standard deviation of U_n which is solution of

$$P_\theta \left\{ \mid U_n - g(\theta) \mid > \varepsilon \right\} = 2 \left[1 - \Phi(\varepsilon/\lambda) \right] \qquad (3.3.2)$$

<u>Example</u> : <u>3.3.3</u> : Let $X_1, X_2, \ldots, X_n$ be i.i.d. $U(0, \theta)$ and

$$T_n = X_{(n)}, \quad U_n = X_{(1)} + X_{(n)}.$$

Consider for T_n, $P_\theta \left[\mid T_n - \theta \mid > \varepsilon \right]$

$$= P_\theta \left[T_n < \theta - \varepsilon \right]$$

$$= \int_0^{\theta-\varepsilon} \frac{n \cdot x^{n-1}}{\theta^n} . dx.$$

$$= (1 - \varepsilon/\theta)^n.$$

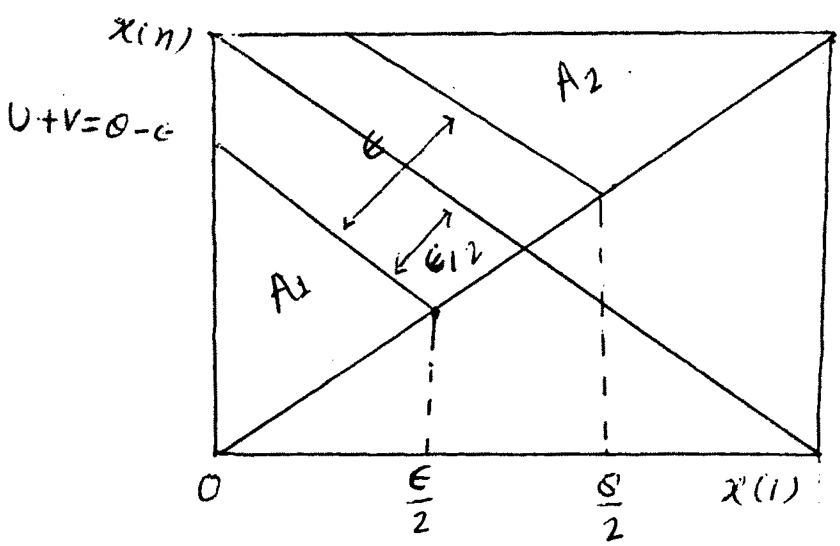
From (3.3.2) $(1 - \varepsilon/\theta)^n = 2 \left[1 - \Phi(\varepsilon/\lambda) \right]$

This implies that $\frac{\varepsilon}{\lambda} = \Phi^{-1}\left[1 - \frac{1}{2} (1 - \varepsilon/\theta)^n \right]$

so $$\lambda(T_n, \varepsilon, \theta) = \frac{\varepsilon}{\Phi^{-1}\left\{1 - \frac{1}{2}(1-\varepsilon/\theta)^n\right\}} \qquad (3.3.3)$$

Now consider $P_\theta \left\{ \mid X_{(1)} + X_{(n)} - \theta \mid > \varepsilon \right\} = P(A)$.

Here $P(A) = P(A_1) + P(A_2)$



$$P(A_1) = \int_{A_1} f(U,\ V)\ du.dv.$$

$$= \int_0^{\frac{\theta-\varepsilon}{2}} \left[\int_u^{\theta-\varepsilon-u} \frac{n(n-1)}{\theta^n} (V - U)^{n-2}\ dv \right] du.$$

$$= \frac{n(n-1}{\theta^n} \int_0^{\frac{\theta-\varepsilon}{2}} \left[\int_0^{\theta-\varepsilon-2u} t^{n-2}\ dt \right] du.$$

$$= \frac{n(n-1)}{\theta^n} \int_0^{\frac{\theta-\varepsilon}{2}} \frac{(\theta-\varepsilon-2u)^{n-1}}{(n-1)}\ du$$

$$= \frac{n}{\theta^n} \int_0^{\frac{\theta-\varepsilon}{2}} \left[\theta - \varepsilon - 2u \right]^{n-1} du.$$

$$= \frac{n}{\theta^n} \int_{\theta-\varepsilon}^{0} Z^{n-1} \left(\frac{d_z}{-2}\right)$$

$$= -\frac{1}{2} \frac{n}{\theta^n} \left[\frac{Z^n}{n} \right]_{\theta-\varepsilon}^{0} = \frac{1}{2} (1 - \varepsilon/\theta)^n.$$

Similarly $P(A_2) = \frac{1}{2} (1 - \varepsilon/\theta)^n$

Hence $P(A) = P(A_1) + P(A_2) = (1 - \varepsilon/\theta)^n.$

That is $P_\theta \left[\mid X_{(1)} + X_{(n)} - \theta \mid > \varepsilon \right] = (1 - \varepsilon/\theta)^n, \quad \text{if } \varepsilon < \theta$

....(3.3.4)

Hence from (3.3.3) and (3.3.4)

$$\lambda(T_n, \varepsilon, \theta) = \lambda(U_n, \varepsilon, \theta) \quad \forall\, n, \quad 0<\varepsilon<\theta, \quad \theta > 0.$$

This implies that $\bar{e}_\theta (T_n, U_n, \theta) = 1$.

So T_n and U_n are equally efficient consistent estimators.

3.4 SUPER EFFICIENCY OF CAN :

3.4.1 : When an sample size grows to infinity estimators which are inefficient may converge in distribution to estimators with asymptotic variances which attains the Cramer rao bound or even become smaller than that, such estimators are called super efficient. (Zack 1970)

Example : 3.4.1 : A CAN estimator which is super efficient.

Let $X_1, X_2, \ldots, X_n$ be a sequence of i.i.d. random variable having an $N(\theta, 1)$ distribution, $-\infty<\theta<\infty$. Then define a estimator

$$T_n = \begin{cases} \dfrac{\bar{X}_n}{2} & \text{if } |\bar{X}| \leq \dfrac{\log n}{\sqrt{n}} \\ \bar{X}_n & \text{otherwise} \end{cases} \tag{3.4.1}$$

Here $\bar{X}_n \longrightarrow N(\theta, \frac{1}{n})$

$$\text{so} \quad E(T_n) = \int_{-\infty}^{\infty} T_n \, f(t)\, dt$$

$$= \int_{-\frac{\log n}{\sqrt{n}}}^{\log n/\sqrt{n}} \frac{t}{2} \frac{\sqrt{n}}{\sqrt{2\pi}} e^{-n/2\,(t-\theta)^2}\, dt. + \int_{-\infty}^{-\frac{\log n}{\sqrt{n}}} - 1 \frac{\sqrt{n}}{\sqrt{2\pi}}$$

$$e^{-\frac{n}{2}(t-\theta)^2}\,dt + \int_{\frac{\log n}{\sqrt{n}}}^{\infty} t.\frac{\sqrt{n}}{\sqrt{2\pi}}\; e^{-\frac{n}{2}(t-\theta)^2}\,dt.$$

$$= \int_{-\infty}^{\infty} \frac{t}{2}.\frac{\sqrt{n}}{\sqrt{2\pi}}\; e^{-\frac{n}{2}(t-\theta)^2}\,dt + \int_{\frac{\log n}{\sqrt{n}}}^{\infty} \frac{t}{2}\,\frac{\sqrt{n}}{\sqrt{2\pi}}\; e^{-\frac{n}{2}(t-\theta)^2}\,dt$$

$$+ \int_{-\infty}^{\frac{\log n}{\sqrt{n}}} \frac{t}{2}\,\frac{\sqrt{n}}{\sqrt{2\pi}}\; e^{-\frac{n}{2}(t-\theta)^2}\,dt$$

$$= \frac{\theta}{2} + \int_{\frac{\log n}{\sqrt{n}}}^{\infty} \frac{t}{2}\, f(t)dt + \int_{-\infty}^{\frac{-\log n}{\sqrt{n}}} \frac{t}{2}\, f(t)dt \qquad (3.4.2)$$

This implies that $E(T_n) \neq \theta$.

We have to find Cramer-Rao bound that is

$$Var(T_n) \geq \varphi'(\theta)/E(-\frac{\partial^2}{\partial\theta^2}\log L).$$

Let here $E(T_n) = \varphi(\theta)$. So from (3.4.2)

$$\varphi(\theta) = \frac{\theta}{2} + \frac{1}{2}\int_{-\infty}^{\frac{-\log n}{\sqrt{n}}} t\; f(t)dt + \frac{1}{2}\int_{\frac{\log n}{\sqrt{n}}}^{\infty} t\; f(t)dt$$

$$= \frac{\theta}{2} + \frac{1}{2}\int_{-\infty}^{\frac{-\log n}{\sqrt{n}}} (t-\theta)\; f(t)dt + \int_{\frac{\log n}{\sqrt{n}}}^{\infty} (t-\theta)\; f(t)dt$$

$$+ \frac{\theta}{2} \int_{-\infty}^{\frac{-\log n}{\sqrt{n}}} f(t)dt + \frac{\theta}{2} \int_{\frac{\log n}{\sqrt{n}}}^{\infty} f(t)dt.$$

$$= \frac{\theta}{2} + \frac{1}{2} \int_{-\infty}^{\frac{-\log n}{\sqrt{n}}} (t - \theta) \frac{\sqrt{n}}{\sqrt{2\pi}} e^{-\frac{n}{2}(t-\theta)^2} dt + \int_{\frac{\log n}{\sqrt{n}}}^{\infty} (t - \theta) e^{-\frac{n}{2}(t-\theta)^2} dt$$

$$+ \frac{\theta}{2} \int_{-\infty}^{\frac{-\log n}{\sqrt{n}}} \frac{\sqrt{n}}{\sqrt{2\pi}} e^{-\frac{n}{2}(t-\theta)^2} dt + \frac{\theta}{2} \int_{\frac{\log n}{\sqrt{n}}}^{\infty} \frac{\sqrt{n}}{\sqrt{2\pi}} e^{-\frac{n}{2}(t-\theta)^2} dt$$

put $t - \theta = y$ and $\frac{\log n}{\sqrt{n}} = C_n$.

$$\text{So } \varphi(\theta) = \frac{\theta}{2} + \frac{1}{2} \left\{ \int_{-\infty}^{-C_n-\theta} \frac{\sqrt{n}}{\sqrt{2\pi}} y.\ e^{-\frac{ny^2}{2}} dy + \int_{C_n-\theta}^{\infty} \frac{\sqrt{n}}{\sqrt{2\pi}} y.e^{-\frac{ny}{2}} dy \right\}$$

$$+ \frac{\theta}{2} \left\{ \int_{-\infty}^{-C_n-\theta} \frac{\sqrt{n}}{\sqrt{2\pi}} e^{-\frac{ny^2}{2}} dy + \int_{C_n-\theta}^{\infty} \frac{\sqrt{n}}{\sqrt{2\pi}} e^{-\frac{ny^2}{2}} dy \right\}$$

$$= \frac{\theta}{2} + \frac{1}{2} \left\{ \int_{-\infty}^{-C_n-\theta} y.\ \frac{\sqrt{n}}{\sqrt{2\pi}} e^{-\frac{ny^2}{2}} dy + \int_{C_n-\theta}^{\infty} \frac{\sqrt{n}}{\sqrt{2\pi}} y.e^{-\frac{ny^2}{2}} dy \right\}$$

$$+ \frac{\theta}{2} \left\{ 1 - \Phi\left[(C_n - \theta)\sqrt{n}\right] \right\} + \frac{\theta}{2} \left\{ \Phi\left[-(-C_n - \theta)\sqrt{n}\right] \right\}$$

$$\text{Thus } \varphi'(\theta) = \frac{1}{2} + \frac{1}{2} \left\{ \frac{-\sqrt{n}}{\sqrt{2\pi}} (-C_n - \theta)\ e^{-\frac{n}{2}(-C_n-\theta)^2} \right.$$

SHIVAJI UNIVERSITY LIBRARY KOLHAPUR

$$+ \frac{\sqrt{n}}{\sqrt{2\pi}} (C_n - \theta)\, e^{-\frac{n}{2}(C_n-\theta)^2} \Big\}$$

$$+ \frac{1}{2} \left[1 - \Phi (C_n - \theta) \sqrt{n} + \Phi(- C_n - \theta) \right]$$

$$+ \frac{\theta}{2} \left\{ \varphi(C_n - \theta) \sqrt{n} - \varphi(- C_n - \theta) \right\} \quad (3.4.3)$$

Hence T_n does not attain C.R. lower bound.

Consider $A_n = \left\{ (X_1, X_2, \ldots, X_n) \; : \; | \bar{X}_n | \leq \frac{\log n}{\sqrt{n}} \right\}$

Here $\frac{\log n}{\sqrt{n}} \longrightarrow 0$ as $n \longrightarrow \infty$. (By using hospitals rule). So A_n is a decreasing to zero.

Here $\bar{X}_n \longrightarrow N(0, 1/n)$ and $C_n = \frac{\log n}{\sqrt{n}}$

So

$$P_0(A_n) = P \left[| \bar{X}_n | \leq C_n \right]$$

$$= \int_{-C_n}^{C_n} \frac{\sqrt{n}}{\sqrt{2\pi}} . e^{-\frac{nt^2}{2}} dt \quad \text{put } \sqrt{n}\, t = u$$

$$= 2 \int_0^{\log n} \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}} du.$$

$$= 2 \left[\Phi (\log n) - \frac{1}{2} \right]$$

$$= 2 \left[1 - \frac{1}{2} \right] \quad \text{as } n \longrightarrow \infty$$

$$= 1$$

Hence $P_0 (A_n) \longrightarrow 1$, as $n \longrightarrow \infty$ when $\theta = 0$

SHIVAJI UNIVERSITY, KOLHAPUR

Now $\theta = \theta_o \neq 0$

$$P_{\theta_o}(A_n) = P_{\theta_o}\left\{ |\bar{X}_n| \leq C_n \right\}$$

$$= \int_{-C_n}^{C_n} \frac{\sqrt{n}}{\sqrt{2\pi}} \cdot e^{-\frac{n}{2}(t-\theta_o)^2} \, dt$$

put $\sqrt{n}(t - \theta_o) = u$

$$= \int_{-\log n - \sqrt{n}\,\theta_o}^{\log n - \sqrt{n}\,\theta_o} \frac{1}{\sqrt{2\pi}} \quad e^{-\frac{u^2}{2}} \, . \, du.$$

$$= \Phi(\log n - \sqrt{n}\,\theta_o) - \Phi(-\log n - \sqrt{n}\,\theta_o)$$

(a) If $\theta_o > 0$ and rate of convergence of $\sqrt{n}\,\theta_o > \log n$

hence $\log n - \sqrt{n}\,\theta_o \longrightarrow -\infty$

(b) If $\theta_o < 0$

$$\log n - \sqrt{n}\,\theta_o \longrightarrow \infty$$

Hence for (a) and (b) both cases

$P_{\theta_o}(A_n) \longrightarrow 0$ as $n \longrightarrow \infty$

$$\text{so} \qquad P(A_n) \longrightarrow \begin{cases} 1 & \text{if } \theta = 0 \\ 0 & \text{if } \theta \neq 0 \end{cases} \qquad (3.4.4)$$

If $X_1, X_2, \ldots, X_n$ are independent and $E(X) < \infty$, we obtain from weak law of large numbers

$$P_\theta\{\mid T_n - \theta \mid \leq \varepsilon \,/\, (X_1, X_2, \ldots, X_n) \in A_n\} \longrightarrow \begin{cases} 1 & \text{if } \mid \theta \mid \leq \varepsilon \\ 0 & \text{if } \mid \theta \mid > \varepsilon \end{cases}$$

as $n \longrightarrow \infty$.

So from definition of T_n, A_n and the limits we have

$$P_\theta \{ \mid T_n - \theta \mid \leq \varepsilon \} \longrightarrow 1 \quad \text{for every } \theta \in (-\infty, \infty) \qquad (3.4.5)$$

But we know that

$$P_\theta\{\mid T_n - \theta \mid < \varepsilon\} = P_\theta\{\mid T_n - \theta \mid \leq \varepsilon, A_n\} + P_\theta\{\mid T_n - \theta \mid \leq \varepsilon, A'n\}$$

from (3.4.5) T_n is consistent estimator for θ and T_n can be written as $T_n = \frac{\bar{X}_n}{2} I_{An}(\bar{X}_n) + \bar{X}_n I_{An}'(\bar{X}_n)$ we know the CAN estimator as

$$\lim_{n\to\infty} P_\theta \left[\sqrt{n} \, (T_n - \theta) \leq x \right]$$

$$= \lim_{n\to\infty} \{P_{\theta=0}[\sqrt{n}. \, T_n \leq x, \, A_n] + P_{\theta=0}[\sqrt{n} \, T_n \leq x, \, A_n']\}$$

but $P_{\theta=0}[\sqrt{n} \, T_n \leq x, \, A_n''] = 0$

$$= P_0 \left[\sqrt{n} \, \frac{\bar{X}_n}{2} \leq x, \, A_n \right]$$

$$= P_0 \left[\sqrt{n} \, \bar{X}_n \leq 2x, \, A_n \right] = \Phi(2x) \qquad (3.4.6)$$

when $\theta = \theta_0 \neq 0$

$$P_{\theta_0} \left[\sqrt{n} \, (T_n - \theta_0) \leq x \right]$$

$$= P_{\theta_0}\left[\sqrt{n} \, (\tfrac{\bar{X}_n}{2} - \theta) \leq x, \, A_n\right] + P_{\theta_0}\left[\sqrt{n} \, (\bar{X}_n - \theta) \leq x, \, A_n'\right]$$

$$= \quad 0 \quad + \quad P_{\theta_o}\left[\sqrt{n}\,(\bar{X}_n - \theta_o) \le x,\ A_n'\right] = \Phi(x) \qquad (3.4.7)$$

Hence from (3.4.6) and (3.4.7).

$$\lim_{n\to\infty} P_\theta\left[\sqrt{n}\,(T_n - \theta) \le x\right] = \begin{cases} \Phi(x) & \text{if } \theta \ne 0 \\ \Phi(2x) & \text{if } \theta = 0 \end{cases} \qquad (3.4.8)$$

where $\Phi(x)$ and $\Phi(2x)$ are standard normal integrals.

Hence the asymptotic variance of T_n is $\frac{1}{n}$ if $\theta \ne 0$ and $\frac{1}{4n}$ if $\theta = 0$. The Craner Rao lower bound for each n is $\frac{1}{n}$. Thus T_n is asymptotically super efficient at $\theta = 0$.

3.5 CONSISTENCY OF TESTING OF HYPOTHESIS :

3.5.1 : Introduction : In this section we discuss the consistency of testing of hypothesis. We can judge the correctness of a null hypothesis of the parameter of a set of probability measures with increasing reliability with increasing sample size. We discuss here important concept based on "infinitely large sample size" one never rejects the 'true' null hypothesis and never accepts the 'false' one. Some definitions which are usefull for further discussion.

3.5.2 : Test of hypothesis : The mapping φ (one and zero) is said to be a test of hypothesis $H_o : \theta \in \Theta_o$ against the alternative $H_1 : \theta \in \Theta_1$ with error probability α if $E_\theta\, \varphi(X) \le \alpha \quad \forall\ \theta \in \Theta_o$.

In short φ is a test for the problem $(\alpha, \Theta_0, \Theta_1)$ let us write $\beta_\varphi(\theta) = E_\theta \varphi(X)$ we find test φ for a given α, $0 \leq \alpha \leq 1$ such that

$$\sup_{\theta \in \Theta_0} \beta_\varphi(\theta) \leq \alpha .$$

L.H.S. is the size of the test φ .

3.5.3 : Power function of the test : Let φ be a test function for the problem $(\alpha, \Theta_0, \Theta_1)$ for every $\theta \in \Theta$ define

$$\beta_\varphi(\theta) = E_\theta \varphi(X) = P_\theta \left\{ \text{Reject } H_0 \right\}.$$

As a function of θ, $\beta_\varphi(\theta)$ is called the power function of the test φ . For any $\theta \in \Theta_1$, $\beta_\varphi(\theta)$ is called the power of φ against the alternative θ.

3.5.4 : Let φ_α be the class of all tests for the problem $(\alpha, \Theta_0, \Theta_1))$. A test $\varphi_0 \in \varphi_\alpha$ is said to be a most powerful (m.p) test against an alternative $\theta \in \Theta_1$ if

$$\beta_{\varphi_0}(\theta) \geq \beta_\varphi(\theta), \quad \forall \ \varphi \in \varphi_\alpha .$$

3.5.5 : Neyman-pearson lemma (N.p.lemma) An test φ of the form

$$\varphi(X) = \begin{cases} 1 & \text{if } f_1(x) > K f_0(X) \\ r(x) & \text{if } f_1(x) = K f_0(X) \\ 0 & \text{if } f_1(x) < K f_0(X) \end{cases}$$

for some $K \geq 0$ and $0 \leq r(x) \leq 1$ is most powerful of it's size for testing $H_0 : \theta = \theta_0$ against $H_1 : \theta = \theta_1$ given α,

$0 \le \alpha \le 1$ such that $E_\theta \varphi(X) = \alpha$. is the N.P. lemma.

and if $K = \infty$ then

$$\varphi(X) = \begin{cases} 1 & \text{if } f_0(x) = 0 \\ 0 & \text{if } f_0(x) > 0 \end{cases}$$

is most powerful of size zero for testing H_0 Vs H_1 .

3.5.6 : Consistency of test can be explained as below. As increasing sample sizes in tesfting procedures we define test for such sample size, so we have sequence of tests. For these sequence of tests type one and type two errors approaches to zero as sample size increases infinitely large. This property of test is called consistency and it can be defined as below : A sequence of tests $\varphi_n(x)$ is called a consistent sequence of tests for the test problem (Θ_0, Θ_1)

$$\text{if} \quad \lim_{n\to\infty} E_\theta(\varphi_n) = 0 , \qquad \text{for } \theta \in \Theta_0$$

$$\text{and} \quad \lim_{n\to\infty} E_\theta(\varphi_n) = 1 \qquad \text{for } \theta \in \Theta_1$$

This implies that for every $\theta \in \Theta_0$ the type one error converges to zero and for every $\theta \in \Theta_1$ power converges to one (That is type two error converges to zero) as the sample size increases infinitely large.

3.5.7 : Lemma : Let P_1 and P_2 be arbitarary p bability measures overs ($\mathcal{R}$, $\mathcal{Y}$). Let f_1 and f_2 be the corresponding

Radon Nikodym (R.N) densities and let

$$\rho(P_1, P_2) = \int_R \sqrt{f_1 f_2} \,. d\mu \quad \text{then} \quad 0 \le \rho(P_1, P_2) \le 1.$$

$\rho(P_1, P_2) = 1$ if $P_1 = P_2$ and $\rho(P_1, P_2) = 0$ if P_1 and P_2 are orthogonal measures (Schmetterer (1974)).

<u>3.5.8</u> : <u>Theorem</u> : Let P_{θ_0} and P_{θ_1} be two probability measures over $(\mathbb{R}_1, \mathcal{B}_1)$. Consider the sequence of sample spaces $\left\{ (\mathbb{R}_n, \mathcal{B}_n) \right\}$ and on them the product measures

$P_{\theta_i}^{(n)} = \prod_{i=1}^{n} P_{\theta_{ij}}$ with $P_{\theta_{ij}}$ $1 \le j \le n$ $i = 0, 1$

Then there always exists a consistent sequence of tests for the problem $\left[\{\theta_0\}, \{\theta_1\} \right]$

<u>Proof</u> : Let μ be a probability measure dominating P_{θ_0} and P_{θ_1}. Denote the R.N. densities of P_{θ_i} w.r.t. μ by f_{θ_i} $i = 0, 1$. Then the R.N. density of $P_{\theta_i}^{(n)}$ w.r.t. $\prod_{j=1}^{n} \mu_j$ with $\mu_j = \mu$, $1 \le j \le n$ is given by

$$(X_1, X_2, \dots, X_n) \longrightarrow \prod_{j=1}^{n} f_{\theta_i} (x_j).$$

Hence $$\rho(P_{\theta_0}^{(n)} \; P_{\theta_1}^{(n)}) = \int_{R_n} \Big[\prod_{j=1}^{n} f_{\theta_0}(x_j) \; f_{\theta_1}(x_j) \Big]^{1/2} d\mu(x_1) \dots d\mu(x_n)$$

$$= \prod_{j=1}^{n} \int_{R_n} \sqrt{f_{\theta_0}(x_j) \; f_{\theta_1}(x_j)} \,. d\mu(x_j).$$

and so $\rho(P_{\theta_o}^{(n)}, P_{\theta_1}^{(n)}) = (\rho(P_{\theta_o}, P_{\theta_1}))^n$.

since $P_{\theta_o} \neq P_{\theta_1}$ we have from above lemma

$$\rho(P_{\theta_o}^{(n)}\ P_{\theta_1}^{(n)}) \longrightarrow 0 \qquad \text{as } n \longrightarrow \infty.$$

Now we choose a sequence of real numbers $\{K_n\}$ and real numbers C_1, C_2 with $0<C_1 \le C_2$ such that for $n = 1.2, \ldots$

$$0<C_1 \le K_n \le C_2$$

for $n \ge 1$ let

$$\varphi_n(X_1, X_2, \ldots, X_n) = \begin{cases} 1 & \text{if } \dfrac{\prod_{j=1}^{n} (f_{\theta_1}(x_j))^{1/2}}{\prod_{j=1}^{n} (f_{\theta_o}(x_j))^{1/2}} > K_n \\ 0 & \text{otherwise.} \end{cases}$$

Then $E_{\theta_o}(\varphi_n) = \int_{R_n} \varphi_n(X_1, X_2, \ldots, X_n) \prod_{j=1}^{n} f_{\theta_o}(x_j) d\mu(x_1) \ldots d\mu(x_n)$

$$\le \frac{1}{K_n} \int_{R_n} \left[\prod_{j=1}^{n} f_{\theta_o}(x_j) \prod_{j=1}^{n} f_{\theta_1}(x_j)\right]^{1/2} d\mu(x_1) \ldots d_\mu(x_n)$$

$$= \frac{1}{K_n} \rho\ (P_{\theta_o}^{(n)}\ P_{\theta_1}^{(n)})$$

$$\le \frac{1}{C_1} \rho\ (P_{\theta_o}^{(n)}\ P_{\theta_1}^{(n)}) \longrightarrow 0 \text{ as } n \longrightarrow \infty$$

similarly

$$E_{\theta_1}(1 - \varphi_n) \leq K_n \int \left[\prod_{j=1}^{n} f_{\theta_0}(x_j) \prod_{j=1}^{n} f_{\theta_1}(x_j) \right]^{1/2} d\mu(x_1)\ldots d\mu(x_n)$$

$$\leq C_2 \, \rho \, (P_{\theta_0}^{(n)} \; P_{\theta_1}^{(n)}) \longrightarrow 0 \quad \text{as} \quad n \longrightarrow \infty.$$

This implies that type one and two error approaches to zero as n increases infinitely large.

From the above theorem (3.5.8) we can say that N.P. test is consistent. Following is an example that establishes the consistency of N.P. test.

Example : 3.5.8 : Let $X_1, X_2, \ldots, X_n$ be i.i.d. with density function $f(x, \theta) = \frac{1}{\theta} \cdot e^{-x/\theta} \quad x \geq 0, \quad \theta > 0$

we have to test $H_0 : \theta_0 = 1$ V/s $H_1 : \theta_1 > 1$.

$$\varphi(X) = \begin{cases} 1 & \text{if } \Sigma x_i > C_n \\ 0 & \text{otherwise.} \end{cases}$$

where C_n is such that $P\left[\Sigma x_i > C_n \right] = \alpha = 0.05$ (say)

Under $H_0 : T = \Sigma x_i \rightsquigarrow G(n)$.

Now $E_{\theta_0}\varphi(X) = P_{\theta_0}\left[T > C_n \right]$

$$= \int_{C_n}^{\infty} f_{\theta_0}(t) \, dt.$$

$$= \int_{C_n}^{\infty} \frac{e^{-x} \, x^{n-1}}{\sqrt{n}} \, dx.$$

Here C_n is such that $\int_{C_n}^{\infty} f_1 (t)\, dt = \alpha = 0.05$ (3.5.8)

and power $\beta(\theta) = \int_{C_n/\theta}^{\infty} f_{\theta_1}(t)\, dt$ (3.5.9)

Let $\theta_1 = 1.5$ and $n = 2$. Then from the equation (3.5.9) $C_n = 3.4$ and

$$\beta(\theta) = 1 - \int_0^{3.4/1.5} f_{\theta_1}(t)\, dt$$

$$= 1 - \int_0^{2.2} f_{\theta_1}(t)\, dt = 1 - 0.8169 = 0.1831.$$

(Karl-pearson F.R.I. (1965))

Similarly we can obtain power as n increases as given in following table. $\theta_0 = 1$, $\theta_i = 1.5$, $\alpha = 0.05$

n	C_n	$\beta(\theta)$
2	3.4	0.1831
3	3.7	0.1936
4	3.9	0.2381
5	4.1	0.2801
10	5.0	0.4028
20	6.3	0.4592
30	7.2	0.7408
40	8.1	0.8213
50	8.8	0.8839
50*	61.60	0.9090
100*	111.64	0.9962

For * values normal approximation is used as

$$\varphi(\underset{\sim}{X}_n) = \begin{cases} 1 & \text{if } \Sigma x_i > C_n \\ 0 & \text{otherwise} \end{cases}$$

$$0.05 = P_{\theta_o}\left[\Sigma x_i > C_n\right] = P_{\theta_o}\left[Z > \frac{C_n - n_\theta}{\sqrt{n\,\theta^2}}\right]$$

$$= P_1\left[Z > \frac{C_n - n}{\sqrt{n}}\right]$$

This implies that $\frac{C_n - n}{\sqrt{n}} = 1.64$

so $C_n = \sqrt{n}\, 1.64 + n$

let $n = 50$ then $C_n = 61.60$

then $\beta(\theta) = P_{\theta_1} \left[Z > \frac{C_n - n\theta}{\sqrt{n\theta^2}} \right]$ but $\theta_1 = 1.5$

$$= P_{1.5} \left[Z > \frac{61.60 - 75}{10.13} \right]$$

$$= P_{1.5} \left[Z > - 1.340 \right]$$

$$= 0.90901.$$

In this way power of N.P. test approches to one as n increases.

Similarly we shall see size of the test approaches to zero as n increases given below :

$$\varphi(\underset{\sim}{X}_n) = \begin{cases} 1 & \text{if } \frac{f_{\theta_1}(\underset{\sim}{x}_n)}{f_{\theta_0}(\underset{\sim}{x}_n)} > K_n = 0.5 \\ 0 & \text{otherwise.} \end{cases}$$

So $$\frac{f_{\theta_1}(\underset{\sim}{x}_n)}{f_{\theta_0}(\underset{\sim}{x}_n)} = \left[\frac{\theta_0}{\theta_1} \right]^n \exp \left[\Sigma x_i \left(\frac{1}{\theta_0} - \frac{1}{\theta_1}\right) \right]$$

consider $\theta_0 = 1$ $\theta_1 = 1.5$

This implies that

$$\Sigma x_i > 3/2 \ n\log(3/2)$$

so test will be

$$\varphi(\underset{\sim}{X}_n) = \begin{cases} 1 & \text{if } \Sigma x_i > \frac{3}{2} n\log(\frac{3}{2}) \\ 0 & \text{otherwise.} \end{cases}$$

$$E_{\theta_o}\varphi(X) = \int_{3/2\, n\log 3/2}^{\infty} f_{\theta_o}(x)dx$$

$$= 1 - \int_0^{3/2\, n\log\, 3/2} \frac{e^{-x} \quad x^{n-1}}{\sqrt{n}}\, dx$$

We calculate α for different values of n as

$$n = 2, \quad \alpha = 1 - \int_0^{3/2.\,2\log(3/2)} \frac{e^{-x} \quad x^{n-1}}{\sqrt{n}}\, dx$$

$$= 1 - \int_0^{1.5 \times 2(0.4054)} \frac{e^{-x} \quad x^{n-1}}{\sqrt{n}}\, dx$$

$$= 1 - \int_0^{1.2} \frac{e^{-x} \quad x^{n-1}}{\sqrt{n}}\, dx = 1 - 0.5058397$$

$$= 0.6941603.$$

Similarly we can obtain α as n increases as given in following tables :

For $K_n = 0.5$ $\theta_o = 1$ and $\theta_2 = 1.5$

n	∝
2	0.6941603
3	0.3973457
4	0.2942299
5	0.2013125
10	0.0075165
20	0

For $K_n = 1$ $\theta_o = 1$ and $\theta_2 = 1.5$

n	∝
2	0.1475118
3	0.046033
4	0.0119601
5	0.0023514
10	0

...Oo*oO...