STATISTICAL INFERENCE. 9/1/23

Estimation

Population : group of individuals where we have interest

$X \sim f(x, \theta)$ → parameter which defines the population.

↓

Random Variable

$B(n, p)$

$P(\lambda)$

$N(\mu, \sigma^2)$

$Exp(\theta)$

Statistic : Any function of sample values, it must be free from any population parameter.

Estimator : It is a statistic, used to estimate unknown value of population.

* every estimator is statistic but not every statistic is estimator

(Sampling distribution : Chi-square, t, f)

for estimating parameter, it can be possible to use more than one statistic.

Criterions of a good Estimator.

i) Unbiasedness

ii) Consistency

iii) Efficiency

iv) Sufficiency.

$\boxed{\text{I}}$ UNBIASEDNESS

$x_1, x_2, x_3 \cdots x_n \overset{r.s}{\sim} f(x, \theta)$

$t = t(x_1, x_2, \cdots x_n) = t(X)$

Defination : An estimator 't' of an unknown parameter θ is said to be an unbiased estimator of θ if expectation of t is $E(t)$.

$E(t) = \theta.$

If $E(t) \neq \theta$, t is said to biased estimator.

Bias of t

$B(t) = (E(t) - \theta) = 0$ [unbiased]

> 0 [+vely biased]

< 0 [-vely biased].

Ex1: $x_1, x_2, \dots x_n \overset{r.s.}{\sim} P(\lambda)$

then show that sample mean $\bar{x}$ is an unbiased estimator of λ.

$$\bar{x} = \frac{1}{n}\sum x_i$$

$$E(\bar{x}) = \frac{1}{n}\sum E(x_i) = \frac{1}{n}\sum_{i=1}^{n} \lambda$$

$$= \frac{1}{n} n\lambda = \underline{\lambda}$$

2: ① $f(x, \theta) = \frac{1}{\theta} e^{-x/\theta}, x > 0$ $[E(x_i) = \theta]$

show $\bar{x}$ is unbiased estimator of θ.

$$E(\bar{x}) = \frac{1}{n}\sum_{i}^{n} E(x)$$

$$E(x_i) = \frac{1}{n}\sum\theta = \frac{1}{n} n\theta = \theta.$$

(ii) In case of $f(x, \theta) = \theta e^{-x\theta}$; then $\left[E(x_i) = \frac{1}{\theta}\right]$

$x_1, x_2, \dots x_n \overset{r.v}{\sim} N(\mu, \sigma^2)$.

show that $\bar{x}$ is unbiased for μ and $s^2 = \frac{1}{n-1}\sum(x_i - \bar{x})^2$ is an unbiased for σ^2.

$$E(\bar{x}) = \frac{1}{n}\sum_{i=1}^{n} E(x_i)$$

$$= \frac{1}{n}\sum_{i}^{n}\mu = \frac{1}{n} \cdot n\mu = \mu.$$

$$s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2 \Rightarrow \frac{1}{n-1}\left[\sum_{i=1}^{n} x_i^2 - n\bar{x}^2\right]$$

$$E(s^2) = \frac{1}{n-1}\left[\sum_{i=1}^{n} E(x_i^2) - nE(\bar{x}^2)\right]$$

$$V(x_i) = E(x_i)^2 - (E(x_i))^2$$

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

11/1/23

$\bar{x}$ is unbiased for μ.

$$s^2 = \frac{1}{n-1}\left[\sum_{i=1}^{n} x_i^2 - n\bar{x}^2\right]$$

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

$$= \frac{1}{\cancel{n-1}}\,(\cancel{n-1})\,\sigma^2 = \underline{\underline{\sigma^2}}$$

→ Suppose $x_1, x_2, x_3 \ldots x_n \sim N(\mu, \sigma^2)$.

$$\sum a x_i \sim N(\sum a_i\mu, \sum a^2\sigma^2)$$

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

$$V\left(\frac{(n-1)s^2}{\sigma^2}\right) = (n-1)$$

→ degree of freedom.
→ distributions.

$$E(s^2) = \sigma^2$$

$x_1, x_2, \ldots x_n \overset{r.s}{\sim} f(x,\theta) = \frac{1}{\theta}$, $0 < x < \theta$ [where $x_1, x_2, \ldots x_n$ are R.V of a random sample].

To Find :- unbiased estimator of θ.

$$E(x) = \int_0^\theta x f(x)\,dx = \int_0^\theta x \frac{1}{\theta}\,dx = \frac{1}{\theta}\left[\frac{x^2}{2}\right]_0^\theta$$

$$= \underline{\underline{\frac{\theta}{2}}}$$

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$$

$$E(\bar{x}) = \frac{1}{n}\sum_{i=1}^{n} E(x_i)$$

$$= \frac{1}{n}\, n \cdot \frac{\theta}{2} = \frac{\theta}{2}$$

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

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

Hence $2\bar{x}$ is an unbiased estimator of θ.

let $x_1, x_2 \ldots x_n \overset{r.s}{\sim} p(x) = p^{x}(1-p)^{1-x}, x = 0, 1$

$0 < p < 1$.

To Show $= \frac{t(t-1)}{n(n-1)}$ is unbiased to p^2.

where $t = \sum_{i=1}^{n} x_i \sim B(n, p)$

$E(t) = np$

$V(t) = npq = np(1-p)$.

$$E\left(\frac{t(t-1)}{n(n-1)}\right) = \frac{E(t^2) - E(t)}{n(n-1)}$$

$$= \frac{V(t) + [E(t)]^2 - E(t)}{n(n-1)}$$

$$= \frac{np(1-p) + n^2p^2 - np}{n(n-1)}$$

$$= \frac{\cancel{np} - np^2 + n^2p^2 - \cancel{np}}{n(n-1)}$$

$$= \frac{n(n-1)p^2}{n(n-1)}$$

$$= \underline{\underline{p^2}}$$

$$E\left[\frac{t(t-1)}{n(n-1)}\right] = p^2.$$

$\therefore \frac{t(t-1)}{n(n-1)}$ is unbiased to p^2.

Suppose $x_1, x_2, x_3 \ldots x_n \overset{r.s}{\sim} N(\mu, 1) \sim N(\mu, \sigma^2)$.

Show that $t = \frac{1}{N} \sum x_i^2$ is an unbiased estimator of $\mu^2 + 1$

ie to show $E(t) = \mu^2 + 1$.

$E(t) = E\left(\frac{1}{N} \sum x_i^2\right) = \frac{1}{N} E(\sum x_i^2) = \frac{1}{N} V(x) + [E(x)]^2$

$$= \frac{1}{n}\sum_{i=1}^{n}\left[1+\mu^2\right]$$

$$= \frac{1}{n} n(1+\mu^2) = 1+\mu^2)$$

$x_1, x_2, \ldots x_n$ x_i's $B(n, p)$

Show that $\frac{\bar{x}}{n}$ is an unbiased estimator of p.

$$E\left(\frac{\bar{x}}{n}\right) = \frac{E(\bar{x})}{n} = \frac{E\left[\frac{1}{n}\sum x_i\right]}{n}$$

$$= \frac{1}{n^2} E(\sum x_i)$$

$$= \frac{1}{n^2} \sum E[x_i]$$

$$= \frac{1}{n^2} \sum np$$

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

Hence $E\left(\frac{\bar{x}}{n}\right) = p$.

$\therefore$ $\frac{\bar{x}}{n}$ is an unbiased estimator of p.

II CONSISTENCY.

An estimator t of a parameter θ is said to be a consistent estimator of θ in $\psi(\theta)$ if $E(t) = \theta$, $V(t) \to 0$ as $n \to \infty$,

→ x_1, x_2 is a random sample $\sim N(\mu, \sigma^2)$.

Show that $\bar{x}$ and s^2 are consistent estimator of μ and σ^2 respectively.

1. $E(\bar{x}) = \mu$, $\bar{x} \sim N(\mu, \frac{\sigma^2}{n})$.

$V(\bar{x}) = \frac{\sigma^2}{n}$. $E(\bar{x}) = \mu$.

as $n \to \infty$, $V(\bar{x}) \to 0$

Hence we can say that it is consistent for μ.

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

$\sum_{i=1}^{n} \frac{(x_i-\bar{x})^2}{\sigma^2} \sim \chi^2_{n-1}$

$E(s^2) =$

$V\left(\frac{(n-1)s^2}{\sigma^2}\right) = 2(n-1).$

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

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

as $n \to \infty$, $V(s^2) \to 0$.

(iii) Show that $S^2 = \frac{1}{n} \sum_{i=1}^{n} (x_i-\bar{x})^2$.

is consistent but not unbiased to σ^2.

Sol:- $\sum_{i=1}^{n} \frac{(x_i-\bar{x})^2}{\sigma^2} \sim \chi^2_{n-1}$

$\frac{ns^2}{\sigma^2} \sim \chi^2_{n-1}$

$E\left(\frac{ns^2}{\sigma^2}\right) = (n-1)$

$E(s^2) = \frac{(n-1)}{n} \sigma^2 \neq \sigma^2.$

Hence it is a biased estimator, but it is consistent

$V\left(\frac{ns^2}{\sigma^2}\right) = \frac{n^2}{\sigma^4} V(s^2) = \frac{2(n-1)}{1}$

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

$E(s^2) = \left(\frac{n-1}{n}\right) \sigma^2$

$= \left(1-\frac{1}{n}\right)\sigma^2$

$V(s^2) = \frac{2\sigma^4}{n}\left(1-\frac{1}{n}\right)$

$V(s^2) \to 0$ as $n \to \infty$.

s^2 is ~~no~~ biased and consistent.

$x_1, x_2 \ldots x_n \sim N(\mu, \sigma^2)$

Show that $\bar{x}$ and $\tilde{x}$ (sample median)

both are unbiased and consistent estimator of ~~$\bar{x}$~~ μ.

$E(\tilde{x}) = \mu$. [$\because$ $\tilde{x}$ is a middle most value of given r.s, ~~the~~ $E(\tilde{x}) = \mu$.]

for normal distribution mean = median = mode].

$$V(\tilde{x}) = \frac{1}{4nf_i^2} = \frac{1}{4n} \sigma^2 2\pi = \frac{\pi}{2n} \sigma^2$$

$V(\tilde{x}) \to 0$ as $n \to \infty$.

13/1/23

$x_1, x_2, \ldots x_n \sim f(x, \theta) = \frac{1}{\theta} e^{-x/\theta}, x > 0$. Show that $\bar{x}$ is constant for θ.

$E(x) = \theta$

$V(x) = \frac{\theta^2}{n}$

$$\bar{x} = \frac{1}{n} \sum_{i=1}^{n} x_i$$

$$E(\bar{x}) = \frac{1}{n} \sum_{i=1}^{n} E(x_i) = \frac{1}{n} \sum_{i=1}^{n} \theta = \frac{n\theta}{n} = \theta$$

$$V(\bar{x}) = V\left(\frac{1}{n} \sum_{i=1}^{n} E(x_i)\right) = \frac{n\theta^2}{n^2} = \frac{\theta^2}{n} \to 0$$

as $n \to \infty$

$x_1, x_2, \ldots x_n \overset{r.s}{\sim} p(x) = p^x (1-p)^{1-x}$

$\bar{x}$ is constant to p.

$$E(\bar{x}) = \frac{1}{n} \sum_{i=1}^{n} E(x_i)$$

$$= \frac{np}{n} = p$$

$$V(x) = \frac{1}{n^2} \sum V(x_i) = \frac{np(1-p)}{n^2} = \frac{p(1-p)}{n^2} \to 0 \text{ as } n \to \infty.$$

III

Efficiency

Defn: In a class of ~~constants~~ consistent estimators of a parameter $\theta \sim \psi(\theta)$, the estimates having least variance among all the variances of the consistent estimates, it is said to be most efficient estimate of θ.

Let t_1 be the most efficient estimate of θ with variance say V_1 and t_2 be any other estimate of θ with variance V_2, the efficiency of the estimate t_2 wrt the most efficient estimate t_1 is defined as $E = \frac{V_1}{V_2}$.

$x_1, x_2 \ldots x_n \overset{iid}{\sim} N(\mu, \sigma^2)$

$\bar{x}$ (sample mean), $\tilde{x}$ (sample median).

is consistent for μ.

$x_1, x_2, x_3 \overset{iid}{\sim} N(\mu, \sigma^2)$

$t_1 = \frac{1}{3}\sum_{i=1}^{n} x_i \qquad V(t_1) = \frac{\sigma^2}{3}$

$t_2 = \frac{x_1 + 2x_2 + 3x_3}{6} \qquad V(t_2) = \frac{V(x_1) + 4V(x_2) + 9V(x_3)}{36}$

$= \frac{1}{36}[\sigma^2 + 4\sigma^2 + 9\sigma^2]$

$= \frac{14}{36}\sigma^2$

$E = \frac{V(t_1)}{V(t_2)} = 0.86$.

Minimum Variance Unbiased Estimator (MVUE)

Defn: Let t be an unbiased estimator of θ.

such that $E(t) = \theta$

$V(t) < V(t')$

where t' is any other unbiased estimator of θ. Then t is said to be a MVUE of θ.

Remark→ MVUE is an unique estimator.

Theorem: A minimum variance unbiased estimator (MVUE) is unique in the sense that if T_1 and T_2 are MVUE's of θ then $T_1 = T_2$.

Proof: We are given $E(T_1) = E(T_2) = \theta$

$$V(T_1) = V(T_2)$$

Consider a new estimator $T = \frac{1}{2}(T_1 + T_2)$, which is also unbiased $E(T) = \frac{1}{2}[E(T_1) + E(T_2)] = \frac{1}{2}[\theta + \theta] = \theta$.

$$V(T) = \frac{1}{4}[V(T_1) + V(T_2) + 2\,cov(T_1 T_2)]$$

$$= \frac{1}{4}\left[V(T_1) + V(T_2) + 2\rho\sqrt{V(T_1)V(T_2)}\right]$$

$$= \frac{1}{4}[V(T_1) + V(T_1) + 2\rho V(T_1)]$$

$$= \frac{1}{4}[2V(T_1) + 2\rho V(T_1)]$$

$$= \frac{1}{2}V(T_1)[1+\rho]$$

$\because$ T_1 is a MVUE, $V(T) \geq V(T_1)$

$$\frac{1}{2}V(T)(1+\rho) \geq V(T_1)$$

$$1+\rho \geq 2$$

$$\rho \geq 1 \quad — (x)$$

but we know $-1 \leq \rho \leq 1$. $|\rho| \leq 1$ — (x^*)

from x and x^* we get that

$$\rho = 1$$

⇒ T_1 and T_2 have the perfect linear relation.

let $T_1 = \alpha + \beta T_2$

$E(T_1) = \alpha + \beta E(T_2)$ (α and β are constants).

$\theta = \alpha + \beta\theta \Rightarrow \alpha = 0$, assuming $\beta = 1$

$V(T_1) = V(\alpha + \beta T_2) = \beta^2 V(T_2)$

$\beta^2 = 1$

$\Rightarrow \beta = \pm 1$

16/1/23

Theorem 2: If T_1 and T_2 be unbiased estimator of θ or $\psi(\theta)$ with efficiencies e_1 and e_2 resp. and ρ be the correlation coefficient between them, then.

$$\sqrt{e_1 e_2} - \sqrt{(1-e_1)\times(1-e_2)} \leq \rho \leq \sqrt{e_1 e_2} + \sqrt{(1-e_1)(1-e_2)}$$

Proof :- Let T be the MVE of θ, we are given $E(T_1) = E(T_2) = \theta$

e.

Let us consider another unbiased estimator T_3.

Such that $T_3 = \lambda T_1 + \mu T_2$

λ and μ are constants.

$E(T_3) = \lambda E(T_1) + \mu E(T_2)$

$= (\lambda + \mu)\theta \Rightarrow \lambda + \mu = 1$

$= \lambda^2 \frac{V}{e_1} + \mu^2 \frac{V}{e_2} + 2\lambda\mu\rho \sqrt{\frac{V}{e_1}}\sqrt{\frac{V}{e_2}}$

$= V^2 \left[\frac{\lambda^2}{e_1} + \frac{\mu^2}{e_2} + \frac{2\lambda\mu\rho}{\sqrt{e_1 e_2}}\right]$

But, $V(T_3) \geq V(T)$

$\Rightarrow V(T_3) \geq V$.

$$v^2\left[\frac{\lambda^2}{e_1}+\frac{\mu^2}{e_2}+\frac{2\lambda\mu\rho}{\sqrt{e_1e_2}}\right]\geq v^2$$

$$\frac{\lambda^2}{e_1}+\frac{\mu^2}{e_2}+\frac{2\lambda\mu\rho}{\sqrt{e_1e_2}}\geq 1.$$

$$\left(\frac{1}{e_1}-1\right)^2\lambda^2+\left(\frac{1}{e_2}-1\right)\mu^2+2\lambda\mu\left(\frac{\rho}{\sqrt{e_1e_2}}-1\right)\geq 0$$

dividing by μ^2

$$\left(\frac{1}{e_1}-1\right)^2\frac{\lambda^2}{\mu^2}+\left(\frac{1}{e_2}-1\right)+\frac{2\lambda}{\mu}\left(\frac{\rho}{\sqrt{e_1e_2}}-1\right)\geq 0$$

This is a quadratic eq^n in $\frac{\lambda}{\mu}$.

$e_1<1$

$\frac{1}{e}>1$

$$\boxed{\begin{gathered}Ax^2+Bx+C\geq 0\ \forall x\\ \text{if } A>0;\ C>0\\ B^2-AC>0\end{gathered}}$$

$$\left(\frac{\rho}{\sqrt{e_1e_2}}-1\right)^2-\left(\frac{1}{e_1}-1\right)\left(\frac{1}{e_2}-1\right)\leq 0$$

$$\Rightarrow\left(\rho-\sqrt{e_1e_2}\right)^2-(1-e_1)(1-e_2)\leq 0$$

$$\Rightarrow \rho^2-2\sqrt{e_1e_2}\,\rho+(e_1+e_2-1)=0$$

Theorem of the quadratic eq.

$$\sqrt{e_1e_2}\pm\sqrt{(e_1-1)(e_2-1)}$$

$$\Rightarrow\sqrt{e_1e_2}-\sqrt{(e_1-1)(e_2-1)}\leq\rho\leq\sqrt{e_1e_2}+\sqrt{(e_1-1)(e_2-1)}$$

$$\Rightarrow\boxed{\sqrt{e_1e_2}-\sqrt{(1-e_1)(1-e_2)}\leq\rho\leq\sqrt{e_1e_2}+\sqrt{(1-e_1)(1-e_2)}}$$

if $e_1=e_1$ and $e_2=e$

we have $\sqrt{e}\leq\rho\leq\sqrt{e}\ \Rightarrow\ \rho=\sqrt{e}$

~~Corollary~~ Corallary: If T_1 is a MVUE of θ and T_2 is any unbiased estimator of θ with efficiency e, then the correlation coefficient between T_1 and T_2 is

$$\boxed{\rho = \sqrt{e}.}$$

RESULT :- : If T_1 is MVUE of θ and T_2 be any other unbiased estimator of θ with efficiency $e<1$, then no unbiased linear combination of T_1 and T_2 can be a MVUE of θ.

Sol^n: Let T be a linear combination of T_1 and T_2.

$$T = l_1 T_1 + l_2 T_2$$

T is unbiased if $l_1 + l_2 = 1$

$$E(T) = l_1 E(T_1) + l_2 E(T_2)$$
$$= (l_1 + l_2)\theta = \theta$$
$$\text{if } l_1 + l_2 = 1.$$

given $E(T_1) = E(T_2) = \theta$

$$e = \frac{V(T_1)}{V(T_2)} = 1, \quad V(T_2) = \frac{V(T_1)}{e}$$

$$\rho_{T_1T_2} = \rho = \sqrt{e}.$$

$$V(T) = l_1^2 V(T_1) + l_2^2 V(T_2) + 2 l_1 l_2 \operatorname{cov}(T_1 T_2)$$
$$= l_1^2 V(T_1) + l_2^2 V(T_2) + 2 l_1 l_2 \rho \sqrt{V(T_1) V(T_2)}$$
$$= l_1^2 V(T_1) + l_2^2 \frac{V(T_1)}{e} + 2 l_1 l_2 \rho \sqrt{V(T_1) \frac{V(T_2)}{e}}$$
$$= V(T_1)\left[l_1^2 + \frac{l_2^2}{e} + 2 l_1 l_2 \frac{\rho}{e}\right]$$

$$V(T) = V(T_1)\left[l_1^2 + \frac{l_2^2}{e} + 2 l_1 l_2\right] \rho_{T_1T_2}$$

$$\geq V(T_1)\left[l_1^2+l_2^2+2l_1l_2\right]$$

$$\geq V(T_1)(l_1+l_2)^2$$

$$V(T)\geq V(T_1).$$

---

(4) If T_1 and T_2 be two unbiased estimator of θ with variances σ_1^2 and σ_2^2 and correlation ρ between them, what is the best linear combination of T_1 and T_2 and what is the variance of such combinations.

Sol^n: Given $E(T_1)=E(T_2)=\theta$

$$V(T_1)=\sigma_1^2,\ V(T_2)=\sigma_2^2$$

$$\rho_{T_1T_2}=\rho$$

Let T be the unbiased linear combination of T_1 and T_2.

$$T=l_1T_1+l_2T_2$$

$$E(T)=\theta \text{ iff } l_1+l_2=1$$

$$V(T)=l_1^2V(T_1)+l_2^2V(T_2)+2l_1l_2\rho\sqrt{V(T_1)}\sqrt{V(T_2)}.$$

$$=l_1^2\sigma_1^2+l_2^2\sigma_2^2+2l_1l_2\rho\sigma_1\sigma_2.$$

$$\frac{\partial V(T)}{\partial l_1}=0\Rightarrow l_1\sigma_1^2+l_2\rho\sigma_1\sigma_2=0 \quad -(1)$$

$$\frac{\partial V(T)}{\partial l_2}=0\Rightarrow l_2\sigma_2^2+l_1\rho\sigma_1\sigma_2=0 \quad -(2)$$

on solving (1) and (2) we have.

$$l_1=\frac{\sigma_2^2-\rho(\sigma_1\sigma_2)}{\sigma_1^2+\sigma_2^2-2\rho\sigma_1\sigma_2}=l_1^*$$

$$l_2=\frac{\sigma_1^2-\rho\sigma_1\sigma_2}{\sigma_1^2+\sigma_2^2-2\rho\sigma_1\sigma_2}=l_2^*$$

5) If T_1 is MVUE of θ and T_2 be any other unbiased estimator of θ with variance $\frac{\sigma^2}{e}$, then prove that the correlation coefficient b.w them is $\rho = \rho_{T_1 T_2} = \sqrt{e}$

Proof = The coefficient of — linear combination of T_1 and T_2 is $T = l_1 T_1 + l_2 T_2$.

$$l_1 = \frac{\sigma_2^2 - \rho\sigma_1\sigma_2}{\sigma_1^2 + \sigma_2^2 - 2\rho\sigma_1\sigma_2} \qquad l_2 = \frac{\sigma_1^2 - \rho\sigma_1\sigma_2}{\sigma_1^2 + \sigma_2^2 - 2\rho\sigma_1\sigma_2}$$

given $T_1 \sim$ MVUE.

$$V(T_1) = \sigma^2$$

$$e = \frac{V(T_1)}{V(T_2)}$$

$$V(T_2) = \frac{V(T_1)}{e} = \frac{\sigma^2}{e}$$

multiplying σ_1^2 by σ^2 and σ_2^2 by $\frac{\sigma^2}{e}$.

Then we get.

$$l_1 = \frac{1 - \rho\sqrt{e}}{D}, \quad l_2 = \frac{e - \rho\sqrt{e}}{D}$$

$$D = 1 + e - 2\rho\sqrt{e}$$

Hence the unbiased estimate T ~~of θ~~ if θ takes the form

$$T = \frac{(1 - \rho\sqrt{e})T_1 + (e - \rho\sqrt{e})T_2}{D}$$

$$V(T) = \frac{1}{\cancel{\rho} D^2}\left((1-\rho\sqrt{e})^2\sigma^2 + (e - \rho\sqrt{e})^2\frac{\sigma^2}{e} + 2(1-\rho\sqrt{e})(e - \rho\sqrt{e})\frac{\sigma \cdot \sigma}{\sqrt{e}}\right.$$

$$V(T) = \frac{\sigma^2}{D^2}\left\{(1 + e - 2\rho\sqrt{e}) - \rho^2(e + 1 - 2\rho\sqrt{e})\right)$$

$$= \frac{\sigma^2(1-\rho^2)(1 + e - 2\rho\sqrt{e})}{(1 + e - 2\rho\sqrt{e})^2} = \frac{\sigma^2(1-\rho^2)}{(1-\rho^2) + (\sqrt{e} - \rho)^2}$$

6.w

$$\frac{V(T)}{\sigma^2} = \frac{(1-\rho^2)}{(1-\rho^2)+(\sqrt{e}-\rho)^2} \leq 1 \quad -(Y)$$

dT_2

But σ^2 is the variance of MVUE T_1.

$$\Rightarrow \frac{V(T)}{\sigma^2} \geq 1 \quad - (X)$$

from (Y) and (X) we have.

$$\frac{V(T)}{\sigma^2} = 1 \quad \text{and} \quad \frac{(1-\rho^2)}{(1-\rho^2)+(\sqrt{e}-\rho)^2} = 1$$

$$\cancel{(1-\rho^2)} = \cancel{(1-\rho^2)} + (\sqrt{e}-\rho)^2$$

$$\boxed{\sqrt{e} = \rho}$$

$$\therefore \rho = \sqrt{e}$$

Aliter.

$$T = \alpha_1 T_1 + \alpha_2 T_2$$

$$\alpha_1 = 1, \alpha_2 = 0$$

Substituting the above, we will get.

$$\Rightarrow \rho = \sqrt{e}$$

Result 6: If T_1 and T_2 are two unbiased estimators of θ (or $\psi(\theta)$), having the same variance and ρ is the correlation coefficient between them, then show that

$$\rho \geq 2e - 1$$

where e is the efficiency of each estimator.

Proof:- Let T be the MVUE of θ. we are given $E(T_1) = E(T_2)$
$V(T_1) = V(T_2)$.

$$e = \frac{V(T)}{V(T_1)} = \frac{V(T)}{V(T_2)}.$$

$$\Rightarrow V(T) = V(T_2) = V(T_1)/e.$$

Consider another unbiased estimator T_3 of θ, which is given as $T_3 = \frac{1}{2}(T_1 + T_2)$.

$$V(T_3) = \frac{1}{4}\left[V(T_1) + V(T_2) + 2\rho\sqrt{V(T_1)}\sqrt{V(T_2)}\right]$$

$$= \frac{1}{4}\left[\frac{V(T)}{e} + \frac{V(T)}{e} + 2\rho\sqrt{\frac{V(T)}{e}\frac{V(T)}{e}}\right]$$

$$V(T_3) = \frac{V(T)}{4e}\left[1 + 1 + 2\rho\right]$$

$$= \frac{(1+\rho)V(T)}{2e}$$

$\because$ T is the MVUE, then $V(T_3) \geq V(T)$.

$$(1+\rho) \geq 2e$$

$$\boxed{\rho \geq (2e-1)}$$

The above 6 results are important.

IV SUFFIENCY

$$x_1, x_2, \ldots x_n \overset{r.s}{\sim} f(x, \theta)$$

$$t(x) = t(x_1, x_2, \ldots x_n)$$

An estimator is said to be sufficient for a parameter if it contains all the info. in the sample regarding the parameter.

More precisely, if $t(x) = t(x_1, x_2, \ldots x_n)$ is an estimator of a parameter θ, based on a sample $x_1, x_2, \ldots x_n$ of size n, from the population with probability density ~~dist.~~ $f(x, \theta)$, the conditional probability distribution of $(x_1, x_2, \ldots x_n)$ is independent of θ, then T is sufficient estimator for θ.

Ex: $x_1, x_2, \cdots x_n \overset{r.s}{\sim} p(x) = p^x(1-p)^{1-x}, x=0,1.$

show that $T = \sum_{i=1}^{n} x_i$ is sufficient to p.

$$x_i = \begin{cases} 1 & \text{with prob } p \\ 0 & \text{with prof } (1-p) = q. \end{cases}$$

$\sum_{i=1}^{n} x_i = k \text{ (say)} \leq n.$

$T = \sum_{i=1}^{n} x_i \sim B(n, p)$

$P(T=k) = p(k) = \binom{n}{k} p^k (1-p)^{n-k}, k = 0, 1, \cdots n.$

The conditional prob. dist. of $(x_1, x_2 \ldots x_n)$ given T_n.