STATISTICAL INFERENCE

BY DR RAJIV SAKSENA
DEPARTMENT OF STATISTICS
UNIVERSITY OF LUCKNOW

We know that statistical data is nothing but a random sample of observations drawn from a population described by a random variable whose probability distribution is unknown or partly unknown and we try to know about the properties of the population on the basis of knowledge of the properties of the sample. This inductive process of going from known sample to the unknown population is called 'Statistical Inference '

Formally, let x be a random variable describing the population under investigation. Suppose X has þ.m.f $f_o(x) = P(x = x)$ or þ d f $f_o(x)$ which depend on some unknown parameter θ (single or vector valued) that may have any value in a set Ω (called the parameters space). We assume that the functional form of $f_o(x)$ is known but not the parameter θ(except that $\theta \in \Omega$). For example, the family of distributions $\{f_\theta(x), \theta \in \Omega\}$ may be the family of Poisson distribution $\{P(\lambda), \lambda \geq 0\}$ or normal distribution $\{N(\mu, \sigma^2), -\infty < \mu < \infty, \sigma \geq 0\}$

Two problem of statistical inference are-

1. To estimate the value of θ – problem of estimation
2. To test a hypothesis about θ - problem of testing of the hypothesis

POINT ESTIMATION

Definition: A random sample of size 'n' from the distribution of X is a set of independent and identically distributed random variables $\{x_1, x_2, \dots, x_n\}$ each of which has the same distribution as that of X. The probability of the sample is given by

$$f_o(x_1, x_2, \dots, x_n) = f_o(x_1) f_o(x_2) \dots f_o(x_n)$$

Definition: A statistic T = T ($x_1, x_2, \dots, x_n$) is any function of the sample values, which does not depend on the unknown parameter θ. Evidently, T is a random variable which has its own probability distribution (called the ' Sampling distribution' of T)

For example, $\bar{x} = \frac{1}{n}\sum_i^n x_i$; $s^2 = \frac{1}{n-1}\sum_i^n (x_i - \bar{x})^2$ $X_{(1)} = \min(x_1, x_2, \dots x_n)$, $X_{(n)} = max(x_1, x_2, \dots x_n)$ are some statistics.

If we use the statistic T to estimate the unknown parameter θ, it is called the estimator (or point estimators) of θ and the value of T obtained from a given sample is its 'estimate'

Remark: Obviously, for T to be a good estimator of θ , the difference $[T - \theta]$ should be as small as possible. However, since T is itself a random variable all that we can hope for is that it is close to θ with high probability.

Theorem : Let ($X_1, X_2,\ldots,X_n$) be a random sample of 'n' observations on X with mean $E(X) = \mu$ and variance $Var(x) = \sigma^2$ Let the sample mean and sample variance be $\bar{x} = \frac{1}{n}\sum_i^n x_i$ and $s^2 = \frac{1}{n}\sum(x_i - x)^2$

Then,

(I)E($\overline{X}$)=μ

(ii)$V(\overline{X})$=$\frac{\sigma^2}{n}$

(iii)E(S^2) $= \frac{n-1}{n}\sigma^2$

Prof: We have

$$E(\overline{X}) = E\left(\frac{1}{n}\sum_i^n x_i\right) = \frac{1}{n}\sum_i^n E(x_i) = \mu$$

$$V(\overline{X}) = V\left(\frac{1}{n}\sum_i^n x_i\right) = \frac{1}{n^2}\sum_i^n V(x_i) = \frac{\sigma^2}{n}$$

$$E(ns^2) = E\sum_i^n (x_i - \bar{x})^2$$

$$= E\sum_1^n [[(x_i - \mu) - (\bar{x} - \mu)]^2]$$

$$= E\left[\sum_i^n (x_i - \mu)^2 - n(\bar{x} - \mu)^2\right]$$

$$= E(x_i - \mu)^2 - nE(\bar{x} - \mu)^2$$

$$= n\sigma^2 - n\sigma^2/n$$

$$= (n-1)\sigma^2$$

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

PROPERTIES OF ESTIMATORS

UNBIASEDNESS:

An estimator T of an unknown parameter θ is called unbiased if

$E(T) = \theta$ for all $\theta \in \Omega$

Example. If $(x_1, x_{2,} \ldots, x_n)$ is a random sample from any population with mean μ and variance σ^2, the sample mean $\bar{x}$ is an unbiased estimator of μ but the sample variance S^2 is not an unbiased estimator of σ^2.

However, $\frac{ns^2}{n-1} = \frac{1}{n-1}\sum_i^n (x_i - \bar{x})^2$ is an unbiased estimator of σ^2.

Ex. if $(x_1, x_{2,} \ldots x_n)$ is a random sample from a normal distribution $N(\mu, I)$ show that $T = \frac{1}{n}\sum_i^n {x_i}^2 - 1$ is an unbiased estimator of μ^2 ,

Soln. $E(T) = E\left[\frac{1}{n}\sum_i^n {x_i}^2 - 1\right] = \frac{1}{n}\sum_i^n E({x_i}^2) - 1$

$E(x_i^2)= V(x)+E(x_i)= (\mu^2 + 1)$

$$= \frac{1}{n}\sum_1^n (\mu^2 + 1) - 1 = \mu^2$$

Example: Let $(x_1, x_{2,} \ldots x_n)$ be a random sample of observation from a Bernoulli distribution $f_\theta(x) = \theta^x(1-\theta)^{1-x}(x = 0,1)$ show that $T = \frac{y(y-1)}{n(n-1)}$ is an unbiased estimator of θ where $y = \sum_i^n x_i$

<u>Soln:</u> We know that $E(x_i) = \theta$ and $V(x_i) = \theta(1-\theta)$ so that $E(Y) = n\,\theta$ and $V(Y) = n\theta(1-\theta)$

Now

$$E(Y(Y-1) = E(Y^2) - E(Y)$$

$$= V(Y) + [E(Y)]^2 - E(Y)$$

$$= n\theta(1-\theta) + n^2\theta^2 - n\theta$$

$$= n(n-1)\theta^2$$

$$E(T)=E\left[\frac{Y(Y-1)}{n(n-1)}\right] = \theta^2$$

Showing it to be an unbiased estimator of θ^2

Example: Show that the mean $\bar{x}$ of a random sample of size n from the exponential distribution $f_\theta(x) = \frac{1}{\theta}\bar{e}^{\frac{x}{\theta}}(x > 0)$ is an unbiased estimator of θ and has variance θ^2/n

Soln: We know that

$E(x_i) = \theta$ and $V(x_i) = \theta^2$ $(i = 1, .., n)$

$E(\overline{X})= \theta$ and $V(\overline{X}) = \theta^2/n$

Example: Let $(x_1, x_{2,} \ldots x_n)$ to a random sample from a normal distribution with mean 0 and variance θ ($0 < \theta < \infty$) so that $T = \sum x_i^2/n$ is an unbiased estimator of θ and has variance $2\theta^2$/n

Sohm we know that

$$E(x_i) = 0, E\left(x_i^2\right) = V(x_i) = \theta$$

$$E(T) = \frac{1}{n}\sum_i^n E\left(x_i^2\right) = \theta$$

Also $$E\left(x_i^4\right) = \mu_4 = 3\theta^2$$

$$V(T) = V\left(\frac{1}{n}\sum_i^n x_i^2\right)$$

$$= \frac{1}{n^2}\sum_i^n V(x_i^2)$$

$$= \frac{1}{n^2}\sum_i^n \left[E\left(x_i^4\right) - \{E\left(x_i^2\right)\}^2\right]$$

$$= \frac{1}{n^2}\sum_i^n [3\theta^2 - \theta^2]$$

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

Example Let $(x_1, x_{2,} \dots x_n)$ be a random sample from the rectangular distribution $R(0, \theta)$ having þ, d, f

$$f_\theta(x) = \begin{cases} \frac{1}{\theta}, & o \le x \le \theta \ (\theta > o) \\ 0 & ,\text{otherwise} \end{cases}$$

Show that $T_1 = 2\overline{x}$, $T_2 = \frac{n+1}{n} Y_n$ and $T_3 = (n+1)\gamma_i$ are all unbiased for θ , where $Y_1 = \min(x_1, x_{2,} \dots x_n)$ and $Y_n = \max(x_1, x_{2,} \dots x_n)$

Soln: We know that

$E(x) = \theta/2$ and $V(x) = \theta^2/12$

$$E(T_I) = E\ 2\left(\frac{\sum_i^n x_i}{n}\right) = \theta \ and \ V(T_I) = \frac{\theta^2}{3n}$$

To obtain the expectation of T_2 and T_3 we need to obtain their distribution.

The $d.f.$ of Y_n is-

$$F_y(y) = P(Y_n \le y)$$

$$= P(\max(x_1, x_{2,} \dots x_n) \le y)$$

$$= P(x_i \leqslant y, x_n \leqslant y)$$

$$= [P(x \leqslant y)]^n$$

$$= \left(\frac{y}{\theta}\right) = \frac{y^n}{\theta^n}$$

$p, d, f\ of\ Y_n$ is- $\quad g Y_n(y) = \begin{cases} \frac{n y^{n-i}, 0 \leqslant y \leqslant \theta}{\theta^n} \\ 0 \quad , elsewhere \end{cases}$

Hence, $\quad E(Y_n) = \int_0^\theta \frac{n y^n}{\theta^n} y = \left(\frac{n}{n+1}\right)\theta$

Or $\quad E\left(\frac{n+1}{n} Y_n\right) = \theta$

So that T_2 is unbiased for θ

[We can check that V (T_2) =$\frac{\theta^2}{n(n+2)}$]

Again, the $d. f.$ of Y_i is-

$$F_{Y_i}(y) = P\{Y_i \leq y\}$$

$$= \mathrm{P}\{\min(x_1, x_{2,} \dots x_n) \leq y\}$$

$$= I - P\{x_1 > y, x_2 > y, \dots x_n > y\}$$

$$= I - [I - P(X < y)]^n$$

$$= I - \left[I - \frac{y}{\theta}\right]$$

p, d, f of $Y_i$is

$$g_{Y_i}(y) = \begin{cases} \frac{n(\theta - y)^{n-1}, 0 \leqslant y \leqslant \theta}{\theta^n} \\ 0, \quad elsewhere \end{cases}$$

Hence, $\quad E(Y_i) = \int_O^\theta \frac{n y (\theta - y)^{n-1}}{\theta^n} d\, y$

$$= \frac{n}{\theta^n}\left\{ -y \frac{(\theta - y)^n}{n} \right|_o^\theta + \frac{1}{n}\int_0^\theta (\theta - y)^n dy$$

$$= \frac{n}{\theta^n}\left[\left[\frac{-1}{n}\frac{(\theta - y)^{n+1}}{n+1}\right]_o^\theta\right.$$

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

So that $\quad E(T_3) = E[(n+1)Y_1] = \theta$

$$\begin{bmatrix} we\ can\ check\ that\ V(T_3) = \dfrac{n}{n+2}\theta^2 \\ so\ that\ V(T_2) < V(T_1) < V(T_3) \end{bmatrix}$$

Example: Let $\left((x_1, x_{2,} \dots x_n\right)$ be a random variable from the Rectangular distribution $R(\theta, 2\theta)$ having p, d, f

$$f(x,\theta) = \begin{cases} \dfrac{1}{\theta}, \theta \le x \le 2\theta \\ 0\ , elsewhere \end{cases}$$

Show that $T_I = \frac{n+1}{2n+1} x_{(n)}, T_2 = \frac{n+1}{n+2} x_{(1)}$

And $T_3 = \frac{n+1}{5n+4}\left[2x_{(n)} + x_{(1)}\right] and\ T_4 = \frac{2}{3}\bar{x}$ are all unbiased

Soln: We can show that the distribution $\uparrow x_{(n)} dx_{(i)}$ have p, d, f given by

$$f_{x_{(n)}}(y) = \frac{n(y-\theta)^{n-1}}{\theta^n} = \theta \le y \le 2\theta$$
$$f_{x_{(1)}}(y) = \frac{n(2\theta - y)^{n-1}}{\theta^n} \theta \le y \le 2\theta$$

Example: Let y_1, y_2, y_3 be the order statistics of a random sample of size 3 from a uniform distribution having $p, d, f\ f(x,\theta) = \frac{1}{\theta}(0 \le x \le \theta)$ show that $4y_1, 2y_2\ , \frac{4}{3}y_3$ are all unbiased estimator of θ. Also obtain their variance.

Soln: We can show that Y_1, Y_2, Y_3 have p, d, f

$$fy_1(y) = \frac{3(\theta - y)}{\theta^3} = o \le y \le \theta$$
$$fy_2(y) = \frac{6y(\theta - y)}{\theta^3} = o \le y \le \theta$$
$$fy_3(y) = \frac{3y^2}{\theta^3} = o \le y \le \theta$$

$E(y_1) = \theta/4, E(y_2) = \theta/2\ , E(X_3) = {}^3/_{4\theta}$

$V(y_1) = 3\theta^2/80, V(y_2) = \theta^2/20, V(y_3) = 3\theta^3/80$

*If $y_1, y_2, \dots, y_n$ are two unbiased estimator with variance σ_1^2, σ_2^2 and correlation coeff. P between than the linear combination which is unbiased and has minimum variance is.

$$Y = \frac{(\sigma_2^2 - P\sigma_1\sigma_2)Y_I + (\sigma_1^2 - \varphi\sigma_1\sigma_2)Y_2}{\sigma_1^2 + \sigma_2^2 - 2\varphi\sigma_1\sigma_2}$$

*If $y_1, y_2, \dots, y_n$ are ind ept unbiased estimators if θ with variance $\sigma_i^2 (i = 1,2..n)$, the linear combination with minimum variance is

$$Y = k_1 y_1 + k_2 k_2 + +k_n k_n$$

Where

$k_i = \frac{1}{\sigma_i^2} / \sum_i^n (i/\sigma_i^2)$

$i.e \quad y = \frac{\frac{I}{\sigma_1^2}y_1 + \frac{I}{\sigma_2^2}y_2 + \cdots + \frac{I}{\sigma_n^2}y_n}{\frac{I}{\sigma_1^2} + \frac{I}{\sigma_2^2} + \cdots + \frac{I}{\sigma_n^2}}$

Example Let 'T' be an unbiased estimator of θ. Does it imply that T^2 and$\sqrt{T}$, are unbiased for $\theta^2 and \sqrt{\theta)}$ respectively?

Soln : $$V(T) = E(T^2) - [E(T)]^2$$

If $E(T^2) = \theta^2$, then $V(T) = O$ so that $P\,(T = \theta) = 1$ which is impossible since T has to be of independent of θ.

Also, $$V(\sqrt{T}) = E(T) - (E\sqrt{T})^2$$

If $E(\sqrt{T}) = \sqrt{\theta}$, then $V(\sqrt{T}) = o$ so that $P(\sqrt{T)} = \sqrt{\theta)} = 1 = P(T = \theta)$ which is impossible.

Example let y_1, y_2, be independent unbiased estimator of θ, having finite variance $(\sigma_1^2, \sigma_2^2, say)$. Obtain a linear combination of y_1, y_2 which is unbiased and has the smallest variance.

Sohn Let $Y = k y_1 + k' y_2$

Evidently, $k + k' = 1$ or $k' = 1 - k$

Then $V(Y) = V[k y_i + (1 - k) y_2]$

$$= k^2 \sigma_1^2 + (I - k)^2 \sigma_2^2$$

Minimising $V(Y)\ w.r.t. k,$ we get

Or $$2k\sigma_1^2 - 2(1 - k)\sigma_2^2 = 0$$

$$k = \sigma_2^2 / (\sigma_1^2 + \sigma_2^2)$$

The linear combination with minimum variance is

$$Y = \left(\frac{\sigma_2^2}{\sigma_1^2 + \sigma_2^2}\right) y_1 + \frac{\sigma_1^2}{(\sigma_1^2 + \sigma_2^2)} y_2 = \frac{\frac{1}{\sigma_1^2} Y_1 + \frac{1}{\sigma_2^2} Y_2}{\frac{1}{\sigma_1^2} + \frac{1}{\sigma_2^2}}$$

Note : if $\sigma_1^2 = 2\sigma_2^2$ then k=1/3

<u>**Remarks:**</u> (i) An unbiased estimator may not exist. Let x be a random variable with Bernoulli distribution.

$$f_\theta(x) = \theta^x (1 - \theta)^{1-x}, x = 0,1$$

It can be shown that no unbiased estimator exists for θ^2.

(ii) Unbiased estimator may be assured.

Let X be a random variable having Poisson distribution $P(x)$ and suppose we want estimator $g(\lambda)$ $=e^{3\lambda}$. Consider a sample of one observation and the estimator T= . Then E(T)= $e^{-3\lambda}$ so that T is an unbiased estimator of $e^{-3\lambda}$ but T(x)= $(-2)^{x}$ for x even and T(x) <0 for x odd, which is absurd since $e^{-3\lambda}$ is always positive.

(iii) Instead of the parameter θ we may be interested in estimating a function $g(\theta)$. $g(\theta)$ is said to be 'estimable' if there exists an estimator T Such that E(T)= $g(\theta), \theta \in \Omega$.

Minimum Variance Unbiased (MVU) estimators : The class of unbiased estimators may, in general, be quite large and we would like to choose the best estimator from this class. Among two estimators of θ which are both unbiased , we would choose the one with smaller variance. The reason for doing this rests on the interpretation of variance as a measure of concentration about the mean. Thus, if T is unbiased for θ, then by Chebyshev's inequality-

$$P\{[T-\theta] \leq \varepsilon\} > 1 - \frac{Var(T)}{\varepsilon^2}$$

Therefore, the smaller $Var(T)$ is, the larger the lower bound of the probability of concentration of T about θ becomes. Consequently, within the restricted class of unbiased estimators we would choose the estimator with the smallest variance.

Definition: An estimator T =T ($X_1,\dots, X_n$) is said to be a uniformly minimum variance unbiased

(UMVU) estimator of θ (or an estimator for $g(\theta)$ if it is unbiased and has the smallest variance within the class of unbiased estimators of θ (or $g(\theta)$,) of all $\theta \in \Omega$. That is if T is any other unbiased estimator of θ, then-

$$Var(T) \leq Var(T') for \text{ all } \theta \in \Omega$$

Suppose we decide to restrict ourselves to the class of all unbiased estimators with finite variance. The problem arises as to how we find an UMVU estimator, if such an estimator exists. For this we would first determine a lower bound for the variances of all estimators (in the class of unbiased estimators under consideration) and then would try to determine an unbiased estimator whose variance is equal to this lower bound. The lower bound for the variances will be given by the Cramer-Rao inequality for which we assume the following regularity conditions:

Let X be a random variable with $p.d.f\ f(x;\theta)\ \theta \in \Omega$

(i) Ω is an open interval (finite or not)

$(ii) f(x;\theta)$ is positive on a set S independent of θ.

$(iii) \frac{\partial}{\partial\theta} f(x;\theta)$ exists for all $\theta \in \Omega$

$(iv) \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x_1, \theta) f(x_2, \theta) \dots f(x_n, \theta) dx_1, x_{2,\dots\dots} d_{x_n}$

May be differentiated under the integral sign.

$$(v) \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} T\,(x_1, x_{2,} \dots x_n) f(x_1; \theta) \dots f(x_n; \theta) dx_1, x_{2,\dots\dots} d_{x_n}$$

May be differentiated under the integral sign where T(X_1, X_n) is any unbiased estimator of θ

Cramer-Rao inequality: Let (X_1,..., X_n) be a random sample of n observations on X with $p.d.f\ f(x; \theta)$ and suppose the above regularity conditions hold. If T is any unbiased estimator of θ, then-

$$Var(T) \leq \frac{1}{nE\left[\frac{\partial}{\partial\theta} log\, f(x; \theta)\right]^2}$$

Proof: We have

$$\int_{-\infty}^{\infty} f(x_i; \theta)\, dx_i = 1\, ;\; i = 1,2..n$$

Which gives, on differentiating.$w.r.t\ \theta$

$$\int_{-\infty}^{\infty} \frac{\partial}{\partial\theta} f(x_i, \theta)\, dx_i = 0$$

Or $\int_{-\infty}^{\infty} \left[\frac{\partial}{\partial\theta} log\, f(x_i; \theta)\right] f(x_i; \theta) dx_i = 0 \dots \dots..(A)$

Or $E\left[\frac{\partial}{\partial\theta} log f(x_i; \theta)\right] = 0 \dots \dots..(1)$

Also, since T is unbiased estimator of $\theta, we\ have$

$$E(T) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} T(x_1, .x_n) f\,(x_i, \theta)..f(x_n, \theta) dx_i \dots dx_n = \theta$$

Which given on differentiation $w.r.t.\theta$

$$E(T) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} T(x_1, .x_n) \frac{\partial}{\partial\theta} \left[\prod_{i=1}^{n} f(x_i, \theta)\right] dx_i \dots dx_n = 1 \dots \dots (2)$$

But

$$\frac{\partial}{\partial\theta} \prod_{i=1}^{n} f(x_i; \theta) = \sum_{i=i}^{n} \left[\frac{\partial}{\partial\theta} f(x_i; \theta) \prod_{i=i} f(x_i; \theta)\right]$$

$$= \sum_{i=i}^{n} \left[\frac{1}{f(x_i; \theta)} \frac{\partial}{\partial\theta} f(x_i; \theta) \prod_{I=i}^{n} f(x_i; \theta)\right]$$

$$= \left[\sum_{i=1}^{n} \frac{\partial}{\partial\theta} logf(x_i;\theta)\right] \prod_{I=i}^{n} f(x_i;\theta)$$

So that (2) becomes

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} T(x_1, \ldots x_n) \left[\sum_{i=1}^{n} \frac{\partial}{\partial\theta} logf(x_i;\theta)\right] f(x_1,\theta) \ldots f(x_n,\theta) dx_i \ldots dx_n = 1$$

Or E (TZ) =I .... (3)

Where

$$Z = \sum_{i=1}^{n} \frac{\partial}{\partial\theta} logf(x_i;\theta)$$

From (1) we immediately get

$$E(Z) = \sum_{i=1}^{n} E\left[\frac{\partial}{\partial\theta} logf(x_i;\theta)\right] = 0 \ldots \ldots . . (4)$$

And

$$Var(z) = \sum_{i=1}^{n} E\left[\frac{\partial}{\partial\theta} logf(x_{1;}\theta)\right]^2$$

$$= nE\left[\frac{\partial}{\partial\theta} \log f(x_1;\theta)\right]^2 \qquad \ldots . . (5)$$

Now, $Cov(TZ) = E(TZ) - E(T)E(Z)$

=1

(i)An unbiased estimator T whose variance equals the lower bound $\frac{1}{nE\left[\frac{\partial}{\partial\theta}\log f(x,\theta)\right]^2}$

If and only if T is if the from $T = \theta + b_\theta^z$ where $z = \sum_{i=1}^{n} \frac{\partial}{\partial\theta} logf(x,\theta)$

Proof:

$$V(T) = \frac{1}{nE\left[\frac{\partial}{\partial\theta} \log f(x,\theta)\right]^2}$$

I*ff*

$$R(T,Z) = 1$$

$i.e$, *iff* T is a linear *f*unction of Z, say

$$T = a_\theta + b_\theta z$$

But $$E(T) = a_\theta = \theta$$

$i.e$ $$T = \theta + b_\theta z$$

Let $(x_1, .. x_n)$ be a random sample from R $(0, \theta)$

$$f(x,\theta) = \frac{1}{\theta}, 0 \leq x \leq \theta$$

$$\frac{\partial}{\partial\theta} log f(x,\theta) = \frac{1}{\theta}$$

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

$$\text{CRB} = \frac{\theta^2}{n}$$

We know that $T = \frac{n+1}{n} X_{(n)}$ is UMVUE whose variance is-

$$V(T) = \frac{\theta^2}{n(n+2)} < \frac{\theta^2}{n}$$

Therefore, we have $$P(T,Z) = \frac{Cov(T,Z)}{V(T)V(Z)} = \frac{1}{V(T)V(Z)}$$

Since $P(T,Z) \leq 1$ we get

$$V(T) \geq \frac{1}{nE\left[\frac{\partial}{\partial\theta}\log f(x.\theta)\right]^2}$$

Remark: (i) the left page

(ii) If $g(\theta)$ is an estimable function for which an unbiased estimator is T $(i.e.\ E(T) = g(\theta))$ then C.R Inequality becomes-

$$V(T) \geq \frac{[g(\theta)]^2}{nE\left[\frac{\partial}{\partial\theta}\log f(x.\theta)\right]^2}$$

(iii) It can be show that

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

(iv) If an unbiased estimator exists which is such that its variance is equal to the lower bound CRB=$\frac{1}{nE\left[\frac{\partial}{\partial\theta}\log f(x.\theta)\right]^2}$ then it will be UMVUE.

(v) If there is no unbiased estimator whose variance equals the C R B it does not mean that UMVUE will not exist. Such estimators can be found (if these exists) by other methods.

(vi) In case of distributions not satisfying the regularity conditions (e.g.: Rectangular distribution) UMVU estimators, if these exists can be found by other methods. For such cases UMVU estimator may have variance less than CRB.

Example: Let $(x_1, \dots x_n)$ be a random sample from a Bernoulli distribution $f(x;\theta) = \theta^x(1-\theta)^{1-x}(x = 0,1), 0 < \theta < 1$

Show that $\bar{x} = \frac{1}{n}\sum_i^n x_i$ is a UMVU of θ

Sohn :

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

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

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

So that

$$E\left[\frac{\partial}{\partial\theta}\log f(x.\theta)\right]^2 = \frac{E(x-\theta)^2}{\theta^2(1-\theta)^2}$$

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

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

By CR inequality we have C R B = $\frac{\theta(1-\theta)}{n}$

Now, $E(\bar{x}) = \theta$ and $Var\,(\bar{x}) = \frac{\theta(1-\theta)}{n}$ that is equal to C R B. Hence $\bar{x}$ is UMVUE of θ

Example: Let x be a random sample having Binomial distribution

$$f(x,\theta) = \binom{m}{x}\theta^x(1-\theta)^{m-x};\ x = 0,1,\dots,m(0 < \theta < 1)$$

Show that $\bar{x}/_m$ is UMVUE of θ.

Soln:

$$log f(x,\theta) = log\binom{m}{x} + x log\theta + (m-x)\log(1-\theta)$$

$$\frac{\partial}{\partial\theta} log f(x,\theta) = \frac{x}{\theta} + \frac{m-x}{1-\theta}$$

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

So that

$$E\left[\frac{\partial}{\partial\theta} log f(x,\theta)\right]^2 = \frac{E(x-m\theta)^2}{\theta^2(1-\theta)^2}$$

$$= \frac{m\theta(1-\theta)}{\theta^2(1-\theta)^2}$$

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

For sample of one observation X let T=T(X) be an unbiased estimator. The C.R.B is $\frac{\theta(1-\theta)}{mn}$. Now $E\left(\frac{\bar{X}}{m}\right) = \theta$ and $Var\left(\frac{x}{m}\right) = \frac{\theta(1-\theta)}{mn}$ so that $\frac{\bar{X}}{m}$ is UMVUE of θ (see left page)

Example: Let (x_1,..., x_n) be a random sample from a Poisson distribution

$$f(x,\theta) = \frac{e^{-\theta}\theta^x}{x}; \quad x = o, 1 \dots\dots (\theta > O)$$

Show that $\bar{x}$ is UMVUE of θ.

Soln:

$$logf(x,\theta) = -\theta + xlog\theta - logx;$$

$$\frac{\partial}{\partial\theta} logf(x,\theta) = -I + \frac{x}{\theta}$$

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

$$E\left[\frac{\partial}{\partial\theta} logf(x,\theta)\right]^2 = \frac{E(x,\theta)^2}{\theta^2}$$

$$= \frac{1}{\theta}$$

The C.R.B = $\theta/_n$

Now $E(\bar{x}) = \theta$ and $Var(\bar{x}) = \frac{\theta}{n}$ so that $\bar{x}$ is UMVUE of θ

Example: Let $(x_1, .., x_n)$ be a random sample from a normal distribution $N(\theta\ ,\sigma^2)$ where variance σ is known show that $\bar{x}$ is UMVUE of θ.

Soln:

$$f(x,\theta) = \frac{1}{\sqrt{2x\sigma}} e^{-\frac{(x-\theta)^2}{2\sigma^2}}$$

$$\log f(x,\theta) = log\left(\frac{1}{\sqrt{2x\sigma}}\right) - \frac{(x-\theta)^2}{2\sigma^2}$$

Or

$$\frac{\partial}{\partial\theta} logf(x,\theta) = \frac{(x-\theta)}{\sigma^2}$$

$$E\left[\frac{\partial}{\partial\theta} logf(x,\theta)\right]^2 = \frac{E(x-\theta)^2}{\sigma^4}$$

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

The C.R.B= $\sigma^2/_n$

Now $E(\bar{x}) = \theta$ and $V(\bar{x}) = \sigma^2/_n$ so that $\bar{x}$ is UMVUE of θ

Example Let $x_1, .., x_n$ be a random sample from a normal distribution $N(\mu, \theta)$ where μ is known and θ is that variance to be estimated. Show that $s^2 = \sum_i^n (x_i - \mu)^2/n$ is UMVUE of θ

Soln:

$$f(x;\theta) = \frac{1}{\sqrt{2\pi\theta}} e^{-\frac{(x-\mu)^2}{2\theta}}$$

$$log\ f(x;\theta) = log\frac{1}{\sqrt{2x}} - \frac{1}{2}log\theta - \frac{(x-\mu)^2}{2\theta}$$

Or

$$\frac{\partial}{\partial\theta} \log \mathrm{f}(\mathrm{x},\theta) = -\frac{1}{2\theta} + \frac{(x-\mu)^2}{2\theta^2}$$

$$= \frac{(x-\mu)^2-\theta}{2\theta^2}$$

$$E\left[\frac{\partial}{\partial\theta}\log f(x,\theta)\right]^2 = \frac{E[(x-\mu)^2-\theta]^2}{4\theta^4}$$

$$= \frac{E(x-\mu)^4 - 2\theta E(x-\mu)^2 + \theta^2}{4\theta^4}$$

$$= \frac{3\theta^2 - 2\theta^2 + \theta^2}{4\theta^4}$$

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

The C.R.B= ${2\theta^2}/{n}$

Consider the estimator $S^2 = \frac{\sum_i^n (x_i-\mu)^2}{n}$ for which E(S²)= θ and V(S²)=$\frac{2\theta^2}{n}$ so that S² is UMVUE of θ

Example An UMVU estimator is unique , in the sense that if T_O and T_I are both UMVU estimator then $T_O = T_I$ almost surely ($i.e\ P(T_O \neq T_I) = 0$)

Soln: Since both T_O and T_I are unbiased

$$E(T_O) = E(T_I) = \theta \text{ for all } \theta \in \Omega$$

And since both are UMVUE,

$$V(T_O) = V(T_I) \text{for all } \theta \in \Omega$$

Consider the new estimator

$$T = \frac{1}{2}(T_O + T_I)$$

Which is also unbiased. Moreover,

$$V(T) = \frac{1}{4}[V(T_O) + V(T_I) + 2\rho\sqrt{V(T_O)V(T_I)}]$$

Where ρ is the corr. Coefficient between T_O *and* T_I

$$V(T) = \frac{I+\rho}{2}V(T_O)$$

By definition, $V(T) \geq V(T_O)$. It follows that $\rho \geq I$. Therefore ρ =I so that, for every θ, T_O and T_I are linearly related, $i.e.$

$$T_O = a + bT_I$$

Where a, b are amstants (may depend on θ) and b≥ 0 . Taking expectation and variance we get

$$\left.\begin{aligned}\theta &= a + b\theta \\ V(T_0) &= b^2 V(T_I)\end{aligned}\right\}$$

Which imply that b=1 and $a = 0$. Therefore

$$T_0 = T$$

CONSISTENCY

Definition: A sequence of estimator $\{T_n\}. n = 1,2, \ldots$ of a parameter θ is said to be consistent if, as n→∞

$T_n \rightarrow p\ \theta$ for each fixed $\theta \in \Omega$ that is , for any $\epsilon(> 0)$

$$T_n\ converges\ to\ \ \theta\ in\ probablity$$

Or $P\{|T_n - \theta| > \epsilon\} \rightarrow 0$

Or $$P\{|T_n - \theta| \leq \epsilon\} \rightarrow 1$$

$as\ n \rightarrow \infty$

Remarks:

($\boldsymbol{i}$) For increase in sample size a consistent estimator will become more and more close to θ

(ii)Consistency is essentially a large sample property. We speak of the consistency of a sequence of estimators rather than that of one estimator.

(iii) If $\{T_n\}$ is a sequence of estimator which is consistent for θ and $\{C_n\}, \{g_n\}$ are sequence of constants such that $C_n \rightarrow 0$ $g \rightarrow 1$ as $n \rightarrow \infty$ then $\{T_n + C_n\}$ and $\{g_n T_n\}$ are sequences of consistent estimators also.

(iv) We will show later that if $\{T_n\}$ is a sequence of estimators such that $E(T_n) \rightarrow \theta$ and $V(T_n) \rightarrow 0$ and $n \rightarrow \infty$ then $\{T_n\}$ is consistent.

Examples:

1. Let $(x_1, \ldots x_n)$ be a random sample from any distribution with finite mean θ. Then it follows from LLN that $\bar{x}$ so that $\overline{x \rightarrow}$ is consistent for θ. If the distribution has finite variance (σ^2, say) $V(\bar{x}) = {}^{\sigma^2}/_{n} \rightarrow 0$ so that it follows from Remark (IV) that $\bar{x}$ is consistent .it can be shown that the sample median is also consistent for θ
2. Suppose $(x_1, \ldots, x_n)$ is a random sample from $N(\mu, \sigma^2)$.
   Let

$$\boldsymbol{\bar{x}} = \sum_{1}^{n} {}^{\boldsymbol{x_i}}/_{\boldsymbol{n}}$$

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

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

4 The following is an example of an estimator which is unbiased but **not** consistent

Let $(x_1, \ldots x_n)$ be a random sample from rectangular distribution. $R(O, \theta)$ and let $Y_i = min(x_1, \ldots x_n)$ consider the estimator $T = (n+1)Y_1$. This is unbiased . Now for a any $E(> 0)$,

$$p\left\{\left[Y_1 - \frac{\theta}{n+1}\right] \leq \frac{\varepsilon}{n+1}\right\}$$

$$= \frac{n}{\theta^n}$$

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

$$=\frac{1}{\theta^n}[-\theta-y)^n]^{\frac{\theta}{n+1}+\frac{\epsilon}{n+1}}_{\frac{\theta}{n+1}-\frac{\epsilon}{n+1}}$$

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

$$=\frac{n^n}{(n+1)^n}\left[\left(1+\frac{\epsilon}{n\theta}\right)^n-\left(1-\frac{e}{n\theta}\right)^n\right]$$

$$\xrightarrow[n\to\infty]{} (e^{\frac{\epsilon}{\theta}}-e^{-\frac{\epsilon}{\theta}})$$

Which is some fixed number

$$P\{[T-\theta]\epsilon\}+1$$

Thus, T is not constant

We can show that

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

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

By remark (iv) above $s^2+s'^2$ are both constant for σ^2, s^2 is biased and s'^2 is unbiased.

3. Let $(x_1, \ldots\ldots x_n)$ be for a random sample for gamma distribution

$$\text{f(x},\theta)=\frac{1}{\theta^b\Gamma(b)}e^{\frac{x}{\theta}x^{b-1}}(x\geq\theta,\theta>0)\ b\ known$$

Show that $\bar{X}/b$ is unbiased and consistent for θ

Soln: $E(\bar{X}/b)=\theta,\ V(\bar{X}/b)=\frac{\theta^2}{nb}\to 0$

$\bar{X}/b$ is unbiased and consistent

<u>Theorem:</u> If $\{T_n\}$_is a sequence of estimators (of θ)such that

$$E(T_n)=\theta_n\to\theta$$

And $$V(T_n)\to 0$$

As n→∞ then $\{T_n\}$ is consistent estimator of θ.

Proof: By Chebyshev's inequality, for any $\epsilon(>0)$we have

$$P\{|T_n-\theta|>\epsilon\}\leq\frac{E(T_n-\theta)^2}{\epsilon^2}$$

$$=\frac{1}{\epsilon^2}E[(T_n-\theta_n)+(\theta_n-\theta)]^2$$

$$=\frac{1}{\epsilon^2}E[(T_n-\theta_n)^2+(\theta_n-\theta)^2+2(\theta_n-\theta)(T_n-\theta)]$$

$$=\frac{1}{e^2}[V(T_n)+(\theta_n-\theta)^2]\to 0$$

As n→∞ by given condition of the theorem so that T_n is consistent for θ.

Theorem: If$\{T_n\}$ is a sequence of consistent estimators of θ and $g(\theta)$ is a continuous function of θ, then $\{g(T_n)\}$ is consistent for $g(\theta)$

Proof: Since T_n is consistent for θ, for any $\epsilon_1(> o)$

$$P\{|T_n-\theta|\le\epsilon_1\}\to 1$$

As n→∞

Also , since $g(\theta)$ is a continuous function , given $\epsilon(> 0)$we can choose $\epsilon_1(> o)$such that

$$|T_n-\theta|\le\epsilon_1\to|g(T_n)-g(\theta)|\le\epsilon$$

Therefore ,

$$P\{|T_n-\theta|\le\epsilon_1\}\le P\{|g(T_n)-g(\theta)|\le\epsilon\}$$

But as n→∞, L.H.S → 1 and, consequently, R.H.S →1, $i,e.$

$$P\{|g(T_n)-g(\theta)|\le\epsilon\}\to 1$$

As n→∞. Hence $g(T_n)$ is consistent for$g(\theta)$.

We can prove the following results:

(i) If $\{T_n\}$ is consistent for, then T_n^2 is consistent for θ^2.

(ii) If $\{T_n\}$ is consistent for θ(R and non-negative) then $\sqrt{T_n}$ is consistent for $\sqrt{\theta}$.

Proof For any $\epsilon(> 0)$ we have

$$P\{|T_n-\theta|\ge\epsilon\}P\{|(\sqrt{T_n}-\sqrt{\theta})(\sqrt{T_n}-\sqrt{\theta})|\ge\epsilon\}$$

$$=P\left\{|\sqrt{T_n}-\sqrt{\theta})|\ge\frac{\epsilon}{\sqrt{T_n}+\sqrt{\theta}}\right\}$$

$$\ge P\left\{|\sqrt{T_n}-\sqrt{\theta}|\ge\frac{\epsilon}{\sqrt{\theta}}\right\}$$

Since L. H. S→0, R. H.S →0 as n→∞

(iii) If $\{T_n\}$ is consistent for θ and $\{T'_n\}$ *is consistent for* θ', then $\{T_n\pm T'_n\}$ is consistent for $\theta+\theta'$.

Proof: for any $\epsilon(> 0)$,we have $P\{|(T_n+T_n')-(\theta+\theta')|\ge\epsilon\}$

$$\le P\{|T_n-\theta|+|T'_n-\theta'|\ge\epsilon\}$$

$$\leq P\left\{|T_n - \theta| \geq \frac{\epsilon}{2}\right] U|T'_n - \theta'| \geq \epsilon\right\}$$

$$\leq P\left\{|T_n - \theta| \geq \frac{\epsilon}{2}\right\} + P\{|T'_n - \theta'| \geq \frac{\epsilon}{2}\} \rightarrow 0$$

As n→∞.

There fore $\{T_n + T'_n\}$ is consistent for $(\theta + \theta')$

(iv)if T_n and T'_n are consistent for θ and θ' respectively , $T_n T'_n$ is consistent for $\theta\theta'$.

<u>Proof</u>: we can write

$$T_n T'_n = \frac{(T_n + T'_n)^2 - (T_n - T'_n)^2}{4}$$

$$\xrightarrow{\text{þ}} \frac{(\theta + \theta')^2 - (\theta - \theta')^2}{4}$$

<u>EFFICIENCY</u>:

If T_1 and T_2 are two unbiased estimators of a parameter θ , each having finite variance T_1 is said to be more efficient then T_2 if $V(T_1)$ >$V(T_2)$. The (relative) efficient of T_1 relative to T_2 is defined by

$$\text{Eff}(T_1/T_2) = \frac{V(T_2)}{V(T_1)}$$

It is used to judge the efficiency of an unbiased estimator by comparing its variance with the Cramer- Rao lower bound (C R B) .

<u>Definition:</u> Assume that the regularity condition of CR inequality hold (we call it a regular situation) for family$\{f(x,\theta), \theta \in \Omega\}$. An unbiased estimator T* of θ is called <u>most efficient</u> if $V(T^*)$ equals the CRB. In this situation, the 'efficiency' of any other unbiased estimator T of θ is defined by

$$\text{Eff(T)} = \frac{V(T^*)}{\text{V(T)}}$$

Where T* is the most efficient estimator defined above

<u>Remarks:</u>

(i)The above definition not proper in—

(a) regular situation when there is no unbiased estimator whose variance equals the CRB but an UMVUE exists and maybe found by other methods.

(b)Non-regular situations when an UMVUE exists and may be found by other methods

(ii)The UMVUE is 'most efficient' estimator in the examples considered earlier all UMVUE, whose variances equalled CRB are most efficient

Example Consider $a, r, s(x_1, \dots x_n)$ from a normal distribution$N(\mu, \theta)$ where mean μ is known and variance $\theta(0 < \theta < \infty)$ is to be estimated

We has seen that $s^2 = \frac{1}{n}\sum_1^n (x_i - \mu)^2$ is UMVUE of θ for which the variance is equal to CRB and consequently, s^2 is most efficient . Let $s'^2 = \frac{1}{(n-1)}\sum_1^n (X_i - \overline{\overline{X}})^2$

Then $E(S'^2) = \theta$ and $V(S'^2) = \frac{2\theta^2}{n-1}$ so that the efficiency of s'^2 is given by

$$\text{Eff}(s'^2) = \frac{2\theta^2/n}{2\theta^2/(n-1)}$$

$$= \frac{\text{n} - 1}{\text{n}}$$

Asymptotic efficiency: As different from the above definition of efficiency we may define efficiency in another way as follows, which may be called asymptotic efficiency.

Let us confine ourselves to consistent estimators which are asymptotically normally distributed. Among this class, the estimator with the minimum asymptotic variance is called the 'most efficient estimator'. It is also called best asymptotically normal (BAN) or consistent asymptotically normal efficient (CANE) estimator it we denote by avar(T*) the asymptotic variance of a BAN estimator T* then the efficiency of any other estimator T (within the class of asymptotically normal estimators) is defined by

$$Eff(T/\text{T}^*) = \frac{avar(\text{T}*)}{avar(T)}$$

Where avar (T) is the asymptotic variance of T.

Example: Let $(x_1, \dots, x_n)$ be a random sample from a normal distribution $N(\mu, \sigma)$, Consider the 'most efficient estimator $\bar{x}$ and another estimator $\bar{x}_{\text{me}}$. It can be show that both are CAN estimator. We have

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

And
$$V(\bar{x}_{me}) = \frac{\pi}{2}\frac{\sigma^2}{n}$$

So that the efficiency of $\bar{x}_{me}$ is given by

$$\text{Eff}(\bar{x}_{me}) = \frac{2}{\pi}$$

Example: let T_1, T_2 be two unbiased estimators of θ, having the same variance. Show that the correlation coefficient ρ between T_1, T_2 cannot be smaller than 2e-1, where e is the efficiency of each estimator,

Proof. Let T_o be the most efficient estimator then

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

Consider the unbiased estimator

$$T = \frac{\boldsymbol{T_1} + \boldsymbol{T_2}}{2}$$

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

$$= \left[\frac{V(T_o)}{e} + \frac{V(T_o)}{e} + 2\rho\frac{V(T_o)}{e}\right]$$

$$= \frac{I + \rho}{2e}V(T_o)$$

Since $T_o$is UMVUE, V (T)$\geq V(T_o)$ which gives

$$\frac{I + \rho}{2e} \geq 1 \quad or \quad \rho \geq 2e - 1$$

Example: let T_o be an UMVME (or most efficient estimator) where T_1 an unbiased with efficiency 'e'. If ρ is the correction coefficient between T_o andT_1, then show that ρ =$\sqrt{e}$.

<u>Soln:</u> we have

$$e = V(T_o)/V(T_1)$$

Or $$V(T_1) = V(T_o)/e$$

Consider the estimator

$$T = \frac{(1 - \rho\sqrt{e})T_o + \sqrt{e}(\sqrt{e} - \rho)T_1}{1 - 2\rho\sqrt{e} + e}$$

(Which the linear combination of T_o, T_1 with minimum variance) then T is also unbiased, having variance

$$V(T) = [(1 + \rho^2 e - 2\rho\sqrt{e})V(T_o) + e(e + \rho^2 - 2\rho\sqrt{e})\frac{V(T_o)}{e}$$

$$\frac{+2\sqrt{e}(\sqrt{e} - \rho - \rho e - \rho^2\sqrt{e})\rho\left[\frac{V(T_o)}{\sqrt{e}}\right]}{[1 - 2\rho\sqrt{e} + e]^2}$$

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

Or $$V(T) = \frac{(1-\rho^2)V(T_o)}{1-2\rho\sqrt{e}+e} = \frac{1-\rho^2}{(1-\rho^2)+(\sqrt{e}-\rho)^2}V(T_o)$$

Since $(1 - \rho^2)$ and$(\sqrt{e} - \rho$ are both non-negative $V(T) \leq V(T_o)$ but since T_o is UMVUE, $V(T) \quad V(T_o)$. therefore $V(T) = V(T_o)$, and $\rho = \sqrt{e}$

<u>SUFFICIENCY CRITERION</u>:

A preliminary choice among statistics for estimating θ , before having for a UMVUE as BAN estimator, can be made on the basic of another enter on suggested by R.A fisher. This is called 'sufficiency' criterion.

<u>Definition:</u> let $(\boldsymbol{x_1}, \dots, \boldsymbol{x_n})$ be a random sample from the distribution of X having þ, $d, f\ f(x, \theta)$ $\theta \epsilon\ \Omega$.A statistic $T = T(\boldsymbol{x_1}, \dots, \boldsymbol{x_n})$ is defined to be **<u>sufficient statistic</u>** if and only if the conditional distribution of $(\boldsymbol{x_1}, \dots, \boldsymbol{x_n})$ given T=t does not depend on θ, for any value t.

[Note: In such a case if we know the value of the sufficient statistic T, then the sample values are not needed to tell us anything more about θ].

Also the conditional distribution of any other statistic T (which is not for Ω tray) given T is independent of θ.

A necessary and sufficient condition for T to be sufficient for θ is that the joint þ. d, f of $(\boldsymbol{x_1}, \dots, \boldsymbol{x_n})$ should be of the form

$$f(x_1, \dots, x_n; \theta) = g(\mathrm{T}, \theta) h(x_1, \dots, x_n)$$

Where the first term on $r, h, s.$, depends on T and θ and the second them is independent of θ. *This* is known as <u>Nyman's Factorisation Theorem</u> which provides a simple method of judging whether a statistic T is sufficient

Remark: Any one to one function of a sufficient statistic is also a sufficient statistic

Example: Consider n Bernoulli trials with probability of success P. The associated Bernoulli random variables $(x_1, \dots, x_n)$ have common distribution given by

$$f(x, p) = p^x (1-p)^{1-x}, x = 0,1$$

The joint probability function of $(\boldsymbol{x_1}, \dots, \boldsymbol{x_n})$ is

$$f(x_1, .., x_n, p) = \mathrm{p}^{\sum_{\mathrm{i}}^{\mathrm{n}} \mathrm{x_i}} (1-p)^{n - \sum_i^n x_i}$$

$$= g\left(\sum \mathrm{x_i}, \mathrm{p}\right) (\mathrm{x_i}, \mathrm{x_n})$$

Where $\qquad g(\sum_i^n x_i, p) = p^{\sum_i^n x_i} (1-p)^{n - \sum_i^n x_i}$

And $h(x_1, \dots, x_n) = 1$

Therefore$\sum_i^n x_i$ is sufficient for p, and, so is$\bar{x} = \sum_i^n x_i / n$.

Example $(x_1, \dots, x_n)$ be a random sample from a position distribution $\mathrm{P}(\lambda) i.e$

$$f(x_i, \lambda) = \frac{e^{-\lambda} \lambda^x}{x!}, \quad x = 0,1 \dots$$

The joint probability function of $(x_1, \dots, x_n)$ *is*

$$f\big((x_1, \dots, x_n), \lambda\big) = \frac{e^{-n\lambda} \lambda^{\epsilon x_i}}{\prod_i^n x_i i}$$

$$= g\left(\sum_{\mathrm{i}}^{\mathrm{n}} \mathrm{x_i}, \lambda\right) h(\mathrm{x_i}, \mathrm{x_n})$$

Where $\qquad g(\sum_{\mathrm{i}}^{\mathrm{n}} \mathrm{x_i}, \lambda) = e^{-n\lambda} \lambda^{\sum_i^n x_i}$

$$h(x_i, x_n) = \frac{1}{\prod_i^n x_i i}$$

Hence.

$$\sum_i^n x_i \quad or \sum_i^n x_i \;/n$$

Are sufficient for λ

Example: let $(x_1, \dots, x_n)$be a random sample from a Normal population $N(\mu, \sigma)$.

<u>**Case I:**</u> μ unknown, σ known (=σ_o)

$$f((x_1, \dots, x_n), \mu) = \frac{1}{(\sigma_o \sqrt{2\pi})^{n e^{-\sum_i^n \{x_i - \mu\}^2 / 2\sigma_0^2}}}$$

$$= \frac{1}{(\sigma_o\sqrt{2\pi})^{n-e^{[\sum_i^n x_1^2+n_\mu^2-2nx\mu]/2\sigma_0^2}}}$$

$$= \mathcal{g}(\bar{x}, \mu)\mathcal{h}(x_1, .. x_n)$$

$$= [2\mu^2 - 2n\bar{x}, \mu]/\ 2\sigma_0^2$$

Where $\mathcal{g}(\bar{x}, \mu) = e - \sum_i^n x_i/\ 2\sigma_0^2$

As

$$\mathcal{h}(x_i, .. x_n) = \frac{1}{(\sigma_o\sqrt{2\pi})^{ne}}$$

Which show that $\bar{x}$ is sufficient for μ .

<u>Case II:</u> μ is know($= \mu_o$), σ unknown

$$f((x_1, \ldots, x_n), \sigma) = \frac{1}{(\sigma_o\sqrt{2\pi})^{ne}} -^{\sum_i^n (x_i-\mu_o)^2} /2\sigma_0^2$$

$$= \mathcal{g}\left[\sum_i^n (x_i - \mu_\theta)^2, \sigma\right] \mathcal{h}(x_i, x_n)$$

Where

$$\mathcal{g}\left[\sum_i^n (x_i - \mu_\theta)^2, \sigma\right] = \frac{1}{(\sigma_o\sqrt{2\pi})^{ne}} -^{\sum_i^n (x_i-\mu_o)^2} /2\sigma_0^2$$

Which show that $\sum_i^n (x_i - \mu_o)^2$ is sufficient for σ

<u>Case III:</u> Both μ and σ are unknown

$$f(x_i, x_n, \mu, \sigma) = \frac{1}{(\sigma_o\sqrt{2\pi})^{ne}} -^{\sum_i^n (x_i-\mu_o)^2} /2\sigma_0^2$$

$$= \frac{1}{(\sigma_o\sqrt{2\pi})^{ne}} -^{[\sum_i^n x_i^2-2\mu\sum_i^n x_i+2_\mu^2]} 2\sigma_0^2$$

Which shows that $[\sum_i^n x_i, \sum_i^n x_1^2]$ an jointly sufficient for $[\mu, \sigma]$ Similarly,[$\bar{x}, \sum(x_i, x)^2$ / n-1]are also sufficient for $[\mu, \sigma]$,

Example let $(x_1, \ldots, x_n)$ be a random sample from a gamer distribution having þ, d, f

$$f(x, \theta, þ) = \frac{1}{\theta^{þ}(þ)} e^{-\frac{x}{\theta}x^{þ-1}, x \geq \theta}$$

We have

$$f(x_i, x_n, \theta, þ) = \frac{1}{\theta^{nþ}(þ)n^e} -^{\sum_i^n x_i/\theta} (\prod_i^n x_i)^{þ-1}$$

<u>Case I</u> θ unknown but þ is known

We can write

$$f\big((x_1,\dots,x_n),\theta\big)=\left[\frac{1}{\theta^{n\text{þ}}([(\text{þ}))n^{e}}^{-\Sigma x_i/\theta}\right]\left[\prod_{i}^{n} x_i\right]^{-\text{þ}-1}$$

So that $\sum_i^n x_i(or\ \bar{x})$ is sufficient for θ.

Case II: θ Known but þ unknown

We can write

$$f\big((x_1,\dots,x_n),\text{þ}\big)=\left[\frac{1}{\theta^{n\text{þ}}([(\text{þ}))n^{e}}\left(\prod_{i}^{n} x_i\right)^{\text{þ}-1}\right]\left[e^{\Sigma x_i/\theta}\right]$$

So that is sufficient for þ

Case III : Both θ and þ are unknown it is seen that $(\sum_i^n x_i, \prod_i^n x_i)$ are jointly sufficient for $(\theta,\text{þ})$

Example: let $(x_i, x_n,)$ be a random sample from the experiential distribution

$$f(x,\theta)=\frac{1}{\theta}e^{-\frac{x}{\theta}}, x\geq\theta$$

It follows from above that $\sum_i^n x_i(or\bar{x})$ is sufficient for θ.

Example let $(x_1,\dots,x_n)$be a random sample from the distribution with þ, d, f

$$f(x,\theta)=\theta x^{\theta-1},\theta\leq x\leq 1$$

We have

$$f\big((x_1,\dots,x_n),\theta\big)=\theta^n\left(\prod x_i\right)^{\theta-1}=\left[\theta^n\left(\prod x_i\right)^{\theta}\right]\left[\frac{1}{\prod_i^n x_i}\right]$$

So that $\prod_{i=1}^n x_i$ is sufficient for θ

Example let$(x_1,\dots,x_n)$be a $a.r.s$ from the Laplace distribution having þ, d, f

$$f(x,\theta)=\frac{1}{2}e^{-[x-\theta]},\infty<x<\infty$$

We have

$$f\big((x_1,\dots,x_n),\theta\big)=\frac{1}{2^n}e^{-\sum_{i=1}^n[x_i-\theta]}$$

For no single statistics T it is possible to express the above in the form $g[\mathrm{T},\theta]h(x_i,x_n,)$. Hence there exists no statistic T which taken alone is sufficient for θ. However the whole set $(x_1,\dots,x_n)$ or the set of order statistics $(x_{(1)},\dots,x_{(n)})$is jointly sufficient for θ

Example let $(x_1,\dots,x_n)$ be a random sample from the Rectangular distribution $R(0,\theta)$ having þ, d, f.

$$f(x,\theta)=\frac{1}{\theta},-\theta\leq x\leq\theta$$

We have

$$f(x_i,x_n,\theta)=\frac{1}{\theta^n}\prod_{i=1}^n I_{[\theta,\theta]}(x_I)$$

Where $I_{A(x)}$ is the indicator function for which

$$I_{A(x)} = \begin{cases} 1 & if \quad x \in A \\ 0 & if \quad x \notin A \end{cases}$$

But $$\prod_{i=1}^{n} I_{[\theta,\theta]}(x_I) = I_{[O,X_{(1)}}(X_{(1)} I_{[x(!),o]}(x_{(n)})$$

Where $X_{(1)}$ and $x_{(n)}$ are the minimum and maximum of sample values$(x_1, \dots, x_n)$

Therefore, we can write

$$f((x_1, \dots, x_n), \theta) = g[\mathrm{x}_{(\mathrm{n})}, \theta] h((x_i, x_n)$$

Where $$g[\mathrm{x}_{(\mathrm{n})}, \theta] = \frac{1}{\theta^n} I_{[\mathrm{x}_{(\mathrm{n})}, \theta]}(x_{(n)})$$

$$h(x_i, x_n) = I_{[O, x_{(n)}]}(x_i)$$

Where shows that $x_{(n)}$ is sufficient for θ

Example : If x has þ, d, f

$$f(x, \theta) = \frac{1}{\theta}; -\theta \leq x \leq \theta$$

We can check that

$$f(x_i, x_n, \theta) = \frac{1}{\theta^n} I_{[-\theta X_{(n)}]}(n_{(n)}) I_{[X_{(1)}, O]} X_{(n)}$$

So that $x_{(1)}$ is sufficient for θ

Example Let $(x_1, \dots, x_n)$ be a random sample from the rectangular distribution $R(\theta_1, \theta_2)$ having þ, d, f

$$f(x, \theta_1, \theta_2) = \begin{cases} \frac{1}{\theta_2 - \theta_1} & \theta_1 \leq x_i \leq \theta_2 \\ o & elsewhere \end{cases}$$

The þ, $d, f((x_1, \dots, x_n))$ is given by

$$f((x_1, \dots, x_n), \theta_1, \theta_2) = \frac{1}{(\theta_2 - \theta_1)^n} \prod_{i=1}^{n} I_{[\theta_1, \theta_2]}(X_I)$$

Where $$I_{[\theta_i \theta_i]}(x_i) = \begin{cases} 1 & if \theta_1 \leq x_i \leq \theta_2 \\ 0 & elsewhere \end{cases}$$

We can write

$$\sum_{i}^{n} I_{[\theta_i \theta_i]}(x_i) = I_{[\theta_i x_{(i)}]}(x_{(i)}) I_{[X_{(I)}, \theta_2]}(x_{(i)})$$

$$= g[x_{(i)}, x_{(n)}, \theta_1 \theta_2] h(x_i, x_n)$$

Where

$$g[x_{(i)}, x_{(n)}, \theta_1 \theta_2] = \boldsymbol{I}_{[\boldsymbol{\theta_1} \cdot \boldsymbol{X_{(n)}}]}(\boldsymbol{x}_{(i)}) \mathrm{I}_{\{\mathrm{X}_{(1)}, \theta_2]}(x_{(i)})$$

And $$h((x_1, \dots, x_n)) = 1$$

Hence $\left[x_{(1)},\ldots,x_{(n)}\right]$ are jointly sufficient for θ_1,θ_2

<u>Corollary :</u> If θ_1 is known $x_{(n)}$ is sufficient for θ_2

If θ_1 is known $x_{(i)}$ is sufficient for θ_1

Example: let $\left((x_1,\ldots,x_n)\right)$ be a,r,s from the rectangular distribution R (-θ,θ).

$$f(x,\theta)=\frac{1}{2\theta},-\theta\leq x\leq\theta$$

Then

$$f\left((x_1,\ldots,x_n),\theta\right)=\frac{1}{(2\theta)^n}\prod_{i=1}^{n}I_{[-\theta,\theta]}(Xi)$$

$$=\frac{1}{(2\theta)^n}I_{[-\theta,,x_{(n)}}(x_{(i)})I_{[X_{(n)},\theta]}x_{(n)}$$

So that $\left[x_{(1)},\ldots,x_{(n)}\right]$ are jointly sufficient for θ

Example: $\left[x_{(1)},\ldots,x_{(n)}\right]$ are jointly sufficient forθ in $R(\theta-\frac{1}{2},\theta+\frac{1}{2})$ and $R(\theta,\theta+1)$

Example: Let $(x_1,\ldots,x_n)$be a random from an exponential distribution

$$f(x)=\lambda_e-^{\lambda(x-\theta)},\theta\leq x<\infty$$

<u>Case I : λ</u> Unknown θ known $(=\theta_o)$

$$f\left((x_1,\ldots,x_n),\lambda\right)=\lambda_e^{n-\lambda}\sum_{i}^{n}(x_i-\theta)\prod_{i=1}^{n}I_{[\theta,\infty)}(x_i)$$

Which show that $\sum_i^n(x_i-\theta_o)$ is sufficient for λor $\bar{x}$ is sufficient for λ

Case II: λ know $(=\lambda_o)$, θ Unknown

$$f((x_i,x_n,\theta)=\lambda_{oe^{-\lambda_o}}\sum_{i}^{n}(x_i-\theta)\prod_{i=1}^{n}I_{[\theta,\infty)}(x_i)$$

$$=\lambda_{oe^{-\lambda_o}}\sum_{i}^{n}x_i+\lambda n\theta\, I_{[\theta,\infty)}(X_{(i)})$$

$$\prod i_x(i),x)(x_{(i)})$$

$$=\left\{e^{\lambda n\theta}I_{[\theta,\infty)}(X_{(i)})\right\}\left\{xoe^{n-\lambda\theta\sum_i^n x_i}I_{[x,(i)\infty)}(X_{(i)})\right\}$$

Which shows that $x_{(i)}$ is sufficient for θ

Case III: Both λ,θ unknown

It is easy to check that $[\sum x,x(i)]$ are jointly sufficient for $[\lambda,\theta]$

<u>METHHODS OF ESTIMATION</u>:

For important methods of obtaining estimators are (I) methods of moments,(II) methods of maximum likelihood (III)method of minimum $\chi 2$ and (IV) method of least squares.

(I)Method of moments

Suppose the distribution of a random variable X has K parameters $(\theta_1, \theta_2, \dots \dots .. \theta_k)$ which have to be estimated. let $\mu_r = E(x^r)$ denote the r^{th} moment of about O. in general μ'_r is a known function of $\theta_1, \dots, \theta_k$ so that $= \mu_r(\theta_1, \dots \theta_k)$Let $(x_1, \dots x_n)$ be a random sample from the distribution of X and let $m_r = \sum_i^n x_i^r / n$ be the r^{th}. Sample moment from the equation

$$m'_r = \mu'_r(\theta_1, \dots \theta_k), r = 1, \dots k$$

Whose solution is say $\hat{\theta}_1 \dots \hat{\theta}_k$, where $\hat{\theta}_i$ is the estimate of $\theta_i (i = 1, .. k)$ Those are the method of moments estimators of the parameters.

Example let

$$x = N(\mu, \sigma)$$

$$\mu'_1 = \mu$$

$$\mu'_2 = \sigma^2 + \mu^2$$

The equation

$$\bar{x} = \mu$$

$$\frac{\sum x_i^2}{n} = \sigma^2 + \mu^2$$

Have the solution

Let

$$\mu = \bar{x}$$

$$\sigma = \sqrt{\frac{\sum_i^n x_i^2}{n} - \bar{x}^2} = \sqrt{\frac{\sum_i^n (x_i - \bar{x})^2}{n}}$$

Example let x ~P (λ) and let $(x_1, .., x_n)$ be random sample from P(λ).

$$\mu'_1 = \lambda$$

The equation

$$\bar{x} = \lambda$$

Provides the estimator

$$\lambda = \bar{x}$$

Example let $(x_1, .., x_n)$ be a random sample from the exponential distribution

$$f(x, \theta) = \theta e^{-\theta x}, x \geqslant \theta$$

$$\mu'_1 = \frac{1}{\theta}$$

The moment equation

$$\bar{x} = \frac{1}{\theta}$$

Provides the estimator

$$\hat{\theta} = \frac{1}{\bar{x}}$$

Remark: (I) the method of moments estimators are not uniquely defined. We may equate the central moments instead of the raw moments and obtain solutions.

(II) These estimator are not, in general, consistent and efficient but will be so only if the parent distributions is of particular form.

(III) When population moments do not exist $(e.g.\ C$auchy population) this method of estimation is inapplicable.

METHOD OF MAXIMUM LIKELIHOOD

Consider$f(x_1, .., x_n, \theta)$, the joint þ, d, f of sample $(x_1, .., x_n)$ of observations of $a, r, s. X$ having the þ, d, f $f(x, \theta)$ whose parameters θ is to be estimated. When the values $(x_1, .., x_n)$ are given, $f(x_1, .., x_n, .. \theta)$ may be looked upon as a function of θ which is called the likelihood function of θ and is denoted by $L(\theta) = L(\theta, x_1, .., x_n)$ it gives the likelihood that the $r, v. (x_1, .., x_n)$assumes the value $(x_1, .., x_n)$ when θ is the parameter.

We want to know from which distribution ($i.e.$ for what value of θ) is the likelihood largest for this set of observations. In other words we want to find the value of θ, denoted by $\hat{\theta}$ which maximizes $L(x_1, .., x_n, \theta)$. The value $\hat{\theta}$ maximizes the likelihood function is in general, a function of $x_1, .., x_n$ say

$$\hat{\theta} = \hat{\theta}(x_1, .., x_n)$$

Such that $$L(\hat{\theta}) = \max \quad L(\theta, , x_1, .., x_n) \quad \theta \in \Omega$$

Then $\hat{\theta}$ is called the maximum likelihood estimator or MLE.

In many cases it would be more convenient to deal with $\log L(\theta)$, rather then $L(\theta)$, since $\log L(\theta)$ is maximized for the some value of θ as $L(\theta)$. For obtaining $m. \ell. e$ we find the value of θ for which

$$\frac{\partial}{\partial\theta}\log L(\theta) = 0 \ \ldots\ldots\ldots (1)$$

We must however, check that this provides the absolute maximum. It the derivate dose not exists at $\theta = \hat{\theta}$ or equation (1) is not solvable this method of solving (1) will fail.

Example: Let $(x_1, .., x_n)$ be a, r, s from the Bernoulli distribution.

$$f(x,\theta) = \theta^x(1-\theta)^{i-x}, x = \theta, 1$$

Then the likelihood $\quad L(\theta, x_1, .., x_n) = \theta^{\sum_i^n x_i}(i-\theta)^{n-\sum_i^n x_i}$

And $\quad \log \ L(\theta) = (\sum_i^n x_i) \log\theta + (n - \sum_i^n x_i) \log(1-\theta)$

Differentiating and equating to zero, we have

$$\frac{\partial}{\partial\theta}\log L(\theta) = 0$$

i, e $\quad \frac{\sum_i^n x_i}{e} - \frac{n-\sum_i^n x_i}{i-\theta} = 0$

Or $\quad \frac{\sum_i^n x_i - n\theta}{\theta(i-\theta)} = 0$

Or $\quad e = \sum_i^n x_i / n = \bar{x}$

$m.\ell.e$ of θ is $\hat{\theta} = \bar{x}$

Example: Let $(x_1, .., x_n)$ be a, r, s from the Poisson's distribution

$$f(x,\lambda) = \frac{e^{-\lambda}\lambda^x}{x!}, x = 0,1,2, \ldots.$$

Then $\quad L(\lambda, x_1, .., x_n) = \frac{e^{-nx}\lambda^{\sum_i^n x_i}}{\prod_i^n x_i!}$

And $\log L(\lambda) = nx + (\sum_i^n x_i)\log\lambda - \sum_i^n log x_i!$

Or $\quad \frac{\partial}{\partial\theta}\log L(\lambda) = -n + \frac{\sum_i^n x_i}{\lambda}$

Equating to zero we get $\lambda = \bar{x}$

$m.\ell.e$ of λ is $\hat{\lambda} = \bar{x}$

Example: Let $x_1,..,x_n)$ be a, r, s from the truncated Binomial distribution having þ, d, f

$$f(x,\theta) = \binom{2}{x}\frac{\theta^x(i-\theta)^{2-x}}{i-[i-\theta]^2}, x = i, 1,2(\theta < \theta < 1$$

Then $$L(\theta, x_i, x_n) = \prod_i^n \binom{2}{x_i}\frac{\theta^{2x_i}(1-\theta)^{2n-2x_i}}{[i-(i-\theta)^{2n}]}$$

And $\log L(\theta) = \sum_i^n log\binom{2}{x_i} + (\sum x_i)\log\theta + (2n - 2x_i)\log(1-\theta) - nlog[1-(1-\theta)^2]$

$$\frac{\partial}{\partial\theta}\log L(\theta) = \frac{\sum_i^n x_i}{\theta} + \frac{\sum_i^n x_i - 2n}{1-\theta} + \frac{2n(1-\theta)}{1-(1-\theta)^2}$$

Equating to zero we get

$$\sum x_i[(1-\theta)\{1-(1-\theta)^2\}] + (\sum x_i - 2n)[\theta\{1-(1-\theta)^2\}]$$

$$-2n\theta(1-\theta)^2] = \theta$$

Or $$\sum x_i[1-(1-\theta)^2] = 2n\theta$$

Or $$\sum x_i[\theta(2-\theta)] = 2n\theta$$

Or $$2 - \theta = \frac{2}{\pi}$$

Or $$\theta = 2 - \frac{2}{\pi}$$

$m.\ell.e$ is $\theta = 2 - \frac{2}{\pi}$

Example: Let $(x_1,..,x_n)$ be a, r, s from the normal distribution $N(\mu, \sigma)$

Case I: μ unknown but $\sigma = \sigma_0$(known)

Then $$L(\mu, x_1,..,x_n) = \frac{1}{(\sigma\sqrt{2\pi})^n} e^{-\sum_i^n(x_i-\mu\theta)^2/2\sigma^2}$$

And $$logL(\mu) = -nlog(\sigma_\theta\sqrt{2\pi}) - \sum_i^n(x_i - \mu)^2/2\sigma_\theta^2$$

Or $$\frac{\partial}{\partial\theta}\log L(\mu) = \frac{\sum_I^N (X_I-\mu)}{\sigma_\theta^2}$$

Equating to zero we get $\mu = \bar{x}$

$m.\ell.e$ Of $\hat{\mu} = \bar{x}$

Case II: $\mu = \mu_0$(known)but σ unknown

Then $$L(\sigma, x_1, .., x_n) = \frac{1}{(\sigma\sqrt{2\pi})^n} e^{-\sum_i^n (x_i-\mu\theta)^2 / 2\sigma^2}$$

And $$\log L(\sigma) = -nlog\sigma - \frac{n}{2}\log(2\pi) - \frac{-\sum_i^n (x_i-\mu\theta)^2}{2\sigma^2}$$

Or $$\frac{\partial}{\partial\sigma} logL(\sigma) = -\frac{n}{\sigma} + \frac{\sum_i^n (x_i-\mu\theta)^2}{\sigma^3}$$

Equating to zero we get $$\sigma = \sqrt{\frac{\sum_i^n (x_i-\mu\theta)^2}{n}}$$

$m.\ell.e$ Of σ is $$\hat{\sigma} = \sqrt{\frac{\sum_i^n (x_i-\mu\theta)^2}{n}}$$

Case III: Both μ and σ are unknown

Then $$L(\mu, \sigma, x_1, .., x_n) = \frac{1}{(\sigma\sqrt{2\pi})^n} e^{-\sum_i^n (x_i-\mu\theta)^2 / 2\sigma^2}$$

And $$logL(\mu,\sigma) = -\frac{n}{2} log\sigma - \frac{n}{2}\log(2\pi) - \frac{\sum (x_i-\mu)^2}{2\sigma^2}$$

Differentiating partially $w.r.t$ μ, σ we get

$$\frac{\partial}{\partial\mu} logL(\mu,\sigma) = \frac{\sum (x_i-\mu)}{2\sigma^2}$$

And $$\frac{\partial}{\partial\sigma} logL(\mu,\sigma) = \frac{n}{\sigma} + \frac{\sum (x_i-\mu)^2}{\sigma^3}$$

Equating to zero both the derivatives and solving the equations we get $\mu = \bar{x}$ and $\sigma = \sqrt{\frac{\sum_i^n (x_i-\bar{x})^2}{n}}$

$m.\ell.e$ are $\hat{\mu} = \bar{x}$ and $\hat{\sigma} = \sqrt{\frac{\sum_i^n (x_i-\bar{x})^2}{n}}$

Example: Let $(x_1, .., x_n)$ be a, r, s from the exponential distribution

$$f(x,\theta) = \frac{1}{\theta} e^{-\frac{x}{\theta}}, x \geqslant \theta$$

Then $$L(\theta, x_1, .., x_n) = \frac{1}{\theta^e} e^{-\sum_i^x x_i/\theta}$$

And $$\log L(\theta) = -nlog\theta - \textstyle\sum_i^x x_i/\theta$$

$$\frac{\partial}{\partial\theta} logL(\theta) = -\frac{n}{\theta} + \frac{\sum_i^n x_i}{\theta^2}$$

Quoting to zero, we get $\theta = \bar{x}$ so that the $m.\ell.e$ of $\theta$is $\hat{\theta} = \bar{x}$

Example: Let $(x_1, .., x_n)$ be a, r, s from the exponential distribution

$$f(x,\theta) = e^{-(x-\theta)}, x \geqslant \theta$$

Then $$L(\theta, x_i, \dots x_n) = e^{-n(x-\theta)}$$

If we differentiate $logL(\theta)$ w, r, t θ and equate to zero we get $n = \theta$ which does not yield any result.

Now $L(\theta)$ is maximized by choosing the maximum value of θ subject to the condition

$$\theta \leq x_{(1)} \leq x_{(2)} \leq, \dots, \leq x_{(n)}$$

Which shows that $\theta = x_{(1)}$ so that the $m.\ell.e$ of $\hat{x} = X_1$

Example: X has $p.d.f$

$$f(x, \lambda, \theta) = \lambda_e{-}^{\lambda(x-\theta)}, x \geqslant \theta$$

$m.\ell.e\hat{\theta} = x_{(i)}$

$$\lambda = \frac{1}{\bar{x} - x_{(1)}}$$

Example: Let$(x_1, .., x_n)$ be a, r, s from the distribution

$$f(x,\theta) = \theta x^{\theta-1}, \qquad 0 \leq x \leq 1 (\theta > 1)$$

Then $$L(\theta, x_1, .., x_n) = \theta^n \left(\textstyle\prod_i^n x_i\right)^{\theta-1}$$

And $\log L(\theta) = nlog\,\theta + (\theta - 1)\sum_i^n logx_i$

Or $\frac{\partial}{\partial\theta} logL(\theta) = \frac{n}{\theta} + \sum_i^n logx_i$

Equating to zero we get $\theta = \frac{n}{-\sum \log x_i}$

$m.\ell.e$ of $\hat{\theta} = \frac{n}{\sum_i^n logx_i}$

Example: Let$(x_1, .., x_n)$ have rectangular distribution $R\ (0, \theta)$ having þ, d, f

$$f(x, \theta) = \frac{1}{\theta}, 0 \leq X \leq \theta$$

Then $L(\theta, x_1, .., x_n) = \frac{1}{\theta^n}, 0 \leq x_{(1)} \leq \cdots \leq x_{(n)} \leq \theta$

Which is maximized when θ is maximum subject to the condition

$$0 \leq x_{(1)} \leq \cdots \leq x_{(n)} \leq \theta$$

The minimum value of θ is $x_{(n)}$ so that

$m.\ell.e$ of θ is$\hat{\theta} = x_{(i)}$

Example: Let$(x_1, .., x_n)$ be a, r, s of the regular distribution $R(-\theta, \theta)$ having þ, d, f

$$f(x, \theta) = \frac{1}{2\theta}, 0 \leq X \leq \theta$$

Then $L(\theta, x_1, .., x_n) = \frac{1}{2^n\theta^n}, -\theta \leq x_{(1)} \leq \cdots \leq x_{(n)} \leq \theta$

When is maximized when θ is minimum subject to the condition $-\theta \leq x_{(1)} \leq \cdots \leq x_{(n)} \leq \theta$

So that since $-\theta \leq x_{(1)}$ or $\theta \geqslant -x_{(1)}$

$m.\ell.e$ of θ is $\hat{\theta} = -x_{(i)}$

Example: Let$(x_1, \ldots x_n)$ be a, r, s from the regular distribution $R(\theta_1, \theta_2)$ having þ, d, f

$$f(x,\theta_1,\theta_2)=\frac{1}{\theta_2-\theta_1},\theta_1\leqslant x\leqslant\theta_2$$

Then $\qquad L(\theta_1,\theta_2,x_ix_n)=\frac{1}{(\theta_2\theta 1)^n},\theta_1\leqslant x_{(i)}\leqslant\cdots x_{(n)}\leqslant\theta_2$

In maximized when $(\theta_2-\theta_1)$ is minimum $i,e\theta_1$ is maximum and θ_2 is minimum subject to the condition

$$\theta_1\leqslant x_{(i)}\leqslant.\leqslant x_{(n)}\leqslant\theta_2$$

We have to take $\theta_2=x_{(n)}$ and $\theta_1=x_{(i)}$ so that $m.\ell.e\ of\theta_1$and θ_2 are $\hat{\theta}_1=x_{(i)}$ and$\hat{\theta}_2=x_{(n)}$

Example: Let$(x_1,\ldots x_n)$ be a,r,s from the regular distribution $R(\theta-c,\theta+c)$ having þ, d,f

$$f(x,\theta)=\frac{1}{(2c)},\theta-c\leqslant x\leqslant\theta+c$$

Then $L(\theta,x_i,x_n)=\frac{1}{(2c)^n},\theta,c\leqslant x_{(i)}\leqslant.\leqslant x_{(n)}\leqslant\theta+c$is maxi zed for any θsuch that

$$\theta-c\leqslant x_{(i)}\leqslant\cdots\leqslant x_{(n)}\leqslant\theta+c$$

$i.e\ \theta-c\leqslant x_{(i)}$ or $\theta\leqslant x_{(i)}c$ and $\theta+c\geqslant x_{(n)}-c$

And $\theta+c\geqslant x_{(n)}$ is $\theta\geqslant x_{(n)}-c$

This shows that any statistics which lies in between $x_{(n)}-c$ and $x_{(i)}+c,e.g\frac{x_{(i)}+x_{(n)}}{2}$is $a,m.l.e$ the $m.l.e$ is not unique in this case

Example 12 It x has $R(\theta,\theta+I)$,any statistics which lies between $x_{(n)}-1$ and $x_{(i)}$ is a $m.l.e$ if θ

Example 13 Let$(x_1,\ldots x_n)$ be a,r,s from the regular distribution $R(\theta,2\theta)$ having þ, d,f

$$f(x,\theta)=\frac{1}{\theta}\theta\leqslant x\leqslant 2\theta$$

Then $\qquad L(\theta,x_i,..x_n)=\frac{1}{\theta^n},\theta\leqslant x_{(i)}\leqslant\cdots\leqslant x_{(n)}\leqslant 2\theta$

Is maxi zed when θ is minimum subject to the condition $\theta\leqslant x_{(i)}\leqslant\cdots\leqslant x_{(n)}\leqslant 2\theta$

$i.e \qquad \theta\leqslant x_{(1)}\ldots\ldots(i)$

And $\qquad \theta\geqslant x_{(n)}\ldots\ldots.(ii)$

Since $$\frac{x_{(n)}}{x_{(i)}} \leqslant \frac{2\theta}{\theta} = 2$$

$i.e$ $$\frac{x_{(n)}}{2} \leqslant x_{(i)}$$

The minimum value of θ satisfying $(i),(ii)$ is $\frac{x_{(n)}}{2}$ so that the $m.l.e$of θ is

$$\hat{\theta} = \frac{x_{(n)}}{2}$$

Example: Let$(x_i, .. x_n)$ be a, r, s from the regular distribution $R(-\theta, \theta)$ having þ, d, f

$$f(x, \theta) = \frac{1}{2\theta}, -\theta \leqslant x \leqslant \theta$$

Then

$$L(\theta, x_i, .. x_n) = \frac{1}{(2\theta)^n}, \theta \leqslant x_{(1)} \leqslant .. \leqslant x_{(n)} \leqslant \theta$$

This is maximized when θ is minimum subject to the condition

$$x_{(n)} \leqslant or\theta \geqslant x_{(n)}$$

And $$-\theta \leqslant x_{(i)} or\theta \geqslant -x_{(i)}$$

This happens when $\theta = \max(-x_{(i)}, x_{(n)})$

m, l, e of $\hat{\theta} = \max(-X_{(1)}, X_{(n)})$

Example: Let$(x_1, ... x_n)$ be a, r, s from the Laplace distribution with þ, d, f

$$f(x, \theta) = \frac{1}{2} e^{-[x-\theta]}, -\infty < x < \infty$$

Then

$$L(\theta, x_i, x_n) = \frac{1}{2} e^{-\sum_i^n [x-\theta]}$$

And $$logL(\theta) = -nlog2 - \sum_i^n [x_i - \theta]$$

Which is maximized when θ is the sample median.

m, l, e of θ is $\hat{\theta} = \hat{x}_{me}$

Example: Let$(x_1, \dots x_n)$ be n independent r, v, s such that x_r has normal distribution $N(r\theta, r^3\sigma^2)$

We have to estimate ***e***and then

$$L(\theta, x_i, x_n) = \prod_{r=i}^{n} \left[\frac{I}{\sqrt{2\pi r^3 \sigma^2}} e^{-\frac{I}{2r^3\sigma^2}(x_r - r\theta)^2}\right.$$

$$= \left(\frac{1}{2\pi\sigma}\right) \frac{I}{(n_i)2^e} - \frac{1}{2\sigma^2} \textstyle\sum_{r1}^{n} \frac{(x_r - r\theta)^2}{r^3}$$

And $$\log L(\theta) = nlog\left(\frac{1}{2\pi\sigma}\right) - \frac{3}{2} logn1 - \frac{1}{2\sigma^2} \sum \frac{(x_\sigma - r\theta)^2}{r^3}$$

Or $$\frac{\partial}{\partial\theta} logL(\theta) = \frac{1}{2\sigma^2} \textstyle\sum_{r=1}^{n} \frac{(x_r - r_\theta)}{r^2}$$

Equating to zero , we get

$$\sum_{i}^{n} \left[\frac{(x_r - r_\theta)}{r^2}\right] = \theta$$

Or $$\theta = \frac{\sum_i^n x_r/r^2}{\sum_i^n i/r^2}$$

$m. l. e$ Of θ is

$$\hat{\theta} = \frac{\sum_i^n x_r/r^2}{\sum_i^n i/r^2}$$

We have $$E(\hat{\theta}) = \theta, V(\hat{\theta}) = \frac{\sigma^2}{\sum_I^n ({}^I/_r)}$$

<u>**Optimum properties of MLE:**</u> (i) If $\hat{\theta}$ is $m. l. e$ of θ and Ψ (θ) is a simple valued function of θ with unique inverse, then $\Psi(\hat{\theta})$ is the $m. l. e$ of Ψ (θ).

(ii) If a sufficient statistics exists for θ $m. l. e$ $\hat{\theta}$ is a function of this sufficient statistics.

(iii) Suppose $f(x, \theta)$ statistics certain regularity conditions and $\hat{\theta}_n = \hat{\theta}_n(x_1, \dots x_n)$ is the $m. l. e$ of a random sample of size n from $f(x, \theta)$

Then- (a) $\{\hat{\theta}_n\}$ is consistent sequence of estimators of θ

(b) $\hat{\theta}_n$ is asymptotically normally distributed with mean θ variance $\frac{1}{nE\left[\frac{\partial}{\partial\theta}logf(x,\theta)\right]^2}$

(c)The sequence of estimators $\hat{\theta}_n$ has the smallest asymptotic variance among all consistent, asymptotically normally distributed estimate of $\theta, i.e\ \hat{\theta}_n$ is BAN or CANE or most efficient.

(iii) **METHOD OF MINIMUM χ2**: Let X be $a.r.v$ with $þ.d.f\ f(x,\theta)$ where parameter to be estimated $\theta = (\theta_1, \ldots\ldots \theta_r)$ Suppose $S_1, S_2....S_k$ are k mutually exclusive classes which from a partition of the range of X. Let the profanity at X falls in S_J be

$$þ,(\theta) = \int_{S_j} f(x,\theta)dx\ , j = 1,2, \ldots k$$

Where $\sum_{j=1}^{k} þ,(\theta) = 1$

Suppose ,in practice ,corresponding to a random sample of n observations from the distribution of X we are given the frequencies $(N_1, \ldots\ldots, N_k)$ where N_j=observed number of sample observations falling in the class $S_J(j = 1,2 \ldots\ldots k)$ such that $\sum_i^k N_j = n$ then the expected number of observation in S_J is $nþ_j(\theta)$, Define

$$\chi 2 = \sum_{j=1}^{k} [n_j - nþ_j(\theta)]^2 / nþ_j(\theta)$$

Where n_j is the observed value of $N_j(j = 1,2 \ldots . k)$ Evidently x^2 will be a function of $\theta(or\ \theta_i, .. \theta_r)$to obtain the estimator of θwe minimise $x^2 w.r,t\ \theta$. The minimise x^2 estimator of θ is that $\hat{\theta}$ which minimise above χ2.

The equation (s) for determining the estimator(s) by this me that are

$$\frac{\partial\chi 2}{\partial\theta} = \theta\ or \frac{\partial\chi 2}{\partial\theta} = 0\ \ (i = 1, \ldots \ldots r)$$

Remarks:

(i)Often it is difficult to obtain $\hat{\theta}$ which minimum χ2, hence χ2 is changed to modified

$$\chi 2 = \sum_{j=1}^{k} [n_j - nþ_j(\theta)]^2\, n_j$$

(If $n_j = 0$, unity is used). The modified minimum χ2 estimator of θ is $\hat{\theta}$ which minimises the modified χ2

(ii) For large n, the minimum χ2 and likelihood equations are identical and, consequently, provide identical minimum χ2maximum likelihood estimators.

(iii) The minimum χ2 estimators are consistent asymptotically normal and efficient .

Example: Let$(x_1, .. x_n)$ be a, r, s from a Bernoulli distribution having $þ. d. f$

$$f(x, \theta) = \theta^x (1-\theta)^{1-x}, x = 0,1$$

Take N_j =the number of observations equal to j for j= 0,1 Here the range of X is pinioned into the two sets consisting of the minimises θ,and i respectively then

$$\left.\begin{array}{c} þ_0(\theta) = P(x=0) = 1-\theta \\ þ_1(\theta) P(x=1) = \theta \end{array}\right\}$$

And

$$\chi 2 = \sum_{j=o}^{i} \frac{[n_j - nþ_j(\theta)]^2}{nþ_j(\theta)}$$

$$= \frac{[n_\theta - n(1-\theta)]^2}{n(1-\theta)} + \frac{[n_i - n\theta]^2}{n\theta}$$

$$= \frac{[n - n_i - n(1-\theta)]^2}{n(1-\theta)} + \frac{[n_i - n_\theta]^2}{n\theta}$$

$$= \frac{[n_i - n_\theta]^2}{n} \frac{1}{\theta(1-\theta)}$$

By inspection $\chi 2 = \theta$ for $\theta = n_i/n$ Therefore $\hat{\theta} = n_i/n$. This is a same as what would be obtained by the method of moments or method of maximum likelihood

(IV) <u>METHOD OF LEAST SQUARES</u> Suppose γ is e random variable whose value depends on the value of a (non-random) variable$x$. For example the weight of a baby (Y) depends on its age(x) , the temperature (Y) of a place at a given time depends on its altitude (x), or the salary (Y) of an individual at a given age depends on the number of years (x) of formal education which he has had the maintenance cost (y) per year of an automobile depends on its age (x) etc.

We assume that the distribution of the $r.v\ Y$ is such that for a given $x, E(\mathrm{Y}/\mathrm{x})$ is a linear function of x while the variance and higher moments of γ are independent of x. It means that we assume the liner model

$$E(\mathrm{Y}/x)=\alpha+\beta x$$

Where d and β and two parameters .We also write

$$\mathrm{Y}=\alpha+\beta x+\epsilon$$

Where ϵ is a, r, u such that $\mathrm{E}\ (\epsilon)=\theta, V(\theta)=\sigma^2$

The problem is to estimate the parameters d and β on the basic of a random sample of n observations$(\gamma, \mathrm{x}_\mathrm{i}), (\gamma_2, \mathrm{x}_2), \dots \dots, (\gamma_\mathrm{n}, \mathrm{x}_\mathrm{n})$

The method of least squares estimations of α and β specifies that we should take as our estimates of d and β those values that minimise

$$\sum_{i=1}^{n}[y_i-\alpha-\mathrm{x}_\mathrm{i}]^2$$

Where y_i is the observed value of y_i and x_i are the associated values of x. This we minimise the sum of squares of the residuals when applying the method of least squares.

The least squares estimators of α and β

Are $$\hat{\beta}=\frac{\sum_{i=1}^{n}(y_i-\bar{y})(x_i-\bar{x})}{\sum_{i=1}^{n}(x_i-\bar{x})^2}$$

And $$\hat{\alpha}=\bar{y}-\hat{\beta}\bar{x}$$

<u>**Remarks:**</u>

The least square estimator do not have any optimum properties ever asymptotically However in linear estimation this method provides good estimation in small simples. These estimators are minimum variance unbiased estimators among the class of linear function of Y′s.

TESTING OF HYPOTHESIS

(NEYMAN PEARSON THEORY)

Let x be a, r, u with þ.$d.f$ $f(x, \theta)$ where θ (may be a vector ($\theta_1, \dots \dots . \theta_k$) is an unknown parameter. A random sample of n observation denoted by $E = (x_1 \dots \dots x_k)$ which takes values in general, in the n-dimensional real space R_n the parameter space (all possible values of the parameter is denoted by Ω, say . For any subset AC R_n we can calculate.

$$p_o(E\ E\ A) = \int_A \left[\prod_{i=1}^{n} f(x, \theta) \right] dx, \dots d_{xn}$$

Which will depend on θ.

Definition: A statistical hypothesis is a statement about the parameter θin the form $H: \theta \in \omega(< \Omega)$

For example consider

$$H: \theta = \theta_O$$

$$or\ H: \theta \geqslant \theta_0\ or\ H: \theta \neq \theta_0\ H: \theta_1 < \theta < \theta_2$$

Definition If a hypotheses specifics an exact value of the parameterθ, it is called a simple hypothesis e, g . $H: \theta = \theta_0$ in this case ω in $H : \theta \in \omega$ is a set of a single point

If a hypothesis does not fully specify the value of θ(but gives a set of possible values only) it is called a composite hypothesis e, g $H: \theta \neq \theta_0$ or $H: \theta \geqslant \theta_0$ etc. In this case ω in $H: \theta \in \omega$ is set of more than one point.

Definition the hypothesis which is being actually tested is called the null hypothesis and other hypothesis which is stated as the alternative to the null hypothesis is called alternative hypothesis. For example, null hypothesis may be $H_o: \theta = \theta_0$ and the alternative may be $H_i: \theta \neq \theta_0$or $H_i: \theta > \theta_0$or $H_i: \theta \leq \theta_0$ etc.

Both null and alternative hypothesis may be simple or composite .For our study, we shall usually take null hypothesis to be simple .

Suppose we want to test a null hypothesis H_O against an alternative hypothesis H_1on the basis of a random sample $E = (X_1, . . X_n)$ in the sense that we have to decide when to reject or accept H_O

<u>Definition</u> A Statistical test of a (null) hypothesis H_O against an alternative hypothesis H_1 is a rule or procedure for deciding when to reject or accept H_O on the basis of the sample $E = (X_I, .. X_n)$ .It specifies a position of the sample space R_n into two disjoint subsets W and $\bar{W} = R_n - W$ such that we reject H_O when $E \epsilon W$and accept H_O when $E \epsilon \overline{W}$ [We note that the rejection of H_O amounts to acceptance of H_1 and vice-versa]

<u>Definition</u> The set W, corresponding to a test T, which is that we reject H_O when $E \epsilon W$ is called the <u>critical region</u> of the test while $\overline{W}$ is called its acceptance region. For different test the critical regions are different.

<u>Two types of errors:</u> In a testing problem we are liable to commit two types of error. Suppose $H_O$is true and get $E \epsilon \overline{W}$ so that we reject H_O this is called the <u>Type I error</u> which occasion when we reject the null hypothesis when it is actually true. On the other hand, suppose H_O is false and $H_i$is true and yet $x \epsilon \bar{w}$ so that we accept H_O this is called the <u>types II errors</u> which occurs when we accept the null hypothesis when it is actually false. We denote by α and β the probability of type I error and type II error, respectively, i, e

$$\alpha = P\{H_o / H_o \text{is true}\}$$

$$= P\{E \epsilon W / \theta \ \epsilon H_o\}$$

And $$\beta = P\{AccepH_o / H_o \text{ is fabe}$$

$$= P\{E \epsilon \bar{W} / \theta \ \epsilon H_o\}$$

<u>Definition</u> The probability of type I error for a test T, denoted by ∝ is called the "<u>size</u>" or <u>level of significance</u> of the test T.

<u>Remark</u> If H_o is simple (say H_o:θ=θ_o) is clearly defined ,when H_o is composite (say H_o: $\theta \epsilon W$)we take

$$\alpha = supP_T\{E \ \epsilon W /: \theta\} \ \theta \epsilon H_o$$

<u>Definition</u> For a test T having the co region w_2 the <u>power function</u> $P_T()$ is defind by

$$P_T(\theta) = P\{RejectH_o / \ \theta\}$$

$$= \mathrm{P}_\theta\{\mathrm{E} \ \epsilon \ W\}$$

As a function of θ

Evidently,

$$P_T(\theta) = \alpha \ \ for\ \theta\ \epsilon\ H_o$$

$$P_T(\theta) = 1 - \beta\ \theta\ \ \epsilon\ H_1$$

If we would find a test of the given hypothesis for which both ∝and $\beta$are minimum it would be the best. Unfortunately, it is not possible to minimise both error simultaneously for a fixed sample size test. Consists two tests T and T_2 defined as follows

T_I always rejects H_0 i,e its critical region $W_1 = R_n$, while T_2 always accepts H_o i,e its cur region $W_2 = \emptyset$ then for T_1, $\propto = 0$ and $\beta = 1$ this shows that if the probability of type I error becomes minimum than the probability of type II error becomes maximum and vice-versa what is done is to fix ∝, taking ∝ to be quite small (in practical ∝= .05 or .01)so that all test of size ∝ are only considered. Among all test of a given size∝ comparison made on the basic of their power function. If T and T are two tests (for the same testing problem) of same size ∝, T is said to be better than T if its power is greater than the power of T for alteration hypothesis (equivalently the probability of type II error for T is less then the probability of type II for T,)

<u>Simple hypothesis against a simple alternation</u>: Consider the testing problem

$$H_O: \theta = \theta_O$$

$$H_1: \theta = \theta_1 (\neq \theta_O)$$

<u>Definition</u> A test T* is called a most powerful test (MP) of size ∝ (0 <∝< 1) if only if the probability of type I error is equal to ∝ and its power $P_T(\theta)$ is not less than the power $P_T(\theta)$ of all other test T of size ∝, i, e

$$(i) P_{T*}(\theta_o) = \propto$$

$$(ii) P_{T*}(\theta_i) = P_T(\theta_i)$$

For any other test T of size ∝

[This means that the probability of type II error for T is less that of IV any other test]

<u>Simple hypothesis against a composite alteration</u>: Consider the testing problem

$$H_o: \theta = \theta_0$$

$$H_i: \theta \neq \theta_0 (or\theta > \theta_o or\theta < \theta_0)$$

<u>**Definition**</u>: A test T is called a uniformly most powerful test (VMP) of size $\propto(0 < \propto < 1)$if its probability type I error is equal to $\propto$ and its power function is such that

$P_{Tx}(\theta) \geqslant P_T(\theta)$for all $\theta \epsilon H_I$and all other test T of size $\propto$

Example Let x be a, r, u having exponential distribution

$$f(x,\theta) = \theta\, e^{-\theta x}(x \geqslant \theta)$$

And we want to test

Against $\left.\begin{matrix} H_0: \theta = 2 \\ H_1: \theta = 1 \end{matrix}\right\}$

Let the sample consist of only one observation X and consider two tests T and T with associated regions $W = \{X \geqslant I\}$ and $W = \{X \leq 0.7\}$ respectively

The probabilities of two error for T are

$$\propto = P\{X \geqslant 1/\theta = 2\} = 2\int_1^{\infty} e^{-2x} dx = 0.135$$

$$\beta = P\{X \geqslant 1/\theta = 1\} = \int_0^{1} e^{-x} dx = 0.635$$

The probabilities of two error for T are

$$\propto = P\{X \geqslant 0.7/\theta = 2\} = 2\int_0^{7} e^{-2x} dx = 0.135$$

$$\beta = P\{X \geqslant 0.7/\theta = 1\} = \int_7^{\infty} e^{-2x} dx = 0.932$$

Obviously T is better than T'.

Example A Two –faced coin is tossed six times for which the probability of getting head in a toss is θ and the probability of getting a tail is $(1-\theta)$. it is required to test the hypothesis.

$$H_0: \theta = \theta_o = {}^1\!/_2$$

Against $H_1: \theta = \theta_o = {}^2/_3$

If the test consists in rejecting H_0 when head appeases nurse then tours times and accepting H_0 otherwise find d, β soln

$$\propto = P\{= \theta_o\} = {}^7/_{2^6}$$

$$\beta = 1 - P\{RejH_0/\theta = \theta_1\} = 1 - {}^{2^8}/_{3^6}$$

Example Let x have an exponential distribution

$$f(x, \theta) = \frac{1}{\theta} e^{-\frac{x}{\theta}}, x \geqslant \theta$$

It is required to test $H_O: \theta = 1$

$$H_1: \theta = 4$$

Find $\propto$ and β for the test having region $C = \{X > 3\}$on the basic of a sample observation

Soln : We have $\propto = P\{RejH_O/\theta{=}\theta_O\}$

$$= P\{X > 3/ \theta = 1\}$$

$$= \int_3^\infty e^{-x} dx = 3^e$$

$$\beta = P\{sce. H_0/\theta = \theta_0$$

$$= 1 - P\{X > 3/\theta{=}4\}$$

$$=1 - \tfrac{1}{4}\int_3^\infty e^{-x/4} dx$$

$$= 1 - e^{-3/4}$$

Power $= 1 - \beta = e^{-3/4}$

Example Let x have the rectangular distribution

$$f(x, o) = \frac{1}{\theta}, o \leqslant x \leqslant \theta$$

It is required to test the hypothesis

$$H_0 : \theta = 1$$

Against $$H_1 : \theta = 2$$

Suppose one observation is taken and the tests having the critical regions (a) $C_1 = \{x \geqslant .7\}$ and (b) $C_2 = \{.8 \leqslant x \leqslant 1.3\}$ obtain the profanities of two types error $\propto$ and β

Soln : (a) $$C_1 = \{x \geqslant .7\}$$

$$\propto = P\{RajH_O/\theta{=}\theta_O\}$$

$$P[X \geqslant .7/\theta{=}1]$$

$$= \int_{7}^{1} i \, dx = .3$$

$$\beta = P\{\sec H_0/\theta{=}\theta_i\}$$

$$= \int_{0}^{.7} \frac{1}{2} dx$$

$$= 35$$

(b) $$C_2 = \{.8 \leqslant x \leqslant 1.3\}$$

$$\propto = P\{.8 \leqslant x \leqslant 1.3/\theta = 1\}$$

$$= \int_{.8}^{1} 1. dx = 2$$

$$1 - \beta = P\{.8 \leqslant x \leqslant 1.3/\ \theta{=}2\}$$

$$= \int_{.8}^{1.3} \frac{1}{2} dx = .25$$

Or $$\beta = .75$$

Example Let x have a Binomial distribution $B(10, \text{þ})$ for which

$$f(x, \text{þ}) = \binom{10}{x} \text{þ}^x (1 - \text{þ})^{10-x}, x = 0,1, \dots .10$$

One observation x is taken for testing H_0 : $\text{þ} = {}^1\!/_2$ against $H_1 : \text{þ} = {}^1\!/_4$. Find $\propto$ *and* β for the test which rejects H_0 when x≤3.

Soln $\propto = P\{x \leq 3/þ = {}^{1}/_{2}\}$

$$= \sum_{x=0}^{3} \binom{10}{x} \left(\frac{1}{2}\right)^{x} \left(\frac{1}{2}\right)^{10-x}$$

$$= \frac{11}{64}$$

$$\beta = 1 - P\{x \leq 3\ þ = {}^{1}/_{4}\}$$

$$1 - \sum_{x=0}^{3} \binom{10}{x} \left(\frac{1}{4}\right)^{x} \left(\frac{3}{4}\right)^{10-x}$$

$$= 1 - 31.\frac{3^8}{4^9}$$

Example Let x have a Poison distribution $P(\lambda)$ and it is required to test the hypothesis $H_O: \lambda = 1$vs $H_i: \lambda = 2$. One observation is taken and a test is considered which reject H_O when X⩾3 . Find $\propto, \beta$

Soln: we have $\propto = P(X \geq 3/\lambda = 1\}$

$$= 1 - \textstyle\sum_{X=0}^{2} \frac{e^{-1}}{x!}$$

$$= 1 - \left[\frac{1}{e} + \frac{1}{e} + \frac{1}{2_e}\right] = 1 - \frac{5}{2e}$$

$$\beta = P(X \geq 3/\lambda = 2\}$$

$$= \sum_{x=0}^{2} \frac{e^{-2} 2^x}{x!}$$

$$= \frac{1}{e^2}[1 + 2 + 2]$$

$$= \frac{5}{e^2}$$

Now we are in a positions to power a the over which helps us to obtain MP tests of a sample hypothesis against a simple alternative. In some special situations, this also gives a UMP test when the alternation is composite.

Let us suppose that we are testing a simple hypothesis against a simple alternative

$$H_o : \theta = \theta_o$$

Us $$H_1 : \theta = \theta_1 (\neq \theta_o)$$

Theorem (Neyman- Pearson Lemma)

let the like hood of the sample E=$(X_1, \dots, X_n)$ under H_o and $H_i$be

$$L(\theta_j) = L(\theta_j, X_1, \dots, X_n)$$

$$= \prod_{L=1}^{n} f(X_i, \theta_j), j = 0,1$$

Let T be a test of size∝,for which the cr. region W is defined by

$$W = \left\{ E / \frac{L(\theta_I)}{L(\theta_0)} \geqslant e \right\}$$

Where e is a constant determined by the size condition

$$P\{E \,\epsilon W / \theta_0\} = \propto$$

Then T is a MP of size ∝for testing H_0 against H_I

Prof Let us write

$$L_0 = L(\theta_o) and\ L_i = L(\theta_i)$$

So that the size and power of any test T with Cr. Regain W are follows:

Size of $$T = \int_W L_0\, dx\ and\ power\ of\ T = \int_w L_i\, dx$$

Where $$dx = d_{x_1} d_{x_2} \dots \dots d_{x_n}$$

Consider the test T (having cr. Region w) and other test T (having is Region since both W are of size ∝ we have w)

$$\int_W L_o dx = \propto = \int_w L_o dx - (1)$$

W W1 W2 W3

Let
$$W_1 = W - W \cap W$$
$$W_2 = W \cap W$$
$$W_3 = W - W$$

We have using (i),

$$\int_{w_1} L_O dx = \int_W L_O dx - \int_{W_2} L_o dx$$

$$= \int_w L_O dx - \int_{w_2} L_O dx = \int_{W_3} L_0 dx - (ii)$$

Sine $W_1 \subset W^*$ and $W_3 \not\subset W^*$ we have, by definition of w^γ and using (i)

$$\int_{w_i} L_i dx \geqslant c \int_{w_i} l_o dx - (ii)$$

And $$\int_{w_3} L_i dx < c \int_{w_3} L_o dx = c \int_{w_1} L_o dx \qquad - (iii)$$

Therefore ,from (ii) $(iii) we get

$$\int_{w_1} L_i dx \geqslant \int_{w_3} l_i dx \qquad - (iv)$$

Adding $\int_{W_2} L_I dx$ on both sides of (iv) we get

$$\int_{w_1 \cup w_2} L_{\,i} dx \geqslant \int_{w_3 \cup w_2} L_O dx$$

Or $$\int_w L_{\,i} dx \int_w L_O dx$$

Or $P_r(Rej H_o/\theta{=}\theta_i)$

Or $P_r(\theta_i) \geqslant P_r(\theta_i)$

Which shows that T is more powerful then T any other test of size $\propto$ Hence T is the MP test

<u>Remarks</u> (1) The constant C for the MP test is determined by using the size condition

$$\int_w L_o dx = \propto$$

Usually, a unique value of C is obtained when the $r.v$ has a continuous distribution.

(2) When X is a discrete $r.v.$ the constant C may not be unique. What is more important is that we may not be able to find a MP critical region with expect size ∝. To get rid of the difficultly the cr. Region is defined by the following

$$\begin{cases} Rej\ H_O\ if\ \dfrac{L(\theta_1)}{L(\theta_2)} > c \\ Rej\ H_O with\ probablity\ r\ if\ \dfrac{L(\theta_1)}{L(\theta_2)} = c \\ Acc\ H_O\ if\ \dfrac{L(\theta_1)}{L(\theta_2)} > c \end{cases}$$

Then the size of test is $P_o\left\{\frac{L(\theta_1)}{L(\theta_o)} > c\right\} + rP_o\left\{\frac{L(\theta_1)}{L(\theta_o)} = c\right\} = \propto$

To any given $\propto, r$ can be determined. Such a test is called the a randomized test

Example Let $(x_1, .. x_5)$ be a random sample from H_0 Bernoulli .distribution

$$f(x,\theta) = \theta^x(1-\theta)^{1-x}, x = 0,1 (0 < \theta < 1)$$

Let us test $H_O\ \theta = 6$ us $H_1: \theta = \theta_1 (> .6)$. The MP test has cr. Region $\{\sum_1^5 x_i \geqslant c\}$

Now $\sum_1^5 x_i$ has Bernoulli. distribution B(5,θ)

From the tables of Bernoulli Distribution we can to tabulate $P_o\{\sum_1^5 x_i \geqslant c / \theta = .6\}$ us follows

C	P($\sum_1^5 x_i \geqslant c$)	P_0
1	0.01024	1.00000
2	0.23040	0.98976
3	0.34560	0.68256
4	0.25420	0.33696
5	0.07776	0.07776

As such, no non-randomized MP test of exact size∝ .05 or 01 exists. However, the randomized MP test of size .35 is given by

$$\begin{cases} Raj & H_O\ if \sum_{1=i}^{5} x_i \quad > \qquad\qquad 3 \\ Raj & H_O\ with\ probability \quad \frac{.01304}{.34560}\ if \sum_{1=i}^{5} x_i = 3 \\ Ace & H_O\ \ if \sum_{1=i}^{5} x_i = 3 \end{cases}$$

(3) Suppose we test the simple hypothesis $H_O: \theta > \theta_o$against a composite alternation $H_i: \theta \neq \theta_o$ or $H_i: \theta > \theta_o$ or $H_i: \theta < \theta_o$ if the MP test for $H_O: \theta = \theta_o$ a gains $H_i: \theta = \theta_i$given by the NP lemma dose not depend on θ_i,the same test with be MP for all alternative values of θ and, therefore it will be a.UMP test.

Example (1) Let x have a Poisson distribution $P(\lambda)$having þ. m . f

$$f(x, \lambda) = \frac{e^{-\lambda} \lambda^x}{x^i}, x = o, 1,2$$

We want to test $H_i : \lambda = \lambda_0$

Against $H_i : \lambda = \lambda_1$

We have

$$L(o) = \prod_{i}^{n} f(x, \lambda) = e^{-n\lambda} \lambda^{\sum_i^n x_i} / \prod_{i}^{n} x_i$$

Therefore, the MP test has the cr region W given by

$$W = \left\{\frac{L(\lambda_1)}{L(\lambda_0)} \geqslant C\right\}, i, e. \text{ inside W we have}$$

$$\frac{L(\lambda_1)}{L(\lambda_0)} = e^{-n(\lambda_1 - \lambda_0)} \left(\frac{\lambda_1}{\lambda_0}\right) \sum_{i}^{n} xi \quad \geqslant C$$

Or $$-n(\lambda_1 - \lambda_0) + (\sum x_i) log\left(\frac{\lambda_1}{\lambda_0}\right) \geqslant C$$

Or $$\sum_1^n x_i \geqslant k$$

Where $$k = \frac{c + n(\lambda_1 - \lambda_0)}{\log(\lambda_1 / \lambda_0)}$$

$$w = \left\{\sum_{i}^{n} x_i \geqslant k\right\}$$

We know that $\sum_i^n x_i$ has Poisson distribution $P(n\lambda)$ so that k can be determined by solving

$$P(\sum_i^n x_i \geqslant k\ /\lambda=\lambda_0) = \propto$$

__Remarks:__ (i) When $\lambda_1 < \lambda_0$ the MP test will be given by $\{\sum_i^n x_i \leqslant k\}$

(ii) Sine the cr region does not depend on the value of λ_1 there are UMP for the alternative $H_i: \lambda > \lambda_0$ as $H_i: \lambda < \lambda_0$, respectively.

(iii)For getting a MP test for an exact size $\propto$ we may have to use randomized test

(2) Let X have an exponential distribution

$$f(x,o) = \theta e^{-\theta x} \ (x \geqslant o)$$

We want to test $$H_o: \theta = \theta_o$$

Us $$H_i: \theta = \theta_i (< \theta_o)$$

We have $$L(\theta) = \prod_i^n f(x,o) = \theta_e^{n-\theta \sum_i^n (x_i-\mu)^2}$$

Therefore, the MP test has the critical region W defined by

$$W = \left\{\frac{L(\theta_i)}{L(\theta_0)} \geqslant c\right\}$$

i, e Inside W $$\frac{L(\mu_i)}{L(\mu_O)} = \frac{e^{-\frac{1}{2\sigma^2}\sum_i^n (x_i-\mu_i)^2}}{e^{-\frac{1}{2\sigma^2}\sum_i^n (x_i-\mu_o)^2}} \geqslant c$$

Or $$e^{-\frac{1}{2\sigma^2}}[\sum_i^n (x_i-\mu_i)^2 - \sum_i^n (x_i-\mu_o)^2] \geqslant c$$

Or $$[\sum_i^n (x_i-\mu_o)^2 - \sum_i^n (x_i-\mu_i)^2] \geqslant 2\sigma^2 \log c$$

Or $$2(\mu_i-\mu_o)\sum_i^n x_i \geqslant 2\sigma^2 \log c + (\mu_1^2 - \mu_0^2)n$$

Or $$\frac{i}{n}\sum_i^n x_i \geqslant \frac{\sigma^2 logc}{n(\mu_i-\mu_o)} + \frac{\mu_i+\mu_o}{2}\begin{pmatrix} since \\ H, 7\mu_o \end{pmatrix}$$

Or $$\bar{x} \geqslant k$$

Whose $k = r, h, s$

: MP test is given by W=$\{\bar{x} \geqslant k\}$ Since $\bar{x} \geqslant N(\mu, \sigma/\sqrt{n})$ we can determine

GRAPH HERE

$$P[Z \geqslant k_\propto] = \propto$$

$k_\propto$ Is called the upper ∝ % point of N (0,1)

$k_\propto$ Is called the lower ∝ % point of N (0,1)

k by solving

$$P_{\mu o}\{\bar{x} \geqslant k\} = \propto$$

Or $$P_{\mu o}\left\{\frac{\bar{x} \geqslant \mu_o}{\sigma\sqrt{n}} \geqslant \frac{k-\mu o}{\sigma\sqrt{n}}\right\} = \propto$$

Or $$T_{\mu_o}\left\{z \geqslant \frac{k-\mu o}{\sigma\sqrt{n}}\right\} = \propto$$

Under H_o, z has N (o,1)and the tables of standard normal distribution provider the value of $k_\propto$such that $k_\propto = \frac{k-\mu o}{\sigma\sqrt{n}}$ or $k + \mu o + k_\propto \frac{\sigma}{\sqrt{n}}$

<u>Remark</u> (1) the power of the MP test given above is

$$P_{\mu i}\{\bar{x} \geqslant k\}$$

$$P_{\mu i}\left\{\frac{\bar{x} - \mu_i}{\sigma\sqrt{n}} \geqslant \frac{k - \mu_i}{\sigma\sqrt{n}}\right\}$$

$$P_{\mu i}\left\{z \geqslant \frac{\sqrt{n}(\mu_o - \mu_i)}{\sigma} + k_\propto\right\}$$

Since $(\mu_o - \mu_i) < o$,it shows that the power is an impressing function of n

(ii) If $\mu i < \mu_o$ the MP test can be shown to have the critical region $\{\bar{x} \geqslant k\}$ where $k = \mu_o + k_\propto \frac{\sigma}{\sqrt{n}}$ such that $P\{Z \leqslant k_\propto\} = \propto$ for a standard normal $r, v(in\ pact\ k_\propto = -k_\propto)$

(iii) We observe that the MP test of $H_o: \mu = \mu_o$ us $H_I: \mu = \mu_i (> \mu_o)$ has a cr region which dose not depend on $\mu_i$the same test will be UMP for testing $H_o: \mu = \mu_o$ against $H_I: \mu > \mu_o$ Similarly the MP test $H_o: \mu = \mu_o$ against $H_I: \mu = \mu_i (> \mu_o)$ is UMP for testing $H_o: \mu = \mu_o$ against$H_I: \mu < \mu_o$

However it can be shown that there is no test which is UMP for$H_o: \mu = \mu_o$ against$H_I: \mu \neq \mu_o$

(4) Let X have a normal distribution $N(\mu, \sigma)$ where μ is a known constant

We want to test

$$H_o : \sigma = \sigma_o$$

Us $$H_i : \sigma = \sigma_i (> \sigma_o)$$

We have

$$L(\sigma) = \frac{1}{(2\pi)^{n/2}\sigma^n} e^{-\frac{1}{2\sigma^2}\sum_i^n (x_i-\mu)^2}$$

Therefore the MP test has the cr region w depend by $W = \left\{\frac{L(\sigma_i)}{L(\sigma_o)} \geqslant c\right\}$

i, e inside W

$$\frac{L(\sigma_i)}{L(\sigma_o)} = \left(\frac{\sigma_o}{\sigma_i}\right)^n e^{-\sum_i^n (x_i-\mu)^2}\left(\frac{1}{2_1^{\sigma 2}} - \frac{1}{2_0^{\sigma 2}}\right) \geqslant c$$

Or $$\sum_i^n (x_i - \mu)^2 \left(\frac{1}{2_1^{\sigma 2}} - \frac{1}{2_0^{\sigma 2}}\right) \geqslant \log c \left(\frac{\sigma_i}{\sigma_o}\right)^n$$

Or $$\sum (x_i - \mu)^2 \geqslant k (since\ \sigma_1 > \sigma_o)$$

Where $$k = \frac{2\{\log e + n log\left(\frac{\sigma_i}{\sigma_o}\right)]}{\left(\frac{1}{\sigma_0^2} - \frac{1}{\sigma_1^2}\right)}$$

MP test cr region is given by $$W = \{\sum_i^n (x_i - \mu)^2 \geqslant k\}$$

Since $\sum_i^n \frac{(x_i-\mu)^2}{\sigma^2} \sim x_n^2$ we can determine $k$by solving

$$P_{\sigma_o}\left\{\sum_i^n (x_i - \mu)^2 \geqslant k\right\} = \propto$$

Or $$P_{\sigma_o}\left\{\sum_i^n \frac{(x_i-\mu)^2}{\sigma^2} \geqslant \frac{k}{\sigma_0^2}\right\} = \propto$$

Or $$P_{\sigma_o}\left\{\gamma \geqslant \frac{k}{\sigma_0^2}\right\} = \propto$$

Where $Y \sim x_n^2$

From the table of x_n^2 we can find $k_\propto$such that $P\{\gamma \geqslant k_\propto\} = \propto$ so that $k = \sigma_0^2 k_\propto$

<u>**Remark**</u> (i) the power if the test is given by

$$P_{\sigma_o}\left\{\sum_{i}^{n}(x_i-\mu)^2 \geqslant k\right\}$$

$$= P_{\sigma_1}\left\{\frac{\sum(x_i-\mu)^2}{\sigma_1^2} \geqslant \frac{k}{\sigma_1^2}\right\}$$

$$= P_{\sigma_1}\left\{Y \geqslant \frac{\sigma_0^2}{\sigma_1^2}k_{\propto}\right\}$$

Where $Y \sim x_n^2$

(ii)If $\sigma_1 < \sigma_0$the MP test can be shown to have the cr region $\{\sum_i^n (x_i-\mu)^2 \leq k'\}$

(iii)Since the MP test of $H_0\colon \sigma = \sigma_o$us $H_i\colon \sigma = \sigma_1(> \sigma_o)$ dose not depend on $\sigma_i$it is UMP for testing $H_0\colon \sigma = \sigma_o$ against $H_i\colon \sigma > \sigma_o$ Similarity the MP test for $H_0\colon \sigma = \sigma_o$ against $H_i\colon \sigma > \sigma_1(> \sigma_o)$is UMP test for $H_0\colon \sigma = \sigma_o$ against$H_i\colon \sigma < \sigma_o$

However, no UMP test exists for alternative $H_1\colon \sigma \neq \sigma_o$

(5) Let X have the distribution with þ, d, f

$$f(x,o) = \theta x^{\theta-1}(0 \leq x \leq 1)$$

We want to test

$$H_0 : \theta = \theta_0$$

Against $H_i : \theta = \theta_1(> \theta_0)$

We have $1(\theta) = \theta^n[\prod x_i]^{\theta-1}$

Therefore, the MP has the cr region W=$\left\{\frac{L(\theta_i)}{L(\theta_0)} \geqslant C\right\} i, e \text{ inside } W$

$$\left(\frac{\theta_i}{\theta_0}\right)^n \left[\prod_{i=1}^{n} x_i\right]^{\theta_i-\theta_o} \geqslant c$$

Or$\prod_{i=1}^{n} x_i \geqslant k$where $k = \left[c\left(\frac{\theta_o}{\theta_i}\right)^n\right]^{1}/\theta_i - \theta_0)$

The MP test has cr region

$$\left\{\prod_{i=i}^{n} x_i \geqslant k\right\}$$

Or $\{-\sum_{i=i}^{n} logx_i \leq k_o\}$ where $k_o = -\log k$

If can be shown that $\gamma = (20)(\sum_{i=i}^{n} logx_i)$ has x_{2n}^2 therefore the constant k_o(and have k) can. Be determined by solving

$$P\{\gamma \leqslant (2\theta_O)k_O\} = \propto$$

Where $\gamma \sim x_{2n}^2$

Remark In the same manner for $H_o : \theta = \theta_O$ against $H_i: \theta = \theta_1 (< \theta_O)$ MP test can be found.

$$\left\{x \frac{f,(x)}{f_o,(x)} \geqslant c\right\}$$

Or $$\sqrt{\frac{2}{x}} \frac{e^{x^2/2}}{1+x^2} \geqslant a$$

Since L.H.S is non decreasing for |x| the cr region is $\{|x| \geqslant k\}$

Where k is dreaming from the size condition

$$P_{H_O} \{|x| \geqslant k\} = \propto$$

Since X~N (0,1) Under $H_O, k = Z_\propto/2$

(7) Suppose X has the following distribution under H_O and H_i will here the critical region

$$\{ x : \sqrt{\frac{\pi}{2}} e^{|x| + x^2/2} \geq C \}$$

Since $\frac{f1(x)}{f0(x)}$ is a non-decreasing function of $|x|$, the critical region is $\{ |x| \geq k\}$ where k= $z\alpha/_2$

(8) Suppose x has the following distribution

$$H_0 : f_0(x) = \frac{1}{\sqrt{2\pi}} e^{x^2/2} ; -\infty < x < +\infty$$

$$H_1 : f_1(x) = \frac{2}{\Gamma\frac{1}{4}} e^{-X^4} ; -\infty < x < +\infty$$

Let us take a single observation. The MP test of H_0 Vs H_1 has the critical region

$\{ x: \frac{f1(x)}{f0(x)} \geq C \}$

Or $e^{-x^4 + x^2/2} \geq C'$

Since L.H.S. is a non-increasing function of $|x|$, the critical region is $\{|x| \leq k\}$ where k = $z_{(1-\alpha)}/_2$

(9) Suppose X has the following distribution

$$H_0: f_0(x) = \begin{cases} 4x ; 0 < x < 1/2 \\ 4(1-x); 1/2 \leq x < 1 \end{cases}$$

H_1: $f_1(x) = 1$; $0<x<1$

Let us take a single observation. The MP test of H_0 VS H_1has the critical region given by

$\frac{f1(x)}{f0(x)} \geq C$

Where $\frac{f1(x)}{f0(x)} = \begin{cases} \frac{1}{4x}; 0 < x < 1/2 \\ \frac{1}{4(1-x)}; 1/2 \leq x < 1 \end{cases}$

We see that $\frac{f1(x)}{f0(x)} \geq C$

If either $x < k_1$ or $x>k_2$

Hence MP or region is

$$\{x < k_1\} U \{x > k_2\}$$

The size of the test is P_{H0} $\{x < k_1\} U \{x>k_2\}$ + $P_{H0}\{x>k_2\} = \alpha$

For simplicity we can take $k_2 = 1-k_1$

(10) Let X have the rectangular distribution R(0,ϴ) having p.d.f.

$f(x,\Theta) = \frac{1}{\theta}$; $0 \leq x \leq \Theta$

We want to test

$$H_0: \Theta = \Theta_0 \text{ Vs}$$

$$H_1: \Theta = \Theta_1 (>\Theta_0)$$

We have

$$L(\Theta) = \frac{1}{\theta^n} I_{[0,X(n)]}(X_{(1)}) I_{[0,\Theta]}(x_{(n)})$$

Therefore the MP test has the critical region $W = \{\frac{L(\Theta 0)}{L(\Theta 1)} \geq C\}$

Now,

$$\frac{L(\Theta 0)}{L(\Theta 1)} = \left(\frac{\Theta 0}{\Theta 1}\right) n \frac{I_{[0,\theta 1]}(x_{(n)})}{I_{[0,\theta 0]}(x_{(n)})}$$

$$= \begin{cases} \left(\frac{\Theta 0}{\Theta 1}\right)^n for\ 0 \leq x_{(n)} \leq \theta_0 \\ \infty\ for\ \theta_0 \leq x_{(n)} \leq \theta_1 \end{cases}$$

This shows that $\frac{L(\Theta 0)}{L(\Theta 1)}$ is an increasing function of $x_{(n)}$ and, therefore

$$\frac{L(\Theta 0)}{L(\Theta 1)} \geq C \ \ x_{(n)} \geq k$$

Hence the MP test has the critical region

$$\{x_{(n)} \geq k\}$$

The value of k is determined by the size condition

$$P\{x_{(n)} \geq k/\theta_0\} = \alpha$$

Since $x_{(n)}$ has p.d.f. $fx_{(n)}(Y) = \frac{ny^{n-1}}{\theta^n}$; $0 \leq y \leq \Theta$

We have $$\frac{n}{\theta^n} \int_k^{\theta_0} y^{n-1} dy = \alpha$$

Remark: the above test is UMP for H_0: $\Theta=\Theta_0$ against H_1:$\Theta>\Theta_0$

As we have remarked, UMP test may not always exist. Therefore we for their restrict the class of tests by considering unbiased tests (defined below) and then try to obtain UMP test in the class of unbiased tests. If such a test exists we call it uniformly not powerful unbiased test (UMPU test)

Definition Suppose we are testing a sample hypothesis Hθ: $\theta = \theta_0$ against a conqurite alternative

$H_i(may\ be\ \theta \neq \theta_0 or\ \theta > \theta_0\ or\ \theta < \theta_0)$ A test T is called unbiased if

$P_o(T) \geqslant \propto$ for all $\theta \epsilon H_i$

Where $\propto$ is the size of T $i, e. P_o(T) = \propto$

Remark: Suppose $\theta = \theta_1$ is one of the alternative value of θ. If the test is not unbiased it may happen that $P_o(T) < \propto = P_{0_0}(T)$ which means that the probability of rejecting H_o when it is false is less then the probability if rejecting H_o when it is true if the test is unbiased it will not happen.

Theorem A MP test or UMP test is unbiased.

Prof Let T be a MP (or UMP) test of size $\propto$. Consider another test T which rejects the null hypothesis HO: $\theta = \theta_0$ with probability $\propto$ irrespective of the sample outcome. We may just toss a coin for which the probability of is $\propto$ and decide to reject the null hypothesis Hθ if we get $\propto$, irrespective if the sample values obtained. Then

$$P_T\{RejectH_o/\text{H}_0 \text{ is } true\} = \propto$$

So that the size of the test T=$\propto$. Also the power of test T is also$\propto$, since

$$P_T\{RejectH_o/\text{H}_0 \text{ is } false\} = \propto$$

But T being MP (or UMP) is such that

$$P_T(\theta) \geqslant P_T(\theta) \text{ for } \theta \epsilon \text{ H}_i$$

Or $$P_T(\theta) \geqslant \propto \text{ for } \theta \neq \theta_0$$

Remark: It may be shown that the following tests are UMPU for two sided alternative $H_i : \theta \neq \theta_0$ in example 1,2 and 3

For example 1, UMPU test is $\{\bar{x} \geqslant k_1 or \bar{x} \leqslant k_2\}$

For example 2, UMPU test is$\{[x] \geqslant k\}$

For example 3, UMPU test is$\{\sum(x_i - \mu)^2 \geqslant k_1\ or \sum(x_i - \mu)^2 \leqslant k_2\}$

The constant k, k_1, k_2 are determined from size condition

Now we consider a produce for constructing tests that has some intuitive appeal and that . Frequently, though not always, leads to UMP or UMPU test. Also the produce leads to test that have decided large sample properties

Suppose we are given a sample $(x_1, \dots, x_n)$ from a distribution with þ, $d, f\ f(x, \theta)$ (where θ may be a vector) and we deice to test the null hypothesis $H_o : \theta \epsilon w(\subset \Omega)$ against the alternative hypothesis $H_i : \theta \epsilon w(\subset \Omega)$ where Ω is the parameter space,

The likelihood function of the sample is given by

$$L(\theta) = l(\theta, x_1, \dots, x_n) = \prod_{i=1}^{n} f(x_i, \theta)$$

Define the <u>likelihood ratio</u>

$$\lambda = \frac{\underset{\theta\,\varepsilon\omega}{\max} L(\theta)}{\underset{\theta\,\varepsilon\Omega}{\max} L(\theta)}$$

Where $\underset{\theta\,\varepsilon\omega}{\max}\ L(\theta)$ denotes the maximum of the likelihood function when θ is restricted to values in w and max $L(\theta)$ denotes the maximum of the likelihood for when θ takes all possible values in Ω

Obviously, $0 \le \lambda \le 1$ and λ is also to 1 of the sample shows that θ lies actually in **ω**.

Definition The likelihood ratio test of H_o against H_i has the critical region

$$w = \{\lambda \leqslant \lambda_o\}$$

When λ_o is determined by the size condition

$$\underset{\theta\epsilon H_o}{\text{Sup}} P\{\lambda \leqslant \lambda_o/\theta\epsilon H_o\} = \propto$$

Remark (1) For testing a simple hypothesis against a simple alternative likelihood ratio test is equivalent to the test given by the Neyman –Pearson lemma.

(ii) if a sufficient statistics exists the L.R test is a function of the sufficient statistics.

(iii) Under some regularity condition -2 $\log_e \lambda$ is asymptotically distributed as a χ2 $r.v.$ with degrees of freedom equal to the difference between the number in **ω**.

Example: (1) Let X be a r.v. having a normal distribution $N(\mu,\sigma)$ where σ (=σ_o) is known

We want to test $H_O{:}\,\mu = \mu_o$

Against $H_1{:}\,\mu \neq \mu_o$

We have the likelihood function

$$L(\mu) = \frac{1}{(\sigma_O\sqrt{2\pi})^n} e^{-\sum_{i=1}^{n}(x_i-\mu)^2}/2\sigma_o^2$$

Then

$$\underset{H_0}{\max} L(\mu) = \frac{1}{(\sigma 0\sqrt{2\pi})^{n}} e^{-\sum_i^n (x_i-\mu_0)^2/2\sigma_0^2}$$

Since MLE of μ is $\hat{\mu} = \bar{x}$, therefore

$$\underset{\mu}{\max} L(\mu) = \frac{1}{(\sigma_0\sqrt{2\pi})^{n}} e^{-\sum_i^n (x_i-\bar{x})^2/2\sigma_0^2}$$

The LR test critical region is given by $\lambda \le \lambda_0$

$$\frac{\underset{H_0}{\max} L(\mu)}{\underset{\mu}{\max} L(\mu)} \le \lambda_0$$

$$\text{Or } \frac{e^{-\sum_i^n (x_i-\mu_0)^2/2\sigma_0^2}}{e^{-\sum_i^n (x_i-\bar{x})^2/2\sigma_0^2}} \le \lambda_0$$

$$e^{\frac{1}{2\sigma_0^2}[\Sigma(x_i-\bar{x})^2 - \Sigma(x_i-\mu_0)^2} \le \lambda_0$$

$$\text{Or } \frac{-n(\bar{x}-\mu_0)^2}{2\sigma_0^2} \le \log \lambda_0$$

$$\text{or } \frac{n(\bar{x}-\mu_0)^2}{\sigma_0^2} \ge k$$

$$\text{or } \frac{|(\bar{x}-\mu_0)|}{\sigma_0/\sqrt{n}} \ge k'$$

Remark (i) the above test is not UMP test since there exists other UMP tests for $H_1: \mu > \mu_0$ and $H_i: \mu < \mu_0$ (II) $\frac{\sqrt{n}(\bar{x}-\mu_o)}{\sigma^o} \sim N(0, I)$ under H_O so that k can k found easily by using size condition

(2) Let x $\sim N(O, I)$ where both μ and σ are unknown we want to test

$$H_O: \mu = \mu_o$$

Against $$H_i: \mu \neq \mu_0$$

We have the likelihood for

$$L(\mu, \sigma) = \frac{1}{(\sigma\sqrt{2\pi})} e^{-\frac{1}{2\sigma^2}\sum_i^n (X_I - \mu)^2}$$

Under $H_O: \mu = \mu_o$, (given) so the MLE of σ is

$$\hat{\sigma} == \sqrt{\sum_i^n \frac{(x_i - \mu o)^2}{n}}$$

In general, m, l, e of μ is $\hat{\mu} = \bar{x}$ and MLEof σ is

$$\hat{\sigma} = s_0 = \sqrt{\sum_i^n \frac{(x_i - \bar{x})^2}{n}}$$

Therefore, we have $$\max L(\mu, \sigma) = \frac{i}{(\&_o\sqrt{2\pi})^n} e^{-\Sigma(x_i - \mu o)^2} / 2\&_o^2$$

$$= \frac{i}{(\&_o\sqrt{2\pi})^n} e^{-\frac{n}{2}}$$

And $$\max_{\mu, \sigma} L(\mu, \sigma) = \frac{I}{(\&\sqrt{2\pi})^n} e^{-\Sigma(x_i - \bar{x})^2} / 2\&^2$$

$$= \frac{I}{(\&\sqrt{2\pi})^n} e^{-n/2}$$

The L.R test critical region is given by

$$\lambda = \frac{\max_{H_o} L(\mu, \sigma)}{\max_{H_o} L(\mu, \sigma)} \leqslant \lambda_O$$

Or $$\left(\frac{\&}{\&_o}\right)^n$$

Or $$\frac{\&^2}{\&_o^2} \leqslant \lambda'_O$$

Or $$\frac{n\&_o^2}{s^2} \geqslant k$$

Since $\&_o^2 = n(\bar{x} - \mu o)^2 + n\&^2$ the above cr region becomes

$$\frac{n(\bar{x} - \mu o)^2}{s^2} \geqslant k'$$

Or $$\frac{\sqrt{n}[(\bar{x} - \mu_o)}{s^2} \geqslant k''$$

Where $$s = \frac{\Sigma(x_i - \bar{x})^2}{n - i} = \frac{ns^2}{n - i},$$

It is know that $\frac{\sqrt{n}(\bar{x}-\mu_o)}{\&}$ has t distribution on $(n-1)d.f$ under H_O There fore the values of k can be found from the size condition

$$P\{|Y| \geqslant k\} = \propto$$

Where $Y \sim t_{n-i}$

(3) Let X ~N(μ, σ) when both μ and σ are unknown we want to test

$$H_O: \sigma = \sigma_O$$

Against

$$H_i: \sigma \neq \sigma_0$$

We have the likelihood function

$$L(\mu, \sigma) = \frac{1}{(\sigma\sqrt{2\pi})^n} e^{-\frac{1}{2\sigma^2}\sum_i^n (x_i-\mu)^2}$$

UnderH_O, the m, l, e of μ is $\hat{\mu} = \bar{x}$

In general, m, l, e of μ is $\hat{\mu} = \bar{x}$ and m, l, e of σ is

$$\hat{\sigma} = s = \sqrt{\frac{\sum(x_i - \bar{x})^2}{n}}$$

Then we have

$$\max_{H_o} L(\mu, \sigma) = \frac{1}{(\sigma_O\sqrt{2\pi})^n} e^{-\sum(x_i-\bar{x})^2}/2\sigma_o^2$$

$$= \frac{1}{(\sigma_O\sqrt{2\pi})^n} e^{-\frac{ns^2}{2\sigma_s^2}}$$

And $$\max_{H_o} L(\mu, \sigma) = \frac{1}{(\sigma_O\sqrt{2\pi})^n} e^{-\sum(x_i-\bar{x})^2}/2s^2$$

$$= \frac{1}{(\&\sqrt{2\pi})^n} e^{-\frac{n}{2}}$$

L.R test cr region is given by $$\lambda = \frac{\max_{H_O} L(\mu,\sigma)}{\max_{\mu,\sigma} L(\mu,\sigma)} < \lambda_O$$

Or $$\left(\frac{s^o}{\sigma_o^2}\right)^{\frac{n}{2}\left(\frac{s^o}{\sigma_o^2}-I\right)} < \lambda_o$$

Or $$y^{\frac{n}{2}} e^{-\frac{n}{2}(y-i)} < \lambda_o \text{ where } y = \frac{s^2}{\sigma_o^2}$$

We note that $y^{\frac{n}{2}} e^{-\frac{n}{2}(y-i)}$has a maximum at $y = 1$

Therefore $\lambda < \lambda_o$ if and only if $y \geqslant k_2$ or $y \geqslant k_1$ that is the critical region is

$$\left\{\frac{s^2}{\sigma_o^2} \leqslant k_2 or \frac{s^2}{\sigma_o^2} \leqslant k_1\right\}$$

$$\left\{\frac{(n)s^2}{\sigma_o^2} \geqslant k_2 or \frac{(n)s^2}{\sigma_o^2} \leqslant k_1\right\}$$

But it is know that $\frac{(n)s^2}{\sigma_o^2} = \frac{\sum_i^n (x_i-\bar{x})^2}{\sigma_o^2}$ has x^2 distribution on (n-i) d, f using the x_{n-i}^2tables and size condition we can get the values of $k_1$and k_2

(3a) suppose in example 3 the value of $\mu(=\mu_o)$is know. Then the L.R cr region because

$$\left\{\frac{ns_o^2}{\sigma_o^2} \geqslant c_1 or \frac{ns_o^2}{\sigma_o^2} \geqslant c_2\right\}$$

Where $$s_o^2 = \sum_i^n (x_i - \mu o)^2/\mathrm{n}$$

In than case $\frac{ns_o^2}{\sigma_o^2} = \frac{\sum_i^n (x_i - \bar{x})^2}{\sigma^2}$ hasx_n^2

(4)Let x have an exponential distribution

$$f(x,o) = \frac{1}{\theta} e^{-\frac{x}{\theta}} (x \geqslant \theta)$$

We want to test $$H_o : \theta = \theta_0$$

Against $$H_i : \theta = \theta_o$$

We have the likelihood function

$$L(\theta) = \frac{1}{\theta^n} e^{-\frac{1}{\theta}\Sigma x_i}$$

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

Then we get

$$\underset{H_o}{maxL(\theta)} = \begin{cases} \frac{1}{(\theta_o)^n} e^{-\frac{n\bar{x}}{\theta_o} for \bar{x} > \theta_o} \\ \frac{i}{(\bar{x})^n} e^{-n \ for \bar{x} \leqslant \theta_o} \end{cases}$$

Also $$\underset{\theta}{maxL(\theta)} = \frac{i}{(\bar{x})^n} e^{-n}$$

Because m, l, e of θ is $\hat{\theta} = \bar{x}$

The LR test cr region is given by $$x \leqslant \lambda_o$$

Where

$$x = \begin{cases} \frac{i}{(\theta_o)^n} e^{-\frac{i}{\theta}\Sigma x_i \bar{x} > \theta_o} \\ \frac{i}{(\bar{x})^n} e^{-n \ for \ \bar{x} \leqslant \theta_o} \end{cases}$$

Since $y^n e^{-n(y-i)}$ at lains maximum at $y - i$ taking $y = \frac{\bar{x}}{\theta_o}$ we see that λ=i if $y = i$ and λ$\leqslant\lambda_o$for $y \geqslant k (o < k < i)$

LR test critical region because

$$\left\{\frac{\bar{x}}{\theta_o} \geqslant k\right\} or \{\bar{x} \geqslant k\}$$

Remark (i) if one take $H_i : \theta = \theta_o$ we shall get the L.R critical region as$\{\bar{x} \geqslant k\}$ in both case of one –sided alternation the L.R test are UMP test.

(2) Since $\sum_i^n x_i$ has gamma distribution we can find the value of k by using size condition

(5) Let $(x_i, .. x_n)$ be a, r, s from $N(\mu, \sigma_i)$ and (γ_i, γ_{n2}) be a, r, s from another $N(\mu_2,, \sigma_2)$ where two samples (distribution) are independent.

We want to test

$$\left.\begin{matrix} H_o: \mu_1 = \mu_2 \\ H_i: \mu_1 \neq \mu_2 \end{matrix}\right\}$$

Where it is assumed that $\sigma_1 = \sigma_2 (= \sigma unkown)$ we that the like hood function

$$1(\mu_1, \mu_2, \sigma) = \frac{1}{(\sqrt{2\pi})^{n_{ith}} \sigma^{n_{ith}}} e^{-\frac{1}{2\sigma^2}[\sum_i^n (x_i - \mu_i)^2 + \sum_i^n (y_i - \mu_i)^2]}$$

In general the m, l, e of μ_1, μ_2 and σ are

$$\widehat{\mu_i} = \bar{x} = \frac{i}{n_1} \sum_i^n x_i, \widehat{\mu_2} = \bar{y} = \frac{i}{n_2} \sum_i^n y_i$$

And $$\widehat{\sigma^o} = \frac{n_1 s_1^2 + n_2 s_2^2}{n_1 + n_2} = s^2 (say)$$

Also $$s_1^2 = \frac{I}{n_1} \sum_i^n (x_i - \bar{x})^2 \text{ and} s_2^2 = \frac{I}{n_1} (y_i - \bar{y})^2$$

Therefore

$$\underset{\mu_1, \mu_2 \sigma}{maxL(\mu_1, \mu_2, \sigma)} = \frac{1}{(2\pi)^{n_1 + n_2} (s^2)^{n_1 + n_2}} e^{-\frac{(n_1 + n_2}{2}}$$

Against the m, l, e under H_O are

$$\widehat{\mu_1} = \widehat{\mu_2} = \frac{n_1 \bar{x} + n_2 \bar{y}}{n_1 + n_2} = m(say)$$

And $$\widehat{\sigma^2} = \frac{1}{n_1 + n_2} [\sum_i^n (x_i - m)^2 + \sum_i^n (x_i - m)^2$$

$$= \frac{1}{n_1 + n_2} [\sum_i^{n_1} \left\{ (x_i - \bar{x}) + (\bar{x} - m) \right\}^2 + \sum_i^{n_2} \{(y_i - \bar{y}) + (\bar{y} - m)\}^2]$$

$$= \frac{1}{n_1 + n_2} \left[\sum_i^{n_1} (x_i - \bar{x})^2 + n_1 (\bar{x} - m)^2 + \sum_i^{n_2} (y_j - \bar{y})^2 + n_2 (\bar{y} - m)\}^2 \right]$$

$$= \frac{1}{n_1 + n_2} \left[\sum_i^{n_1} (x_i - \bar{x})^2 + \sum_i^{n_2} (y_i - \bar{y})^2 + \frac{n_1 n_2}{n_1 n_2} (\bar{x} - \bar{y})^2 \right]$$

$$= s^o + \frac{n_1 n_2}{(n_1 + n_2)} (\bar{x} - \bar{y})^2 = s_o^2 (say)$$

Therefore $$\underset{H_o}{maxL(\mu_1, \mu_2, \sigma)} = \frac{I}{(\sqrt{2\pi})^{n_1 + n_2} (s_o^2)^{n_1 + n_2}} e^{-\frac{n_1 + n_2}{2}}$$

So that the LR cr region is given by

$$\lambda = \left(\frac{S_o^2}{S_o^2} \right)^{n_1 + n_2} \leqslant \lambda_o$$

Or $$\frac{s_o^2}{s_o^2} \leqslant k$$

Or $$\frac{(\bar{x} - y)^2}{(n_1 + n_2)^{s^2} \left(\frac{i}{n_1} + \frac{i}{n_2} \right)}$$

Or $$\frac{(\bar{x} - y)^2}{s^2 \left(\frac{i}{n_1} + \frac{i}{n_2} \right)}$$

Where $$s^2 = \frac{n_1 s_1^2 + n_2 s_2^2}{n_1 + n_2 - 2} = \frac{n_1 + n_2}{(n_1 + n_2 - 2)} s^2$$

The cr region can be within as

$$\left\{\frac{[\bar{x}-\bar{y}]}{s^{o}\sqrt{\frac{i}{n_1}+\frac{i}{n_2}}} \geqslant k\right\}$$

Since under find k such that $P\{\gamma \geqslant k\} = \propto$

Where $\gamma \sim t_{n_1+n_2-2}$

(6)Let $(X_I, . X_{n_I})$ be a, r, s from N (μ, σ_i) and$(\gamma_1, . \gamma_{n_2})n$ N (μ_2, σ_2) where two samples (and two distributions) are indecent

We want to test

Against
$$\left.\begin{matrix} H_O: \sigma_1 = \sigma_2 \\ H_i : \sigma_1 \neq \sigma_2 \end{matrix}\right\}$$

We have the likelihood function

$$1(\mu_1, . \mu_2 \sigma^{2^{\sigma}}) = \frac{1}{(2\pi)^{n_1+n_2}\sigma_1^{n1}\sigma_2^{n2}} e^{-\frac{1}{2}\left[\frac{\sum_i^{n_1}(x_i-\mu i)^2}{\sigma_1^2}+\frac{\sum_i^{n_2}(y_i-\mu i)^2}{\sigma_2^2}\right]}$$

In general, be m, l, e of $\mu_1, \mu_2, \sigma_1, \sigma_2$ are

$$\widehat{\mu_1} = \bar{x}, \widehat{\mu_2} = \bar{y}, \widehat{\sigma_1} = \frac{1}{n_1}\sum_i^{n_1}(x_i - \bar{x})^2, \widehat{\sigma_2} = \frac{1}{n_2}\sum_i^{n_2}(y_i - \bar{y})^2$$

So that
$$= s_1^2(say) = s_2^2(say)$$

$$\max L(\mu_1, \mu_2, \sigma_1, \sigma_2) = \frac{1}{(2\pi)^{n_1+n_2}(s_1^2)^{\frac{n}{2}}(s_2^2)^{n_2/2}} e^{-\frac{n_1+n_2}{2}}$$

Against, the m, l, e under H_O are

$$\widehat{\mu_1} = \bar{x}, \widehat{\mu_2} = \bar{y}, \widehat{\sigma_1} = \widehat{\sigma_2} = \hat{\sigma} = \frac{1}{n_1 + n_2}\left[\sum_i^{n_1}(x_i - \bar{x})^2 + \sum_i^{n_2}(y_i - \bar{y})^2\right]$$

$$= \frac{n_1 s_1^2 + n_2 s_2^2}{n_1 + n_2} = s^2(say)$$

So that
$$\max L(\mu_1, \mu_2, \sigma_1, \sigma_2) = \frac{1}{(2\pi)^{n_1+n_2}(s^2)^{\frac{n_1+n_2}{2}}} e^{-\frac{n_1+n_2}{2}}$$

Therefore, the LR cr region is given by

$$\lambda = \frac{(s_1^2)^{\frac{n_1}{2}}(s_2^2)^{\frac{n_2}{2}}}{(s^2)^{\frac{n_1+n_2}{2}}} \leqslant \lambda_o$$

Or
$$\frac{s_1^2)^{\frac{n_1}{2}}(s_2^2)^{\frac{n_2}{2}}}{\left(\frac{n_1 s_1^2+n_2 s_2^2}{n_1+n_2}\right)^{\frac{n_1+n_2}{2}}} \leqslant \lambda_o$$

Or
$$\frac{\left[\frac{(n_1-1)}{(n_2-1)}f\right]^{\frac{n_1}{2}}}{\left[1+\frac{(n_1-1)}{(n_2-1)}f\right]^{\frac{n_1+n_2}{2}}} \leqslant \lambda_o$$

Where
$$f = \frac{n_1 s_1^2}{(n_1-1)} / \frac{n_1 s_2^2}{(n_2-1)}$$

Setting $g(f) for$the $L.H.S$ of (i) we have g(o)=o and g(f)→o∞. Furthermore g(f) attains its maximum for $\frac{f}{may} = \frac{n_1(n_2-1)}{n_2(n_1-1)}$, it is impressing between o and f may and derision in (f may, ∞).

Therefore $g(f) \leqslant \lambda_o$ if and only if $f < k_1$ or $f >$ the LR cr region can be within as $\{F < k_1 \; or \; F > k_2\}$

Where $$F = \frac{n_1 s_1^2/(n_1-1)}{n_2 s_2^2/(n_2-1)}$$

But under $H_o, F \sim F n_{1-1,n_2-1}$,

Hence k_1, k_2 can be obtained from the size condition $P\{f > k_1 \; or \; F < k_2\} = \propto$ whese $F \sim F_{n_{1-1,n_2-1}}$

Some distribution: X^2, t and F

Definition: A $r, v. x$ is said to have a Gamma distribution $G(\propto, \beta)$ of its $þ. d. f.$ is given by

$$f(x) = \frac{\beta^{\propto}}{\Gamma(\propto)} x^{\propto-1} e^{-\beta\propto} \quad ;x \geq 0$$

$$= 0 \quad ; x < 0$$

$$(\propto > 0, \beta > 0)$$

We have m, g, f $M_x(t) = (1 - \frac{t}{\beta})^{-\propto}$, $\quad t < \beta$

$$E(X) = \propto/\beta$$

$$V(X) = \propto/\beta^2$$

If $\propto = 1$ we get the exponential distribution

$$f(x) = \beta e^{-\beta x} \quad , x \geqslant 0 (\beta > 0)$$

$$E(X) = 1/\beta$$

$$V(X) = 1/\beta^2$$

If $\propto = n/2$(n a positive integer) $\beta = 1/2$ we get the x^2 distribution on n, d, f where $þ, d, f$ is

$$f(x) = \frac{1}{2^{\frac{n}{2}} I(^n/_2)} x^{\frac{n}{2}-i} e^{-\frac{x}{2}}, x \geqslant o$$

We have m, g, f $M_x(t) = (1 - 2t)^{-n/2}$

$$\left.\begin{array}{l} E(x) = n \\ v(x) = 2n \end{array}\right\}$$

Definition: A r, v X is said to have a t −distribution on n, d, f if its $þ, d, f$ is given by

$$f(x) = \frac{\Gamma\left(\frac{n+1}{2}\right)}{\Gamma\left(\frac{n}{2}\right)\sqrt{nx}} \left(1 + \frac{x^2}{n}\right)^{-\frac{n+1}{2}}, -\infty < x < \infty$$

If X~$n(o - i), \gamma \sim$x^2(n) and x and γ are inept then $T = X/\sqrt{^{\gamma}/_n}$ has $t(n)$

Define: A r, vX is said to have a $F -$ distribution on$(m, n) d, f$ if its $þ, d, f$ is given by

$$f(x) = \frac{\Gamma\left(\frac{m+n}{2}\right)}{\Gamma\left(\frac{m}{2}\right)\Gamma\left(\frac{n}{2}\right)} \left(\frac{m}{n}\right)^{\frac{m}{2}} \frac{\frac{m}{2} - 1}{(I + \frac{m}{n} x)^{\frac{m+n}{2}}}, x \geqslant 0$$

$$=0 \qquad , x<0$$

Of $x \sim x^2(m)$ and $\gamma \sim x^2(n)$ where x and γ are independent $z = \frac{x/m}{Y/n}$ has $F(m,n)$

Percentage points the upper$\propto -$ percent point of the $x^2(n)$ distribution is $x^2 n, \propto$ where

$$P(x^2(n) > x^2 n, \propto) = \propto$$

The upper$\propto -$ percent point of the $t_{(n)}$ distribution $is, tn, \propto$ where

$$P\left(t_{(n)} > tn, \propto\right) = \propto$$

Since t-distribution is symmetrical

$$P\left(\left[t_{(n)}\right] > t_{n, \propto/2}\right) = \propto$$

The upper$\propto -$ percent point of the F(m ,n, $\propto$) distribution is Fm ,n, $\propto$ where

$$P(F(m,n) > F, m.n, \propto) = \propto$$

Note that $$Fm, n, i - \propto = \frac{i}{Fn_{,n\propto}}$$

Use of $x^2 t$ and $\bar{t}$ distribution in testing problem

Use of x^2 distribution (i) **Testing the variance of a of a distribution**: Given a sample $(x_i, \ldots x_n)$ of size n from a normal distribution $N(\mu, \sigma)$ where σ is unknown, we would like to test $H_o: \sigma = \sigma_o$ against alternative $\sigma > \sigma_o$ or $\sigma < \sigma_o$ or $\sigma \neq \sigma_o$ the tests are summarised n follows

Case I μ know

Alternative	**reject H_o at level $\propto$ if**
$H_0: \sigma > \sigma_O$	$\sum_i^n (x_i - \mu)^2 / \sigma_o^2 \geqslant x^2 n, \propto$
$H_o: \sigma < \sigma_O$	" " $\leqslant x^2 n, \propto$
$H_o: \sigma \neq \sigma_O$	" " $\leqslant x^2 n, \propto x^2_{n-i,-\propto/2}$ or $\geqslant x^2_{n-i,-\propto/2}$

Case II μ know

Alternative	**reject H_o at level $\propto$ if**
$H_0: \sigma > \sigma_O$	$(n-i)(s)^2 \geqslant x^2 n - i \propto,$
$H_o: \sigma < \sigma_O$	" " $\leqslant x^2 n, \propto$
$H_o: \sigma \neq \sigma_O$	" " $\leqslant x^2 n, \propto x^2_{n-i,-\propto/2}$ or $\geqslant x^2_{n-i,-\propto/2}$

Where $(s)^2 = \frac{1}{n-1} \sum_i^n (x_i - \bar{x})^2$

(2) Testing proportions in k(>2) classes Suppose a, r, v takes values in one of k(>2)mutually exclusive classes $A_I, \ldots\ldots A_k$ with $þ = P(x \epsilon A_I), 1,2, \ldots\ldots k, \sum_I^k þ_I = I$ we want to test the hypotheses that

$$H_o: þ_i = þ_i^o (i = 1, \ldots . k)$$

Against $$H_i: þ_i \neq þ_i^o \text{ for all}$$

For a random (x, … x_n) of n observation let the observed frequencies in the k classes be o_1, o_2, $o_n (\sum_i^n o_i = n)$and the expected frequencies under the H_o be $e_1, e_2, \ldots \ldots . e_k$ $(\sum_i^n e_i = n)$ where $e_i = n\text{þ}_i$ calculate

$$\chi 2 = \sum_i^{k} \frac{(o_i - e_i)^2}{e_i}$$

Them, for large sample, x^2has $x^2(k - i)$ the test of H_o has the cr. region

$$x^2 \geqslant x^2_{k-i,k}$$

Note: it we want to test $H_o \text{þ}_1 = \text{þ}_2, \ldots \ldots . = \text{þ}n$ we take $\text{þ}_i^o = \frac{1}{k}$to any

(3) <u>Testing goodness of fit:</u> given a sample $(x_1, . . x_n)$ of Observation on $a.r.v$ X arranged in the form of a frequencies distribution having k classes $A_I, \ldots . . A_k$ we would like to test the hypothesis that distribution of X has a specified from with $\text{þ}, d, f (or\ \text{þ}, m, f) f_o(x, \theta)$ the parameter θ be a simple one or a vector $(o_i, \ldots . . \theta_e)$

Let the observed frequencies in the k classes be $o_1, o_2 \ldots . , o_k, \sum_i^k o_i = n$ and the expected frequencies under H_o be $e_i, e_2, \ldots . e_n \sum_i^k o_i = n$

Such that $e_i = P_{H_o}(x \epsilon A_i)$ Calculate

$$\chi 2 = \sum_i^{k} \frac{(o_i - ei))^2}{ei} = \sum_i^{k} \frac{o_i^2}{e_i} - n$$

Then, for large sample, x^2has $x^2(k - i)$ the test of H_o has the cr. Region

$$x^2 \geqslant x^2_{k-i,\propto}$$

Note if $r(of\ \ell)$ parameters in θ are estimated from the sample then χ2 has $\chi 2(k - r - i)$if any expected frequency is lass then 5 we pool this class with the adjoining class and denote by$k$ the effective number ʅ classes after paroling

(4) **<u>Testing independence of two attributes in a $k x \ell$ contingency table</u>**

In a $(k x \ell)$ contingency table for two attributes, we want to test

H_o: Two attributes are independent

Against H_o: Two attributes are not independent

Let O_{ij} = observed frequency in the (i, j) the cell

And e_{ij} = expected a=$(ith\ row\ total\ x^{jth} colum\ total)$n " " "under H_o

Calculate

$$\chi 2 = \sum_{i=1}^{k} \sum_{i=1}^{k} \frac{(o_{ij} - eij)^2}{e_{ij}}$$

$$= \sum_{i=1}^{k} \sum_{i=1}^{k} \frac{o_{ij}^2}{e_{ij}} - n$$

Where n=total frequency. Then x^2 has x^2 on$(k - i)x(\ell - i)d.f$ the test of H_o has the cr. Region

$$x^2 \geqslant x^2_{k-i,\ell-i,\propto}$$

(5) Testing the homogeneity of $k(> 2)$ correlation coefficients.

Suppose $r_1, \ldots. r_k$ are k sample correlation coefficients corresponding to k normal

Distribution with population correlation coefficients $p_i, \ldots p_k$ we want to test

$$H_O : p_1, \ldots p_k =, .. p_k$$

Us H_I: all correlation coefficients are not equal we use the friskers z-trans function of correlation coefficients given by $z = \frac{1}{2} log_e \frac{i+r}{i+r}, S = \frac{1}{2} log_e \frac{i+p}{i-p}$ so that

$$z \sim N\left(S, \frac{1}{\sqrt{n-3}}\right)$$

Where n is the sample size.

We calculate $z_1, z_2, \ldots\ldots z_k$ corresponding to $r_1, r_2, \ldots\ldots. r_k$ having sample size $n_1, n_2, \ldots. n_k$ and define

$$\bar{z} = \sum_{i}^{k} (n_i - 3) z_i / (\sum_{i}^{k} (n_i - 3)$$

And $$x^2 = \sum_i^k (n_i - 3)(z_i - \bar{z})^2$$

Then $\chi 2$has $\chi 2$on $(k - i)d.f$ and the test of H_o has cr. Region

$\chi 2 \geqslant \chi 2_{(n-i),\propto}$

Remark: if H_o is accepted we may obtain an estimate of the common corresponding coefficients $\rho^*(say)$ by solving

$$\bar{z} = \frac{1}{2} log_e \frac{1 + \rho^*}{1 - \rho^*}$$

Uses if t-distribution:

(i)**Testing the mean of a single population**: let $(x_1, \ldots\ldots.. x_n)$ be a sample of size n from a normal population $N(\mu, \sigma^2)$ and, as usual, $\bar{x}$ and s^2 are the sample mean and sample variance. We would like to let the null hypothesis $H_o: \mu = \mu_o$ against alterative $\mu > \mu_o$ or $\mu < \mu_o$ or $\mu \neq \mu_o$ the tests are summarised as follows:

(2) **Testing the equality of two population means:** let $(x_1, \ldots.. x_{n2})$ and $(y_1, \ldots \ldots y_{n2})$ be two samples from in dept normal populations $N(\mu_1, \sigma_1)$ and $N(\mu_2, \sigma_2)$ respectively let $\bar{x}, \bar{y}, s_1^2, s_2^2$ be as usual and let

$$(s)^2 = \frac{(n_1 - 1)(s_1)^{\,2} + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}$$

Be the pooled variance.

We would like to test $H_O: \mu_1 = \mu_2$ against alternative $\mu_1 < \mu_2$ or $\mu_1 \neq \mu_2$ the test are summarised as follows:

Case I

Alternative | Reject H_O at level $\propto if$

$H_i: \mu_1 > \mu_2$ $\qquad \frac{\bar{x} - \bar{y}}{\sqrt{\left\{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}\right\}}} \geqslant z_\propto$

$H_i: \mu_1 < \mu_2$ $\qquad$ " $\leqslant -z_\propto$

$H_i: \mu_1 \neq \mu_2$ $\qquad \frac{[\bar{x} - \bar{y}]}{\sqrt{\left\{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}\right\}}} \geqslant z_{\propto/2}$

Case II σ_1, σ_2 unknown $(\sigma_1 = \sigma_2)$ essential corruption

Alternative | Reject H_O at level $\propto if$

Remark: if we want to test $H_O = \mu_1 - \mu_2 = (\neq o) we$ use the statistics

$$\frac{(\bar{x} - \bar{y}) - (\delta)}{\& \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

Uses of F-distribution:

(1)**Testing equality of two population variances**:

Let two samples of sizes n_1 and n_2 be given from two independent normal population $N(\mu_1, \sigma_1)$ and $N(\mu_2, \sigma_2)$, respectively .Let $s_1^2, s_{\,2}^{\,2}$ be the two sample variance. We would like to test the null hypothesis $H_o: \sigma_1 = \sigma_2$ against$H_i: \sigma_1 \neq \sigma_2$ The test are cr follows:

Case I μ_1, μ_2 known

Reject H_o if either $\frac{\sum_{i=1}^{n_1} (x_i - \mu 1)^2}{\sum_{i=1}^{n_2} (y_i - \mu 2)^2} \geqslant \frac{n_1}{n_2} F_{n_1, n_2, \propto/2}$

Or $\qquad \frac{\sum_{i=1}^{n_2} (y_i - \mu 2)^2}{\sum_{i=1}^{n_2} (x_i - \mu 1)^2} \geqslant \frac{n_2}{n_1} F_{n_1, n_2, \propto/2}$

Case II I μ_1, μ_2 known

Reject H_o if either $\frac{(s1)^2}{(s2)^1} \geqslant F_{n_1-1,n_2-1,\propto/2}$ If $s_1 > s_2$

Or $\frac{(s1)^2}{(s2)^1} \geqslant F_{n_2-1,n_1-1,\propto/2}$ If $s_2 > s_1$

(2) **Testing the multiple correlation coefficient**: Given a sample of size or from a bivariate normal population (x_1, x_2, x_3) with multiple correlation coefficient $R_{1(23)}$ of $x_1 or (x_2, x_3)$ we would like to test the null hypotheses $H_O R_{1(23)} = 0$ let the sample multiple correlation coefficient be $R_{1(23)}$. The test is to reject H_O at level $\propto$ if

$$\frac{r_{(23)}^2}{1 - r_{1(23)}^2} \cdot \frac{n-3}{2} \geqslant F_{2,n-3,\propto}$$

(3) **Testing the equality of means of k normal distribution ($k > 2$)[see left page]**

Farceur's z-transformation of correlation coefficient: Suppose a sample of size n is drawn from a bivariate population with correlation coefficient the variables Fisher intruded the transformation

$$z = \frac{1}{2} log_e \frac{1+r}{1-r}$$

Where r is a sample correlation coefficient Though the population correlation coefficient P may be widely different from zero, the new statistics z may be amounted to be normally distributed even when n is as small as 10 it has hen show that z has approximate mean

$$\xi = \frac{1}{2} log_e \frac{1+p}{1-p}$$

And approximate mean $^1/_{(n-3)}, i.e$

$$\sqrt{n-3}(z - \xi) \sim N(o, 1)$$

(I)For testing $H_o : P = P_o$ against $H_i : P \neq P_o$ we reject H_o if

$$\sqrt{n-3}[z - \xi_o] \geqslant N_{\propto/2}$$

Where $\xi_o = \frac{1}{2} log_e \frac{1+p}{1-p}$ and $N_\propto$ is the appear $\propto$ % point of normal distribution $N(O, 1)$

(ii)For testing $H_o : p_1 = p_2$ against $H_i : p_1 \neq p_2$ involving two populations, let r_1, r_2 be the sample correlation coefficient for two independent sample of size n_1, n_2 from the two populations and let $z_{1,} z_2$ be there transformed values, i, e

$$z_i = \frac{1}{2} log_e \frac{1+r_i}{1-r_i} (i = 1,2)$$

The test is to reject H_o at level $\propto$ if

$$\frac{|z_1 - z_2|}{\sqrt{\frac{1}{n_{1-3}} + \frac{1}{n_{2-3}}}} \geqslant N_{\propto/2}$$

(iii)Let $r_1, r_2 \ldots\ldots r_k$ be sample correlation coefficient for k sample of sizes $n_1, n_2 \ldots n_k$ drown from k independent vicariate normal population with correlation coefficients $p_1 p_2 \ldots\ldots p_k$. Let z_1, z_k be the transformed values and let

$$\bar{z} = \frac{\sum_{i=1}^{k}(n_i - 3)z_i}{\sum_{i=1}^{k}(n_i - 3)}$$

The test is to reject H_o at level $\propto$ if

$$\sum_{i=1}^{k}(n_i - 3)(z_i - \bar{z})^2 \geqslant x^2_{k-1,\propto}$$

If H_o is accepted an estimate of common correlation coefficient p is p^r where$\bar{z}$ is the transformed values of p^*(x) For large sample

$$þ \sim N\left(p\frac{\sqrt{P(I - P}}{n}\right)$$

Large sample tests so for we have considered tests of hypothesis which contain assumptions regarding the population are satisfied .Now we consider some approximate test which are valid only for sufficiently large samples, but they have wide applicability and hold for all populations satisfying certain general conditions rather than being valid for some particular populations only (e.g. normal)

(i)Testing a proportion: Suppose in a population is the proportion of members with a qualitative character A. Let p be the proportion of members with A in a random sample of size n. we would like to test the hypothesis H_0: $P=P_0$ .The test is to reject H_0 at level α if

$$\frac{[þ - P_o]}{\sqrt{p_o(1 - p_o)/n}} \geqslant N_{\propto/2}$$

(ii)Testing the equality of two population proportions: Let p_1, p_2 be two population proportions and $þ_1, þ_2$ be the two sample proportions dream from there indecent population the test of $H_o:, P_1, P_2$ is to reject $H_o$at level ∝if

$$\frac{[þ_1 - þ_2]}{\sqrt{þ(i - þ)\left\{\frac{1}{n_1} + \frac{1}{n_2}\right\}}} \geq N_{\propto/2}$$

Where

$$þ = \frac{n_1 þ_1 + n_2 n_2}{n_1 + n_2}$$

(iii)Testing for a st. deviation: let s be the st. Deviation of a sample of observation of size drown from a population with st. Deviation σ(x) the test of $H_O: \sigma = \sigma_o$ is to reject H_O at level ∝if

$$\frac{[s - \sigma_o]}{\sigma_o/\sqrt{2n}} \geqslant N_{\propto/2}$$

(iv)Testing for equality of two population st. Deviation Let s_1, s_2 be the st. Deviation of two sample of sprees n_1, n_2 from two independent population with st. Deviation σ_1, σ_2 Let

$$s^2 = \frac{n_1 s_1^2 + n_2 s_2^2}{n_1 + n_2}$$

The test of $H_o: \sigma_1, \sigma_2$ is to reject the at level $\propto$if

$$\frac{[s_1 - s_2]}{s\sqrt{\frac{1}{2n_1} + \frac{1}{2n_2}}} \geqslant N_{\propto/2}$$

Definition:- For a random sample $(x_1, \dots, x_n)$ from the distribution of a $r.v.x$ having $þ, d, f\ \ f(x, \theta)$ Let $L_1 L_1(x_1, \dots, x_n)$and $L_2(x_1, \dots, x_n)$ be two statistics such that $L_1 \leq L_2$. The interval $[L_1, L_2]$ is a confidence interval for θ with. Confidence coefficient $1{-}\propto$ $(0 < \propto < 1)$ if $P_\theta[L_1 \leq \theta \leq L_2] = 1{-}\propto$ for all $\theta \in \Omega$ L_1 and L_2 are called the lower and upper confidence limits, respectively at least one of them should not be a constant.

Interval Estimation

Estimation of a parameter by a sample value is known as point estimation. An alternation produce is to give an interval within which the parameter may be supposed to lie with high probability. This is called interval estimation and the interval is called the confidence for the parameter

Suppose a, r, v x has Normal distribution $N(\mu, \sigma)$ with unknown mean μ and known st. Deviationσ. Let $(x_i, \dots, x_n)$ be the values of a random sample of size or from then distribution .We know that the sample mean $\bar{x} \sim N\left(\mu, \frac{\sigma}{\sqrt{n}}\right)$ and, hence$\frac{\sqrt{n}(x-\mu)}{\sigma} \sim N(o, i)$. It follows that

$$P\left\{-1.96 \leqslant \frac{\sqrt{n}(x-\mu)}{\sigma} \leqslant 1.96\right\} = 0.95$$

Or, equivalently,

$$P\left\{\bar{X} - 1.96 \leqslant \mu \leqslant \bar{X} + 1.96\frac{\sigma}{\sqrt{n}}\right\} = 0.95$$

This shows that, in respected sampling the probability is 0.95 that the interval

$$\left\{\bar{X} - 1.96\frac{\sigma}{\sqrt{n}}; \bar{X} + 1.96\frac{\sigma}{\sqrt{n}}\right\}$$

Will include μ, We say that above is a confidence interval for μ with confidence coefficient ,95. The two end points are known as 95% confidence limits forμ.

Let us now consider the general problem Let a, r, v x has distribution depending on an unknown parameter θ which is to be estimated. Suppose Z is a statistics (usually it is a function of a sufficient statistics if it exists) which is a function of θ but whose distribution does not depend onθ. Such a statistics z is called a ploetal function Let λ_1 and λ_2 be two numbers such that

$$P\{\lambda_1 \leq Z \leq \lambda_2\} = 1{-}\propto \qquad -\ (1)$$

For a specified $\propto$ $(o < \propto < 1)$

The above inequality can be solved such that it assumes the from

$$P\{\theta_1((x_1, \dots, x_n)) \leq \theta \leq \theta_2(\lambda_1, .. \lambda_2)\} = 1{-}\propto$$

For all θ where $\theta_1$and $\theta_2$are random variables which do not depend onθ.Finally, if we astute the sample value $[\theta_1((x_1, \dots, x_n)), \theta_2((x_1, \dots, x_n))]$ becomes a confidence interval for θ with desired confidence coefficient $1{-}\propto$.

Remark: the numbers λ_1, λ_2 may be chosen in several ways, giving rise to several confidence intervals. We usually choose confidence intervals of shortest length.

Example (i) $X \sim N(\mu, \sigma)$ where σ is Known and μ is to be estimated

$$Let\ z = \frac{\sqrt{n}(\bar{x} - \mu)}{\sigma}$$

Which has $N(O, I)$ distribution For a specified$\propto$ let $N_{\propto/2}$ be the $\frac{\alpha}{2}$ % critical value of $N(o, 1)$then

$$P\left\{-N_{\propto/2} \leqslant \frac{\sqrt{n}(\bar{x} - \mu)}{\sigma} \leqslant N_{\propto/2}\right\} = I{-}\propto$$

Or $$P\left\{\bar{x} - N_{\propto/2}\frac{\sigma}{\sqrt{n}} \leqslant \mu \leqslant \bar{x} + N_{\propto/2}\frac{\sigma}{\sqrt{n}}\right\} = i{-}\propto$$

So that $$P\left\{\bar{x} - N_{\propto/2}\frac{\sigma}{\sqrt{n}}\bar{x} + N_{\propto/2}\frac{\sigma}{\sqrt{n}}\right\}$$

Isa confidence interval of μ with confidence coefficient $(i{-}\propto)$

(2) . $x \sim N(\mu, \sigma)$, σ unknown and μ to be estimated

$Let\ z = \frac{\sqrt{n(\bar{x}-\mu)}}{s}$where $s^2 = \frac{i}{i-1}\sum_i^n(x - \bar{x})^2$

Then z has t(n-i) distribution , so that for a specified $\propto$,

$$P\left\{t_{n-1,\propto/2} \leqslant \frac{\sqrt{n}(\bar{x} - \mu)}{s} \leqslant t_{n-1\propto/2}\right\} = i{-}\propto$$

Or $$P\left\{\bar{X} - t_{n-1,\propto/2}\frac{S}{\sqrt{n}} \leqslant \bar{\mu} \leqslant \bar{X} + t_{n-1,\propto/2}\frac{S}{\sqrt{n}}\right\} = i{-}\propto$$

So that $$\left\{\bar{X} - t_{n-1,\propto/2}\frac{S}{\sqrt{n}}, \bar{X} + t_{n-1,\propto/2}\frac{S}{\sqrt{n}}\right\}$$

Is a confidence interval of μ with confidence coefficient $(1{-}\propto)$

(3) $x \sim N(\mu, \sigma)$, μ known and σ is to be estimated

$$Let\ z = \sum_i^n (x_1 - \mu)^2$$

Then z has $x^2(n)$ distribution, so that for a specified $\propto$

$$P\left\{X^2_{n,i-\propto/2} \leqslant \frac{\sum(x_i - \mu)^2}{\sigma^2} \leqslant X^2_{n,i-\propto/2}\right\} = 1{-}\propto$$

Or $$P\left\{\frac{\sum(x_i-\mu)^2}{X^2_{n,1-\propto/2}} \leqslant \sigma^2 \leqslant \frac{\sum(x_i-\mu)^2}{X^2_{n,1-\propto/2}}\right\} = 1{-}\propto$$

There, $a(1{-}\propto)$% confidence interval of σ^2

$$\left\{\frac{\sum(x_i - \mu)^2}{X^2_{n,1-\propto/2}}, \frac{\sum(x_i - \mu)^2}{X^2_{n,1-\propto/2}}\right\}$$

(4) $x \sim N(\mu, \sigma), \mu$ Unknown and σ is to be estimated

$$Let\ z = \frac{(n-1)s^2}{\sigma^2}\text{ Where } s^2 = \frac{1}{n=1}\sum_i^n(x_1 - \bar{x})^2$$

Then z has $x^2(n)$ distribution, such that

$$P\left\{X^2_{n,i-\propto/2} \leqslant \frac{(n-1)s^2}{\sigma^2} \leqslant X^2_{n,i-\propto/2}\right\} = 1-\propto$$

Or $$P\left\{\frac{(n-1)s^2}{X^2_{n,i-\propto/2}} \leqslant \sigma^2 \leqslant \frac{(n-1)s^2}{X^2_{n,i-\propto/2}}\right\} = 1-\propto$$

Therefore, a $(i-\propto)\%$ confidence interval of σ^2 is

$$\left\{\frac{(n-1)s^2}{X^2_{n,i-\propto/2}}, \frac{(n-1)s^2}{X^2_{n,i-\propto/2}}\right\}$$

(5) Let x have an exponential distribution with parameter λ which is to be estimated

$$Let\ z = 2\lambda n\bar{x}$$

Then Z has $x^2(2n)$ ditribution, so that for a specified $\propto$

$$P\{X^2_{2n,1-\propto/2} \leqslant 2\lambda n\bar{x} \leqslant X^2_{2n,1-\propto/2}\} = i-\propto$$

Or $$P\left\{\frac{X^2_{2n,1-\propto/2}}{2n\bar{x}} \leqslant, \leqslant \frac{X^2_{2n,,\propto/2}}{2n\bar{x}}\right\}$$

Therefore, a $(i-\propto)\%$ confidence interval of λ is

$$\left\{\frac{X^2_{2n,1-\propto/2}}{2n\bar{x}} \leqslant, \leqslant \frac{X^2_{2n,,\propto/2}}{2n\bar{x}}\right\}$$

(6) Let X ~N(μ, σ) and γ ~N$(\mu_2, \sigma_2)$where $\sigma_1 = \sigma_2(unkown\ )$. We want a confidence for $(\mu_1 - \mu_2)$

$$Let\ z = \frac{(\bar{x}-\bar{y}) - (\mu_1 - \mu_2)}{s\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

Where $\bar{x}, \bar{y}, s$ are usually defined $(s^2 = \frac{(n_1-1)s_1^2 + (n_2-1)s_2^2}{n_1+n_2-2}$

Then Z has $t(n_1 + n_2 - 2)$distribution, such that

$$P\left\{t_{n_1+n_2-2,\propto/2} \leqslant \frac{(\bar{x}-\bar{y}) - (\mu_1 - \mu_2)}{s\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} t_{n_1+n_2-2,\propto/2}\right\} = i-\propto$$

Or $P\left\{(\bar{x}-\bar{y}) - t_{n_1+n_2-2,\propto/2s}s\sqrt{\frac{1}{n_1}+\frac{1}{n_2}} \leqslant (\mu_1 - \mu_2) \leqslant (\bar{x}-\bar{y}) + t_{n_1+n_2-2,\propto/2s}xs\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}\right\} = i-\propto$

So that a confidence interval for $\mu_1 - \mu_2$ is

$$\left\{(\bar{x}-\bar{y}) - t_{n_1+n_2-2,\propto/2s}s\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}\right\}(\bar{x}-\bar{y}) + t_{n_1+n_2-2,\propto/2s}xs\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}$$

With confidence coefficient 1−∝

(7)Let X ~N (μ_1, σ_1) and γ ~N (μ_2, σ_2) where μ_1, μ_2 are unknown and it is requested to obtain a confidence interval of $\frac{\sigma_1^2}{\sigma_2^2}$

Let Z = $\frac{S_1^2/\sigma_1^2}{S_2^2/\sigma_2^2}\ (S_1^2 > S_2^2)$

So that Z has F distribution on $(n_1 - i, n_2 - i)d, f$

Then

$$P\left\{F_{n_1-i,n_2-i,i-\propto/2} \leqslant \frac{S_1^2/\sigma_1^2}{S_2^2/\sigma_2^2} \leqslant F_{n_1-i,n_2-i,i-\propto/2}\right\} = i{-}\propto$$

Or $$P\left\{\frac{S_1^2/S_2^2}{F_{n_1-i,n_2-i,i-\propto/2}} \leqslant \frac{\sigma_1^2}{\sigma_2^2} \leqslant \frac{S_1^2/S_2^2}{F_{n_1-i,n_2-i,i-\propto/2}}\right\} = i{-}\propto$$

Or $$P\left\{\frac{S_1^2/S_2^2}{F_{n_1-i,n_2-i,i-\propto/2}} \leqslant \frac{\sigma_1^2}{\sigma_2^2} \leqslant \frac{S_1^2/S_2^2}{F_{n_1-i,n_2-i,i-\propto/2}}\right\} = i{-}\propto$$

So that $$\left\{\frac{1}{F_{n_1-i,n_2-i,i-\propto/2}}\frac{S_1^2}{S_2^2}, F_{n_1-i,n_2-i,i-\propto/2}\frac{S_1^2}{S_2^2}\right\}$$

Is a confidence interval of $\frac{\sigma_1^2}{\sigma_2^2}$ with confidence coefficient $i{-}\propto$

(8) Simultaneous confidence region for $(\mu,\sigma)$for a normal distribution.

Let $x\sim N(\mu,\sigma),\mu,\sigma$ with unknown

One many chose a confidence region for (μ,σ) using the two relations

$$P\left\{\bar{x} - t_{n-1,\propto/2}\frac{s}{\sqrt{n}} \leqslant \mu \leqslant \bar{x} - t_{n-1,\propto/2}\frac{s}{\sqrt{n}}\right\} = i{-}\propto$$

Diagrammatically shown as the shaded region below

Where $t_a = \bar{x} - t_{n-i,\propto/2}\frac{s}{\sqrt{n}} etc$

$$x_a = \frac{(n-1)s^2}{x_{n-1,\propto/2}^2}$$

But it is difficult to find the probability of the sample to full in the shaded region (confidence region)

Alternatively, using the independence of $\bar{x}$ and s^2 we chose the cofidence region by the help of relation

$$P\left\{-N_{\propto_1/2} \leqslant \frac{\bar{x}-\mu}{\sigma/\sqrt{n}} \leqslant N_{\propto_1/2}\right\} = 1-\propto_1$$

A,d $$P\left\{x_{n-i,\propto/2}^2 \leqslant \frac{(n-1)s^2}{\sigma^2} \leqslant x_{n-i,\propto/2}^2\right\} = 1-\propto_2$$

Since $\bar{x},s^2$ are indept

$$P\left\{N_{\propto_1/2} \leqslant \frac{\bar{x}-\mu}{\sigma/\sqrt{n}} \leqslant N_{\propto_1/2}, x_{n-i,\propto/2}^2 \leqslant \frac{(n-1)s^2}{\sigma^2} \leqslant x_{n-i,\propto/2}^2\right\} = (1-\propto_1),(1-\propto_2)$$

Chosing $\propto_1,\propto_2$ such that $(I-\propto_1),(i-\propto_2) = i{-}\propto$ we can

Obtain the boundaris of the confidence .region without difficully this is shown by the shaded region below

Where $$q = N_{\propto_1/2}$$

$$q_1 = x_{n-i,i,\propto/2}^2$$

Approximate confidence intervals(for large samples)

Let x be bernoulli $r.v$ with

$P(X=1)=P, P(x=o)=1-p$ we want to find confidence interval for P.

For lage sample size ,n, we have

$$\frac{p-p}{\sqrt{P(i-P)/n}} \sim N(o,1)$$

Or

$$\frac{p-p}{\sqrt{P(i-P)/n}} \sim N(o,1)$$

Where þ is the sample propostion

Them , approxi mately ,

$$P\left\{-N_{\propto/_2} \leq \frac{þ-þ}{\sqrt{þ(i-þ)/n}} \leq N_{\propto/_2}\right\} = 1-\propto$$

Or

$$P\left\{þ - N_{\propto_2}\sqrt{\frac{þ(I-þ)}{n}} \leqslant þ + N_{\propto/_2}\sqrt{\frac{þ(I-þ)}{n}}\right\} = 1-\propto$$

So that

$$\left\{þ - N_{\propto/_2}\sqrt{\frac{þ(1-þ)}{n}}, þ + N_{\propto/_2}\sqrt{\frac{þ(1-þ)}{n}}\right\}$$

Is a $(1-\propto)\%$ confidence interval for P

(II) For two sample we can similerly find a confidence interval for P_1, P_2 as follows:

$$P\left\{N_{\propto_2} \leqslant \frac{(þ_1, þ_2) = (P_1, P_2)}{\sqrt{[þ(I-þ)\left(\frac{1}{n_1}+\frac{1}{n_2}\right)]}} \leqslant N_{\propto/_2}\right\} = 1-\propto$$

Where $þ = \frac{n_1 þ_1 + n_2 þ_2}{n_1+n_2}$

So that $\left\{(þ_1, þ_2) - N_{\propto/_2}\sqrt{[þ(I-þ)}\left(\frac{1}{n_1}+\frac{1}{n_2}\right) þ_1, þ_2) - N_{\propto/_2}\sqrt{[þ(I-þ)}\left(\frac{1}{n_1}+\frac{1}{n_2}\right)\right\}$

Is a $(i-\propto)\%$ confidence interval for $p_1 - p_2$

(iii) Let x be a, r, v having mean μ, variance σ^2 and we want a confidence interval for σ

For that approximately .

$$P\left\{-N_{\propto_2} \leqslant \frac{s-\sigma}{s/\sqrt{2n}} \leqslant N_{\propto_2}\right\} = 1-\propto$$

Or

$$P\left\{s - N_{\propto_2}\frac{s}{\sqrt{n}} \leqslant \sigma \leqslant s + N_{\propto_2}\frac{s}{\sqrt{n}}\right\} = 1-\propto$$

Then

$$P\left\{s - N_{\propto_2}\frac{s}{\sqrt{n}}, s + N_{\propto_2}\frac{s}{\sqrt{n}}\right\}$$

Is a $(i-\propto)\%$ confidence for σ

(iv) For two sample we an similerly find a cofidence interval for $\sigma_1 - \sigma_2$ as follows:

$$P\left\{-N_{\propto/2} \leqslant \frac{(s_1 - s_2) - (\sigma_1 - \sigma_2)}{s\sqrt{\frac{1}{2n_1} + \frac{1}{2n_2}}} \leqslant N_{\propto/2}\right\} = 1{-}\propto$$

Where $s^2 = \frac{n_1 s_1^2 + n_2 s_2^2}{n_1 + n_2}$

So that

$$\left\{(s_1 - s_2) - N_{\propto/2} s\sqrt{\frac{1}{2n_1} + \frac{1}{2n_2}}, (s_1 - s_2) + N_{\propto/2} s\sqrt{\frac{1}{2n_1} + \frac{1}{2n_2}}\right\}$$

Is a $(i{-}\propto)\%$ confidence interval for $(\sigma_1 - \sigma_2)$

(v) Let (x, y) have a bivanate normal distribution with coefficient P and me want to find a confidence region for P.

By using Fisher,s Z transformation

$$\xi = \frac{1}{2} log_e \frac{1+p}{1-p}$$

and $$z = \frac{1}{2} log_e \frac{1+r}{1-r}$$

whose r is the corr crofficient in a sample of size n

Then $$\frac{Z-3}{1\sqrt{n-3}} \sim N(O, I)$$

So that

$$P\left\{-N_{\propto/2} \leqslant \sqrt{n-3}(Z-3) \leqslant N_{\propto/2}\right\} = I{-}\propto$$

Or $$P\left\{Z - \frac{1}{\sqrt{n-3}} N_{\propto/2} < 3 \leqslant Z + \frac{1}{\sqrt{n-3}} N_{\propto/2}\right\} = I{-}\propto$$

So that

$$\left\{z - \frac{1}{\sqrt{n-3}} N_{\propto/2}, z + \frac{1}{\sqrt{n-3}} N_{\propto/2}\right\}$$

Gives a $(i{-}\propto)\%$ confidence interval for ξ.From this we can earily obtain the corrponding confidence interval for P.

NON-PARAMETRIC INFERENCE

In all problems of statictics inference considered so fan we assumed that the distribution of the random variable breing sampled is know n except for some parameters . in pratice however the functional from in the distribution is seldom if ever , known if is therefore desivable to devise some produres that are free from this assumption concering distribution such produres are commonly refered to as distribution free or non-parametric methods the term distribution free refers to the fact that no assumptions are made about the underlying distribution execpt that the distribution function being sampled is absolutely continuous or purely discrete. The term non-parametric refers to the factors that there are no parameters involved in the traditional sense of the parameter used so for.

We will consider only the inferential problem of testing of hypothesis and dercribe a few non-parametrictests

Single- sample problems : **(a)The problem of fit** : the problem of fit is to test the hypothesis that a sample of obsevations (x_i, x_n) is from some specified distribution against the alternative that it is from some other distribution.Thus we have to test

$$H_o: x \sim F_o(x) = F_o(x)$$

Against $$H_o: x \sim F(X) \neq F_o(x) \text{for some } x$$

(i)Chi- square test: Let there be k categories and let p_i be the probality of a random obsevation from $F_o(x)$ to fall in the ith category $(i = 1,2, \ldots . n)$.For a sample of size n, Let o_i be the obsevarved freqnecy in the ith category and let $e_i = np_i$ be the expected frequency in the ith category under H_o.

To test H_o we use the chi-square statics

$$x^2 = \sum_{i=1}^{n} \frac{(o_i - ei)^2}{e_i}$$

The larger the value of x^2 the more likely it is that the $o_{i,s}$ did not come from $F_o(x)$. The x^2 −statistic for large samples has a x^2 distribution on $(k - 1)$d.f .Thus an approximate level ∝ test is provided by rejecting H_o if

$$x^2 > x^2_{k-1\propto,}$$

(ii)**Kolmogoror – Smironv one sample test** : For the sample $(x_i, \ldots x_n)$let the empirical distribution function F_n(x) be given by

$$F_n(x) \begin{cases} o & if\ x < \quad x_{(i)} \\ k/n & if\ x_{(k)} \leqslant x < \quad x_{(k-i)} \\ i & if\ x \geqslant \quad x_{(n)} \end{cases}$$

$(k = 1,2, \ldots n, -1)$ whese $x_{(1)}, x_{(2)}, \ldots . x_{(n)}$ are the order statistic , Evidently ,

$$F_n^Y(x) = \frac{number\ of\ x_k, s\ (I, \leqslant k \leqslant n) \leqslant x}{n}$$

For testing $H_O: F_{(x)} = F_o(x)$ against the two sided alternative $H_i: F_{(x)} \neq F_o(x)$ we use the Kolmogoror – Smironv statictic

$$D_n = \sup_x [F_n^Y(x) - F_o(x)]$$

It can be shown that the K-S statistic D_n is completely distribution free for any continouns distribution $F_o(x)$

At level ∝, Kolmogoror – Smironv test rejects H_O if

$$D_n > D_{n,\propto}$$

Whese $$P(D_n > D_{n,\propto}) \leqslant \propto$$

Tables of $D_{n,\propto}$ for given ∝ and n are available

Remark1: For testing$H_O: F_{(x)} = F_o(x)$ against one-sided alternatives $H_1: F_{(x)} > F_o(x)$ or $H_2: F_{(x)} < F_o(x)$ based on one-sided K.S statistics $D_n^+$and D_n^- are also available

Remark 2: For small sample x^2 −test is not available but K.S test can be applied. For discrete distibution K.Stest is not availible but x^2 −test can be appled K.S test is more powerful then x^2 −test.

(B) **The problem of Location:** Let $(x_i, \ldots . x_n)$ be a radom sample from a distribution $F_{(x)}$ with unknown median ξ ,where $F_{(x)}$ is assumed to be continus in the neigbourhood of ξ. By definition of median $(P(x \geqslant \xi) = \frac{1}{2}$.We would like to test the hypothesis

(x) If n>25, normal appronimution may be used

We take $\frac{R-n/2}{n/4} \sim N(o, i)$

$H_o: \xi = \xi o$ against one sided or two sided alternatioes

Sign Test: We from the n differences $(x_i - \xi_O)I = 1{,}2 \ldots \ldots \ldots n$ and find out the number, R,of position differences (differences having postive signs) i, e when $(x_i - \xi_O) > o$.

If H_O is true, $P(X_i - \xi_O \geqslant O) = \frac{1}{2}, i = 1{,}2, \ldots . n$ and R has a Biomial distribution with paramer$\frac{1}{2}$. We may use an exect test ofH_O based on the Biomial Distribution. In the case of one-sided alternative

$$H_i: \xi > \xi o$$

The sample will have an excess of positive signs and in the case of

$$H_i: \xi > \xi o$$

The sample will have a small number of postive signs

The signs test based on R, for testing H_O can be summarised as follows :

The critical values $R_{1\propto}, R_{2\propto}, R_{\propto/2}, R_{\propto/2}$ are calculate from tables of Biomaial distribution

Rajred –sample signs test: Here we assume that we have a random sample of n pains (x_n, x_n) giving the the differences

$$D_i = x_i - y_i \quad , i = 1, \ldots n$$

It is assumed that the distribution of D=X-Y is absolutely continous with median ξ

We have , now a single sample $D_I, \ldots \ldots D_n$ and we can test $H_o: \xi = \xi_o$ which can be taken to be oby the sign test descrited above.

Remark the above two sign test s are , repectively aralogoun to single sample $t - test$ and paired t-test for testing location of a normal distribution ,

Two sample problems : let $(x_i, \ldots \ldots x_n)$ and $(\gamma_i, \ldots \ldots \gamma_n)$ be independent random sample s from two absolutely continous distribution $F_x(x)$ and $F_\gamma(y)$, respectively

Suppose we want to test

$$H_o: F_x(x) = F_\gamma(y) \text{ for all } x$$

Against $H_i: F_x(x) \neq F_\gamma(y)$ for same x

Run test(Wald –Wolfowitz): we assarge the m, x's and n $\gamma's$ in increasing order of size $\underline{XYYXXXYYYXY}$and count the numbers of runs .if H_o is true the (m+n) values will be well mixed up and we expect that R, the total number of runs , will be relatively large. But R will be small if the samples come from differernt popaltions i, e H_o is false in the extreme case , if all the value of y are greater than all the value of x, or vice – vera , there will be only two runs

The run test of H_o against H_i at level ∝ is to reject H_o if

$$R \leqslant R_\propto$$

Where $R_{\propto}$ is the largest interteger such that

$$P(R \leqslant R_{\propto}/H_o) \leqslant \propto$$

It can be show that distribution of R, under H_o is given by

$$P(R = 2 \propto /H_o) = 2\binom{m-i}{\propto -i}\binom{n-i}{\propto -i}/\binom{m+n}{m}$$

And $$P(R = 2 \propto +i/H_o) = \binom{m-i}{\propto}\binom{n-i}{\propto -i} + \binom{m-i}{\propto -i}\binom{n-i}{\propto}$$

Tables of critical values of R based on above have been given by swed and Eisenhant

For large m,n(both greater then 10), Ris asymptohcally Normally distributed with

$$E(R) = \frac{2mn}{m+n} + 1$$

And $$V(R) = \frac{2mn(2mn-m-n)}{(m+n)^2(m+n-i)}$$

<u>Median it test</u>: We arrange the x's and y's in asscending order of size and find the median M of the contied sample let

$$V = number\ of\ x'swhich\ are\ \leqslant median\ M$$

If V is large it is reasomable to conclude that the actual median of x is smaller than the median of Y $i, e\ H_o: F_x(x) = F_Y(x)$ is respected

Hown of $H_i: F_x(x) > F_Y(x) -$

On the other hand , if V is too small it is reamable to condude that the actual median of X is greater than the median of y $i.e\ \ H_o: F_{x(x)} = F_y(x)$is respected in fovoues of $H_i: F_x(x) < F_Y(x)$

For the two sided alternative , we use the two sided test .

The median test can be summarised as follows:

It can be shown that the distribution of V, under H_o is given by

$$P(V = u/H_o) = \frac{\binom{m}{u}\binom{n}{þ-u}}{\binom{m+n}{þ}}, u = o, 1 \ldots \ldots.., n$$

Where $m + n = 2þ, þ$ positive integer

And

$$P(V = u/H_o) = \frac{\binom{m}{u}\binom{n}{þ-u}}{\binom{m+n}{þ}}, u, 1 \ldots \ldots. \min(m, þ)$$

Where $m + n = 2þ + 1, þis\ a$ positive integer

<u>Wilcoxon- Mann –Whitney U test</u>: This is the most widely used two- sample non-parametric test and is a useful alternative to the t-test assumotions.

The test is like the run test based on the pattern of $m, x's$ and $n, y's$ arranged in ascending order of size . The Main- Whitney U statistic is defined as the number of times as X preades $a\ Y$ In the combined sample of size $m + n$. We define

$$z_{ij} = \begin{pmatrix} 1, x_i < y_j \\ 0\ , x_i > y_j \end{pmatrix} \begin{pmatrix} i = 1, \ldots \ldots m \\ j = 1 \ldots \ldots. n \end{pmatrix}$$

And write

$$U=\sum_{i=1}^{m}\sum_{i=1}^{n} z_{ij}$$

Note that $\sum_{i=1}^{m} z_{ij}$ is the number of $y_{j's}$ that are larger than x_i and hence U is the number of values of $x_i, \dots \dots \dots x_n$ that are smaller than each of $y_i, \dots \dots \dots, y_n$. For example , suppose the contined sample when ordered is as follows :

$$X_2 < X_1 < Y_3 < Y_2 < X_4 < Y_1 < X_3$$

Then U=7, becouse there are three values of X<Y_1, two values of X<Y_2 and two values of X<Y_3

It is obseved that U=0 if all the$x_i's$ are larger than all $y_i's$ and U=mn of all the $x_i's$ are smaller than all the $y_i's$. Thus $o \leqslant U \leqslant mn$. If U is large the values of y tend to be larger than X (Y is stochastically larger than X) and this supposts the alternative $F_x(x) > F_y(x)$. Similarly, if U is small, the values of Y tend to be smaller than X and this supposts the alternative $F_x(x) > F_y(x)$.

Thereforer , U-test can be summarised as follows:

H_o	H_i	$RejectH_o \ if$
$F_x(x) = F_y(x)$	$F_x(x) > F_y(x)$.	$U \geqslant C_1$
$F_x(x) = F_y(x)$	$F_x(x) < F_y(x)$	$U \leqslant C_2$
$F_x(x) = F_y(x)$	$F_x(x) \neq F_y(x)$	$U \geqslant C_3 or\ U \leqslant C_4$

It can be shown that Under H_O

$$E(U)=\frac{mn}{2}$$

And $$V(U)=\frac{mn(m+n+1)}{12}$$

The tables of distribution of U for small samples are given by table and Mann-Whitney. For large samples U has asymptotic normal distribution,i,e

$$\frac{U-\frac{mn}{2}}{\sqrt{\frac{mn(m+n+1)}{12}}} \sim N(O,I)$$

APPENDIX

Distribution of function of random variables (transformations method)

Therom: suppose Xis a continuous r,u with $\text{þ},d,f\ f_x(x)$. Set $x=\{x,f_x(x)>o\}$.Let

(i) $y = g(x)$ difine a d.f transformation of x anto x

(ii) the derivative of $x = g^{-1}(x)\ \ \omega.r.t\ \ y$ is continous and non-zero for $y \epsilon x$, where $g^{-1}(y)$ is the inverse for of $y(x)\ i,e\ \ g^{-1}(y)$ isthat x for which $g(x) = y$

Then $\gamma = g(x)$ is a cont. r,u with $\text{þ},d,f$.

$$f_y(y) = f_x(g^{-1}(y))\left[\frac{d}{dy}g^{-1}(y)\right]$$

Therom : let x_1 and x_2 be jointly continous $r.u.s$ with þ$,d,f$ $f_{x_1,x_2}(x_1,x_2)$. Set $x = \{(x_1,x_2): f((x_1,x_2) > o\}$Assumu that

(i)$y_1, = g_1(x_1,x_2)$ and $y_2, = g_2(x_1,x_2)$ defines i:i transformation of x onto x.

(ii)The first partical of derivatives of $x_1 = g_i^{-1}(y_1 y_2)$ and $x_2 = g_i^{-1}(y_1 y_1)$ are continous over x.

(iii) The jacebian of transformation is non-zero for $(y_1 y_1)\epsilon x$.Then the joint þ$,d,f$ of $\gamma_1 = g,(x_1,x_2)$and $\gamma_2 = g,(x_1,x_2)$is given by

$$f\gamma_{1\gamma_2}(y_1,y_2) = fx_{1x_2}\{g_1^{-1}(y_1,y_2)g_2^{-1}(y_1,y_2\}iji$$

Where

$$IJI = \begin{bmatrix} \frac{\propto x_i}{\propto y_i} & \frac{\propto x_1}{\propto y_2} \\ \frac{\propto x_2}{\propto y_1} & \frac{\propto x_2}{\propto y_2} \end{bmatrix}$$

X²- distribution

Definition : A continous r,u,x is said to have the X²- distribution on n degrees of freedom if its þ$,d,f$ is given by

$$f(x) = \frac{1}{x^{n/2}1(^n/_2)} x^{\frac{n}{2}-1} e^{-x/2}, \quad x \geqslant o$$

$$= o \qquad\qquad x < o$$

The m,g,f of x is given by

$$M_x(t) = Ee^{tx}$$

$$= \frac{1}{x^{n/2}1(^n/_2)} \int_o^\infty x^{\frac{n}{2}-1}\, e^{x(1-2t)/2dx}$$

$$= \frac{1}{x^{n/2}1(^n/_2)} \frac{1(^n/_2)}{\left(\frac{1-2t}{2}\right) \quad n/2}$$

$$= (1-2b)^{-n/2}$$

From this we can earily show that

$$E(X) = n \; and \; v(x) = 2n$$

For $n \leqslant 2$the þ$,d,f$ of $x^2(n)$ steadily dencress as x iscrese while for $n > 2$ there is a uniqne maximum at $x = n - 2$

Theorom : Let $x_1, x_2 \ldots \ldots .. x_n$ be n independent standand normal r,v,s $i.e \;\; x_i \sim N(o,1), i = 1, \ldots n$ Then $y = \frac{n}{2} x_i^2$ has a X²- distribution on n,d,f.

Proof: Let X be $N(o,1)$ the m,g,f of x^2 is given by

$$M_{x2} = E\left(e^{tx^2}\right)$$

$$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{tx^2} - x^{2/2}dx$$

$$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{tx^2} - x^{2(1-2t)/2}dx$$

$$= \frac{\sqrt{2\pi}}{\sqrt{(1-2t)}} \frac{1}{\sqrt{2\pi}}$$

$$= (1-2t)^{-1/2}$$

Which show that $x^2 \sim x^2(1)$ Then , the $m, g,$f of $\gamma = \sum_i^n x_i^2$ is given by

$$M_{X2}(t) = [M_{X2}(t)]^n = (1-2t)^{-n/2}$$

Which shows that $\gamma \sim x^2(n)$

Therom : Let $\gamma_1, \gamma_2 \ldots . \gamma_n$ be indepent r, u, s with X^2- distribution on $n_i, \ldots \ldots . n_k$ degrees of freedom resp .

Then $z = \sum_i^k \gamma_i \sim x^2(n_1 + n_2 + .. + n_k)$

Proof :the m, g, f **Z**

$$M_Z(1) = Ee^{tz}$$

$$= Ee^t \sum_i^k Y_e$$

$$= \prod_{i=1}^k E(e^{tye})$$

$$= (1-2t)^{-(n_i + .. + n_k)}/2$$

Which about that y$\sim x^2(n_i + \cdots + n_k)$

Crollanj : Let $(x_i, \ldots \ldots x_n)$be a random simple from a Normal distributuion $N(\mu, \sigma)$.Then $\sum_{i=1}^n \frac{(x_i-\mu)^2}{\sigma^2}$ has x^2 distribution on n, d, f.

Therom: Let $(x_i, \ldots \ldots x_n)$be a random simple from a Normal distributuion $N(\mu, \sigma)$ Let $\bar{x} = \sum_i^n x_i/n$

And $s^2 = \frac{1}{n-i} \sum_i^n (x_i - \bar{x})^2$ be the sample mean and sample variance. Then $\frac{(n-i)s^2}{\sigma^2}$ has x^2 distribution on $(n-i)d, f$.

Therom: For large $n, \sqrt{2x^2}$ can be shown to be approximately normally distributred with mean $\sqrt{2n-1}$ and st-dearation unity.

Therom: Assume that y has distribution function F_Y which satifies some regularity conditions ad which has r-unknown parameters $\theta_1, \theta_2 \ldots . \theta_r$ and that $(y_i, .. y_n)$ is a random sample of y.Let $\widehat{\theta_\iota}, \widehat{\theta_r}$ be the $m. \ell, e$ of $\theta's$ .Suppose the sample is distribution in k non-orerlapping intervals $\{I_J\}$
where $I_J = \{y: a_{j-i} < y < a_{j-i}\}, j = 1, \ldots k(a_o = -\infty a_k = \infty$and . Let $x_i, \ldots \ldots x_k$ be the number of sample values falling in these inervals, respectively if me define

$$\widehat{p_J} = P\{Y falls\ in I_J\}, j = 1, \ldots k$$

Where $\widehat{\theta_\iota}, \widehat{\theta_k}$ replace θ_i, θ_k in F_y ,then the distribution of the statistics $z = \sum_{j=1}^k \frac{(x_j - n\widehat{p_J})^2}{n\widehat{P_J}}$Lerger is appoximately distributed as x^2on $k - r - i\ d, f$ as n gets

<u>Students t-distribution</u>

Definintion : A Continous r, u, x is said to have the t-distribution on n, d, f if its þ$, d, f$ is given by

$$f(x)\frac{[\left(\frac{n+1}{2}\right)}{[\left(\frac{n}{2}\right)\sqrt{n\pi}}\frac{1}{(1+\frac{x^2}{n})^{\frac{n+1}{2}}}, -\infty < x < \infty$$

Remark : For $n = i$the $þ, d, f$

$$f(x) = \frac{1}{\pi}\frac{1}{i + x^2}, -\infty < x < \infty$$

Which shows that it is a couchy distribution We will therefore, assume that $n > i$

Remark:the $þ, d, f$ of t-distribution is symmctric about again. For large n the t-distribution tends to Normal distribution. For small n hawever t-distribution deviates considerally from the normal in fact if $T \sim t_{(n)}$and $z \sim N(o, i)$

$$P\{[T] \geqslant t_o\} \geqslant P\{[Z] > t_o\}$$

Moments : Since the distribution a symmetrial about origin $\mu_{2r} + 1 = 0$

For 2r<n

$$\mu_{2r} = E(X^{2r})$$

$$= \frac{2[\left(\frac{n+1}{2}\right)}{[\left(\frac{n}{2}\right)\sqrt{n\pi}}\int_o^\infty \frac{X^{2r}}{\left(1+\frac{x^{2r}}{n}\right)\frac{n+1}{2}^{dx}}$$

Therom : Let $x \sim N(o, 1)$ and $y \sim x^2(n)$ and Let $x$and y be independent .Then $U = \frac{x}{\sqrt{y/n}}$