: Unit-4 - Testing of Hypothesis :

Testing of Hypothesis

: Hypothesis :

① Null Hypothesis (H_0)

Alternative hypothesis (H_1)

population
↓
Characteristics (Parameters)

ERROR:

Type-1 error: Probability of rejecting the null hypothesis when it is true. α

Type-2 error: Probability of accepting the null hypothesis when it is false.

Power of the test: Probability of rejecting a null hypothesis when it is false.

Level of significance (α)

Maximum probability of rejecting the null hypothesis when it is true.

1. The fraction of defecting item in a large lot is p. To test the null hypothesis $H_0 : p = 0.2$, one considers the numbers of defecting x in a sample of 8 items and failes to reject the null hypothesis if $x \leq 6$ & reject the hypothesis otherwise.

(i) What is the amount of type-1 error?

(ii) What is the amount of type-2 error? (corresponding to $p = 0.1$)

(iii) What is the power of the test $p = 0.01$?

$\Rightarrow$ (i) $x \leq 6 \longrightarrow$ accept $\quad x > 6 \longrightarrow$ Reject

$x \sim B(8, 0.2)$

$\boxed{{}^{n}c_{x} \cdot p^{x} \cdot q(1-p)^{n-x}}$

$n = 8$
$p = 0.2$

Type - 1 - error

$= P(x > 6 / p = 0.2)$

$= P(x = 7 / p = 0.2) + P(x = 8 / p = 0.2)$

$= {}^{8}c_{7}(0.2)^{7} \cdot (0.8)^{1} + {}^{8}c_{8}(0.2)^{8} \cdot (0.8)^{0}$

$= \frac{8!}{7!} \cdot (0.2)^{7} \cdot (0.8) + \frac{8!}{8!}(0.2)^{8}$

$= 8(0.2)^{7} \cdot (0.8) + (0.2)^{8}$

$= 0.00008192 + 0.00000256$

$= 0.00008448$.

(Am-Q) 2.

(ii) Type - 2 - error:

$P(x \le 6 / p = 0.1)$

$= 1 - [P(x=7/p=0.1) + P(x=8/p=0.1)]$

$= 1 - [{}^{8}C_{7} \cdot (0.1)^7 \cdot (0.9)^1 + {}^{8}C_{8} (0.1)^8 \cdot (0.9)^0]$

$= 1 - [\frac{8!}{7!} \cdot (0.1)^7 \cdot (0.9) + \frac{8!}{8!} (0.1)^8]$

$= 1 - [0.000000072 + 0.00000001]$

$= 0.99999$

(iii) Power of the test

1 - type - II error

$= 1 - 0.9999$

$= 0.00001$

2. To examine the claim of reputed published the λ, the average number ~~number~~ misprint, per page of a book is one. If a particular page of that book contains more then (two) misprint, then the null hypothesis $\lambda = 1$ is rejected.

(i) What is the amount of type-1 error?

(ii) Find the power of the test when the alternative hypothesis is $\lambda = 2$. given that $e^{-1} = 0.368$.

$\Rightarrow$ let, x be ~~the~~ number of misprint per page of a book.

Here, $X \sim P(\lambda)$

$$\boxed{\frac{e^{-\lambda} \lambda^x}{x!}}$$

(i) Type-1 error: $\qquad \dfrac{e^{-\lambda}\cdot\lambda^{x}}{x!}$

$= P(X>2/\lambda=1)$

$= 1-P(X\le 2/\lambda=1)$

$= 1-[P(X=0/\lambda=1)+P(X=1/\lambda=1)+P(X=2/\lambda=1)]$

$= 1-\left[\dfrac{e^{-1}\cdot 1^{0}}{0!}+\dfrac{e^{-1}\cdot 1^{1}}{1!}+\dfrac{e^{-1}\cdot 1^{2}}{2!}\right]$

$= 1-\left[e^{-1}+e^{-1}+\dfrac{e^{-1}}{2}\right]$

$= 1-\left[\dfrac{e^{-1}+2e^{-1}+e^{-1}}{2}\right] = 1-\dfrac{4e^{-1}+e^{-1}}{2}$

$= \cancel{1-\dfrac{1.3e^{-1}}{2}} \qquad = 1-\dfrac{5e^{-1}}{2}$

$= \cancel{1-\dfrac{3}{2}(0.368)} \qquad = 1-\dfrac{5}{2}(0.368)$

$\qquad = 0.08.$

$= \cancel{1-0.552}$

$= \cancel{0.448}.$

(ii) Power of the test:

$= P(X>2/\lambda=2)$

$= 1-P(X\le 2/\lambda=2)$

$= 1-[P(X=0/\lambda=2)+P(X=1/\lambda=2)+P(X=2/\lambda=2)]$

$= 1-\left[\dfrac{e^{-2}\cdot 2^{0}}{0!}+\dfrac{e^{-2}\cdot 2^{1}}{1!}+\dfrac{e^{-2}\cdot 2^{2}}{2!}\right]$

$= 1-\left[e^{-2}+2e^{-2}+\dfrac{4e^{-2}}{2}\right]$

$= 1-[5e^{-2}]$

$= 1-5\times(e^{-1})^{2} = 1-5\times(0.368)^{2} = 0.1354$

3. If $x \geq 1$ is the critical region for testing $H_0 : \theta = 2$ against the alternative $\theta = 1$ on the basis of the single observation from the population $f(x:\theta) = \theta \cdot e^{-\theta x};\ 0 < x < \infty$.

Obtain the value of type-1 & type-2 error.

→ <u>type-1 error:</u> $P(x \geq 1 / \theta = 2)$

$$= \int_0^\infty 2 \cdot e^{-2x}\, dx$$

$$= 2\int_0^\infty e^{-2x}\, dx$$

$$= 2 \cdot \left[\frac{e^{-2x}}{-2}\right]_1^\infty$$

$$= -\left[e^{-2\infty} - e^{-2}\right]$$

$$= e^{-2}$$

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

<u>type-2 error:</u> $P(x < 1 / \theta = 1)$

$$= \int_0^1 e^{-x}\, dx$$

$$= \left[\frac{e^{-x}}{-1}\right]_0^1$$

$$= -1\left[e^{-1} - e^{-0}\right]$$

$$= -1\left(\frac{1}{e} - 1\right)$$

$$= 1 - \frac{1}{e}$$

4. Let, p be the probability that a coin will fall head in a single toss in order to test $H_0 : p = \frac{1}{2}$; Against $H_1 : p = \frac{3}{4}$.
The coin is tost 5 times & H_0 is rejected if more then 3 heards are obtain find the amount of type-1 error & power of the test.

→ type-1 error :

$n=5$
$p=\frac{1}{2}$.

$P(X>3 / p=\frac{1}{2})$

$= P(X=4/p=\frac{1}{2}) + P(X=5/p=\frac{1}{2})$

$= {}^5C_4 \cdot (\frac{1}{2})^4 \cdot (\frac{1}{2})^1 + {}^5C_5 \cdot (\frac{1}{2})^5 \cdot (\frac{1}{2})^0$

$= 5 \cdot \frac{1}{16} \cdot \frac{1}{2} + 1 \cdot \frac{1}{32}$

$\cancel{= 0.1562 + 0.0312}$

$\cancel{= 0.00488}$.

$= \frac{5}{32} + \frac{1}{32} = \frac{6}{32} = \frac{3}{16} = 0.1875$.

Power of the test :

$P(X>3 / p=\frac{3}{4})$

$= P(X=4/p=\frac{3}{4}) + P(X=5/p=\frac{3}{4})$

$= {}^5C_4 \cdot (\frac{3}{4})^4 \cdot (\frac{1}{4})^1 + {}^5C_5 \cdot (\frac{3}{4})^5 \cdot (\frac{1}{4})^0$

$= 5 \cdot (\frac{3}{4})^4 \cdot \frac{1}{4} + 1 \cdot (\frac{3}{4})^5$

$= (\frac{3}{4})^4 \left[\frac{5}{4} + \frac{3}{4}\right]$

$= (\frac{3}{4})^4 \times \frac{8}{4}\,2$

$= \frac{81}{256} \times 2 = \frac{81}{128}$.

Level of significance :

5. A coin is to be tested for unbaisedness. The hypothesis that it is unbaised is rejected if more toses of the coin out of 10 toses result in heads. Can we take 1% as level of significance.

$Bin(10, \frac{1}{2})$

$H_0 : p = \frac{1}{2}$

→ **type - 1 error:**

$$p(x \geq 9 / p = \tfrac{1}{2})$$

$$= p(x = 9 / p = \tfrac{1}{2}) + p(x = 10 / p = \tfrac{1}{2})$$

$$= {}^{10}C_9 \cdot \left(\tfrac{1}{2}\right)^9 \cdot \left(\tfrac{1}{2}\right)^1 + {}^{10}C_{10} \left(\tfrac{1}{2}\right)^{10} \cdot \left(\tfrac{1}{2}\right)^0$$

$$= 10 \cdot \left(\tfrac{1}{2}\right)^{10} + \left(\tfrac{1}{2}\right)^{10}$$

$$= \left(\tfrac{1}{2}\right)^{10} \times 11$$

$$= \frac{11}{1024}$$

1%

$\frac{1}{100}$

$= 0.01$

$$= 0.01074 > 0.01$$

Since, the value of type- 1 error is greater then 0.01 error.

Test statistics:

A statistics a function of sample observation, used ~~the~~ to test the null hypothesis is known as a test statistic. A concludtion to reject or not to reject H_0 is baised on the value of the test statistic, observed in a sample selected.

Critical Region:

The entire ~~set of~~ posible values of the test statictics is devided into 2 sets of region. One region consisting of the values that supont the alternative hypothesis & leds to rejection of H_0 is called the critical region or region of rejection. The other consisting of the values that suport H_0 is called the non-rejection region. The value of test statistic which lies to the bounded critical region & accept ~~the critical value~~ once. Region is know a critical region. The critical value dependent on the level of significance (α).

1 sample mean test:

Statement: To test the null hypothesis, $H_0 : \mu = \mu_0$. μ_0 is a specified value of μ.

Against the alternatives,

(i) $H_1 : \mu \neq \mu_0$.

(ii) $H_1 : \mu > \mu_0$.

(iii) $H_1 : \mu < \mu_0$.

Case - 1 : σ^2 known

Let, $x_1, x_2, \ldots\ldots, x_n$ be a random sample of size 'n' drawn from $N(\mu, \sigma^2)$ population. To test the hypothesis $H_0 : \mu = \mu_0$.

Against the above ~~alternative~~ mention alternatives we use, $\bar{x} = \frac{1}{n}\sum x_i$, sample mean as an appropriate statistic of population mean μ.

we known that $\bar{x} \sim N(\mu, \frac{\sigma^2}{n})$

Here, the test statistic defined by

$$\tau \text{ or } Z = \frac{\bar{x} - \bar{\mu}}{\sigma/\sqrt{n}} \sim N(0,1)$$

under $H_0 : \mu = \mu_0$ the test statistic is

$$Z = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}}$$

decision:

(i) If $|Z| \geq Z_{\alpha/2}$ we reject $H_0 : \mu = \mu_0$ aginst the alternative $H_1 : \mu \neq \mu_0$ otherwise we may accept the null hypothesis.

Two test

(ii) If $Z \geq Z_\alpha$ we reject $H_0 : \mu = \mu_0$ against the alternative $H_1 : \mu > \mu_0$, othewise we may accept the null hypothesis.

one test

(iii) If $Z \leq -Z_\alpha$ we reject $H_0 : \mu = \mu_0$ against the alternative $H_1 : \mu < \mu_0$, otherwis we may accept the null hypothesis.

Case - 2 : (σ^2 is Unknown)

If σ^2 is unknown we may estimate σ^2 in two ways, either we can use the sample variance $s^2 = \frac{1}{n}\sum(x_i - \bar{x})^2$ or the unbiased estimate for population variance

$$S^2 = \frac{1}{n-1}\sum_i (x_i - \bar{x})^2.$$

If we used sample variance as an estimate of σ^2 then the statistic under Ho is defind by $t = \frac{\bar{x} - \mu_0}{s/\sqrt{n-1}} \sim t_{n-1}$ d.f.

If one used the unbiased estimator of population variance as an estimator of population variance then the test ~~statics~~ statistic

$$t = \frac{\bar{x} - \mu_0}{S/\sqrt{n}} \sim t_{n-1} \text{ d.f.}$$

Decision :

(i) If $|t| \geq t_{\alpha/2, (n-1)}$ we reject $H_0 = \mu = \mu_0$, against the alternative $H_1 : \mu \neq \mu_0$. otherwise we may accept the null hypothesis.

(ii) If $t \geq t_\alpha$ we reject $H_0 : \mu = \mu_0$ against the alternative $H_1 : \mu > \mu_0$, otherwise we may accept the null hypothesis.

(iii) If $t \leq -t_{\alpha, n-1}$ we reject $H_0 : \mu = \mu_0$ against the alternative $H_1 : \mu < \mu_0$, otherwise we may accept the null hypothesis.

Two sample mean test:

A	B
μ_1	μ_2
σ_1^2	σ_2^2

Let, $x_{11}, x_{12}, \ldots x_{1n}$ be a random sample of size n_1 draw from $N(\mu_1, \sigma_1^2)$ population.

Let, $x_{21}, x_{22}, \ldots x_{2n}$ be a another random sample of size n_2 draw from $N(\mu_2, \sigma_2^2)$ population.

Both the sample are independent. Here, we one to test null hypothesis $H_0: \mu_1 = \mu_2$, against, any one of the following there alternative hypothesis $H_1: \mu_1 \neq \mu_2$, $H_1: \mu_1 > \mu_2$, $H_1: \mu_1 < \mu_2$. at the pre-spacifide value of α the null hypothesis $H_0: \mu_1 = \mu_2$ alternatively expresed as $H_0: \mu_1 - \mu_2 = 0$. Here, the sample mean $\bar{x}_1 \sim N(\mu_1, \sigma_1^2/n_1)$ & $\bar{x}_2 \sim N(\mu_2, \sigma_2^2/n_2)$ to test the hypothesis we can take $(\bar{x}_1 - \bar{x}_2)$ as an approporeate estimator of $(\mu_1 - \mu_2)$ and

$$\bar{x}_1 - \bar{x}_2 \sim N\left(\mu_1 - \mu_2, \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}\right)$$

Case-1: (σ_1^2, σ_2^2 are known)

In this case we use the test statistics

$$z = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \sim N(0,1)$$

Under $H_0: \mu_1 = \mu_2$,

the test statistics becomes,

$$Z = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

Decision :

(i) If $|Z| \geq Z_{\alpha/2}$ we reject $H_0: \mu_1 = \mu_2$ against the alternative $H_1: \mu_1 \neq \mu_2$, otherwise we accept $(H_0: \mu_1 = \mu_2)$ the null hypothesis.

(ii) If $Z \geq Z_\alpha$ we reject $H_0: \mu_1 = \mu_2$, against the alternative $H_1: \mu_1 > \mu_2$, otherwise we accept the null hypothesis.

(iii) If $Z \leq -Z_\alpha$ we reject $H_0: \mu_1 = \mu_2$, against the alternative $H_1: \mu_1 < \mu_2$, otherwise we accept the null hypothesis.

Case-2 (σ_1^2 & σ_2^2 are unknown):

σ_1^2 & σ_2^2 are unknown but known to be equal i.e $\sigma_1^2 = \sigma_2^2 = \sigma^2$.

Here, we use a pulled estimator of σ^2

i.e $\hat{\sigma}^2 = s^2 = \dfrac{\Sigma(x_{1i} - \bar{x}_1)^2 + \Sigma(x_{2j} - \bar{x}_2)^2}{n_1 + n_2 - 2}$ → (dependent)

To test the hypothesis we used the test statistic

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{s\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \sim t_{n_1 + n_2 - 2}.$$

Under null hypothesis it becomes

$$\frac{(\bar{x}_1 - \bar{x}_2)}{s\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}.$$

Note:

(i) If we define sample variance of the first sample as $s_1^2 = \frac{1}{n} \sum (x_{1i} - \bar{x}_1)^2$

and for 2nd sample $s_2^2 = \frac{1}{n_2} \sum (x_{2j} - \bar{x}_2)^2$.

then S^2, i.e. pooled estimator of σ^2 is,

$$S^2 = \frac{n_1 s_1^2 + n_2 s_2^2}{n_1 + n_2 - 2}$$

(ii) If we define the unbaised estimator of population variance of the 1st population as

$$s_1^2 = \frac{1}{n_1 - 1} \sum (x_{1i} - \bar{x}_1)^2$$

& for the 2nd population,

$$s_2^2 = \frac{1}{n_2 - 1} \sum (x_{2j} - \bar{x}_2)^2$$

then, S^2, i.e pooled estimator of σ^2 is,

$$S^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}$$

Decison:

(i) If $|t| \geq t_{\alpha/2, n_1+n_2-2}$ we reject $H_0: \mu_1 = \mu_2$, against $H_1: \mu_1 \neq \mu_2$, otherwise we may accept the null hypothesis.

(ii) If $t \geq t_{\alpha, n_1+n_2-2}$ we reject $H_0: \mu_1 = \mu_2$, against $H_1: \mu_1 > \mu_2$, otherwise we may accept the null hypothesis.

(iii) If $t \leq -t_{\alpha, n_1+n_2-2}$ we reject $H_0: \mu_1 = \mu_2$, against $H_1: \mu_1 < \mu_2$, otherwise we may accept the null hypothesis.

Two sample variance Test:

let, $x_1, x_2, \ldots x_n$ be a random sample of size n drawn from a normal population with parameter μ & σ^2. Here we are to test null hypothesis $H_0 : \sigma^2 = \sigma_0^2$, against any one of the following three alternative hypothesis

$H_1 : \sigma^2 \neq \sigma_0^2$; $H_1 : \sigma^2 > \sigma_0^2$; $H_1 : \sigma_0^2 < \sigma_0^2$.

Case-1 : (Unknown) :

To test the hypothesis we use the test statistic $\chi^2 = \frac{\sum(x_i - \mu)^2}{\sigma^2} \sim \chi^2_n$, under $H_0 : \sigma^2 = \sigma_0^2$ the test statistics becomes $\chi^2 = \frac{\sum(x_i - \mu)^2}{\sigma_0^2} \sim \chi^2$

n degrees of freedom.

Decision :

(i) If $\chi^2 \geq \chi^2_{\alpha/2, n}$ or $\chi^2 \leq \chi^2_{1-\alpha/2}$, we reject $H_0 : \sigma^2 = \sigma_0^2$, against the alternative hypothesis $H_1 : \sigma^2 \neq \sigma_0^2$, otherwise we accept the null hypothesis.

(ii) If $\chi^2 \geq \chi^2_{\alpha, n}$ we reject $H_0 : \sigma^2 = \sigma_0^2$, against the alternative hypothesis $H_1 : \sigma^2 > \sigma_0^2$, otherwise we may accept the null hypothesis.

(iii) If $\chi^2 \leq \chi^2_{1-\alpha, n}$, we reject the null hytusis, against the alternative hypothesis $H_1 : \sigma^2 < \sigma_0^2$, otherwise we may accept the null hypothesis.

Case - II : (μ is known) :

(i) If μ is unknown then it is estimated by sample mean & it define by $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$.

(ii) Here test the null hypothesis $H_0 : \sigma^2 = \sigma_0^2$ we use the test statistics

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

under H_0 the test statistics becomes

$$\chi^2 = \frac{\sum (x_i - \bar{x})^2}{\sigma_0^2} \sim \chi^2_{n}$$

Decision:

(i) If $\chi^2 \geq \chi^2_{\alpha/2, n-1}$ or $\chi^2 \leq$ 1-

we reject $H_0 : \sigma^2 = \sigma_0^2$ against the alternative hypothesis $H_1 : \sigma^2 \neq \sigma_0^2$, otherwise we accept the null hypothesis.

(ii) If $\chi^2 \geq \chi^2_{\alpha, n-1}$ we reject $H_0 : \sigma^2 = \sigma_0^2$ against $H_1 : \sigma^2 > \sigma_0^2$, otherwise we accept the null hypothesis.

(iii) If $\chi^2 \leq \chi^2_{1-\alpha, n-1}$, we reject the null hypothesis against $H_1 : \sigma^2 < \sigma_0^2$, otherwise we may accept the null hypothesis.

* (6) **Theorem :**

Let $x_1, x_2, \ldots, x_n$ be a r.s of size 'n' drawn from a Normal population with mean μ & var σ^2, then

(i) $\bar{x} \sim N(\mu, \frac{\sigma^2}{n})$

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

(iii) $\bar{x}$ & $\sum_{i=1}^{n} \left(\frac{x_i - \bar{x}}{\sigma}\right)^2 = \sum_{i=2}^{n} \frac{y_i^2}{\sigma^2}$ are indepedent

$\boxed{dF(x) = f(x)\,dx}$

$\rightarrow$ The joint probability differenciation of $x_1, x_2, \ldots x_n$.

$$dP(x_1, x_2, \ldots x_n) = \left(\frac{1}{\sigma\sqrt{2\pi}}\right)^n \cdot e^{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i - \mu)^2} dx_1 dx_2 \cdots dx_n$$

Now, let us transform to the variables $Y_i\,(i = 1, 2, \ldots, n)$ by mean of a linier orthogonal transformation of the form $Y = AX$

~~let, us chose in particular,~~

$\boxed{\sum y_i^2 = \sum x_i^2 \quad \sum_{i=1}^{n} a_{ij} = 1}$

Ref

$$y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}, \quad A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix} \quad x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

$$y_1 = a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n$$

$$y_2 = a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n$$

$$\vdots \quad \cdots \quad \cdots$$

$$y_n = a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n.$$

let, us chose in particular,

$a_{11} = a_{12} = \cdots \cdots = a_{1n} = \frac{1}{\sqrt{n}}$

$\therefore y_1 = \frac{x_1 + x_2 + \cdots + x_n}{\sqrt{n}}$

or, $y_1 = \frac{n\bar{x}}{\sqrt{n}} = \sqrt{n}\,\bar{x}$

$\therefore dy_1 = \sqrt{n}\,d\bar{x}$ ——— ①

$\because$ the transformation is orthogonal,

$\therefore \sum_{i=1}^{n} y_i^2 = \sum_{i=1}^{n} x_i^2 - n\bar{x}^2 + n\bar{x}^2)$

$= \sum_{i=1}^{n} (x_i - \bar{x})^2 + y_1^2$

or, $\sum_{i=1}^{n} (x_i - \bar{x})^2 = \sum_{i=2}^{n} y_i^2$ ——— ⑪

$\sum (x_i - \bar{x})^2$

$= \sum (x_i^2 - 2x_i\bar{x} + \bar{x}^2)$

$= \sum (x_i^2 - 2\bar{x}\sum x_i + n\bar{x}^2)$

$= \sum x_i^2 - 2n\bar{x}^2 + n\bar{x}^2$

$= \sum x_i^2 - n\bar{x}^2$

Now,

$\sum_{i=1}^{n} (x_i - \mu)^2 = \sum_{i=1}^{n} \{(x_i - \bar{x}) + (\bar{x} - \mu)\}^2$

$= \sum_{i=1}^{n} (x_i - \bar{x})^2 + n(\bar{x} - \mu) \cdot$ [$\because$ sum of deviation from mean is zero]

$= \sum_{i=2}^{n} y_i^2 + n(\bar{x} - \mu)$ ($\because$ by ⑪)

the jacobian of the transformation is $J = \pm 1$

$\therefore |J| = 1$

$\therefore$ The join probabily deffenentical $Y_1, Y_2, \cdots Y_n$ is,

$dG(y_1, y_2 \cdots y_n)$

$= \left(\frac{1}{\sigma\sqrt{2\pi}}\right)^n \cdot e^{-\frac{1}{2\sigma^2}\left\{\sum_{i=2}^{n} y_i^2 + n(\bar{x} - \mu)^2\right\}} \cdot 1 \cdot dy_1\, dy_2 \cdots dy_n$

$= \left(\frac{1}{\sigma\sqrt{2\pi}}\right)^n \cdot e^{-\frac{n(\bar{x} - \mu)^2}{2\sigma^2}} \cdot e^{-\frac{\sum_{i=2}^{n} y_i^2}{2\sigma^2}} \cdot \sqrt{n}\,d\bar{x} \cdot dy_2 \cdots dy_n$

$$= \left\{ \frac{1}{\sigma/\sqrt{n}\sqrt{2\pi}} \cdot e^{-\frac{1}{2}\left(\frac{\bar{x}-\mu}{\sigma^2/n}\right)^2} d\bar{x} \right\} \left\{ \left(\frac{1}{\sigma\sqrt{2\pi}}\right)^{n-1} \cdot e^{-\frac{1}{2}\sum_{i=2}^{n} y_i^2/\sigma^2} dy_2 \cdots dy_n \right\}$$

$$= \left\{ g_1(\bar{x})\, d\bar{x} \right\} \left\{ g_2\left(\sum_{i=2}^{n} y_i^2/\sigma^2\right) dy_2 \cdots dy_n \right\}$$

i.e $\bar{X}$ & $\sum_{i=2}^{n}(y_i^2/\sigma^2) = \sum_{i=1}^{n}\left(\frac{x_i-\bar{X}}{\sigma}\right)^2$ are independently distributed.

Moreover, $\bar{x} \sim N(\mu, \frac{\sigma^2}{n})$ (1st part)

Again, y_i $(i=2,3,\ldots n)$ are independent normal variat with mean 0 & var σ^2.

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

<u>Note:</u> The sample variance is defined by

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

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

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

The unbaised estimate of population variance is define by

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

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

$$\boxed{\therefore \frac{(n-1)S^2}{\sigma^2} \sim \chi^2_{n-1}}$$

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

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

$\neq \sigma^2$

(1) Let, $x_{11}, x_{12}, \cdots, x_{1n_1}$ be a r.s of size n_1 drawn from a normal population with mean μ_1 & unknown var σ_1^2.

$\Rightarrow$ Let, us consider another independent r.s of size n_2 cases of the unique $x_{21}, x_{22}, \cdots, x_{2n_2}$ from another normal population with mean μ_2 & unknown var σ_2^2.

$\rightarrow$ Here, we are to test the hypothesis $H_0 : \sigma_1^2 = \sigma_2^2$ against any one of the following three alternatives : $H_1 : \sigma_1^2 \neq \sigma_2^2$, $H_1 : \sigma_1^2 > \sigma_2^2$, $H_1 : \sigma_1^2 < \sigma_2^2$.

at any pre-specify level of α.

Case-1 : (μ_1 & μ_2 are known)

If μ_1 & μ_2 are known then to test the null hypothesis $H_0 : \sigma_1^2 = \sigma_2^2$ we may use the test statistics.

$$F = \frac{\dfrac{\sum (x_{1i} - \mu_1)^2}{\sigma_1^2} \Big/ n_1}{\dfrac{\sum (x_{2j} - \mu_2)^2}{\sigma_2^2} \Big/ n_2}$$

Under $H_0 : \sigma_1^2 = \sigma_2^2$, F becomes

$$F = \frac{\sum_i (x_{1i} - \mu_1)^2}{\sum_j (x_{2j} - \mu_2)^2} \cdot \frac{n_2}{n_1} \sim F_{n_1, n_2}$$

Decision :

(i) If $F \geq F_{\alpha/2, n_1, n_2}$ or $F \leq F_{1-\alpha/2, n_1, n_2}$ we reject $H_0 : \sigma^2 = \sigma_1^2$ against the alternative $H_1 : \sigma_1^2 \neq \sigma_2^2$, otherwi-se we accept the null hypothesis.

(ii) If $F \geq F_{\alpha/2}, n_1, n_2$ $H_0 : \sigma_1^2 = \sigma_2^2$ against the alternative $H_1 : \sigma_1^2 > \sigma_2^2$ otherwise we accept the null hypothesis.

(iii) If $F \geq F_{1-\alpha/2}, n_1, n_2$ we reject $H_0 : \sigma_1^2 = \sigma_2^2$ against, the alternative $H_1 : \sigma_1^2 < \sigma_2^2$ otherwise we accept the null hypothesis.

Case-II: (μ_1 & μ_2 are unknown)

If μ_1 & μ_2 are unknown then they are estimated by there respective sample mean that is $\bar{x}_1$ & $\bar{x}_2$. Here we use to statistics,