SV
7/4/2021

Testing:

UNIT – IV

(1) Large Sample test → Proportions, Means

(2) Small Sample test → Means, Variance

(3) Chisquare Test

SU

Procedure for Testing Hypothesis:-

Step-I:- Hypothesis [Assumptions]

Null Hypothesis (H_0 or H_N):

$LL_c = LL_U$ [Region of rejection is known as N.H)

Alternative Hypothesis (H_1 or H_A):

Converse of null H is known as A.H

$A \neq B$

$L_c \neq L_U$ → Two Tailed Test

$A < B$

$A > B$

(Right) $L_c > L_U$ → Single tailed Test (one tailed test)

$L_c < L_U$ → (left)

II : degrees of freedom (d.f):

" The no of independent observations are actively participating of Testing Hypothesis".

eg:- if size is n ⇒ d.f is "n-1"

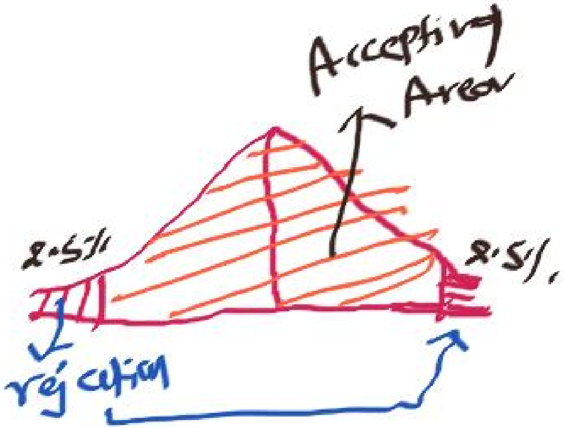
III :- Levels of Significance (LOS)

90%, 95%, 98%, and 99%.

↳ Accepting

10%, 5%, 2%, & 1%.

↳ rejection



IV : Test statistics

$\boxed{(n) < 30}$ ⇒ Small sample test

If the number of observations (n) less than 30 then it is known as small sample test.

Defined as

$$\boxed{t = \frac{x - E(x)}{S.E(x)} \overset{df}{\sim} t_{n-1}}$$

$SE = \sqrt{\text{Sample Variance}}$

Standard Error

Calculated Part

Tabulated Part

Comparision b/w

Large sample test:-

If the no of observations $\boxed{(n) \geq 30}$, then the test is known as Large sample test

$$\boxed{Z = \frac{x - E(x)}{s.d(x)} \sim N(0,1)}$$

Calculated part

Tabulated part

V Calculations / procedure

VI Inference / conclusion

If Calculated value $<$ Table value at different d.f in different L.O.S our Null Hypothesis (Ho) will accepted otherwise

Calculated value $>$ Table value at different d.f in different L.O.S our Null Hypothesis will be rejected

Types of Errors :-

Type-I Error :- [Producer's Risk]

Def :- Rejecting H_0 when H_0 is "true" is known as Type-I error $\Rightarrow$ which is denoted by "α".

Type-II error :- [Consumer's Risk]

Accepting H_0 when H_0 is "false". is known as Type-II error

Large sample test table values $\left[z = \frac{x - E(x)}{s.d(x)} \right]$

Two tailed Table value

LOS:		
90%	10%	1.645
95%	5%	1.96
98%	2%	2.33
99%	1%	2.58

NOTATIONS

	Sample	Population
Size	n	N
Mean	$\bar{x}$	μ
Variance	s^2	σ^2
Proportion	p	P

Sample: -

A part of the population is known as sample.

Estimation: -

Studying samples and drawing the conclusions about the population is known as estimation:

Types of Estimation: -

(1) Point Estimation (2) Interval Estimation

→ By studying single observation and drawing the conclusion about populations is known as "Point Estimation"

⇒ By studying the set of observations and drawing the conclusion about a population is known as "interval Estimation"

N

SV

27/4/21
2.20 to 4.30 PM

Testing the significance difference in single sample mean (Large sample σ – known)

$H_0: \bar{x} = \mu$ [i.e: There is no significant difference b/w sample mean & population]

$H_1: \bar{x} \neq \mu$ (Two tailed test)

LOS: at different levels.

Test Statistics:-

calculated

$$Z = \frac{x - E(x)}{S.d} \sim N(0,1)$$

→ Table value

Procedure:- Let $x_1, x_2, \ldots x_n$ are n-random samples drawn from the same normal population Mean 'μ' and variance 'σ^2/n'. $\left[\because \bar{x} \sim N(\mu, \sigma^2/n)\right]$

Here we are testing our N.H using Large sample tests.

But r.v is single sample mean $(\bar{x})$

$$Z = \frac{\bar{x} - \overset{?}{E(\bar{x})}}{?\; s.d(\bar{x})} \sim N(0,1)$$

Now $E(\bar{x}) = \mu$ $\quad\quad (\because \bar{x} \sim N(\mu, \sigma^2/n))$

$Var(\bar{x}) = \sigma^2/n$ $\quad\quad (\quad\quad "\quad\quad)$

Finally

$$Z = \frac{\bar{x} - \mu}{\sqrt{\sigma^2/n}} \sim N(0,1)$$

Sample mean

population Mean

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

s.d

size

Inference :- Since $Z_{cal} < Z_{tab}$ our H_0 will accepted otherwise rejected.

Testing the significance difference b/w two sample Mean (Large sample test)

$H_0: \mu_1 = \mu_2$

$H_1: \mu_1 \neq \mu_2$

LOS:- at different levels

Test statistics:-

$$Z = \frac{x - E(x)}{S.d(x)} \sim N(0,1)$$

Procedure :-

Let x_1, x_2 are random samples with size n_1, n_2 drawn from the two different Normal population with means μ_1, μ_2 and variances $\frac{\sigma_1^2}{n_1}, \frac{\sigma_2^2}{n_2}$ respectively.

Here, testing N.H using Large sample tests.

Now, r.v is difference b/w two sample means $(\bar{x}_1 - \bar{x}_2)$

Test statistics $$Z = \frac{(\bar{x}_1 - \bar{x}_2) - (E(\bar{x}_1 - \bar{x}_2))}{\sqrt{Var(\bar{x}_1 - \bar{x}_2)}} \sim N(0,1)$$

let us determine $E(\bar{x}_1 - \bar{x}_2)$ & $V(\bar{x}_1 - \bar{x}_2)$

$\left(\bar{X}_1 \sim N(\mu_1, \sigma_1^2/n_1) \right.$

$\bar{x}_2 \sim N(\mu_2, \sigma_2^2/n_2))$

$(H_0 ; \mu_1 = \mu_2)$

Now $E(\bar{x}_1 - \bar{x}_2) = E(\bar{x}_1) - E(\bar{x}_2)$

$= \mu_1 - \mu_2$

$= 0$

$$V(\bar{x}_1 - \bar{x}_2) = V(\bar{x}_1) + V(\bar{x}_2)$$

$$= \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}$$

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

finally, test statistics rewritten as

$$\left[Z = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \sim N(0,1) \right]$$

Case:- $\sigma_1^2 = \sigma_2^2$

$$\left[Z = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\sigma^2 \left(1/n_1 + 1/n_2 \right)}} \sim N(0,1) \right]$$

Inference :- Since Z cal < Z tabulated, null hypothesis will be accepted.
otherwise rejected.

—— X ——

Testing the significance difference in single sample Proportion (Large sample test)

(ratio)

$H_0: p = P$

$H_1: p \neq P$

Los: at different levels

Test statistics:-

$$Z = \frac{x - E(x)}{s.d(x)} \sim N(0,1)$$

<u>Procedure:-</u> let 'x' is a random sample size 'n' following Binomial disn with parameter (n, P)

Here, we are N.H using large sample test and also r.v is single sample proportion (p) $\left(p = \frac{x}{n}\right)$

$$\begin{bmatrix} E(x) = n \cdot P \\ V(x) = nPQ \end{bmatrix}$$

$$Z = \frac{p - \overset{?}{E}(p)}{?\sqrt{V(p)}} \sim N(0,1)$$

let us compute the values $E(p)$ & $V(p)$

Now $E(p) = E(x/n) = \frac{1}{n} E(x)$
$= \frac{1}{\cancel{n}}(\cancel{n}P) = P$

$$V(p) = V(x/n) = \frac{1}{n^2} V(x)$$

$$= \frac{1}{n^{\cancel{2}}} (\cancel{n} PQ) = \frac{PQ}{n}$$

There fore, test statistics is

$$\boxed{Z = \frac{p - P}{\sqrt{\frac{PQ}{n}}} \sim N(0,1) \qquad Q = 1 - P} \quad \left(p = \frac{x}{n}\right)$$

if 'p' is not given in the data, we should consider "P = 0.5"

Inference:- Since $Z_{cal} < Z_{tab}$, N.H will be accepted. otherwise rejected.

— x —

Testing the significance differences b/w two sample proportions

(Large sample test):

$H_0: P_1 = P_2 = P$

$H_1: P_1 \neq P_2$

Los: at different levels.

Test statistics:

$$Z = \frac{x - E(x)}{s.d(x)} \sim N(0,1).$$

Procedure:- let x_1, x_2 are two random samples with sizes n_1, n_2 drawn from the same Binomial population the population parameter P, Q:

Here, we are testing N.H using Large sample test

But our r.v is difference b/w two sample proportion denoted by $(p_1 - p_2)$,

Now the test statistics for Large sample defined as

$$Z = \frac{(p_1 - p_2) - E(p_1 - p_2)}{\sqrt{Var(p_1 - p_2)}} \sim N(0,1)$$

let us find the values of $E(p_1 - p_2)$ and $Var(p_1 - p_2)$

$$\left[\begin{array}{l} p_1 = \frac{x_1}{n_1} \\ p_2 = x_2/n_2 \end{array}\right]$$

$$E(p_1 - p_2) = E(p_1) - E(p_2)$$

$$= E\left(\frac{x_1}{n_1}\right) - E\left(\frac{x_2}{n_2}\right)$$

$$= \frac{1}{n_1} E(x_1) - \frac{1}{n_2} E(x_2)$$

$$= \frac{1}{\cancel{n_1}} \cancel{n_1} P - \frac{1}{\cancel{n_2}} \cancel{n_2} P = 0$$

$$
\begin{aligned}
Var(p_1 - p_2) &= Var(p_1) + Var(p_2) \\
&= Var(x_1/n_1) + Var(x_2/n_2) \\
&= \frac{1}{n_1^2} Var(x_1) + \frac{1}{n_2^2} Var(x_2) \\
&= \frac{1}{n_1^{\cancel{2}}} (\cancel{n_1} PQ) + \frac{1}{n_2^{\cancel{2}}} (\cancel{n_2} PQ) \\
&= \frac{PQ}{n_1} + \frac{PQ}{n_2} \\
&= PQ(1/n_1 + 1/n_2)
\end{aligned}
$$

$$Var(p_1 - p_2) = PQ\left(\frac{1}{n_1} + \frac{1}{n_2}\right)$$

Therefore, the test statistics of difference b/w two proportion is

$$\boxed{Z = \frac{p_1 - p_2}{\sqrt{PQ\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} \sim N(0,1)}$$

Here $p_1 = \frac{x_1}{n_1}$, $p_2 = \frac{x_2}{n_2}$

$$\hat{P} = \frac{x_1 + x_2}{n_1 + n_2} \text{ or } \frac{n_1 p_1 + n_2 p_2}{n_1 + n_2}$$

$$Q = 1 - P$$

Inference:- Since $Z_{cal} < Z_{tab}$, N.H will be accepted, otherwise rejected.

Problems :-

① The average life of the computer is 1150 Hours and s.d is 42 Hours. A random samples 500 computer were tested whose average life is 1155 Hours

Test where overall average life is significant or not at 5% LOS ?

Sol:- $H_0: \mu = 1150$ → pop Mean

$H_1: \mu \neq 1150$

LOS: 0.05 ✓

Test statistics:-

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

data:-

$\bar{x} = 1155$ $\mu = 1150$

$\sigma = 42$ $n = 800$

Calculation

$$Z_{cal} = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}} = \frac{1155 - 1150}{42/\sqrt{800}} = 3.36$$

$$Z_{cal} = 3.36$$
$$Z_{tab} = 1.96$$

Inference:- Since $Z_{cal} > Z_{tab}$ at 5% LOS.

our N.H will be rejected.

i.e average life of computers is not equal to

1150 Hours ?

Prob:- Average Indian life expectency is 72 years accouding medial station. A random sample 50 people observed whose averege lif was 71.9 years and s.d is 11.5 years. Test whore average life of Indians is significant or not at 5% LOS?

Sol:- $H_0 : \mu = 72$

$H_1 : \mu \neq 72$

LOS; 5% LOS = 0.05

Test statistics:

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

data:-

$\bar{x} = 71.9 \quad \mu = 72 \quad \sigma = 11.5 \quad n = 500$

Calculation:-

$$Z_{cal} = \frac{\bar{x} - \mu}{\sigma / \sqrt{n}} = \frac{71.9 - 72}{11.5 / \sqrt{500}} = -0.194$$

$$|Z_{cal}| = 0.194$$

Z_{tab} at 5% LOS $= 1.96$

$|Z_{cal}| = 2.194$

Inference:- Since $Z_{cal} < Z_{tab}$, our N.H will be accepted.

i.e The average Indian life is 72 years.

③ The average length of bolts is (3.6) inches and s.d is (0.3) in, another sample average length is (3.72) inch and s.d (1.1) inch whose sample size is (55). Test whether both samples are drawn from same population or not?

Sol:— $H_0 : \mu_1 = \mu_2$

$H_1 : \mu_1 \neq \mu_2$

LOS: 5% LOS or 0.05

Test statistics:-

$$Z = \frac{\overline{x_1} - \overline{x_2}}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \sim N(0,1)$$

data:

$\overline{x_1} = 3.6$ $\sigma_1 = 0.35$

$\overline{x_2} = 3.72$ $\sigma_2 = 1.1$

$n_1 = n_2 = 55$

Calculation

$$Z_{cal} = \frac{\overline{x_1} - \overline{x_2}}{\sqrt{\frac{\sigma_1^2 + \sigma_2^2}{n}}}$$

$$Z_{cal} = \frac{3.6 - 3.72}{\sqrt{\frac{(0.35)^2 + (1.1)^2}{55}}}$$

$= -3$

$$|Z_{cal}| = 3$$

$$\left.\begin{array}{c} Z_{tab} \\ \text{at } 5\% \\ \text{LOS} \end{array}\right\} = 1.96$$

Inference:- Since $Z_{cal} > Z_{tab}$ at 5% LOS,

our H_0 will be rejected.

i.e both samples are not belongs same population

—x—

SW
4/5/21
2-20 to
430PM

Single Sample proportion:

$$Z = \frac{p - P}{\sqrt{PQ/n}} \sim N(0,1)$$

Here $p = x/n$

Two Sample Proportion:

$$Z = \frac{p_1 - p_2}{\sqrt{PQ\left(1/n_1 + 1/n_2\right)}} \sim N(0,1)$$

$p_1 = \frac{x_1}{n_1}$ $p_2 = \frac{x_2}{n_2}$

$$P = \frac{x_1 + x_2}{n_1 + n_2} = \frac{n_1 p_1 + n_2 p_2}{n_1 + n_2}$$

$$Q = 1 - P$$

— x —

Problems on proportion test:

A random sample of 400 items, out of 39 are defective, company claims that 6.5% are defective. Check whether company claim can be accepted or not at 5% ?

Sol

$H_0: P = 0.065$

$H_1: P \neq 0.065$

LOS: 5% or 0.05

Test statistics:

$$Z = \frac{p - P}{\sqrt{PQ/n}} \sim N(0,1)$$

data:- $x = 39$: $n = 400$

$P = 0.065$ $\quad Q = 1 - 0.065$

$= 0.935$

Cal :-

$$Z_{cal} = \frac{p - P}{\sqrt{PQ/n}} \sim N(0,1)$$

$$p = \frac{x}{n} = \frac{39}{400} = 0.0975$$

$$Z_{cal} = \frac{0.0975 - 0.0625}{\sqrt{\frac{0.0625 \times 0.935}{400}}} = 3.27$$

$$\left.\begin{matrix} Z_{tab} \\ \text{at } 5\% \\ LOS \end{matrix}\right\} = 1.96$$

Inferenca :-

Since $Z_{cal} > Z_{tab}$ at 5% LOS our H_0 will be rejected.

$\therefore P \neq 0.0625$.

In a T|F a students answers 70 ques correct out of 130. Whether True | False question are uniformly distributed or not?

Sol:- $H_0: P = 0.5$

$H_1: P \neq 0.5$

LOS: 5% or 0.05

Test statistics:

$$Z = \frac{p - P}{\sqrt{PQ/n}} \sim N(0,1)$$

data:

$n = 130$: $x = 70$,

$$p = \frac{x}{n} = \frac{70}{130} = 0.538$$

$P = 0.5$, $Q = 0.5$

$$Z_{cal} = \frac{p - P}{\sqrt{PQ/n}} = \frac{0.538 - 0.5}{\sqrt{\frac{0.5 \times 0.5}{130}}}$$

$Z_{cal} = 0.883$

$Z_{cal} = 0.883$

$\left.\begin{array}{l} Z_{tab} \\ \text{at } 5\% \text{ LOS} \end{array}\right\} = 1.96$

Inferene :- Since $Z_{cal} < Z_{tab}$ at 5% LOS, our N.H will be accepted.

i.e The true/False question equal.

A dice rolled 350 times, out of 210 times face 3 or 4 are occred. Check whether die is unbaised or not?

Sol:- $H_0: P = \left[\frac{1}{6} + \frac{1}{6}\right] = 2/6 =$

$H_1: P \neq 2/6$

LOS: 5% LOS

Test statistics:

$$Z = \frac{p - P}{\sqrt{PQ/n}} \sim N(0,1)$$

data:

$x = 210 ; \quad n = 350$

$p = \frac{x}{n} = \frac{210}{350} = 0.6$

$P = 2/6 = 0.33$

$Q = 0.67$

$$Z_{cal} = \frac{p - P}{\sqrt{\frac{PQ}{n}}} = \frac{0.6 - 0.33}{\sqrt{\frac{0.33 \times 0.67}{350}}} =$$

$= 10.75$

$Z_{tab} = 1.96$

5% LOS

Inferene: $Z_{cal} > Z_{tab}$ at 5% LOS our Ho will be rejected.

i.e die is not unbaised.

— x —

A survey is conduted regarding online examination system. 300 boys are favour out of 320 and 290 Girls are favour out of 350. Check whether opinion is uniform or not?

Sol:- $H_0: P_1 = P_2$ $H_1: P_1 \neq P_2$

LOS: 5% or 0.05

Test statistics:

$$Z = \frac{p_1 - p_2}{\sqrt{PQ\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}} \sim N(0,1)$$

data:-

$n_1 = 320$: $x_1 = 300$

$n_2 = 350$: $x_2 = 290$

$$p_1 = \frac{x_1}{n_1} = \frac{300}{320} = 0.937$$

$$p_2 = \frac{x_2}{n_2} = \frac{290}{350} = 0.829$$

$$P = \frac{x_1 + x_2}{n_1 + n_2} = \frac{300+290}{320+350}$$

$$= 0.880$$

$$Q = 1 - 0.880 = 0.12$$

$$Z_{cal} = \frac{p_1 - p_2}{\sqrt{PQ\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}}$$

$$= \frac{0.937 - 0.829}{\sqrt{0.88 \times 0.12\left(\frac{1}{320}+\frac{1}{350}\right)}}$$

$Z_{cal} = 4.32$

$Z_{tab} = 1.96$

Inference: Since $Z_{cal} > Z_{tab}$ at 5% LOS: our H_0 will be rejected.

— x —

	Pro	Number
S_1	36%	480
S_2	45%	520

Test at 5% LOS ?

Sol: $p_1 = 0.36$; $p_2 = 0.45$

$$P = \frac{n_1 p_1 + n_2 p_2}{n_1 + n_2} = \frac{480 \times 0.36 + 520 \times 0.45}{480 + 520}$$

$$= 0.4068$$

$$Q = 1 - 0.4068 = 0.5932$$

$$Z_{cal} = \frac{p_1 - p_2}{\sqrt{PQ\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} = \frac{0.36 - 0.45}{\sqrt{0.4068 \times 0.5932\left(\frac{1}{480} + \frac{1}{520}\right)}}$$

$$= -2.90$$

$$|Z| = 2.90$$

$$Z_{tab} = 1.96$$
at 5%

Inference :- Since $Z_{cal} > Z_{tab}$ at 5% LOS

our Ho will be rejected.

Test the significance difference in single sample Mean (σ unknown) (small test) ($n<30$)

$H_0: \bar{x} = \mu$ (no significance difference b/w S.M and P.M)

$H_1: \bar{x} \neq \mu$

must d.f :- $n-1$

Los: at different levels

Test statistics:

$$t = \frac{x - E(x)}{S.E(x)} \overset{\text{d.f}}{\sim} t_{n-1}$$

Procedure:-

Let $x_1, x_2, --- x_n$ are the random samples drawn from the same normal population with mean μ and variance σ^2/n with same size n.

Here, we are testing N.H using small sample test.

But r.v is <u>single sample Mean $(\bar{x})$</u>

Now the test statistics is

$$t = \frac{\bar{x} - E(\bar{x})}{S.E(\bar{x})} \overset{df}{\sim} t_{n-1}.$$

?

?

Let us find $E(\bar{x})$: $V(\bar{x})$

Now $E(\bar{x}) = \mu$ $\quad \left(\because \bar{x} \sim N\left(\mu, \frac{\sigma^2}{n}\right)\right)$

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

$S \cdot E(\bar{x}) = \sqrt{\sigma^2/n}$

$= \sqrt{s^2/n}$ $\quad$ (σ unknown)

$S \cdot E(\bar{x}) = s/\sqrt{n}$

$$\boxed{\begin{aligned} \bar{x} &= \frac{\sum x_i}{n} \\ s^2 &= \frac{\sum (x_i - \bar{x})^2}{n-1} \end{aligned}}$$

$\therefore$ Test statistics

$$\boxed{t = \frac{\bar{x} - \mu}{s/\sqrt{n}} \overset{df}{\sim} t_{n-1}}$$

Inference:

Since t cal value is < t tab at differnt d.f in diffent LOS our Ho will be accepted otherwise rejected.

Testing the significance differences b/w two sample Mean (small sample) σ Unknown

$H_0: \mu_1 = \mu_2 : \sigma_1^2 = \sigma_2^2$

$H_1: \mu_1 \neq \mu_2 : \sigma_1^2 \neq \sigma_2^2$

d.f:- $n_1 - 1 + n_2 - 1 = n_1 + n_2 - 2$

Los: at different levels.

Test statistics:

$$t = \frac{x - E(x)}{S.E(x)} \overset{df}{\sim} t_{n-1}$$

Procedure:- Let x_1, x_2 are two random samples with sizes n_1, n_2 drawn from two different normal populations with means μ_1, μ_2 and variances $\frac{\sigma_1^2}{n_1}, \frac{\sigma_2^2}{n_2}$ respectively.

Here we are testing H_0 using small sample test.

But our r.v is $(\overline{x_1} - \overline{x_2})$

Now the test statistics is written as

$$t = \frac{(\overline{x_1} - \overline{x_2}) - \underset{?}{E}(\overline{x_1} - \overline{x_2})}{\underset{?}{S.E}(\overline{x_1} - \overline{x_2})} \overset{df}{\sim} t_{n_1+n_2-2}$$

Let us find $E(\overline{x_1}-\overline{x_2})$ and $V(\overline{x_1}-\overline{x_2})$

$\overline{x_1} \sim N(\mu_1, \frac{\sigma_1^2}{n_1})$

$\overline{x_2} \sim N(\mu_2, \frac{\sigma_2^2}{n_2})$

$$E(\overline{x_1}-\overline{x_2}) = E(\overline{x_1}) - E(\overline{x_2})$$

$$= \mu_1 - \mu_2$$

$(H_0 : \mu_1 = \mu_2)$

$$= 0$$

$$V(\overline{x_1}-\overline{x_2}) = V(\overline{x_1}) + V(\overline{x_2})$$

$$= \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}$$

$(H_0 : \sigma_1^2 = \sigma_2^2)$

$$= \sigma^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)$$

$$= s^2\left(\frac{1}{n_1} + 1/n_2\right)$$

$(\sigma$ unknown$)$

Finally the test statistics

$$\boxed{t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{s^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} \overset{df}{\sim} t_{n_1+n_2-2}}$$

Here $\bar{x}_1 = \frac{\Sigma x_{1i}}{n_1}$ $\quad \bar{x}_2 = \frac{\Sigma x_{2j}}{n_2}$

$$\underset{\substack{\Downarrow \\ \text{(pooled sample Variance)}}}{s^2} = \frac{\Sigma(x_{1i}-\bar{x}_1)^2 + \Sigma(x_{2j}-\bar{x}_2)^2}{n_1+n_2-2} = \frac{(n_1-1)s_1^2 + (n_2-1)s_2^2}{n_1+n_2-2}$$

<u>Inference:</u>

Since t cal $<$ ttab at different levels in different d.f's our H_0 will be accepted otherwise rejected.

Equality of Variance (F – test)

$H_0: \sigma_1^2 = \sigma_2^2$

$H_1: \sigma_1^2 \neq \sigma_2^2$

$d.f: (n_1 - 1, n_2 - 1)$

always $F_{cal} > 1$

Los: at diffent level.

Test statistics:

$$F = \frac{S_1^2}{S_2^2} \sim F(\underset{(Nr)}{n_1 - 1}, \underset{(Dr)}{n_2 - 1})$$

s.t

s.t

$Nr > Dr$

(or) $F = \frac{S_2^2}{S_1^2}$

$Nr > Dr$

$$F = \frac{S_2^2}{S_1^2} \sim F(n_2 - 1, n_1 - 1)$$

Inference: Since F cal < Ftab at diffent d.f in different levels our Ho will be accepted otherwise rejected.

— x —

A company claims life of the average life Computer is 1150 hours. A random sample of 6 computer were observed who average lives are 1160, 1175, 1140, 1110, 1200, 1185. Test at 2% LOS whereas company average life is significant or not?

Sol:- $H_0: \mu = 1150$

$H_1: \mu \neq 1150$

d.f: $n = 6$

$n - 1 = 5$

Los: 0.02

Test statistics:

$$t = \frac{\bar{x} - \mu}{s/\sqrt{n}} \overset{df}{\sim} t_{n-1}$$

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

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

Calculations:-

x_i	$(x_i - \bar{x})^2$
1160	
1175	
1140	
1110	
1200	
1185	

$\sum (x_i - \bar{x})^2 = 5333.28$

$\bar{x} = 1161.67$

$$s^2 = \frac{\sum (x_i - \bar{x})^2}{n-1} = \frac{5333.28}{5} = 1066.65$$

$$s = \sqrt{1066.65} = 32.659$$

$$t_{cal} = \frac{\bar{x} - \mu}{s/\sqrt{n}} = \frac{1161.67 - 1150}{32.659/\sqrt{6}}$$

$$= 0.8752$$

$$\left.\begin{matrix} t_{tab} \\ 5\,df \\ 0.02 \end{matrix}\right\} = 3.365$$

Inference: Since $t_{cal} < t_{tab}$ at 5% LOS in 5 df our H_0 will be accepted. i.e the average life compon[...] is 1150 Hours.

05/5/21
JT3 —
— 1PM to
3.20PM

A random sample of 11 students I.Q are 110, 109, 111, 108, 111, 110, 112, 113, 110, 110, 109. The average I.Q level is 111.5, Test whether it is significant or not at 2% LOS?

Sol:-

$H_0 : \mu = 111.5$

$H_1 : \mu \neq 111.5$

d.f: $n = 11$

$n - 1 = 10$

LOS: 0.02 or 2%

Test Statistics

$$t = \frac{\bar{x} - \mu}{s/\sqrt{n}} \overset{df}{\sim} t_{n-1}$$

Cal:

$n = 11$: $\mu = 111.5$

x_i	$(x_i - \bar{x})^2$ $(x_i - 110.27)^2$
110	$(110 - 110.27)^2 =$
109	
111	
108	
111	
110	
112	
113	
110	
110	
109	$(109 - 110.27)^2 =$

$\Sigma x_i = 1213$ $\qquad \Sigma (x_i - \bar{x})^2 = 20.1819$

$$\bar{X} = \frac{1213}{11} = 110.27$$

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

$$= \frac{20.1819}{10}$$

$$s^2 = 2.01819$$

$$s = \sqrt{2.01819}$$

$$= 1.4206$$

$$t_{cal} = \frac{\bar{x} - \mu}{s/\sqrt{n}}$$

$$t_{cal} = \frac{110.27 - 111.5}{1.4206/\sqrt{11}}$$

$$= -2.8719$$

$$|t_{cal}| = 2.8719$$

$\left.\begin{array}{l} t_{tab} \\ 10 \text{ d.f} \\ 0.02 \end{array}\right\} = 2.764$

two tailed test.

Inf:- Since $t_{cal} > t_{tab}$ at 10 d.f in 0.02 LOS, H_0 will be rejected. i.e $\mu \neq 110$.

Test the equality of Means as well as Variance for following data set:

Sales	8	11	15	4	7	6	3	9
Sales after Adv	11	10	15	6	10	8	10	14

Test whether sales are uniform or not at 5% LOS ?

Sol:- $H_0: \mu_1 = \mu_2 : \sigma_1^2 = \sigma_2^2$

$H_1: \mu_1 \neq \mu_2 : \sigma_1^2 \neq \sigma_2^2$

d.f:- $n_1 + n_2 - 2 = 8+8-2 = 14$

LOS: at 5% or 0.05

Test statistics

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{S^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} \text{ df } \sim t_{n_1+n_2-2}$$

Cal:-

x_1	x_2	$(x_1-\bar{x}_1)^2$ $(x_1-7.875)^2$	$(x_2-\bar{x}_2)^2$ $(x_2-10.5)^2$
8	11		
11	10		
15	15		
4	6		
7	10		
6	8		
3	10		
9	14		
63	84	$\Sigma(x_i-\bar{x}_1)^2 = 105.0154$	$\Sigma(x_2-\bar{x}_2)^2 = 60$

$$\bar{x}_1 = \frac{63}{8} = 7.875$$

$$\bar{x}_2 = \frac{84}{8} = 10.5$$

$$S_1^2 = \frac{\Sigma(x_1-\bar{x}_1)^2}{n-1}$$

$$= \frac{105.0154}{7} =$$

$$S_1^2 = 15.022$$

$$S_2^2 = \frac{\Sigma(x_2-\bar{x}_2)^2}{n-1}$$

$$S_2^2 = \frac{60}{7} = 8.5714$$

$$S^2 = \frac{\Sigma(x_i-\bar{x}_1)^2 + \Sigma(x_2-\bar{x}_2)^2}{n_1+n_2-2}$$

$$= \frac{105.0154+60}{14} = 11.786$$

$$t_{cal} = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{s^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} = \frac{7.875 - 10.5}{\sqrt{11.786\left(\frac{1}{8} + \frac{1}{8}\right)}} = -1.529$$

$|t_{cal}| = 1.529$

$\left.\begin{array}{l} t_{tab}\ 14df \\ 0.05\ LOS \end{array}\right\} = 2.145$

at 2 tailed

Inf:- since $t_{cal} < t_{tab}$ at 14 df in 0.05 LOS, Ho will be accepted.

i.e there no significance difference b/w the population Means (Sales _ uniform)

F-test:

$$F = \frac{s_1^2}{s_2^2} \overset{df}{\sim} F(n_1 - 1, n_2 - 1)$$

$N_r > D_r$

$$F_{cal} = \frac{s_1^2}{s_2^2} = \frac{15.022}{8.5714} = 1.7525$$

$\left.\begin{array}{l} F_{tab}\ (7,7) \\ \text{at } 5\%\ LOS \end{array}\right\} = 3.79$

Inf:- since $F_{cal} < F_{tab}$ at $(7,7)$ df in 0.05 LOS, Ho will be accepted.

Test the equality of Means and equality Variances for the following data at 5% LOS:

x_1	7	14	3	8	2	6	5	1	4
x_2	8	3	6	2	1	7	9		

Sol:- $H_0: \mu_1 = \mu_2 : \sigma_1^2 = \sigma_2^2$

$H_1: \mu_1 \neq \mu_2 : \sigma_1^2 \neq \sigma_2^2$

d.f: $n_1 + n_2 - 2 = 9 + 7 - 2$
$= 14$

LOS: 0.05

Te

Test Statistics:-

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{s^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} \overset{df}{\sim} t_{n_1+n_2-2}$$

x_1	x_2	$(x_1-\bar{x}_1)^2$ $(x_1-5.55)^2$	$(x_2-\bar{x}_2)^2$ $(x_2-5.14)^2$
7	8		
14	3		
3	6		
8	2		
2	1		
6	7		
5	9		
1			
4			
50	36	$\Sigma(x_1-\bar{x}_1)^2 = 122.22$	$\Sigma(x_2-\bar{x}_2)^2 = (58.85) =$

$$\bar{x}_1 = \frac{50}{9} = 5.55$$

$$\bar{x}_2 = \frac{36}{7} = 5.14$$

$$\boxed{S_1^2 = \frac{\Sigma(x_1-\bar{x}_1)^2}{n_1-1} = \frac{122.22}{8} = 15.277}$$

$$\boxed{S_2^2 = \frac{\Sigma(x_2-\bar{x}_2)^2}{n_2-1} = \frac{58.85}{6} = 9.8095}$$

$$S^2 = \frac{\Sigma(x_1-\bar{x}_1)^2 + \Sigma(x_2-\bar{x}_2)^2}{n_1+n_2-2}$$

$$= \frac{122.22 + 58.85}{14} = 12.933$$

$$\therefore \; t_{cal} = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{S^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} = \frac{5.55 - 5.14}{\sqrt{12.933\left(\frac{1}{9} + \frac{1}{7}\right)}} = 0.226$$

$t_{cal} = 0.226$

$$\left.\begin{array}{l} t_{tab} \\ \text{at } 14\,df \\ \text{in } 0.05 \end{array}\right\} = 2.145$$

Inf:- Since t cal < t-tab at 14 d.f in 0.05 LOS our Ho will be accepted

So, no significance differences b/w populations Means

F-test:

$$F = \frac{S_1^2}{S_2^2} \sim F(n_1 - 1, n_2 - 1)$$

$(Nr > Dr)$

$$F_{cal} = \frac{15.277}{9.8095} = 1.557$$

$$\left.\begin{array}{l} F_{tab} \\ (9-1, 7-1) \\ \text{at } 5\% \end{array}\right\} = 4.15$$

Inf:- Since Fcal < Ftab at (8, 6) d.f in 0.05 LOS, Ho will be accepted.

Test the equality of Means and equality of Variance following results at 5% LOS:

	Size	Sum of the observation	Sum of Squares of observation from Mean
S_1	9	450	1785
S_2	11	510	2145

Sol:-

$H_0: \mu_1 = \mu_2 : \sigma_1^2 = \sigma_2^2$

$H_1: \mu_1 \neq \mu_2 : \sigma_1^2 \neq \sigma_2^2$

df: $n_1 + n_2 - 2 = 9 + 11 - 2$
$= 18$

LOS: 5% LOS

Test statistics:

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{s^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} \overset{df}{\sim} t_{n_1+n_2-2}$$

data:

$$\begin{cases} n_1 = 9 & \sum x_{1i} = 450 & \sum (x_{1i} - \bar{x}_1)^2 = 1785 \\ n_2 = 11 & \sum x_{2j} = 510 & \sum (x_{2j} - \bar{x}_2)^2 = 2145 \end{cases}$$

$$\bar{x}_1 = \frac{\sum x_{1i}}{n_1} = \frac{450}{9} = 50$$

$$\bar{x}_2 = \frac{\sum x_{2j}}{n_2} = \frac{510}{11} = 46.36$$

$$s_1^2 = \frac{\sum (x_{1i} - \bar{x}_1)^2}{n_1 - 1} = \frac{1785}{8} = 223.125$$

$$s_2^2 = \frac{\sum (x_{2j} - \bar{x}_2)^2}{n_2 - 1} = \frac{2145}{10} = 214.5$$

$$S^2 = \frac{\sum (x_{1i} - \bar{x}_1)^2 + \sum (x_{2j} - \bar{x}_2)^2}{n_1 + n_2 - 2}$$

$$= \frac{1785 + 2145}{18}$$

$$= 218.333$$

$$t_{cal} = \frac{\overline{x_1} - \overline{x_2}}{\sqrt{s^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} = \frac{50 - 46.36}{\sqrt{218.33\left(\frac{1}{9} + \frac{1}{11}\right)}} = 0.548$$

$$\left.\begin{array}{c} t_{tab} \\ \text{at } 18 \, df \\ 0.05 \end{array}\right\} = 2.101$$

Inf:- Since $t_{cal} < t_{tab}$ at 18 d.f 0.05 LOS, Ho will be accepted.

F-test:-

$$F = \frac{s_1^2}{s_2^2} \sim F(n_1 - 1, n_2 - 1)$$

$Nr > Dr$

$$F_{cal} = \frac{s_1^2}{s_2^2} = \frac{223.125}{214.5} = 1.0402$$

$$\left.\begin{array}{c} F_{tab}(9-1, 11-1) \\ \text{at } 0.05 \end{array}\right\} = 3.07$$

Inf:- Since $F_{cal} < F_{tab}$ at (8, 10) df in 0.05 LOS Ho will be accepted.

Chisquare Test (χ^2 - test) (Goodness of fit)

$H_0: P_1 = P_2 = P_3 = \cdots = P_n$

$H_1: P_1 \neq P_2 \neq P_3 \neq \cdots \neq P_n$

d.f:- $n-1$

$(\underbrace{r-1}_{\text{no of rows}} \times \underbrace{c-1}_{\text{no of columns}})$

LoS:- at different levels

Test statistics:-

$$\boxed{\chi^2 = \sum_{i=1}^{n} \frac{(O_i - E_i)^2}{E_i} \sim_{df} \chi^2_{(n-1)}}$$

O - observed freq's (given data)

E: Expected. freq (estimated freq)

χ^2 – test table:

O_i	E_i	$(O_i - E_i)^2$	$\frac{(O_i - E_i)^2}{E_i}$
			$\sum \frac{(O_i - E_i)^2}{E_i} =$

Total of the last column itself calculated value

We should learn how find expected

Inference:- Since $\chi^2_{cal} < \chi^2_{tab}$ at different d.f in different levels our H_0 will be accepted otherwise rejected.

def → Chisquare:-

Sum of the squares of standard normal r.v's is known as Chisquare r. variate.

Methods:-

1) χ^2 – test for independent attributes

General Method

fitting of the disn

(2) χ^2 – test for r x c tables

They have different procedures for computing expected values.

Method - I :-

χ^2 - test for independent of attributes:

$$\left.\begin{array}{l}\text{Test}\\ \text{statistics}\end{array}\right\} = \sum \frac{(O_i - E_i)^2}{E_i} \overset{df}{\sim} \chi^2_{n-1}$$

Here E_i = Expected value $= \frac{\sum O_i}{n} = \overline{O}$

i.e "Every observed value has same expected value"

- x -

A die rolled, the following table shows different face value and freq, Test whether die is unbaised or not at 5% LOS

Sv
$I T_3$
11/5/21
2.20 to 4.30

face value	1	2	3	4	5	6
freq	110	115	140	165	95	125

Sol:—

$\Sigma f_i = 750$

$\bar{x} = \frac{\Sigma f_i}{6} = 125$

$\bar{O} = 125$

Expected freq $(e_i) = 125$

Chisquare table

O	E	$(O-E)^2$	$\frac{(O-E)^2}{E_i}$
110	125	225	1.8
115	125	100	0.8
140	125	225	1.8
165	125	1600	12.8
95	125	900	7.2
125	125	0	0
			24.4

$$\chi^2_{cal} = 24.4$$

$$\left.\begin{matrix}\chi^2_{tab\ 5\,d.f} \\ 0.05\ LOS\end{matrix}\right\} = 11.070$$

Inference:- Since $\chi^2_{cal} > \chi^2_{tab}$ at 5 d.f in 0.05 LOS

our Ho will be rejceted.

i.e Given die is not unbaised

—x—

The air craft accidents during the week days are given below. Test whether accidents are uniform or not at 5% LOS ?

days x	M	T	W	Th	F	Sat	Sun
f	3	5	2	1	4	4	5

Sol:-

$$\Sigma f_i = \Sigma o_i = 24$$

$$\bar{o} = \frac{24}{7} = 3.42$$

O	E	$(O-E)^2$	$\frac{(O-E)^2}{E_i}$
3	3.42	0.1764	0.0515
5	3.42	2.4964	0.7299
2	3.42	2.016	0.5894
1	3.42	5.8564	1.7123
4	3.42	0.1764	0.0515
4	3.42	0.1764	0.0515
5	3.42	2.4964	0.7299

$\chi^2_{cal} = 3.916$

$$\left.\begin{array}{l}\chi^2_{tab} \\ 6\ d.f \\ 0.05\end{array}\right\} = 12.592$$

Inference:- Since $\chi^2_{cal} < \chi^2_{tab}$ at 6 d.f in 0.05 our Ho will accepted. $\therefore$, aircraft accidents are uniform.

Method: 2 Test the Chisquare Goodness of fit by fitting data with Poisson Disn:

$$\left[P(x) = \frac{e^{-\lambda}\lambda^x}{x!} \quad \lambda > 0 \right]$$

x	0	1	2	3	4	5	6
f	1150	900	749	550	410	295	78

$N = 4132$

Sol:- Table - I (Mean Table)

x	f	fx

$\Sigma fx = 7631$

$\Sigma f = 4132$

$$\bar{x} = \frac{\Sigma fx}{\Sigma f} = \frac{7631}{4132}$$

$= 1.846$

$\bar{x} = \lambda = 1.846$

Table - 2:- Fitting of the poisson dis^n

$[\lambda = 1.846]$

[To find expected values]

x	f	$P(x) = \dfrac{e^{-\lambda} \cdot \lambda^x}{x!}$	$E(x) = N \cdot P(x) = 4132\, P(x)$
0	1150	$P(0) = \dfrac{e^{-1.846} (1.846)^0}{0!} = 0.1578$	$E(0) = 4132 \times 0.1578 = 652.02$
1	900	$P(1) = \dfrac{P(0) \cdot \lambda}{1} = 0.2914$	$E(1) = 1204.06$
2	749	$P(2) = \dfrac{P(1) \cdot \lambda}{2} = 0.2689$	$E(2) = 1111.09$
3	550	$P(3) = \dfrac{P(2) \cdot \lambda}{3} = 0.1655$	$E(3) = 683.846$
4	410	$P(4) = \dfrac{P(3) \cdot \lambda}{4} = 0.0763$	$E(4) = 314.03$
5	295	$P(5) = \dfrac{P(4) \cdot \lambda}{5} = 0.0281$	$E(5) = 116.109$
6	78	$P(6) = \dfrac{P(5) \cdot \lambda}{6} = 0.0086$	$E(6) = 35.5352$

Table 3:- Chisquare Goodness of fit

O	E	$(O-E)^2$	$\frac{(O-E)^2}{E_i}$
1150	652.02	247984.08	380.332
900	1204.06	92452.4836	76.783
749	1111.09		118.0004
550	683.846		26.197
410	314.03		29.329
295	116.109		275.620
78	33.5352		58.956

$$\sum \frac{(O-E)^2}{E_i} = 965.2174$$

$\therefore \chi^2_{cal} = 965.2174$

$$\left.\begin{matrix}\chi^2_{tab} \\ 6\ df \\ 0.05\end{matrix}\right\} = 12.592$$

<u>Inference:</u>– Since $\chi^2_{cal} > \chi^2_{tab}$ at 6 d.f in 0.05 LOS our Ho will be rejected.

–x–

Fit the poisson dis^n and test the Chisquare Goodness of fit:

x	0	1	2	3	4	5	6
f	850	710	640	520	445	310	98

$N = 3573$

Sol:- Mean Table

x	f	fx

$\Sigma fx = 8318$

$\Sigma f = 3573$

$$\bar{x} = \frac{\Sigma fx}{N} = \frac{8318}{3573}$$

$$= 2.380$$

$$\bar{x} = \lambda = 2.380$$

Table (2):- fitting of the poisson Dis^n

x	f	$P(x) = \dfrac{e^{-\lambda} \cdot \lambda^x}{x!}$ $\quad \lambda = 2.380$	$E(x) = N \cdot P(x)$ $= 3573 P(x)$
0		$P(0) = \dfrac{e^{-2.380} (2.380)^0}{0!} = 0.0925$	$E(0) = 333.683$
1		$P(1) = \dfrac{P(0) \cdot \lambda}{1} = 0.2206$	$E(1) = 786.595$
2		$P(2) = \dfrac{P(1) \cdot \lambda}{2} = 0.2619$	$E(2) = 936.049$
3		$P(3) = \dfrac{P(2) \cdot \lambda}{3} = 0.20784$	$E(3) = 742.576$
4		$P(4) = \dfrac{P(3) \cdot \lambda}{4} = 0.1236$	$E(4) = 441.556$
5		$P(5) = \dfrac{P(4) \cdot \lambda}{5} = 0.0588$	$E(5) = 210.212$
6		$P(6) = \dfrac{P(5) \cdot \lambda}{6} = 0.0233$	$E(6) = 83.336$

Table -3:- Chisquare Table

O	E	$(O-E)^2$	$\frac{(O-E)^2}{E_i}$	
850	333.683		798.911	
710	786.595		7.458	
640	936.049		93.632	
520	742.376		66.611	
445	441.556		0.0268	
310	210.212		0.0564	
98	83.336		2.580	

$$\sum \frac{(O_i - E_i)^2}{E_i} = 969.275$$

$\therefore \chi^2_{cal} = 969.275$

$\left.\begin{matrix}\chi^2_{tab} \\ 6\,df\ 0.05\end{matrix}\right\} = 12.592$

Inf :- Since $\chi^2_{cal} > \chi^2_{tab}$ at 6 d.f in 0.05 LOS our Ho will be rejcuted.

Method - II :- χ^2 method (r x c tables)

Test statistics:-

$$\chi^2 = \sum \frac{(O_i - E_i)^2}{E_i} \sim \chi^2_{\underbrace{(r-1)}_{\text{no of rows}} \times \underbrace{(c-1)}_{\text{no of columns}}}$$

2x2 Contingency table:-

	S-II: α	β	Row Total
S-I: A	a	b	a+b
B	c	d	c+d
Column Total	a+c	b+d	$\boxed{a+b+c+d = N}$

Grand Total =

= Sum Row Totals =

= Sum Column Total

How find the expected freq (r x c method)

$$E(a) = \frac{R.T \times C.T}{G.T} = \frac{(a+b) \times (a+c)}{a+b+c+d} = \frac{(a+b) \times (a+c)}{N}$$

$$E(b) = \frac{R.T \times C.T}{G.T} = \frac{(a+b) \times (b+d)}{N}$$

$$E(c) = \frac{R.T \times C.T}{G.T} = \frac{(c+d) \times (a+c)}{N}$$

$$E(d) = \frac{R.T \times C.T}{G.T} = \frac{(c+d) \times (b+d)}{N}$$

Test the chi-square goodness of fit at 5% LOS for the following cells :

	α	β
I	5	15
II	10	9
III	4	6

Sol:-

	α	β	RT
I	5 (7.7551)	15 (12.244)	20
II	10 (7.36)	9 (11.6326)	19
III	4 (3.8775)	6 (6.122)	10
	19	30	49 = N

Let us find expected freq's

$$E(5) = \frac{20 \times 19}{49} = 7.7551$$

$$E(15) = \frac{20 \times 30}{49} = 12.244$$

$$E(10) = \frac{19 \times 19}{49} = 7.36$$

$$E(9) = \frac{19 \times 30}{49} = 11.6326$$

$$E(4) = \frac{10 \times 19}{49} = 3.8775$$

$$E(6) = \frac{10 \times 30}{49} = 6.122$$

Chi Square Table:

O	E	$(O-E)^2$	$\frac{(O-E)^2}{E_i}$
5	7.9551		0.9787
15	12.244		0.6191
10	7.36		0.9465
9	11.6326		0.595
4	3.8775		0.0038
6	6.122		0.0024
			3.1455

$\therefore \chi^2_{cal} = 3.1445$

χ^2_{tab}

$\left.\begin{array}{l}(3-1)\times(2-1) = 2\,df \\ \text{at } 0.05\,LOS\end{array}\right\} = 5.991$

<u>Infer:</u> Since $\chi^2_{cal} < \chi^2_{tab}$ at 2 d.f in 0.05 LOS our Ho will be accepted.

test the chi square Goodness of fit for the following data at 5% LOS ?

	α	β	γ
A	18	21	19
B	15	17	10
C	11	8	10

SW
IT3
12/5/21
1.20 to 3.30PM

Sol:-

	α	β	γ	R.T
A	18 (19.78)	21 (20.68)	19 (17.53)	58
B	15 (14.32)	17 (14.97)	10 (12.69)	42
C	11 (9.89)	8 (10.34)	10 (8.76)	29
C.T	44	46	39	129 = N

$$E(18) = \frac{58 \times 44}{129} = 19.78$$

$$E(21) = \frac{58 \times 46}{129} = 20.68$$

$$E(19) = \frac{58 \times 39}{129} = 17.53$$

$$E(15) = \frac{42 \times 44}{129} = 14.32$$

$$E(17) = \frac{42 \times 46}{129} = 14.97$$

$$E(10) = \frac{42 \times 39}{129} = 12.69$$

$$E(11) = \frac{29 \times 44}{129} = 9.89$$

$$E(8) = \frac{29 \times 46}{129} = 10.34$$

$$E(10) = \frac{29 \times 39}{129} = 8.76$$

Chisquare table:-

O_i	E_i	$(O-E)^2$	$\frac{(O-E)^2}{E}$
18	19.78		0.1610
21	20.68		0.0049
19	17.53		0.1232
15	14.32		0.0322
17	14.97		0.2752
10	12.69		0.5702
11	9.89		0.1245
8	10.34		0.5295
10	8.76		0.1755

$$\chi^2_{cal} = \Sigma \frac{(O_i - E_i)^2}{E_i} = 1.9962$$

$$\left.\begin{array}{l}\chi^2_{tab} \\ (3-1)\times(3-1)=4 \\ \text{at } 0.05\end{array}\right\} = 9.488$$

Inference:- Since $\chi^2_{cal} < \chi^2_{tab}$ at 4 d.f in 0.05 LOS

our N.H will be accepted.

(i.e) there is no significance differences with in the cells.

Key points:-

1. Test statistics for single mean (σ known) $\Rightarrow$ $Z = \dfrac{\bar{x} - \mu}{\sigma/\sqrt{n}} \sim N(0,1)$

2. Test statistics for difference of two mean (σ known):

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

3. Test statistics for single sample proportion

$$Z = \frac{p - P}{\sqrt{PQ/n}} \sim N(0,1)$$

4. Test statistics difference b/w two sample proportions

$$Z = \frac{p_1 - p_2}{\sqrt{PQ\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} \sim N(0,1)$$

$p_1 = \frac{x_1}{n_1}$. $p_2 = \frac{x_2}{n_2}$

$P = \frac{x_1 + x_2}{n_1 + n_2}$ or $\frac{n_1 p_1 + n_2 p_2}{n_1 + n_2}$

$(Q = 1 - P)$

5. Test statistics for single sample mean (σ unknown)

$$t = \frac{\bar{x} - \mu}{s/\sqrt{n}} \overset{df}{\sim} t_{n-1} \qquad S = \sqrt{\frac{\Sigma(x_i - \bar{x})^2}{n-1}}$$

6. Test statistics for difference b/w two sample mean (σ unknown)

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{s^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} \overset{df}{\sim} t_{n_1+n_2-2} \qquad S^2 = \frac{\Sigma(x_{1i} - \bar{x}_1)^2 + \Sigma(x_{2j} - \bar{x}_2)^2}{n_1 + n_2 - 2}$$

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

7. F-test :-

$$F = \frac{s_1^2}{s_2^2} \sim F(\underset{\nu_1}{n_1-1}, \underset{\nu_2}{n_2-1})$$

$(N_r > D_r)$

$$= \frac{s_2^2}{s_1^2} \sim F(\underset{\nu_1}{n_2-1}, \underset{\nu_2}{n_1-1})$$

8. Chisqaure Test (Goodness fit)

(a) $$\chi^2 = \Sigma \frac{(O-E)^2}{E_i} \sim \chi^2_{n-1}$$ [Independent attributes]

(b) $$\chi^2 = \Sigma \frac{(O-E)^2}{E} \sim \chi^2_{(r-1 \times c-1)}$$ [$r \times c$]