Student's t-test for difference of means:

To test the significant difference between the sample means $\bar{x}$ and $\bar{y}$ of two independent samples of sizes n_1 and n_2 with the same variance, we use statistic

$$t = \frac{\bar{x}-\bar{y}}{\sqrt{s^2\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}}$$

where $\bar{x} = \frac{1}{n_1}\Sigma x_i$, $\bar{y} = \frac{1}{n}\Sigma y_i$ & $s^2 = \frac{1}{n_1+n_2-2}\left[\Sigma(x_i-\bar{x})^2 + \Sigma(y_i-\bar{y})^2\right]$

or $s^2 = \frac{1}{n_1+n_2-2}\left[(n_1-1)s_1^2 + (n_2-1)s_2^2\right]$

→ Samples of two types of electric light bulbs were tested for length of life and following data were obtained:

	Type I	Type II
Sample number	$n_1=8$	$n_2=7$
Sample mean,	$\bar{x}=1234$ hrs	$\bar{y}=1036$ hrs
Sample S.D	$S_1=36$ hrs	$S_2=40$ hrs

Is the difference in the means sufficient to warrant that type I is superior to type II regarding length of life.

soln:

Null hypothesis: H_0: The two types I and II of electric bulbs are identical.

Alternative hypothesis H_1: $\mu_1 > \mu_2$

Since the two sample means $\bar{x}$ and $\bar{y}$ are given and also sample standard deviations s_1 and s_2 are given, we use the statistic

$$t = \frac{\bar{x}-\bar{y}}{s\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}}$$

where $s^2 = \frac{1}{n_1+n_2-2}\left[(n_1-1)s_1^2 + (n_2-1)s_2^2\right]$

$= \frac{1}{8+7-2}\left[7(36)^2 + 6(40)^2\right] = 1436.31$

$\therefore t = \frac{1234-1036}{\sqrt{1436.31\left(\frac{1}{8}+\frac{1}{7}\right)}} = 10.09$

Degrees of freedom (d.f) $= n_1+n_2-2 = 8+7-2 = 13$

Tabulated value of t for 13 d.f at 5% level is 1.77

Since calculated t > tabulated t

we reject null hypothesis H_0.

i.e Two types I and II of electric bulbs are not identical.

→ The means of two random samples of sizes 9 and 7 are 196.42 and 198.82 respectively. The sum of the squares of the deviations from the mean are 26.94 and 18.73 respectively. Can the sample be considered to have been drawn from the same normal population.

Sol: Given $n_1 = 9$, $n_2 = 7$, $\bar{x} = 196.42$, $\bar{y} = 198.82$

night math.

$\Sigma(x_i-\bar{x})^2 = 26.94$ $\quad \Sigma(y_i-\bar{y})^2 = 18.73$

$\therefore S^2 = \frac{\Sigma(x_i-\bar{x})^2 + \Sigma(y_i-\bar{y})^2}{n_1+n_2-2} = \frac{26.94+18.73}{9+7-2} = 3.26$

$\therefore S = 1.81$

1, Null hypothesis H_0: The two samples are drawn from the same population i.e $\mu_1 = \mu_2$

2, Alternative hypothesis H_1: $\mu_1 \neq \mu_2$

3, Level of significance $\alpha = 0.05$

9, Test statistic is $t = \dfrac{\bar{x}-\bar{y}}{S\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}} = \dfrac{196.42-198.82}{1.81\sqrt{\frac{1}{9}+\frac{1}{7}}}$

$= \dfrac{-2.4}{0.912} = -2.63$

calculated value of $|t| = 2.63$

Tabulated value of t for $9+7-2=14$ d.f at 5% level of significance is 2.15. since calculated vale of t > tabulated vale of t, we reject the null hypothesis H_0.

i.e the two samples are not drawn from the same population.

→ ~~Below are given the gain in weights (in lbs) of pigs fed on~~

→ Two horses A and B were tested according to the time (in seconds) to run a particular track with the following results.

Horse A	28	30	32	33	33	29	34
Horse B	29	30	30	24	27	29	

Test whether the two horses have the same running capacity?

Sol: Given $n_1 = 7$, $n_2 = 6$

we first compute the sample means and S.d's

$\bar{x}$ = mean of first sample

$= \frac{1}{7}(28+30+32+33+33+29+34) = \frac{219}{7} = 31.286$

$\bar{y}$ = mean of second sample $= \frac{1}{6}(29+30+30+24+27+29)$

$= \frac{169}{6} = 28.16$

x	$x-\bar{x}$	$(x-\bar{x})^2$	y	$y-\bar{y}$	$(y-\bar{y})^2$
28	-3.286	10.8	29	0.84	0.7056
30	-1.286	1.6538	30	1.84	3.3856
32	0.714	0.51	30	1.84	3.3856
33	1.714	2.94	24	-4.16	17.3056
33	1.714	2.94	27	-1.16	1.3456
29	-2.286	5.226	29	0.84	0.7056
34	2.714	7.366			
219		31.4358	169		26.8336

Now $S^2 = \frac{1}{n_1+n_2-2}\left[\Sigma(x_i-\bar{x})^2 + \Sigma(y_i-\bar{y})^2\right]$

$= \frac{1}{11}\left[31.4358 + 26.8336\right]$

$= \frac{1}{11}[58.2694] = 5.23$

$S = \sqrt{5.23} = 2.3$

1, Null hypothesis $H_0: \mu_1 = \mu_2$

2, Alternative hypothesis: $H_1: \mu_1 \neq \mu_2$

3, Level of significance $\alpha = 0.05$

4, the test statistic is $t = \frac{\bar{x}-\bar{y}}{S\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}} = \frac{31.286 - 28.16}{2.3\sqrt{\frac{1}{7}+\frac{1}{6}}}$

$= 2.443$

Tabulated value of t for $7+6-2 = 11$ d.f at 5% level of significance is 2.2.

Since cal value of t > Table t, we reject the null hypothesis H_0 i.e both horses A and B do not have the same running capacity.

→ To test the claim that the resistance of electric wire can be reduced by atleast 0.05 ohm by alloying, 25 values obtained for each alloyed wire and standard wire produced the following results:

	Mean	standard deviation
Alloyed wire	0.083 ohm	0.003 ohm
Standard wire	0.136 ohm	0.002 ohm

Test at 5% level whether or not the claim is substantiated.

Sol: Null hypothesis H_0: $\mu_1 - \mu_2 \geq 0.05$ (i.e the claim is substantiated)

Alternative hypothesis H_1: $\mu_1 - \mu_2 < 0.05$ (Left tailed test)

Test statistic: Under H_0, the test statistic is

$$t = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{s^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

where $s^2 = \frac{n_1 s_1^2 + n_2 s_2^2}{n_1 + n_2 - 2} = \frac{25 \times (0.003)^2 + 25 \times (0.002)^2}{25 + 25 - 2}$

$$= \frac{0.000225 + 0.0001}{48} = 0.0000067$$

$$\therefore t = \frac{(0.083 - 0.136) - 0.05}{\sqrt{0.0000067\left(\frac{1}{25} + \frac{1}{25}\right)}} = \frac{-0.103}{0.00071} = -145.07$$

The tabulated value of t for 48 d.o.f at 5% level of significance for left tailed test is -1.645.

Conclusion: Since calculated value of t is much less than tabulated value of t, it falls in the rejection region. We, therefore, reject the null hypothesis and conclude that the claim is not substantiated.

Paired Sample t-test for Difference of means:

Let us now consider the case when

i) the sample sizes are equal i.e $n_1 = n_2 = n$ (say)

ii) the two samples are not independent but the sample observations are paired together i.e the pair of observations (x_i, y_i), $i = 1, 2 - - - n$ corresponds to the same (i^{th}) sample unit.

The problem is to test if the sample means differ significantly or not.

For example, suppose we want to test the efficacy of a particular drug, say, for inducing sleep. Let x_i and y_i $(1, 2 - - n)$ be the readings, in hours of sleep, on the i^{th} individual, before ~~the drug~~ and after the drug is given respectively. Here instead of applying the difference of the means ~~test~~ test, we apply the paired t-test given below.

Here we consider the increments $d_i = x_i - y_i$ $(i = 1, 2 - - n)$

Under the null hypothesis, H_0 that increments are due to fluctuations of sampling i.e the drug is not responsible for these increments, the statistic $t = \frac{\bar{d}}{S/\sqrt{n}}$

where $\bar{d} = \frac{1}{n}\sum_{i=1}^{n} d_i$ & $S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(d_i - \bar{d})^2$

follows student's t-distribution with $(n-1)$ d.o.f.

→ Ten workers were given a training programme with a view to study their assembly time for a certain mechanism. The results of the time and motion studies before and after the training programme are given below.

workers	1	2	3	4	5	6	7	8	9	10
X_1	15	18	20	17	16	14	21	19	13	22
Y_1	14	16	21	10	15	18	17	16	14	20

X_1 = Time taken for assembling before training

Y_1 = Time taken for assembling after training

Test whether there is significant difference in assembly times before and after training.

Sol: Let μ be the mean of population of differences.

1. Null hypothesis $H_0: \mu = 0$ i.e the training is not useful

2. Alternative hypothesis $H_1: \mu > 0$ i.e training is useful in assembly time.

3. Level of significance $\alpha = 0.05$

4. computation: Difference d_i's (before and after training) are

$1, 2, -1, 7, 1, -4, 4, 3, -1, 2$

$\bar{d}$ = mean of differences of sample data $= \frac{14}{10} = 1.4$

$$s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(d_i - \bar{d})^2$$

$$= \frac{1}{9}\Big[(1-1.4)^2 + (2-1.4)^2 + (-1-1.4)^2 + (7-1.4)^2 + (1-1.4)^2 + (-4-1.4)^2 + (4-1.4)^2 + (3-1.4)^2 + (-1-1.4)^2 + (2-1.4)^2\Big]$$

$$= \frac{1}{9}\Big[0.16 + 0.36 + 5.76 + 31.36 + 0.16 + 29.16 + 6.76 + 2.56 + 5.76 + 0.36\Big]$$

$$= \frac{82.4}{9} = 9.1555 \qquad \therefore s = 3.026$$

5, The test statistic is $t = \frac{\bar{d}-\mu}{S/\sqrt{n}} = \frac{1.4-0}{3.026/\sqrt{10}} = \frac{(1.4)(3.1623)}{3.026} = 1.46$

$\therefore$ calculated $|t| = 1.46$

Tabulated $t_{0.05}$ with $10-1=9$ degrees of freedom is 1.833

Since calculated $t < t_{0.05}$, we accept the null hypothesis

i.e training is not useful.

Thus there is no ~~de~~ significante difference

in assembly times before and after training.

→ Scores obtained in a shooting competetion by 10 soldiers before and after intensive training are given below.

Before	67	24	57	55	63	54	56	68	33	43
After	70	38	58	58	56	67	68	75	42	38

Test whether the intensive training is useful at 0.05 level of significance.

Sol: Let μ be the mean of population of differences.

1, Null hypothesis H_0: $\mu=0$ i.e not useful.

2, Alternative hypothesis H_1: $\mu>0$ intensive training is useful.

3, Level of significance $\alpha=0.05$

4, computation:

Differences d_i's (before and after training) are

$-3, -14, -1, -3, 7, -13, -12, -7, -9, 5$

$\bar{d} = \frac{1}{10}\sum_{i=1}^{10} d_i = \frac{-50}{10} = -5$

$S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(d_i-\bar{d})^2 = \frac{1}{9}\sum_{i=1}^{10}(d_i-\bar{d})^2$

$= \frac{1}{9}\left[(2)^2+(-9)^2+(4)^2+(2)^2+(12)^2+(-8)^2+(-7)^2+(-2)^2+(-4)^2+(10)^2\right]$

$= \frac{1}{9}\left[4+81+16+4+144+64+49+4+16+100\right]$

$= \frac{462}{9} = 51.5555$

$\therefore s = 7.22$

$\therefore$ The test statistic is $t = \frac{\bar{d}-\mu}{s/\sqrt{n}} = \frac{-5-0}{7.22/\sqrt{10}} = -2.16$

$\therefore |t| = 2.16$

Tabulated $t_{0.05}$ with $10-1=9$ degrees of freedom is 1.83.

Since calculated $t >$ tabulated t, we reject the null hypothesis

i.e the intensive training is useful.

→ The Blood pressure of 5 women before and after intake of a certain drug are given below:

Before	110	120	125	132	125
After	120	118	125	136	121

Test whether there is significant change in Blood pressure at 1% level of significance.

Sol: Let μ be the mean of population of differences

1, Null hypothesis $H_0: \mu = 0$ i.e no change in B.P

2, Alternative hypothesis $H_1: \mu \neq 0$ i.e change in B.P

3, Level of significance $\alpha = 0.01$

4, Computation: Difference d_i's (before and after drug) are

$-10, 2, 0, -4, 4$

$\bar{d} = \frac{-10+2+0-4+4}{5} = \frac{-8}{5} = -1.6$

$s^2 = \frac{1}{4}\sum_{i=1}^{5}(d_i - \bar{d})^2 = \frac{1}{4}\left[(-10+1.6)^2 + (2+1.6)^2 + (0+1.6)^2 + (-4+1.6)^2 + (4+1.6)^2\right]$

$= \frac{1}{4}\left[70.56 + 12.96 + 2.56 + 5.76 + 31.36\right] = \frac{123.20}{4} = 30.8$

$s = \sqrt{10.8} = 5.55$

So the test statistic is $t = \dfrac{\bar{d} - \mu}{s/\sqrt{n}}$

$$t = \frac{-1.6}{5.55/\sqrt{5}} = -0.645$$

$\therefore$ calculated $|t| = 0.645$

Tabulated $t_{0.01}$ with $5-1=4$ degrees of freedom is 3.747.

Since calculated $t < t_{0.01}$, we accept the null hypothesis

i.e there is no significant change in Blood pressure.

→ In a certain experiment to compare two types of animal foods A and B the following results of increase in weights were observed in animals:

	Animal number	1	2	3	4	5	6	7	8	Total
Increase weight in lb	Food A	49	53	51	52	47	50	52	53	407
	Food B	52	55	52	53	50	54	54	53	423

i) Assuming that the two samples of animals are independent, can we conclude that food B is better than food A?

ii) Also examine the case when the same set of eight animals were used in both the foods.

Sol: Null hypothesis H_0: If the increase in weights due to foods A and B are denoted by X and Y respectively, then $H_0 = \mu_X = \mu_Y$ i.e there is no significant difference in increase in weights due to diets A and B.

Alternative hypothesis, $H_1: \mu_X < \mu_Y$

i) If the two samples of animals be assumed to be independent, then we will apply t-test for difference of means to test H_0

Test statistic under $H_0: \mu_x = \mu_y$, the test criterion is

$$t = \frac{\bar{x} - \bar{y}}{\sqrt{s^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

Food A			Food B		
x_i	$x_i - \bar{x}$	$(x_i - \bar{x})^2$	y	$y_i - \bar{y}$	$(y_i - \bar{y})^2$
49	-1	1	52	0	0
53	3	9	55	3	9
51	1	1	52	0	0
52	2	4	53	1	1
47	-3	9	50	-2	4
50	0	0	54	2	4
52	2	4	54	2	4
53	3	9	53	1	1
	7	37		7	23

$$s^2 = \frac{1}{n_1 + n_2 - 2}\left[\Sigma(x_i - \bar{x})^2 + \Sigma(y_i - \bar{y})^2\right]$$

$$= \frac{1}{14}\left[30.875 + 16.875\right] = 3.41$$

$$t = \frac{\bar{x} - \bar{y}}{\sqrt{s^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} = \frac{50.875 - 52.875}{\sqrt{3.41\left(\frac{1}{8} + \frac{1}{8}\right)}} = -2.17$$

Tabulated $t_{0.05}$ for $(8+8-2) = 14$ d.o.f for one tail test is 1.76.

Conclusion: The critical region for the left tail test is $t < -1.76$. Since calculated t is less than -1.76, Ho is rejected at 5% level of significance. Here we conclude that the foods A and B differ significantly as regards their effect on increase in weight.

ii) If the same set of animals is used in both the cases, then the readings X and Y are not independent but they are paired together and we apply the paired t-test for testing H_0.

under $H_0: \mu_x = \mu_y$

the statistic is $t = \frac{\bar{d}}{s/\sqrt{n}}$.

x	49	53	51	52	47	50	52	53	
y	52	55	52	53	50	54	54	53	
$d = x - y$	-3	-2	-1	-1	-3	-4	-2	0	-16
d^2	9	4	1	1	9	16	4	0	44

$\bar{d} = \frac{\Sigma d}{n} = \frac{-16}{8} = -2$

$s^2 = \frac{1}{n-1}\left[\Sigma d^2 - \frac{(\Sigma d)^2}{n}\right]$

$= \frac{1}{7}\left[44 - \frac{256}{8}\right] = 1.714$

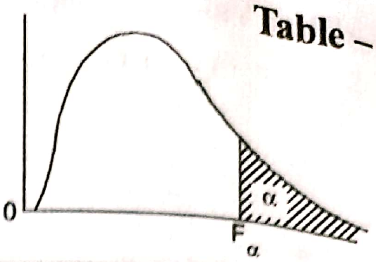
$|t| = \frac{|\bar{d}|}{\sqrt{s^2/n}} = \frac{2}{\sqrt{1.7143/8}} = \frac{2}{0.4629} = 4.32$

Tabulated $t_{0.05}$ for $(8-1) = 7$ d.o.f for one tail test is 1.90

Conclusion: Here also the observed value of t is significant at 5% level of significance and we conclude that food B is superior to food A.

Critical Values of the F-Distribution

Table – 5



Values of $F_{0.05}(\nu_1, \nu_2)$

ν_2 \ ν_1	1	2	3	4	5	6	7	8	9
1	161.4	199.5	215.7	224.6	230.2	234.0	236.8	238.9	240.5
2	18.51	19.00	19.16	19.25	19.30	19.33	19.35	19.37	19.38
3	10.13	9.55	9.28	9.12	9.01	8.94	8.89	8.85	8.81
4	7.71	6.94	6.59	6.39	6.26	6.16	6.09	6.04	6.00
5	6.61	5.79	5.41	5.19	5.05	4.95	4.88	4.82	4.77
6	5.99	5.14	4.76	4.53	4.39	4.28	4.21	4.15	4.10
7	5.59	4.74	4.35	4.12	3.97	3.87	3.79	3.73	3.68
8	5.32	4.46	4.07	3.84	3.69	3.58	3.50	3.44	3.39
9	5.12	4.26	3.86	3.63	3.48	3.37	3.29	3.23	3.18
10	4.96	4.10	3.71	3.48	3.33	3.22	3.14	3.07	3.02
11	4.84	3.98	3.59	3.36	3.20	3.09	3.01	2.95	2.90
12	4.75	3.89	3.49	3.26	3.11	3.00	2.91	2.85	2.80
13	4.67	3.81	3.41	3.18	3.03	2.92	2.83	2.77	2.71
14	4.60	3.74	3.34	3.11	2.96	2.85	2.76	2.70	2.65
15	4.54	3.68	3.29	3.06	2.90	2.79	2.71	2.64	2.59
16	4.49	3.63	3.24	3.01	2.85	2.74	2.66	2.59	2.54
17	4.45	3.59	3.20	2.96	2.81	2.70	2.61	2.55	2.49
18	4.41	3.55	3.16	2.93	2.77	2.66	2.58	2.51	2.46
19	4.38	3.52	3.13	2.90	2.74	2.63	2.54	2.48	2.42
20	4.35	3.49	3.10	2.87	2.71	2.60	2.51	2.45	2.39
21	4.32	3.47	3.07	2.84	2.68	2.57	2.49	2.42	2.37
22	4.30	3.44	3.05	2.82	2.66	2.55	2.46	2.40	2.34
23	4.28	3.42	3.03	2.80	2.64	2.53	2.44	2.37	2.32
24	4.26	3.40	3.01	2.78	2.62	2.51	2.42	2.36	2.30
25	4.24	3.39	2.99	2.76	2.60	2.49	2.40	2.34	2.28
26	4.23	3.37	2.98	2.74	2.59	2.47	2.39	2.32	2.27
27	4.21	3.35	2.96	2.73	2.57	2.46	2.37	2.31	2.25
28	4.20	3.34	2.95	2.71	2.56	2.45	2.36	2.29	2.24
29	4.18	3.33	2.93	2.70	2.55	2.43	2.35	2.28	2.22
30	4.17	3.32	2.92	2.69	2.53	2.42	2.33	2.27	2.21
40	4.08	3.23	2.84	2.61	2.45	2.34	2.25	2.18	2.12
60	4.00	3.15	2.76	2.53	2.37	2.25	2.17	2.10	2.04
120	3.92	3.07	2.68	2.45	2.29	2.17	2.09	2.02	1.96
∞	3.84	3.00	2.60	2.37	2.21	2.10	2.01	1.94	1.88

(Continued) Critical Values of the F-Distribution

Values of $F_{0.05}(\nu_1, \nu_2)$

ν_2 \ ν_1	10	12	15	20	24	30	40	60	120	∞
1	241.9	243.9	245.9	248.0	249.1	250.1	251.1	252.2	253.3	254.3
2	19.40	19.41	19.43	19.45	19.45	19.46	19.47	19.48	19.49	19.50
3	8.79	8.74	8.70	8.66	8.64	8.62	8.59	8.57	8.55	8.52
4	5.96	5.91	5.86	5.80	5.77	5.75	5.72	5.69	5.66	5.62
5	4.74	4.68	4.62	4.56	4.53	4.50	4.46	4.43	4.40	4.36
6	4.06	4.00	3.94	3.87	3.84	3.81	3.77	3.74	3.70	3.67
7	3.64	3.57	3.51	3.44	3.41	3.38	3.34	3.30	3.27	3.23
8	3.35	3.28	3.22	3.15	3.12	3.08	3.04	3.01	2.97	2.93
9	3.14	3.07	3.01	2.94	2.90	2.86	2.83	2.79	2.75	2.71
10	2.98	2.91	2.85	2.77	2.74	2.70	2.66	2.62	2.58	2.54
11	2.85	2.79	2.72	2.65	2.61	2.57	2.73	2.49	2.45	2.40
12	2.75	2.69	2.62	2.54	2.51	2.47	2.43	2.38	2.34	2.30
13	2.67	2.60	2.53	2.46	2.42	2.38	2.34	2.30	2.25	2.21
14	2.60	2.53	2.46	2.39	2.35	2.31	2.27	2.22	2.18	2.13
15	2.54	2.48	2.40	2.33	2.29	2.25	2.20	2.16	2.11	2.07
16	2.49	2.42	2.35	2.28	2.24	2.19	2.15	2.11	2.06	2.01
17	2.45	2.38	2.31	2.23	2.19	2.17	2.10	2.06	2.01	1.96
18	2.41	2.34	2.27	2.19	2.15	2.11	2.06	2.02	1.97	1.92
19	2.38	2.31	2.23	2.16	2.11	2.07	2.03	1.98	1.93	1.88
20	2.35	2.28	2.20	2.12	2.08	2.04	1.99	1.95	1.90	1.84
21	2.32	2.25	2.18	2.10	2.05	2.01	1.96	1.92	1.87	1.81
22	2.30	2.23	2.15	2.07	2.03	1.98	1.94	1.89	1.84	1.78
23	2.27	2.20	2.13	2.05	2.01	1.96	1.91	1.86	1.81	1.76
24	2.25	2.18	2.11	2.03	1.98	1.94	1.89	1.94	1.79	1.73
25	2.24	2.16	2.09	2.01	1.96	1.92	1.87	1.82	1.77	1.71
26	2.22	2.15	2.07	1.99	1.95	1.90	1.85	1.80	1.75	1.69
27	2.20	2.13	2.06	1.97	1.93	1.88	1.84	1.79	1.73	1.67
28	2.19	2.12	2.04	1.96	1.91	1.87	1.82	1.77	1.71	1.65
29	2.18	2.10	2.03	1.94	1.90	1.85	1.81	1.75	1.70	1.64
30	2.16	2.09	2.01	1.93	1.89	1.84	1.79	1.75	1.68	1.62
40	2.08	2.00	1.92	1.84	1.79	1.74	1.69	1.64	1.58	1.51
60	1.99	1.92	1.84	1.75	1.70	1.65	1.59	1.53	1.47	1.39
120	1.91	1.83	1.75	1.66	1.61	1.55	1.50	1.43	1.35	1.25
∞	1.83	1.75	1.67	1.57	1.52	1.46	1.39	1.32	1.22	1.00

(Continued) Critical Values of the F-Distribution

Values of $F_{0.01}(\nu_1, \nu_2)$

	ν_1								
ν_2	1	2	3	4	5	6	7	8	9
1	4052	4999.5	5403	5625	5764	5859	5928	5981	6022
2	98.50	99.00	99.17	99.25	99.30	99.33	99.36	99.37	99.39
3	34.12	30.82	29.46	28.71	28.24	27.91	27.67	27.49	27.35
4	21.20	18.00	16.69	15.98	15.52	15.21	14.98	14.80	14.66
5	16.26	13.27	12.06	11.39	10.97	10.67	10.46	10.29	10.16
6	13.75	10.92	9.78	9.15	8.75	8.47	8.26	8.10	7.98
7	12.25	9.55	8.45	7.85	7.46	7.19	6.99	6.84	6.72
8	11.26	8.65	7.59	7.01	6.63	6.37	6.18	6.03	5.91
9	10.56	8.02	6.99	6.42	6.06	5.80	5.61	5.47	5.35
10	10.04	7.56	6.55	5.99	5.64	5.39	5.20	5.06	4.94
11	9.65	7.21	6.22	5.67	5.32	5.07	4.89	4.74	4.63
12	9.33	6.93	5.95	5.41	5.06	4.82	4.64	4.50	4.39
13	9.07	6.70	5.74	5.21	4.86	4.62	4.44	4.30	4.19
14	8.86	6.51	5.56	5.04	4.69	4.46	4.28	4.14	4.03
15	8.68	6.36	5.42	4.89	4.56	4.32	4.14	4.00	3.89
16	8.53	6.23	5.29	4.77	4.44	4.20	4.03	3.89	3.78
17	8.40	6.11	5.18	4.67	4.34	4.10	3.93	3.79	3.68
18	8.29	6.01	5.09	4.58	4.25	4.01	3.84	3.71	3.60
19	8.18	5.93	5.01	4.50	4.17	3.94	3.77	3.63	3.52
20	8.10	5.85	4.94	4.43	4.10	3.87	3.70	3.56	3.46
21	8.02	5.78	4.87	4.37	4.04	3.81	3.64	3.51	3.40
22	7.95	5.72	4.82	4.31	3.99	3.76	3.59	3.45	3.35
23	7.88	5.66	4.76	4.26	3.94	3.71	3.54	3.41	3.30
24	7.82	5.61	4.72	4.22	3.90	3.67	3.50	3.36	3.26
25	7.77	5.57	4.68	4.18	3.85	3.63	3.46	3.32	3.22
26	7.72	5.53	4.64	4.14	3.82	3.59	3.42	3.29	3.18
27	7.68	5.49	4.60	4.11	3.78	3.56	3.39	3.26	3.15
28	7.64	5.45	4.57	4.07	3.75	3.53	3.36	3.23	3.12
29	7.60	5.42	4.54	4.04	3.73	3.50	3.33	3.20	3.09
30	7.56	5.39	4.51	4.02	3.70	3.47	3.30	3.17	3.07
40	7.31	5.18	4.31	3.83	3.51	3.29	3.12	2.99	2.89
60	7.08	4.98	4.13	3.65	3.34	3.12	2.95	2.82	2.72
120	6.85	4.79	3.95	3.48	3.17	2.96	3.79	2.66	2.56
∞	6.63	4.61	3.78	3.32	3.02	2.80	2.64	2.51	2.41

Critical Values of the F-Distribution

Values of $F_{0.01}(\nu_1, \nu_2)$

ν_2 \ ν_1	10	12	15	20	24	30	40	60	120	∞
1	6056	6106	6157	6209	6235	6261	6287	6313	6339	6366
2	99.40	99.42	99.43	99.45	99.46	99.47	99.47	99.48	99.49	99.50
3	27.23	27.05	26.87	26.69	26.60	26.50	26.41	26.32	26.22	26.13
4	14.55	14.37	14.20	14.02	13.93	13.84	13.75	13.65	13.56	13.46
5	10.05	9.89	9.72	9.55	9.47	9.38	9.29	9.20	9.11	9.02
6	7.87	7.72	7.56	7.40	7.31	7.23	7.14	7.06	6.97	6.88
7	6.62	6.47	6.31	6.16	6.07	5.99	5.91	5.82	5.74	5.65
8	5.81	5.67	5.52	5.36	5.28	5.20	5.12	5.03	4.95	4.86
9	5.26	5.11	4.96	4.81	4.73	4.65	4.57	4.48	4.40	4.31
10	4.85	4.71	4.56	4.41	4.33	4.25	4.17	4.08	4.00	3.91
11	4.54	4.40	4.25	4.10	4.02	3.94	3.86	3.78	3.69	3.60
12	4.30	4.16	4.01	3.86	3.78	3.70	3.62	3.54	3.45	3.36
13	4.10	3.96	3.82	3.66	3.59	3.51	3.43	3.34	3.25	3.17
14	3.94	3.80	3.66	3.51	3.43	3.35	3.27	3.18	3.09	3.00
15	3.80	3.67	3.52	3.37	3.29	3.21	3.13	3.05	2.96	2.87
16	3.69	3.55	3.41	3.26	3.18	3.10	3.02	2.93	2.84	2.75
17	3.59	3.46	3.31	3.16	3.08	3.00	2.92	2.83	2.75	2.65
18	3.51	3.37	3.23	3.08	3.00	2.92	2.84	2.75	2.66	2.57
19	3.43	3.30	3.15	3.00	2.92	2.84	2.76	2.67	2.58	2.49
20	3.37	3.23	3.09	2.94	2.86	2.78	2.69	2.61	2.52	2.42
21	3.31	3.17	3.03	2.88	2.80	2.72	2.64	2.55	2.46	2.36
22	3.26	3.12	2.98	2.83	2.75	2.67	2.58	2.50	2.40	2.31
23	3.21	3.07	2.93	2.78	2.70	2.62	2.54	2.45	2.35	2.26
24	3.17	3.03	2.89	2.74	2.66	2.58	2.49	2.40	2.31	2.21
25	3.13	2.99	2.85	2.70	2.62	2.54	2.45	2.36	2.27	2.17
26	3.09	2.96	2.81	2.66	2.58	2.50	2.42	2.33	2.23	2.13
27	3.06	2.93	2.78	2.63	2.55	2.47	2.38	2.29	2.20	2.10
28	3.03	2.90	2.75	2.60	2.52	2.44	2.35	2.26	2.17	2.06
29	3.00	2.87	2.73	2.57	2.49	2.41	2.33	2.23	2.14	2.03
30	2.98	2.84	2.70	2.55	2.47	2.39	2.30	2.21	2.11	2.01
40	2.80	2.66	2.52	2.37	2.29	2.20	2.11	2.02	1.92	1.80
60	2.63	2.50	2.35	2.20	2.12	2.03	1.94	1.84	1.73	1.60
120	2.47	2.34	2.19	2.03	1.95	1.86	1.76	1.66	1.53	1.38
∞	2.32	2.18	2.04	1.88	1.79	1.70	1.59	1.47	1.32	1.00

F-Test:

This is very useful in testing the equality of population means by comparing sample variances.

To test whether there is any significanet difference between two estimates of population variance (σ^2) To test if the two samples have come from the same population, we use the F-test.

In this case we setup null hypothesis

$H_0 = \sigma_1^2 = \sigma_2^2$ i.e population variances are same

under H_0, the test statistic is $F = \frac{S_1^2}{S_2^2}$

where $S_1^2 = \frac{\Sigma(x_i - \bar{x})^2}{n-1}$ [n_1 - first sample size]

$S_2^2 = \frac{\Sigma(y_i - \bar{y})^2}{n-1}$ [n_2 - second sample size]

and $S_1^2 > S_2^2$

The degrees of freedom are $V_1 = n_1 - 1$ & $V_2 = n_2 - 1$

Note: 1, we will take greater of the variances S_1^2 or S_2^2 in the numerator and adjust for the degree of freedom accordingly.

i.e $F = \frac{\text{Greater variance}}{\text{Smaller variance}}$

2, If sample variance S^2 is given, we can obtain population variance σ^2 by $n\sigma^2 = (n-1)S^2$ & vice-versa.

3 $F_\alpha(V_1, V_2)$ is the value of F with ν_1 and ν_2 degrees of freedom such that the area under F-distribution to the right of F_α is α. In tables F_α tabulated for $\alpha = 0.05$ and $\alpha = 0.01$ for various combinations of the d.o.f ν_1 and ν_2.

→ In one sample of 8 observations, the sum of the squares of deviations of sample values from the sample mean was 84.4 and in the other sample of 10 observations was 102.6. Test whether this difference is significant at 5% level, given that the 5% point of F for $n_1 = 7$ and $n_2 = 9$ degrees of freedom is 3.29.

Sol: Here $n_1 = 8$, $n_2 = 10$, $\sum(x_i - \bar{x})^2 = 84.4$, $\sum(y_i - \bar{y})^2 = 102.6$

$$S_x^2 = \frac{\sum(x_i - \bar{x})^2}{n-1} = \frac{84.4}{8-1} = 12.057$$

$$S_y^2 = \frac{\sum(y_i - \bar{y})^2}{n-1} = \frac{102.6}{10-1} = 11.4$$

Null hypothesis $H_0: S_x^2 = S_y^2$ i.e the estimates of variances given by the samples are homogeneous.

Now $F = \frac{S_x^2}{S_y^2} = \frac{12.057}{11.4} = 1.057$

i.e calculateedF = 1.057

Tabulated value of F at 5% level for (7,9) degrees of freedom is 3.29.

i.e $F_{0.05}(7,9) = 3.29$

Since calculated F < tabulated F, we accept the null hypothesis.

→ In one sample of 10 observations from a normal population, the sum of the squares of the deviations of the sample values from the sample mean is 102.4 and in other sample of 12 observations from another normal population, the sum of the squares of the deviations of the sample values from the sample mean is 120.5. Examine whether the two normal populations have the same variance.

Sol:

Given $n_1 = 10$ $n_2 = 12$

$\Sigma(x_i - \bar{x})^2 = 102.4$ $\Sigma(y_i - \bar{y})^2 = 120.5$

$s_1^2 = \frac{\Sigma(x_i - \bar{x})^2}{n-1}$ $s_2^2 = \frac{\Sigma(y_i - \bar{y})^2}{n-1}$

$= \frac{102.4}{9} = 11.37$ $= \frac{120.5}{11} = 10.95$

Null hypothesis H_0: The two normal populations have the same variance

i.e $\sigma_1^2 = \sigma_2^2 = \sigma^2$

The test statistic is $F = \frac{s_1^2}{s_2^2}$

$= \frac{11.37}{10.95} = 1.038$

i.e calculated $F = 1.038$

Tabulated value of F for (9, 11) d.o.f at 5% level of significance = 2.91.

Since calculated F < tabulated F, we accept null hypothesis H_0 i.e the two normal populations have same variances.

→ Pumpkins were grown under two experimental conditions. Two random samples of 11 and 9 pumpkins show the sample standard deviations of their weights as 0.8 and 0.5 respectively. Assuming that the weight distributions are normal, test hypothesis that the true variances are equal.

Sol: Given $n_1 = 11$, $n_2 = 9$, $\sigma_1 = 0.8$, $\sigma_2 = 0.5$

Here the sample S.D's are given.

we have to determine the population variances σ_1^2 & σ_2^2 by using the relations.

$n_1 \sigma_1^2 = (n_1-1)s_1^2$ $\qquad$ $n_2 \sigma_2^2 = (n_2-1)s_2^2$

$s_1^2 = \frac{n_1\sigma_1^2}{n_1-1} = \frac{11(0.8)^2}{10} = 0.704$ $\qquad$ $s_2^2 = \frac{n_2\sigma_2^2}{n_2-1}$

$= \frac{9(0.5)^2}{8} = 0.281$.

Null hypothesis H_0: The variances are equal $\sigma_1^2 = \sigma_2^2$

The test statistic $\therefore$ $F = \frac{s_1^2}{s_2^2} = \frac{0.704}{0.281} = 2.5 \, [\because s_1^2 > s_2^2]$

$\therefore$ calculated $F = 2.5$

Tabulated value of F for (10, 8) d.o.f at 5% level of significance $\therefore 3.35$.

Since calculated F < tabulated F, we accept null hypothesis H_0 i.e the true variances are equal.

→ Two random samples gave the following results:

Sample	size	Sample mean	sum of squares of Deviations from the mean
1	10	15	90
2	12	14	108

Test whether the samples came from the same normal population.

Sol:

Null hypothesis H_0: The two samples have been drawn from the same normal population.

i.e $H_0: \mu_1 = \mu_2$ and $\sigma_1^2 = \sigma_2^2$

Here we have to use two tests

i) To test equality of variances by F-test

ii) to test equality of means by t-test.

i, F- test (equality of variances)

Given $n_1 = 10$ $n_2 = 12$

$\bar{x}_1 = 15$ $\bar{x}_2 = 14$

$\Sigma(x_i - \bar{x})^2 = 90$ $\Sigma(y_i - \bar{y})^2 = 108$

$S_1^2 = \frac{\Sigma(x_i - \bar{x})^2}{n_1 - 1}$ $S_2^2 = \frac{\Sigma(y_i - \bar{y})^2}{n_2 - 1}$

$= \frac{90}{9} = 10$ $= \frac{108}{12-1} = \frac{108}{11} = 9.82$

$\therefore F = \frac{S_1^2}{S_2^2} = \frac{10}{9.82} = 1.018$

i.e calculated $F = 1.018$

Tabulated F at 5% level for (9, 11) degrees of freedom

is 2.90 i.e $F_{(0.05)}(9, 11) = 2.90$

Since calculated F < tabulated F, we accept the hypothesis

i.e $\sigma_1^2 = \sigma_2^2$

i.e the samples come from the same normal populations

with same variance.

ii, t-test (To test equality of means)

Null hypothesis $H_0: \mu_1 = \mu_2$

Given $\bar{x}_1 = 15$ $\bar{x}_2 = 14$

$n_1 = 10$ $n_2 = 12$

$\Sigma(x_i - \bar{x})^2 = 90$ $\Sigma(y_i - \bar{y})^2 = 108$

$S^2 = \frac{1}{n_1 + n_2 - 2}\left[\Sigma(x_i - \bar{x})^2 + \Sigma(y_i - \bar{y})^2\right]$

$= \frac{1}{10 + 12 - 2}[90 + 108] = 9.9$

$S = 3.15$

test statistic : $t = \dfrac{\bar{x}-\bar{y}}{s\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}} = \dfrac{15-14}{2.15\sqrt{\frac{1}{10}+\frac{1}{12}}} = 0.74$

$\therefore$ calculated value of $|t| = 0.74$

Tabulated value of t for 20 d.o.f (n_1+n_2-2) at 5% level of significance : 2.086.

since calculated value of $t <$ tabulated value of t, we accept the null hypothesis H_0: i.e $\mu_1 = \mu_2$.

Hence from i, & ii, the given samples have been drawn from the same normal population.

Hence we accept null hypothesis that $\mu_1 = \mu_2$ and $\sigma_1^2 = \sigma_2^2$.

→ The time taken by workers in performing a job by method I and method II : given below:

Method I	20	16	26	27	23	22	
Method II	27	33	42	35	32	34	38

Do the data show that the variances of time distribution from population from which these samples are drawn do not differ significantly?

Sol: calculation of sample variances

x_i	$x_i-\bar{x}$	$(x_i-\bar{x})^2$	y_i	$y_i-\bar{y}$	$(y_i-\bar{y})^2$
20	-2.3	5.29	27	-7.4	54.76
16	-6.3	39.69	33	-1.4	1.96
26	3.7	13.69	42	7.6	57.76
27	4.7	22.09	35	0.6	0.36
23	0.7	0.49	32	-2.4	5.76
22	-0.3	0.09	34	-0.4	0.16
			38	3.6	12.96
134		81.34	241		133.72

Given $n_1 = 6$ $n_2 = 7$

$$\bar{x} = \frac{\sum x_i}{n_1} = \frac{134}{6} = 22.3$$

$$\bar{y} = \frac{\sum y_i}{n_2} = \frac{241}{7} = 34.4$$

and $\sum(x_i - \bar{x})^2 = 81.34$, $\sum(y_i - \bar{y})^2 = 133.72$

$$\therefore S_1^2 = \frac{\sum(x_i - \bar{x})^2}{n_1 - 1} = \frac{81.34}{5} = 16.26$$

$$\text{and } S_2^2 = \frac{\sum(y_i - \bar{y})^2}{n_2 - 1} = \frac{133.72}{6} = 22.29$$

Null hypothesis H_0: There is no significant difference between the variances. ~~So~~

Since $S_2^2 > S_1^2$, we use

$$F = \frac{S_2^2}{S_1^2} = \frac{22.29}{16.268} = 1.3699 \approx 1.37$$

Tabulated value of F for $(n_1 - 1, n_2 - 1)$ d.o.f i.e $(6, 5)$ d.of at 5% level of significance is 4.95.

Since calculated F < tabulated F, we accept the null hypothesis ~~o~~ H_0 at 5% level of significance i.e there is no significant difference between the variances of the time distribution by the workers.

→ The nicotine contents in milligrams in two samples of tobacco were found to be as follows:

Sample A	24	27	26	21	25	
Sample B	27	30	28	31	32	36

Can it be said that two samples from same normal population.

CHI-SQUARE (χ^2) TEST

Definition

If a set of events $A_1, A_2, A_3, \ldots A_n$ are observed to occur with frequencies $O_1, O_2, O_3 \ldots O_n$ respectively & according to probability rules $A_1, A_2, \ldots A_n$ are expected to occur with frequencies & $E_1, E_2, \ldots E_n$ respectively with $O_1, O_2, \ldots O_n$ are called observed frequencies & $E_1, E_2, \ldots E_n$ are called expected frequencies.

If O_i $(i = 1, 2, \ldots n)$ is a set of observed (experimental) frequencies & E_i $(i = 1, 2, \ldots n)$ is the corresponding set of expected (theoretical) frequencies, then

χ^2 is defined as $$\boxed{\chi^2 = \sum_{i=1}^{n} \frac{(O_i - E_i)^2}{E_i}}$$ with $(n-1)$ d.o.f.

χ^2 is used to test whether differences between observed & expected frequencies are significant.

Note:-

If the data is given in a series of 'n' numbers then

d.o.f $= n-1$

In case of Binomial distribution d.o.f $= n-1$

In case of Poison distribution d.o.f $= n-2$

In case of Normal distribution d.o.f $= n-3$

Since the equation of χ^2-curve does not involve any parameters of the population, this distribution does not depend on the form of the population & is therefore very useful in a large number of problems.

Chi-square Distribution is an important continuous Probability distribution & it is used in both large & small tests. In χ^2 distribution is mainly used

i) to test the goodness of fit.

ii) to test the independence of attributes.

χ^2 Test as a test of Goodness of fit.

We use this test to decide whether the difference between the theoretical & observed values can be attributed to chance or not.

χ^2 distribution is used to test how well an observed frequency distribution fits a theoretical distribution. It is called the test of the goodness of fit.

Let the N.H H_0 be that there is no significant difference between the observed values & the corresponding expected values.

A.H H_1 is that the above difference is significant.

Let $O_1, O_2, \ldots O_n$ be a set of observed frequencies & $E_1, E_2, \ldots E_n$ the corresponding set of expected frequencies. Then the test statistic χ^2 is given by.

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

Test statistic χ^2 follows chi-square distribution with $(n-1)$ d.o.f. where

$$\boxed{\sum_{i=1}^{n} O_i = \sum_{i=1}^{n} E_i}$$

$$\Rightarrow \sum_{i=1}^{n} (O_i - E_i) = 0$$

Conclusion:- If the calculated value of χ^2 > tabulated value of χ^2 at α level, the N.H H_0 is rejected. Otherwise, H_0 is accepted.

χ^2 test enables us to assertain how well the theoretical distributions such as Binomial, Poisson, Normal, etc. fit the distributions obtained from sample data.

Conditions of validity

Following are the conditions which should be satisfied before χ^2 test can be applied.

1) The sample observations should be independent.
2) The total frequency N is large i.e > 50.
3) The constraints on the cell frequencies must be linear.
4) No theoretical (expected) frequency should not be small in less than 10. when the cell frequencies are small in two or more cells, they should be combined in to a single cell.
5) The sum of observed frequencies must be always equal to the sum of expected frequencies $\boxed{\Sigma O_i = \Sigma E_i}$

1) A sample analysis of examination results of 500 students was made. It was found that 220 students had failed, 170 had secured a third class, 90 were placed in second class & 20 got a first class. Do these figures commensurate with the general examination result which is in the ratio of 4:3:2:1 for the various categories respectively.

Solution: Given Total frequency N = 500

Expected frequencies are in the ratio of 4:3:2:1

If we divide the total frequency 500 in the ratio 4:3:2:1, we get the expected frequencies as 200, 150, 100, 50.

i.e $\frac{4}{10} \times 500 = 200$, $\frac{3}{10} \times 500 = 150$, $\frac{2}{10} \times 500 = 100$, $\frac{1}{10} \times 500 = 50$

1) Null Hypothesis (N.H) H_0: The observed results commensurate with the general examination results.

2) Alternative Hypothesis (A.H) H_1: The observed results does not commensurate with the general examination results.

3) L.O.S = α = 5%.

4) Test statistic $\boxed{\chi^2 = \sum_{i=1}^{n} \frac{(O_i - E_i)^2}{E_i}}$

Class	observed frequency (O_i)	Expected frequency (E_i)	$O_i - E_i$	$\frac{(O_i - E_i)^2}{E_i}$
Failed	220	200	20	200
Third	170	150	20	2.667
Second	90	100	-10	1.0
First	20	50	-30	18.00
	500	500		23.667

$$\boxed{\chi^2 = \sum_{i=1}^{n} \frac{(O_i - E_i)^2}{E_i}} = 23.667$$

$d.o.f = \nu = n - 1 = 4 - 1 = 3$

Tabulated value of χ^2 at 5% L.os for $\nu = n-1 = 3$ d.o.f is 7.81

Conclusion:- $\because$ Calculated $\chi^2_{0.05} >$ tabulated $\chi^2_{0.05}$, we reject the N.H i.e the observed results are not commensurate with the general examination results.

2) The following figures show the distribution of digits in numbers chosen at random from a telephone directory.

Digits	0	1	2	3	4	5	6	7	8	9
frequency	1026	1107	997	966	1075	933	1107	972	964	853

Test whether the digits may be taken to occur equally frequently in the directory.

Solution:-

The expected frequencies for each of the digits $0, 1, 2, \ldots 9$ is $\frac{10{,}000}{10} = 1000$.

Digits	observed frequency (O_i)	Expected frequency (E_i)	$(O_i - E_i)^2$	$\frac{(O_i - E_i)^2}{E_i}$
0	1026	1000	676	0.676
1	1107	1000	11449	11.449
2	997	1000	9	0.009
3	966	1000	1156	1.156
4	1075	1000	5625	5.625
5	933	1000	4489	4.489
6	1107	1000	11449	11.449
7	972	1000	784	0.784
8	964	1000	1296	1.296
9	853	1000	21609	21.609
Total	10000	10,000		58.542

Null Hypothesis H_0: The digits occur equally frequently in the directory.

Alternative Hypothesis H_1: The digits do not occur equally frequently.

L.O.S $= \alpha = 5\%$

Test statistic $\boxed{\chi^2 = \Sigma \frac{(O_i - E_i)^2}{E_i}} = 58.542.$

degrees of freedom (d.o.f) $= n-1 = 10-1 = 9$

The tabulated value of χ^2 for 9 d.o.f at 5% L.O.S $= 16.919$.

Conclusion:- $\because$ calculated χ^2 > tabulated χ^2, we reject the N.H H_0 at 5% L.O.S & conclude that the digits do not occur equally frequently in the directory.

3) A pair of dice are thrown 360 times & the frequency of each sum is indicated below:

Sum	2	3	4	5	6	7	8	9	10	11	12
Frequency	8	24	35	37	44	65	51	42	26	14	14

Would you say that the dice are fair on the basis of the chi-square test at 0.05 L.O.S?

Solution:-

The probabilities of getting a sum 2, 3, 4, 5, 6, 7, 8, 9, 10, 11 & 12 are

$X=x_i$	2	3	4	5	6	7	8	9	10	11	12
$P(x_i)$	$1/36$	$2/36$	$3/36$	$4/36$	$5/36$	$6/36$	$5/36$	$4/36$	$3/36$	$2/36$	$1/36$

1) Null Hypothesis: H_0: The dice are fair

2) Alternative Hypothesis H_1: The dice are not fair.

3) L.O.S: $\alpha = 5\%$.

Sum	observed frequency (O_i)	Expected frequency (E_i) $E_i = 360 \cdot P(x)$	$(O_i - E_i)^2$	$\frac{(O_i - E_i)^2}{E_i}$
2	8	$E_i = 360 \times \frac{1}{36} = 10$	4	0.4
3	24	$E_i = 360 \times \frac{2}{36} = 20$	16	0.8
4	35	30	25	0.833
5	37	40	9	0.225
6	44	50	36	0.72
7	65	60	25	0.417
8	51	50	1	0.02
9	42	40	4	0.1
10	26	30	16	0.53
11	14	20	36	1.8
12	14	10	16	1.6
	360	360		7.445

Test statistic $\boxed{\chi^2 = \sum \frac{(O_i - E_i)^2}{E_i}} = 7.445$.

d.o.f $= n - 1 = 11 - 1 = 10$

The tabulated value of χ^2 for 10 d.o.f at 5% L.o.S is 18.307

Conclusion :- $\because$ calculated χ^2 < tabulated χ^2, we accept the N.H H_0 at 5% L.o.S i.e the dice are fair.

4) 4 coins were tossed 160 times & the following results were obtained.

No. of Heads :	0	1	2	3	4
observed Frequencies :	17	52	54	31	6

under the assumption that coins are balanced, find the expected frequencies of 0, 1, 2, 3, or 4 heads. & test the goodness of fit ($\alpha = 0.05$).

Solution:- Given $N = 160$, $n = 4$.

p = Probability of getting a head $= \frac{1}{2}$.

$q = 1 - p = 1 - \frac{1}{2}$.

Binomial distribution $\boxed{P(x) = nC_r\, p^r\, q^{n-r}}$.

Expected frequency $\boxed{N \cdot P(r)}$.

The expected frequencies of getting 0, 1, 2, 3, 4 are the successive terms of the binomial expansion

$= \boxed{160 \cdot P(r)}$.

$= 160\left[P(r=0) + P(r=1) + P(r=2) + P(r=3) + P(r=4)\right]$

$= 160\left[{}^4C_0\left(\frac{1}{2}\right)^4 + 4C_1\left(\frac{1}{2}\right)^1\left(\frac{1}{2}\right)^3 + 4C_2\left(\frac{1}{2}\right)^2\left(\frac{1}{2}\right)^2 + 4C_3\left(\frac{1}{2}\right)^3\left(\frac{1}{2}\right) + 4C_4\left(\frac{1}{2}\right)^4\right]$

$= 160\left[\left(\frac{1}{2}\right)^4 + 4\left(\frac{1}{2}\right)^4 + 6\left(\frac{1}{2}\right)^4 + 4\left(\frac{1}{2}\right)^4 + \left(\frac{1}{2}\right)^4\right]$.

$= 160\left(\frac{1}{2}\right)^4\left[1 + 4 + 6 + 4 + 1\right]$

$= \cancel{160}\left(\frac{1}{\cancel{16}}\right)\left[\cancel{16}\right] = \left[10 + 40 + 60 + 40 + 10\right]$.

$\therefore$ The theoretical or expected frequencies are 10, 40, 60, 40, 10.

$$\therefore \chi^2 = \Sigma \frac{(O_i - E_i)^2}{E_i}$$

$$= \frac{(17-10)^2}{10} + \frac{(52-40)^2}{40} + \frac{(54-60)^2}{60} + \frac{(31-40)^2}{40} + \frac{(6-10)^2}{10}$$

$$= \frac{1}{120}(588+432+72+243+192)$$

$$= \frac{1527}{120} = 12.725$$

Degrees of freedom $(\nu) = n-1 = 5-1 = 4$

$\therefore \chi^2_{0.05} = 9.49$ (From tables).

Conclusion :- $\because$ the calculated value > Tabulated value

$\therefore$ we conclude that the hypothesis that the data follows the binomial is rejected.

5) A survey of 800 families with four children each revealed the following distribution:

No. of boys :	0	1	2	3	4
No. of girls :	4	3	2	1	0
No. of families :	32	178	290	236	64

Is this result consistent with the hypothesis that male & female births are equally probable?.

Solution:-

Null Hypothesis : The data are consistent with the hypothesis of equal probability for male & female births i.e P = Prob. of male birth $= \frac{1}{2} = q$.

Alternative Hypothesis : $P \neq \frac{1}{2}$.

L.O.S : $\alpha = 5\%$

$\boxed{P(r) = nc_r \, p^r \, q^{n-r}}$.

$\boxed{\text{Expected frequency} = N \cdot P(r)}$

Given $n = 4$, $N = 800$.

$p = \frac{1}{2}$, $q = \frac{1}{2}$

If $r=0$, $P(r) = P(r=0) = 4C_0 (\frac{1}{2})^0 (\frac{1}{2})^4$.

$= \frac{1}{16}$.

$E.F = N.P(r) = 800 \times \frac{1}{16} = 50$.

If $r=1$, $P(r) = P(r=1) = 4C_1 (\frac{1}{2})^1 (\frac{1}{2})^3$.

$= 4 \times \frac{1}{16} = \frac{1}{4}$.

$E.F = N.P(r) = 800 \times \frac{1}{4} = 200$.

If $r=2$, $P(r) = P(r=2) = 4C_2 (\frac{1}{2})^2 (\frac{1}{2})^2$.

$= 6 \times (\frac{1}{16}) = \frac{3}{8}$.

$E.F = N.P(r) = 800 \times \frac{3}{8} = 300$

If $r=3$, $P(r) = P(r=3) = 4C_3 (\frac{1}{2})^3 (\frac{1}{2})$

$= 4 (\frac{1}{16}) = \frac{1}{4}$.

$E.F = N.P(r) = 800 \times \frac{1}{4} = 200$

If $r=4$, $P(r) = P(r=4) = 4C_4 (\frac{1}{2})^4 (\frac{1}{2})^0$

$= \frac{1}{16}$.

$E.F = N.P(r) = 800 \times \frac{1}{16} = 50$.

Calculations for χ^2

No. of male births	observed frequency (O_i)	Expected frequency (E_i)	$(O_i - E_i)^2$	$\frac{(O_i - E_i)^2}{E_i}$
0	32	50	324	6.48
1	178	200	484	2.42
2	290	300	100	0.33
3	236	200	1296	6.48
4	64	50	196	3.92
Total	800	800		19.63

$\therefore \boxed{\chi^2 = \sum \frac{(O_i - E_i)^2}{E_i}} = 19.63$.

$\boxed{d.o.f\,(\nu) = n - 1} = 5 - 1 = 4$.

Tabulated $\chi^2_{0.05}$ for 4 d.o.f is 9.488

conclusion:- $\because$ calculated value of $\chi^2 >$ Tabulated value.

$\therefore$ we reject the Null Hypothesis at 5% L.O.S & we conclude that male & female births are not equally probable.

6) When the first proof of 392 pages of a book of 1200 pages were read, the distribution of printing mistakes were found to be as follows:

No. of mistakes in a page (x) :	0	1	2	3	4	5	6
No. of pages (f) :	275	72	30	7	5	2	1

Fit a poisson distribution to the above data & test the goodness of fit.

Solution :

Null Hypothesis : "From the data, we take the mean (parameter) λ of the poisson distribution equal to the mean of the given distribution".

Alternative Hypothesis : The mean of the poisson distribution is not equal to the mean of the given distribution.

L.O.S : $\alpha = 5\%$.

$\boxed{P(r) = \frac{e^{-\lambda} \lambda^r}{r!}}$, Expected frequency $= N \cdot P(r)$.

λ is the A.M of the Poisson distribution.

In order to fit a poisson distribution to the given data, we take the mean of the poisson distribution equal to the mean of the given distribution.

i.e $\boxed{\bar{x} = \cancel{\sum f(x)} \frac{\sum fx}{N}} = \frac{189}{392} = 0.482$

$\Rightarrow \lambda = \bar{x} = 0.482$

when $r=0$, $P(r) = P(r=0) = \frac{e^{-0.482} \times (0.482)^0}{0!} = 0.6176.$

$E.F. = N.P(r) = 392 \times 0.6176 = 242.1$

when $r=1$, $P(r=1) = \frac{e^{-0.482} \times (0.482)^1}{1!} = 242.1 \times 0.482.$

$E.F = N.P(r) = 392 \times 242.1 \times 0.482$

$= 116.69.$

when $r=2$, $P(r=2) = \frac{e^{-0.482} \times (0.482)^2}{2!}$

$E.F = N.P(r=2) = \frac{392 \times 242.1 \times 0.482}{2}$

$= 28.12$

when $r=3$, $P(r=3) = \frac{e^{-0.482} (0.482)^3}{3!}$

$E.F = N.P(r=3) = 4.518.$

when $r=4$, $P(r=4) = \frac{e^{-0.482} (0.482)^4}{4!}$

$E.F = N.P(r=4) = 0.544$

when $r=5$, $P(r=5) = \frac{e^{-0.482} (0.482)^5}{5!}$

$E.F = N.P(r=5) = 0.052$

when $r=6$, $P(r=6) = \frac{e^{-0.482} (0.482)^6}{6!}.$

$E.F = N.P(r=6) = 0.004$

∵ frequencies are always integers by converting them to nearest integers, we get:

x	0	1	2	3	4	5	6
observed frequency	275	72	30	7	5	2	1
Expected frequency (theoretical frequency)	242	117	28	5	0.5	0.1	0

calculations for χ^2.

Mistakes per page (x)	observed frequency (O_i)	Expected frequency (E_i)	$(O_i - E_i)^2$	$\frac{(O_i - E_i)^2}{E_i}$
0	275	242	1089	4.5
1	72	117	2025	17.3
2	30	28	4	0.14
3	7	5		
4	5 } 15	0.5 } 5.6	88.36	15.77
5	2	0.1		
6	1	0		
				37.71

$$\boxed{\chi^2 = \Sigma \frac{(O_i - E_i)^2}{E_i}} = 37.71$$

$$\boxed{d.o.f = \nu = n - 2} = 4 - 2 = 2.$$

Conclusion:- Tabulated value of χ^2 for 2 d.of at 5% L.O.S is = 5.99 ∵ calculated value > Tabulated value, we reject the N.H at 5% L.O.S & conclude that Poisson distribution is not a good fit to the given data.

CHI-SQUARE TEST FOR INDEPENDENCE OF ATTRIBUTES

An attribute means a quality or characteristic. Examples of attributes are drinking, smoking, blindness, honesty, etc.

Let the observations be classified according to two attributes & the frequencies O_i in the different categories be shown in a two-way table called contigency table. We have to test on the basis of cell frequencies whether the two attributes are independent or not.

Null Hypothesis H_0: The two attributes are independent.
Alternative Hypothesis H_1: The two attributes are not independent.

L.O.S : $\alpha = 5\%$ or 1%.

O_i - observed frequency
E_i - Expected frequency.

Test statistic $\boxed{\chi^2 = \Sigma \frac{(O_i - E_i)^2}{E_i}}$

d.o.f = (No. of rows − 1) × (No. of columns − 1).

The expected frequencies (E_i) of any cell

$$\boxed{E_i = \frac{\text{Row total} \times \text{Column total}}{\text{Grand total}}}$$

If the calculated value of x^2 is less than the table value at a specified level of significance. the hypothesis holds good. i.e the attributes are independent. other wise we reject the N.H.

Let us consider two attributes A & B. A is divided in to two classes & B is divided in to two classes. The various cell frequencies can be expressed in the following table known as 2x2 contigency table:

A	a	b	a+b
B	c	d	c+d.
	a+c	b+d	N=a+b+c+d

The expected frequencies are given by

$E(a) = \frac{(a+b)(a+c)}{N}$	$E(b) = \frac{(a+b)(b+d)}{N}$
$E(c) = \frac{(c+d)(a+c)}{N}$	$E(d) = \frac{(c+d)(b+d)}{N}$

1) On the basis of information given below about the treatment of 200 patients suffering from a disease, state whether the new treatment is comparatively superior to the conventional treatment.

	Favourable	Not favourable	Total
New	60	30	90
conventional	40	70	110

Solution:-

a	b	a+b
c	d	c+d
a+c	b+d	N = a+b+c+d

The expected frequencies are given by

$E(a) = \frac{(a+c)(a+b)}{N}$	$E(b) = \frac{(b+d)(c+d)}{N}$	a+b
$E(c) = \frac{(a+c)(c+d)}{N}$	$E(d) = \frac{(b+d)(c+d)}{N}$	c+d
a+c	b+d	N

$$\boxed{\text{Expected frequency} = \frac{\text{Row total} \times \text{Column total}}{\text{Grand total}}}$$

			Total.
New	60 (a)	30 (b)	90 (a+b)
Conventional	40 (c)	70 (d)	110
Total	100 (a+c)	100 (b+d)	200 (N)

Expected frequency = $\dfrac{\text{Row total} \times \text{column total}}{\text{Grand total}}$.

$E(60) = \dfrac{90 \times 100}{200} = 45$	$E(30) = \dfrac{90 \times 100}{200} = 45$	90
$E(40) = \dfrac{100 \times 110}{200} = 55$	$E(70) = \dfrac{100 \times 110}{200} = 55$	110
100	100	200

Null Hypothesis H_0: No difference between new & conventional treatment.

Alternative Hypothesis H_1: Difference between New & conventional treatment.

L.O.S: $\alpha = 5\%$

Test statistic: $\boxed{\chi^2 = \Sigma \dfrac{(O-E)^2}{E}}$

calculation of χ^2.

Observed frequency (O_i)	Expected frequency (E_i)	$(O_i - E_i)^2$	$\dfrac{(O_i - E_i)^2}{E_i}$
60	45	225	5
30	45	225	5
40	55	225	4.09
70	55	225	4.09
200	200		18.18

$\therefore \boxed{\chi^2 = \Sigma \frac{(O-E)^2}{E} = 18.18}$

Tabulated χ^2 for 1 d.o.f at 5% L.O.S is 3.841

Conclusion:- $\because$ calculated χ^2 > tabulated χ^2.

$\therefore$ we reject the N.H H_0 at 5% L.O.S. & we conclude that the new treatment is comparatively superior to conventional treatment.

2) Two sample polls of votes for two candidates A & B for a public office are taken, one from among the residents of rural areas. The results are given in the adjoining table. Examine whether the nature of the area is related to voting preference in this election.

Area	Votes for A	Votes for B	Total
Rural	620	380	1000
Urban	550	450	1000
Total	1170	830	2000

Solution:-

$$\boxed{\text{Expected frequency} = \frac{\text{Row total} \times \text{Column total}}{\text{Grand total}}}$$

$E(620) = \frac{1000 \times 1170}{2000} = 585$,	$E(380) = \frac{1000 \times 830}{2000} = 415$	1000
$E(550) = \frac{1000 \times 1170}{2000} = 585$	$E(450) = \frac{1000 \times 830}{200} = 415$	1000
Total 1170	830	2000

Null Hypothesis: The nature of the area is related to voting preference in the election.

Alternative Hypothesis: The nature of the area is not related to voting preference in the election.

L.O.S = α = 5%.

Test statistic: $\boxed{\chi^2 = \Sigma \frac{(O_i - E_i)^2}{E}}$.

Calculation of χ^2.

observed frequency (O_i)	Expected frequency (E_i)	$(O_i - E_i)^2$	$\frac{(O_i - E_i)^2}{E_i}$
620	585	1225	2.094
380	415	1225	2.951
550	585	1225	2.094
450	415	1225	2.951
2000	2000		10.09

$\therefore \chi^2 = \sum \frac{(O_i - E_i)^2}{E_i} = 10.09$

d.o.f = (No. of rows − 1) × (No. of columns − 1).

$= (2-1) \times (2-1) = 1$

Tabulated χ^2 for 1 d.o.f at 5% L.o.S is 3.841

Conclusion :- $\because$ calculated χ^2 > tabulated χ^2.

$\therefore$ we reject the N.H H_0 at 5% L.O.S.

3) Out of 8,000, graduates in a town 800 are females, out of 1,600 graduate employees 120 are females. Use χ^2 to determine if any distinction is made in appointment on the basis of sex. value of χ^2 at 5% level for one degree of freedom is 3.84.

Solution :-

Table on observed frequencies

	Employed	unemployed	Total
Male	1480	5720	7200
Female	120	680	800.
Total	1600	6400	8000

$$\boxed{\text{Expected frequency} = \frac{\text{Row total} \times \text{Column total}}{\text{Grand total}}}$$

Expected frequencies

$E(1480) = \frac{7200 \times 1600}{8000} = 1440$	$E(5720) = \frac{7200 \times 6400}{8000} = 5760$
$E(120) = \frac{800 \times 1600}{8000} = 160$	$E(680) = \frac{800 \times 6400}{8000} = 640$
1600.	6400 = 800

calculations for χ^2.

observed frequency (O_i)	Expected frequency (E_i)	$(O_i - E_i)^2$	$\frac{(O_i - E_i)^2}{E_i}$
1480	1440	1600	1.11
5720	5760	1600	0.28
120	160	1600	10.0
680	640.	1600	2.50
8000	8000		13.89.

$$\boxed{\chi^2 = \Sigma \frac{(O_i - E_i)^2}{E_i}} = 13.89$$

Null Hypothesis: No distinction is made in appointment on the basis of sex.

Alternative Hypothesis: Distinction is made in appointment on the basis of sex.

L.O.S : $\alpha = 5\%$

Test statistic $\boxed{\chi^2 = \Sigma \frac{(O_i - E_i)^2}{E_i}} = 13.89$.

$\boxed{d.o.f = (\text{No. of rows} - 1) \times (\text{No. of Columns} - 1)}$.

conclusion:- Tabulated χ^2 for 1 d.o.f at 5% L.O.S is 3.841.

$\therefore$ calculated χ^2 > tabulated χ^2.

$\therefore$ we reject the N.H H_0 at 5% L.O.S & we conclude that distinction is made in appointment on the basis of sex.

~~From the following data, find whether there is any significant liking in the habit of taking soft drinks among the categories of employees.~~

4) A random sample of students of XYZ university was selected and asked their opinions about 'autonomous colleges'. The results are given below. The same number of each sex was included within each class-group. Test the hypothesis at 5% level that opinions are independent of the class groupings:

class	Numbers: Favouring autonomous colleges	Numbers: opposed to autonomous colleges	Total
B.A/B.Sc/B.Com Part I	120	80	200
B.A/B.Sc/B.Com Part II	130	70	200
B.A/B.Sc/B.Com Part III	70	30	100
M.A/M.Sc/M.Com	80	20	100
Total	400	200	600

Solution :-

Expected frequency = $\frac{\text{Row total} \times \text{Column total}}{\text{Grand total}}$

		Total
$E(120) = \frac{200 \times 400}{600} = 133.33$	$E(80) = \frac{200 \times 200}{600} = 66.67$	200
$E(130) = \frac{200 \times 400}{600} = 133.33$	$E(70) = \frac{200 \times 200}{600} = 66.67$	200
$E(70) = \frac{100 \times 400}{600} = 66.67$	$E(30) = \frac{100 \times 200}{600} = 33.33$	100
$E(80) = \frac{100 \times 400}{600} = 66.67$	$E(20) = \frac{100 \times 200}{600} = 33.33$	100
400	200	600

Calculations for χ^2

Observed frequency (O_i)	Expected frequency (E_i)	$(O_i - E_i)^2$	$\frac{(O_i - E_i)^2}{E_i}$
120	133.33	177.688	1.332
130	133.33	11.088	0.083
70	66.67	11.088	0.166
80	66.67	177.688	2.665
80	66.67	177.688	2.665
70	66.67	11.088	0.166
30	33.33	11.088	0.332
20	33.33	177.688	5.331
400	400		12.74

Null Hypothesis: The opinions about autonomous colleges are independent of the class-groupings

Alternative Hypothesis:- The opinions about autonomous colleges are not independent of the class-groupings

L.O.S : $\alpha = 5\%$

Test statistic $\boxed{\chi^2 = \Sigma \frac{(O_i - E_i)^2}{E_i}}$

$= 12.74$

d.o.f = (Rows − 1) × (No. of columns − 1)

$= (4-1)(2-1) = 3.$

Tabulated value of χ^2 for 3 d.o.f at 5% L.O.S is 7.815.

Conclusion: ∵ calculated value of χ^2 > Tabulated χ^2.

∴ we reject the N.H H_0 at 5% L.O.S. i.e the opinions about autonomous college are dependent on the class-groupings.

5) Two researchers adopted different sampling techniques while investigating the same group of students to find the number of students falling in different intelligence levels. The results are as follows.

Researcher	No. of students in each level				Total
	Below Average	Average	Above Average	Genius.	
X	86	60	44	10	200
Y	40	33	25	2	100.
Total	126	93	69	12	300

Would you say that the sampling techniques adopted by the two researchers are significantly different?

(Given 5% value of χ^2 for 2 d.o.f & 3 d.o.f are 5.991 & 7.82 respectively).

Solution :-

$$\boxed{\text{Expected Frequency} : \frac{\text{Row total} \times \text{Column total}}{\text{Grand total}}}$$

$E(86) = \frac{200 \times 126}{300} = 84$		$E(60) = \frac{200 \times 93}{300} = 62$	$E(44) = \frac{200 \times 69}{300} = 46$	$200 - 192 = 8$
	$126 - 84 = 42$	$93 - 62 = 31$	$69 - 46 = 23$	$12 - 8 = 4.$
Total	126	93	69	12 $= 300$

calculations for χ^2.

observed frequency (O_i)	Expected frequency (E_i)	$(O_i - E_i)^2$	$\frac{(O_i-E_i)^2}{E_i}$
86	84	4	0.048
60	62	4	0.064
44	46	4	0.087
10	8	4	0.500
40	42	4	0.095
33	31	4	0.129
25, 2 } 27	23, 4 } 27	0	0
300	300		0.923

$$\boxed{\chi^2 = \sum \frac{(O_i - E_i)^2}{E_i}} = 0.923.$$

Null Hypothesis: There is no significant difference between the sampling techniques used by the two researchers for collecting the required data.

Alternative Hypothesis :- There is significant difference between the sampling techniques used by the two researchers for collecting the required data.

L.O.S : $\alpha = 5\%$

Test Statistic $\boxed{\chi^2 = \sum \frac{(O_i - E_i)^2}{E_i}} = 0.923$.

d.o.f $= [(2-1) \times (4-1)] - 1 = 3 - 1 = 2$.

$\because$ 1 d.o.f is lost in the method of pooling (grouping two values).

Tabulated value of x^2 for 2 d.o.f at 5% L.O.S is 5.991 .

conclusion:- ∴ Calculated value of x^2 < Tabulated x^2.

∴ we accept the N.H at 5% L.O.S. & we conclude that there is no significant difference in the sampling techniques used by the two researchers.

6) From the following data, find whether there is any significant liking in the habit of taking soft drinks among the categories of employees.

Employees

soft drinks	clerks	Teachers	officers
Pepsi	10	25	65
Thumpsup	15	30	65
Fanta	50	60	30

solution:-

$$\boxed{\text{Expected frequency} = \frac{\text{Row total} \times \text{Colum total}}{\text{Grand total}}}$$

soft drinks	clerks	Teachers	officers	Total
Pepsi	10	25	65	100
Thumsup	15	30	65	110
Fanta	50	60	30	140
Total	75	115	160	350

$E(10) = \frac{100 \times 75}{350} = 21.4$	$E(25) = \frac{100 \times 115}{350} = 32.9$	$E(65) = \frac{100 \times 160}{350} = 45.7$
$E(15) = \frac{110 \times 75}{350} = 23.6$	$E(30) = \frac{110 \times 115}{350} = 36.1$	$E(65) = \frac{110 \times 160}{350} = 50.3$
$E(50) = \frac{140 \times 75}{350} = 30$	$E(60) = \frac{140 \times 115}{350} = 46$	$E(30) = \frac{140 \times 160}{350} = 64$

calculations for x^2.

Observed frequency (O_i)	Expected frequency (E_i)	$(O_i - E_i)^2$	$\frac{(O_i - E_i)^2}{E_i}$
10	21.4	129.96	6.073
25	32.9	62.41	1.897
65	45.7	372.49	8.151
15	23.6	73.96	3.134
30	36.1	37.21	1.031
65	50.3	216.09	4.3
50	30	400	13.333
60	46	196	4.261
30	64	1156	18.062
350	350.		60.2425

Null Hypothesis H_0: There is a significant liking in the habit of taking soft drinks.

Alternative Hypothesis H_1: There is no significant liking.

L.O.S : $\alpha = 5\%$

Test statistic $\boxed{\chi^2 = \Sigma \frac{(O_i - E_i)^2}{E_i}}$

$= 60.2425$.

conclusion :-

d.o.f = (No. of rows − 1) × (No. of columns − 1).

$= (3-1)\ (3-1) = 4$

Tabulated χ^2 for 4 d.o.f at 5% L.O.S is 9.488.

$\because$ calculated χ^2 > Tabulated χ^2.

$\therefore$ we reject the N.H H_0 at 5% L.O.S.