30/6/22

UNIT-4

ESTIMATION & TESTING OF HYPOTHESIS

Estimation

Def: Estimation is a method or process of estimating population parameter (mean, variance, S.D, etc.) with the help of sample information

→ Estimation is classified into 2 types:

i) point estimation

ii) Interval estimation

Point Estimation

Point Estimation of population parameter is estimated by a single numerical value by the sample information

Interval Estimation

Interval Estimation explains how to find the interval whose end values are estimated by sample data

Formulas

1) Maximum Error : E_{max} (or) $E = Z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}$

Sample Size $(n) = \left(Z_{\alpha/2} \cdot \frac{\sigma}{E} \right)^2$

confidence intervals or confidence limits

$$\left(\bar{x} \pm Z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}} \right)$$

NOTE If population S.D (σ) is not given we replace by sample S.D (S)

Critical Values for Z (for two tail)

level (α)	value of Z
1%	2.58
2%	2.33
5%	1.96
10%	1.645

Hypothesis

While taking a decision about population we make assumptions about the population parameters. Such assumptions or statements is known as Hypothesis.

→ A Hypothesis is may or may not be true.

The Procedure which enables to decide whether the

→ Sample size $(n) = \left(\frac{Z_{\alpha/2} \cdot \sigma}{E} \right)^2$

→ confidence intervals or confidence limits

$$\left(\bar{x} \pm Z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}} \right)$$

NOTE: If population S.D (σ) is not given use replace by sample S.D (S)

Critical Values for Z (for two tail)

level (α)	value of Z
1%	2.58
2%	2.33
5%	1.96
10%	1.645

Hypothesis

While taking a decision about population we make assumptions about the population parameters. Such assumptions or statements is known as Hypothesis.

→ A Hypothesis is may or may not be true.

The Procedure which enables to decide whether the hypothesis is true or false is known as testing of hypothesis.

Example: 1) majority of the men in hyderabad are smokers.

2) Most of the people in Delhi are chapathi eaters.

→ Hypothesis is classified into 2 types:

i) Null Hypothesis

ii) Alternate Hypothesis

Null Hypothesis:

For applying any test of significance, we setup a hypothesis which is a definite statement about population. This Hypothesis is also known as Hypothesis of no difference. We call it as null hypothesis. It is represented by H_0

Example: To check the significant difference b/w $\bar{x}, \mu$ then null hypothesis is $\bar{x} = \mu$

Alternate Hypothesis:

A Hypothesis which contradicts null hypothesis is known as Alternate Hypothesis. It is represented by H_1.

Example: If we want to check significant diff b/w $\bar{x}, \mu$ then alternate hypothesis $\bar{x} \neq \mu$

5/1/23 **One tailed test**

In testing of hypothesis, if the alternate hypothesis is either right tail or left tail is known as one tailed test

Example: If we want to check the significant diff b/w $\bar{x}, \mu$ the alternate hypothesis which is defined as $\bar{x} < \mu$ (Left tail) or $\bar{x} > \mu$ (right tail) is known as one-tailed test.

Two tailed test

In testing of hypotheses, if the alternate hypothesis is both right tail and left tail is known as two tailed test.

Example: If we want to check the significant diff b/w $\bar{x}, \mu$ the alternate hypothesis which is defined as $\bar{x} \neq \mu$. is known as two-tailed test.

Type - I Error

Reject null hypothesis (H_0) when it is true

i.e. the null hypothesis H_0 is rejected but it is true

The error so defined is known as type-I error.

Type-II Error

Accept Null hypothesis H_0 when it is false.
The null hypothesis is accepted but it is false.
The error so obtained is known as type-II error.

Critical Values of Z

Level α	I-Tail	II-Tail
1%	2.33	2.58
2%	1.96	2.33
5%	1.645	1.96
10%	1.28	1.645

Procedure for testing of hypothesis

we have following steps in testing of hypothesis

1. Null Hypothesis
2. Alternate Hypothesis
3. Level of Significance
4. Test Statistic
5. Conclusion

Null Hypothesis

Define or set up a null Hypothesis (H_0) taking into consideration the nature of the problem and data involved.

Alternate Hypothesis

Define Alternate Hypothesis H_1 which contradicts null Hypothesis. Based on Alternate hypothesis we decide the data is one-tail or two-tail

Level of Significance

The level of significances is generally specified before a test procedure so that the results obtained may not influence our decision.

Generally, we take level 1%, 2%, 5%, 10%.

Test Statistic

Based on the given data we choose suitable test statistics. By using this we get the calculated value of the data.

Conclusion

In the conclusion we decide whether the null hypothesis is correct or wrong using the following cases:

case i : If the calculated value is less than the tabulated value at given level we accept null hypothesis.

case ii : If the calculated value is greater than the tabulated value at given level we reject null hypothesis

* Assuming that $\sigma = 20$ how large a random sample be taken to assert with probability 0.95 that the sample mean will not differ from the true mean by more than 3 points.

Given that $\sigma = 20$

$\alpha = 5\%$

$E = 3$

we know that, max error $E = Z_{\alpha/2} \frac{\sigma}{\sqrt{n}}$

$$n = \left(Z_{\alpha/2} \frac{\sigma}{E}\right)^2$$

$$n = \left(1.96 \times \frac{20}{3}\right)^2$$

$$n = 170.76 \simeq 171$$

* It is desired to estimate the mean no. of hours of continuous use until a certain computer will

first require repair. If it can be assumed that $\sigma = 48$ hours, how large is a sample we needed so that one will be able to assert 90% of confidence that the sample mean is atmost 10 hours.

Given that $\sigma = 48$

$\alpha = 10$

$E = 10$

wkt $E = Z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}$

$$n = \left(Z_{\alpha/2} \frac{\sigma}{E}\right)^2$$

$$n = \left(1.645 \times \frac{48}{10}\right)^2$$

$$n = 62.34 \approx 62$$

7/1/22 * What is the maximum error one can expect to make with probability 0.90 when using the mean of a random sample of size $n = 64$ to estimate the mean of population with $\sigma^2 = 2.56$

Given that $\alpha = 10\%$

$n = 64$

$\sigma^2 = 2.56 \Rightarrow \sigma = 1.6$

$\therefore$ Max Error $E = Z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}} = 1.645 \times \frac{1.6}{\sqrt{64}}$

$E = 0.329$

* A random sample of size 100 has a S.D of 5. what can you say about maximum error with 95% of confidence

Given that $n=100$, $\alpha=5\%$, $s=5$

$$E = Z_{\alpha/2} \cdot \frac{s}{\sqrt{n}}$$

$$E = 1.96 \times \frac{5}{10} = 0.98$$

* The mean & S.D of a population are 11,795 & 14,054 resp what can one assert with 95% of confidence about the maximum error. If $\bar{x}=11,795$ & $n=50$ construct 95% confidence interval.

Given that $\bar{x} = \mu = 11,795$

$\sigma = 14,054$

$n = 50$

$\alpha = 5\%$

wkt, maximum error $E = Z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}$

$$E = 1.96 \times \frac{14054}{\sqrt{50}} = 3895.57$$

IIIly the confidence intervals are

$$\bar{x} \pm Z_{\alpha/2} \frac{\sigma}{\sqrt{n}}$$

$\Rightarrow 11,795 \pm 3895.57$

$\Rightarrow 15690.57, 7899.413$

$\Rightarrow (7899.413, 15690.57)$

* A random sample of size 81 was taken whose variance is 20.25 and the mean is 32. Construct 98% confidence interval and find maximum error.

$\alpha = 2\%$.

Given that $n = 81$

$\bar{x} = 32$

$s^2 = 20.25 \Rightarrow s = 4.5$

wkt, maximum Error $E = Z_{\alpha/2} \frac{s}{\sqrt{n}}$

$E = 2.33 \times \frac{4.5}{\sqrt{81}} = 1.165$

Illy the confidence intervals are

$\bar{x} \pm Z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}$

32 ± 1.165

$\Rightarrow (30.835, 33.165)$

* A research worker wants to determine the average time it takes a mechanic to rotate the tyres of a car and he wants to be able to assert with 95% confidence. The mean of this sample is

off by atmost 0.5 minutes. If we can presume from past experience that $\sigma = 1.6$ minutes. How large a sample will have to take.

Given that $E = 0.5$, $\sigma = 1.6$, $n = ?$, $\alpha = 5\%$

wkt $E = z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}$

$$n = \left(z_{\alpha/2} \cdot \frac{\sigma}{E}\right)^2 = \left(1.96 \times \frac{1.6}{0.5}\right)^2 = 39.33 \simeq$$

<u>Method-1</u> : Test of for significance for Single mean of Large Sample

1. we apply this method to check whether the population mean (μ) has a specified value (μ_0) or not.

2. we use the test statistic $|z| = \frac{|\bar{x} - \mu|}{\frac{\sigma}{\sqrt{n}}}$

3. If population S.D 'σ' is not given then replace it by sample SD (S)

4. If calculated value < tabulated value then accept null hypothesis.

5. If calculated value exceeds tabulated value then reject null hypothesis

1/7/22 * According to the Norm's established for a mechanical appitude test, persons who are 18 yrs old have an average height of 73.2 cms with a S.D of 8.6 cms. If a sample of 40 students of that same age group have the average height 76.7 cms. Test the hypothesis whether $\mu = 73.2$ cms or not at 5% level.

Given that $n = 40$

$\bar{x} = 76.7$

$\mu = 73.2$

$\sigma = 8.6$

1) Null Hypothesis : Let $\mu = 73.2$

2) Alternate Hypothesis : $\mu \neq 73.2$ (Two Tail Test)

3) Level of Significance : $\alpha = 5\%$

4) Test statistic : $|z| = \dfrac{|\bar{x} - \mu|}{\frac{\sigma}{\sqrt{n}}}$

$$\Rightarrow \frac{76.7 - 73.2}{\left(\frac{8.6}{\sqrt{40}}\right)} = 2.57$$

5) Conclusion : Here calculated value of z (2.57) is greater than the tabulated value (1.96) for two Tail Test at 5% level

The Null hypothesis is rejected i.e. $\mu \neq 73.2$

* A sample of 400 items is taken from a population whose SD is 10. The mean of the sample is 40. Test whether the sample has come from a population with mean 38. Also calculate 95% confidence interval

Given that $n = 400$

$\sigma = 10$

$\bar{x} = 40$

$\mu = 38$

$\alpha = 5\%$

i) Null Hypothesis. Let $\mu = 38$

ii) Alternate Hypothesis: $\mu \neq 38$ (Two tail test)

iii) Level of significance: $\alpha = 5\%$

iv) Test statistic: $|z| = \dfrac{|\bar{x} - \mu|}{\frac{\sigma}{\sqrt{n}}} = \dfrac{40 - 38}{\frac{10}{\sqrt{400}}} = 4$

v) Conclusion: Here calculated value of Z (4) is more than the tabulated value (1.96) for Two tail test at 5% level.

$\therefore$ The Null Hypothesis is rejected i.e. $\mu \neq 38$

W.K.T confidence interval is $\left(\bar{x} \pm z_{\alpha/2} \frac{\sigma}{\sqrt{n}}\right)$

$$\left(40 \pm (1.96) \cdot \frac{10}{\sqrt{400}}\right) = (39.02, 40.98)$$

* An Ambulance service claims that it takes on average less than 10 min to reach the destination in emergency calls. A sample of 36 calls as a mean of 11 mins & the variance of 16 mins. Test the claim at 5% level.

Given that $\mu = 10$

$n = 36$

$\bar{x} = 11$

$s^2 = 16 \Rightarrow s = 4$

i) Null Hypothesis: $\mu = 10$

ii) Alternate Hypothesis: $\mu < 10$ (one tail test)

iii) LOS: α: 5%

iv) Test statistics: $|z| = \frac{|\bar{x} - \mu|}{\frac{s}{\sqrt{n}}} = \frac{11-10}{\frac{4}{6}} = 1.5$

v) Conclusion: Here calculated value of z (1.5) is less than the tabulated value (1.645) for one-tail test at 5% level.

$\therefore$ The Null Hypothesis is accepted i.e. $\mu = 10$

8/7/23 Method 2: Test of significance of difference of means of large samples

1. We apply this method to check whether the two samples drawn from same population or not.

2. We use the test statistics $|Z| = \frac{|\bar{x}_1 - \bar{x}_2|}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$

3. If population variances (σ_1^2, σ_2^2) are not given, replace with sample variance (s_1^2, s_2^2)

NOTE: If only one population variance σ^2 is given then assume $\sigma_1^2 = \sigma_2^2 = \sigma^2$

* The means of two large samples of sizes 1000 and 2000 members are 67.5 inches and 68.0 inches resp. Can the samples be regarded as drawn from the same population of S.D 2.5 inches. Given that

$n_1 = 1000$

$n_2 = 2000$

$\bar{x}_1 = 67.5$

$\bar{x}_2 = 68.0$

$\sigma = 2.5$

i) Null Hypothesis:

Let the two samples drawn from same population

ii) Alternate Hypothesis:

Let two samples not drawn from same population

(Two-tail test)

iii) Level of significance : $\alpha = 5\%$ (Assume)

iv) Test statistic : $$|z| = \frac{|\bar{x}_1 - \bar{x}_2|}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

$$|z| = \frac{|67.5 - 68.0|}{2.5\sqrt{\frac{1}{1000} + \frac{1}{2000}}} = 5.16$$

v) Conclusion : calculated value of $|z|$ (5.16) is more than to the tabulated value (1.96) for two-tail test at 5% level.

$\therefore$ Reject Null Hypothesis (H_0) i.e. the two samples are not drawn from same population.

* Samples of students were drawn from 2 universities & their weights in kg, mean & S.D are calculated & shown below. Make a large sample test to test the significance of the difference b/w the means.

	mean	S.D	size of sample
U-A	55	10	400
U-B	57	15	100

Given that $n = 400$ $\quad \bar{x}_1 = 55$ $\quad S_1 = 10$

$n_2 = 100$ $\quad \bar{x}_2 = 57$ $\quad S_2 = 15$

i) Null Hypothesis: There is no difference b/w means

ii) Alternate Hypothesis: There is difference b/w means (2-tail test)

iii) Level of significance: $\alpha = 5\%$ (Assume)

iv) Test statistic:

$$|Z| = \frac{|\bar{x}_1 - \bar{x}_2|}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

$$|Z| = \frac{|55-57|}{\sqrt{\frac{100}{400} + \frac{225}{100}}} = 1.26$$

v) Conclusion: calculated value of $|Z|$ (1.26) is less than the tabulated value (1.96) for two-tail test at 5% level

$\therefore$ Accept Null-Hypothesis i.e. ~~both the~~ there is no difference b/w the means.

$\star$ The ~~Licoting~~ Nicotine content (in mg) of two samples of tobacco were found to be as follows. Test whether there is any difference b/w 2 means at 0.05 level

Sample A	24	27	26	23	25	-
Sample B	29	30	30	31	24	36

and construct the confidence interval. Given that

$n_1 = 5$

$n_2 = 6$

$$\bar{x}_1 = \frac{24+27+26+23+25}{5} = 25$$

$$\bar{x}_2 = \frac{29+30+30+31+24+36}{6} = 30$$

$$S_1^2 = \sum_{i=1}^{n} \frac{(x_i - \bar{x}_1)^2}{n_1 - 1} = \frac{1+4+1+4+0}{4} = \frac{10}{4} = 2.5$$

$$S_2^2 = \sum_{i=1}^{n} \frac{x_i - \bar{x}_2}{n_2 - 1} = 14.8$$

i) Null Hypothesis: There is no difference b/w means

ii) Alternate Hypothesis: There is difference b/w means (2-tail test)

iii) Level of Significance $\alpha = 5\%$

iv) Test statistic : $$|Z| = \frac{|\bar{x}_1 - \bar{x}_2|}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}}$$

$$|Z| = \frac{|25-30|}{\sqrt{\frac{2.5}{5} + \frac{14.8}{6}}} = 2.90$$

v) Conclusion: calculated value of $|Z|$ (2.90) is more than the tabulated value (1.96) for two-tail test at 5% level

$\therefore$ Reject Null Hypothesis i.e. there is difference b/w means.

12/7/22 **Method 3: Test of Significance for proportions of large samples**

1. we apply this method to check whether the population has a specified value $P = P_0$ or not

2. We use the test statistic $|Z| = \frac{|p - P|}{\sqrt{PQ/n}}$ where $Q = 1 - P$

* In a sample of 1000 people in Karnataka, 540 are rice eaters & the rest are wheat eaters. Can we assume that both rice & wheat eaters are equally popular in this state at 1% level

Given that $n = 1000$

$x = 540$

$$\Rightarrow p = \frac{x}{n} = \frac{540}{1000} = 0.54$$

population proportion of rice eaters $P = \frac{1}{2}$

$$\Rightarrow Q = 1 - P = \frac{1}{2}$$

i) Null Hypothesis (H_0): Let Rice eaters and wheat eaters are equally popular.

ii) Alternate Hypothesis (H_1): Let rice and wheat eaters are not equally popular (2-tail test)

iii) LOS: $\alpha = 0.01$

iv) Test statistic : $|Z| = \frac{|p-P|}{\sqrt{PQ/n}}$

$$|Z| = \frac{0.54-0.5}{\sqrt{\frac{0.5\times0.5}{1000}}} = 2.52$$

v) Conclusion : The calculated value of $|Z|$ (2.52) is less than the tabulated value (2.58) for 2-tail test at 1% level.

∴ Accept Null Hypothesis i.e. both rice & wheat eaters are equally popular.

* In a big city 325 men out of 600 were found to be smokers. Does this info support the conclusion that majority of men in this city are smokers.

$x = 325$, $n = 600$

Given that $p = \frac{325}{600} = 0.54$

sample proportion

Population proportion of smokers $P = \frac{1}{2}$

$\Rightarrow Q = \frac{1}{2}$

i) Null Hypothesis : let Both smokers & non-smokers are equal

ii) Alternate Hypothesis : Majority of the men in city are smokers (one-tail test).

iii) LOS: $\alpha = 5\%$ (assumed)

iv) Test statistic: $|Z| = \frac{|p-P|}{\sqrt{PQ/n}}$

$$|Z| = \frac{|0.54 - 0.5|}{\sqrt{\frac{0.5 \times 0.5}{600}}} = 1.95$$

v) Conclusion: The calculated value of $|Z|$ (1.95) is greater than tabulated value (1.645) for one-tail test at 5% level

$\therefore$ Reject Null Hypothesis i.e. majority of men in city are smokers

★ A manufacturer claimed that atleast 95% of the equipment which is supplied to a factory are confirmed to specifications. An examination of a sample of 200 pieces of equipment revealed that 18 were fault. Test is claimed at 5% level.

Given that $P = 0.95 \Rightarrow Q = 1 - P = 0.05$

sample proportion of good items $= p = \frac{182}{200} = 0.91$

i) **Null Hypothesis**: Let 95% of the items are confirmed for specification

ii) **Alternate Hypothesis**: $P < 0.95$ (one-tail test)

iii) **LOS**: $\alpha = 5\%$

iv) **Test statistic**: $|Z| = \frac{|p - P|}{\sqrt{PQ/n}}$

$$|Z| = \frac{0.91 - 0.95}{\sqrt{\frac{0.95 \times 0.05}{200}}} = 2.59$$

v) **Conclusion**: The calculated value of $|Z|$ (2.59) is greater than tabulated value (1.645) for one-tail test at 5% level

$\therefore$ Reject Null Hypothesis i.e. $P < 0.95$

13/7/22 **Method 4**: Test of Significance for two proportions of large samples

1. we apply this method to check whether the two samples drawn from same population or not.

2. we apply this method to check there is any diff b/w 2 proportions P_1, P_2

3. we use the test statistic : $|Z| = \dfrac{|P_1 - P_2|}{\sqrt{PQ\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}}$

where P_1 = First sample proportion $= \dfrac{x_1}{n_1}$

P_2 = ~~Second sample~~ common proportion $= \dfrac{x_2}{n_2}$

P = common proportion $= \dfrac{n_1 P_1 + n_2 P_2}{n_1 + n_2}$

$Q = 1 - P$

* Random samples of 400 men & 600 women were asked whether they would like to have a fly over near their residence. 200 men & 325 women were in favor of proposal. Test the hypothesis that proportions of men and women in favor of the proposal are same at 1% level.

Given that $n_1 = 400$

$n_2 = 600$

$x_1 = 200$

$x_2 = 325$

$\Rightarrow P_1 = \dfrac{x_1}{n_1} = \dfrac{200}{400} = 0.5$

$P_2 = \dfrac{x_2}{n_2} = \dfrac{325}{600} = 0.54$

i) Null Hypothesis : Let $P_1 = P_2$

ii) Alternate Hypothesis : Let $P_1 \neq P_2$ (Two tail test)

iii) LOS : $\alpha = 1\%$

iv) Test statistic : $$|z| = \frac{|P_1 - P_2|}{\sqrt{PQ\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

Here, common proportion, $P = \frac{n_1 P_1 + n_2 P_2}{n_1 + n_2}$

$$P = \frac{(400)(0.5) + (600)(0.54)}{400 + 600} = 0.524 \Rightarrow Q = 0.476$$

$$\Rightarrow |z| = \frac{|0.5 - 0.54|}{\sqrt{(0.524)(0.476)\left(\frac{1}{400} + \frac{1}{600}\right)}} = 1.24$$

v) Conclusion : The calculated value of $|z|$ (1.24) is less than the tabulated value (2.58) for two tail test at 1% level.

$\therefore$ Accept Null Hypothesis i.e. $P_1 = P_2$

[* A cigarette manufacturing company claims that this brand 'A' cigarette outsells this brand 'B'] ×

* In two large populations there are 30% & 25% resp. of faired haired people. Is this difference likely

to be hidden in samples of 1200 & 900 resp. for the two populations

Given that $P_1 = 30\% = 0.3$

$P_2 = 25\% = 0.25$

$n_1 = 1200$

$n_2 = 900$

i) Null Hypothesis: Let $P_1 = P_2$

ii) Alternate Hypothesis: Let $P_1 \neq P_2$ (2-tail test)

iii) LOS: $\alpha = 5\%$ (assumed)

iv) Test statistic: $|Z| = \dfrac{|P_1 - P_2|}{\sqrt{PQ\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$

Here $P = \dfrac{n_1 P_1 + n_2 P_2}{n_1 + n_2} = \dfrac{(1200)(0.3) + (900)(0.25)}{900 + 1200}$

$\Rightarrow Q = 0.73$

$\Rightarrow |Z| = \dfrac{|0.3 - 0.25|}{\sqrt{(0.27)(0.73)\left(\frac{1}{1200} + \frac{1}{900}\right)}} = 2.55$

v) Conclusion: The tabulated value of $|Z|$ (2.55) is greater than the calculated value (1.96) for two tail test at 5% level

$\therefore$ Reject Null hypothesis i.e. $P_1 \neq P_2$

4/7/22 Small Sample Test

If sample size n is less than or equal to 30, that sample is known as a small sample.

Following are the some steps for small sample:

i) t-distribution (or) students t-test

ii) F-distribution

iii) χ^2 (chi square) distribution)

t-distribution (or) students t-test

If $x_1, x_2, x_3 \dots x_n$ be any random sample which is taken from a population with mean μ & SD σ then the ~~std~~ t-distribution is defined as

$$t = \frac{\bar{x}-\mu}{\frac{s}{\sqrt{n}}} \quad (\text{or}) \quad t = \frac{\bar{x}-\mu}{\frac{s}{\sqrt{n-1}}}$$

with degree of freedom $\mu = n-1$.

Degree of freedom

Degree of freedom is a no. which indicate how many of the values of a variable may be freely chosen.

Properties of t-distribution

1) The shape of t-distribution is bell shaped.

which is similar to the normal distribution

ii) The t- distribution curve is symetrical about the line $t=0$

iii) The form of the probability curve varing with degree of freedom.

iv) The mean, median, mode of t-distribution are same.

Applications of t-distributions

1) we apply t-distribution to test the significance of sample mean $\bar{x}$ when population variance ~~mean~~ is not given

2) we apply t-distribution to check the significance difference b/w sample mean $\bar{x}$ & population mean μ

3) we apply t-distribution to check the significance difference b/w 2 sample means

15/7/22

Method 1: Test of significance for single mean of small sample

1. we use this method to check whether the sample is drawn from the given population or not.

2. we apply this method to check the significant difference between sample mean $\bar{x}$, population mean 'μ'

we use the test statistic $t = \frac{\bar{x} - \mu}{\frac{s}{\sqrt{n}}}$ where s is sample S.D. $\sqrt{\frac{\Sigma (x_i - \bar{x})^2}{n-1}}$

$\therefore$ Degree of freedom $\gamma = n-1$

The average ~~basic sense~~ breaking strength of steel rods is specified to be 18.5 thousand pounds. To test the sample of 14 rods were tested. The mean and S.D obtained were 17.85 and 1.955 resp. Is the result of exp significant?

Given that $n = 14$

$\bar{x} = 17.85$

$s = 1.955$

$\mu = 18.5$

$\gamma = n-1 = 14-1 = 13$

i) Null Hypothesis: Let the result of exp is significant

ii) Alternate Hypothesis: Let the result of exp is not significant. (Two-tail test)

iii) LOS: $\alpha = 5\%$ (assumed)

iv) Test statistic: $t = \frac{\bar{x} - \mu}{\frac{s}{\sqrt{n}}} = \frac{17.85 - 18.5}{\frac{1.955}{\sqrt{14}}} = -1.24$

v) Conclusion: Here, calculated value of $|t| = 1.27$ is less than tabulated value (2.160) for two tail test of at 5% level.

$\therefore$ Accept Null Hypothesis i.e. the result of exp significant.

* A random sample of 6 steel beams has a mean compressive strength of 58,392 pounds per square inch with S.D of 648 P.S.i. Use this info of the level of significance $\alpha = 0.05$ to test whether the avg compressive strength of steel rod is 58,000 C.S.I.

Given that $n = 6$

$\bar{x} = 58,392$

$S = 648$

$\mu = 58,000$

$\nu = n-1 = 6-1 = 5$

i) Null Hypothesis: Let $\mu = 58,000$

ii) Alternate Hypothesis: Let $\mu \neq 58,000$ (2-tail test)

iii) LOS: $\alpha = 5\%$

iv) Test statistic: $t = \dfrac{\bar{x}-\mu}{\frac{S}{\sqrt{n}}} = \dfrac{58392 - 58000}{\frac{648}{\sqrt{6}}} = 1.48$

v) Conclusion: Here, calculated value of $|t| = 1.48$ is less than tabulated value (2.571) for two-tail test at 5% level

$\therefore$ Accept Null Hypothesis i.e. $\mu = 58,000$

Note:- If the sample mean $\bar{x}$, sample S.D 's' is not given directly we have to find by using

$$\bar{x} = \frac{x_1 + x_2 + x_3 + \cdots + x_n}{n}$$

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

* A random sample of 10 boys has the following I.Qs : 70, 120, 110, 101, 88, 83, 95, 98, 107 and 100

i) Do this data support the assumption of a population mean IQ of 100?

ii) Construct the confidence interval at 5% level.

Given that $n = 10$ $\qquad \mu = 100$

$$\gamma = n - 1 = 10 - 1 = 9$$

$$\bar{x} = \frac{70 + 120 + 110 + \cdots 100}{10} = 97.2$$

$$S = \sqrt{\frac{\Sigma (x_i - \bar{x})^2}{n-1}} = \sqrt{\frac{(70-97.2)^2 + \cdots (100-97.2)^2}{9}} = 14.27$$

i) Null Hypothesis : $\mu = 100$

ii) Alternate Hypothesis : $\mu \neq 100$ (2-tail test)

iii) LOS : $\alpha = 5\%$

iv) Test statistic : $t = \dfrac{\bar{x} - \mu}{\frac{s}{\sqrt{n}}} = \dfrac{97.2 - 100}{\frac{14.27}{\sqrt{10}}}$

$t = -0.62$

v) Conclusion : Here, calculated value of $|t| = 0.62$ is less than tabulated value (2.262) for two tail test at 5% level

$\therefore$ Accept Null Hypothesis i.e. $\mu = 100$.

(ii) confidence interval $\left(\bar{x} \pm t_{\alpha/2} \frac{s}{\sqrt{n}}\right)$

$$97.2 \pm 2.262\left(\frac{14.27}{\sqrt{10}}\right)$$

$$= (86.9,\ 107.4)$$

18/7/22 Method 2 : Test of significance for difference of two means of small samples

1. we apply this method to check whether the two samples drawn from same population or not

2. We apply this method to check the significant diff b/w 2 means.

3. We use the test statistic $|t| = \dfrac{|\bar{x}-\bar{y}|}{\sqrt{S^2\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}}$

Here, the common variance $S^2 = \dfrac{n_1 s_1^2 + n_2 s_2^2}{n_1+n_2-2}$

and $S_1^2 = \dfrac{\Sigma(x_i-\bar{x})^2}{n_1-1}$ $\quad S_2^2 = \dfrac{\Sigma(x_i-\bar{x})^2}{n_2-1}$

4. The degree of freedom in this method is

$\nu = n_1+n_2-2$

* Samples of 2 types of electric light bulbs were tested for length of life & following data was obtained: ~~& following do~~

Type - I	Type - II
$n_1 = 8$	$n_2 = 7$
mean = 1234 Hrs	mean = 1036 Hrs
SD = 36 Hrs	SD = 40 Hrs

Is the difference in the mean significant to warrant that type-I is superior to type-II regarding length of life.?

Given that $n_1 = 8$

$n_2 = 7$

$\bar{x} = 1234$ Hrs

$\bar{y} = 1036$ Hrs

$s_1 = 36$ Hrs

$s_2 = 40$ Hrs

degree of freedom

$\gamma = n_1 + n_2 - 2$

$= 8 + 7 - 2$

$= 13$

i) Null Hypothesis (H_0): Let there is no diff in mean life of type-I and type-II

ii) Alternate Hypothesis (H_1): mean life of type-I is superior to type-II (one-tail test)

iii) LOS: let $\alpha = 5\%$ (assumed)

iv) Test statistic:

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

Here common variance $s^2 = \frac{n_1 s_1^2 + n_2 s_2^2}{n_1 + n_2 - 2} = 1659.07 \simeq 1659$

$$\therefore t = \frac{|1234 - 1036|}{\sqrt{(1659)\left(\frac{1}{8}+\frac{1}{7}\right)}} = 9.39$$

v) Conclusion: Here, calculated value of t (9.39) is greater than tabulated value (1.771) for one-tail test at 5% level and 13 degree of freedom.

$\therefore$ Reject Null Hypothesis i.e. mean life of type-I is superior to type-II

* Two horses A and B were tested according to the time (in sec) to run a particular track with the following result:

H-A	28	30	32	33	33	29	34
H-B	29	30	30	24	27	29	-

Test whether the 2 horses have the same running capacity

Given that $n_1 = 7$

$n_2 = 6$

$$\bar{x} = \frac{28+30+\cdots+34}{7} = \cancel{27.4}\ 31.28$$

$$\bar{y} = \frac{29+30+\cdots+29}{6} = 28.16$$

$$s_1 = \sqrt{\frac{\Sigma(x_i-\bar{x})^2}{n_1-1}} = \sqrt{\frac{(28-31.28)^2+(30-31.28)^2+\cdots+(34-31.28)^2}{7-1}}$$

$$= 2.28 \Rightarrow s_1^2 = 5.19$$

$$s_2 = \sqrt{\frac{\Sigma(y_i-\bar{y})^2}{n_2-1}} = \sqrt{\frac{(29-28.16)^2+\cdots+(29-28.16)^2}{6-1}}$$

$$= 2.31 \Rightarrow s_2^2 = 5.33$$

Here, common variance

$$s^2 = \frac{n_1 s_1^2 + n_2 s_2^2}{n_1+n_2-2} = \cancel{24.68}\ 6.21$$

degree of freedom, $\gamma = n_1+n_2-2 = 7+6-2 = 11$

i) Null hypothesis: let The 2 horses have same running capacity

ii) Alternate hypothesis: let The 2 horses doesn't have same running capacity (2-tail test)

iii) LOS: $\alpha = 5\%$ (assumed)

iv) Test statistic:

$$t = \frac{|\bar{x}-\bar{y}|}{\sqrt{s^2\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}} = \frac{|31.78 - 28.16|}{\sqrt{(\cancel{37.63}\ 6.21)\left(\frac{1}{7}+\frac{1}{6}\right)}} = \cancel{2.99}\ 2.29$$

v) Conclusion: Here, calculated value (2.29) is ~~greater~~ lesser than tabulated value (2.201) for 2-tail test at 5% level and 11 degree of freedom

$\therefore$ ~~Accept~~ Reject Null Hypothesis i.e. the 2 horses does not have same running capacity.

19/7/23 **Paired Sample t-test**

→ we apply this method whenever we have paired observations i.e. before & after giving a drug to the patients, before & after giving an intensive training to the students, etc...

→ we use the test statistic $t = \frac{\bar{d}-\mu}{\frac{s}{\sqrt{n}}}$ (or)

$t = \frac{\bar{d}}{\frac{s}{\sqrt{n}}}$ where

$\bar{d}$ = mean of the differences

s = S.D of the differences

n = sample size

→ Degree of freedom, $\nu = n-1$

* The blood pressure of 5 women, before & after intake of a certain drug are given below:

Before	110	120	123	132	125
After	120	118	125	136	121

Test whether there is any significant difference in BP at 1% level.

Given that $n = 5$

$d = -10, 2, -2, -4, 4$

$$\Rightarrow \bar{d} = \frac{\Sigma d}{n} = \frac{-10+2+\cdots+4}{5} = \cancel{-10}\ -2$$

$$\Rightarrow S = \sqrt{\frac{\Sigma(d_i-\bar{d})^2}{n-1}} = 5.47$$

Here, $\nu = n-1 = 5-1 = 4$

i) Null hypothesis: Let there is no difference in BP

ii) Alternate hypothesis: Let there is difference in B.P (2-tail test)

iii) LOS: $\alpha = 1\%$

iv) Test statistic : $|t| = \frac{|\bar{d}|}{\frac{s}{\sqrt{n}}} = \frac{1.9}{\frac{5.41}{\sqrt{5}}} = 0.81$

v) Conclusion : Here, the calculated value of $|t|$ (0.81) is less than the tabulated value (4.604) for 2-tail test at 1% level and 4 degree of freedom.

$\therefore$ Accept Null Hypothesis (H_0) i.e. there is no significant diff in B.P before & after intake of drug.

* Memory capacity of 10 students were tested before & after training. State whether the training program was effective or not from the following scores.

Before Training	12	14	11	8	7	10	3	0	5	6
After Training	15	16	10	7	5	12	10	2	3	8

Given that $n = 10$

$\nu = n-1 = 10-1 = 9$

$d = -3, -2, 1, 1, 2, -2, -7, -2, 2, -2$

$$\bar{d} = \frac{-3-2+\cdots-2}{10} = -1.2$$

$$S = \sqrt{\frac{\Sigma(d_i - \bar{d})^2}{n-1}} = \sqrt{\frac{(3+1.2)^2 + \cdots (-2+1.2)^2}{10-1}} = 2.78$$

i) Null Hypothesis: Let the training program was effective

ii) Alternate Hypothesis: Let training program was not effective. (2-tail test)

iii) LOS: Let $\alpha = 5\%$ (assumed)

iv) Test statistic:
$$|t| = \frac{|\bar{d}|}{\frac{S}{\sqrt{n}}} = \frac{|-1.2|}{\frac{2.78}{\sqrt{10}}} = 1.365$$

v) Conclusion: Here, calculated value of $|t|$ (1.365) is less than the tabulated value (2.262) for 2-tail test at 5% level & 9 degree of freedom.

$\therefore$ Accept Null Hypothesis i.e. the training program was effective.

20/7/23 F-Distribution

1) We apply F-Distribution to check the significance difference b/w 2 variances

2) We use the test statistic $F = \frac{S_1^2}{S_2^2}$ if $S_1^2 > S_2^2$

where $S_1^2 = \sum_{i=1}^{n} \frac{(x_i - \bar{x})^2}{n_1 - 1}$ (or) $F = \frac{S_2^2}{S_1^2}$ if $S_2^2 > S_1^2$

$S_2^2 = \sum_{i=1}^{n} \frac{(y_i - \bar{y})^2}{n_2 - 1}$

3) We use the degree of freedom $(\nu_1, \nu_2) = (n_1 - 1, n_2 - 1)$

properties of F-Distribution

i) F-Distribution is free from population parameters & depends upon degree of freedom

ii) F-Distribution curve lies entirely in I^{st} quadrant

iii) The mode of F-Distribution is less than unity.

* In one sample of 8 observations from a normal population the ~~sample~~ sum of the squares of the deviations of the sample values from the sample mean is 84.4 and in another sample of 10 observations it was 102.6. Test at 5% level whether the 2 samples have the same variances.

Given that $n_1 = 8$

$n_2 = 10$

$\Sigma(x_i - \bar{x})^2 = 84.4$

$\Sigma(y_i - \bar{y})^2 = 102.6$

w.k.t $s_1^2 = \frac{\Sigma(x_i - \bar{x})^2}{n_1 - 1} \Rightarrow \frac{84.4}{7} \Rightarrow s_1^2 = 12.05$

$s_2^2 = \frac{\Sigma(y_i - \bar{y})^2}{n_2 - 1} \Rightarrow s_2^2 = \frac{102.6}{9} = 11.4$

here $s_1^2 > s_2^2$

i) Null Hypothesis : Let both the variances are same ($s_1^2 = s_2^2$).

ii) Alternate Hypothesis : Let both the variances are not same ($s_1^2 \neq s_2^2$)

iii) LOS : $\alpha = 5\%$

iv) Test statistic : $F = \frac{s_1^2}{s_2^2} = \frac{12.05}{11.4} = 1.05$

v) Conclusion : Here, calculated value of F (1.05) is less than tabulated value (3.29)

$\therefore$ Accept Null Hypothesis i.e. both the variances are

The Nicotin contains in mg in 2 samples of tobacco were found to be as follows

Sample - A	24	27	26	21	25	-
Sample - B	27	30	28	31	22	36

check whether there is any difference in the variances of 2 samples.

F - Distribution

Given that $n_1 = 5$
$n_2 = 6$

Here $\bar{x} = \frac{24+27+\cdots+25}{5} = 24.6$

$\bar{y} = \frac{27+30+\cdots+36}{6} = 29$

$\Rightarrow s_1^2 = \frac{\Sigma(x_i-\bar{x})^2}{n_1-1} = \frac{(24-24.6)^2+\cdots(25-24.6)^2}{4} = 5.3$

$s_2^2 = \frac{\Sigma(y_i-\bar{y})^2}{n_2-1} = 21.6$

w.k.t $F = \frac{s_2^2}{s_1^2} = \frac{21.6}{5.3} = 4.07$

The calculated value of F (4.07) is less than to the tabulated value (6.26) for 2-tail test at 5% and (5, 4) degree of freedom.

$\therefore$ Accept Null Hypothesis

t - Distribution

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

Here $S^2 = \dfrac{n_1 s_1^2 + n_2 s_2^2}{n_1+n_2-2} = 17.34$ (common variance)

$\therefore |t| = \dfrac{|24.6 - 29|}{\sqrt{(17.34)\left(\frac{1}{5}+\frac{1}{6}\right)}} = 1.74$

Here, the calculated value of |t| (1.74) is lesser than the tabulated value (2.262) for 2-tail test at 5% level & 9 degree of freedom.

$\therefore$ Accept Null Hypothesis i.e. there is no difference in variance of 2 samples.

21/7/23 χ^2 Distribution (Chi-Square Distribution)

i) we apply χ^2 distribution to check the significant difference b/w 2 frequencies (observed, expected)

ii) we use the test statistic $\chi^2 = \sum_{i=1}^{n} \frac{(O_i - E_i)^2}{E_i}$

iii) Degree of freedom $\nu = n-1$

properties of χ^2 Distribution

i) χ^2 distribution curve is not symmetric

ii) χ^2 curve lies entirely in the 1st quadrant & it varies from 0 to ∞.

iii) The value of χ^2 depends only on degree of freedom

iv) The mean of χ^2 distribution is 'ν' & variance is '2ν'

☆ The no of automobile accidents per week in a certain community are as follows: 12, 8, 20, 2, 14, 10, 15, 6, 9, 4. Are these frequencies in argument with the belief that accident conditions were the same during this 10 week period.

Given that $n = 10$

observed frequencies = 12, 8, 20, 2, 14, 10, 15, 6, 9, 4

wkt Expected frequencies $E = \frac{12 + 8 + \cdots + 9 + 4}{10}$

$E = 10$

Degree of freedom. $\gamma = n - 1 = 10 - 1 = 9$

i) Null Hypothesis: Let the accident conditions were same during 10 week period

ii) Alternate Hypothesis: Let the accident conditions were not same during 10 week period

iii) LOS: Let $\alpha = 5\%$ (assumed)

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

O_i	E_i	$O_i - E_i$	$(O_i - E_i)^2$	$\frac{(O_i - E_i)^2}{E_i}$
12	10	2	4	0.4
8	10	-2	4	0.4
20	10	10	100	10
2	10	-8	64	6.4
14	10	4	16	1.6
10	10	0	0	0
15	10	5	25	2.5
6	10	-4	16	1.6
9	10	-1	1	0.1
4	10	[illegible]	[illegible]	[illegible]

v) Conclusion: The calculated value of χ^2 (26.6) is greater than tabulated value (16.919) for 2-tail test at 5% with 9 degree of freedom.

∴ Reject null hypothesis i.e. the accidents conditions were not same during 10 week period.

* A die is thrown with the following results show that the die is biased.

No. appeared on the die	1	2	3	4	5	6
frequency	40	32	28	58	54	52

Given that $n = 6$

observed frequencies - 40, 32, 28, 58, 54, 52

wkt Expected frequencies $E = \frac{40 + \cdots + 52}{6}$

$E = 44$

Degree of freedom, $\gamma = n-1 = 6-1 = 5$

i) Null hypothesis: Let the die is unbiased

ii) Alternate hypothesis: Let the die is biased

iii) LOS: Let $\alpha = 5\%$ (assumed)

iv) Test statistic: $\chi^2 = \sum_{i=1}^{n} \frac{(O_i - E_i)^2}{E_i} = 17.63$

v) Conclusion: Here, the calculated value of χ^2 (17.63) is greater than tabulated value (11.070) for 2-tail test at 5% with 5 degree of freedom.

$\therefore$ Reject Null Hypothesis i.e. the die is biased.

* A survey of 240 families with 4 children each reveal the following distribution:

Male Births	4	3	2	1	0
observed frequency	10	55	105	58	12

Check whether the frequency of male & female births are equal or not.

Given that $n = 5$ (sample size)

$$\gamma = 5 - 1 = 4$$

Observed frequencies: 10, 55, 105, 58, 12

The Expected frequencies are the successive terms of binomial frequency distribution:

$$N(p+q)^n$$

here $N = 240$, $n = 4$

Let p be the probability of getting a male birth

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

The expected frequency $E_i = 240\left(\frac{1}{2} + \frac{1}{2}\right)^4$

$$= 240\left[{}^4C_4\left(\tfrac{1}{2}\right)^4\left(\tfrac{1}{2}\right)^0 + {}^4C_3\left(\tfrac{1}{2}\right)^3\left(\tfrac{1}{2}\right)^1 + {}^4C_2\left(\tfrac{1}{2}\right)^2\left(\tfrac{1}{2}\right)^2 + {}^4C_1\left(\tfrac{1}{2}\right)^1\left(\tfrac{1}{2}\right)^3 + {}^4C_0\left(\tfrac{1}{2}\right)^0\left(\tfrac{1}{2}\right)^4\right]$$

$$= 240\left(\tfrac{1}{2}\right)^4\left[1 + 4 + 6 + 4 + 1\right]$$

$= 15(16)$ i.e. the expected frequency of male births are

$= 240$

15, 60, 90, 60, 15

i) Null Hypothesis: Let the frequency of male & female births are equal

ii) Alternate Hypothesis: Let the frequency of male & female births are not equal.

iii) LOS: Let $\alpha = 5\%$ (assumed)

iv) Test Statistic: $\chi^2 = \sum_{i=1}^{n} \frac{(O_i - E_i)^2}{E_i} = 9.23$

v) Conclusion: Here, the calculated value (5.23) is less than tabulated value (9.488) for 2-tail test at 5% with 4 degree of freedom.

∴ Accept Null Hypothesis i.e. the frequency of male & female births are equal.

28/7/23 χ^2 - Test for independent of Attributes

Literally an attribute means quality or characteristic. Examples of Attributes are drinking, honesty, smoking etc.

If the observed frequencies are given in matrix form

	I	II	
I	a	b	(a+b)
II	c	d	(c+d)
	(a+c)	(b+d)	$\boxed{G.T.}$

we obtain expected frequency of each observed frequency as

$$E(a) = \frac{\text{Row sum} \times \text{Column sum}}{\text{Grand Total}} \Rightarrow \frac{(a+b)\times(a+c)}{GT}$$

lly $E(b) = \frac{(a+b)\times(b+d)}{GT}$; $E(d) = \frac{(c+d)\times(b+d)}{GT}$

$E(c) = \frac{(c+d)\times(a+c)}{GT}$

→ we take degree of freedom in this method as (No. of rows − 1) × (No. of columns − 1)

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

~~Soft Drinks~~ Employees

Soft Drinks	clerks	Teachers	officers	
Pepsi	10	25	65	100
Thumps up	15	30	65	110
Fanta	50	60	30	140
	75	115	160	350

i) Null Hypothesis : Let there is no significant difference in taking of soft drinks.

ii) Alternate Hypothesis: Let there is a significant difference in taking of soft drinks (2-tail test)

iii) LOS: Let $\alpha = 5\%$ (assumed)

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

Degree of freedom is $(3-1)\times(3-1) = 4$

The expected frequency of each observed frequency is

$E(10) = \frac{100 \times 75}{350} = 21.4$

$E(25) = \frac{100 \times 115}{350} = 32.8$

$E(65) = \frac{100 \times 160}{350} = 45.7$

$E(15) = \frac{110 \times 75}{350} = 23.6$

$E(30) = \frac{110 \times 115}{350} = 36.1$

$E(65) = \frac{110 \times 160}{350} = 50.3$

$$E(50) = \frac{140 \times 75}{350} = 30$$

$$E(60) = \frac{140 \times 115}{350} = 46$$

$$E(30) = \frac{140 \times 160}{350} = 64$$

O_i	E_i	$O_i - E_i$	$(O_i - E_i)^2$	$(O_i - E_i)^2/E_i$
10	21.4	-11.4	129.96	6.073
25	32.8	-7.8	67.41	1.899
65	45.7	19.3	372.49	8.151
15	23.6	-8.6	73.96	3.134
30	36.1	-6.1	37.21	1.031
65	50.3	14.7	216.09	4.3
50	30	20	400	13.333
60	46	14	196	4.261
30	64	-34	1156	18.062
				60.2425

v) Conclusion: Here, calculated value of χ^2 (60.2425) is greater than to the tabulated value (9.488) for 2-tail test at 5% level with 4 degree of freedom.

∴ Reject Null Hypothesis i.e. there is difference in taking of soft drinks.

* ~~In~~ 1000 students in college level were graded acc to their IQ & economic conditions of their home. Use χ^2 test to find out whether there is any association b/w condition at home and IQ

Economic Condition	IQ High	Low	
Rich	460	140	600
Poor	240	160	400
	700	300	$\boxed{1000}$

i) Null Hypothesis: Let there is no association b/w condition at home & IQ

ii) Alternate Hypothesis: Let there is association b/w condition at home & IQ (2-tail test)

iii) LOS: Let $\alpha = 5\%$ (assumed)

iv) Test statistic :

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

Degree of freedom is $(2-1) \times (2-1) = 1$

The expected frequency of each observed frequency is

$$E(460) = \frac{600 \times 700}{1000} = 420$$

$$E(140) = \frac{600 \times 300}{1000} = 180$$

$$E(240) = \frac{400 \times 700}{1000} = 280$$

$$E(160) = \frac{400 \times 300}{1000} = 120$$

O_i	E_i	$O_i - E_i$	$(O_i - E_i)^2$	$(O_i - E_i)^2/E_i$
460	420	40	1600	3.80
140	180	-40	1600	8.89
240	280	-40	1600	5.71
160	120	40	1600	13.34
				31.74

v) Conclusion : Here, the calculated value (31.74) is greater than the tabulated value (3.841) at 5% level & '1' degree of freedom.

$\therefore$ Reject Null Hypothesis i.e there is association b/w condition at home & IQ.