Test of Hypothesis.

parameters → population - Mean (μ) → variance (σ^2) → deviation (σ)

statistics → sample ($\bar{x}$) → Mean and variance (S^2) → deviation (S).

Population:-

The collection of objects about a particular characteristic which we are going to study is called population.

Ex:- 1. All Engineering colleges in Andhra pradesh.

2. Number of vehicles in Tirupati.

population Size:-

Number of objects in the given population is defined as population size. It is denoted with 'N'. Depend upon population size. There are two types.

1. Finite population.

2. Infinite population.

1. Finite population:-

A population is said to be finite if the population size is uncountable.

Ex: Number of pens.

2. Infinite population:-

A population is said to be infinite if the population size is uncountable.

Ex: Stars in the sky.

Sample:-

A finite subset of a population is called a sample.

Ex: 1. Number of engineering colleges in Tirupathi

2. Number of two wheelers in Tirupati

Sample Size:-

Number of objects in a given sample is defined as sample size. It is denoted with 'n'. Depend upon sample size, there are two types.

1. Large sample.
2. Small sample.

1. Large Sample:-

If the sample size $n \geq 30$, then the sample is called large sample.

2. Small sample:-

If the sample size $n < 30$, then the sample is called small sample.

Parameter:-

The statistical constants obtained by using population data is called a parameter.

Ex: 1. population mean $= \mu$

2. population variance $= \sigma^2$

3. population standard deviation $= \sigma$.

Statistic:-

The statistical constants obtained by using statistical data is called statistics.

Ex: 1. Sample Mean $= \bar{x}$ (or) μ_0.

2. sample variance $= S^2$

3. deviation $= S$.

Statistical Hypothesis :-

- To Make decisions about a population with the basis of sample information. Such decisions may or may not be true are called a statistical hypothesis.
- It is a simply statement about the statistical hypothesis. They are of two types.

1. Null Hypothesis.
2. Alternative Hypothesis.

1. Null Hypothesis (H_0) :-

A define statement about the population parameter is called the Null hypothesis (H_0). H_0 is a commonly accepted fact.

- A Null hypothesis is the hypothesis which asserts that there is no significant difference between the statistic and the population parameter and whatever observed difference is there merely due to fluctuations in sampling from the same population.

$$H_0 \rightarrow \bar{x} = \mu \text{ (or) } \mu = \bar{x}, \quad \sigma = \sigma_0.$$

$$\left[\begin{array}{l} \mu = \mu_0 \text{ (or)} \\ \sigma = \sigma_0 \text{ (or)} \end{array}\right]$$

2. Alternative Hypothesis (H_1 or H_A) :-

A hypothesis which is complementary to the null hypothesis is called the alternative Hypothesis (H_1)

$$H_1 \rightarrow \bar{x} \neq \mu \text{ (or) } \mu = \mu_0(\bar{x}), \quad \sigma \neq \sigma_0 \text{ [two tailed test]}$$

$\rightarrow$ ~~$\bar{x} > \mu$~~ (or) ~~$\bar{x}$~~

$\rightarrow \mu > \mu_0$ (Right tailed test) (or) [$\bar{x} > \mu$

study something

$\rightarrow \mu < \mu_0$ (Left tailed test) (or).

• Based on the alternative Hypothesis, the test is divided into two types.

i. Two tailed test.

ii. one tailed test.

i) Two tailed test :-

If the alternative Hypothesis is "not equal to" type then the critical region completely lies on both sides of the curve having area '$\alpha/2$' on both sides.

ii) one-tailed test :-

one-tailed test is of two types.

a) Right - tailed test.

b) Left - tailed test.

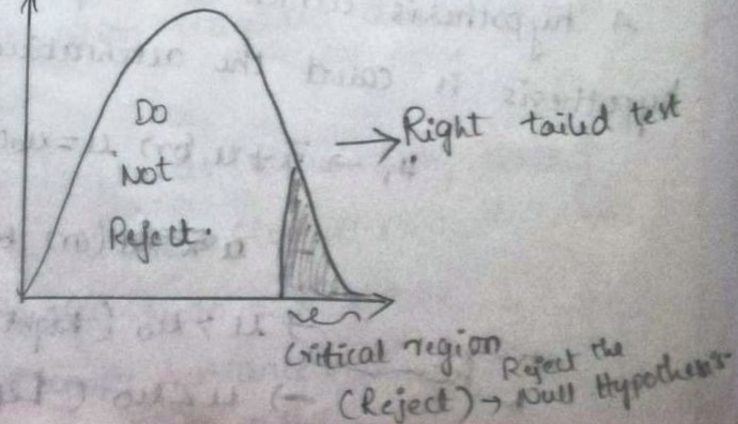
a) Right - Tailed test :-

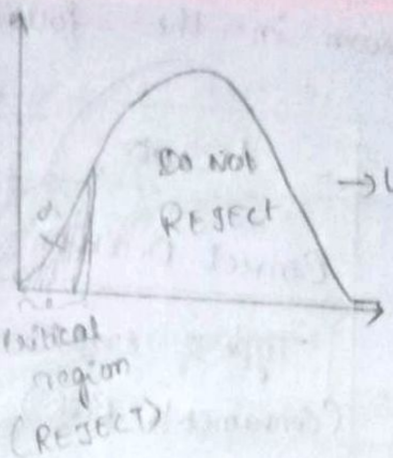
If the alternative Hypothesis is of "greater than" type $(H_1: \mu > \mu_0)$, then the critical region completely lies on the right side of the curve having area "α".

b) Left - Tailed test :-

If the alternative Hypothesis is of "less than" type $(H_1: \mu < \mu_0)$, then the critical region completely lies on the left side of the curve handling area "α".

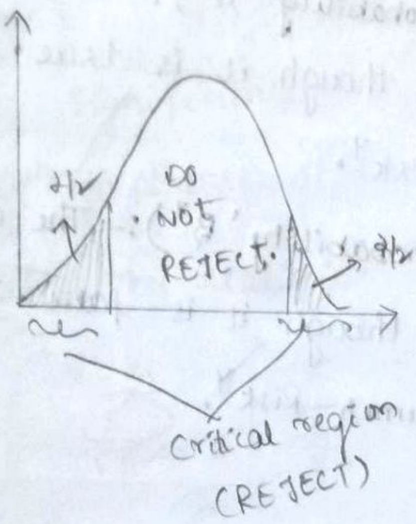
ONE - Sided - Test





→ Left Tailed test.

Two-Sided-Test (Two tailed)



ERRORS IN SAMPLING:-

The main objective of sampling theory is to draw a obtain a valid inference (information) about population when a statistical hypothesis is tested, there are four possibilities.

1. The Hypothesis is true and Test is accepted (correct Decision).

2. The Hypothesis is false and the Test is Rejected (Correct Decision).

3. The Hypothesis is true and the test is Rejected (Type 1 error)

4. The Hypothesis is false and the test is accepted (Type 2 error).

This can be shown in the following table.

Ho	Accept	Reject.
True	Correct Decision	Type 1 Error (producer's Risk)
False	Type 2 error (Consumer's Risk)	Correct Decision.

Note:-

1. Type 1 Error (probability 'α'):- The error involves rejecting Ho even though it is true. It is also known as "producer's Risk".

2. Type 2 Error (probability 'β'):- The error involves accepting Ho even though it is false. The error is also known as "Consumer Risk".

Critical Region:-

The region in which the null hypothesis is rejected is called critical region.

Acceptance Region:-

The region in which the null hypothesis is accepted is called the acceptance region.

LEVEL OF SIGNIFICANCE:-

The maximum probability of Type 1 Error is α.
Here α is called Level of significance.

- It generally expressed in %.
- It measures associated with making decisions.

we take $\alpha = 1\%$ which is used for high accuracy.

$\alpha = 5\%$ which is used for Moderate accuracy.

$\alpha = 10\%$ Can also be used.

Procedure for Testing of Hypothesis

1. Set up Null Hypothesis.

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

2. Set up alternate Hypothesis.

$\mu \neq \bar{x} \rightarrow$ Two tailed test.

$\mu > \bar{x} \rightarrow$ Right tailed

$\mu < \bar{x} \rightarrow$ Left tailed.

3. Fix Level of Significance.

Fix Level of Significance at $\alpha = 1\%$, 5% or 10%.

4. Choose suitable Test statistic.

$$z = \frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}} = \frac{\bar{x} - \mu}{\frac{S}{\sqrt{n}}} \quad d = \frac{\bar{x}}{\frac{S}{\sqrt{n}}} \quad (\text{when } \mu = 0)\ (n \geq 30)$$

For Small samples

$$t = \frac{\bar{x} - \mu}{\frac{S}{\sqrt{n}}} = \frac{\bar{x}}{\frac{S}{\sqrt{n}}}.$$

5. Interpretation.

If $n \geq 30$

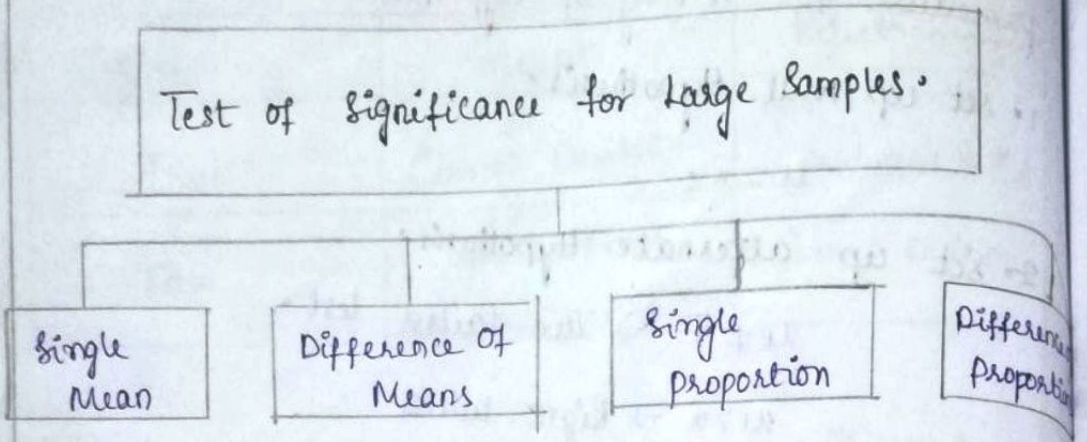
$|z|_{calculated} > |z|_{table} \rightarrow$ Reject Null Hypothesis

$|z|_{calculated} < |z|_{table} \rightarrow$ Accept Null Hypothesis.

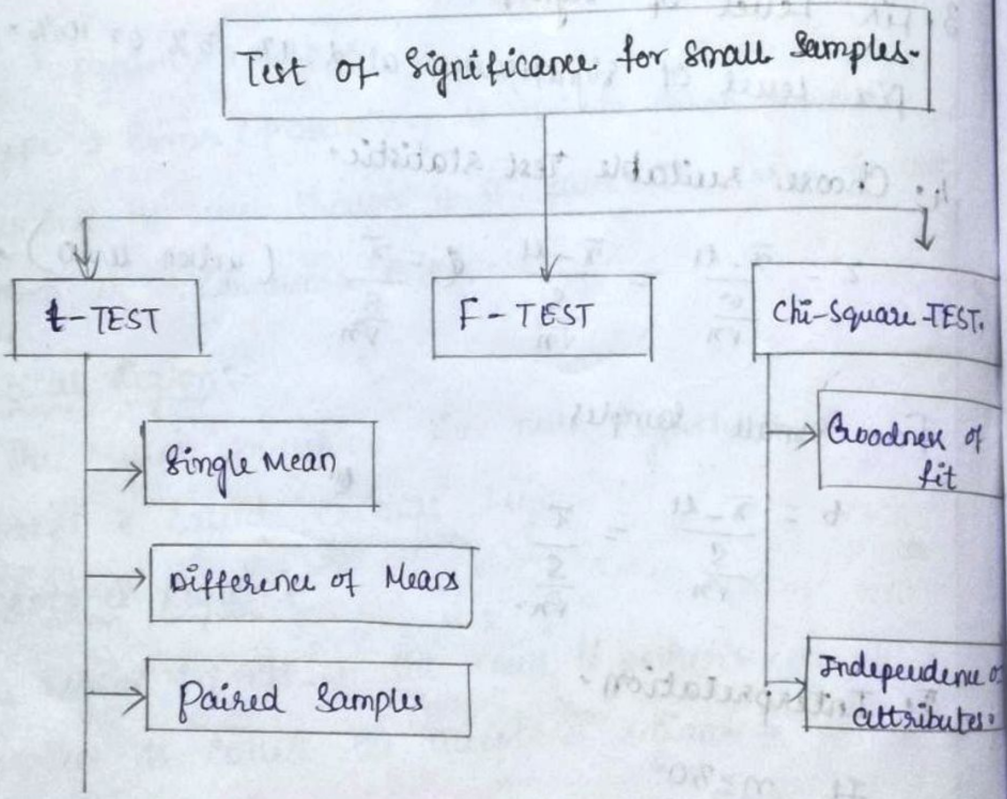
Note:

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

Test of Significance for Large Samples:



Test of Significance for Small Samples:-



Z-test for Single Mean:-

Let $\bar{x}$ be the sample mean of the large sample size 'n' drawn from the normal population having mean μ and standard deviation 'σ'.

Null Hypothesis:-(H_0)

There is no significant difference between the sample Mean and population Mean. $\bar{x}=\mu$ (or) The sample is drawn from normal population.

Alternative Hypothesis:- (H_1 or H_A):-

There is significant difference between the Sample Mean and population Mean.

$\bar{x} \neq \mu$ (Two Tailed)

$\left.\begin{array}{l}\bar{x} < \mu \\ \bar{x} > \mu\end{array}\right\}$ one tailed.

The Sample is not drawn from normal population.

Level of Significance:-

Choose the level of significance based on problem.

Default 5%.

Test statistic:-

SE = Standard Error $z = \frac{|\bar{x}-\mu|}{SE(x)} = \frac{|\bar{x}-\mu|}{\frac{\sigma}{\sqrt{n}}}$, if σ is known

$= \frac{|\bar{x}-\mu|}{\frac{S}{\sqrt{n}}}$, if σ is not known.

Conclusion:-

If $|z|$ calculated $<$ $|z|$ critical value or table value at α% level of significance then we accept H_0, otherwise we reject H_0.

Confidence limits or Fudicial limits:-

$z = \bar{x} \pm z_\alpha SE(\bar{x})$.

$= (\bar{x} - z_\alpha SE(\bar{x}), \bar{x} + z_\alpha SE(\bar{x}))$

	S.No	1%	5%	10%
Two-Tailed	$\bar{x} = \mu$	2.58	1.96	1.645
Right tailed	$\bar{x} > \mu$	2.33	1.64	1.29
Left tailed	$\bar{x} < \mu$	-2.33	-1.64	-1.29.

$\frac{\alpha}{2}$ $\frac{0.99}{2}$ 0.49 0.99

Problems :-

P1. A Sample of 64 students have a mean weight of 70 kgs, can this regarded as a sample from a population with mean weight 56 kgs and standard deviation 25 kgs?

Sol) Given $n = 64$, $\bar{x} = 70$, $\mu = 56$, $\sigma = 25$

i) Null Hypothesis (H_0)

(or).

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

The Sample is drawn from normal population.

ii) Alternate Hypothesis (H_1)

$$H_1 \rightarrow \bar{x} \neq \mu. \quad (or)$$

The Sample is not drawn from normal population.

iii) Level of significance.

Let $\alpha = 5\%$

$= 0.05.$

iv) Test Statistic.

$$z = \frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}} = \frac{70 - 56}{\frac{25}{\sqrt{64}}} = \frac{70 - 56}{\frac{25}{8}} \Rightarrow 70 - 56 \times \frac{8}{25}$$

$$= \frac{14 \times 8}{25} = 4.48.$$

Here $|z|$ calculated $= 4.48$.

$|z|$ table $= 1.96$.

Conclusion.

Here $|z|$ calculated $>$ $|z|$ table value

i.e $4.48 > 1.96$, Therefore, we can reject Null Hypothesis and we can accept Alternative Hypothesis; $\therefore \bar{x} \neq \mu$.

P2. An oceanographer wants to check ~~the~~ whether the depth of the ocean in a certain region is 57.4 fathoms as had previously been recorded. what can he conclude at 0.05 level of significance, if readings taken at 40 random locations in the region yielded a mean of 59.1 fathoms with a standard deviation of 5.2 fathoms.

Sol) Given $n = 40$, $\bar{x} = 59.1$, $\mu = 57.4$, $s = 5.2$, $\alpha = 0.05$.

i) Null Hypothesis(H_0):-
The Sample is drawn from normal population (or)

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

ii) Alternate Hypothesis (H_1):-
the Sample is not drawn from normal population

$$\bar{x} \neq \mu.$$

iii) Level of Significance:-

Here $\alpha = 0.05$.

iv) Test statistics.

$$z = \frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}} = \frac{\bar{x} - \mu}{\frac{s}{\sqrt{n}}} = \frac{59.1 - 57.4}{\frac{5.2}{\sqrt{40}}} = 1.7 \times \frac{6.32}{5.2}$$

$$= 2.06.$$

Here $|z|$ calculated $= 2.06$

$|z|$ table $= 1.96$.

v) Conclusion.

Here $|z|$ calculated $>$ $|z|$ table i.e $2.06 > 1.96$,

$\therefore$ we can reject Null Hypothesis i.e $\boxed{\bar{x} \neq \mu}$

P3. A Sample of 900 has a mean of 3.4 cms and S.D 2.61 cms. Is this sample has been taken from a large population of mean 3.25 cms and S.D is 2.61 cms. If the population is normal then test the hypothesis with 95% fidicial limits.

Sol) Given, $n = 900$, $\bar{x} = 3.4$, $S = 2.61$

$\mu = 3.25$, $\sigma = 2.61$

$\alpha = 1 - 0.95$

$= 0.05$.

i) Null Hypothesis:-

This Sample is drawn from normal population (or).

$\bar{x} = \mu$.

ii) Alternate Hypothesis:-

The Sample is not drawn from normal population (or)

$\bar{x} \neq \mu$.

iii) Level of Significance:-

Given $\alpha = 0.05$.

3.57052

iv) Test Statistic.

$$z = \frac{|\bar{x} - \mu|}{\frac{\sigma}{\sqrt{n}}} = \frac{|3.4 - 3.25|}{\frac{2.61}{\sqrt{900}}} = \frac{0.15 \times 30}{2.61}$$

$= 1.72$.

Here $|z|$ calculated $= 1.72$

$|z|$ table $= 1.96$

Conclusion :-

Here $|z|$ calculated $\nless |z|$ table . i.e $1.72 \nless 1.96$,

we can ~~reject~~ accept Null Hypothesis. i.e $\boxed{\bar{x} = \mu}$.

Confidence limits :-

$$Z = \left(\bar{x} - Z_{\frac{\alpha}{2}} SE(\bar{x}),\ \bar{x} - Z_{\frac{\alpha}{2}} SE(\bar{x})\right)$$

$$= \left(\bar{x} - Z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}},\ \bar{x} + Z_{\alpha/2} \frac{\sigma}{\sqrt{n}}\right)$$

$$= \left(3.4 - (1.96)\left(\frac{2.61}{\sqrt{900}}\right),\ 3.4 + (1.96)\left(\frac{2.61}{\sqrt{900}}\right)\right)$$

$$= (3.4 - 0.17052,\ 3.4 + 0.17052)$$

$$= (3.22,\ 3.57052).$$

P4. A Sample 400 items is taken from a population whose S.D 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 limits for the population.

Sol: Given $n = 400$, $\sigma = 10$, $\bar{x} = 40$, $\mu = 38$,

$$\alpha = 1 - 0.95$$
$$= 0.05.$$

i) Null Hypothesis.

The Sample drawn from normal population (or)

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

ii) Alternate Hypothesis :-

The Sample is not drawn from normal population (or)

$$\bar{x} \neq \mu.$$

iii) Level of Significance:-

$\alpha = 0.05$.

iv) Test Statistics:-

$$z = \frac{|\bar{x} - \mu|}{\frac{\sigma}{\sqrt{n}}} = \frac{|40-38|}{\frac{10}{\sqrt{400}}}$$

$$= 2 \times \frac{2\cancel{0}}{1\cancel{0}} \Rightarrow 4.$$

Here $|z|$ calculated $= 4$

$|z|$ table $= 1.96$.

~~Here |z|c,~~

Conclusion:- $\mu < \mu_0$

Here $|z|$ calculated $> |z|$ table i.e $4 > 1.96$, we can reject Null Hypothesis i.e $\boxed{\bar{x} \neq \mu}$

Confidence limits:-

$$z = \left(\bar{x} - z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}, \bar{x} + z_{\frac{\alpha}{2}}\left(\frac{\sigma}{\sqrt{n}}\right)\right)$$

$$= \left(40 - (1.96)\left(\frac{10}{\sqrt{400}}\right), 40 + (1.96)\left(\frac{10}{\sqrt{400}}\right)\right)$$

$$= (40 - (1.96)(0.5), 40 + (1.96)(0.5))$$

$$= (40 - 0.98, 40 + 0.98)$$

$$z = (39.02, 40.98).$$

P5. A Ambulance Service claims that it takes on the average less than 10 minutes to reach its destination of emergency calls. A sample of 36 calls has a mean of 11 minutes and the variance of 16 minutes Test the Claim at 0.05 level of Significance

Sol) Given $n = 36$, $\bar{x} = 11$, $s^2 = 16$ and $\mu = 10$, $\alpha = 0.05$
$s = 4$

(i) Null Hypothesis:-
The Sample is drawn from normal population (or)
$$\bar{x} = \mu_0 \text{ (or) } \mu = \bar{x}$$

ii) Alternate Hypothesis:-
$$\mu < \mu_0 (\bar{x}) \quad \text{(Left tailed test)}$$
$$\text{i.e } 10 < \mu$$

iii) Level of Significance (α):
Given $\alpha = 0.05$.

iv) Test statistic:-
$$z = \frac{|\bar{x} - \mu|}{\frac{\sigma}{\sqrt{n}}} = \frac{|\bar{x} - \mu|}{\frac{s}{\sqrt{n}}} = \frac{|11 - 10|}{\frac{4}{\sqrt{36}}}$$
$$= 1 \times \frac{6}{4}$$
$$|z| = 1.5$$

Here $|z|$ calculated $= 1.5$

$|z|$ table $= 1.64$

Conclusion:-
Here $|z|$ calculated $<$ $|z|$ Table i.e $1.5 < 1.64$,
we can accept Null Hypothesis i.e $\boxed{\bar{x} = \mu}$

P5) An insurance agent has claimed that the average age of policy holders who issue through him is less than the average ~~life~~ for all agents which is 30.5 years. A random sample of 100 policy holders who had issues through him have the following age distribution.

Age	16-20	21-25	26-30	31-35	36-40
persons	12	22	20	30	16

Sol Given

$n = 100$, $\mu = 30.5$

$\bar{x} = \frac{\text{Sum of number of persons}}{\text{~~Total no. of~~ no. of persons}}$

$\bar{x} = \frac{12+22+20+30+16}{5} = \frac{100}{5} = 20$

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

$= \frac{1}{4} \left[(12-20)^2 + (22-20)^2 + (20-20)^2 + (30-20)^2 + (16-20)^2 \right]$

$\Rightarrow \frac{1}{4} \left[(-8)^2 + (2)^2 + (0)^2 + (10)^2 + (-4)^2 \right]$

$\Rightarrow \frac{1}{4} \left[64+4+0+100+16 \right]$

$= \frac{1}{4} [184]$

$S^2 = 46$

$S^2 = 46$

$\therefore S = \sqrt{46}$

$= 6.78.$

(i) Null Hypothesis.

The Sample is drawn from normal population (or)

$\bar{x} = \mu.$

ii) Alternate Hypothesis.

The Sample is not drawn from normal population (or)

$\mu < \bar{x}$ (left tailed test)

iii) Level of Significance:-

Let $\alpha = 5\%$

$= 0.05$.

iv) Test Statistics

$$z = \frac{|\bar{x} - \mu|}{\frac{\sigma}{\sqrt{n}}} = \frac{|\bar{x} - \mu|}{\frac{s}{\sqrt{n}}}$$

$$= \frac{|20 - 30.5|}{\frac{6.78}{\sqrt{100}}}$$

$$= 10.5 \times \frac{10}{6.78}$$

$$= 15.48$$

Here $|z|$ calculated $= 15.48$

$|z|$ table $= 1.64$.

v) Conclusion:-

Here $|z|$calculated $>$ $|z|$table i.e $15.48 > 1.64$, we can reject null hypothesis i.e $\boxed{\mu < \bar{x}}$.

Z Test for Difference of two means:-

Let $\bar{x}_1$ and $\bar{x}_2$ be two sample means of the large samples sizes 'n_1' and 'n_2' drawn from two normal population having mean μ_1 and μ_2 and standard deviation σ_1 and σ_2.

Null Hypothesis:-

There is no significant difference between the two population means $\mu_1 = \mu_2$ (or) The two samples are drawn from same population.

Alternative Hypothesis:-

There is a significant difference between the two population mean

$\mu_1 \neq \mu_2$ (Two Tailed)

$\left.\begin{matrix} \mu_1 < \mu_2 \\ \mu_1 > \mu_2 \end{matrix}\right\}$ (One Tailed).

The two samples are drawn from same population.

Level of significance:-

Choose the level of significance based on the problem. Default 5%.

Test statistic:-

SE = standard error

Fudicial limits:-

$z = |\bar{x}_1 - \bar{x}_2| \pm z_\alpha SE(\bar{x}_1 - \bar{x}_2)$

$$z = \frac{|\bar{x}_1 - \bar{x}_2|}{SE(\bar{x}_1 - \bar{x}_2)}$$

$$= \frac{|\bar{x}_1 - \bar{x}_2|}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \quad \text{if } \sigma_1, \sigma_2 \text{ is known}$$

$$= \frac{|\bar{x}_1 - \bar{x}_2|}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}} \quad \text{if } \sigma_1, \sigma_2 \text{ is not known.}$$

$$z = \frac{|\bar{x}_1 - \bar{x}_2|}{\sigma\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}}$$ if $\sigma_1 = \sigma_2 = \sigma$ is known.

Conclusion:-

If Z calculated value < Z critical value or table value at α% level of significance then we accept H_0. Otherwise, we reject H_0.

Problems:-

P1. The Mean yield of wheat from a district A was 210 pounds with S.D 10 pounds per acre from a sample of 100 plots. In another district B the mean yield was 220 pounds with S.D is 12 pounds from a sample of 150 plots. Assuming that the S.D of yield in the entire state was 11 pounds, test whether there is any significant difference between the mean yields of crops in the two districts?

Sol) Given $n_1 = 100$, $n_2 = 150$

$\bar{x}_1 = 210$, $\bar{x}_2 = 220$, $S_1 = 10$, $S_2 = 12$

and $\sigma_1 = \sigma_2 = \sigma = 11$.

i) Null Hypothesis:

$\mu_1 = \mu_2$ (or)

The two samples are drawn from same population.

ii) Alternative Hypothesis:

$\mu_1 \neq \mu_2$ (or)

The two samples are not drawn from same population.

iii) Level of significance:

Let $\alpha = 5\% = 0.05$.

Test statistic:-

$$z = \frac{|\bar{x}_1 - \bar{x}_2|}{\sigma\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}}$$

$$= \frac{|210-220|}{11\sqrt{\frac{1}{100}+\frac{1}{150}}} = \frac{|-10|}{11\sqrt{\frac{1}{100}+\frac{1}{150}}}$$

$$= \frac{10}{11\sqrt{0.0166}}$$

$$= \frac{10}{11 \times 0.1288}$$

$$= \frac{10}{1.41}$$

$$= 7.092.$$

Here $|z|$ calculated $= 7.092$

$|z|$ table $= 1.96.$

Conclusion.

Here $|z|$calculated $>$ $|z|$table i.e $7.092 > 1.96$

$\therefore$ we can reject Null Hypothesis i.e $\mu_1 \neq \mu_2$.

P2. The research investigator is interested in studying whether there is a significant difference in the Salaries of MBA grades in two metropolitan cities. A random sample size 100 from Mumbai yields of average income of Rs. 20,150/- Another sample of 60 from chennai 20,250. if the variances of both the populations are given as Rs. 40,000 and Rs. 32,400 respectively.

Sol) Given $n_1 = 100$, $n_2 = 60$

$\bar{x}_1 = 20,150$, $\bar{x}_2 = 20,250$

Here $S_1^2 = 40,000$, $S_2^2 = 32,400$.

$S_1 = 200$, $S_2 = 180$.

Let $\alpha = 0.05$.

i) Null Hypothesis:-

$\mu_1 = \mu_2$ (or)

The two Samples are drawn from Same population.

ii) Alternate Hypothesis

$\mu_1 \neq \mu_2$ (or)

The two Samples are not drawn from Same population.

iii) Level of Significance.

Let $\alpha = 0.05$.

iv) Test statistics :-

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

$$= \frac{|20150 - 20,250|}{\sqrt{\frac{40,000}{100} + \frac{32,400}{60}}}$$

$$= \frac{|20150 - 20,250|}{30.659}$$

$$= \frac{100}{30.659}$$

$$Z = 3.26$$

Here $|Z|$ calculated $= 3.26$

$|Z|$ table $= 1.96$.

Conclusion:-

Here $|z|_{calculated} > |z|_{table}$ i.e $3.26 > 1.96$.

$\therefore$ we can reject Null Hypothesis i.e $\mu_1 \neq \mu_2$.

P3. A Simple Sample of the height of 6400 Englishmen has a mean of 67.85 inches and S.D of 2.56 inches while a simple sample of heights of 1600 Australians has a mean of 68.55 inches and S.D of 2.52 inches. Do the Data indicate the Australians on the average taller than the Englishmen? (use 0.01).

Sol) Let Australian Sample

$n_1 = 1600, \ S_1 = 2.52, \ \bar{x}_1 = 68.55$ and

Englishmen Sample

$n_2 = 6400, \ S_2 = 2.56, \ \bar{x}_2 = 67.85$.

Let $\alpha = 0.01$

i) Null Hypothesis.

$\mu_1 = \mu_2$ (or)

The two Samples are drawn from Same population.

ii) Alternate Hypothesis. (Right Tailed)

$\mu_1 > \mu_2$. (or)

The Australians are on the average taller than the Englishmen.

iii) Level of Significance.

Let $\alpha = 0.01$.

iv) Test Statistic.

$$z = \frac{|\bar{x}_1 - \bar{x}_2|}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}}$$

$$= \frac{|68.55 - 67.85|}{\sqrt{\frac{(2.52)^2}{1600} + \frac{(2.56)^2}{6400}}}$$

$$= \frac{0.7}{0.070}$$

$$= 10.$$

Here $|z|$ calculated $= 10.$

$|z|$ table $= 2.33.$

Conclusion :-

Here $|z|$ calculated $> |z|$ table value i.e $10 > 2.33$

$\therefore$ we can reject Null Hypothesis i.e $\mu_1 > \mu_2$ which means the Australians on the average taller than the Englishmen.

P4. The average mark scored by 32 boys is 72 with a SD of 8, while that for 36 girls is 70 with a SD of 6. Does this indicate that the boys perform better than girls at level of significance 0.05.

Ans) Given, $n_1 = 32$, $S_1 = 8$, $\bar{x}_1 = 72$ and

$n_2 = 36$, $S_2 = 6$, $\bar{x}_2 = 70$ ~~and~~.

$\alpha = 0.05.$

i) Null Hypothesis :

$\mu_1 = \mu_2$ (OR)

The two samples are drawn from same population.

ii) Alternate Hypothesis :-

$\mu_1 > \mu_2$ (OR)

The Boys perform better than girls at level of significance 0.05.

iii) Level of significance.

Let $\alpha = 0.05$.

iv) Test Statistics:-

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

$$= \frac{|72 - 70|}{\sqrt{\frac{(8)^2}{32} + \frac{(6)^2}{36}}}$$

$$= \frac{2}{1.732}$$

$$Z = 1.15.$$

Here $|Z|$ calculated $= 1.154$

$|Z|$ table value $= 1.645$

Conclusion:-

Here $|Z|$ calculated $< |Z|$ table value i.e $1.154 < 1.645$.

$\therefore$ we can accept null hypothesis i.e $\boxed{\mu_1 = \mu_2}$

P5. A Company claims that its bulbs are superior to those of its main competitor. If a study showed that a sample 40 of its bulbs have a mean life time of 647 hrs of continuous use with a SD of 27 hrs. while a sample of 40 bulbs made by its main competitor had a mean life time of 638 hrs of continuous use with a SD of 31 hrs. Test the significance between the difference of two means at 5% level and fudicial limits.

(d) Given

$n_1 = n_2 = 40,$

$\bar{x}_1 = 647, \quad s_1 = 27$ and $\bar{x}_2 = 638, s_2 = 31$

$\alpha = 5\% = 0.05.$

i) Null Hypothesis:-

$\mu_1 = \mu_2$ (or)

The two samples are drawn from same population.

ii) Alternate Hypothesis:-

$\mu_1 > \mu_2$ (or)

The Company bulbs are superior to those of its main main Competitor.

iii) Level of significance.

$\alpha = 0.05.$

iv) Test statistic:-

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

$$= \frac{|647 - 638|}{\sqrt{\frac{(27)^2}{40} + \frac{(31)^2}{40}}}$$

$$= \frac{9}{6.5}$$

$z = 1.38.$

$|z|$ calculated $= 1.38$

$|z|$ table value $= 1.645$

Conclusion:-

Here $|z|$ calculated $<$ $|z|$ table value i.e $1.38 < 1.645$

$\therefore$ We can accept Null Hypothesis i.e $\mu_1 = \mu_2.$

Fudicial limits:-

$$z = |\bar{x}_1 - \bar{x}_2| \pm z_\alpha SE(\bar{x}_1 - \bar{x}_2).$$

$$= (|\bar{x}_1 - \bar{x}_2| - z_\alpha SE(\bar{x}_1 - \bar{x}_2), (\bar{x}_1 - \bar{x}_2) + z_\alpha SE(\bar{x}_1 - \bar{x}_2))$$

$$= (|647-638| - (1.645)(6.5), (|647-638| + (1.645)(6.5))$$

$$= (9 - (1.645)(6.5), 9 + (1.645)(6.5))$$

$$= (9 - 10.6925, 9 + 10.6925)$$

$$z = (-1.69, 19.6925).$$

Z- Test for single proportion.

Let p be the sample proportion of the large sample having size 'n' drawn from the normal population proportion 'P'.

Null hypothesis:

There is no significant difference between the sample proportion and population proportion $p = P$ (or)

The sample is drawn from the normal population.

Alternative hypothesis:-

There is a significant difference between the sample proportion and population proportion.

$p \neq P$ (Two Tailed)

$p < P$ (or) $p > P$ (one tailed)

The two samples are drawn from same population.

Level of significance:

Choose the level of significance based on the problem. Default 5%.

Test Statistic:-

SE = Standard Error

$Q = 1 - P$, $q = 1 - p$.

where $p = \frac{x}{n}$.

$$z = \frac{|p-P|}{SE(p)} = \frac{|p-P|}{\sqrt{\frac{PQ}{n}}} \text{, if } P \text{ is known}$$

$$\text{(or)} \quad z = \frac{|p-P|}{\sqrt{\frac{pq}{n}}} \text{, if } P \text{ is not known.}$$

Conclusion:-

If Z-calculated value $\leq$ Z-critical value or Table value at α% level of significance then we accept H_0.

Confidence limits or Fudicial limits.

18" $\quad Z = P \pm Z_{\alpha} SE(p)$. (some doubt)

P1 A manufacturer claimed that atleast 95% of the equipment which he supplied to a factory conformed to specifications. An Examination of sample of 200 pieces of equipment revealed that 18 were faulty. Test his claim at 5% level of significance?

Sol Given, $n = 200$.

Here population proportion $(P) = 95\%$

$= 0.95$.

$\therefore Q = 1 - P = 0.05$.

An Examination of sample of 200 pieces of equipment revealed that 18 were faulty $(p) = \frac{18}{200}$

$= 0.09$

$q = 1 - p$

$\Rightarrow 1 - 0.09 = 0.91$.

Null Hypothesis:

There is no significant difference between Sample proportion and population proportion $p = P_1$

Alternative Hypothesis:

There is significant difference between Sample and population proportion $p \neq P$.

Level of significance (α):

Given $\alpha = 0.05$.

Test Statistics:

$$z = \frac{|p - P|}{\sqrt{\frac{PQ}{n}}} = \frac{|0.09 - 0.95|}{\sqrt{\frac{(0.95)(0.05)}{200}}}$$

$$= \frac{|0.09 - 0.95|}{0.0154} \quad \text{(some doubt).}$$

$$= 55.8.$$

$|z|$ calculated $= 55.8$

$|z|$ table $= 1.64$.

Conclusion: Here $|z|$ calculated $>$ $|z|$ table. Hence we reject Null Hypothesis H_0 i.e. $p \neq P$.

Pr. In a Sample of 1000 people in karnataka 540 are rice eaters and the rest are wheat eaters. Can we assume that both rice and wheat are equally popular in this state at 1% level of significance?

Sol) Given $n = 1000$

540 are rice eaters among 1000 $(p) = \frac{540}{1000}$

$= 0.54$

$q = 1 - 0.54 = 0.46$, $P = 0.5$

Null Hypothesis :-

There is no significance difference between sample and population proportion $p = P$

Alternate Hypothesis:

$$p \neq P.$$

Level of significance

Given $\alpha = 0.01$.

Test statistic

$$z = \frac{|p - P|}{\sqrt{\frac{PQ}{n}}}$$

$$= \frac{|0.54 - 0.5|}{\sqrt{\frac{(0.54)(0.46)}{1000}}}$$

$$= \frac{|0.54 - 0.5|}{0.0154}$$

$$= \frac{0.04}{0.0154}$$

$$|z| = 2.593.$$

$|z|$ calculated $= 2.59$

$|z|$ table value $= 2.58$.

Conclusion:-

Here $|z|$ calculated $<$ $|z|$ table value, therefore we accept null hypothesis i.e $p = P$.

Z- test for Difference of two proportions:-

Let p_1 and p_2 be the two samples proportions of the large samples sizes 'n_1' and 'n_2' drawn from two normal populations having proportions P_1 and P_2.

Null Hypothesis:-

There is no significant difference between the population proportions $P_1 = P_2$ (or)

The two samples are drawn from same population.

Alternative Hypothesis:-

There is a significant difference between the two population proportions.

$$P_1 \neq P_2 \text{ (or)}$$

$$P_1 < P_2 \text{ or } P_1 > P_2 \text{ (one tailed).}$$

(or)

The two samples are not drawn from same population.

Level of significance (α)

Choose the level of significance based on the product. Default 5%.

Test statistic:-

SE = Standard Error, $Q_1 = 1-P_1$, $q_1 = 1-p_1$, $Q_2 = 1-P_2$, $q_2 = 1-p_2$

$$z = \frac{|p_1 - p_2|}{SE(p_1-p_2)}$$

$$= \frac{|p_1 - p_2|}{\sqrt{\frac{P_1Q_1}{n_1} + \frac{P_2Q_2}{n_2}}}, \text{ if } P_1, P_2 \text{ is known}$$

$$= \frac{|p_1 - p_2|}{\sqrt{\frac{p_1q_1}{n_1} + \frac{p_2q_2}{n_2}}}, \text{ if } P_1, P_2 \text{ is not known.}$$

$$z = \frac{|p_1 - p_2|}{\sqrt{pq\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}}$$,if P_1, P_2 is not known (method of pooling).

where $$p = \frac{n_1 p_1 + n_2 p_2}{n_1 + n_2}$$

Conclusion:-

If Z-calculated value < Z critical value or Table value at 2% level of significance then we accept H_0; otherwise, we reject H_0.

Confidence Limits or Fudicial Limits.

$$z = |p_1 - p_2| \pm z_\alpha \, SE(p_1 - p_2).$$

Q. Random Samples of 400 men and 600 women were asked whether they would like to have a flyover near their residence. 200men and 325 women were in favour of the proposal. Test the Hypothesis that proportions of men and women in favour of the proposal are same at 5% level?

Sol Given $n_1 = 400$, $n_2 = 600$

200men out of 400 are favour in proposal $(p_1) = \frac{200}{400}$

$= 0.5$

325 women out of 600 are favour in proposal $(p_2) = \frac{325}{600}$

$= 0.54$

$q_1 = 1 - p_1$

$= 1 - 0.5 = 0.5$

$q_2 = 1 - p_2$

$= 1 - 0.54$

$= 0.46.$

Null Hypothesis -

$P_1 = P_2$ (or)

The two samples are drawn from same population.

Alternate Hypothesis :-

$P_1 \neq P_2$ (or)

The two samples are not drawn from same population.

Level of significance(α)

Given $\alpha = 0.05$.

Test statistic

$$z = \frac{|P_1 - P_2|}{\sqrt{\frac{P_1 q_1}{n_1} + \frac{P_2 q_2}{n_2}}}$$

$$= \frac{|0.5 - 0.54|}{\sqrt{\frac{(0.5)(0.5)}{400} + \frac{(0.54)(0.46)}{600}}}$$

$$= \frac{|0.5 - 0.54|}{0.0322}$$

$$= \frac{0.04}{0.0322}$$

$$= 1.242.$$

$|z|_{calculated} = 1.242$

$|z|$ table value $= 1.96$.

Conclusion:

Here $|z|$calculated < $|z|$table value ∴ Hence we accept Null Hypothesis ∴ $P_1 = P_2$.

B. On the basis of their total scores, 200 candidates of a civil service Examination are divided into two groups, the upper 30% and the remaining 70%. Consider the first question of the Examination. Among the first group 40 had the correct answer whereas among the second group 80 had the correct answer. On the basis of these results can one conclude that the first question is not good at discriminating ability of the type being examined here?

Sol) Given

upper 30% of students among 200 students are in civil service examination $(n_1) = \frac{30}{100} \times 200$

$= 60.$

remaining 70% of students among 200 students are in civil service examination $(n_2) = \frac{70}{100} \times 200$

$= 140.$

And also

40 students got correct in first group among 60 and

students $(P_1) = \frac{40}{60} = 0.666$

80 students got correct in second group among 140

students $(P_2) = \frac{80}{140} = 0.5714.$

$q_1 = 1 - P_1$

$= 1 - 0.666 = 0.334$,

$q_2 = 1 - P_2$

$= 1 - 0.5714$

$= 0.4286.$

Let $\alpha = 0.05$

Null hypothesis:

$P_1 = P_2$ (or)

The two samples are drawn from same population

Alternate hypothesis.

$P_1 \neq P_2$ (or)

The two samples are not drawn from same population.

Level of significance (α)

Let $\alpha = 0.05$.

Test statistic.

$$z = \frac{|P_1 - P_2|}{\sqrt{\frac{P_1 q_1}{n_1} + \frac{P_2 q_2}{n_2}}}$$

$$= \frac{0.666 - 0.5714}{\sqrt{\frac{(0.666)(0.334)}{60} + \frac{(0.5714)(0.428)}{140}}}$$

$$= \frac{0.666 - 0.5714}{0.073}$$

$$= 1.29$$

$|z|$ calculated $= 1.29$

$|z|$ table $= 1.96$

Conclusion:-

Here $|z|$ calculated $< |z|$ table value, Hence we accept Null Hypothesis. i.e $P_1 = P_2$.

P3. In a city A, 20% of a random sample of 900 school boys has a certain slight physical defect. In another city B, 18.5% of a random sample of 1600 school boys has same defect. Is the difference between the proportions significant at 0.05 level of significance.

Sol. Given $n_1 = 900$, $n_2 = 1600$

$$P_1 = \frac{20}{100} = 0.2, \quad q_1 = 1 - P_1 = 0.8$$

$$P_2 = \frac{18.5}{100} = 0.185, \quad q_2 = 0.815.$$

Given $\alpha = 0.05$.

Null hypothesis:-

$$p_1 = p_2 \text{ (or)}$$

The two samples are drawn from same population.

Alternate hypothesis

$$p_1 \neq p_2$$

The two samples are not drawn from same population.

Level of significance (α):-

Given $\alpha = 0.05$.

Test statistic:-

$$z = \frac{|p_1 - p_2|}{\sqrt{\frac{p_1 q_1}{n_1} + \frac{p_2 q_2}{n_2}}}$$

$$= \frac{|0.2 - 0.185|}{\sqrt{\frac{(0.2)(0.8)}{900} + \frac{(0.185)(0.815)}{1600}}}$$

$$= \frac{|0.2 - 0.185|}{0.0164}$$

$$= \frac{0.015}{0.0164} = 0.914.$$

$|z|$ calculated $= 0.914$.

$|z|$ table value $= 1.96$

Conclusion:-

Here $|z|$ calculated $<$ $|z|$ table, Hence we accept null hypothesis H_0 i.e. $p_1 = p_2$.

P4. A machine puts out 9 imperfect articles in a sample of 200 articles. After the machine is overhauled (service) it puts out 5 imperfect articles in the sample of 700 articles. Test at 5% level whether the machine is improved?

SOD Given $n_1 = 200$, $n_2 = 700$

$$p_1 = \frac{9}{200}, \quad p_2 = \frac{5}{200}$$

$$= 0.045, \quad p_2 = 0.025$$

$$q_1 = 1 - 0.045, \quad q_2 = 1 - 0.025$$

$$= 0.955 \qquad = 0.975.$$

$$\alpha = 0.05.$$

Null Hypothesis:

$$P_1 = P_2$$

The two samples are drawn from same population.

Alternative Hypothesis:-

$$P_1 \neq P_2$$

The two samples are not drawn from same population.

Level of significance.

$$\alpha = 0.05.$$

Test Statistic:-

$$Z = \frac{|P_1 - P_2|}{\sqrt{\frac{P_1 q_1}{n_1} + \frac{P_2 q_2}{n_2}}}$$

$$= \frac{|0.045 - 0.025|}{\sqrt{\frac{(0.045)(0.955)}{200} + \frac{(0.025)(0.975)}{700}}}$$

$$= \frac{|0.045 - 0.025|}{0.0158}$$

$$= \frac{0.02}{0.0158} = 1.26.$$

$|z|$ calculated $= 1.26$, $|z|$ table $=$ ~~1.96~~ 1.64

Conclusion:

$|z|$calculated $<$ $|z|$Tablevalue i.e $1.26 < 1.96$. Hence we accept Null Hypothesis.

13. Before an increase on excise duty on tea 500 people out of a sample 900 found to have the habit of having tea. After an increase on excise duty 250 are have the habit of ~~tea~~ among 1100. Is there in decrease in the consumption of tea?

Sol) Given $n_1 = 900$, $n_2 = 1100$

$p_1 = \frac{500}{900}$, $p_2 = \frac{250}{1100}$

$p_1 = 0.555$, $p_2 = 0.2272$, $q_1 = 0.445$, $q_2 = 0.7728$

$\alpha = 0.05$.

Null Hypothesis:-

$p_1 = p_2$ (or)

The two samples are drawn from same population.

Alternate Hypothesis

$p_1 \neq p_2$

The two samples are not drawn from same population.

Level of significance:-

Let $\alpha = 0.05$.

Test statistic:-

$$z = \frac{|0.55 - 0.2272|}{\sqrt{\frac{(0.555)(0.445)}{900} + \frac{(0.2272)(0.7728)}{1100}}}$$

$$= \frac{0.55 - 0.2272}{0.020} = \frac{0.322}{0.020} = 16.1$$

$|z|$ calculated $= 18.1$

$|z|$ table value $= 1.64$.

Conclusion:

Here $|z|$ calculated $>$ $|z|$ table Hence we can reject null hypothesis. i-e $P_1 \neq P_2$

6. During a country wide investigation the incidence of tuberculosis was found to be 1%. In a college of 400 students 3 reported to be affected whereas in another college of 1200 students 10 were affected. Does this indicate any significant difference?

Sol. Here $n_1 = 400$, $n_2 = 1200$

$$P_1 = \frac{3}{400} = 0.0075, \quad P_2 = \frac{10}{1200} = 0.0083$$

$$q_1 = 0.9925, \quad q_2 = 1 - 0.083 = 0.9917.$$

$$P = 1\% = 1/100 = 0.01.$$

$$Q = 0.99.$$

Null Hypothesis:-

The two samples are drawn from same population

$$P_1 = P_2.$$

Alternate Hypothesis.

The two samples are not drawn from same population proportions

$$P_1 \neq P_2.$$

honor 7x

Level of Significance:-

Let $\alpha = 0.05$.

Test statistic

$$z = \frac{|p_1 - p_2|}{\sqrt{\frac{p_1 q_1}{n_1} + \frac{p_2 q_2}{n_2}}}$$

$$= \frac{|0.0075 - 0.0083|}{\sqrt{\frac{(0.0075)(0.9925)}{400} + \frac{(0.0083)(0.9917)}{1200}}}$$

$$= \frac{0.00088}{0.00504}$$

$|z| = 0.15$

Here $|z|$ calculated $= 0.15$.

$|z|$ Table value $= 1.96$.

Conclusion:- $|z|$Table value i.e $0.15 < 1.96$

Here $|z|$calculated $<$ $|z|$Table ... i.e $P_1 = P_2$.

Hence we can accept Null Hypothesis

5.7444×10^{-3}

$\sqrt{3.3 \times 10^{-5}}$

Student's t-test for single mean.

Let $\bar{x}$ be the sample mean of the small sample size drawn from the normal population having mean μ and standard deviation σ.

Null hypothesis:-

There is no significant difference between the two population mean i.e sample mean and population mean.

$$\bar{x} = \mu \text{ (or)}$$

The sample is drawn from the normal population

Alternative Hypothesis H_1

There is a significant difference between the sample mean and population mean.

$$\bar{x} \neq \mu \text{ (Two Tailed)}$$

(or)

$$\left.\begin{matrix}\bar{x} < \mu \\ \bar{x} > \mu\end{matrix}\right\} \text{(one Tailed)}$$

The sample is not drawn from same population.

Level of significance (α):-

Choose the level of significance based on the problem. Default 5%

Test statistic:

$$t = \frac{|\bar{x} - \mu|}{SE(\bar{x})} \text{ ; t-distribution with } n-1 \text{ degrees of freedom.}$$

$$z = \frac{|\bar{x} - \mu|}{\frac{\sigma}{\sqrt{n}}} \text{ , if } \sigma \text{ is known}$$

$$t = \frac{|\bar{x} - \mu|}{\frac{s}{\sqrt{n}}} \text{, if } \sigma \text{ is not known.}$$

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

Conclusion:

If t-Calculated < t critical value or table value at α% level of significance with n-1 degrees of freedom then accept H_0.

Otherwise we reject H_0.

Confidence limits or Fudicial limits

$$t = \bar{x} \pm t_\alpha SE(\bar{x}).$$

P1. A mechanist is making engine parts with axle diameters of 0.7 inch. A random sample shows diameter of 0.742 inch with SD of 0.04 inch. Complete the statistic you would use to test whether the work is meeting the specification at 0.05% level of significance.

Sol Given $\mu = 0.7$, $\bar{x} = 0.742$, $S = 0.004$, $\alpha = 0.05$

$n = 10$; degrees of freedom $= n-1 = 9.$

Null Hypothesis.

$\bar{x} = \mu$ (or)

The Sample is drawn from same population.

Alternate Hypothesis:

$\bar{x} \neq \mu$. (or)

The Sample is not drawn from Same population.

Level of significance:

Given $\alpha = 0.05$.

Test statistic

$$t = \frac{|\bar{x} - \mu|}{\frac{s}{\sqrt{n}}}$$

$$= \frac{|0.742 - 0.7|}{\frac{0.04}{\sqrt{10}}}$$

$$= \frac{|0.742 - 0.7|}{0.0126}$$

$$= 3.333$$

Here $|t|$ calculated $= 3.333$

(t) table value for 9 degrees of freedom

for two tailed test at $\frac{0.05}{2} = 0.025 = 2.262$.

Conclusion:

Here $|t|$ calculated $> |t|$ table $\therefore$ we can

reject Null Hypothesis.

P> A sample of 26 bulbs gives a mean life of 990 hours with a SD of 20 hrs. The manufactura Claims the mean life bulbs is 1000 hrs. Is the Sample not upto the standard?

Sol) Given $n = 26$, $\bar{x} = 990$, $\mu = 1000$, $S = 20$, $\alpha = 0.05$,

degrees of freedom $= n - 1$

$= 26 - 1$

$= 25$.

Null Hypothesis:

The sample is drawn from same population

$\bar{x} = \mu$.

honor 7x

Alternate Hypothesis:-

$\bar{x} \neq \mu$.

The sample is not drawn from same population.

Level of Significance:-

$\alpha = 0.05$.

Test statistic:-

$$t = \frac{|\bar{x} - \mu|}{\frac{s}{\sqrt{n}}}$$

$$= \frac{|990 - 1000|}{\frac{20}{\sqrt{26}}}$$

$$= \frac{|-10|}{3.922}$$

$$|t| = 2.549.$$

Here |t| calculated = 2.549

|t| table for 25 degrees of freedom = 2.060.

Conclusion:-

Here |t| calculated > |t| table, Hence we reject Null Hypothesis i.e $\bar{x} \neq \mu$.

P3. A random sample of 10 boys had the following IQ's 70, 120, 110, 101, 88, 83, 95, 98, 107 and 100.

a) Does this data support the assumption of a population mean?

b) Find a reasonable range in which most of them IQ values of 10 boys lie.

Sol) Given data

$n = 10$, $\mu = 100$.

70, 120, 110, 101, 88, 83, 95, 98, 107 and 100.

$$\mu = \frac{70+120+110+101+88+83+95+98+107+100}{10}$$

$= 97.2.$

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

$$= \frac{1}{10-1} \Big[\Sigma (70-97.2)^2 + (120-97.2)^2 + (110-97.2)^2 + (101-97.2)^2 + (88-97.2)^2 + (83-97.2)^2 + (95-97.2)^2 + (98-97.2)^2 + (107-97.2)^2 + (100-97.2)^2 \Big]$$

$S^2 = 203.73$

$S = \sqrt{203.73}$

$S = 14.27.$

degrees of freedom $= n-1$
$= 10-1 = 9.$

a) Null Hypothesis:-

$\bar{x} = \mu$. (or)

The Sample is drawn from same population.

Alternate Hypothesis

$\bar{x} \neq \mu$ (or)

The Sample is not drawn from same population.

Level of Significance:

$\alpha = 1\%$

$= 0.01.$

Test Statistic.

$$t = \frac{|\bar{x} - \mu|}{\frac{s}{\sqrt{n}}} = \frac{|97.2-100|}{\frac{14.27}{\sqrt{10}}}$$

$$= \frac{|97.2 - 100|}{4.512}$$

$|t| = 0.620.$

Here $|t|$ calculated $= 0.620$

$|t|$ table value at 9 degrees of freedom

at $\frac{0.01}{2} = 0.0005 = 3.250.$

Conclusion:-

Here $|t|$ calculated $<$ $|t|$ table value. Hence, we accept the Null Hypothesis.

b) Reasonable Range is

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

$$= (97.2 - (3.250)(4.512),\ 97.2 + (3.250)(4.512))$$

$$= (97.2 - 14.664,\ 97.2 + 14.664)$$

$$= (82.53,\ 111.864).$$

$\therefore$ The Reasonable Range is $(82.53, 111.864)$.

A producer of gutkha claims that the nicotine content in gutkha on the average is 18.3 mg. Can this claim accepted if a random sample of 8 gutkha of this type have the nicotine contents of 2, 1.7, 2.1, 1.9, 2.2, 2.1, 2, 1.6 mg? use a 0.05 level of significance?

Sol: Given $n = 8$, $\mu = 1.83$ mg, $\alpha = 0.05$

$$\bar{x} = \frac{2 + 1.7 + 2.1 + 1.9 + 2.2 + 2.1 + 2 + 1.6}{8}$$

$= 1.95.$

$$s^2 = \frac{1}{8-1}\left[\Sigma (2 - 1.95)^2 + (1.7 - (1.95))^2 + (2.1 - 1.95)^2 + (1.9 - 1.95)^2 + (2.2 - 1.95)^2 + (2.1 - 1.95)^2 + (2 - 1.95)^2 + (1.6 - 1.95)^2\right.$$

$S^2 = 0.0428.$

$= 0.20688$

Null Hypothesis :-

$\bar{x} = \mu$ (or)

The sample is drawn from same population.

Alternate Hypothesis.

$\bar{x} \neq \mu$ (or)

The sample is not drawn from same population.

Level of significance.

$\alpha = 0.05.$

Test statistic :-

$$t = \frac{|\bar{x} - \mu|}{\frac{s}{\sqrt{n}}}$$

$$= \frac{|1.95 - 1.83|}{\frac{0.20688}{\sqrt{8}}}$$

$$= \frac{(1.95 - 1.83)}{0.073}$$

$= 1.64.$

Here $|t|$ calculated $= 1.64.$

$|t|$ table at degrees of freedom 7 is

at two tailed test $\frac{0.05}{2} = 0.025 = 2.365$

Conclusion.

Here $|t|$ calculated $<$ $|t|$ table $\therefore$ we can accept

the Null Hypothesis i.e $\bar{x} = \mu$.

A random sample of 8 envelopes is taken from the letter box of a post office and their weights in grams are found to be : 12.1, 11.9, 12.4, 12.3, 11.5, 11.6, 12.1 and 12.4

a) Does this sample indicate 1% level that the average weight of envelopes received at their post office is 12.35 gms.

b) Find 95% Confidence limits for the mean weight of the envelopes received at that post office.

Ans) Given $n=8$, $\mu = 12.35$,

$$\bar{x} = \frac{12.1 + 11.9 + 12.4 + 12.3 + 11.5 + 11.6 + 12.1 + 12.4}{8}$$

$$= 12.03$$

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

$$= \frac{1}{8-1}\Big[\Sigma\,(12.1-12.03)^2 + (11.9-12.03)^2 + (12.4-12.03)^2 + (12.3-12.03)^2 + (11.5-12.03)^2 + (11.6-12.03)^2 + (12.1-12.03)^2 + (12.4-12.03)^2\Big]$$

$$= \frac{1}{7}\left[0.8392\right]$$

$s^2 = 0.11988$

$s = \sqrt{0.11988}$

$s = 0.346.$

Now $\alpha = 1 - 0.95$
$= 0.05.$

a) Null hypothesis:

$\bar{x} = \mu$ (or)

Sample is drawn from sample population.

b) Alternate Hypothesis:

$\bar{x} \neq \mu$.

The sample is not drawn from same population.

c) Level of significance.

here $\alpha = 0.05$.

d) Test statistic:

$$t = \frac{|\bar{x} - \mu|}{\frac{s}{\sqrt{n}}}$$

$$= \frac{|12.03 - 12.35|}{\frac{0.346}{\sqrt{8}}} \quad 0.122$$

$$= \frac{0.32}{0.122}$$

$$= 2.622.$$

$|t|_{calculated} = 2.622$

$|t|_{table}$ at degrees of freedom $8-1 = 7$ at 0.05 level of significance i.e $\frac{0.05}{2} = 0.025 = 2.365$.

Conclusion:

here $|t|_{calculated} > |t|_{table}$ i.e $2.365 > 2.622$

$\therefore$ we can reject Null hypothesis i.e $\bar{x} \neq \mu$.

b) Confidence limits =

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

$$\Rightarrow \left(\bar{x} - t_{\alpha/2}\frac{s}{\sqrt{n}},\ \bar{x} + t_{\alpha/2}\frac{s}{\sqrt{n}}\right)$$

$$= (12.03 - (2.365)(0.122)),\ 12.03 + (2.365)(0.122))$$

$$= (12.03 - 0.288,\ 12.03 + 0.288)$$

$$= (11.742,\ 12.318).$$

Random sample of 10 bags of pesticide are taken whose weights are 50, 49, 52, 44, 45, 48, 46, 45, 49, 45. Test whether the average packing can be taken to be 50 kgs.

Given $n = 10$, $\mu = 50$.

$$\bar{x} = \frac{50+49+52+44+45+48+46+45+49+45}{10}$$

$$\bar{x} = 47.3$$

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

$$= \frac{1}{10-1}\Big[\Sigma\big[(50-47.3)^2 + (49-47.3)^2 + (52-47.3)^2 + (46-47.3)^2 + (48-47.3)^2 + (44-47.3)^2 + (45-47.3)^2 + (45-47.3)^2 + (45-47.3)^2 + (49-47.3)^2\big]\Big]$$

$$= \frac{1}{9}[64.1]$$

$$= 7.122$$

$$s = \sqrt{7.122}$$

$$= 2.668$$

$\alpha = 0.05$, $\nu = 10 - 1 = 9$ $\Rightarrow \frac{0.05}{2} = 0.025$

Null Hypothesis:

$\bar{x} = \mu$ (or)

The sample is drawn from same population.

Alternate Hypothesis:

$\bar{x} \neq \mu$.

The sample is not drawn from same population.

Level of significance

$\alpha = 0.05$.

Test statistic:-

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

$$= \frac{|147.3 - 50|}{\frac{2.668}{\sqrt{10}}}$$

$$= \frac{|147.3 - 50|}{0.8436}$$

$$= \frac{2.7}{0.8436}$$

$$|t| = 3.200.$$

Here $|t|_{calculated} = 3.200$

$|t|_{table} = 2.262.$

Conclusion:-

Here $|t|_{calculated} > |t|_{table}$, Hence we reject Null Hypothesis i.e $\bar{x} \neq \mu$.

Student's t-test for Difference of two Means:-

Let $\bar{x}_1$ and $\bar{x}_2$ be two sample means of the small sample sizes "n_1" and "n_2" drawn from two normal population having mean μ_1 and μ_2 and standard deviation σ_1 and σ_2.

Null Hypothesis:-

There is no significant difference between the two population means $\mu_1 = \mu_2$ (or) The two samples are drawn from same population.

Alternate Hypothesis

There is a significant difference between the two population means.

$\mu_1 \neq \mu_2$ (Two Tailed)

$\mu_1 < \mu_2$
} one Tailed.
$\mu_1 > \mu_2$

The two samples are not drawn from same population.

Level of significance:

Choose the level of significance based on the problem

Default 5%.

Test statistic:-

$$t = \frac{|\bar{x}_1 - \bar{x}_2|}{SE(\bar{x}_1 - \bar{x}_2)}, \text{ t- distribution } n_1 + n_2 - 2 \text{ degrees of freedom.}$$

$$= \frac{|\bar{x}_1 - \bar{x}_2|}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} \quad \text{if } \sigma_1, \sigma_2 \text{ is not known.}$$

$$= \frac{|\bar{x}_1 - \bar{x}_2|}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \quad \text{if } \sigma_1, \sigma_2 \text{ is known.}$$

$$= \frac{|\bar{x}_1 - \bar{x}_2|}{S\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \quad \text{if } s_1 = s_2 = s$$

$$\text{where } S = \sqrt{\frac{n_1 s_1^2 + n_2 s_2^2}{n_1 + n_2 - 2}} \quad \text{(or)}$$

$$S = \sqrt{\frac{1}{n_1 + n_2 - 2}\left[\Sigma(x_i - \bar{x}_1)^2 + \Sigma(x_i - \bar{x}_2)^2\right]}$$

Conclusion:-

If t-Calculated value < t critical value or table value at α % level of significance with $n_1 + n_2 - 1$ degrees of freedom then we accept otherwise reject the

Confidence limits or Fudicial limits.

$$t = |\bar{x}_1 - \bar{x}_2| \pm t_\alpha \, SE(\bar{x}_1 - \bar{x}_2)$$

F-Test :-

When testing the significance of the difference of two means of two samples, we assumed that the two samples came from the same population or from populations with equal variances. If the variances of, the population are not equal, a significant difference in the means may arise. Hence, before we apply the t-test for the significance of the difference of two means, we have to test for the equality of population variances using F-Test of significance.

If S_1^2 and S_2^2 are the two samples variances with sample size n_1 and n_2 drawn from the normal population.

Null Hypothesis :-

There is no significant difference between the two population variances $\sigma_1^2 = \sigma_2^2$ (or)

The samples are drawn from same population.

Alternative Hypothesis :-

There is a significant difference between two population variances

$\sigma_1^2 \neq \sigma_2^2$ (Two Tailed) (or)

$$\left.\begin{array}{l}\sigma_1^2 < \sigma_2^2 \\ \sigma_1^2 > \sigma_2^2\end{array}\right\} \text{(One Tailed)}$$

The samples are not drawn from same population.

Level of significance:-

Choose the level of significance based on the problem. Default 5%.

Test statistic:

$$F = \frac{\text{Greater variance}}{\text{Smaller variance}}$$

$F = \frac{S_1^2}{S_2^2}$ if $S_1^2 > S_2^2 \sim$ F-distribution with (n_1-1, n_2-1) degrees of freedom.

$F = \frac{S_2^2}{S_1^2}$ if $S_2^2 > S_1^2 \sim$ F distribution with (n_2-1, n_1-1) degrees of freedom.

where $S_1^2 = \frac{1}{n_1-1}\left[\Sigma (x_i - \bar{x}_1)^2\right.$

$S_2^2 = \frac{1}{n_2-1}\left[\Sigma (y_i - \bar{y})^2\right]$

$\Sigma (x_i - \bar{x})^2$ = sum of squares of the deviations from the sample mean.

Conclusion:-

If F-calculated value < F critical value or table value at α% level of significance with (n_1-1, n_2-1) or (n_2-1, n_1-1) degrees of freedom then we accept H_0.