UNIT - III

Testing of Hypothesis

Samples

- Large Sample (Sample size $n \geq 30$)
- Small Sample (Sample size $n < 30$)

Large sample	Small Sample
1) Test for single sample mean	1) t - test
2) Test for difference of two sample means	2) F - test
3) Test for single sample proportion	3) chi-square (χ^2) test

procedure for testing of hypothesis:-

Step 1: Set up Null Hypothesis H_0

H_0: There is no significant differense between sample statistic and population parameter

Step 2

Set up Alternative Hypothesis H_1

H_1: There is a difference between sample statistic and population parameter

Step 3

To fix LoS (Level of Significance) α.

$\alpha = 1\%$ or 5% (level)

[Acceptance or Rejection level]

Step 4

Compute test statistic z about the sample with the help of statistical formula

Step 5

See the table value at 1% or 5% level.

Step 6

Write the conclusion based on the following decisions.

(i) Accept null hypothesis Ho

$$\boxed{\text{Calculated value of } Z < \text{Table value of } Z_\alpha}$$

(ii) Reject null Hypothesis Ho

$$\boxed{\text{Calculated value of } Z > \text{Table value of } Z_\alpha}$$

Two tailed and One tailed test

Let θ_0 be the population parameter and θ be the sample statistic.

Two tailed test

Null Hypothesis $H_0 : \theta = \theta_0$

Alternative Hyp. $H_1 : \theta \neq \theta_0$

One-tailed test

Null Hypothesis $H_0 : \theta = \theta_0$

Alternative Hypo $H_1 : \theta < \theta_0$ (Left-tailed)

(or) $\theta > \theta_0$ (Right-tailed)

Error

Type I Error : Reject H_0 when it is true

Type II Error : Accept H_0 when it is wrong.

Table value of large samples

Test	Level of significance (α)	
	1%	5%
Two Tailed	$\lvert Z_\alpha \rvert = 2.58$	$\lvert Z_\alpha \rvert = 1.96$
One Tailed	$\lvert Z_\alpha \rvert = 2.33$	$\lvert Z_\alpha \rvert = 1.645$

Large Sample Test

Test - I : Test for significance difference between sample proportion and population proportion

procedure:-

Null Hypothesis $H_0 : p = P$

[There is no difference b/w sample and population proportion]

Alternative Hypothesis $H_1 : p \neq P$

LOS α : 1% or 5% level

Test Statistic $$Z = \frac{p - P}{\sqrt{\frac{PQ}{n}}}$$

Where $p \rightarrow$ sample proportion

$P \rightarrow$ population proportion

$n \rightarrow$ sample size

Conclusion

calculated value of Z $<$ Table value of Z_α [Accept H_0]

calculated value of Z $>$ Table value of Z_α [Reject H_0]

(7) In a city, a sample of 1000 people were taken out of them 540 are vegetarians and the rest are non vegetarians. Can we say that both habits of eating are equally popular in the city at 5% level of significance.

<u>Solution</u>:

sample size $n = 1000$

Sample proportion $p = \frac{540}{1000} = 0.54$ (sample proportion of veg)

population proportion $P = \frac{1}{2}$ (population prop. of veg)

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

Null Hypothesis $H_0 : p = P$ [Both habits are ~~are~~ equally popular in the city]

Alternative Hypothesis $H_1 : p \neq P$ [TWO TAILED TEST]

LOS : 5% level

Test Statistic $Z = \frac{p - P}{\sqrt{\frac{PQ}{n}}} = \frac{0.54 - 0.5}{\sqrt{\frac{0.5 \times 0.5}{1000}}}$

$$Z = 2.5298$$

calculated value of Z	Table value of $Z_{5\%}$
2.5298	1.96

$\therefore \boxed{Z > Z_{5\%}}$

$\therefore$ Reject null Hypothesis H_0.

Hence both habits of eating are not equally popular in the city.

(2). The fatality rate of typhoid patients is believed to be 17.26 %. In a certain year 640 patients suffering from typhoid were treated in a metropolitan hospital and only 63 patients died. Can you consider the hospital is efficient?

Solution:

Sample size $n = 640$

Sample proportion $p = \frac{63}{640} = 0.0984$

population proportion $P = 0.1726$

$Q = 1 - P = 0.8274$

Null Hypothesis $H_0 : p = P$ [The hospital is not efficient]

Alternative Hypo% $H_1 : p < P$

(one tailed test)

LOS : 5% level

Test statistic $Z = \dfrac{p - P}{\sqrt{\dfrac{PQ}{n}}}$

$= \dfrac{0.0984 - 0.1726}{\sqrt{\dfrac{0.1726 \times 0.8274}{640}}}$

$= -4.96$

$|Z| = 4.96$

Cal. value of Z	Table value of $Z_{5\%}$ for one tailed test
4.96	1.645

$\therefore Z > Z_{5\%}$

$\therefore$ Reject null hypothesis H_0.

Hence the hospital is efficient

Test-II (Test for significance difference between two sample proportions)

procedure:

Null Hypothesis $H_0 : p_1 = p_2$ [There is no difference b/w two sample proportions]

Alternative Hypothesis $H_1 : p_1 \neq p_2$

LOS α : 1% or 5% level

Test statistic $$Z = \frac{p_1 - p_2}{\sqrt{PQ\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}}$$

where p_1 = sample I proportion

p_2 = Sample II proportion

n_1 = Sample I size

n_2 = Sample II size

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

$$Q = 1 - P$$

(1) In a large city A, 20% of a random sample of 900 school boys had a slight physical defect. In another large city B, 18.5% of a random sample of 1600 school boys had the same defect. Is the difference between the proportions are significant?

Solution:

$$n_1 = 900, \quad p_1 = 20\% = 0.2$$

$$n_2 = 1600, \quad p_2 = 18.5\% = 0.185$$

$$P = \frac{n_1 p_1 + n_2 p_2}{n_1 + n_2} = \frac{(900 \times 0.2) + (1600 \times 0.185)}{900 + 1600}$$

$$= 0.1904$$

$$Q = 1 - P = 1 - 0.1904 = 0.8096$$

Null Hypothesis $H_0 : p_1 = p_2$

Alternative Hyp% $H_1 : p_1 \neq p_2$

LOS α : 5% level

$$\text{Test statistic } Z = \frac{p_1 - p_2}{\sqrt{PQ\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

$$= \frac{0.2 - 0.185}{\sqrt{0.1904 \times 0.8096 \left(\frac{1}{900} + \frac{1}{1600}\right)}}$$

$$\boxed{Z = 0.92}$$

Calculated value of Z	Table value of $Z_{5\%}$
0.92	1.645

$\therefore \quad Z < Z_{5\%}$

$\therefore$ Accept Null Hypothesis H_0.

(2). Out of a sample of 1000 persons 800 were found to be coffee drinkers. Subsequently, the excise duty on coffee was increased. After the increase in excise duty of coffee seeds, 800 people were found to take coffee out of sample 1200. Test whether there is any significant decrease in the consumption of coffee after the increase in excise duty.

<u>solution:</u>

$$n_1 = 1000, \quad p_1 = \frac{800}{1000} = 0.8$$

$$n_2 = 1200, \quad p_2 = \frac{800}{1200} = 0.67$$

$$P = \frac{n_1 p_1 + n_2 p_2}{n_1 + n_2} = \frac{4}{9} = \cancel{0.4444}\ 0.7273$$

$$Q = 1 - P = \cancel{0.5555}\ 0.2727$$

Null Hypothesis $H_0 : p_1 = p_2$

Alternative Hypothesis $H_1 : p_1 > p_2$

Test Statistic $$Z = \frac{p_1 - p_2}{\sqrt{PQ\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

$$Z = \frac{0.8 - 0.67}{\sqrt{0.7273 \times 0.2727 \left(\frac{1}{1000} + \frac{1}{1200}\right)}} = 6.82$$

Calculated value of Z	Table value of $Z_{5\%}$
6.82	1.645

$\therefore Z > Z_{5\%}$

$\therefore$ $\boxed{\text{Reject } H_0}$

Test - III : Test for significance difference between single sample mean and population mean:

procedure

Null Hypothesis $H_0 : \bar{x} = \mu$ [There is no significant difference b/w sample and population mean]

Alternative Hypo/. $H_1 : \bar{x} \neq \mu$

LOS : 1% or 5% level

Test Statistic $$Z = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}}$$

where $\bar{x} \rightarrow$ sample mean

$\mu \rightarrow$ population mean

$\sigma \rightarrow$ population S.D

$n \rightarrow$ sample size

Note: If the population S.D is not known, then we use the statistic $$Z = \frac{\bar{x} - \mu}{s/\sqrt{n}}$$

where s is the sample S.D

(1) A Sample of 100 students is taken from a large population. The mean height of the students in this sample is 160 cm. Can it be reasonably regarded that, in the population, the mean height is 165 cm and the S.D is 10 cm?

<u>Solution.</u>

Here $n = 100$, sample mean $\bar{x} = 160$

population mean $\mu = 165$

population S.D $\sigma = 10$

Null Hypothesis $H_0 : \bar{x} = \mu$

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

LOS α : 5% level

Test statistic $Z = \dfrac{\bar{x} - \mu}{\sigma/\sqrt{n}}$

$= \dfrac{160 - 165}{10\sqrt{100}}$

$= -5$

$|Z| = 5$

Calculated value of Z	Table value of $Z_{5\%}$
5	1.96

$\therefore Z > Z_{5\%}$

$\therefore \boxed{\text{Reject } H_0}$

(2) The mean breaking strength of the cables supplied by a manufacturer is 1800 with a S.D of 100. By a new technique in the manufacturing process, it is claimed that the breaking strength of the cable has increased. In order to test this claim, a sample of 50 cables is tested and it is found that the mean breaking strength is 1850. Can we support the claim at 1% level of significance.

Solution:-

$$n = 50, \quad \mu = 1800, \quad \sigma = 100$$

$$\text{and} \quad \bar{x} = 1850$$

Null Hypothesis H_0 : $\bar{x} = \mu$ (Breaking strength is not increased)

Alternative Hypo/. H_1 : $\bar{x} > \mu$ (one tailed test)

LOS α : 1% level

$$\text{Test Statistic } Z = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}} = \frac{1850 - 1800}{100\sqrt{50}}$$

$$Z = 3.54$$

Cal. value of Z	Table value of Z at 1% for one tailed test
3.54	2.33

$\therefore Z > Z_{1\%}$ So we $\boxed{\text{Reject } H_0}$

Test IV: Test for significance difference between the difference of two sample means.

procedure:

Null Hypothesis H_0 : $\overline{x}_1 = \overline{x}_2$

Alternative Hypo/. H_1 : $\overline{x}_1 \neq \overline{x}_2$

LOS : 1% or 5% level

Test Statistic $\boxed{Z = \dfrac{\overline{x}_1 - \overline{x}_2}{\sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}}}$

$\overline{x}_1$ – Sample I mean

$\overline{x}_2$ – Sample II mean

σ_1^2 – Population I S.D

σ_2^2 – population II S.D

n_1 – Sample I size

n_2 – Sample II size

Note: If the samples have been drawn from the same population, then $\sigma_1^2 = \sigma_2^2 = \sigma^2$

$$\therefore \boxed{Z = \frac{\overline{x}_1 - \overline{x}_2}{\sqrt{\dfrac{\sigma^2}{n_1} + \dfrac{\sigma^2}{n_2}}}}$$

If σ is not known directly, we have samples S.D s_1 and s_2 then we estimate $\boxed{\sigma^2 = \dfrac{n_1 s_1^2 + n_2 s_2^2}{n_1 + n_2}}$

(1) In a random sample of size, the mean is found to be 20. In another independent sample of size 400, the mean is 15. Could the samples have been drawn from the ~~samp~~ same population with S.D is 4?

Solution:-

Sample size I $n_1 = 500$

Sample size II $n_2 = 400$

Sample I mean $\bar{x}_1 = 20$

Sample II mean $\bar{x}_2 = 15$

population S.D $\sigma = 4$

Null Hypothesis $H_0 : \bar{x}_1 = \bar{x}_2$

Alternative Hypo/. $H_1 : \bar{x}_1 \neq \bar{x}_2$

LOS α : 1% or 5% level

$$\text{Test Statistic } Z = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{\sigma^2}{n_1} + \frac{\sigma^2}{n_2}}}$$

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

$$= \frac{20 - 15}{4\sqrt{\frac{1}{500} + \frac{1}{400}}}$$

$$= 18.6$$

Calculate value of Z	Table value of Z at 5% for two-tailed test
$Z = 18.6$	$Z_{5\%} = 1.96$

$$\therefore \quad Z > Z_{5\%}$$

So we reject null Hypothesis H_0.

(2) The average marks scored by 32 boys is 72 with a S.D of 8, while that for 36 girls is 70 with a S.D of 6. Test at 1% LOS whether the boys perform better than girls.

Solution:

Sample I size $n_1 = 32$

Sample II size $n_2 = 36$

Sample I mean $\bar{x}_1 = 72$

Sample II mean $\bar{x}_2 = 70$

Sample I S.D $s_1 = 8$

Sample II S.D $s_2 = 6$

Null Hypothesis $H_0 : \bar{x}_1 = \bar{x}_2$

Alternative Hypo $H_1 : \bar{x}_1 > \bar{x}_2$ (one-tailed test)

LOS $\alpha = 1\%$

Now Test Statistic $Z = \dfrac{\bar{x}_1 - \bar{x}_2}{\sigma\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$

Here $\sigma = \dfrac{n_1 s_1^2 + n_2 s_2^2}{n_1 + n_2}$

$= \dfrac{(32 \times 64) + (36 \times 36)}{32 + 36} = 36$

$$\therefore \; Z = \frac{72 - 70}{36\sqrt{\frac{1}{32} + \frac{1}{36}}}$$

$$= 1.15$$

Calculated value of Z	Table value of Z at 1% LOS for one-tailed test
1.15	2.33

$$\therefore \; Z < Z_{1\%}$$

So we accept H_0.

$\therefore$ Boys and girls have equal performance.