Large Sample :- If the Sample Size $n \geq 30$ is called large Sample

Mean & variance of Sample :-

$$\text{Mean } (\bar{x}) = \frac{\Sigma x_i}{n}, \quad n = \text{no. of observations}$$

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

Mean & variance of population :-

$$\text{Mean } (\mu) = \frac{\Sigma x_i}{N}, \quad N = \text{no. of observation}$$

$$\text{Variance } (\sigma^2) = \frac{\Sigma (x_i - \mu)^2}{N}$$

Method 1 : Test of Significance for Single proportion

Suppose a large random Sample of Size n has a Sample proportion p of members possessing a Certain attribute.

To test the hypothesis that the proportion P in the population has a Specified value.

P_0 : The test significance is

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

where p = Sample proportion

P = Population proportion

n = Sample Size, $n \geq 30$

here $p = \frac{x}{n}$

x = Success of occurence

$$P + Q = 1 \Rightarrow Q = 1 - P$$

	α level of significance			
	1%	5%	10%	2%
Two-tailed	$\lvert Z_{\alpha/2}\rvert = 2.58$	1.96	1.645	2.33
Right-tailed	$Z_\alpha = 2.33$	1.645	1.28	
left-tailed	$Z_\alpha = -2.33$	-1.645	-1.28	

Prob:- 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 data,

Sample Size $(n) = 1000$,

$x = 540$,

$\alpha = 1\%$.

$$\text{Sample proportion } (p) = \frac{x}{n} = \frac{540}{1000} = 0.54$$

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

$$Q = 1 - P = 1 - \frac{1}{2} = \frac{1}{2} = 0.5$$

N.H $[H_0]$: Both rice and wheat eaters are equally popular in Karnataka state.

i.e, $P = P_0 = 0.5$

A.H $[H_1]$: Both rice and wheat eaters are not popular in karnataka state.

i.e, $P \neq 0.5$

clearly it is 2-tailed test.

Los (α) Given $\alpha = 1\% = 0.01$

Test statistic (Z_{cal}):

we know that,

$$Z_{cal} = \frac{p - P}{\sqrt{PQ/n}} = \frac{0.54 - 0.5}{\sqrt{\frac{(0.5)(0.5)}{100}}}$$

$$Z_{cal} = 2.532$$

Tabulated value (Z_{tab}):

from table value at 1% level of Significance

$$Z_{tab} = Z_{\alpha/2} = Z_{\frac{0.01}{2}} = Z_{0.005} = 2.58$$

$$Z_{tab} = 2.58$$

Conclusion:

$$|Z|_{cal} = 2.532 \quad , \quad Z_{tab} = 2.58$$

$\Rightarrow$ $|Z| < Z_{tab}$ then

we can accept the N.H at 1% level of Significance.

$\therefore$ Both rice and wheat eaters are equally popular in Karnataka State.

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

Sol:- Given $n = 600$

No. of Smokers $= 325$

p = Sample proportion of smokers $= \frac{325}{600} = 0.5417$

P = population proportion of smokers in the city $= \frac{1}{2} = 0.5$

$Q = 1 - P = 1 - 0.5 = 0.5$

Testing of Hypothesis:

Null hypothesis [H_0]: The number of smokers and non-smokers are equal in the city.

Alternative hypothesis: $P > 0.5$ (right tailed)

The Test statistic is $Z = \frac{p - P}{\sqrt{\frac{PQ}{n}}} = \frac{0.5417 - 0.5}{\sqrt{\frac{0.5 \times 0.5}{600}}} = 2.04$

$\therefore$ Calculated value of Z (Z_{cal}) $= 2.04$

Tabulated value of Z at 5% level of significance for right tail test is 1.645.

Since, Calculated value of $Z >$ tabulated value of Z, we reject the null hypothesis and conclude that the majority of men in the city are smokers.

Prob: A die was thrown 1000 times and of these 3220 yielded 3 or 4. Is this Consistent with the hypothesis that the die was unbiased.

Sol: Given data $n = 9000$, $x = 3220$

$$p = \frac{x}{n} = \frac{3220}{900} = 0.3578$$

P = Population proportion of Success

$= P(\text{getting } 3 \text{ or } 4)$

$= P(x=3) + P(x=4)$

$= \frac{1}{6} + \frac{1}{6}$

$= \frac{2}{6}$

$= 0.3333$

$Q = 1 - P = 0.6667$

Testing of Hypothesis :—

Null Hypothesis $[H_0]$: the die was unbiased.

ie $P = 0.3333$

Alternative hypothesis $[H_1]$: The die was unbiased.

$P \neq 0.3333$

Clearly it is two-tailed test

level of Significance (α) : Assume $\alpha = 5\% = 0.05$

Test Statistic $[Z_{cal}]$:

we know that $Z_{cal} = \frac{p-P}{\sqrt{PQ/n}} = \frac{0.3578 - 0.3333}{\sqrt{\frac{(0.3333)(0.6667)}{9000}}}$

$Z_{cal} = 4.94$

Test Tabulated value $[Z_{tab}]$:

from the table at 5% of Los

$Z_{tab} = Z_{\alpha/2} = Z_{\frac{0.05}{2}} = 1.645$

Conclusion :

$|Z|_{cal} = 4.94$ & $Z_{tab} = 1.645$

$|Z| > Z_{tab}$

we reject the N.H at 5% level of significance.

we Accept the A.H at 5% level of significance.

∴ The die was biased.

Prob: In a random sample of 125 cool drinkers, 68 said they prefer thumsup to pepsi. Test the null hypothesis $P = 0.5$ against the alternative hypothesis $P > 0.5$

Sol: Given data

$n = 125$, $x = 68$, $P = 0.5$, $Q = 1 - P = 0.5$

$$p = \frac{x}{n} = \frac{68}{125} = 0.544$$

Test of Hypothesis :-

N.H $[H_0]$: $P \neq 0.5$

A.H $[H_1]$: $P > 0.5$

clearly it is right tailed test.

LOS (α) :- Assume $\alpha = 5\% = 0.05$

Test Statistic $[Z_{cal}]$:-

We know that $$Z = \frac{p - P}{\sqrt{\frac{PQ}{n}}}$$

$$= \frac{0.544 - 0.5}{\sqrt{\frac{(0.5)(0.5)}{125}}}$$

$$= 0.9839$$

Tabulated value $[Z_{tab}]$:- from the table at 5% level of significan

$$Z_{tab} = Z_\alpha = Z_{0.05} = 1.645$$

Conclusion :-

$|Z|_{cal} = 0.9839$ & $Z_{tab} = 1.645$

$|Z|_{cal} < Z_{tab}$ then we accept N.H at 5% lo

$\therefore P = 0.5$

Prob: Experience had shown that 20% of a manufactured product is of the top quality. In one day's production of 400 articles only 50 are of top quality. Test the hypothesis at 0.05 level.

Sol: Given that $n = 400$, $x = 50$, $P = 20\% = 0.2$

$Q = 1 - P = 0.8$, $\alpha = 0.05$, $p = \frac{x}{n} = \frac{50}{400} = 0.125$

N.H (H_0): $P = 0.2$

A.H (H_1): $P \neq 0.2$

clearly it is two tailed test

LOS (α): Given $\alpha = 0.05$

Test statistic $[Z_{cal}]$:

$$Z = \frac{p - P}{\sqrt{\frac{PQ}{n}}} = \frac{0.125 - 0.2}{\sqrt{\frac{(0.2)(0.8)}{400}}} = -3.75$$

Tabulated value $[Z_{tab}]$:

from the table value at 5% LOS

$$Z_{tab} = Z_{\alpha/2} = Z_{\frac{0.05}{2}} = 1.96$$

Conclusion

$|Z_{cal}| = |-3.75| = 3.75$

$Z_{tab} = 1.96$

$|Z_{cal}| > Z_{tab}$

We reject H_0 at 5% LOS, and accept H_1 at 5% LOS $\therefore P \neq 0.2$

Method - II : Test of Significance for difference between two Sample Proportions (Large Samples)

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

To test whether the two samples have been drawn from the same population.

$$\boxed{Z = \frac{p_1 - p_2}{\sqrt{pq\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}}}, \quad \text{here } p_1 = \frac{x_1}{n_1}$$

$$p_2 = \frac{x_2}{n_2}$$

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

$$q = 1 - p$$

Prob: Random samples of 400 men and 600 women were asked whether they would like to have a flyover near their residence. 200 men 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 data $n_1 = 400$ $\quad x_1 = 200$

$n_2 = 600$ $\quad x_2 = 325$ $\quad \alpha = 5\%$

Sample proportions, $p_1 = \frac{x_1}{n_1} = \frac{200}{400} = 0.5$

$$p_2 = \frac{x_2}{n_2} = \frac{325}{600} = 0.54$$

esting of hypothesis :-

Null hypothesis (H_0) :- There is no difference between men Proposal and women proposal.

i.e, $H_0 : P_1 = P_2 = P$

Alternative hypothesis [H_1] :- There is a difference between men Proposal and women proposal.

i.e $H_1 : P_1 \neq P_2$

clearly, it is 2-tailed test.

Level of Significance (α) :- Given $\alpha = 5\% = 0.05$

Test Statistic [Z_{cal}] :- we know that,

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

$$= \frac{0.5-0.54}{\sqrt{(0.52)(0.48)\left(\frac{1}{400}+\frac{1}{600}\right)}}$$

$$= -1.26$$

$$p = \frac{n_1 P_1 + n_2 P_2}{n_1 + n_2}$$

$$= \frac{(400 \times 0.5) + (600 \times 0.54)}{400 + 600}$$

$$= 0.52$$

$$q = 1 - p = 1 - 0.52 = 0.48$$

Tabulated value [Z_{tab}] :- from the table at 5% level of significance :-

$$Z_{tab} = Z_{\alpha/2} = Z_{\frac{0.05}{2}} = 1.96$$

Conclusion :-

$$|Z_{cal}| = |-1.26| = 1.26$$

$$Z_{tab} = 1.96$$

$\Rightarrow |Z_{cal}| < Z_{tab}$ then we can accept the null hypothesis at 5% level of significance.

$\therefore$ There is no difference between men proposal and women proposal.

Prob:- 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 that,

Total students $= 200$

$$n_1 = 30\% \text{ of } 200 = \frac{30}{100} \times 200 = 60$$

$$n_2 = 70\% \text{ of } 200 = \frac{70}{100} \times 200 = 140$$

$x_1 = 40$, $x_2 = 80$

$$p_1 = \frac{x_1}{n_1} = \frac{40}{60} = \frac{2}{3} = 0.667$$

$$p_2 = \frac{x_2}{n_2} = \frac{80}{140} = \frac{4}{7} = 0.571$$

Testing of Hypothesis

Null Hypothesis [H_0] :- There is no difference b/w two groups

i.e, $P_1 = P_2$

Alternative Hypothesis [H_1] :- There is difference between 2 groups

i.e, $P_1 \neq P_2$

clearly, it is 2-tailed test.

Level of Significance [α] :- Assume $\alpha = 5\% = 0.05$

Test Statistic [Z_{cal}] :- we know that,

$$Z = \frac{P_1 - P_2}{\sqrt{pq\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} = \frac{0.667 - 0.571}{\sqrt{(0.6)(0.4)\left(\frac{1}{60} + \frac{1}{140}\right)}}$$

$$= 1.27$$

where $p = \dfrac{n_1 P_1 + n_2 P_2}{n_1 + n_2}$

$$= \frac{(60)(0.667) + (140)(0.571)}{60 + 140}$$

$$= 0.6$$

$$q = 1 - 0.6 = 0.4$$

Tabulated Value (Z_{tab}):– From the table at 5% Los

$$Z_{tab} = Z_{\alpha/2} = Z_{\frac{0.05}{2}} = 1.96$$

Conclusion:– $|Z_{cal}| = 1.27$ & $Z_{tab} = 1.96$

$$|Z_{cal}| < Z_{tab}$$

Then, we accept the N.H at 5% level of Significance.

$$\therefore P_1 = P_2$$

prob:– 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 the Same defect. Is the difference between the proportions Significant at 0.05 level of Significance.

Sol:– Given that, $n_1 = 900$ & $n_2 = 1600$ $\alpha = 0.05$

$$x_1 = 20\% \text{ of } 900 = \frac{20}{100} \times 900 = 180$$

$$x_2 = 18.5\% \text{ of } 1600 = \frac{18.5}{100} \times 1600 = 296$$

$$p_1 = \frac{x_1}{n_1} = \frac{180}{900} = 0.2$$

$$p_2 = \frac{x_2}{n_2} = \frac{296}{1600} = 0.185$$

Testing of hypothesis:

Null hypothesis $[H_0]$: There is no difference b/w 2 cities

ie, $P_1 = P_2$

Alternative hypothesis $[H_1]$:- there is a difference b/w 2 cities.

i.e, $P_1 \neq P_2$

Clearly, it is 2-tailed test.

Level of Significance $[\alpha]$:- Given $\alpha = 0.05$

Test Statistic $[Z_{cal}]$:- we know that,

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

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

$$= \frac{(900)(0.2) + (1600)(0.185)}{900 + 1600}$$

$$= 0.1904$$

$$q = 1 - p = 0.8096$$

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

$$= 0.916$$

Tabulated value $[Z_{tab}]$:- from the table at 5% Los

$$Z_{tab} = Z_{\alpha/2} = Z_{\frac{0.05}{2}} = 1.96$$

Conclusion $|Z_{cal}| = 0.916$; $Z_{tab} = 1.96$

$$|Z_{cal}| < Z_{tab}$$

Then we can accept the N.H at 5% Los.

There is no difference b/w two cities.

$$\therefore P_1 = P_2$$

Prob: In a random Sample of 1000 persons from town A, 400 are found to be Consumer of wheat. In a Sample of 800 from town B, 400 are found to be Consumers of wheat. Do these data reveal a Significant difference between town A and town B, So far as the Proportion of wheat Consumers is Concerned?

Sol: Given that,

n_1 = Sample Size of town A = 1000

n_2 = " " " town B = 800

x_1 = No. of Consumers of wheat from town A = 400

x_2 = No. of Consumers of wheat from town B = 400

P_1 = Proportion of Consumers of wheat in town A = $\frac{x_1}{n_1} = \frac{400}{1000} = 0.4$

P_2 = " " " " " " town B = $\frac{x_2}{n_2} = \frac{400}{800} = 0.5$

$$P = \frac{x_1+x_2}{n_1+n_2} = \frac{400+400}{1000+800} = \frac{800}{1800} = \frac{4}{9}$$

$$q = 1-P = 1-\frac{4}{9} = \frac{5}{9}$$

Testing of Hypothesis :-

Null hypothesis $[H_0]$:- $H_0: P_1 = P_2$ ie, there is no difference

Alternative hypothesis $[H_1]$:- $H_1: P_1 \neq P_2$ ie, there is a difference

level of Significance (α) :- Assume $\alpha = 5\% = 0.05$

Test Statistic $[Z_{cal}]$:- $Z = \frac{P_1 - P_2}{\sqrt{pq(1/n_1 + 1/n_2)}} = \frac{0.4-0.5}{\sqrt{4/9 \times \frac{5}{9}\left(\frac{1}{1000}+\frac{1}{800}\right)}} = -4.242$

Tabulated value $[Z_{tab}]$:- from the table at 5% of LOS

$$Z_{tab} = Z_{\alpha/2} = Z_{\frac{0.05}{2}} = 1.96$$

Conclusion :-

$$|Z_{cal}| = \cancel{0.976} \; (|-4.242|) = 4.242 \;\; \& \;\; Z_{tab} = 1.96$$

$$|Z_{cal}| > Z_{tab}$$

then we can accept the A.H at 5% LOS

There is ~~no~~ difference betwen town A and town B, as the Proportion of wheat consumers is concerned.

prob:- In two large ~~proportion~~ populations, there are 30% and 25% respectively of fair haired people. Is this difference likely to be hidden in samples of 1200 and 900 respe- from two populations.

sol:- Given that, $n_1 = 1200$, $n_2 = 900$

Population $P_1 = \frac{30}{100} = 0.3$

$P_2 = \frac{25}{100} = 0.25$

Testing of Hypothesis :-

Null hypothesis $[H_0]$:- There is no difference b/w two populations

ie, $P_1 = P_2$

Alternative hypothesis $[H_1]$:- There is a difference b/w two

Populations. i.e $P_1 \neq P_2$

clearly, it is two-tailed test.

Los (α) :- Assume $(\alpha) = 5\% = 0.05$

Test Statistic $[Z_{cal}]$:- we know that,

$$Z = \frac{P_1 - P_2}{\sqrt{\frac{P_1 Q_1}{n} + \frac{P_2 Q_2}{n_2}}}$$

$Q_1 = 1 - P_1 = 0.7$

$Q_2 = 1 - P_2 = 0.75$

$$Z_{cal} = \frac{0.3 - 0.25}{\sqrt{\frac{(0.3)(0.7)}{1200} + \frac{(0.25)(0.75)}{900}}}$$

$$= 2.56$$

Tabulated Value $[Z_{tab}]$:- from the table value at 5% Los

$Z_{tab} = Z_{\alpha/2} = Z_{\frac{0.05}{2}} = 1.96$

Conclusion :- $|Z_{cal}| = 2.56$ & $Z_{tab} = 1.96$

$|Z_{cal}| > Z_{tab}$

Then, we reject N.H at 5% Los.

we accept alternative hypothesis at 5% Los.

Prob :- A machine produced 20 defective articles in a batch of 400. After overhaulting it produced 10 defectives in a batch of 300. has the machine being improved after overhaulting?

Ans :- Hint :- N.H (H_0) : $P_1 = P_2$ T.S $Z_{cal} = 1.0844$

A.H (H_1) : $P_1 > P_2$ Conclusion Accept N.H at 5% Los.