Unit - V

Tests of Hypothesis

Hypothesis :- The value of parameter whether to accept or reject a statement about parameter then that Statement is called hypothesis.

ex :- 1. The majority of men in the City are smokers

2. A drunk chemist is to decide whether a new drug is really effective in Curing a disease.

Test of Hypothesis :-

The procedure which enables us to decide on the basis of Sample results whether a hypothesis is true or not, is called test of hypothesis.

There are two types of hypothesis

(i) Null hypothesis $[H_0]$

(ii) Alternative hypothesis $[H_1]$

Null Hypothesis (H_0) :- we assume that there is no difference between the procedures.

i.e, No difference between the statistic and the population Parameter. It is denoted by H_0.

$$H_0 : \mu = \mu_0$$

Alternative Hypothesis (H_1) :-

Any hypothesis which contradicts the null hypothesis is called an alternative hypothesis, usually denoted by H_1.

Alternative hypothesis would be

(i) $H_1: \mu \neq \mu_0$ [ie either $\mu > \mu_0$ or $\mu < \mu_0$]

(ii) $H_1: \mu > \mu_0$ [Right tailed]

(iii) $H_1: \mu < \mu_0$ [left tailed]

The alternative hypothesis (i) is known as a two tailed alternative. (ii) is known as right tailed alternative. (iii) is known as left tailed alternative.

→ The setting of alternative hypothesis is very important to decide whether we have to use a single tailed or two tailed test.

Level of Significance :- The level of significance is denoted by α is the confidence with we rejects or accept the null hypothesis H_0. The level of significance is generally specified before a test procedure.

Test Statistic :- There are several test of significance like z, t, f etc. first we have to select the right test depending on the nature of the information given in the problem.

Errors of Sampling :- we decide to accept or to reject the lot after examining a sample from it. As such we have two types of errors.

Type - I Error :- Reject H_0 when it is true.

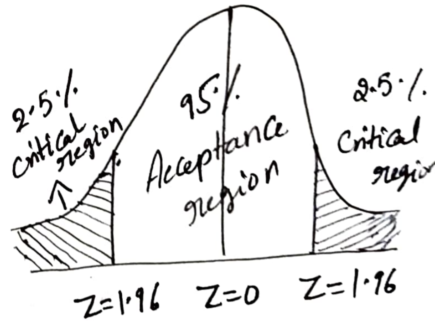
If the null hypothesis H_0 is true but it rejected by test procedure, then the error made is called type I error or α error.

Type - II Error :- Accept H_0, when it is wrong. i.e, Accept H_0 when H_1 is true.

If the null hypothesis is false, but it is accepted by test, then error committed is called type II error or β error.

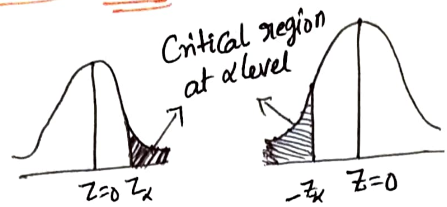
Critical Region (or) Rejection region :-

A Critical region also known as the rejection region, is a set of values for the test statistic for which the null hypothesis is rejected.



i.e, if the observed test statistic is in the critical region then we reject the null hypothesis and accept the alternative hypothesis.

One - Tailed Tests :-



We have to test whether the population mean μ has a specified value μ_0, then the null hypothesis $H_0: \mu = \mu_0$ and the alternative hypothesis may be

$H_1: \mu \neq \mu_0$ $[\mu < \mu_0$ or $\mu > \mu_0]$

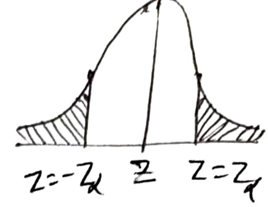
$H_1: \mu > \mu_0$ (Right tailed)
$H_1: \mu < \mu_0$ (left tailed) } one tailed test

The alternative hypothesis in (i) is known as a two-tailed and (ii) & (iii) are known as right tailed and left tailed respe-

In the right tail test $H_1: \mu > \mu_0$ the critical region (or rejection region) $Z > Z_\alpha$ lies entirely in the right tail of the sampling distribution of example mean $\bar{x}$ with area equal to the level of significance α.

Similarly in the left tailed test $H_1: \mu < \mu_0$, the critical region $Z < -Z_\alpha$ lies entirely in the left tail of the sampling distribution $\bar{x}$.

Two tailed test :-



Suppose we want to test null hypothesis $H_0: \mu=\mu_0$ against the alternative hypothesis $H_1: \mu \neq \mu_0$

The critical region under the curve is equally distributed on both side of the mean.

Procedure for Testing of hypothesis :-

Step 1 :- Null Hypothesis $[H_0]$: Define a N.H (H_0) taking into Consideration the nature of the problems and data involved.

Step 2 :- Alternative hypothesis $[H_1]$: Define the alternative hypothesis H_1 so that we could decide whether we should use one-tailed (or) two-tailed test.

Step 3 :- Level of significance (α) select the appropriate level of significance (or) depending on the reliability of estimates. It is not given in the problem usually we choose 5% level of significance.

Step 4 :- Test statistic $[Z_{cal}]$: find the test statistic $[Z_{cal}]$ under the N.H using appropriate formula.

Step 5 :- Tabulated value $[Z_{tab}]$: find the tabulate value of Z at the given level of significance

Step 6 :- Conclusion

Case 1 : $|z|_{cal} < Z_{tab}$ then we accept the null hypothesis at α% level of significance.

Case 2 : $|z|_{cal} > Z_{tab}$ then we accept the alternative hypothesis at α% level of significance. Reject the null hypothesis.

Test of Significance for large Samples :-

Under the large sample tests, we will see four important tests to test the significance.

Method 1 : Test of significance for single proportion.
Method 2 : Test of significance for difference proportion.
Method 3 : Test of significance for single mean.
Method 4 : Test of significance for difference means.

Population :- A population is the collection of objects. Population may be finite or infinite according to the no.of objects in that population. It is denoted by N.

Sample A finite subset of population is called sample. It is denoted by n.

Sample are two types

Small Sample :- If the sample size $n < 30$ is called a small sample.

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 $(n) = 1000$,

$x = 540$,

$\alpha = 1\%$

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

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 :

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

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

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

$= \frac{0.544 - 0.5}{\sqrt{\frac{(0.5)(0.5)}{12.5}}}$

$= 0.9839$

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 $\therefore P \neq 0.2$