UNIT- IV

TESTING OF HYPOTHESIS

Basic Definitions :

Population :

A population is used to refer any collection of individual it may be finite or infinite.

Sample :

A Sample is a small portion selected from the population and the process of drawing a sample from a population is called sampling.

Sample size :

The no. of individual in a selected sample is called the sample size.

Parameter and Statistics :

Any statistical method amputed from population data is known as parameter and any statistical method computed from sample data is known as statistics.

NOTATIONS :

MEASURE		POPULATION ~~SAMPLE~~		SAMPLE
Size	→	N	→	n
Mean	→	μ	→	$\bar{x}$
Standard deviation	→	σ	→	s
proportion	→	P	→	p'
Variance	→	σ^2	→	s^2

Sampling Distribution:

The various value of statistics so obtained may be arrange as a frequency distribution which is known as sampling distribution.

Standard Error:

The standard deviation of samplying distribution of a statistics is known as its standard error, abbreviated as S.E (ie avg. amount of variability from the observation of sampling distribution.)

Statistical Hypothesis :

In attempting to each decision about population on the basis of sample observations, we make assumptions on the basis of sample observations, we make assumptions about population, which are not necessarily true, are called stastical hypothesis.

Null hypothesis:

Null hypothesis is the hypothesis which is tested for possible rejection under the assumption that is true and is denoted by H_0

[(ie) Hypothesis of no difference]

Alternative Hypothesis:

A hypothesis that is complementary to Null hypothesis is called alternative hypothesis and it is denoted by H_1.

A procedure for deciding whether to accept or reject the null ~~tp~~ hypothesis is called the alternative hypothesis and is denoted by H_1.

A procedure for deciding whether to accept or reject the null hypothesis is called the list of hypothesis.

Level of Significance:

A is the probability level below which the null hypothesis is rejected generally. 5% and 1% level of significance are used.

Critical region (or) Region of rejection:

The critical region of a test of statistical hypothesis is that region which leads to ~~region~~ rejection of null hypothesis, H_0. Those region which leads to the acceplance of H_0 is called acceptance region.

Error in Sampling :

Errors are type I , type II errors.

type I error : Reject H_0 when it is ~~error~~ true

type II error : Accept H_0 when it is false.

$$P(\text{Type I error}) = \alpha \text{ \& } P(\text{Type II error}) = \beta$$

One tail or two tail test :

If μ_0 is population parameter & μ is the sample stastistics , then the null hypothesis is given by

$$H_0 : \mu = \mu_0$$

Alternative hypothesis is given by ,

$$H_1 : \mu \neq \mu_0 \text{ (Two - tailed)}$$

$$H_1 : \mu > \mu_0 \text{ (Right tailed)(One tail)}$$

$$H_1 : \mu < \mu_0 \text{ (Left tailed)(one tail)}$$

Procedure for testing a hypothesis :

1) Formulate H_0 and H_1
2) choose the level of significance α
3) compute the test statistic , using the data available
4) Pick out the critical value from the tabulation
5) Conclusion :

Compare the computed value of the test statistic with the critical value at the given level of significance.

Scanned with CamScanner

Sampling

- Large Sample ($n \geq 30$)
  - Test for single mean
  - Test for difference of mean.
  - Test for single Proportion
  - Test for difference proportion;
  - Test for standard Mean.
- Small Samples ($n < 30$)
  - t - Test
    - single mean
    - Difference mean
  - F-Test
  - χ^2 - Test.

Large samples ($n \geq 30$)

Critical Values (or) Significant values:

The sample values of the statistic beyond which the null hypothesis will be rejected are called critical values (or) significant values.

Natures of Test	Level of Significance 1%	5%	10%
Two tailed test (z_α)	2.58	1.96	1.645
One tailed test (z_α)	2.33	1.645	1.28 (Right)
	-2.33	-1.645	-1.28 (Left)

Test of significance of large samples

Test for single mean.

: Null hypothesis $H_0 = \mu = \mu_0$

Test stastistics $z = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}}$ (or) $z = \frac{\bar{x} - \mu}{s/\sqrt{n}}$

Example 1

A sample of 900 members is found to have a mean of 3.4 cm and S.D is 2.61 cm is a sample from a population of Mean 3.25 cm and SD is 2.61 cm If the population is normal & its mean is unknowns. Find the 95% confidential limits of true mean.

Given :

$n = 900$

$\bar{x} = 2.61$ cm

$\mu = 3.25$ cm

$\sigma = 2.61$ cm

Step 1 :

Formulate H_0 & H_1

$H_0 : \mu = 3.25$ cm

$H_1 : \mu \neq 3.25$ cm (Two tailed)

Step 2 :

level of significance $\Rightarrow$ 5% = 0.05

Step 3:

Test the statistics

$$Z = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}} \Rightarrow \frac{3.4 - 3.25}{2.61/\sqrt{900}} = \frac{0.15}{2.61/30}$$

$$= \frac{4.5}{2.61}$$

$$= 1.724$$

Step 4:

Critical value at 5% is $Z_\alpha = 1.96$

Step 5: Conclusion

$|Z| = 1.724 < 1.96 = Z_\alpha$

H_0 is accepted at 5% level of significance.

$\therefore$ The sample is taken from the population

Example 2:

A random sample of 200 employees at a large corporation, should their avg to be 42.8 yrs with a standard deviation of 6.89 yrs. Test the hypothesis $H_0 : \mu = 40$, $H_1 : \mu > 40$ at $\alpha = 0.01$ lvl of significance.

Given:

$n = 200$

$\bar{x} = 42.8$ yrs. $\quad \sigma = 6.89$ yrs.

$\mu = 40.$

Step 1: Formulate H_0 & H_1

$H_0 : \mu = 40$

$H_1 : \mu > 40$ (right one tail) - Test

Step 2: Level of significance $\alpha = 0.01$

Step 3: Test the statistics.

$$z = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}} = \frac{42.8 - 40}{6.89/\sqrt{200}}$$

$$z = 5.747$$

Step 4: Critical Value at 0.01 $\Big\}$ $z_\alpha = 2.33$

Step 5: Conclusion

$$|z| = 5.747 > 2.33 - z_\alpha$$

$\therefore$ H_0 is Rejected at 1% ~~level of~~ level of significanc

$\therefore$ The hypothesis $H_1 : \mu > 40$ is accepted.