Unit-4

Hypothesis Testing / Test of Significance

Some Basic Concept

1. Hypotheses: Statistical hypotheses are hypothesis means any statement or assumption about the statistical population or about the population parameter.
For eg: The ~~average~~ percentage of students securing A grade will be much higher in this semester.

2. Hypotheses testing / test of significance
Hypothesis testing or test of significance is an important aspect of sampling theory which enables us to know on the basis of the sample result ~~either~~ whether (I) the difference between the statistic & population ~~or~~ parameter

2. (ii) whether the difference between the 2 statistics are significant or not.

Statistical Hypothesis

- Null Hypothesis (H_0)
- Alternative Hypothesis (H_1)

3. Null hypothesis - It is a hypothesis of no difference and it is usually denoted by the symbol H_0. According to professor Ronald. A. fisher, null hypothesis is a hypothesis which is tested for possible rejection under the assumption that it is true.

Example: If we want to test whether the population mean (μ), has a specified value (μ_0), then the null hypothesis can be set up as follows

$H_0 : \mu = \mu_0$

4. Alternative hypothesis: A hypothesis which is contradictory to null hypothesis is called alternative hypothesis. It is a hypothesis of difference & it is denoted by the symbol H_1. For example: if we have a null hypothesis of the form $H_0: \mu = \mu_0$, then the corresponding alternative hypothesis could be any one of the following 3 cases

$$H_1: \mu \neq \mu_0$$

or $H_1: \mu > \mu_0$

or $H_1: \mu < \mu_0$

5. One Tailed Test-

In case of hypothesis testing, if we have only one alternative hypothesis, then it is called one tailed or single tailed test. For example,

Let, $H_0: \mu = \mu_0$

against $H_1: \mu > \mu_0$ [Right tailed Test]

or $H_1: \mu < \mu_0$ [Left tailed test]

6/ Two Tailed Test

In case of hypothesis testing, if we have the possibilities of 2 alternatives then it is called 2 tailed test.

Let $H_0 : \mu = \mu_0$

$H_1 : \mu \neq \mu_0$ [Two tailed Test]

7] Critical Region or ~~Acceptance region~~ region of rejection

A region in the Sample Space which amount to the rejection of the null hypothesis is called critical region or region of rejection.

8) Acceptance Region

A region in the Sample Space which amount to acceptance of the null hypo-sis i.e H_0 is called acceptance regio

9) Type ~~One~~ I error & Type II error

Type I error arises when we reject a true null hypothesis. It is usually denoted by α. In terminology

of Statistical Quality Control, Type I error is popularly known as ~~consumers~~ risks ~~producers~~ risks.

Type II error is an error which arises when we accept a wrong null hypothesis. It is usually denoted by 'β'. In the terminology of Statistical Quality Control, Type II error is known as consumers risks.

10/ Test Statistic

Test Statistic in content of hypothesis testing is a phenomenon which value helps us to determine whether the null hypothesis is going to be accepted or rejected.

Ex- If 't' is any statistic, then the test statistic is defined as

$$T = \frac{t - E(t)}{S.E(t)} \sim N(0,1)$$

Level of S

$E U_1^2 = E U_2^2 = \ldots - E U_n^2 = \sigma^2$

or $E U_i^2 = \sigma^2$

11/ Level of Significance { :: same as Type I error

The probability of rejecting a true null hypothesis is called level of significance. Thus, it stands for the size of type I error or the probability of committing type I error. It is denote usually by the symbol 'α'.

12/ General procedure of hypothesis Testing

The steps for hypotheses testing -

Step 1 - Determine or set up the null hypothesis (H_0).

Step-2: Set up the corresponding alternative hypotheses (H_1), which will tell us whether we are using one tailed or two tailed test.

Step-3: Here we will fix the level of significance may be at 1% or 5% level. Sometimes the level of

significance can be fixed at 10% level also.

Step-4: Here under the assumption that the (H_0) is true, we will calculate the following test statistic.

$$T = \frac{t - E(t)}{S.E(t)} \sim N(0, 1)$$

Step-5: After calculating the value of the test statistic it has to be compared with its tabulated value at a particular level of significance. After comparison if we find

$T_c < T_t$, we may always write may while accept accepting the H_0 and conclude that there is no significant difference b/w the statistic and population parameter.

On the other hand if we find $T_c > T_t$, we will reject the H_0 & and conclude that there is significant difference b/w statistic & population parameter.

always use may accept while which accepting H_0 because T_c may vary with level of significant.

Statistical test

Parametric test

Non parametric Test

Large sample test (Z test)

Small sample test or Exact sampling tests

= Application of Large sample test / Z test / T test

Tau

① Z test is used to test the significance of single population mean.

② Z test is used to test the significance of the difference between 2 sample means.

③ Z test is used to test the significant

of a single population proportion.

④ Z test is also used to test the significance of the difference betw 2 sample proportions.

***

However ~~it is to be noted that~~ all the above mention application of Z test are only valid for large ~~test~~ sample i.e $n \geq 30$

Application of T test

① T test is used to test the significance of single population mean, ~~when sample size 'n' < 30~~

② T test is used to test the significance of the differenca betw 2 sample means

3) T test is used to test the significance of simple correlation coefficient, partial correlation coefficient & multiple correlation

4) It is used to the test the significance of regression coefficient.

* It is to be noted that these applications are generaly valid for small samples i.e for $n < 30$.

Application of χ^2 test

(i) Chi-square (χ^2) test is used to test the significance of population variance

(ii) χ^2 test is used as a measure of goodness of fit.

(iii) χ^2 test is used to test the independence of attributes

(iv) It is used to test the homogeneity of correlation coefficient.

However, it is only the 1st application of χ^2 test which falls in the arena of parametric test. The rest 3 applications of χ^2 test are actually non parametric test.

Applications of F test

1. It is used to test the equality of 2 population varieances
2. It is used to test the overall significance of a regression model
3. F test is used to test the significance of sample correlation ratio.
4. F test is used to test the linearity of a regression model.
5. It is also used for testing the equality of several sample means

Table of critical or tabulated values of ~~the~~ for Z test in different circumstances

Critical or tabulated values (Z_t)	Level of significance		
	1%	5%	10%
Two tailed test	2.58	1.96	1.645
Right tailed test	2.33	1.645	1.28
~~Left~~ Left tailed test	-2.33	-1.645	-1.28