HYPOTHESIS TESTING

Definition:

"*A hypothesis is a quantitative statement about a population. It may or may not be true. By testing a hypothesis we can find out whether it deserves acceptance or rejection".*

Some important terms used in hypothesis testing are discussed below before going into the details of the testing hypothesis.

1. **Test of significance**: the test of hypothesis is based on the test of significance enables us to decide based on sample results. If i) the deviation between the observed samples statistic and hypothetical parameter value or (ii) the deviation between two sample statistics is significant or might be attributed due to sampling fluctuations.
   It is well known that if the sample size is large almost all distribution can be approximated by the normal distribution. So for large samples, one can safely use the normal test of significance and form small samples t-test, F-test, etc.
2. **Null hypothesis**: for applying the test of significance one has to set up a hypothesis. i.e., a definite statement about the population parameter. Such hypothesis is a hypothesis of no difference. This hypothesis is known as Null hypothesis and denoted by H_0.
   Definition: *Null hypothesis is the hypothesis, which is tested for possible rejection under the assumption that it is true. It is usually the case that the null hypothesis is either accepted or rejected basing on the test of significance*.
3. **Alternative hypothesis**: any hypothesis, which is complementary to null hypothesis is called alternative hypothesis. This is usually denoted by H_1 or H_A. Suppose we are interested to verify that for a population that mean μ0 or not. Then we set up the null hypothesis –

   $$H_0 = \mu = \mu_0$$

   Then the alternative hypothesis could be any one of the statements

   (i) $H_1 = \mu \neq \mu_0$ (i.e. $\mu > \mu_0$ or $\mu < \mu_0$)

   (ii) $H_1 = \mu > \mu_0$

   (iii) $H_1 = \mu < \mu_0$

   The alternative hypothesis in (i) is known as two tailed alternative, whereas the other two (ii) and (iii) are known as right and left tailed alternatives respectively. In general they are called as one tailed alternatives.

   The setting of alternative hypothesis is very important since on the basis of alternative hypothesis only, we decide whether a one tailed (right or left) or two tailed test to be used.
4. **Errors in testing of hypothesis**:

After applying a test, a decision is taken about the acceptance and rejection of null hypothesis vis-a-vis the alternative hypothesis. There is always some possibility of committing an error in taking a decision about the null hypothesis. These errors are of two types.

Type I error: Reject the null hypothesis when it is true (H_0)

Type II error: accept the null hypothesis when it is wrong (H_0)

Now we write –

Prob. [rejecting H_0 when it is true] = α

Prob. [Accepting H_0 when it is wrong] = β

Here the α and β are called the size of the type I and type II errors respectively.

5. **Level of Significance**: it is the quantity of risk of the type – I error, which can be readily tolerated in making a decision about the null hypothesis H0. In other words, it is the probability of type I error (α) is tolerable. The level of significance is denoted by α and conveniently chosen as 0.05 or 0.01. $\alpha = 0.01$ is used for high precision and $\alpha = 0.05$ for moderate precision.
6. **Critical region**: a statistic is used to test the hypothesis H_0. The test statistic follows a known as distribution. In a test, the area under the probability density curve is divided into two regions. i.e., the region of acceptance and the region of rejection. The region of rejection is the region in which H0 is rejected. It indicates that if the value of test statistic lies in this region, H_0 will be rejected. This region is called critical region. The area of the critical region is equal to the level of significance α. The critical region is always on the tail of the distribution curve. It may be on both sides or on one tail (right or left), depending upon the alternative hypothesis.
7. **Degrees of Freedom**: since the sample is a part of the population with N number of items, it is desirable to get an estimate of the population parameter from the sample statistic. It becomes possible to draw an estimate of the various population parameters from the sample statistic by attaching a restriction to the n number of sample observations. Thus, the formula for sample variance $s^2 = \Sigma(x - \bar{x})^2 /n-1$ which is an appropriate estimate of the population variance $\sigma 2$. Hence in all calculations of the estimates of population variance one should divide the appropriate value by n-1 instead of n. However, in the samples of large size where n is (say 1000) estimate of population parameter value will show little or negligible difference if the division is made either by n or n-1. Hence, in all calculations of sample variance or sample standard deviation where the sample size is less than 30, the division should be by n-1.

   When the number of restrictions is subtracted from the total number of items, we get the degree of freedom.

Procedures for Hypothesis testing:

1) The first step is to set up a hypothesis called Null hypothesis (H_0).
2) The complement of the null hypothesis is Alternative hypothesis (H_1). As against the null hypothesis assumes that the difference is not due to sampling fluctuations, but they are real.
3) The third step is to set up a level of significance (α) depending on reliability of the estimates and permissible risks at which the hypothesis is to be tested. This is to be decided before samples is drawn i.e. α is fixed in advance. It means the level of confidence with which a particular hypothesis is to be accepted or rejected. The level of confidence determines the probability of our being right or wrong in accepting or rejecting the hypothesis. Usually the α value is fixed in 0.01 or 0.05 or 0.1 i.e. same thing as setting confidence coefficient in 99% or 95% or 90% respectively.
4) The fourth step is to decide the test statistic. It is a statistic computed from a sample. It is generally based on some probability distribution from a sample. It is generally based on some probability distribution like Z, t, χ^2 or F etc.
5) In the previous step we get the computed value of test results (example Z), and let us denote that Z_c. Noiw we compare the value of Z_c with the significant value of Z_α at a given level of significance 'α'. (Z_α is called as tabulate value of α at percent of level of significance). If $|Z_c| \leq |Z\alpha|$ i.e. if the calculated value of modulus value of Z is less than Z_α we say that it is not significant and vice versa. Then we say that it is significant and the null hypothesis H_0 is rejected at level of significance α or confidence coefficient (1-α). The acceptance or rejection of the hypothesis at the predetermined level of significance depends on the region, where the test statistic falls.

Z TEST / NORMAL DISTRIBUTION TEST FOR LARGE SAMPLE TEST.

When the number of samples exceed 30, then it is regarded as a large sample let us discuss each of the cases one by one.

A-1 Test of significance for single mean:

H0: $\bar{x} = \mu$

H1: $\bar{x} \neq \mu$ (or $\bar{x} > \mu$ or $\bar{x} < \mu$)

Test statistic

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

Where, $\bar{x}$ = sample mean; μ = population mean; σ = population standard deviation and n = size of the sample

Remark: if σ is unknown, then we can use its estimate s, which wil be calucated from the sample as

$$S = \sqrt{[\left\{\frac{1}{n-1}\right\} x\ \Sigma(x - \bar{x})^2]}$$

Example 1:

The following data were obtained regarding the weights of 1000 number of certain species of fish. Calculated the standard error of the mean. If the averages of the population were 42, what conclusion can you arrive at about the reliability of the sample?

Weight in gram	No. of fishes
0 -10	50
10-20	100
20-30	150
30-40	200
40-50	200
50-60	100
60-70	100
70- 80	100

Solution:

The table prepared below is self-explanatory

Weight in gram	No. of fishes (f)	Class mid value (m)	fm	m^2	fm^2
0 -10	50	5	250	25	1250
10-20	100	15	1500	225	22500
20-30	150	25	3750	625	93750
30-40	200	35	7000	1225	245000
40-50	200	45	9000	2025	405000
50-60	100	55	5500	3025	302500
60-70	100	65	6500	4225	422500
70- 80	100	75	7500	5625	562500
Total	**Σf = 1000**		**Σfm = 41000**		**Σfm^2 = 2055000**

It is a case of large samples because n>30
Mean = 41000/1000 = 41
Variance = [2055000 – (41000^2/1000)] / 1000 = 374
standard deviation = s = 19.34
Hence the standard error of mean = $19.34/\sqrt{1000}$ = 0.61

Step 1:
Null hypothesis: H_0 : the sample has been drawn from the population with mean 42 (μ= 42)
Alternative hypothesis (H_1): the sample not drawn from the mean population mean 42 (μ≠42)

Step 2: level of significance α = 0.05
Step 3: as the sample size (n) > 30, we apply the Z – test which is based on normal distribution.
Computation of Test statistic:

$$\mathbf{Z} = \frac{|\bar{x} - \mu|}{\sigma/\sqrt{n}}$$

$$= \frac{|41 - 42|}{0.61}$$

$$\mathbf{Z_c} = \mathbf{1.64}$$

Step 4: Decision
At the 5% level of significance (α =0.05), the critical value Z_α = 1.963 (read from the table). But the computed value is 1.64 is less than the critical or tabulated value. Hence the computed value falls under the acceptance region.

Step 5: Conclusion
Thus the null hypothesis at 5% level of significance may be accepted. So the sample drawn might have come from the population with mean 42. And thus the conclusion can be made that the sample drawn is reliable.