Testing of Statistical Hypothesis

Definition:

It is an assumption or statement (may or may not be true) concerning one or more populations.

When the evidence from the sample is inconsistent with the stated hypothesis, then there is a **rejection** of hypothesis, otherwise there is an **acceptance** of hypothesis. However, acceptance does not necessarily mean that the hypothesis is true. It means that there is not sufficient evidence from the samples drawn to reject it. On the other hand, rejection of a hypothesis means that there exists evidence that it is false. Thus, statisticians generally hope to reject a formulated hypothesis. The procedure which enables us to decide whether a certain hypothesis is true or not, is called **Test of Significance** or **Test of Hypothesis**.

Types:

(i) Null Hypothesis: The hypothesis that one formulates with the hope of rejection is called **null hypothesis (H_0)** under the assumption that it is true.

This is the statistical hypothesis which is actually tested for rejection or acceptance.

(ii) Alternative Hypothesis: Any hypothesis, which is complementary to null hypothesis, is called **alternative hypothesis (H_1)**

This is never tested, but its acceptance (rejection) depends upon the rejection (acceptance) of null hypothesis.

(iii) Simple Hypothesis: A statistical hypothesis which completely specifies the population is called **simple hypothesis**.

Eg: H: $\mu = 0$, $\sigma = 1$ in the case of standard normal distribution

(iv) <u>Composite Hypothesis:</u> A statistical hypothesis that does not completely specify the population is called **composite hypothesis**

Eg: H: $\mu = 0$, σ is unknown

Statistical Decision:

The decisions or conclusions about the population parameters on the basis of a random sample drawn from the population are called statistical decisions.

Test Statistic:

A function of sample observations whose computed value determines the final decision regarding acceptance or rejection of H_0, is called a **test statistic**.

There are four main test statistics:

(a) standard normal distribution : z
(b) chi square distribution : χ^2
(c) students' t distribution : t
(d) Snedecor's F distribution : F

Errors:

Null Hypothesis (H_0)	Accept H_0	Reject H_0
H_0 is true	correct decision	type-I error
H_0 is false	type-II error	correct decision

Maximum probability of type-I error is denoted by : α

Maximum probability of type-II error is denoted by : β

<u>Note:</u> (i) Acceptance of a false statement is more risky than rejection of a correct statement. Hence, type-II errors are to be minimized after fixing the type-I error.

(ii) Power of a test is given by : $1 - \beta$

Level of Significance (α):

In any test, both type I and type II errors cannot be minimized simultaneously as they are inter-connected. Therefore, generally type-I error is fixed and then the probability of type-II error is minimized. The maximum probability with which one is ready to risk a type-I error is called the **level of significance** of the test and it is denoted by α.

Generally, $\alpha = 0.05$ is taken to conclude decisions (for higher precision, $\alpha = 0.01$ is used). We say that the null hypothesis is rejected at 5% level of significance if our computed probability is found to be less that $\alpha = 0.05$. This means that there are 5 chances out of 100 that the null hypothesis is rejected when it is true. That is, we are 95% confident that our decision is correct. This is similar to say that the probability that we could be wrong is 0.05.

Level of Confidence:

This is complementary to the level of significance. It is given by $(1 - \alpha)$. If the level of significance is 1%, then the **level of confidence** will be 99%.

Degree of Freedom:

The number of independent observations of the sample is called the **degrees of freedom,** $v = n – k$, $v > 0$

= sample size – number of independent constraints imposed on the observations in the sample

Eg: (a) Choose 5 observations. Here $n = 5$, k = number of restrictions on these observations = 0. Hence, d.o.f = 5 – 0 = 5.

(b) Choose 5 observations such that the total sum is 100. Here $n = 5$ but a restriction on their total is given to be 100. This means $k = 1$. So d.o.f = 5 – 1 = 4.

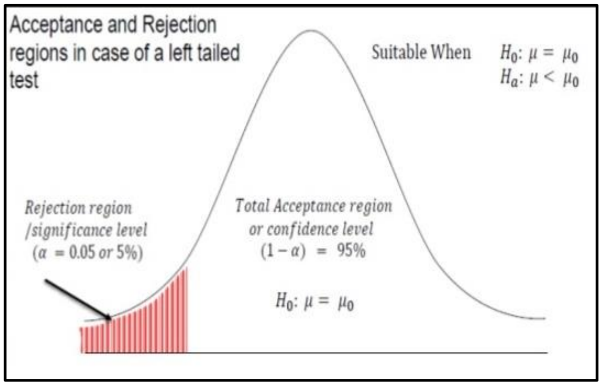
Critical Region:

The set of values of the test statistic which lead to the rejection of the null hypothesis is called the **critical region** or the **rejection region** of the test.

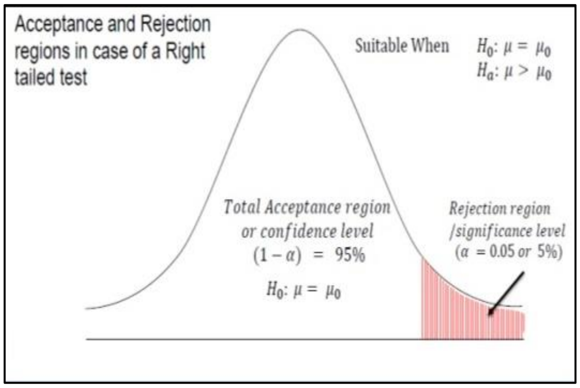
The probability with which a true null hypothesis is rejected by the test, i.e, α, is referred to the **size of the critical region**. This region generally lies on one or both sides of the tails of the distribution. The value of the test statistic, which separates the critical region from the acceptance region is called the **critical value**.

One-tailed and Two-tailed Tests:

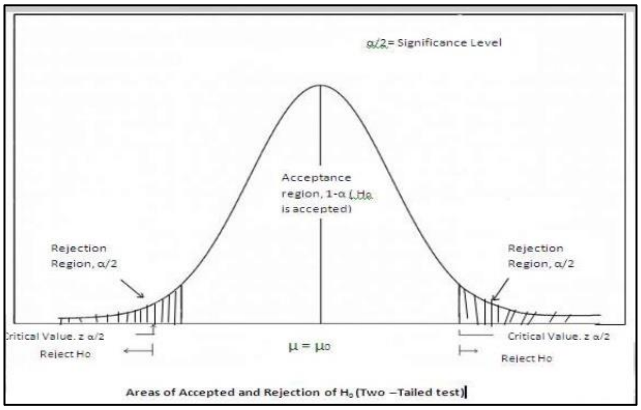
Suppose that under a given hypothesis, the sampling distribution of the test statistic 't' is a normal distribution with mean μ_t and standard deviation σ_t. Then $z = \frac{t-\mu_t}{\sigma_t}$ is a standard normal variable with mean 0 and standard deviation 1. Now, the critical value of a test statistic depends upon the nature of H_1. If H_1 carries a greater than sign (>), then the test hypothesis is known as **right one-tailed test.** If H_1 carries a less than sign (<), then it is called a **left one-tailed test**. If H_1 carries a not equal to sign (≠), then it is called a **two-tailed test**.



Rejection / Critical region: z ≤ - z_α (area: α)



Rejection / Critical region: z ≥ z_α (area: α)



Rejection / Critical region: z ≥ $z_{\alpha/2}$ U z ≤ - $z_{\alpha/2}$ (area: α)

Rules of Decision:

1. Formulate null and alternative hypothesis. Generally H_0 specifies some value of the parameters involved in the population : $H_0 : \theta = \theta_0$. Then H_1 will be either of : $H_1 : \theta \neq \theta_0$ (two-tailed), $H_1 : \theta > \theta_0$ (right one-tailed) and $H_1 : \theta < \theta_0$ (left one-tailed)

2. State the appropriate test statistic 't' and also its sampling distribution, when H_0 is true. In large sample tests, the statistic $z = \frac{t-\theta_0}{S.E.}$ is mostly used, which approximates a standard normal distribution. In small sample tests, the population is taken as normal and various test statistics are used, which may follow standard normal, chi square, t or F distributions

3. Choose the level of significance α of the test. If this is not given in the problem, then 5% level of significance can be used

4. Determine degrees of freedom (for t-test, F-test and chi square tests)

5. Determine the critical region

6. Compute the test statistic from the sample data

7. Decision or conclusion is made : If the value of the test statistic falls in the critical region, then the null hypothesis H_0 is rejected and H_1 is accepted

Central Limit Theorem:

If random samples of sizes n are drawn from an underlying non-normal population having mean μ and standard deviation σ, then the sampling distribution of mean $\bar{x}$ is approximately distributed with mean $\mu_{\bar{x}} = \mu$ and standard deviation $\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$, provided $n \geq 30$ (large). Hence the variable $z = \frac{\bar{x}-\mu}{\frac{\sigma}{\sqrt{n}}}$ will have a standard normal distribution with mean 0 and standard deviation 1.