Module 6: Tests of Hypothesis

Contents

- Null and Alternative hypotheses
- Types of errors
- Neyman–Pearson lemma
- Most Powerful (MP) and Uniformly Most Powerful (UMP) tests.

Introduction to Hypothesis Testing

- Hypothesis testing is one of the most important concepts in statistics.
- Heavily used by **Statisticians**, **Machine Learning Engineers**, and **Data Scientists.**
- Statistical tests are used to check whether a **null hypothesis** is **rejected** or **not rejected (accepted)**.
- Statistical tests assume a null hypothesis **of no relationship or no difference between groups.**

- **Definition**

A hypothesis is defined as a formal statement, which gives the explanation about the relationship between two or more variables of a specified population.

Example: *Based on sample data, we may wish to decide whether a serum is really effective in curing Corona.*

Types of hypothesis

- Simple
- Complex
- Null
- Alternative
- Empirical
- Statistical

What is test of hypothesis?

- Assume that a particular hypothesis is true, we find that results observed in a random sample differ markedly from those expected. We say that observed differences are significant and we reject the hypothesis.
- Procedures that enable us to decide to accept or reject hypothesis are called **test of hypothesis, test of significance, decision rules.**

Type I & Type II errors

- **Type I error:** *Rejecting a hypothesis when it happens to be true.*
- **Type II error:** *Accepting a hypothesis when it is to be rejected.*
- These errors have to be minimized, but the decrease in one causes the increase in the other.
- The best solution is to increase the sample size.

Characteristics of hypothesis

The important characteristics of hypothesis are:

- It should be short and precise.
- It should be specific.
- It must be related to the existing body of knowledge.
- It should be capable of verification.

Statistical hypothesis

- It is a guess or assumption about the parameters of population distribution.
- It is established beforehand and may or may not be true.
- Statistical hypothesis can be either
  - Null hypothesis (H_0)
  - Alternative hypothesis (H_1) or (H_a)

Null Hypothesis (H_0)

- It is a statistical hypothesis which is to be actually tested for acceptance or rejection.
- It is a hypothesis which is tested for possible rejection under the assumption that it is true.
- It is expressed in the form of equality.
- **Example:** *Independent variables have no effect on the dependent variables*.

Examples of Null Hypothesis (H_0)

- Null hypothesis is always a simple hypothesis stated as an equality specifying an exact value of the parameter.
- **Examples:**
  - *Population mean equals to a specified constant μ_0*
  - *The difference between the sample means equals to a constant.*

Alternate Hypothesis (H_1) or (H_a)

- It is any other hypothesis other than null hypothesis (H_0)
- It is expressed in the form of >,<, not =
- We can accept alternative hypothesis if there is sufficient evidence.
- This was originated by Neyman.
- **Example:** *Independent events or variables have effect on dependent variables.*
- H_1: $\mu > \mu_0$

Critical Region

- In any test of hypothesis, a test statistic S*, calculated from the sample data, is used to **<u>accept</u>** or **<u>reject</u>** null hypothesis of the test.
- The area under the probability curve of the sampling distribution of the test statistic S* which follows some known given distributions.
- This area under probability curve is divided into two regions, **region of rejection** where null hypothesis is rejected and **region of acceptance.**

- The **critical region** is the region of rejection of null hypothesis.
- The area of critical region equals to the level of significance (α).
- Critical region lies on the tail(s) of the distribution.
- Depending upon the nature of alternate hypothesis (H_a), critical region may lie on one side or both sides of the tail(s).

Test of Significance

- This is the procedure to decide whether to **accept** or **reject null hypothesis (H_0)**.
- This test is used to determine whether observed samples differ significantly from expected results.
- **Acceptance** of hypothesis merely indicates that the **data did not give sufficient evidence to reject the hypothesis.**

- However **rejection** of hypothesis is **a firm conclusion that the sample evidence rejects it.**
- **When null hypothesis (H_0) is accepted, the result is said to be non-significant,** which means the observed differences are due to chance caused by the process of sampling.
- **When null hypothesis (H_0) is rejected, which means the alternate hypothesis (H_1) is accepted and the result is said to be significant.**
- Since the test is based on sample observation, the decision of acceptance or rejection of null hypothesis is subject to some error or risk.

Level of Significance (α)

- Represented by **α**.
- This is the probability of committing **Type I error**.
- It measures the amount of risks associated in taking decisions.
- This factor has to be chosen before sample information is collected.
- It is either **0.01** or **0.05**.

How to compute the level of significance?

- To measure the level of statistical significance of the result, the investigator first needs to calculate **p-value.**
- It defines the probability of identifying an effect which provides that null hypothesis (H_0) is true.

When p-value is less than the level of significance (α), the null hypothesis is rejected.

Interpretation of p-value based on level of significance (10%)

- If $p > 0.1$, then there will be no assumption for null hypothesis
- If $p > 0.05$ and $p \leq 0.1$, it means that there will be a low assumption for null hypothesis.
- If $p > 0.01$ and $p \leq 0.05$, then there must be a strong assumption about null hypothesis.
- If $p \leq 0.01$, then a very strong assumption about null hypothesis is indicated.

Rejection rule of Null Hypothesis (H_0)

- If $p < \alpha$, then one must reject null hypothesis
- If $p > \alpha$, then one should not reject (i.e., accept) null hypothesis.

Power of test

- α probability of committing Type I error

$= P(\text{reject } H_0/H_1)$

and β probability of committing Type II error

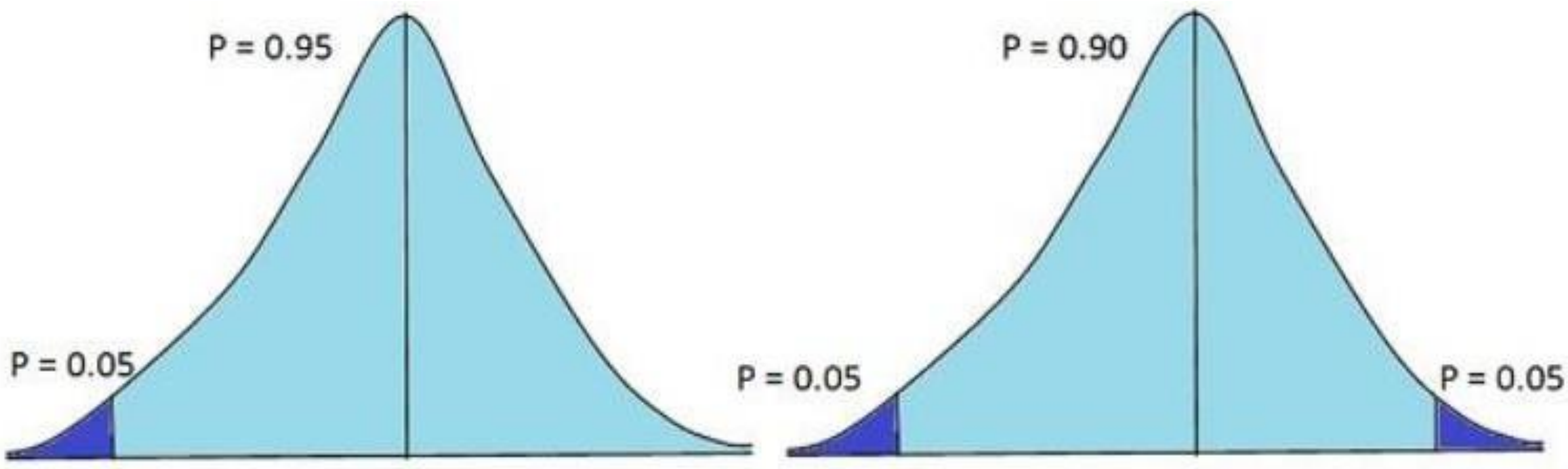
$= P(\text{accept } H_0/H_1)$

Power of test = (1-β)

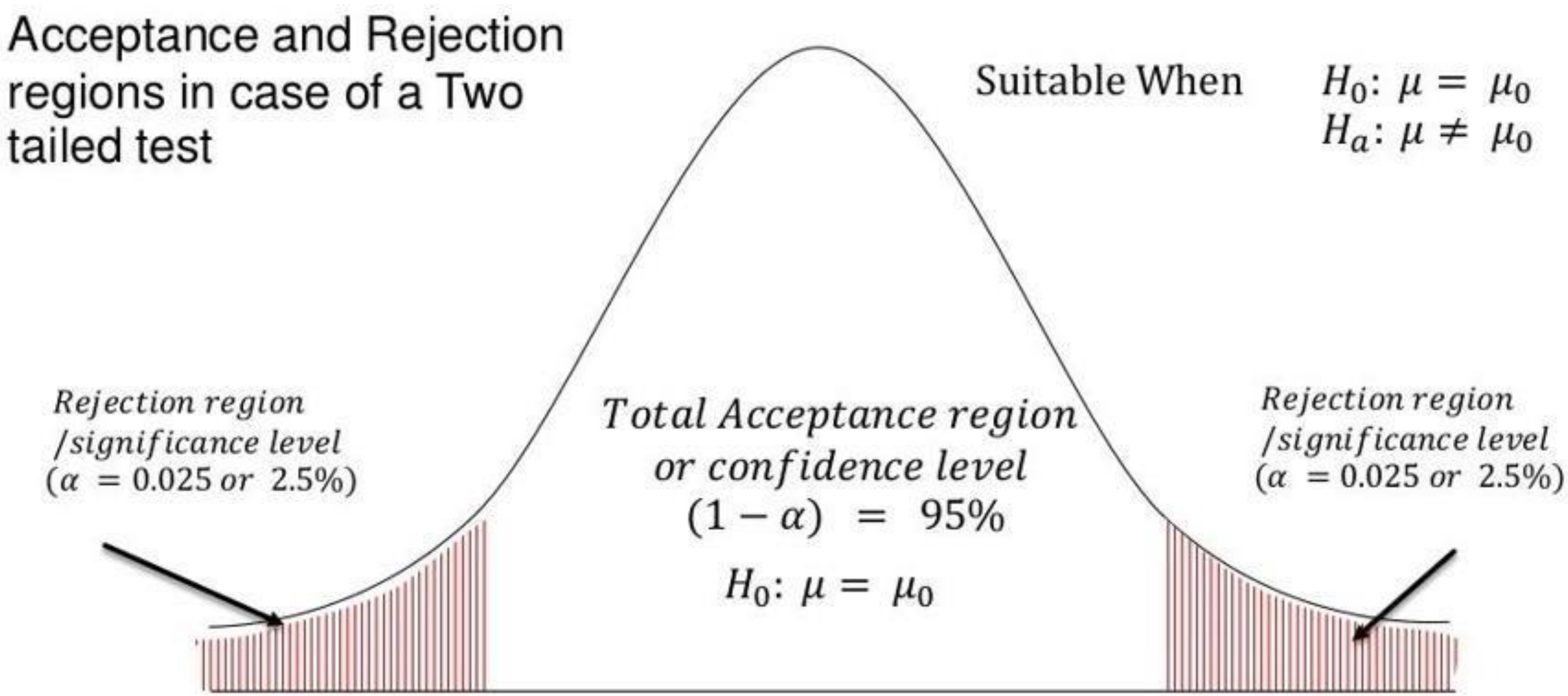
Critical values or significant values

- It is the value of test statistic S_a* which separates the area under the probability curve into critical region and non-critical region.
- ***Note: Critical region is the rejection region, non-critical region is the acceptance region.***

One-tailed test and two-tailed tests



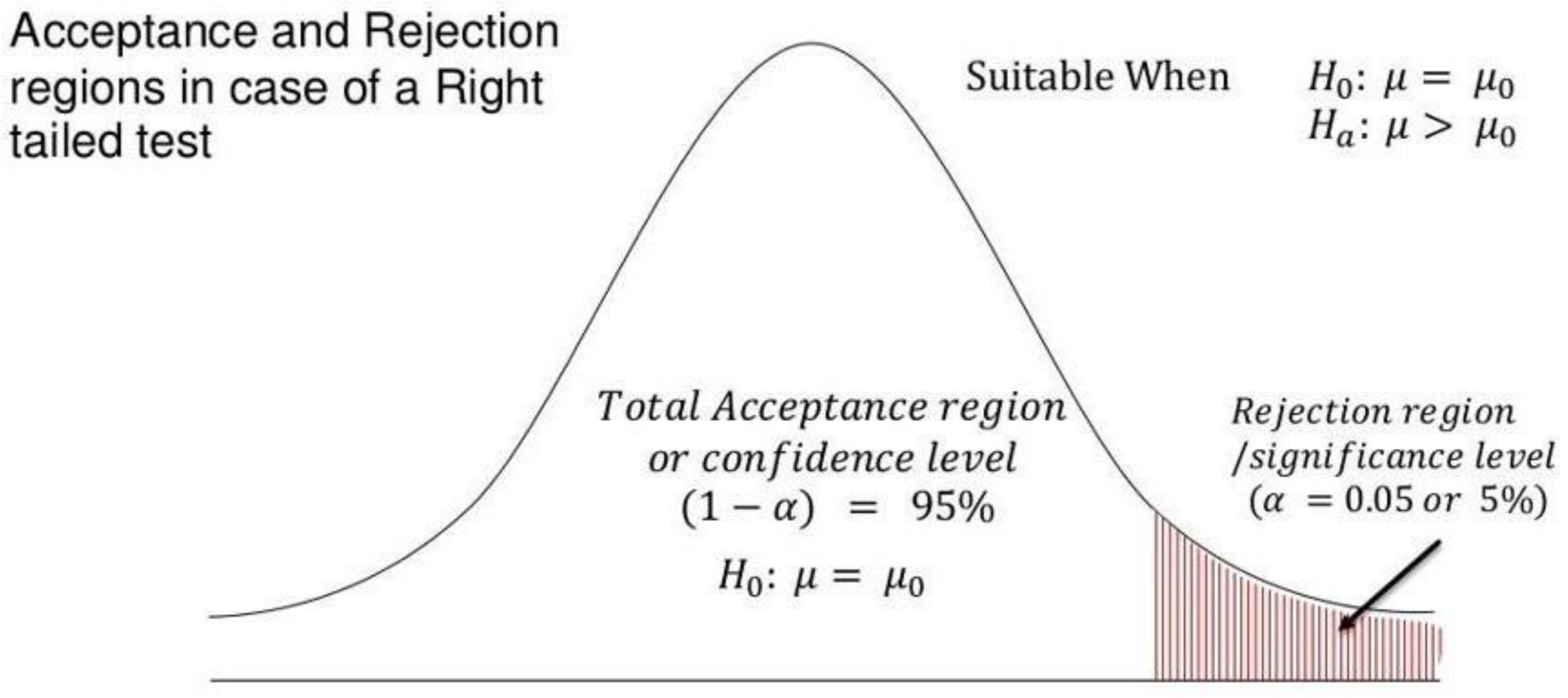
Two-tailed test



One-tailed test

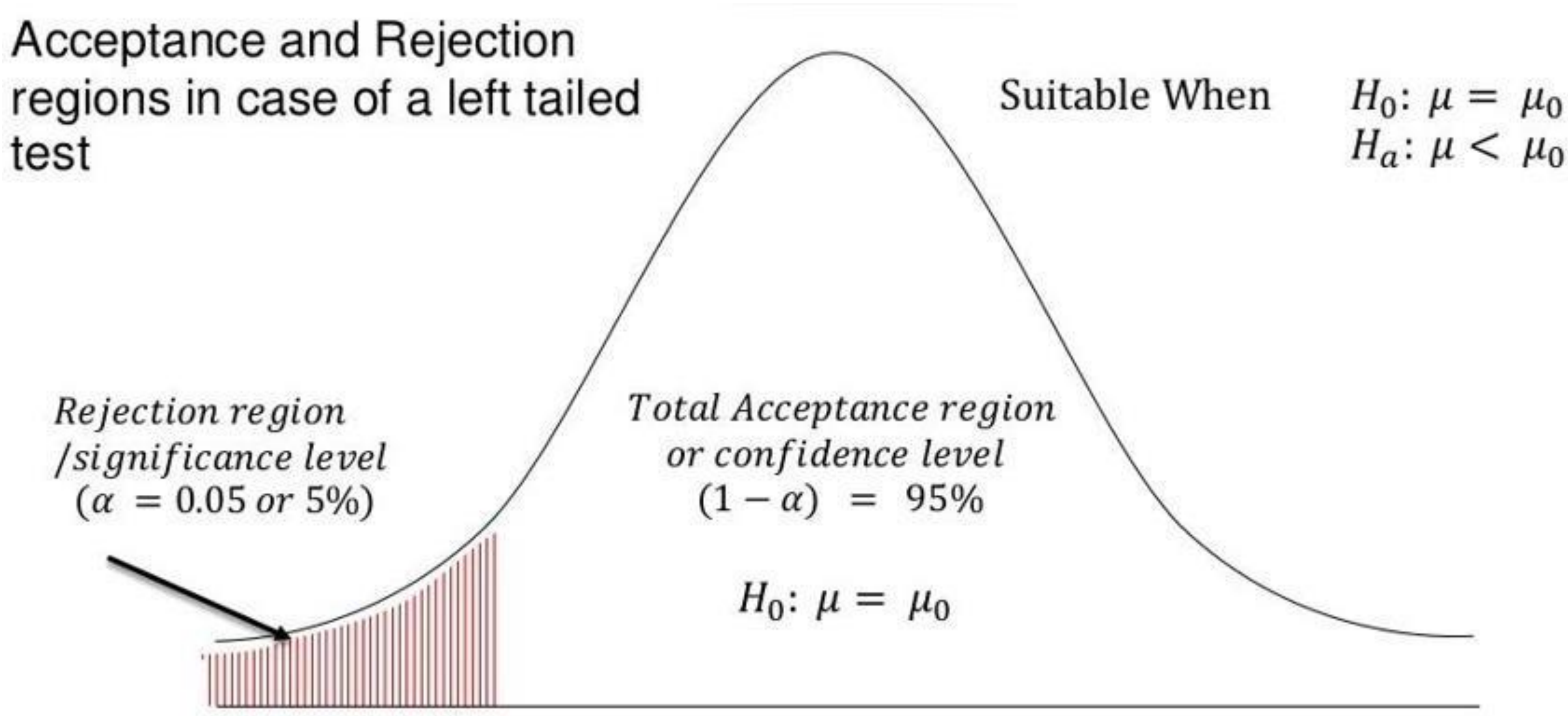
- One-tailed test can be right one-tailed test and left one-tailed test.
- When the alternative hypothesis is of the greater than type H_a: $\mu>\mu_0$, then the entire critical region of area α lies on the right side of the curve.

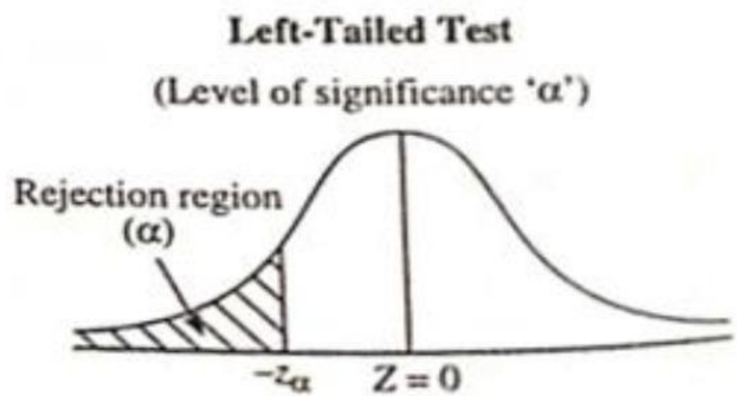
One-sided (right tailed) test



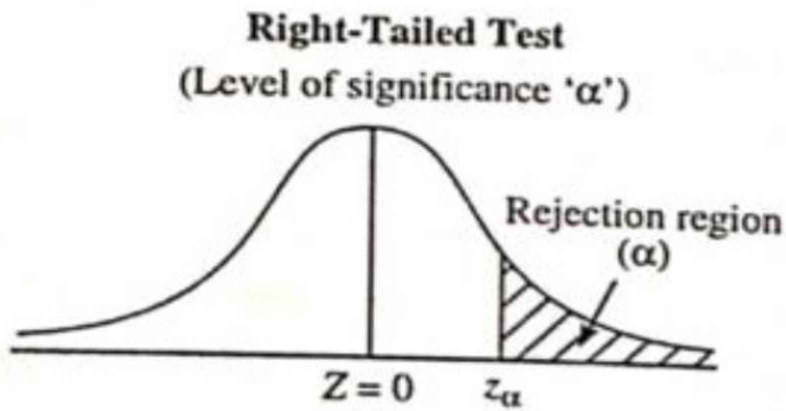
One-sided (left tailed) test

- When the alternative hypothesis is of the less than type H_a: $\mu_1 < \mu_0$, then the entire critical region of area α lies on the left side of the curve.





For Left-tailed Test : $P(Z < -z_\alpha) = \alpha$



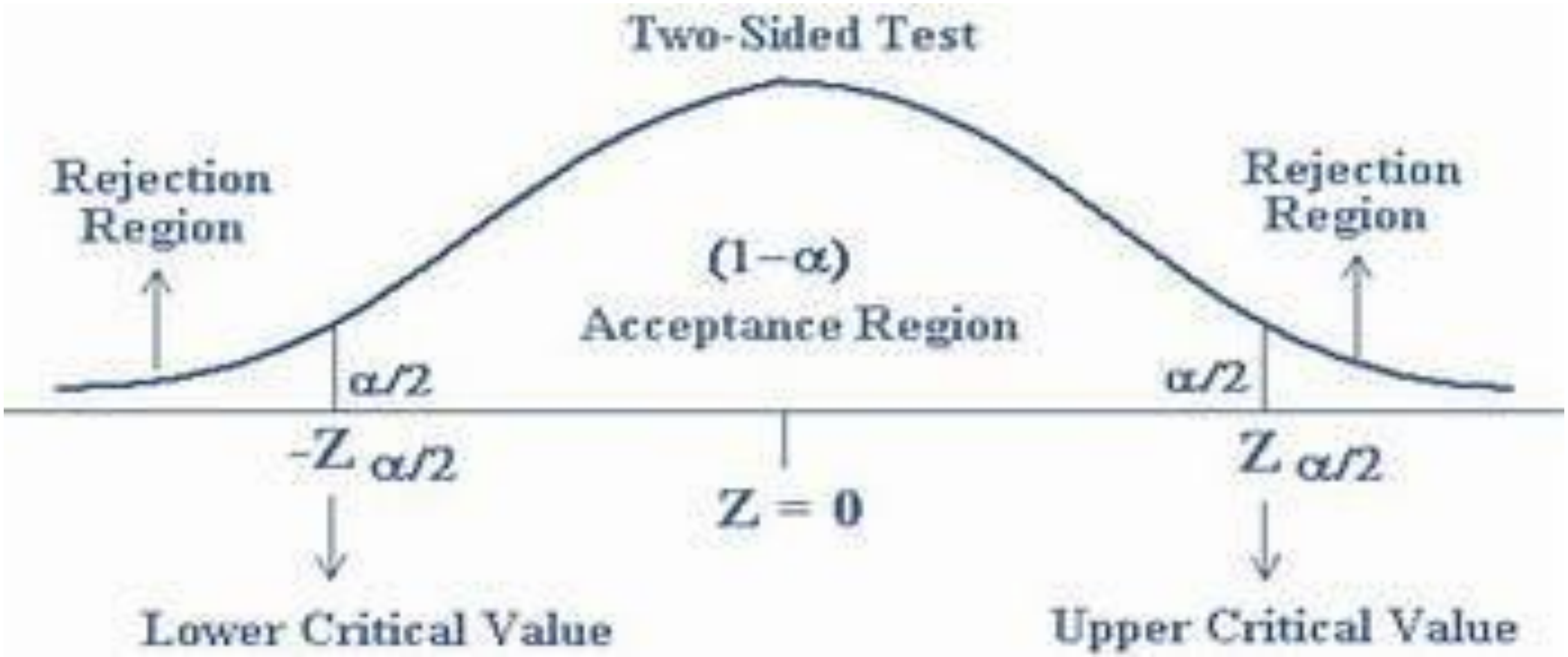
For Right-tailed Test : $P(Z > z_\alpha) = \alpha$

- One-tailed test of hypothesis is used when one talks of **type I error.**
- A hypothesis test is also called as **one-sided test** and is designed to identify a difference from a hypothesized value in only one direction.
- It is also called directional test, because it includes the directional prediction in the statement of hypothesis and the location of the critical region.

- The critical region for a one-sided test is the set of values less than the critical value of the test or the set of values greater than the critical value of the test
- A one-tailed test is one where H_1 is directional and includes < or >
- A one-tailed test looks for an increase or decrease in the parameter.
- If we **reject the null hypothesis** at 5% level of significance, then there is significant evidence to reject the hypothesis at 5% level.

Two Tailed Test

- If alternative hypothesis is of the not equals type i.e., $H_1: \mu_1 \neq \mu_2$
- The critical region lies on both sides of the right and left tails of the curve such that the critical region of area $\alpha/2$ lies on the right tail and critical region of area $\alpha/2$ lies on the left tail.



- A two-tailed test is one where H_1 has no direction.
- The values for which we can reject the hypothesis are located in both tails of the probability distribution

Comparison Chart

BASIS OF COMPARISON	ONE-TAILED TEST	TWO-TAILED TEST
Meaning	A statistical hypothesis test in which alternative hypothesis has only one end, is known as one tailed test.	A significance test in which alternative hypothesis has two ends, is called two-tailed test.
Hypothesis	Directional	Non-directional
Region of rejection	Either left or right	Both left and right
Determines	If there is a relationship between variables in single direction.	If there is a relationship between variables in either direction.
Result	Greater or less than certain value.	Greater or less than certain range of values.
Sign in alternative hypothesis	$>$ or $<$	$\neq$

Steps for test of hypothesis

- Formulate Null Hypothesis (H_0)
- Formulate Alternative hypothesis (H_1)
- Choose level of significance α
- Critical region (CR) is determined by the critical value S_a* and the kind of alternate hypothesis.
- Compute the test statistic S* using the sample data.
- Decision: Accept or reject H_0 depending on the relation between S* and S_a*

Quick revision of Population and Sample

- **Population:** A set or collection or totality of objects under study. Size of population *N* is the number of objects in the population. Parameters are *mean*, *median*, *variance*.

- **Sample:** Finite subset of population. Size is *n*. Sample should be a representative of the general population. Done by random sampling.