Hypothesis

In Statistics, a hypothesis is defined as a formal statement, which gives the explanation about the relationship between the two or more variables of the specified population. It helps the researcher to translate the given problem to a clear explanation for the outcome of the study. It clearly explains and predicts the expected outcome.

Types- 1) Null hypothesis 2) Alternative hypothesis

*__Null hypothesis__

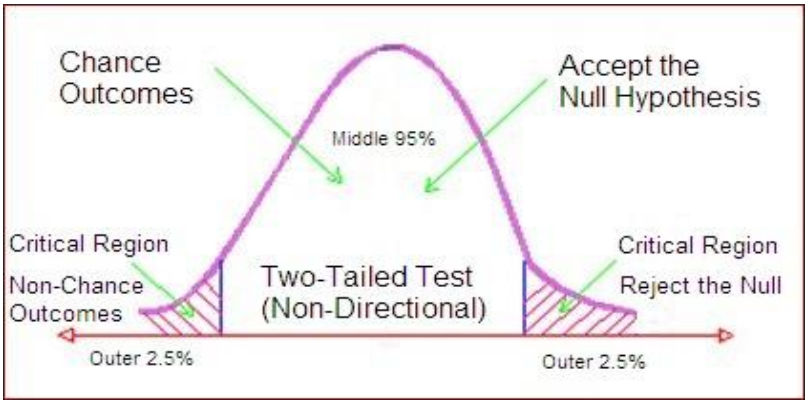
The hypothesis which is tested for possible rejection under the assumption that it is true. This hypothesis is denoted by H_0.

__Alternative hypothesis__

Any hypothesis which is mutually exclusive and complementary to the null hypothesis is called an alternative hypothesis. This hypothesis is denoted by either H_a or by H_1.

Critical region

The region of the sample space which specifies the values of the test statistic for which the null hypothesis H_0 is rejected is called critical region.



Type I and Type II errors

***Type I error**: The rejection of null hypothesis, when it is true, is called type I error. The probability of a type I error is denoted by α. This α is also known as size of the critical region.

α=P (type I error)

=P (reject H_0 when H_0 is true)

*__Type II error__

If false null hypothesis H_0 is accepted, it is said to be type II error. The probability of this type of error is denoted by β.

β=P (type II error)

= P (accept H_0 when H_0 is false)

Action or Decision	State of Nature	
	H_0 is true	H_0 is false
Reject H_0	Type I error	Correct decision
Accept H_0	Correct decision	Type II error

It is noted that as the accepting of rotten egg is more harmful than the rejecting of a good egg, the degree of impact of the type II error is more harmful or dangerous than that of the type I error.

Power of a test

The probability of rejecting null hypothesis H_0, when it is false, is called power of a test. That is

Power of a test= P (reject H_0 when H_0 is false)

= 1- P (accept H_0 when H_0 is false)

=1- P (type II error)

=1-β

The power of a test will be increased when β is minimized to a very small value in the test. But for any fixed or given sample size, α & β are inversely related in any test i.e. when α is reduced, β will be increased and when β is minimized, α will be increased. Therefore, the simultaneous reduction of α & β to very small values is impossible, in practice. The only method to reduce both α & β simultaneously is to increase the sample size n. but this method is not always possible. Therefore, to overcome this difficulty, it is customary to fix α, in advance, at a certain level and select a test which minimize β or equivalently, select a critical region which maximize the power of the test, 1-β. In general, we fix at 5% level or 1% level.

Level of significance

Probability of rejecting null hypothesis H_0 when it is true is called level of significance. In other words, level of significance is the probability of type I error or the size of the critical region.

If we fix α at 5% it implies that

- We are (1-α) =95% confident that our decision of rejecting H_0 is correct.
- We are ready to take a 5% risk of rejecting true null hypothesis H_0.

P-value

The P-value is known as the probability value. It is defined as the probability of getting a result that is either the same or more extreme than the actual observations. The P-value is used as an alternative to the rejection point to provide the least significance at which the null hypothesis would be rejected. If the P-value is small, then there is stronger evidence in favor of the alternative hypothesis.

P-value	Decision
P-value > 0.05	The result is not statistically significant and hence don't reject the null hypothesis.
P-value < 0.05	The result is statistically significant. Generally, reject the null hypothesis in favor of the alternative hypothesis.
P-value < 0.01	The result is highly statistically significant, and thus rejects the null hypothesis in favor of the alternative hypothesis.