Hypothesis Testing:

Hypothesis Testing helps to test an assumption that we make about the outcome of an objective/experiment from the sample data.

- ***Null Hypothesis (H0)*** is the status quo, and tested at equality.
- ***Alternate Hypothesis (Ha)*** is what we want to prove, and never tested at equality.

Example,

H0: The person is a liar;

Ha: The person is not a liar

H0 and Ha are mutually exclusive

Once the desired test completes, the conclusions are made in reference to the null hypothesis as either '*reject the null hypothesis*', or '*fail to reject the null hypothesis*'.

The rejection of H0 is based on the alternate hypothesis.

- When Test-Statistic value is either too small or too large than the test-critical value, then reject H0
- Or, when p-value < alpha (α), then reject H0.

If there are any errors in in collecting the data, or performing the test, then the conclusion maybe incorrect leading to Type 1 and Type 2 Errors.

Decision/ Reality	H_o True (Should not reject)	H_o False (Should reject)
Reject H_o	Type I Error (α)	Correct Rejection (No error)
Fail to Reject H_o	Correct Decision (No error)	Type II Error (β)

1- β
Power of the test

1- α
Confidence Level

How sharply H0 is distinguished from Ha

Type I Error: Rejection of null hypothesis when it should not have been rejected.

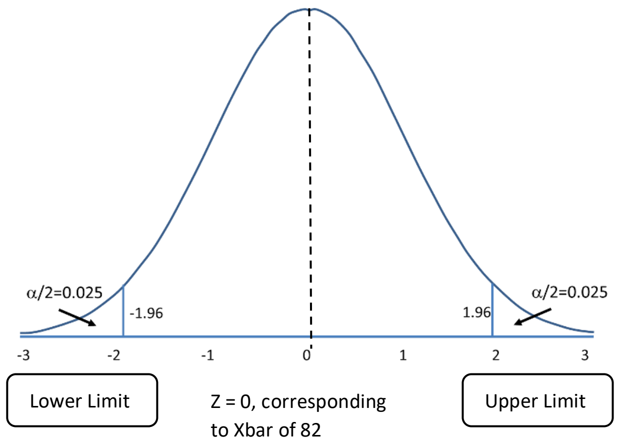
α is the probability of Type 1 Error, also called as the level of significance, or the significance level (it indicates the accepted risk level defaulted at 5% probability).

Type II Error: Failure to reject the null hypothesis, when it should have been rejected. β is the probability of Type 2 error.

p value – It is the actual risk level calculated from the data. p-value provides information on the probability of the observations, given that the null hypothesis is correct. A p-value less than 0.05 is considered statistically significant. It indicates strong evidence against the null hypothesis, as there is less than a 5% probability that the null is correct.

Confidence Level – Indicates how much % confident are we in the decision, usually at 95% level. This mean if we repeat the test multiple times, 95% of the time the results will match the population result.
Confidence Interval - The most common value for α is 0.05 and typically 95% confidence intervals are constructed.
Confidence interval provides an interval, or a range of values, which is expected to cover the true unknown parameter values. This provides richer information in comparison to point estimate, where we have only a single value, thereby exposing any vulnerability in the single estimate.

Confidence Interval is calculated as $\boxed{\bar{X} \pm Z * (\sigma / \sqrt{n})}$
Where $\bar{X}$ is the sample mean, **Z** is the critical test statistic, **σ** is the population standard deviation, and **n** is the sample size.



Confidence Interval Calculation:

Confidence Interval is calculated as

Where Xbar is the sample mean, sigma is the population standard deviation, and n is the sample size.

Sigma/sqrt(n) is the Standard Error.

Standard Error is the Standard Deviation in a Sampling Distribution.

Z value for 95% confidence level from the Z distribution table is 1.96.

Example: Assume for a sample size of 1000 students, their average marks are 82 with a population standard deviation of 10 marks.

Here, n = 1000

Xbar = 82

Sigma = 10

Step 1: Calculate the Standard Error

Standard Error, S.E = 10/sqrt(1000) **=> 0.316**

Step 2: Calculate the Margin of Error

Margin of Error, M.o.E => +- (1.96 * 0.316)

M.o.E => +0.619 and -0.619

Step 3: Add the Sample mean to Margin of Error

Lower Confidence Interval, C.I Lower = 82-0.619 => 81.381

Upper Confidence Interval, C.I Upper = 82+0.619 = > 82.619

Conclusion: We are 95% confident that the mean of the student's marks will be between 81.4 and 82.6

Note: For a T-test, instead of Population Standard deviation, we take the sample standard deviation, and by calculating the degrees of freedom (n-1), we can get the corresponding T value at 95% confidence level instead of Z value.

greatlearning
Power Ahead

Assumptions for t-test

- To perform the hypothesis testing, the following assumptions must hold:
- The variables must follow continuous distribution.
- The sample must be randomly collected from the population.
- The underlying distribution must be normal. Alternatively, if the data is continuous, but may not be assumed to follow a normal distribution, a reasonably large sample size is required. Central Limit Theorem (CLT) asserts that sample mean follows a normal distribution, even if the population distribution is not normal, when sample size is at least 30.
- For 2-sample t-test the population variances of the two distributions must be equal.

Simple example to understand the concept.

When your bike breakdown you will make an educated guess that there may be not **enough petrol** or may be some **technical problem**. Then you will take bike to the nearest workshop to validate your guess/assumption/hypothesis. Depend on mechanic answer you will reject one hypothesis and accept another hypothesis.

- Null hypothesis **(H0)** is **"not enough petrol**"
- Alternate hypothesis **(H1)** is may be some **"technical problem".**

A hypothesis test examines two opposing hypotheses about a population

How to formulate Hypothesis Tests with solved example.

1. Determine the null hypothesis and the alternative hypothesis.
2. Collect and summarize the data into a test statistic.
3. Use the test statistic to determine the p-value.
4. The result is statistically significant if thep-value is less than or equal to the level of significance.

Solved Example

Did only Swimming lose more fat than the Gym?

1. Swimming Only:

- sample mean = 5.9 kg
- sample standard deviation = 4.1 kg
- sample size = n = 42
- standard error = SEM_1 = 4.1/ sqrt (42) = 0.633

2. Gym Only:

- sample mean = 4.1 kg
- sample standard deviation = 3.7 kg
- sample size = n = 47
- standard error = SEM_2 = 3.7/ sqrt (47) = 0.540

measure of variability = $\sqrt{[(0.633)^2 + (0.540)^2]}$

= 0.83

Measure of variability – A measure of variability is a summary statistic that represents the amount of dispersion in a dataset. How spread out are the values? While a measure of central tendency describes the typical value, measures of variability define how far away the data points tend to fall from the center. Here in this example we are using 2 sample so we are using MOV.

greatlearning
Power Ahead

Step 1. Determine the null and alternative hypotheses.

- Null hypothesis: No difference in average fat lost in population for two methods. Population mean difference is zero.
- Alternative hypothesis: There is a difference in average fat lost in population for two methods. Population mean difference is not zero.

Step 2. Collect and summarize data into a test statistic.

- The sample mean difference = 5.9 – 4.1 = 1.8 kg
- and the standard error of the difference is 0.83.
- So, the test statistic: z = = 1.8 – 0 / 0.83 = 2.17

Step 3. Determine the p-value.

- Recall the alternative hypothesis was two-sided.
- p-value = 2 * [proportion of bell-shaped curve above 2.17]
- p-value = proportion is about 2 * 0.015
- p-value = 0.03.

Step 4. Make a decision

The **p-value of 0.03 is less than or equal to 0.05**, so we conclude that there is a **statistically significant difference between average fat loss for the two methods**. i.e. we reject the Null Hypothesis.