Defination :

Inferential statistics is a branch of statistics that involves making inferences or conclusions about
populations based on sample data. It utilizes the principles of probability theory to draw conclusions
about a population parameter based on sample statistics.

Hypothesis Testing

Definition:

Hypothesis testing is a statistical method used to make inferences about population parameters based on sample data. It involves evaluating two competing hypotheses about a population parameter: the null hypothesis (H0) and the alternative hypothesis (H1). The goal is to determine whether there is enough evidence in the sample data to reject the null hypothesis in favor of the alternative hypothesis.

Key Concepts:

Null Hypothesis (H0): The null hypothesis represents the default assumption or claim about a population parameter. It is often denoted as H0 and is assumed to be true unless there is sufficient evidence to reject it.

Alternative Hypothesis (H1): The alternative hypothesis is the assertion that contradicts the null hypothesis. It is denoted as H1 and is accepted if there is enough evidence to reject the null hypothesis.

Significance Level (α): The significance level, denoted as α (alpha), specifies the probability of rejecting the null hypothesis when it is actually true. Commonly used significance levels include 0.05 (5%) and 0.01 (1%).

Test Statistic: A test statistic is a numerical value calculated from sample data that is used to determine whether to reject the null hypothesis. The choice of test statistic depends on the hypothesis being tested and the nature of the data.

P-value: The p-value is the probability of observing the test statistic (or a more extreme value) if the null hypothesis is true. It quantifies the strength of evidence against the null hypothesis. A smaller p-value indicates stronger evidence against the null hypothesis.

Critical Region: The critical region is the set of values of the test statistic for which the null hypothesis is rejected. It is determined based on the significance level and the distribution of the test statistic.

Decision Rule: The decision rule specifies whether to reject the null hypothesis based on the observed value of the test statistic and the critical region. If the test statistic falls within the critical region, the null hypothesis is rejected; otherwise, it is not rejected.

Steps in Hypothesis Testing:

State the Hypotheses: Formulate the null hypothesis (H0) and the alternative hypothesis (H1) based on the research question.

Choose the Significance Level (α): Determine the significance level, which represents the maximum probability of committing a Type I error (rejecting the null hypothesis when it is true).

Select the Test Statistic: Choose an appropriate test statistic based on the type of data and the hypothesis being tested.

Compute the Test Statistic: Calculate the value of the test statistic using the sample data.

Determine the Critical Region: Determine the critical region of the test statistic based on the significance level and the distribution of the test statistic.

Compare the Test Statistic and Critical Region: Compare the observed value of the test statistic with the critical region. If the test statistic falls within the critical region, reject the null hypothesis; otherwise, do not reject the null hypothesis.

Draw Conclusion: Based on the decision made in step 6, draw a conclusion about the null hypothesis and interpret the results in the context of the research question.

Example:
Suppose a researcher wants to test whether a new drug is more effective than an existing drug for treating a certain medical condition. The null hypothesis (H0) states that there is no difference in effectiveness between the two drugs, while the alternative hypothesis (H1) asserts that the new drug is more effective than the existing drug. The researcher collects data on the recovery rates of patients treated with both drugs and performs a hypothesis test using an appropriate statistical test (e.g., t-test or z-test). Based on the test results and the chosen significance level, the researcher decides whether to reject the null hypothesis and conclude whether the new drug is indeed more effective.

Conclusion:
Hypothesis testing is a powerful statistical tool used in various fields to make decisions and draw conclusions based on sample data. By following a systematic approach and considering the significance level, test statistic, and critical region, researchers can make informed decisions about hypotheses and contribute to the advancement of knowledge in their respective domains.

Let's consider an example to illustrate hypothesis testing in a real-world scenario.

Scenario:
Suppose a beverage company wants to determine whether a new marketing campaign has led to an increase in sales of their flagship product. The company's marketing team launched the campaign last month, aiming to boost sales by attracting new customers and increasing repeat purchases.

Hypotheses:

Null Hypothesis (H0): The new marketing campaign has not led to a significant increase in sales.
Alternative Hypothesis (H1): The new marketing campaign has led to a significant increase in sales.
Significance Level:
Let's choose a significance level (α) of 0.05, indicating that we are willing to accept a 5% chance of making a Type I error (rejecting the null hypothesis when it is actually true).

Data Collection and Test Statistic:
The company collects sales data for the past month, both before and after the launch of the marketing campaign. They calculate the difference in average daily sales between the two periods as the test statistic.

Critical Region:
Based on the chosen significance level and the distribution of the test statistic, the critical region is determined.

Hypothesis Test:

State the Hypotheses:

Null Hypothesis (H0): The new marketing campaign has not led to a significant increase in sales.
Alternative Hypothesis (H1): The new marketing campaign has led to a significant increase in sales.
Choose the Significance Level (α):

α = 0.05 (5%)
Select the Test Statistic:

Difference in average daily sales before and after the marketing campaign.

Compute the Test Statistic:
Calculate the difference in average daily sales for both periods.

Determine the Critical Region:
Based on the chosen significance level and the distribution of the test statistic.

Compare the Test Statistic and Critical Region:
If the difference in average daily sales falls within the critical region, reject the null hypothesis; otherwise, do not reject the null hypothesis.

Draw Conclusion:

Based on the decision made in step 6, draw a conclusion about the effectiveness of the marketing campaign in increasing sales.
Conclusion:
After performing the hypothesis test and analyzing the results, the company can determine whether the new marketing campaign has led to a significant increase in sales. If the null hypothesis is rejected, it provides evidence in support of the alternative hypothesis, indicating that the marketing campaign has been effective. Conversely, if the null hypothesis is not rejected, it suggests that the campaign may not have had a significant impact on sales, and further analysis or adjustments to the marketing strategy may be necessary.

Type I Error:

=> Type I error, also known as a false positive, occurs when we reject the null hypothesis when it is actually true.
=> In hypothesis testing, it means concluding that there is a significant effect or difference when, in reality, there is no such effect or difference.
=> The probability of committing a Type I error is denoted by α (alpha) and is known as the significance level of the test.
=> A lower significance level reduces the probability of Type I error but may increase the probability of Type II error.

Type II Error:

=> Type II error, also known as a false negative, occurs when we fail to reject the null hypothesis when it is actually false.
=> In hypothesis testing, it means failing to detect a significant effect or difference when, in reality, there is such an effect or difference.
=> The probability of committing a Type II error is denoted by β (beta).
=> The power of a statistical test (1 - β) is the probability of correctly rejecting the null hypothesis when it is false.
=> Increasing the sample size or adjusting the significance level can reduce the probability of Type II error but may increase the probability of Type I error.

Example:
Consider a medical diagnostic test for a particular disease:

Type I Error: Concluding that a patient has the disease when they do not (false positive).
Type II Error: Failing to diagnose the disease in a patient who actually has it (false negative).
In the context of the medical test:

A Type I error might lead to unnecessary treatment or stress for the patient.
A Type II error might result in delayed treatment, allowing the disease to progress without intervention.

Here's how significance testing works:

1. Formulate Hypotheses:

    Null Hypothesis (H0): This hypothesis assumes that there is no effect or difference in the population.
    Alternative Hypothesis (H1): This hypothesis suggests that there is a significant effect or difference in the population.

2. Select Significance Level (α):

    The significance level (α) determines the threshold for accepting or rejecting the null hypothesis.
    Commonly used values for α include 0.05 (5%) or 0.01 (1%).

3. Choose a Test Statistic:

    The choice of test statistic depends on the type of data and the hypothesis being tested.
    Common test statistics include t-test, chi-square test, ANOVA, etc.

4. Calculate the P-value:

The p-value is the probability of obtaining the observed data or more extreme data if the null hypothesis is true.
A low p-value indicates that the observed effect is unlikely to occur by chance alone.

5. Make a Decision:

If the p-value is less than the significance level (α), we reject the null hypothesis in favor of the alternative hypothesis.
If the p-value is greater than or equal to the significance level, we fail to reject the null hypothesis.

6. Draw Conclusions:

If we reject the null hypothesis, we conclude that there is sufficient evidence to support the alternative hypothesis.
If we fail to reject the null hypothesis, we conclude that there is not enough evidence to support the alternative hypothesis.