|  |  |
| --- | --- |
| Subject: Data Analytics and Visualization Lab | Course ID: CSL-601 |
| Semester: VI | Course: AI & DS |
| Laboratory: 407 | Name of teacher: Prof. Gitanjali Korgaonkar |
| Name of Student: Ayush Gupta | Roll No: VU2S2324001 |

**EXPERIMENT NO. 5**

**Aim:**

To understand and implement Hypothesis Testing.

**Theory:**

**Hypothesis Testing**

Hypothesis testing is a statistical method used to draw conclusions about a population based on a sample. It involves formulating assumptions, collecting data, and determining whether the observed results provide enough evidence to support or reject a claim.

**Steps in Hypothesis Testing:**

1. **State the Hypotheses**
   * **Null Hypothesis (H0​)**: The default assumption that there is no effect or no difference.
   * **Alternative Hypothesis (H1 or Ha​)**: The claim we want to test, which suggests a significant effect or difference.
2. **Choose the Significance Level (α)**
   * Defines the probability of rejecting the null hypothesis when it is actually true (Type I error).
   * Common values are **0.05 (5%)** or **0.01 (1%)**.
3. **Select the Appropriate Statistical Test**
   * The test depends on the type of data and hypothesis.
   * Examples include:
     + **t-test**: Compares the means of two groups.
     + **Chi-square test**: Used for categorical data.
     + **ANOVA**: Compares means of more than two groups.
     + **Z-test**: Used when the sample size is large and population variance is known.
4. **Collect Data and Compute the Test Statistic**
   * A test statistic is calculated based on the chosen test (e.g., **t-statistic, z-statistic**).
   * This value is compared against a critical value or used to determine the **p-value**.
5. **Determine the Critical Region or P-value**
   * The **critical region** is the range of values that lead to rejecting H0​.
   * The **p-value** represents the probability of obtaining the observed result if H0​ is true.
6. **Make a Decision**
   * If the **p-value < α**, **reject H0​** (statistically significant difference).
   * If the **p-value > α**, **fail to reject H0​** (no significant difference).

**Types of Hypothesis Tests:**

1. **One-Tailed vs. Two-Tailed Tests**

* **One-Tailed Test**: Tests if a parameter is **greater than or less than** a certain value.
* **Two-Tailed Test**: Tests if a parameter is **different from** a certain value, without specifying direction.

1. **Parametric vs. Non-Parametric Tests**

* **Parametric Tests**: Assume data follows a normal distribution (e.g., t-test, ANOVA).
* **Non-Parametric Tests**: Do not assume a specific distribution (e.g., Mann-Whitney U test, Kruskal-Wallis test).

**Errors in Hypothesis Testing**

* **Type I Error (False Positive)**: Rejecting H0​ when it is actually true. (Controlled by α).
* **Type II Error (False Negative)**: Failing to reject H0​ when it is actually false. (Controlled by sample size and effect size).

**Learning Objectives**

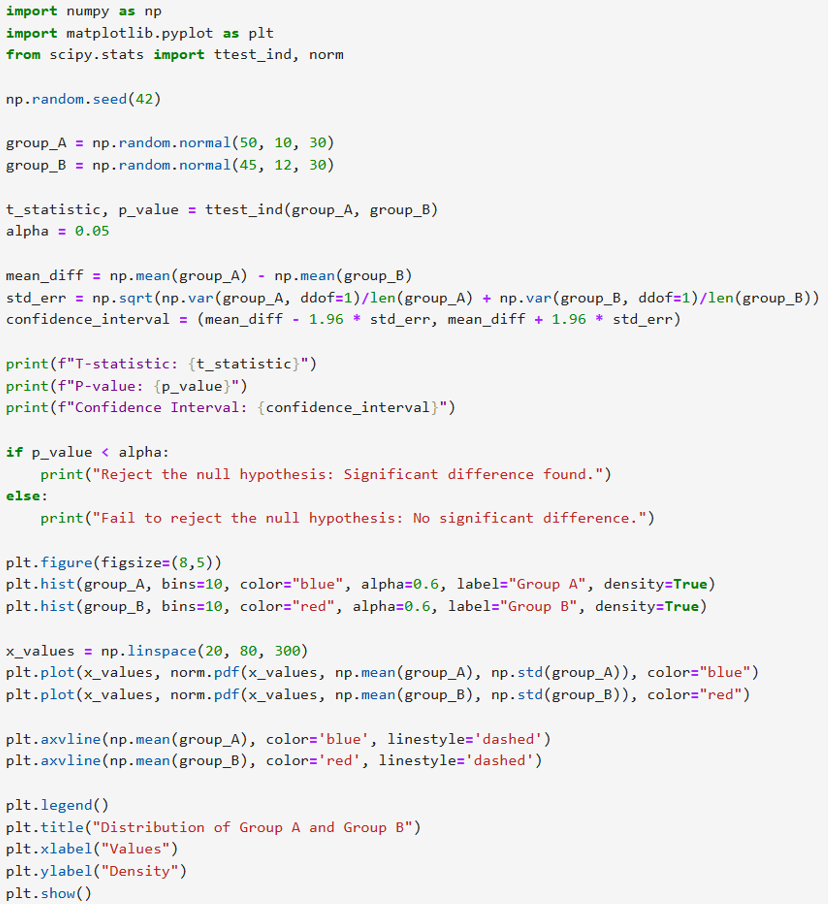
* Understand the **concept** and **importance** of hypothesis testing.
* Learn to apply hypothesis testing using **Python**.
* Differentiate between **null and alternative hypotheses**.
* Interpret **p-values** and **statistical significance**.
* Understand **errors** in hypothesis testing (Type I & Type II).

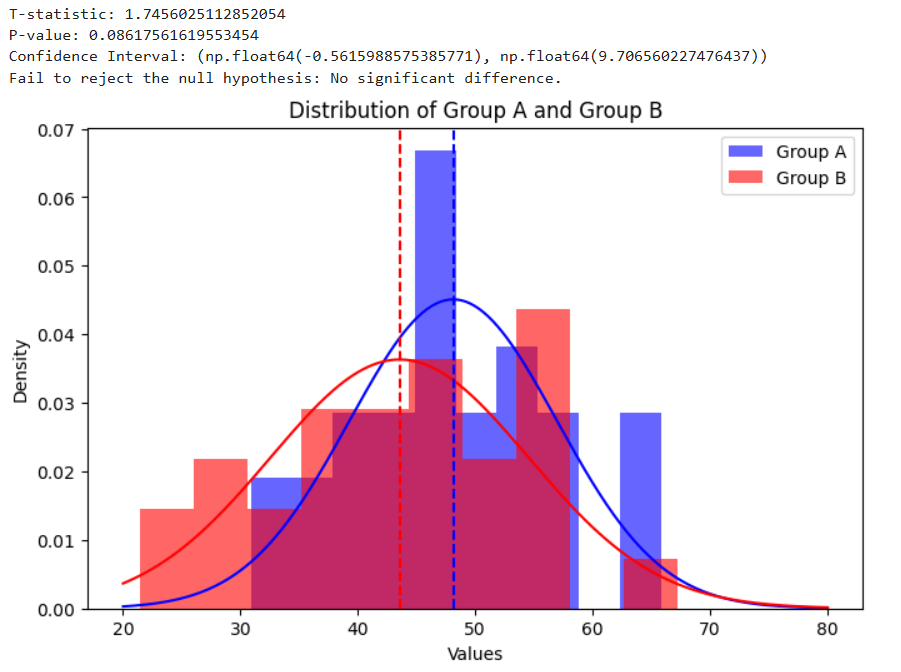
**Conclusion:**

In this experiment, we implemented **hypothesis testing** using a **two-sample t-test** in Python. We formulated **null and alternative hypotheses**, computed the **test statistic and p-value**, and made **decisions** based on the significance level. This process is fundamental in statistical analysis and research, helping to determine if observed differences are statistically significant.



**Program and Output:**

****

****