How to Conduct a Two Sample T-Test in Python

A two sample t-test is used to test whether or not the means of two populations are equal.

This lesson explains how to conduct a two sample t-test in Python.

Example: **Two Sample t-Test in Python**

Researchers want to know whether or not two different subjects have the same mean. To test this, they collect a simple random sample of 435 students from each subject – We use our school dataset for Writing and English subjects

Use the following steps to conduct a two sample t-test to determine if the two subjects have the same mean.

Step 1: **Create the data**.

First, we read our school dataset.

```
import pandas
data = pandas.read_csv("https://modcom.co.ke/data/datasets/schoolcleaned.csv")
data
```

Step 2: **Conduct a two sample t-test**.

Next, we'll use the ttest_ind() function from the scipy.stats library to conduct a two sample t-test, which uses the following syntax:

ttest_ind(**a**, **b**)

where:

- **a**: an array of sample observations for group 1
- **b**: an array of sample observations for group 2

Thus, we can proceed to perform the two sample t-test with equal variances:

The two hypotheses for this particular two sample t-test are as follows:

H0: μ1 = μ2 (the two population means are equal)

HA: μ1 ≠μ2 (the two population means are *not* equal)

```
import scipy.stats as stats

# perform two sample t-test
stats.ttest_ind(a=group1, b=group2)

TtestResult(statistic=-7.705680077982495, pvalue=3.7955905916649034e-14
```

The p-value is `0.00000000000003795590591664903`4.

Step 3: **Interpret the results**.

The two hypotheses for this particular two sample t-test are as follows:

H0: μ1 = μ2 (the two population means are equal)

HA: μ1 ≠μ2 (the two population means are *not* equal)

Because the p-value of our test (`0.00000000000003795590591664903`4) is less than alpha = 0.05, we reject the null hypothesis of the test. We accept the alternative (**HA**: μ1 ≠μ2 (the two population means are *not* equal). We do not have sufficient evidence to say that the mean of both Writing and English are same.

Notebook

https://colab.research.google.com/drive/1WAt2Ml2xbs-Q7S0e32kQ5RUz3erbFT4c?usp=sharing

Assignments

Researchers want to know whether or not two different variables StudyTime and SleepTime has the same mean. To test this, they collect a simple random sample of 435 students from each subject – We use our school dataset for SleepTime and StudyTime.

Use the following steps to conduct a two sample t-test to determine if the two have the same mean.