

How do you conduct a paired samples t-test in Python?

Authored by
stats writer

April 17, 2024

RECOMMENDED CITATION

stats writer (2024). *How do you conduct a paired samples t-test in Python?*.

PSYCHOLOGICAL SCALES. Retrieved from <https://scales.arabpsychology.com/?p=136184>

A paired samples t-test is a statistical method used to compare the means of two related groups. In Python, this can be conducted using the "ttest_rel" function from the "scipy.stats" library. This function takes in two arrays representing the two groups and calculates the t-statistic and p-value. The t-statistic can then be compared to a critical value to determine if there is a significant difference between the two groups. This method is commonly used in research and data analysis to assess the impact of a treatment or intervention on a specific outcome.

Conduct a Paired Samples T-Test in Python

A is used to compare the means of two samples when each observation in one sample can be paired with an observation in the other sample.

This tutorial explains how to conduct a paired samples t-test in Python.

Example: Paired Samples T-Test in Python

Suppose we want to know whether a certain study program significantly impacts student performance on a particular exam. To test this, we have 15 students in a class take a pre-test. Then, we have each of the students participate in the study program for two weeks. Then, the students retake a test of similar difficulty.

To compare the difference between the mean scores on the first and second test, we use a paired samples t-test

because for each student their first test score can be paired with their second test score.

Perform the following steps to conduct a paired samples t-test in Python.

Step 1: Create the data.

First, we'll create two arrays to hold the pre and post-test scores:

```
pre = post =
```

Step 2: Conduct a Paired Samples T-Test.

Next, we'll use the `ttest_rel` from the `scipy.stats` library to conduct a paired samples t-test, which uses the following syntax:

```
ttest_rel(a, b)
```

where:

a: an array of sample observations from group 1
b: an array of sample observations from group 2

Here's how to use this function in our specific example:

```
import scipy.stats as stats
```

```
#perform the paired samples t-test
```

```
stats.ttest_rel(pre, post)
```

```
(statistic=-2.9732, pvalue=0.0101)
```

The test statistic is -2.9732 and the corresponding two-sided p-value is 0.0101.

In this example, the paired samples t-test uses the following null and alternative hypotheses:

H0: The mean pre-test and post-test scores are equal

HA:The mean pre-test and post-test scores are *not* equal

Since the p-value (0.0101) is less than 0.05, we reject the null hypothesis. We have sufficient evidence to say that the true mean test score is different for students before and after participating in the study program.