

How do you perform an F-Test in Python?

Authored by
stats writer

April 17, 2024

RECOMMENDED CITATION

stats writer (2024). *How do you perform an F-Test in Python?*. PSYCHOLOGICAL SCALES.
Retrieved from <https://scales.arabpsychology.com/?p=136213>

The F-Test is a statistical test used to compare the variances of two or more groups of data. In Python, the F-Test can be performed using the "f_oneway" function from the "scipy.stats" module. This function takes in the data from each group as separate arrays and returns the F-statistic and p-value. The F-statistic can then be compared to the critical F-value to determine if there is a significant difference in variances between the groups. Additionally, the p-value can indicate the level of significance for the results. Overall, performing an F-Test in Python involves importing the necessary modules, preparing the data, and using the "f_oneway" function to obtain the F-statistic and p-value.

Perform an F-Test in Python

An F-test is used to test whether two population variances are equal. The null and alternative hypotheses for the test are as follows:

H0: $\sigma_1^2 = \sigma_2^2$ (the population variances are equal)

H1: $\sigma_1^2 \neq \sigma_2^2$ (the population variances are *not* equal)

This tutorial explains how to perform an F-test in Python.

Example: F-Test in Python

Suppose we have the following two samples:

$x = y =$

We can use the following function to perform an F-test

to determine if the two populations these samples came from have equal variances:

```
import numpy as np
```

```
#define F-test functiondef f_test(x, y):
```

```
x = np.array(x)
```

```
y = np.array(y)
```

```
f = np.var(x, ddof=1)/np.var(y, ddof=1) #calculate F test statistic
```

```
dfn = x.size-1 #define degrees of freedom numerator
```

```
dfd = y.size-1 #define degrees of freedom denominator
```

```
p = 1-scipy.stats.f.cdf(f, dfn, dfd) #find p-value of F test statistic return f, p
```

```
#perform F-test
```

```
f_test(x, y)
```

```
(4.38712, 0.019127)
```

The F test statistic is 4.38712 and the corresponding p-value is 0.019127. Since this p-value is less than .05, we would reject the null hypothesis. This means we have sufficient evidence to say that the two population variances are *not* equal.

Notes

The F test statistic is calculated as s_1^2 / s_2^2 . By default, `numpy.var` calculates the population variance. To calculate the sample variance, we need to specify `ddof=1`. The p-value corresponds to $1 - \text{cdf}$ of the F distribution with numerator degrees of freedom = $n_1 - 1$ and denominator degrees of freedom = $n_2 - 1$. This function only works when the first sample variance is larger than the second sample variance. Thus, define the two samples in such a way that they work with the function.

When to Use the F-Test

The F-test is typically used to answer one of the following questions:

1. Do two samples come from populations with equal variances?
2. Does a new treatment or process reduce the variability of some current treatment or process?

How to Perform an F-Test in R