ABOUT STUDY BASIC STATS MACHINE LEARNING SOFTWARE TUTORIALS > TOOLS >

O DECEMBER 18, 2020 💄 BY ZACH

How to Perform Tukey's Test in Python

A one-way ANOVA is used to determine whether or not there is a statistically significant difference between the means of three or more independent groups.

If the overall p-value from the ANOVA table is less than some significance level, then we have sufficient evidence to say that at least one of the means of the groups is different from the others.

However, this doesn't tell us which groups are different from

each other. It simply tells us that not all of the group means are equal. In order to find out exactly which groups are different from each other, we must conduct a post hoc test.

One of the most commonly used post hoc tests is **Tukey's**

Test, which allows us to make pairwise comparisons between the means of each group while controlling for the family-wise error rate.

perform Tukey's Test in Python.

Step 1: Load Necessary Packages and

This tutorial provides a step-by-step example of how to

Functions First, we'll load the necessary packages and functions in

Python:

```
import pandas as pd
import numpy as np
from scipy.stats import f_oneway
from statsmodels.stats.multicomp import pairwise_tukeyh
sd
```

The following code shows how to create a fake dataset with

Step 2: Fit the ANOVA Model

three groups (A, B, and C) and fit a one-way ANOVA model to the data to determine if the mean values for each group are equal:

```
#enter data for three groups
a = [85, 86, 88, 75, 78, 94, 98, 79, 71, 80]
b = [91, 92, 93, 90, 97, 94, 82, 88, 95, 96]
c = [79, 78, 88, 94, 92, 85, 83, 85, 82, 81]

#perform one-way ANOVA
f_oneway(a, b, c)

F_onewayResult(statistic=5.167774552944481, pvalue=0.01
2582197136592609)
```

Since this is less than .05, we have sufficient evidence to say that the mean values across each group are not equal.

We can see that the overall p-value from the ANOVA table

Thus, we can proceed to perform Tukey's Test to determine

Step 3: Perform Tukey's Test

the pairwise_tukeyhsd() function from the statsmodels library:

, 79, 71, 80,

To perform Tukey's test in Python, we can use

exactly which group means are different.

is **0.01258**.

#create DataFrame to hold data
df = pd.DataFrame({'score': [85, 86, 88, 75, 78, 94, 98]

```
91, 92, 93, 90, 97, 94, 82, 8
                           79, 78, 88, 94, 92, 85, 83, 8
                  'group': np.repeat(['a', 'b', 'c'],
 repeats=10)})
 # perform Tukey's test
 tukey = pairwise_tukeyhsd(endog=df['score'],
                        groups=df['group'],
                        alpha=0.05)
 #display results
 print(tukey)
  Multiple Comparison of Means - Tukey HSD, FWER=0.05
  _____
 group1 group2 meandiff p-adj lower upper reject
      a b 8.4 0.0158 1.4272 15.3728 True
      a c 1.3 0.8864 -5.6728 8.2728 False
            c -7.1 0.0453 -14.0728 -0.1272 True
Here's how to interpret the output:

    P-value for the difference in means between a and b:
```

.8864P-value for the difference in means between b and c:

.0158

.0453

• P-value for the difference in means between a and c:

Thus, we would conclude that there is a statistically significant difference between the means of groups *a* and *b* and groups *b* and *c*, but not a statistically significant

difference between the means of groups a and c.

Additional Resources

View all posts by Zach

Email *

Website

POST COMMENT

How to Perform a One-Way ANOVA in Python

How to Perform a Two-Way ANOVA in Python

Published by Zach

How to Perform a Repeated Measures ANOVA in Python

PREV	NEXT
Comparing Mean Absolute Deviation vs Standard Deviation	How to Drop Duplicate Rows in a Pandas DataFrame
Leave a Reply	
	e published. Required fields are marked *
Tour officer addition from the first	o pastrollo al regalica licias al ciliantea
Comment	
	//
Name *	

SEARCH

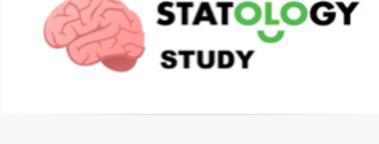
Search ...

ABOUT

Statology is a site that makes learning statistics easy by explaining topics in simple and straightforward ways. Learn more about us.

STATOLOGY STUDY

online statistics study guide that helps you understand all of the core concepts taught in any elementary statistics course and makes your life so much easier as a student.



Try out our free online statistics

CALCULATORS

calculators if you're looking for some help finding probabilities, pvalues, critical values, sample sizes, expected values, summary statistics, or correlation coefficients.

How to Drop Rows by Index in

RECENT POSTS

Pandas (With Examples)

How to Drop Columns by Index in

Pandas

How to Combine Two Columns in

Pandas (With Examples)