

ABOUT BASIC STATS MACHINE LEARNING SOFTWARE TUTORIALS * TOOLS * STUDY

JULY 13, 2020 BY ZACH

How to Perform an ANCOVA in Python

An ANCOVA ("analysis of covariance") is used to determine whether or not there is a statistically significant difference between the means of three or more independent groups, after controlling for one or more covariates.

This tutorial explains how to perform an ANCOVA in Python.

Example: ANCOVA in Python

A teacher wants to know if three different studying techniques have an impact on exam scores, but she wants to account for the current grade that the student already has in the class.

She will perform an ANCOVA using the following variables:

- Factor variable: studying technique
- Covariate: current grade
- Response variable: exam score

Use the following steps to perform an ANCOVA on this dataset:

Step 1: Enter the data.

First, we'll create a pandas DataFrame to hold our data:

```
import numpy as np
import pandas as pd
#create data
df = pd.DataFrame({'technique': np.repeat(['A', 'B', 'C'], 5),
                   'current_grade': [67, 88, 75, 77, 85,
                                    92, 69, 77, 74, 88,
                                    96, 91, 88, 82, 80],
                   'exam_score': [77, 89, 72, 74, 69,
                                 78, 88, 93, 94, 90,
                                 85, 81, 83, 88, 79]})
#view data
df
               current_grade exam_score
   technique
                                      77
                                      89
                          75
                          77
                                      74
                          85
                                       69
                          92
                                      78
                          69
                                       88
                          77
                                       93
                          74
                                       94
                          88
                                       90
10
                          96
                                       85
11
                          91
                                       81
12
                          88
                                       83
13
                          82
                                       88
14
                          80
                                      79
```

Step 2: Perform the ANCOVA.

Next, we'll perform an ANCOVA using the ancova() function from the pingouin library:

```
pip install pingouin
from pingouin import ancova
#perform ANCOVA
ancova(data=df, dv='exam_score', covar='current_grade', between='technique')
        Source
                       SS
                                                                      np2
                                                         p-unc
                       390.575130
                                                                      0.46653
        technique
                                              4.80997
                                                         0.03155
                       4.193886
                                                                      0.00930
        current_grade
                                               0.10329
                                                         0.75393
        Residual
                       446.606114
                                       11
                                               NaN
                                                         NaN
                                                                      NaN
```

Step 3: Interpret the results.

From the ANCOVA table we see that the p-value (p-unc = "uncorrected p-value") for study technique is 0.03155. Since this value is less than 0.05, we can reject the null hypothesis that each of the studying techniques leads to the same average exam score, even after accounting for the student's current grade in the class.

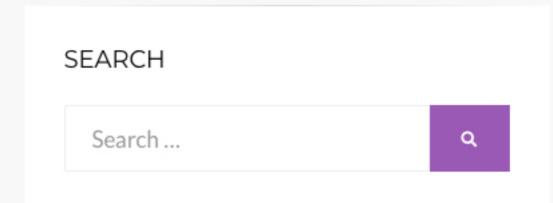


Published by Zach

View all posts by Zach

POST COMMENT

PREV How to Perform the Friedman Test in Python	NEXT How to Calculate Moving Averages in Python
Leave a Reply Your email address will not be published. Re	equired fields are marked *
Comment	
Name *	
Email *	
Website	



ABOUT

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

STATOLOGY STUDY

Statology Study is the ultimate 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.



CALCULATORS

Try out our free online statistics calculators if you're looking for some help finding probabilities, p-values, critical values, sample sizes, expected values, summary statistics, or correlation coefficients.

RECENT POSTS

Bray-Curtis Dissimilarity: Definition & Examples

How to Perform a Bonferroni Correction in Excel

What is a Segmented Bar Chart? (Definition