How can I calculate the Intraclass Correlation Coefficient using Python?

Authored by

stats writer

April 27, 2024

The process of calculating the Intraclass Correlation Coefficient (ICC) using Python involves using statistical functions and libraries to determine the level of agreement between two or more variables in a dataset. This measure is commonly used in research and data analysis to assess the reliability and consistency of measurements. By inputting the necessary data and parameters into the appropriate Python functions, the ICC value can be calculated, providing valuable insights into the relationship between variables and aiding in the interpretation of results. This method allows for efficient and accurate determination of the ICC, making it a valuable tool for researchers and data analysts.

Calculate Intraclass Correlation Coefficient in Python

An (ICC) is used to determine if items or subjects can be rated reliably by different raters.

The value of an ICC can range from 0 to 1, with 0 indicating no among raters and 1 indicating perfect reliability.

The easiest way to calculate ICC in Python is to use the function from the , which uses the following syntax:

pingouin.intraclass_corr(data, targets, raters, ratings)

where:

data: The name of the dataframe
targets: Name of column containing the "targets" (the things being rated)
raters: Name of column containing the raters

ratings: Name of column containing the ratings

This tutorial provides an example of how to use this function in practice.

Step 1: Install Pingouin

First, we must install Pingouin:

```
pip install pingouin
```

Step 2: Create the Data

Suppose four different judges were asked to rate the quality of six different college entrance exams. We can create the following dataframe to hold the ratings of the judges:

```
import pandas as pd

#create DataFrame
df = pd.DataFrame({'exam': ,
'judge': ,
'rating': })

#view first five rows of DataFrame
df.head()
```

exam judge rating
0 1 A 1
1 2 A 1
2 3 A 3
3 4 A 6
4 5 A 6

Step 3: Calculate the Intraclass Correlation Coefficient

Next, we'll use the following code to calculate the intraclass correlation coefficient:

```
import pingouin as pg

icc = pg.intraclass_corr(data=df, targets='exam', raters='judge', ratings='rating')

icc.set_index('Type')
```

Description ICC F df1 df2 pval CI95%
Type
ICC1 Single raters absolute 0.505252 5.084916 5 18 0.004430
ICC2 Single random raters 0.503054 4.909385 5 15 0.007352
ICC3 Single fixed raters 0.494272 4.909385 5 15

0.007352
ICC1k Average raters absolute 0.803340 5.084916 5 18 0.004430
ICC2k Average random raters 0.801947 4.909385 5 15 0.007352
ICC3k Average fixed raters 0.796309 4.909385 5 15 0.007352

This function returns the following results:

Description: The type of ICC calculated
ICC: The intraclass correlation coefficient (ICC)
F: The F-value of the ICC
df1, df2: The degrees of freedom associated with the F-value
pval: The p-value associated with the F-value
CI95%: The 95% confidence interval for the ICC

Notice that there are six different ICC's calculated here. This is because there are multiple ways to calculate the ICC depending on the following assumptions:

Model: One-Way Random Effects, Two-Way Random Effects, or Two-Way Mixed Effects
Type of Relationship: Consistency or Absolute

Agreement

Unit: Single rater or the mean of raters

For a detailed explanation of these assumptions, please refer to .