Statistics. Simplified.

STATOLOGY

☰ MENU

🕐 AUGUST 3, 2021   👤 BY ZACH

How to Calculate Spearman Rank Correlation in R

In statistics, **correlation** refers to the strength and direction of a relationship between two variables. The value of a correlation coefficient can range from -1 to 1, with the following interpretations:

- **-1:** a perfect negative relationship between two variables
- **0:** no relationship between two variables
- **1:** a perfect positive relationship between two variables

One special type of correlation is called **Spearman Rank Correlation**, which is used to measure the correlation between two ranked variables. (e.g. rank of a student's math exam score vs. rank of their science exam score in a class).

To calculate the Spearman rank correlation between two variables in R, we can use the following basic syntax:

```
corr <- cor.test(x, y, method = 'spearman')
```

The following examples show how to use this function in practice.

Example 1: Spearman Rank Correlation Between Vectors

The following code shows how to calculate the Spearman rank correlation between two vectors in R:

```
#define data
x <- c(70, 78, 90, 87, 84, 86, 91, 74, 83, 85)
y <- c(90, 94, 79, 86, 84, 83, 88, 92, 76, 75)

#calculate Spearman rank correlation between x and y
cor.test(x, y, method = 'spearman')

        Spearman's rank correlation rho

data:  x and y
S = 234, p-value = 0.2324
alternative hypothesis: true rho is not equal to 0
sample estimates:
       rho
-0.4181818
```

From the output we can see that the Spearman rank correlation is **-0.41818** and the corresponding p-value is **0.2324**.

This indicates that there is a negative correlation between the two vectors.

However, since the p-value of the correlation is not less than 0.05, the correlation is not statistically significant.

Example 2: Spearman Rank Correlation Between Columns in Data Frame

The following code shows how to calculate the Spearman rank correlation between two column in a data frame:

```
#define data frame
df <- data.frame(team=c('A', 'B', 'C', 'D', 'E', 'F', 'G', 'H', 'I', 'J
                 points=c(67, 70, 75, 78, 73, 89, 84, 99, 90, 91),
                 assists=c(22, 27, 30, 23, 25, 31, 38, 35, 34, 32))

#calculate Spearman rank correlation between x and y
cor.test(df$points, df$assists, method = 'spearman')

        Spearman's rank correlation rho

data:  df$points and df$assists
S = 36, p-value = 0.01165
alternative hypothesis: true rho is not equal to 0
sample estimates:
      rho
0.7818182
```

From the output we can see that the Spearman rank correlation is **0.7818** and the corresponding p-value is **0.01165**.

This indicates that there is a strong positive correlation between the two vectors.

Since the p-value of the correlation is less than 0.05, the correlation is statistically significant.

Additional Resources

How to Calculate Partial Correlation in R
How to Calculate Autocorrelation in R
How to Calculate Rolling Correlation in R
How to Report Spearman's Correlation in APA Format