

How can I calculate a Phi Coefficient in R?

Authored by
stats writer

April 23, 2024

RECOMMENDED CITATION

stats writer (2024). *How can I calculate a Phi Coefficient in R?*. PSYCHOLOGICAL SCALES. Retrieved from <https://scales.arabpsychology.com/?p=138465>

The Phi Coefficient is a statistical measure used to determine the degree of association between two categorical variables. In order to calculate this coefficient using R, you will need to first import the necessary data into the program. Then, you can use the "cor()" function with the method parameter set to "phi" to calculate the Phi Coefficient. This will return a value between -1 and 1, where a value of 0 indicates no association, a negative value indicates a negative association, and a positive value indicates a positive association. This method can be useful in analyzing relationships between variables and drawing conclusions from categorical data.

Calculate a Phi Coefficient in R

A Phi Coefficient (sometimes called a *mean square contingency coefficient*) is a measure of the association between two binary variables.

For a given 2×2 table for two random variables x and y:

	y = 0	y = 1
x = 0	A	B
x = 1	C	D

The Phi Coefficient can be calculated as:

$$\Phi = (AD-BC) / \sqrt{(A+B)(C+D)(A+C)(B+D)}$$

Example: Calculating a Phi Coefficient in R

Suppose we want to know whether or not gender is

associated with political party preference so we take a survey of 25 voters and survey them on their political party preference.

The following table shows the results of the survey:

	Dem	Rep
Male	4	9
Female	8	4

We can use the following code to enter this data into a 2x2 matrix in R:

```
#create 2x2 table
```

```
data = matrix(c(4, 8, 9, 4), nrow = 2)
```

```
#view dataset
```

```
data
```

```
4 9
```

```
8 4
```

We can then use the function from the psych package

to calculate the Phi Coefficient between the two variables:

```
#load psych package  
library(psych)
```

```
#calculate Phi Coefficient  
phi(data)
```

```
-0.36
```

The Phi Coefficient turns out to be -0.36.

Note that the phi function rounds to 2 digits by default, but you can specify the function to round to as many digits as you'd like:

```
#calculate Phi Coefficient and round to 6 digits  
phi(data, digits = 6)
```

```
-0.358974
```

How to Interpret a Phi Coefficient

-1 indicates a perfectly negative relationship between the two variables. 0 indicates no association between

the two variables.1 indicates a perfectly positive relationship between the two variables.

In general, the further away a Phi Coefficient is from zero, the stronger the relationship between the two variables.

In other words, the further away a Phi Coefficient is from zero, the more evidence there is for some type of systematic pattern between the two variables.

ARABPSYCHOLOGY.COM