How do you perform a Cramer-Von Mises test in R? Can you provide some examples?

Authored by

stats writer

June 26, 2024

RECOMMENDED CITATION

stats writer (2024). *How do you perform a Cramer-Von Mises test in R? Can you provide some examples?*. PSYCHOLOGICAL SCALES. Retrieved from https://scales.arabpsychology.com/?p=153144

ARABPSYCHOLOGY.COM

The Cramer-Von Mises test is a statistical test used to determine whether a given set of data follows a specific distribution. In R, this test can be performed using the "cvm.test" function from the "stats" package. This function takes in two arguments - the data and the theoretical distribution to be tested against. It then calculates the Cramer-Von Mises statistic and its associated p-value, which can be used to determine the goodness of fit between the data and the chosen distribution. Some examples of using this test in R include testing whether a set of data follows a normal distribution or a exponential distribution.

Perform a Cramer-Von Mises Test in R (With Examples)

The Cramer-Von Mises test is used to determine whether or not a sample comes from a normal distribution.

This type of test is useful for determining whether or not a given dataset comes from a normal distribution, which is a used in many statistical tests including regression, ANOVA, t-tests, and many others.

We can easily perform a Cramer-Von Mises test using the cvm.test() function from the goftest package in R.

The following example shows how to use this function in practice.

Example 1: Cramer-Von Mises Test on Normal Data

The following code shows how to perform a Cramer-Von Mises test on a dataset with a sample size n=100:

ARABPSYCHOLOGY.COM

```
library(goftest)

#make this example reproducible
set.seed(0)

#create dataset of 100 random values generated from a normal distribution
data <- rnorm(100)

#perform Cramer-Von Mises test for normality
cvm.test(data, 'pnorm')

Cramer-von Mises test of goodness-of-fit
Null hypothesis: Normal distribution
Parameters assumed to be fixed

data: data
omega2 = 0.078666, p-value = 0.7007
```

The p-value of the test turns out to be 0.7007.

Since this value is not less than .05, we can assume the sample data comes from a population that is normally distributed.

This result shouldn't be surprising since we generated

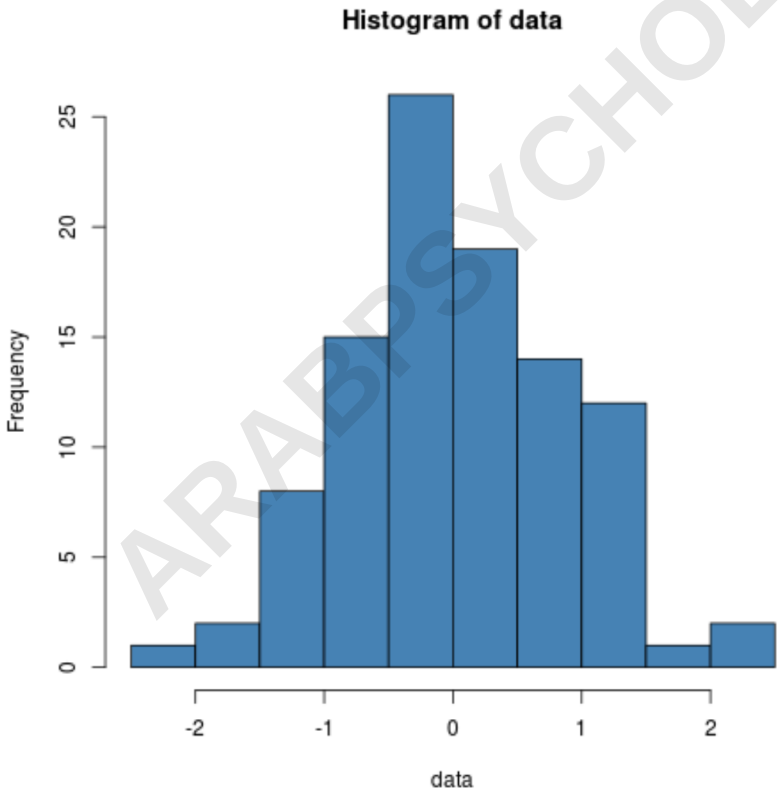
ARABPSYCHOLOGY.COM

the sample data using the rnorm() function, which generates random values from a .

Related:A Guide to dnorm, pnorm, qnorm, and rnorm in R

We can also produce a histogram to visually verify that the sample data is normally distributed:

hist(data, col='steelblue')



ARABPSYCHOLOGY.COM

We can see that the distribution is fairly bell-shaped with one peak in the center of the distribution, which is

typical of data that is normally distributed.

Example 2: Cramer-Von Mises Test on Non-Normal Data

The following code shows how to perform a Cramer-Von Mises test on a dataset with a sample size of 100 in which the values are randomly generated from a **Poisson distribution**:

```
library(goftest)

#make this example reproducible
set.seed(0)

#create dataset of 100 random values generated from a Poisson distribution
data <- rpois(n=100, lambda=3)

#perform Cramer-Von Mises test for normality
cvm.test(data, 'pnorm')

Cramer-von Mises test of goodness-of-fit
Null hypothesis: Normal distribution
Parameters assumed to be fixed

data: data
omega2 = 27.96, p-value < 2.2e-16
```

ARABPSYCHOLOGY.COM

The p-value of the test turns out to be extremely small.

Since this value is less than .05, we have sufficient evidence to say that the sample data does *not* come from a population that is normally distributed.

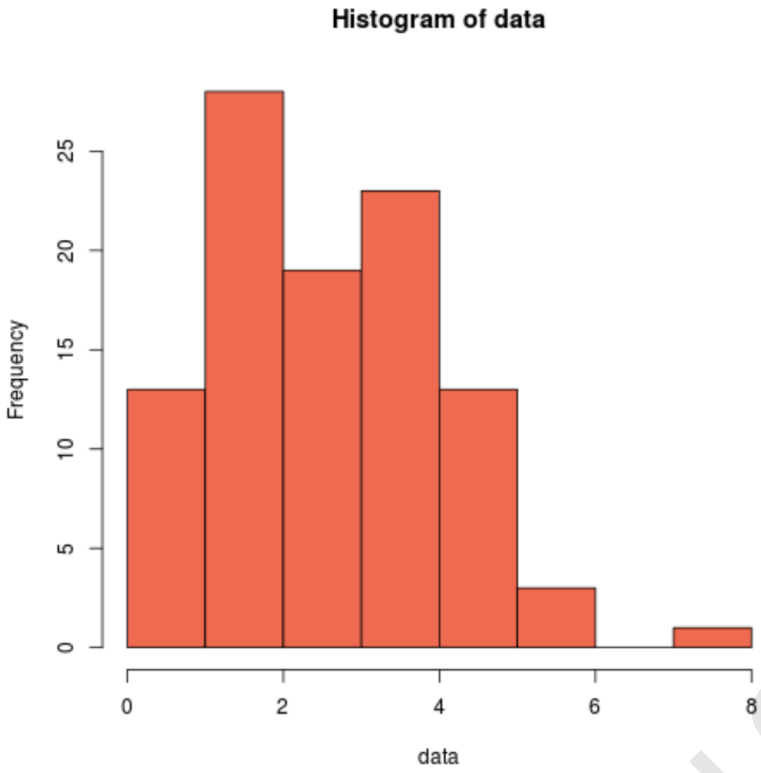
This result shouldn't be surprising since we generated the sample data using the rpois() function, which generates random values from a Poisson distribution.

Related:A Guide to dpois, ppois, qpois, and rpois in R

We can also produce a histogram to visually see that the sample data is not normally distributed:

```
hist(data, col='coral2')
```

ARABPSYCHOLOGY.COM



ARABPSYCHOLOGY.COM

We can see that the distribution is and doesn't have the typical "bell-shape" associated with a normal distribution.

Thus, our histogram matches the results of the Cramer-Von Mises test and confirms that our sample data does not come from a normal distribution.

What to Do with Non-Normal Data

If a given dataset is *not* normally distributed, we can often perform one of the following transformations to make it more normal:

1. Log Transformation: Transform the response variable from y to log(y).

2. Square Root Transformation: Transform the response variable from y to √y.

3. Cube Root Transformation: Transform the response variable from y to y1/3.

By performing these transformations, the response variable typically becomes closer to normally distributed.

Refer to this tutorial to see how to perform these transformations in practice.

The following tutorials explain how to perform other normality tests in R:

How to Conduct an Anderson-Darling Test in R

ARABPSYCHOLOGY.COM