

How can a Shapiro-Wilk Test be performed in Python?

Authored by
stats writer

April 19, 2024

RECOMMENDED CITATION

stats writer (2024). *How can a Shapiro-Wilk Test be performed in Python?*.

PSYCHOLOGICAL SCALES. Retrieved from <https://scales.arabpsychology.com/?p=136836>

The Shapiro-Wilk test is a statistical test used to determine if a given dataset follows a normal distribution. This test can be performed in Python by using the "shapiro" function from the "scipy.stats" library. The function takes in the dataset as an input and returns two values: the W statistic and the p-value. The W statistic measures the deviation of the dataset from a normal distribution, while the p-value indicates the likelihood of the dataset being normally distributed. A p-value less than 0.05 suggests that the dataset is not normally distributed. By performing the Shapiro-Wilk test in Python, one can easily determine the normality of a dataset, which is crucial in many statistical analyses.

Perform a Shapiro-Wilk Test in Python

The Shapiro-Wilk test is a test of normality. It is used to determine whether or not a sample comes from a normal distribution.

To perform a Shapiro-Wilk test in Python we can use the scipy.stats.shapiro() function, which takes on the following syntax:

```
scipy.stats.shapiro(x)
```

where:

x: An array of sample data.

This function returns a test statistic and a corresponding p-value.

If the p-value is below a certain significance level, then

we have sufficient evidence to say that the sample data does not come from a normal distribution.

This tutorial shows a couple examples of how to use this function in practice.

Example 1: Shapiro-Wilk Test on Normally Distributed Data

Suppose we have the following sample data:

```
from numpy.random import seed
from numpy.random import randn

#set seed (e.g. make this example reproducible)
seed(0)

#generate dataset of 100 random values that follow a
standard normal distribution
data = randn(100)
```

The following code shows how to perform a Shapiro-Wilk test on this sample of 100 data values to determine if it came from a normal distribution:

```
from scipy.stats import shapiro

#perform Shapiro-Wilk test
```

```
shapiro(data)
```

```
ShapiroResult(statistic=0.9926937818527222,  
pvalue=0.8689165711402893)
```

From the output we can see that the test statistic is 0.9927 and the corresponding p-value is 0.8689.

Since the p-value is not less than .05, we fail to reject the null hypothesis. We do not have sufficient evidence to say that the sample data does not come from a normal distribution.

This result shouldn't be surprising since we generated the sample data using the `randn()` function, which generates random values that follow a standard normal distribution.

Example 2: Shapiro-Wilk Test on Non-Normally Distributed Data

```
from numpy.random import seed  
from numpy.random import poisson  
  
#set seed (e.g. make this example reproducible)  
seed(0)
```

```
#generate dataset of 100 values that follow a Poisson  
distribution with mean=5  
data = poisson(5, 100)
```

The following code shows how to perform a Shapiro-Wilk test on this sample of 100 data values to determine if it came from a normal distribution:

```
from scipy.stats import shapiro
```

```
#perform Shapiro-Wilk test  
shapiro(data)
```

```
ShapiroResult(statistic=0.9581913948059082,  
pvalue=0.002994443289935589)
```

From the output we can see that the test statistic is 0.9582 and the corresponding p-value is 0.00299.

Since the p-value is less than .05, we reject the null hypothesis. We have sufficient evidence to say that the sample data does not come from a normal distribution.

This result also shouldn't be surprising since we generated the sample data using the `poisson()` function,

which generates random values that follow a **Poisson distribution**.

The following tutorials explain how to perform other normality tests in various statistical software:

[How to Perform a Shapiro-Wilk Test in R](#)

[How to Perform an Anderson-Darling Test in Python](#)

[How to Perform a Kolmogorov-Smirnov Test in Python](#)

ARABPSYCHOLOGY.COM