

How can I calculate skewness and kurtosis using Python?

Authored by
stats writer

April 20, 2024

RECOMMENDED CITATION

stats writer (2024). *How can I calculate skewness and kurtosis using Python?*.
PSYCHOLOGICAL SCALES. Retrieved from <https://scales.arabpsychology.com/?p=137303>

To calculate skewness and kurtosis using Python, one can use the `scipy.stats` library which provides functions for both these measures. Skewness is a measure of the asymmetry of a distribution, while kurtosis measures the degree of peakedness or flatness of a distribution. The skewness and kurtosis functions take in an array of data as input and return the corresponding values for these measures. Alternatively, one can also use the `pandas` library which has built-in functions for calculating skewness and kurtosis. These measures can be useful in understanding the shape and characteristics of a dataset, and can be easily computed using Python for data analysis and statistical purposes.

Calculate Skewness & Kurtosis in Python

In statistics, skewness and kurtosis are two ways to measure the shape of a distribution.

Skewness is a measure of the asymmetry of a distribution. This value can be positive or negative.

A negative skew indicates that the tail is on the left side of the distribution, which extends towards more negative values. A positive skew indicates that the tail is on the right side of the distribution, which extends towards more positive values. A value of zero indicates that there is no skewness in the distribution at all, meaning the distribution is perfectly symmetrical.

Kurtosis is a measure of whether or not a distribution is heavy-tailed or light-tailed relative to a normal distribution.

The kurtosis of a normal distribution is 3. If a given distribution has a kurtosis less than 3, it is said to be *platykurtic*, which means it tends to produce fewer and less extreme outliers than the normal distribution. If a given distribution has a kurtosis greater than 3, it is said to be *leptokurtic*, which means it tends to produce more outliers than the normal distribution.

Note: Some formulas (Fisher's definition) subtract 3 from the kurtosis to make it easier to compare with the normal distribution. Using this definition, a distribution would have kurtosis greater than a normal distribution if it had a kurtosis value greater than 0.

This tutorial explains how to calculate both the skewness and kurtosis of a given dataset in Python.

Example: Skewness & Kurtosis in Python

Suppose we have the following dataset:

data =

To calculate the sample skewness and sample kurtosis of this dataset, we can use the `skew()` and `kurt()` functions from the Scipy Stats library with the following

syntax:

skew(array of values, bias=False)kurt(array of values, bias=False)

We use the argument bias=False to calculate the sample skewness and kurtosis as opposed to the population skewness and kurtosis.

Here is how to use these functions for our particular dataset:

data =

**#calculate sample skewness
skew(data, bias=False)**

0.032697

**#calculate sample kurtosis
kurtosis(data, bias=False)**

0.118157

The skewness turns out to be 0.032697 and the kurtosis turns out to be 0.118157.

This means the distribution is slightly positively skewed and the distribution has more values in the tails compared to a normal distribution.

Additional Resource: [Skewness & Kurtosis Calculator](#)

You can also calculate the skewness for a given dataset using the [Statology Skewness and Kurtosis Calculator](#), which automatically calculates both the skewness and kurtosis for a given dataset.

ARABPSYCHOLOGY.COM