

Understanding Skewness and Kurtosis: A Practical Guide with R Examples

Authored by
Mohammed loot

November 6, 2025

RECOMMENDED CITATION

Mohammed loot (2025). *Understanding Skewness and Kurtosis: A Practical Guide with R Examples*. PSYCHOLOGICAL STATISTICS. Retrieved from <https://statistics.arabpsychology.com/?p=11940>

In modern [statistics](#), analyzing and summarizing complex datasets efficiently requires robust descriptive measures. While measures of central tendency and variability are foundational, they often fail to capture the entire picture of the data's composition. To truly understand the underlying structure of a dataset, analysts must evaluate the fundamental shape and symmetry of its probability [distribution](#). This is where the powerful concepts of [skewness](#) and [kurtosis](#) become indispensable tools. Grasping these metrics is critical for validating assumptions, particularly the assumption of normality, which underpins numerous advanced statistical tests and modeling techniques.

Understanding Skewness: Measuring Asymmetry

Skewness serves as a quantitative measure of the degree of asymmetry present in a data [distribution](#). Conceptually, if a distribution is perfectly symmetrical--such as the theoretical standard [Normal Distribution](#)--its skewness value will precisely equal zero. However, datasets derived from real-world phenomena rarely achieve this perfect symmetry, leading to calculated values that are either positively or negatively skewed, indicating a directional tilt in the data's mass.

The calculation of skewness typically involves standardizing the third central moment of the data, providing a numerical indicator whose sign dictates the direction of the asymmetry. Crucially, the sign of the skew corresponds to the direction of the "tail" of the distribution, which is the stretched-out portion, not the location of the majority of the data. This distinction is vital for accurate interpretation, as the tail represents the extreme values that pull the mean away from the median.

A **Negative Skew** (or left skew) is observed when the distribution's tail extends more prominently towards the left, indicating lower values. In such a scenario, the mean is typically found to be less than the median, pulled by the negative outliers.

A **Positive Skew** (or right skew) occurs when the tail stretches towards the right side of the distribution, encompassing higher positive values. This generally implies that the mean is greater than the median, influenced by extreme positive values.

A value of **Zero** signifies that the dataset is perfectly symmetrical, meaning the data points are balanced around the measure of central tendency.

By assessing [skewness](#), analysts can quickly ascertain if the concentration of data is heavily weighted toward one side of the central point. This provides immediate, valuable insight into potential data quality issues, the presence of influential outliers, or the underlying processes that govern the variable under observation, often guiding the choice of appropriate analytical methods.

Understanding Kurtosis: Measuring Tail Heaviness

While [skewness](#) addresses asymmetry, [kurtosis](#) provides a measure of how the tails of a data [distribution](#) differ from the tails of a standard [Normal Distribution](#). Essentially, kurtosis quantifies

the "tail heaviness"--the probability of observing extreme values--and the corresponding peakedness or flatness around the central mass. This metric is profoundly important in fields like finance and engineering, where the risk associated with extreme, infrequent events (outliers) must be rigorously quantified.

It is important to acknowledge that there are two conventional definitions for kurtosis: Pearson's definition, where the standard normal distribution yields a kurtosis value of 3; and Fisher's definition, which defines the normal distribution's excess kurtosis as 0. Most modern statistical software packages, including the popular R library we will utilize, commonly employ a definition standardized relative to the normal curve, often focusing on the excess kurtosis (Kurtosis - 3).

A distribution with **mesokurtic** properties has a kurtosis of 3 (or 0 excess kurtosis), matching the characteristics of the standard [Normal Distribution](#).

A distribution classified as *platykurtic* exhibits a kurtosis **less than 3** (negative excess kurtosis). These distributions are characterized by lighter tails and a flatter peak than the normal curve, suggesting they are less likely to generate extreme outliers.

A distribution labeled as *leptokurtic* possesses a kurtosis **greater than 3** (positive excess kurtosis). These distributions feature significantly heavier tails and a sharper peak than the normal distribution, indicating a higher propensity for generating outliers or extreme values.

A crucial technical detail concerning calculation methods relates to the concept of **excess kurtosis**. Formulas adhering to **Fisher's definition** subtract 3 from the raw [kurtosis](#) value. This adjustment simplifies the interpretive process: any calculated excess kurtosis value greater than 0 immediately signifies that the distribution possesses heavier tails than the Gaussian (normal) curve. Understanding this convention is vital for accurately interpreting the output from statistical software.

Practical Calculation: Calculating Skewness and Kurtosis in R

The statistical programming language [R](#) offers a robust, high-performance environment for computing these descriptive [statistics](#). While the base installation of [R](#) lacks direct, built-in functions for calculating skewness and [kurtosis](#), the widely utilized and reliable **moments** package provides the necessary computational tools. The following steps demonstrate the straightforward process of calculating these measures using a representative sample dataset.

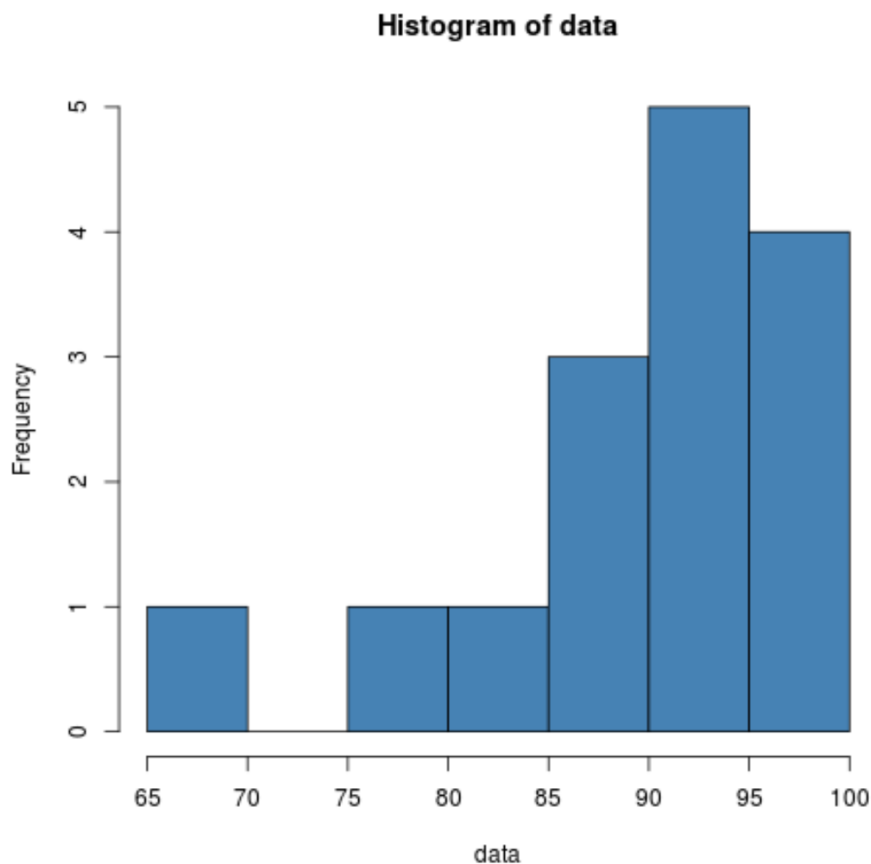
We begin by defining a hypothetical dataset. This array of fifteen observations might represent test scores, performance metrics, or any continuous variable collected during a study:

```
data = c(88, 95, 92, 97, 96, 97, 94, 86, 91, 95, 97, 88, 85, 76, 68)
```

Before proceeding to numerical calculation, best practice dictates a preliminary visual assessment

of the data. Visualizing the [distribution](#) allows analysts to form an initial hypothesis regarding its shape, symmetry, and tail behavior. We can quickly generate a simple histogram in [R](#) to inspect the data visually:

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



Interpreting the R Results and Visual Confirmation

The visual inspection of the histogram provides immediate confirmation that the bulk of the data is concentrated at the higher end of the scale. A discernible tail extends towards the lower (left) values, which is the characteristic visual signature of a **negatively skewed**, or left-skewed, [distribution](#).

To obtain the rigorous numerical values for [skewness](#) and [kurtosis](#), we must first activate the **moments** library. If this package is not yet installed in your R environment, you must execute the command `install.packages("moments")` in the R console. Once the library is active, we can leverage its dedicated functions, `skewness()` and `kurtosis()`, to process the data:

```
library(moments)
```

```
# Calculate skewness
skewness(data)

-1.391777

# Calculate kurtosis
kurtosis(data)

4.177865
```

The calculated **skewness** value is **-1.391777**. This quantitative result unequivocally confirms our visual assessment: the distribution is significantly negatively skewed. A value this far from zero indicates substantial asymmetry, confirming that the tail extends far into the negative domain. Furthermore, the computed **kurtosis** value is **4.177865**.

Since this value is greater than the mesokurtic benchmark of 3 (using Pearson's definition), the dataset is formally classified as *leptokurtic*. This result signals that the distribution exhibits heavier tails and a sharper central peak compared to a [Normal Distribution](#), implying a greater likelihood of observing extreme, outlying values.

Testing for Normality: The Jarque-Bera Test

While descriptive [statistics](#) like skewness and kurtosis provide numerical summaries, analysts often require a formal statistical test to determine if the observed characteristics significantly deviate from those expected under the assumption of a standard [Normal Distribution](#). The **moments** library again facilitates this crucial step through the `jarque.test()` function, which performs the well-known [Jarque-Bera Normality Test](#). This test is a powerful goodness-of-fit statistic specifically designed to assess whether the sample data's calculated skewness and kurtosis jointly align with the parameters of a normal distribution.

The formal statistical hypotheses underpinning the [Jarque-Bera Normality Test](#) are crucial for interpreting the results:

Null Hypothesis (H₀): The dataset possesses skewness and kurtosis values that are statistically consistent with a normal distribution.

Alternative Hypothesis (H₁): The dataset exhibits skewness and/or kurtosis values that significantly deviate from a normal distribution.

We execute the test on our sample data within [R](#) by calling the function on the data vector:

```
jarque.test(data)
```

Jarque-Bera Normality Test

data: data

JB = 5.7097, p-value = 0.05756

alternative hypothesis: greater

The output provides two key values: the Jarque-Bera statistic (JB) and the corresponding **p-value**. In this specific analysis, the resulting **p-value** is **0.05756**. To reach a statistically sound decision, we must compare this computed value against a predetermined significance level, α (alpha), conventionally established at 0.05 (or 5%).

Since the calculated **p-value** (0.05756) is not less than the chosen significance level ($\alpha = 0.05$), we must **fail to reject the null hypothesis (H?)**. Although the descriptive **skewness** and **kurtosis** metrics indicated non-zero values, the magnitude of this deviation was statistically insufficient to conclusively reject the assumption of normality at the 5% level. This highlights the necessity of formal testing, as descriptive statistics alone can sometimes suggest a deviation that lacks statistical significance.

Summary and Further Resources

Measuring skewness and kurtosis provides profoundly essential insight into the geometric shape of a data distribution, moving beyond the fundamental metrics of central tendency and variability. These calculations are foundational requirements for validating the underlying assumptions of parametric statistical tests and for accurately constructing models, especially in fields like finance or quality control where the presence and likelihood of extreme values are paramount to risk assessment.

The powerful **R** programming environment, leveraged through specialized packages such as **moments**, renders the calculation and formal testing of these measures both straightforward and highly efficient. By systematically combining visual assessment (using tools like histograms) with precise numerical calculation and formal hypothesis testing (such as the **Jarque-Bera Test**), analysts can achieve a comprehensive and nuanced understanding of their dataset's entire characteristic profile.

For those interested in delving deeper into the underlying mathematical definitions, algorithms, or exploring additional functionality offered by the package, the complete technical documentation for the **moments** library is readily available for reference:

*You can find the complete documentation for the **moments** library [here](#).*

Bonus: Online Skewness & Kurtosis Calculator

For situations requiring rapid preliminary analysis or quick verification without the immediate use of a statistical programming environment, external tools serve as excellent resources. You can swiftly calculate the skewness and kurtosis for any raw dataset using the [Statology Skewness and Kurtosis Calculator](#), which provides automated results for both descriptive **statistics** based on user-inputted data. This tool is beneficial for immediate exploratory data analysis.