

How do I Perform Multivariate Normality Tests in R?

Authored by
stats writer

December 20, 2025

RECOMMENDED CITATION

stats writer (2025). *How do I Perform Multivariate Normality Tests in R?*. PSYCHOLOGICAL SCALES. Retrieved from <https://scales.arabpsychology.com/?p=108187>

In applied statistics and data analysis, confirming whether a set of variables adheres to a joint multivariate normal distribution is a critical preliminary step. This assessment, performed via a statistical test, determines if the underlying mathematical assumptions required for many advanced multivariate techniques--such as multivariate analysis of variance (MANOVA), linear discriminant analysis, and structural equation modeling--are met. The absence of this property, often referred to as multivariate non-normality, can severely compromise the validity and reliability of subsequent inferential statistics.

The R programming language provides robust tools for conducting these complex diagnostic procedures. While packages like `mvnortest` offer classic univariate checks extended to a multivariate context (such as generalized versions of the Shapiro-Wilk or Kolmogorov-Smirnov tests), researchers often prefer dedicated tests that specifically address the joint distribution of multiple variables simultaneously. These specialized procedures, including Mardia's Test and the Energy Test, are designed to capture dependencies and distributional nuances that simple univariate checks might miss. Ensuring your data is in the correct matrix format is paramount before attempting to execute any of these sophisticated tests.

Understanding the Concept of Multivariate Normality

When dealing with statistical analysis, the assumption of normality is frequently encountered. If we are examining only a single variable, the process is straightforward: we can create a Q-Q plot to visually inspect the distribution against theoretical quantiles, or we can apply formal statistical tests like an Anderson Darling Test or a Jarque-Bera Test. These methods are tailored to assess the distribution along a single dimension.

However, statistical reality often involves multiple, correlated variables. In this scenario, assessing normality requires more than just confirming that each variable individually follows a normal distribution (univariate normality). We must determine whether the joint distribution of all variables together conforms to multivariate normality. This joint property is crucial because it accounts for the covariance structure and interdependencies among the variables, which is a requirement for parametric statistical models that rely on the Maximum Likelihood Estimation (MLE) method.

Therefore, when we aim to determine whether *several* variables are normally distributed as a group, we must employ a dedicated multivariate normality test. These tests assess two main components of the joint distribution: multivariate skewness and multivariate kurtosis. Failing to meet the multivariate normality assumption can lead to inflated Type I error rates or reduced statistical power in subsequent analyses, making these diagnostic checks indispensable for rigorous quantitative research.

Key Multivariate Normality Tests in R

Several powerful statistical procedures have been developed specifically for testing multivariate normality. These tests vary in their sensitivity to different types of non-normality (e.g., heavy tails vs. asymmetry) and their robustness across different sample sizes. Selecting the appropriate test often depends on the specifics of the data and the analytical goals.

The following sections will detail how to perform and interpret two of the most widely recognized and robust multivariate normality tests available within the R environment. We will focus on tests that evaluate the simultaneous distribution rather than relying solely on marginal distributions:

Mardia's Test: A classical approach that focuses on generalizing univariate skewness and kurtosis measures to multiple dimensions.

Energy Test: A modern, powerful non-parametric test based on the concept of statistical energy.

Multivariate Skew and Kurtosis Tests: Specific component tests derived from generalized moments.

Before proceeding with these tests, it is also important to consider if the dataset contains influential data points. If we need to identify potential outliers in a multivariate setting, a prerequisite step is often calculating the Mahalanobis distance, which measures the distance of a point from the mean center of the distribution, accounting for the covariance structure.

Detailed Example: Mardia's Test in R

Mardia's Test, formalized by Kanti Mardia, is perhaps the most established method for assessing whether a group of variables jointly follows a multivariate normal distribution. It operates by calculating two statistics: a generalized measure of skewness and a generalized measure of kurtosis. The test is highly sensitive to deviations in the shape and tail heaviness of the joint distribution.

The underlying hypotheses for Mardia's Test are formulated within the framework of statistical tests:

H₀ (Null Hypothesis): The variables follow a multivariate normal distribution (i.e., generalized skewness and kurtosis statistics do not significantly deviate from expected values).

H_a (Alternative Hypothesis): The variables *do not* follow a multivariate normal distribution (i.e., at least one of the generalized statistics is significant).

To execute this test in R, we typically utilize packages designed for advanced psychometric and multivariate analysis, such as QuantPsyc. The following code demonstrates the necessary steps, including the creation of a synthetic dataset drawn from a standard normal distribution to illustrate

the expected outcome of acceptance of the null hypothesis:

library(QuantPsyc)

```
# Create a synthetic dataset of three normally distributed variables (n=50)
```

```
set.seed(0)
```

```
data <- data.frame(x1 = rnorm(50),
```

```
x2 = rnorm(50),
```

```
x3 = rnorm(50))
```

```
# Perform Mardia's Multivariate normality test
```

```
mult.norm(data)$mult.test
```

```
Beta-hat kappa p-val
```

```
Skewness 1.630474 13.5872843 0.1926626
```

```
Kurtosis 13.895364 -0.7130395 0.4758213
```

Interpreting the Results of Mardia's Test

The **mult.norm()** function returns diagnostic statistics for both skewness and kurtosis. Skewness measures the symmetry of the distribution, while kurtosis assesses the peakedness and tail weight compared to the theoretical normal distribution. In the output provided, we receive two distinct p-values, one corresponding to the multivariate skewness statistic (Beta-hat) and one for the multivariate kurtosis statistic.

For both measures, the interpretation follows standard hypothesis testing procedures. We compare the calculated p-value against a predefined significance level, typically $\alpha = 0.05$. In the example above, the p-value for Skewness is 0.1926626, and the p-value for Kurtosis is 0.4758213. Since both these p-values are considerably greater than 0.05, we lack sufficient statistical evidence to reject the null hypothesis (H_0).

The conclusion drawn is that the three variables in our dataset--x1, x2, and x3--do not significantly deviate from a multivariate normal distribution based on Mardia's criteria. This finding confirms that the assumptions necessary for various parametric multivariate analyses are likely met, allowing the researcher to proceed with confidence. If either p-value were less than 0.05, it would indicate a significant departure from normality, potentially necessitating data transformation or the use of non-parametric methods.

Detailed Example: The Energy Test in R

While Mardia's test is highly effective, it is based on sample moments (like skewness and kurtosis) and can sometimes be less powerful for certain non-normal distributions, particularly those that are not easily characterized by moments alone. The **Energy Test**, or Energy Test of Multivariate Normality, offers an alternative, highly powerful approach. This test is based on the concept of 'statistical energy,' which measures the distance between the empirical distribution of the sample data and the theoretical multivariate normal distribution.

The Energy Test is particularly appealing because it is non-parametric and consistent against all alternatives to multivariate normality, meaning it can detect a broader range of deviations than moment-based tests. The null and alternative hypotheses remain consistent with the goal of multivariate assessment:

H0 (Null Hypothesis): The variables follow a multivariate normal distribution.

Ha (Alternative Hypothesis): The variables *do not* follow a multivariate normal distribution.

We perform this powerful statistical test in R using the dedicated `energy` package. This package provides the `mvnorm.etest()` function, which often relies on permutation or bootstrapping methods to accurately calculate the p-value, especially crucial when dealing with complex distributions or smaller sample sizes, as shown below:

library(energy)

```
# Create the identical synthetic dataset for comparison
```

```
set.seed(0)
```

```
data <- data.frame(x1 = rnorm(50),  
x2 = rnorm(50),  
x3 = rnorm(50))
```

```
# Perform Multivariate normality test using the Energy Test
```

```
mvnorm.etest(data, R=100)
```

Energy test of multivariate normality: estimated parameters

data: x, sample size 50, dimension 3, replicates 100

E-statistic = 0.90923, p-value = 0.31

Interpreting the Energy Test Results and Bootstrapping

The output from `mvnorm.etest()` provides the E-statistic (the measure of statistical energy distance) and the associated p-value. In this specific example, the resulting p-value of the test is **0.31**. Similar to Mardia's test, since 0.31 is not less than the conventional significance level of $\alpha = 0.05$, we do not reject the null hypothesis of multivariate normality.

This consistent finding across both Mardia's Test and the Energy Test strengthens the confidence that the data originates from a multivariate normal distribution. Had the p-value been below the threshold (e.g., 0.01), it would indicate strong evidence against normality, suggesting the joint distribution is significantly non-Gaussian.

It is important to note the argument `R=100` specified in the function call. This argument dictates the number of bootstrapped replicates used to estimate the p-value, which is often necessary because the exact null distribution of the E-statistic is challenging to derive analytically. For datasets with smaller sample sizes, or when greater accuracy is required, researchers may increase this number (e.g., `R=500` or `R=1000`) to produce a more robust and reliable estimate of the test statistic and its corresponding p-value. Higher replicate counts increase computational time but reduce the Monte Carlo error associated with the p-value estimation.

The Role of Generalized Skewness and Kurtosis

Multivariate normality is defined not just by the marginal distributions being normal, but by the combined structure respecting specific moment properties. Specifically, multivariate skewness and kurtosis are generalized counterparts of the univariate measures, designed to capture asymmetry and tail structure in multidimensional space.

Generalized skewness measures the overall asymmetry of the cloud of data points relative to its center, considering all variable interactions. A high level of multivariate skewness suggests that the data points are not distributed symmetrically around the mean vector, which is a key characteristic of the normal distribution. Mardia's coefficient of multivariate skewness, $b_{\{1, p\}}$, transforms this complex measure into a tractable statistical test.

Generalized kurtosis measures how concentrated the data is in the center and in the tails, relative to the theoretical multivariate normal distribution. A high positive kurtosis (leptokurtic) indicates heavy tails and sharp peak, suggesting the presence of outliers or influential observations, while negative kurtosis (platykurtic) indicates light tails. The simultaneous testing of both skewness and kurtosis provides a comprehensive assessment against different types of distributional violations.

Addressing Non-Normality and Further Diagnostics

If a multivariate normality test yields a significant result (i.e., rejection of H_0), it signals that the data deviates significantly from the Gaussian ideal. The immediate next step involves diagnostics to understand the nature of this deviation. Researchers must first check for and handle multivariate outliers, often identified using the Mahalanobis distance, as influential observations can severely inflate skewness and kurtosis statistics.

If outliers are not the primary cause, or if removing them does not resolve the issue, data transformation techniques might be necessary. Common transformations, such as log transformations or square root transformations, can sometimes normalize individual variables and improve the overall multivariate normality. However, transformations should be applied cautiously, as they can complicate the interpretation of the model parameters.

Alternatively, if transformation is not feasible or desirable, researchers may need to pivot to non-parametric or robust statistical methods that do not rely on the assumption of multivariate normality. Techniques like bootstrapping for parameter estimation or using robust regression methods are increasingly viable solutions when working with non-normal multivariate data in R.

Conclusion on Multivariate Testing in R

The ability to accurately test for multivariate normality is essential for sound multivariate statistical inference. The R environment, through packages like `QuantPsyc` and `energy`, provides specialized functions that go far beyond simple checks of marginal distributions. Mardia's Test offers a classic, moment-based assessment based on generalized skewness and kurtosis, while the Energy Test provides a powerful, non-parametric alternative.

By employing both visual diagnostics (like Q-Q plots for marginal distributions) and formal statistical procedures, researchers can confidently determine whether their data meets the stringent assumptions required for complex parametric models. Understanding the output of these tests--especially the p-values relative to the significance level--is key to making informed decisions about whether to proceed with standard analyses or opt for robust alternatives.

For those seeking deeper insight into related univariate and graphical methods, the following resources provide additional details on foundational diagnostic techniques:

[How to Create & Interpret a Q-Q Plot in R](#)

[How to Conduct an Anderson-Darling Test in R](#)

[How to Conduct a Jarque-Bera Test in R](#)

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