

# How to Perform Mood's Median Test in R: A Step-by-Step Guide

Authored by  
**stats writer**

March 14, 2026

## RECOMMENDED CITATION

stats writer (2026). *How to Perform Mood's Median Test in R: A Step-by-Step Guide*. PSYCHOLOGICAL SCALES. Retrieved from <https://scales.arabpsychology.com/?p=135630>

In the realm of quantitative research, the **Mood's Median Test** stands as a pivotal **non-parametric** statistical procedure. It is specifically designed to evaluate whether the **medians** of two or more independent samples exhibit a statistically significant difference. This test is frequently employed by data scientists and researchers as a robust alternative to parametric methods, such as the **t-test** or Analysis of Variance (ANOVA), particularly when the underlying data violates the strict assumption of a **normal distribution**. By focusing on the median rather than the mean, this test provides a measure of central tendency that is less sensitive to extreme outliers, making it an essential tool for analyzing skewed datasets or small sample sizes within the **R** programming environment.

## Perform Mood's Median Test in R

### Understanding the Foundations of Mood's Median Test

The **Mood's Median Test** operates by testing the **null hypothesis** that the medians of the populations from which the samples are drawn are identical. Unlike parametric tests that rely on the calculation of means and variances, this non-parametric approach transforms continuous data into categorical counts based on their position relative to the grand median of all combined observations. This methodology ensures that the test remains valid even when the data is ordinal or when the interval data is heavily skewed, providing a versatile solution for various experimental designs across multiple scientific disciplines.

Conceptually, the test constructs a **contingency table** that categorizes each observation into one of two groups: those that fall above the grand median and those that fall at or below it. By applying a **chi-squared distribution** analysis to this table, the researcher can determine if the distribution of values relative to the median differs significantly across the independent groups. This focus on binary classification relative to a central threshold allows the **Mood's Median Test** to maintain high integrity in scenarios where traditional assumptions of homoscedasticity or normality are entirely absent.

Furthermore, the **Mood's Median Test** is particularly advantageous when dealing with datasets that contain influential **outliers** that would otherwise disproportionately pull a sample mean. Because the median is a resistant statistic, the presence of a few extreme values does not drastically alter the outcome of the test. Researchers often prefer this method over the **Mann-Whitney U test** when the shapes of the distributions for the compared groups are fundamentally different, as the latter test assumes that the distributions possess a similar spread and shape.

### Implementing the Test via the Coin Library

To execute this analysis efficiently in **R**, the **coin library** is considered the most authoritative and

comprehensive resource. The **median\_test** function within this package offers a modernized implementation of the Brown-Mood procedure, allowing for flexible formula-based inputs that align with standard **R** modeling syntax. By utilizing the **coin package**, practitioners gain access to asymptotic and permutation-based inference methods, which enhance the reliability of the resulting **p-value** across diverse sample constraints.

The standard syntax for performing the test involves specifying a relationship between a continuous response variable and a categorical grouping variable. The general structure of the command is as follows:

**median\_test(response ~ group, data)**

In this expression, the **response** represents a vector containing the numerical values of interest, while the **group** represents a factor or vector that defines the independent categories being compared. The **data** parameter points to the **data frame** where these variables are stored, ensuring that the function can correctly map the values to their respective groups during the calculation of the grand **median** and subsequent contingency counts.

Precision in defining the **data frame** is crucial for the successful execution of the **median\_test**. Users must ensure that the grouping variable is correctly identified as a factor to prevent errors during the internal partitioning of the response values. Moreover, the **coin library** provides detailed diagnostic outputs, including Z-scores and **p-values**, which allow for a nuanced interpretation of whether the observed differences are due to random chance or reflect a genuine divergence in the central tendencies of the populations.

## Example Case: Analyzing Educational Outcomes

Consider a pedagogical study where an educator aims to determine if two distinct instructional strategies result in significantly different performance levels on a standardized examination. In this experimental design, the teacher randomly allocates twenty students into two equal groups, with each group utilizing a different studying method for a duration of two weeks. Following the intervention, all students undergo the same assessment to generate a set of test scores that can be statistically analyzed to detect differences in the **median** outcomes.

The teacher selects the **Mood's Median Test** because the sample size is relatively small and there is no guarantee that the exam scores follow a **normal distribution**. By adopting a **non-parametric** approach, the researcher ensures that the findings are robust against potential violations of parametric assumptions. The analysis begins with the systematic preparation of the dataset within the **R** environment, structuring the scores and their corresponding labels into a coherent format for processing.

This case study serves as a practical demonstration of how statistical theory is translated into actionable insights. By comparing the medians rather than the averages, the educator can gain a clearer understanding of the "typical" student's performance under each method, thereby avoiding the distortions caused by exceptionally high or low individual scores. The following sections detail the computational steps required to transform this raw educational data into a definitive statistical conclusion.

## Step 1: Constructing the Data Frame in R

The first technical phase of the analysis involves the creation of a **data frame** to store the experimental observations. In **R**, we utilize the **rep** function to generate labels for the two studying methods and combine them with the vector of exam scores. This structured approach ensures that each score is correctly associated with its respective group, facilitating accurate group-wise comparisons during the subsequent testing phase.

### # Create the dataset for analysis

```
method = rep(c('method1', 'method2'), each=10)
```

```
score = c(75, 77, 78, 83, 83, 85, 89, 90, 91, 97, 77, 80, 84, 84, 85, 90, 92, 92, 94, 95)
```

```
examData = data.frame(method, score)
```

```
# Display the structure of the data frame
```

```
examData
```

```
method score
```

```
1 method1 75
```

```
2 method1 77
```

```
3 method1 78
```

```
4 method1 83
```

```
5 method1 83
```

```
6 method1 85
```

```
7 method1 89
```

```
8 method1 90
```

```
9 method1 91
```

```
10 method1 97
```

```
11 method2 77
```

```
12 method2 80
```

```
13 method2 84
```

```
14 method2 84
```

```
15 method2 85
```

```
16 method2 90
```

17 method2 92  
18 method2 92  
19 method2 94  
20 method2 95

Upon reviewing the **data frame**, we observe that the scores for 'method1' range from 75 to 97, while 'method2' scores range from 77 to 95. While a visual inspection might suggest some overlap, informal observation is insufficient for scientific rigor. The **Mood's Median Test** will provide the necessary objective framework to determine if the discrepancies between these two sets of values are large enough to be considered statistically significant at a chosen **significance level**.

Ensuring the data is tidy and correctly formatted is a prerequisite for any statistical analysis in **R**. By verifying the **data frame** contents before proceeding, the researcher can identify any potential entry errors or missing values that might compromise the validity of the test. Once the data integrity is confirmed, the environment is prepared for the execution of the **median\_test** function from the **coin library**.

## Step 2: Executing and Interpreting the Median Test

With the data properly structured, the next phase involves loading the necessary computational tools and executing the **Mood's Median Test**. The **library(coin)** command initializes the environment, granting access to the **median\_test** function. This function will calculate the grand median of the combined scores and then determine how many scores from each studying method fall above or below that central value.

### # Load the required coin library

```
library(coin)
```

```
# Execute Mood's Median Test
```

```
median_test(score~method, data = examData)
```

```
# Output of the statistical test
```

```
Asymptotic Two-Sample Brown-Mood Median Test
```

```
data: score by method (method1, method2)
```

```
Z = -0.43809, p-value = 0.6613
```

```
alternative hypothesis: true mu is not equal to 0
```

The output reveals a **p-value** of **0.6613**. In statistical inference, this value represents the probability of observing the data--or something more extreme--given that the **null hypothesis** of equal

medians is true. Since this **p-value** is significantly higher than the standard **significance level** of 0.05, we fail to reject the null hypothesis. Consequently, we conclude that there is no statistically significant evidence to suggest a difference in the median exam scores between the two studying methods.

This result implies that, based on the current sample, neither studying method can be claimed as superior to the other in terms of central performance. It is important to note that failing to reject the null hypothesis does not prove that the medians are identical; rather, it indicates that the data does not provide enough evidence to conclude they are different. Such findings are valuable in educational research, as they may suggest that other factors, such as student motivation or prior knowledge, may have a greater impact on outcomes than the specific studying method used.

### Refining the Analysis: Handling Observations at the Median

One critical technical detail of the **Mood's Median Test** is how the algorithm handles observations that are exactly equal to the grand **median**. By default, the **median\_test** function in the **coin package** assigns a score of 0 to these observations, effectively placing them in the "at or below median" category. However, researchers may wish to adjust this behavior to see if it influences the **p-value**, especially in datasets with many tied values.

The **mid.score** argument allows the user to specify alternative values for these median-equivalent observations, such as 0.5 or 1. This flexibility is particularly useful when the researcher wants to be more conservative or when the distribution of data points around the median is highly dense. Adjusting the **mid.score** can provide a sensitivity analysis, ensuring that the final conclusion is not overly dependent on the arbitrary classification of values that land exactly on the threshold.

In the following example, we re-run the **Mood's Median Test** while assigning a value of 0.5 to observations that match the grand median. This effectively splits the influence of these points between the two categorical bins in the **contingency table**.

#### # Perform the test with an adjusted mid.score

```
median_test(score~method, mid.score="0.5", data = examData)
```

```
# Updated output
```

```
Asymptotic Two-Sample Brown-Mood Median Test
```

```
data: score by method (method1, method2)
```

```
Z = -0.45947, p-value = 0.6459
```

```
alternative hypothesis: true mu is not equal to 0
```

With the **mid.score** adjusted to 0.5, the **p-value** shifts slightly to **0.6459**. Despite this minor

change, the overall statistical conclusion remains the same: the **p-value** still exceeds 0.05, and we continue to fail to reject the **null hypothesis**. This consistency reinforces the reliability of the initial finding and demonstrates how **R** provides the tools necessary for a thorough and robust **non-parametric** analysis.

## Summary and Best Practices for Non-Parametric Testing

The **Mood's Median Test** serves as a fundamental component of the statistician's toolkit, offering a reliable path to inference when data conditions are less than ideal. Its ability to compare independent groups without requiring **normal distribution** or equal variances makes it indispensable for real-world data analysis, where outliers and skewness are common. By utilizing the **coin library** in **R**, researchers can ensure their results are derived from a high-quality, modern implementation of this classic test.

When performing the **Mood's Median Test**, it is essential to remember that while it is robust against **outliers**, it generally possesses less **statistical power** than the **Mann-Whitney U test**. Therefore, it is most appropriate when the focus is strictly on the **median** or when the distributional shapes of the groups being compared are significantly different. Analysts should always report the test statistic, the degrees of freedom (if applicable), and the **p-value** to provide a complete picture of the evidence.

In conclusion, the integration of **non-parametric** tests into **R** workflows enhances the integrity of scientific research. Whether evaluating educational techniques, medical treatments, or economic trends, the **Mood's Median Test** provides a clear and effective way to compare central tendencies across groups. By following the structured steps of data preparation, execution, and sensitivity testing via **mid.score**, researchers can confidently communicate their findings with statistical rigor and clarity.