Learn How to Perform Mood’s Median Test in R for Comparing Group Medians

Authored by

Mohammed looti

November 8, 2025

The comparison of central tendency across independent groups is a fundamental task in statistical analysis. When the data cannot satisfy the strict assumptions of parametric tests, such as normality or homogeneity of variance, statisticians often turn to robust, non-parametric methods. Among these, the **Mood's Median Test**, also known as the Brown-Mood Median Test, stands out as a powerful tool. This test is specifically designed to assess whether the medians of two or more independent samples are significantly different. Unlike the Analysis of Variance (ANOVA), which focuses on means, Mood's test focuses exclusively on the median, making it highly resistant to outliers and skewness often found in real-world data distributions.

The core principle behind the **Mood's Median Test** involves calculating the overall median of all observations combined, regardless of group membership. Once the grand median is established, the data is transformed into a contingency table. Each observation is classified into one of four categories: (1) below the grand median and in Group A, (2) above the grand median and in Group A, (3) below the grand median and in Group B, and (4) above the grand median and in Group B (and so on, for multiple groups). Essentially, the test checks whether the proportion of scores above and below the grand median is consistent across all groups. A significant deviation from this expectation suggests that the population medians are indeed different.

Performing this test efficiently requires specialized statistical software. The environment of R offers numerous packages, and for accurate non-parametric testing, the coin library provides the necessary functionality. Utilizing standard functions within R, such as those provided by the **coin** package, ensures that researchers can apply these robust methods without needing to manually construct complex contingency tables or calculate test statistics, thereby streamlining the process of data analysis and interpretation.

Understanding Mood's Median Test

The **Mood's Median Test** is a foundational element in non-parametric statistics, serving as an alternative to the one-way ANOVA when distributional assumptions are violated. It is particularly useful when dealing with ordinal data or data where the assumption of normally distributed errors is untenable. The test operates by partitioning the data based on the combined overall median. The resulting counts are then analyzed using a chi-squared test, which determines if the observed distribution of scores (above or below the overall median) across the groups significantly deviates from the distribution expected if all groups shared the same population median. This methodology makes the test highly effective for detecting shifts in the center of the distribution that might be missed by tests focused solely on means.

The underlying null hypothesis of the Mood's Median Test states that the medians of all populations from which the samples were drawn are equal. Conversely, the alternative hypothesis suggests that at least two of the population medians differ. Given its reliance on median values

rather than means, the test is inherently less sensitive to extreme outliers, providing a more stable measure of central tendency for skewed datasets. Understanding this trade-off--sacrificing some statistical power compared to parametric tests for increased robustness--is crucial when selecting the appropriate analysis technique for a given research question.

While highly valuable, applying the test correctly requires ensuring that the samples are truly independent and that the measurement scale is at least ordinal. When these conditions are met, the test provides a clear, interpretable result regarding the equality of central location. Its primary limitation stems from its focus purely on the median; it does not utilize the full magnitude of the scores, potentially leading to a slight reduction in power compared to parametric tests like ANOVA or even other rank-based non-parametric tests like the Kruskal-Wallis H test, especially when the data is nearly normal. Nonetheless, for data that is severely skewed or contains clear outliers, the **Mood's Median Test** remains an excellent choice for a robust comparison of group centers.

The Role of the coin Package in R

To execute the **Mood's Median Test** within the R statistical environment, we rely on the powerful **coin** library (Conditional Inference Procedures). This package is designed to provide implementations of a wide range of non-parametric procedures, focusing on permutation and asymptotic tests derived from conditional inference frameworks. The primary function we utilize for this specific test is `median_test()`. The use of a dedicated package like **coin** simplifies the coding process significantly, as it handles the complex calculations of the grand median, the contingency table construction, and the final chi-squared or Z-statistic calculation internally, allowing the user to focus on data input and result interpretation.

The general syntax for the `median_test` function is straightforward and follows R's standard formula notation, which is intuitive for most R users: `median_test(response~group, data)`. This formula structure clearly defines the relationship being tested: the distribution of the `response` variable is being evaluated conditional upon the different levels of the `group` variable. The simplicity of this command belies the statistical rigor applied behind the scenes, offering both asymptotic approximations and exact permutation results depending on the sample size and specific test parameters chosen.

The required arguments for the `median_test` function are explicitly defined to ensure clarity in the analysis setup. These arguments map directly to the structure of the data frame being analyzed:

response: This is a vector containing the numerical values of the outcome variable (e.g., test scores, reaction times, etc.).
group: This is a vector that defines the independent categories or groups whose medians are being compared (e.g., treatment groups, methods used).
data: This specifies the data frame object in R that contains both the response and group vectors,

ensuring the function can correctly access and map the observations.

By structuring the analysis using this function, researchers gain a high degree of confidence that the non-parametric assumptions underlying the **Mood's Median Test** are correctly implemented, providing a robust and valid inferential statement regarding the population medians. The next section illustrates the application of this function using a common pedagogical example involving educational testing methods.

Practical Example: Comparing Study Methods

Consider a practical scenario in educational research where a teacher seeks to empirically determine if two distinct studying methodologies yield different outcomes in student performance. The teacher randomly assigns twenty students--ten to use Method 1 and ten to use Method 2--to control for potential confounding variables. After a standardized two-week intervention period, all students take the same comprehensive examination. Given the potential for non-normal score distributions often associated with educational data (e.g., ceiling or floor effects), the teacher wisely chooses to employ the Mood's Median Test to compare the central tendency of the exam scores between the two groups. This choice ensures that the analysis is robust even if the score distributions are skewed or contain mild anomalies.

The initial crucial step in R is the preparation and structuring of the dataset. This involves creating a data frame that clearly delineates the independent variable (the study method) and the dependent variable (the exam score). Proper data structuring ensures that the statistical function can correctly identify which scores belong to which comparison group. The following code demonstrates the creation of the necessary data frame, which aggregates the scores obtained by students under both methods into a single, organized structure ready for statistical processing.

Step 1: Create the data frame.

```
#create data
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)

#view data
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

Once the `examData` frame is constructed, we have a clear representation of twenty observations, where the `method` column serves as the grouping factor and the `score` column is the response variable. This organized structure is paramount for subsequent statistical testing, setting the stage for the application of the `median_test()` function from the **coin** package to formally assess whether Method 1 and Method 2 lead to significantly different median exam performances.

Executing and Interpreting the Test Results

With the data prepared, the next logical step is to load the necessary coin package and execute the **Mood's Median Test**. This involves specifying the relationship between the exam scores and the study methods using the formula notation defined earlier. The test will calculate the overall median score across all twenty students and then determine if the distribution of scores relative to that median is statistically independent of the assigned study method. If the medians truly differ, we would expect a greater proportion of high scores (above the grand median) to cluster in the group associated with the superior method.

Step 2: Perform Mood's Median Test.

#load the coin library
library(coin)

#perform Mood's Median Test

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

#output
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 provides the results of the Asymptotic Two-Sample Brown-Mood Median Test. The critical value for interpretation is the resulting p-value, which quantifies the probability of observing the current data (or more extreme data) if the null hypothesis (that the medians are equal) were true. In this example, the p-value is reported as **0.6613**. Standard statistical practice typically compares this value to a significance level ($alpha$) of 0.05. Since 0.6613 is substantially greater than 0.05, we lack sufficient evidence to reject the null hypothesis.

The conclusion drawn from this analysis is that, based on the collected data, there is no statistically significant difference in the median exam scores between students utilizing Method 1 and students utilizing Method 2. While there may be slight numerical differences in the sample medians, these differences are likely attributable to random sampling variability rather than a genuine effect of the study methods. Therefore, the teacher cannot confidently conclude that one study method is superior to the other in terms of improving the central tendency of student performance.

Adjusting for Ties: The mid.score Argument

A common complexity in non-parametric testing arises when observations are exactly equal to the calculated median, leading to "ties." The way these tied observations are handled can subtly affect the calculated test statistic and the resulting p-value. By default, the `median_test()` function in the coin package assigns a score of 0 to any observation that falls exactly on the overall grand median. This approach effectively excludes tied observations from contributing to the "above" or "below" counts used in the contingency table calculation.

However, the **coin** library provides flexibility through the `mid.score` argument, allowing the user to specify alternative handling for these tied values. For instance, assigning a value of 0.5 to tied observations is a standard approach that distributes the tie equally between the 'above' and 'below' categories. This adjustment ensures that observations exactly equal to the grand median are partially included in the analysis, which some statisticians argue provides a slightly more accurate estimate of the true population effect, especially when the number of ties is large. Other options, such as setting `mid.score` to 1, would treat all tied scores as being "above" the median.

The following code demonstrates how to modify the test to utilize the **0.5** assignment for

observations equal to the median. Although the resulting interpretation often remains the same, understanding the effect of the mid.score argument is essential for rigorous statistical practice and sensitivity analysis, particularly when working with discrete or heavily rounded data where ties are common.

#perform Mood's Median Test with mid.score adjustment
median_test(score~method, mid.score="0.5", data = examData)

#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 00

In this specific instance, adjusting the mid.score to 0.5 results in a slightly lower Z-statistic (-0.45947 vs. -0.43809) and a slightly lower p-value (0.6459 vs. 0.6613). Crucially, the substantive conclusion remains unchanged: we still fail to reject the null hypothesis. This demonstrates the stability of the test, but highlights the importance of documentation and justification regarding the handling of tied data points when reporting non-parametric results.

When to Use Mood's Median Test (Assumptions and Alternatives)

The decision to utilize the Mood's Median Test should be guided by specific characteristics of the data and the research question. The primary assumption is that the samples are **independent** and that the measurement variable is continuous or at least ordinal. Furthermore, while the test does not assume normality, it implicitly assumes that the shape and spread (variance) of the distributions across the groups are relatively similar, or at least that any differences in central tendency are the primary feature of interest. If the distributions differ drastically in spread (heteroscedasticity) without a change in median, the test might still detect a difference, though the interpretation becomes more complex.

A key strength of this test is its robustness against severe violations of normality and the presence of outliers. Because it only classifies observations as being relative to the grand median, extreme values do not exert the same disproportionate influence they would in mean-based tests like ANOVA or the standard t-test. This makes it an ideal preliminary test or a definitive test for fields such as psychology or ecological studies where data often exhibit high skewness.

However, if the data is interval or ratio scale and the assumptions of normality and homogeneity of variance are reasonably met, the parametric one-way ANOVA is generally preferred due to its greater statistical power. If the data violates parametric assumptions but you wish to retain more

information about the magnitude of the ranks, the **Kruskal-Wallis H Test** (for three or more groups) or the **Mann-Whitney U Test** (for two groups) are superior non-parametric alternatives, as they use the rank order of the data rather than simply the dichotomy of "above/below" the median. Choosing the correct test requires careful consideration of the data's distribution characteristics, the measurement level, and the specific parameter (mean, median, or rank) that the researcher wishes to compare across populations.