

How to Easily Perform a Mann-Whitney U Test in SAS

Authored by
stats writer

December 1, 2025

RECOMMENDED CITATION

stats writer (2025). *How to Easily Perform a Mann-Whitney U Test in SAS*.

PSYCHOLOGICAL SCALES. Retrieved from <https://scales.arabpsychology.com/?p=103379>

Mastering the Mann-Whitney U Test within the Statistical Analysis System (**SAS**) environment is essential for researchers dealing with non-normally distributed data. While the process is statistically robust, its execution in SAS is surprisingly streamlined. The fundamental preparatory step involves meticulous organization of your raw data, ensuring it is correctly categorized into two distinct groups and properly formatted into a SAS data set. Once this data preparation is complete, the powerful PROC NPAR1WAY procedure handles the heavy lifting.

This procedure is specifically designed for nonparametric comparisons between groups and will swiftly calculate the necessary statistics, including the U statistic, the z-score approximation, and most critically, the p-value. These outputs are the foundation upon which you determine whether the observed differences between your two comparison groups achieve statistical significance. Understanding the logic behind this test and the precise SAS implementation allows for accurate and reliable hypothesis testing.

1. Introduction to the Mann-Whitney U Test

The Mann-Whitney U test, often referred to interchangeably as the **Wilcoxon rank-sum test**, serves a crucial role in comparative statistics. This test is employed when researchers need to assess differences between two independent samples but cannot satisfy the stringent assumptions required by parametric tests, such as the independence of observations and, most importantly, the assumption of a normal distribution for the dependent variable within each group.

It is classified as a nonparametric procedure because it relies not on the actual raw scores of the observations, but rather on the ranks of the data when combined across both groups. By ranking the data, the test is less sensitive to outliers and extreme deviations from normality, making it highly robust for ordinal data or continuous data suspected of being highly skewed. This makes it an indispensable tool, particularly when dealing with smaller sample sizes (e.g., typically $n < 30$ per group), where confirming normality is statistically challenging.

Historically and practically, the Mann-Whitney U test stands as the primary nonparametric alternative to the two-sample independent t-test. While the t-test compares the means of two populations, the Mann-Whitney U test fundamentally compares the distributions of the two populations. Specifically, it tests the median difference, evaluating whether a randomly selected observation from one population is likely to be larger or smaller than a randomly selected observation from the second population. This tutorial will walk through the exact methodology required to execute this powerful analysis using **SAS** software.

2. When to Apply the Mann-Whitney U Test

Choosing the correct statistical test hinges upon understanding the underlying characteristics of

your data and your research question. The decision to use the Mann-Whitney U test is typically driven by two key factors: the nature of the independent variable and the distributional properties of the dependent variable. The independent variable must be categorical, defining exactly two groups (e.g., Treatment vs. Control, Male vs. Female, Site A vs. Site B). Crucially, these two groups must be independent of one another.

The most compelling reason for selecting this test, however, relates to the dependent variable. If preliminary data analysis, such as graphical inspections (histograms, Q-Q plots) or formal tests (Shapiro-Wilk or Kolmogorov-Smirnov tests), indicates that the dependent variable scores are not distributed according to a normal distribution, or if the variance between the two groups is severely unequal, the assumptions of the parametric t-test are violated. Using the t-test under these conditions can lead to inflated Type I error rates or reduced statistical power.

Therefore, when faced with violations of normality or when working with data that is inherently ordinal (ranked data), the Mann-Whitney U test provides a reliable mechanism for hypothesis testing. It is a robust tool that sacrifices a small amount of statistical power relative to the t-test (if t-test assumptions are met), in exchange for far greater flexibility and fewer restrictive assumptions regarding the population distribution. Researchers must always confirm these distributional properties before committing to either a parametric or nonparametric approach.

3. Setting Up the Experimental Scenario

To illustrate the application of this test in **SAS**, let us consider a classic experimental scenario. Suppose a team of automotive researchers is investigating the efficacy of a new fuel additive designed to improve fuel efficiency. The primary research question is whether applying this specific fuel treatment leads to a statistically significant change in the average miles per gallon (mpg) achieved by standard vehicle models. This involves a comparative study between treated and untreated vehicles.

To rigorously test this, the researchers conduct a controlled experiment using a small sample size-- a total of 24 identical vehicles. They randomly assign 12 cars to the experimental group, which receives the specialized fuel treatment, and 12 cars to the control group, which operates without the additive (the untreated group). The researchers then meticulously measure the miles per gallon (mpg) for each of the 24 cars under standardized driving conditions. The use of a small sample ($n=12$ per group) immediately raises concerns about definitively establishing normal distribution, motivating the choice of a nonparametric procedure.

The collected data, which pairs the continuous dependent variable (mpg) with the categorical independent variable (treatment group status), is presented in the structure below. This structure clearly shows the outcome variable and the binary grouping variable necessary for the Mann-Whitney U Test. Given the small sample sizes and the suspected lack of normality, the researchers

wisely choose the Mann-Whitney U test to determine if the difference in mpg between the treated and untreated cars is statistically meaningful.

Treated	Untreated
24	20
25	23
21	21
22	25
23	18
18	17
17	18
28	24
24	20
27	24
21	23
23	19

4. Preparing Data for SAS Analysis

The first and most critical technical step in executing any analysis in **SAS** is the creation and definition of the data set. For the Mann-Whitney U test, the data must be structured in a "long format," meaning that you require two primary variables: one variable to identify the group (the categorical independent variable) and one variable containing the outcome measurements (the continuous dependent variable, mpg). This structure allows SAS procedures, particularly PROC NPAR1WAY, to correctly identify which observations belong to which population.

We use the **DATA** step in SAS to establish the dataset named `mpg_data`. Within this step, the **INPUT** statement defines the two variables: `group`, which must be defined as a character variable (indicated by the dollar sign `$`) to hold the categorical labels ('treated' or 'untreated'), and `mpg`, which holds the numerical miles-per-gallon values. Following the variable definitions, the **DATALINES** statement signals to SAS that the raw data immediately follows, row by row, until the semicolon and the **RUN** statement are encountered.

The following code block meticulously inputs the 24 observations collected during the experiment, ensuring each mpg score is correctly paired with its corresponding treatment status. This foundational step of accurate data entry prevents errors in subsequent statistical computations and is paramount for valid results:

```
/*create dataset: Step 1 - Define and Input Data*/
```

```
data mpg_data;  
input group $ mpg;  
datalines;  
treated 24  
treated 25  
treated 21  
treated 22  
treated 23  
treated 18  
treated 17  
treated 28  
treated 24  
treated 27  
treated 21  
treated 23  
untreated 20  
untreated 23  
untreated 21  
untreated 25  
untreated 18  
untreated 17  
untreated 18  
untreated 24  
untreated 20  
untreated 24  
untreated 23  
untreated 19  
;  
run;
```

5. Executing the Test using PROC NPAR1WAY

Once the dataset is properly loaded into the SAS environment, the statistical analysis is conducted using the PROC NPAR1WAY procedure. This procedure is the standard SAS utility for conducting one-way analyses involving nonparametric methods, including the Wilcoxon (Mann-Whitney U) test, the Kruskal-Wallis test, and others. The syntax required is straightforward yet powerful, enabling SAS to perform the necessary ranking and calculation of the U statistic.

The primary command, `proc npar1way`, specifies the dataset to be used (`data=mpg_data`). Crucially, we must include the **WILCOXON** option. While the procedure defaults to certain nonparametric tests, explicitly including `WILCOXON` ensures that the output focuses on the statistics relevant to the Mann-Whitney U test (which, as noted, is statistically identical to the Wilcoxon rank-sum test for two independent samples). Without this option, SAS might produce only general output rather than the specific two-sample test results needed.

Following the primary PROC statement, two essential sub-statements are required: the **CLASS** statement and the **VAR** statement. The `CLASS` statement identifies the grouping variable--in this case, `group`--which defines the two populations being compared (treated vs. untreated). The `VAR` statement identifies the dependent variable--`mpg`--which contains the measurements being ranked and analyzed. The combination of these commands instructs SAS exactly how to perform the rank-sum comparison:

```
/*perform Mann Whitney U test: Step 2 - Call the Procedure*/
```

```
proc npar1way data=mpg_data wilcoxon;
```

```
class group;
```

```
var mpg;
```

```
run;
```

6. Interpreting the SAS Output and Key Statistics

Upon execution, the `PROC NPAR1WAY` procedure generates a detailed output, including descriptive statistics, calculated rank sums, and the final test statistics. The most critical section for the Mann-Whitney U test interpretation is the "Wilcoxon Two-Sample Test" table. This table summarizes the core findings, linking the calculated test statistic back to the final probability assessment. The output typically includes the sums of ranks for both groups and the corresponding U statistic.

The U statistic itself is a measure of the degree of separation between the two samples based on ranks. A smaller U value generally suggests that the two groups overlap more significantly in their distributions. However, for practical decision-making, researchers focus almost exclusively on the associated p-value. SAS calculates the exact p-value for small samples and provides an asymptotic approximation (based on the normal distribution) for larger samples. The output displayed below provides this crucial information:

The NPAR1WAY Procedure

Wilcoxon Scores (Rank Sums) for Variable mpg Classified by Variable group					
group	N	Sum of Scores	Expected Under H0	Std Dev Under H0	Mean Score
treated	12	172.0	150.0	17.203387	14.333333
untreat	12	128.0	150.0	17.203387	10.666667
Average scores were used for ties.					

Wilcoxon Two-Sample Test					
Statistic	Z	Pr > Z	Pr > Z	t Approximation	
				Pr > Z	Pr > Z
172.0000	1.2498	0.1057	0.2114	0.1120	0.2240
Z includes a continuity correction of 0.5.					

Kruskal-Wallis Test		
Chi-Square	DF	Pr > ChiSq
1.6354	1	0.2010

From the results generated by SAS, we observe the specific value for the two-sided p-value of the test, which is determined to be **0.2114**. This single statistic is the gateway to hypothesis testing and dictates the final interpretation of the experimental results. Before drawing conclusions, it is necessary to formally state the hypotheses being tested by the Mann-Whitney U Test.

7. Drawing Conclusions from the P-Value

Statistical hypothesis testing involves comparing the observed results against the expected results under the assumption that no effect exists. For the Mann-Whitney U test, the hypotheses are framed around the medians or, more accurately, the probability that an observation from one group is larger than an observation from the other group. The formal hypotheses are defined as follows:

H0: The two population distributions are identical, meaning the two populations have the same median. This is the Null Hypothesis, suggesting no effect of the fuel treatment.

HA: The two population distributions are not identical. This implies that the medians of the two populations are different, indicating a statistically significant effect of the fuel treatment on mpg.

The decision to reject or fail to reject the Null Hypothesis hinges on comparing the calculated p-value (0.2114) to the pre-established level of significance (α). In social sciences and many engineering applications, the standard significance level is $\alpha = 0.05$. If the p-value is less

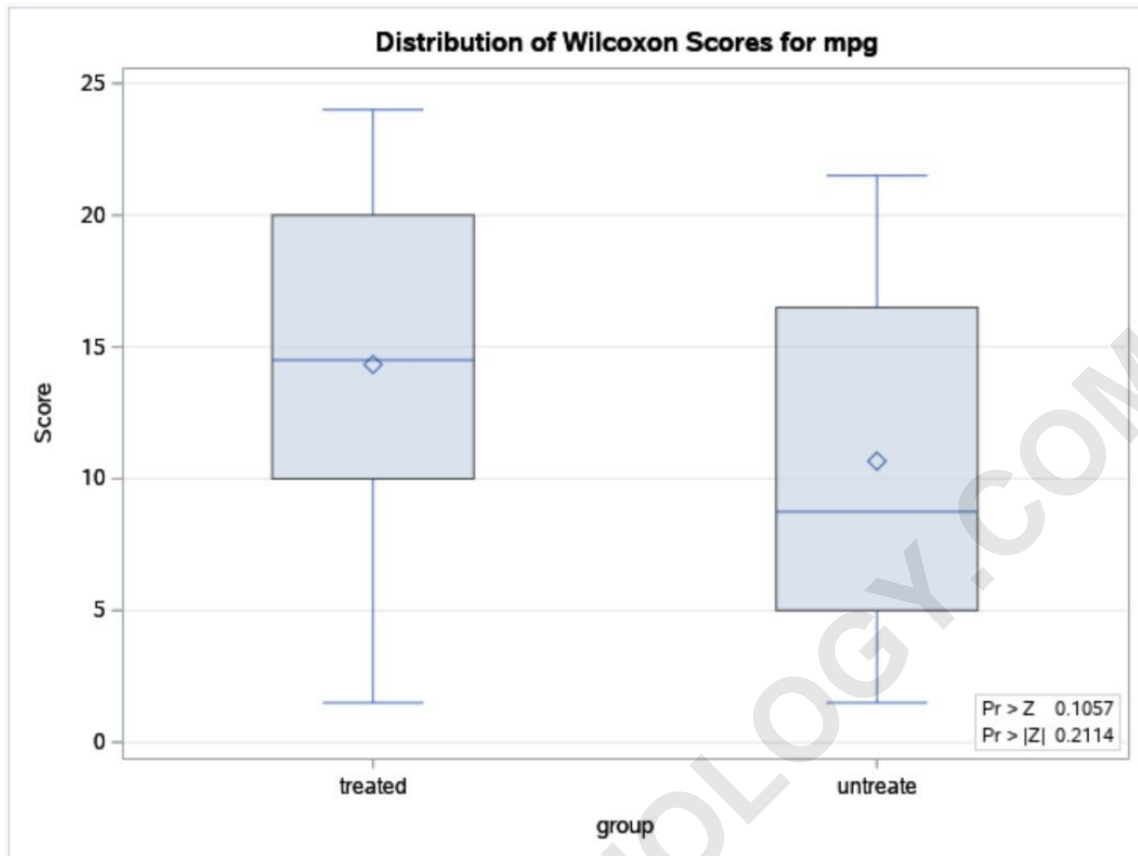
than α , we reject H_0 ; otherwise, we fail to reject H_0 .

In this specific example, since the p-value (**0.2114**) is substantially greater than the significance level of 0.05, we must consequently fail to reject the Null Hypothesis. This statistical outcome signifies that the researchers do not possess sufficient statistical evidence, at the 5% significance level, to conclude that the fuel treatment caused a statistically significant difference in the miles per gallon achieved by the cars. While there might be a numerical difference in the sample medians, the variation within the samples is large enough that the observed difference could easily be attributed to random chance.

8. Visualizing Data Distributions with SAS Boxplots

While the numerical output from the Mann-Whitney U test provides the formal statistical conclusion, visualization of the data distributions is crucial for full understanding. PROC NPAR1WAY often generates graphical output automatically, including boxplots, which offer a clear comparison of the central tendency, spread, and skewness of the two groups without requiring additional coding.

The boxplot visualizes the five-number summary (minimum, first quartile, median, third quartile, and maximum) for both the treated and untreated groups. This allows the researcher to quickly assess overlap and identify potential differences in dispersion that the rank test accounts for. The visualization below shows the distributions for the mpg values in both groups:



From a purely descriptive perspective derived from the boxplot, the cars that received the fuel treatment appear to exhibit slightly higher median mpg values, and potentially a slightly greater range, compared to the untreated vehicles. However, the Mann-Whitney U test result provides the necessary inferential context. The overlapping distributions shown in the plot, combined with the high p-value (0.2114), definitively confirm that the apparent difference observed visually is not strong enough to be declared statistically significant. The benefit of combining both the inferential test and the descriptive visualization is clear: one tells you if the difference is real (statistically), and the other shows you what that difference looks like (visually).

9. Summary and Further Nonparametric Procedures

The successful execution of the Mann-Whitney U test in **SAS**, leveraging the **PROC NPAR1WAY** procedure, underscores the accessibility of nonparametric statistics within this powerful software environment. This methodology provides a robust alternative when the stringent requirements of parametric tests, such as the two-sample independent t-test, are not met--a common occurrence with small sample sizes or highly skewed data.

The core takeaway from this exercise is the importance of linking data structure, procedure selection, and output interpretation. Accurate data preparation (Step 1) is non-negotiable, and

selecting the correct procedure options (e.g., `WILCOXON` in `PROC NPAR1WAY`) ensures the statistical validity of the final result. Most importantly, the interpretation of the resulting p-value against the predefined alpha level provides the clear, definitive answer to the original research question.

For researchers seeking to explore other comparative methods when normality assumptions are violated, SAS offers a suite of procedures designed for nonparametric analysis. These tests extend the capability demonstrated here to scenarios involving matched pairs or more than two independent groups. Mastering these fundamental steps sets the stage for more complex statistical modeling:

The following tutorials explain how to perform other common statistical tests in SAS:

ARABPSYCHOLOGY.COM