

A Step-by-Step Guide to the Wilcoxon Signed-Rank Test in Stata

Authored by
Mohammed looti

November 8, 2025

RECOMMENDED CITATION

Mohammed looti (2025). *A Step-by-Step Guide to the Wilcoxon Signed-Rank Test in Stata*. PSYCHOLOGICAL STATISTICS. Retrieved from <https://statistics.arabpsychology.com/?p=13652>

The [Wilcoxon Signed Rank Test](#) is a fundamental and robust [non-parametric](#) statistical procedure. It serves as the primary alternative to the traditional [paired t-test](#) when analyzing dependent data. This test is meticulously employed by researchers to determine if a statistically significant difference exists between the median values of two related samples, typically involving repeated measurements taken from the same subjects or matched pairs. Crucially, the Wilcoxon test is selected when the underlying assumption of the paired t-test--that the differences between the paired observations follow a normal distribution--cannot be met, or when the sample size is too small to reliably assess normality. By focusing on the magnitude and direction of the differences between pairs, it provides a powerful and resilient method for comparing dependent groups even under non-ideal distributional conditions.

This comprehensive, expert-level guide is designed to walk you through the precise steps and commands necessary to successfully conduct a Wilcoxon Signed Rank Test using the powerful statistical software package, [Stata](#). We will navigate the entire process, starting from the initial data loading and review, progressing through the execution of the primary command, and culminating in the accurate interpretation and formal reporting of the results. Our goal is to ensure clarity, accuracy, and best practices are followed at every stage of the analysis.

The Mechanics and Assumptions of the Wilcoxon Test

The Wilcoxon Signed Rank Test is an indispensable tool across diverse applied statistics fields, including medicine, psychology, and engineering, where assessing the impact of an intervention through pre- and post-measurements is routine. Unlike the parametric [paired t-test](#), which mandates that the differences be normally distributed and the data be measured on an interval or ratio scale, the Wilcoxon test is far less restrictive. It only requires that the data be measured on an [ordinal scale](#) or higher. This enhanced flexibility makes it essential when dealing with datasets that are small or exhibit significant skewness or kurtosis, conditions that fundamentally violate the necessary assumptions for traditional [parametric testing](#).

The fundamental principle underpinning the test involves a rigorous three-step mechanism. First, the differences between each pair of observations are calculated. Second, the absolute values of these differences are ranked from smallest to largest. Third, the ranks are summed separately based on the sign (positive or negative) of the original difference. The resulting sum of ranks forms the basis for the test statistic, which quantifies the magnitude and consistency of the change observed across the pairs.

The primary objective of the test is to evaluate the [null hypothesis](#) (H_0) that the median difference between the pairs in the population is zero. If the calculated test statistic suggests that the observed differences are highly unlikely to occur by chance alone, we reject the null hypothesis. Rejecting H_0 implies that there is a statistically significant shift in the

population median following the intervention or between the two related conditions. For large samples, the test statistic is often converted to a standardized Z-score. In modern statistical software like [Stata](#), this decision is typically made by comparing the calculated [p-value](#) to a predetermined significance level (α), ensuring that valid conclusions can be drawn even when strict parametric assumptions are violated.

Selecting and Structuring the Data for Analysis

To provide a clear, practical demonstration of the Wilcoxon Signed Rank Test, we will utilize the standard Stata example dataset known as **fuel**. This dataset is ideally structured for this procedure because it contains paired observations: specifically, the miles per gallon (mpg) measurements for 12 distinct automobiles, recorded at two different time points--once **before** and once **after** a specialized fuel treatment was administered. The central goal of our forthcoming analysis is to statistically ascertain whether this treatment induced a significant change in the vehicles' mpg performance.

This scenario perfectly exemplifies a classic dependent samples design, where each individual vehicle effectively serves as its own control. The Wilcoxon test is therefore highly suitable for comparing these two dependent measurements: mpg before treatment versus mpg after treatment. Successful execution of the test is predicated on ensuring the data is correctly structured and loaded within the [Stata](#) environment.

A dataset appropriate for this test must contain at least two variables representing the paired measurements. It is absolutely essential that the observation in row *i* for the first variable corresponds directly and accurately to the observation in row *i* for the second variable. Any misalignment in this pairing will render the subsequent analysis meaningless. The following steps will guide you through accessing this standard example data and preparing it for the statistical command.

Step 1: Loading and Reviewing the Dataset in Stata

The initial required step is to load the necessary data file into the active Stata session. This is achieved through the use of the fundamental `use` command, followed by the specific path or URL of the dataset. For the purposes of this tutorial, we will directly access the **fuel** dataset from the Stata Press online repository, which ensures that we are utilizing the standard example data referenced in official documentation. Execute the command below precisely as shown in the Command box within the Stata interface:

```
use http://www.stata-press.com/data/r13/fuel
```

Once the data has been successfully loaded into memory, it is critically important to visually

inspect the raw observations. This preliminary review, or data audit, verifies that the variables have been correctly interpreted by [Stata](#) and confirms the expected pairing structure. To display the entire dataset within the Results window, utilize the simple but powerful `list` command:

list

```
. use http://www.stata-press.com/data/r13/fuel
```

```
. list
```

	mpg1	mpg2
1.	20	24
2.	23	25
3.	21	21
4.	25	22
5.	18	23
6.	17	18
7.	18	17
8.	24	28
9.	20	24
10.	24	27
11.	23	21
12.	19	23

By reviewing the listed data, we can clearly identify the two critical variables involved in our dependent comparison: **mpg1**, which records the miles per gallon measurement for each of the 12 cars **before** the fuel treatment intervention, and **mpg2**, which records the miles per gallon measurement for the same 12 cars **after** the fuel treatment was applied. Each row distinctly represents a single vehicle, confirming the necessary paired structure required for accurate execution of the [Wilcoxon Signed Rank Test](#).

Step 2: Executing the Wilcoxon Signed Rank Test Command

With the dataset prepared and the paired structure verified, the next logical step is to execute the statistical procedure itself. In Stata, the [Wilcoxon Signed Rank Test](#) is initiated using the `signrank` command. This command requires the specification of the two variables containing the paired observations. The syntax is straightforward and highly intuitive: type `signrank`, followed by the variable representing the 'before' measurement, an equals sign, and then the variable representing the 'after' measurement. While the order can be logically reversed (e.g., `signrank mpg2 = mpg1`), the interpretation of the resulting test statistic, specifically its sign (positive or negative), will change

accordingly.

To formally test the [null hypothesis](#) that there is no difference in the median mpg before and after the treatment, we input the following command into the Stata console:

```
signrank mpg1 = mpg2
```

Upon execution, Stata automatically performs the necessary calculations: determining the differences between paired observations, ranking the absolute values of these differences, summing the positive and negative ranks (often denoted T^+ and T^-), and ultimately calculating the test statistic and the associated [p-value](#). This calculation is the core of the analysis, allowing us to objectively determine whether the observed magnitude of change in mpg is likely attributable to random chance or if it represents a genuine, statistically significant effect of the specialized fuel treatment.

```
. signrank mpg1=mpg2
```

Wilcoxon signed-rank test

sign	obs	sum ranks	expected
positive	3	13.5	38.5
negative	8	63.5	38.5
zero	1	1	1
all	12	78	78

```
unadjusted variance      162.50
adjustment for ties      -1.63
adjustment for zeros     -0.25
```

```
adjusted variance       160.63
```

```
Ho: mpg1 = mpg2
```

```
z = -1.973
Prob > |z| = 0.0485
Exact Prob = 0.0479
```

Interpreting the Stata Output and Decision Making

The output generated by the `signrank` command is meticulously structured, providing several critical pieces of information required for robust statistical inference. The upper section of the output typically summarizes the descriptive statistics of the ranks, detailing the number of positive and negative differences (e.g., how many cars showed an increase versus a decrease in mpg) and

the corresponding sum of ranks for each category. This section provides initial insight into the direction of the observed change.

The lower section, labeled **Hypothesis test**, contains the formal results that dictate our decision regarding the [null hypothesis](#). The output clearly restates the premise of the analysis, $H_0: \text{mpg1} = \text{mpg2}$, asserting that the population median difference is zero. Stata provides the standardized test statistic, conventionally denoted as **z**, which for this specific analysis is reported as **-1.973**. This Z-score is derived from the calculated rank sums and is used to estimate the probability of observing such an extreme difference if, in fact, the null hypothesis were true.

The most decisive element for statistical decision-making is the corresponding [p-value](#), which Stata reports as **0.0485** (labeled as **Pr > |z|**). This value represents the probability of obtaining a test statistic as extreme as -1.973 under the assumption that the fuel treatment had no effect. Following conventional standards in statistical reporting, we compare this calculated p-value to the established statistical significance level (α), typically set at 0.05. Since **0.0485 is less than 0.05**, we have obtained sufficient statistical evidence to confidently reject the null hypothesis. The rejection of H_0 leads us to the formal conclusion that there is a statistically significant difference in the median miles per gallon of the vehicles before and after the administration of the specialized fuel treatment.

Step 3: Formal Reporting of the Results

The final phase of any rigorous statistical endeavor is the clear, concise, and accurate reporting of the findings. Reporting the results of a [Wilcoxon Signed Rank Test](#) requires the inclusion of several key elements: the specific test name, the context of the variables being compared, the sample size (N), and the core test statistics (the Z-score and the [p-value](#)). It is imperative to seamlessly transition from the technical statistical finding (rejecting the [null hypothesis](#)) to a practical, meaningful conclusion regarding the phenomenon under study.

A well-structured report ensures that the results are not only accurate but also accessible and comprehensible to a broad audience. The following example illustrates the formal presentation of the findings derived from the preceding Stata analysis, incorporating all necessary metrics and concluding remarks:

A Wilcoxon Signed Rank Test was conducted to assess whether a specialized fuel treatment resulted in a statistically significant change in the median miles per gallon (mpg) of 12 dependent automobiles. The analysis included **N = 12** paired observations.

The test revealed a statistically significant difference in median mpg between the pre-treatment and post-treatment conditions. The calculated Z test statistic was **z = -1.973**, with a corresponding probability of **p = 0.0485**.

Based on these results, we conclude that the fuel treatment had a statistically significant positive effect on the mpg performance of the cars included in the sample, indicating that the treatment successfully altered the distribution of mpg compared to baseline.

By meticulously following these outlined steps, researchers can confidently execute and interpret the [Wilcoxon Signed Rank Test](#) within [Stata](#), ensuring sound statistical inferences when dealing with paired data that does not conform to the strict assumptions of [parametric testing](#).