How to Perform a Friedman Test in Stata: A Step-by-Step Guide

Authored by

stats writer

March 9, 2026

Understanding the Fundamental Principles of the Friedman Test

The **Friedman test** represents a vital **non-parametric statistical test** utilized primarily to detect differences in treatments across multiple test attempts. Developed by the economist Milton Friedman, this procedure serves as the non-parametric alternative to the **one-way repeated measures ANOVA**. It is particularly useful when the assumptions of **normality** or homogeneity of variance are violated, or when the data collected is of an **ordinal** nature. By focusing on the **ranking** of data within each block or subject, the test mitigates the influence of extreme **outliers** that would otherwise compromise the validity of a parametric analysis.

In practice, the **Friedman test** is employed when a researcher has measured the same subjects under three or more distinct conditions. Because the subjects are the same across all groups, the observations are **dependent**, necessitating a test that accounts for the relatedness of the samples. The **null hypothesis** for this test posits that there is no difference between the distributions of the groups, meaning that any observed differences in ranks are due to **sampling error** or random chance. Conversely, the alternative hypothesis suggests that at least one treatment or condition results in a distribution that is significantly different from the others.

The versatility of the **Friedman test** extends across various fields, including psychology, medicine, and marketing research. For instance, a clinician might use it to compare the efficacy of four different medications on the same group of patients over several weeks. Because it does not require the data to follow a **Gaussian distribution**, it provides a robust fallback for small **sample sizes** where normality is difficult to verify. Using **Stata**, researchers can execute this test with precision, ensuring that their findings are statistically sound even when dealing with non-ideal datasets.

Comparing Non-Parametric Methods to Parametric ANOVA

When deciding between a **one-way repeated measures ANOVA** and the **Friedman test**, the primary consideration is the distribution of the **residuals**. Parametric tests require that the data follow a **normal distribution** and exhibit **sphericity**, which refers to the equality of variances of the differences between all combinations of related groups. If these rigorous assumptions are not met, the **Type I error** rate can increase significantly, leading to false-positive results. The **Friedman test** bypasses these requirements by transforming raw scores into ranks, thereby focusing on the relative position of values rather than their absolute magnitude.

One of the trade-offs in selecting **non-parametric statistics** is a potential loss of **statistical power**. If the data actually does meet the assumptions for a parametric test, the **ANOVA** is more likely to detect a true effect if one exists. However, in many real-world scenarios, particularly in clinical settings with small cohorts, the **Friedman test** is the more honest choice. It prevents the

researcher from drawing conclusions based on skewed data or **variance** inconsistencies that are common in repeated measures designs.

Furthermore, the **Friedman test** is inherently more resilient to the presence of **outliers**. In a parametric **ANOVA**, a single extreme value can drastically pull the **mean** of a group, potentially masking a real trend or creating an artificial one. Because **ranking** limits the influence of any single data point to its position in the sequence, the **Friedman test** maintains its integrity. For **Stata** users, understanding this distinction is crucial for selecting the appropriate analytical pathway for their specific research questions.

Initializing the Analysis Environment in Stata

To begin our practical walkthrough, we must first prepare the **Stata** environment. Data analysis in **Stata** is highly efficient, but it requires the user to follow a specific workflow to ensure **reproducibility**. We will utilize a classic dataset provided by **Stata Press** to demonstrate the test. This dataset, known as *t43*, is structured to show the reaction times of five distinct patients across four different drug treatments. This structure is the quintessential **repeated measures design**, as each subject (patient) is exposed to every level of the independent variable (drugs).

The first step involves loading the dataset directly from the official web repository. By using the **use** command, we can pull the data into our active memory. It is a best practice to clear any existing data in the workspace before starting a new analysis to avoid **variable** conflicts or memory errors. Once the data is loaded, it is essential to perform a preliminary inspection. Visualizing the data allows the researcher to understand the **data types** and ensure that the subjects and treatments are correctly aligned in the **wide or long format** required for the specific command being used.

After loading the data, we will use the **browse** command to open the **Data Editor**. This step is more than just a formality; it allows you to verify that there are no **missing values** that could impact the **Friedman test**. Since this test relies on complete blocks (every subject must have a score for every drug), missing data can lead to the exclusion of entire subjects, reducing your **sample size** and the overall **power** of the test.

use http://www.stata-press.com/data/r14/t43

br

	person	drug	score
1	1	1	30
2	1	2	28
3	1	3	16
4	1	4	34
5	2	1	14
6	2	2	18
7	2	3	10
8	2	4	22
9	3	1	24
10	3	2	20
11	3	3	18
12	3	4	30
13	4	1	38
14	4	2	34
15	4	3	20
16	4	4	44
17	5	1	26
18	5	2	28
19	5	3	14
20	5	4	30

Installing Necessary Extensions for Non-Parametric Testing

While **Stata** is incredibly powerful out of the box, some specific **non-parametric** procedures require external packages developed by the statistical community. To perform the **Friedman test** in a modern **Stata** environment, the **emh** (Extended Mantel-Haenszel) package is highly recommended. This package provides a generalized framework for **stratified** analysis of response variables, which includes the **Friedman test** as a specific application when using **rank transformations**.

Installing packages in **Stata** is a straightforward process thanks to the **SSC** (Statistical Software Components) archive. By executing a simple **ssc install** command, **Stata** will communicate with the servers at Boston College, download the necessary files, and integrate them into your local system. This process ensures that you have the most up-to-date version of the **algorithm**, including any bug fixes or performance enhancements contributed by the package authors. It is important to have an active internet connection during this step.

Once the **emh** package is successfully installed, it becomes part of your **Stata** library, meaning you do not need to reinstall it for future sessions. This package is favored over older commands

because of its flexibility in handling **ties** and its ability to provide **asymptotic** results. For researchers who frequently work with **categorical** or **ordinal** data, the **emh** command is a versatile addition to their statistical toolkit, bridging the gap between simple tests and complex modeling.

ssc install emh

Executing the Friedman Test via Rank Transformation

With the environment prepared and the necessary tools installed, we can now proceed to the execution of the **Friedman test**. The syntax for the **emh** command is specific but logical. In our example, the **dependent variable** is *score* (the reaction time), and the **independent variable** is *drug*. To account for the repeated measures nature of the data, we must specify the **strata**, which in this case is the *person* variable. This tells **Stata** to compare the drug scores within each individual patient.

The core of the **Friedman test** logic within this command is found in the **transformation(rank)** option. By specifying this, we instruct **Stata** to convert the raw reaction times into **ranks** within each stratum (person) before performing the **chi-squared** analysis. The **anova** option is also included to produce the relevant **test statistic**. This multifaceted command provides a comprehensive look at the data, ensuring that the **dependency** between observations is fully respected throughout the calculation.

Running this command will generate an output table that includes the **degrees of freedom**, the **chi-square statistic** (often denoted as Q), and the associated **p-value**. Unlike some manual methods of calculating the **Friedman test**, the **emh** command in **Stata** handles **ties** (where two or more measurements are identical) automatically. This ensures a higher degree of accuracy in the resulting **significance** levels, which is vital for rigorous scientific reporting.

emh score drug, strata(person) anova transformation(rank)

```
. emh score drug, strata(person) anova transformation(rank)

Extended Mantel-Haenszel (Cochran-Mantel-Haenszel) Stratified Test of Association

ANOVA (Row Mean Scores) Statistic:
Q (3) = 13.5600, P = 0.0036
Transformation: Ranks
```

Interpreting the Statistical Output and Significance

The interpretation of the **Friedman test** output focuses on two primary values: the Q-statistic and the **p-value**. In the provided example, the Q-statistic (listed as Q(3)) is **13.5600**. This value represents the standardized difference between the observed **ranks** and the ranks we would expect to see if the **null hypothesis** were true. The number in parentheses, 3, indicates the **degrees of freedom**, which is calculated as the number of groups minus one (4 drugs - 1 = 3).

The most critical component for decision-making is the **p-value**, which in our results is **0.0036**. In **hypothesis testing**, we compare this value to a pre-determined **alpha level**, typically 0.05. Since 0.0036 is considerably lower than 0.05, we have sufficient evidence to reject the **null hypothesis**. This suggests that the type of drug administered has a **statistically significant** effect on the reaction times of the patients. It is not merely a result of random variation; the drugs are performing differently.

However, it is important to note that a significant **Friedman test** result is "omnibus" in nature. This means it tells you that a difference exists somewhere among the groups, but it does not specify which particular pairs of drugs are different from each other. To determine the specific nature of these differences--for example, if Drug A is significantly faster than Drug B--researchers must perform **post-hoc analysis**. This often involves **Wilcoxon signed-rank tests** with a **Bonferroni correction** to control for the **family-wise error rate**.

Formal Reporting of Results for Academic Standards

When documenting the results of a **Friedman test** in an academic or professional report, clarity and precision are paramount. A standard report should include the **sample size**, the nature of the repeated measures, the **test statistic**, the **degrees of freedom**, and the **p-value**. Providing the **median** reaction times for each drug condition is also highly recommended, as it offers a descriptive context that complements the **non-parametric** results.

Following our **Stata** example, a formal summary might be phrased as follows: "A **Friedman test** was conducted on five individuals to investigate the influence of four distinct drug treatments on reaction time. Each participant was exposed to all four drug conditions. The analysis revealed a **statistically significant** difference in reaction times across the drug groups, $\chi 2(3) = 13.56$, $p = .0036$. These results indicate that the physiological response to the drugs varies significantly depending on the specific medication administered."

By adhering to these reporting standards, researchers ensure that their work is transparent and easily interpreted by peers. In **Stata**, the transition from raw output to a formal report is made easier by the software's clear labeling of statistics. It is always beneficial to accompany these results with a visual aid, such as a **box plot**, to illustrate the **distribution** and **median** values for

each group, providing a holistic view of the experimental outcomes.

Summary of Best Practices in Stata

Conducting the **Friedman test** in **Stata** is a robust way to handle related-samples data when parametric assumptions fail. To ensure the most accurate results, researchers should always prioritize the following steps:

Verify the **repeated measures** structure of the data, ensuring subjects are correctly identified.
Check for **missing observations**, as the **Friedman test** requires complete blocks for each subject.
Use the **emh** package for its modern handling of **rank transformations** and ties.
Always report the **chi-square statistic** along with the **p-value** for full transparency.
Follow up significant results with **post-hoc tests** to pinpoint specific group differences.

By integrating these practices into your **data analysis** workflow, you can confidently utilize **Stata** to uncover meaningful insights from complex, non-normal datasets. Whether you are a student or a seasoned statistician, the **Friedman test** remains a cornerstone of **non-parametric** inquiry, providing a reliable path to discovery when traditional methods fall short.