How to Perform a Kruskal-Wallis Test to Compare Groups

Authored by

stats writer

December 31, 2025

The Kruskal-Wallis test is a fundamental non-parametric statistical test designed to determine if there are statistically significant differences among two or more independent group medians. This method is employed when comparing multiple population groups derived from independent samples, particularly when the data is measured on an ordinal or continuous scale.

The utility of the Kruskal-Wallis test lies in its direct comparison to the one-way analysis of variance (ANOVA). However, unlike ANOVA, Kruskal-Wallis does not require the strict assumption that the data are normally distributed within each group. This robustness makes it an invaluable tool for researchers dealing with skewed distributions, small sample sizes, or unequal variances, ensuring reliable results even when classic parametric assumptions are violated.

Understanding the Kruskal-Wallis Test

A **Kruskal-Wallis test**, formally known as the Kruskal-Wallis H test, serves as a crucial tool for determining whether a statistically significant difference exists between the medians of three or more independent groups. Since it operates without relying on distributional assumptions, it is correctly classified as the non-parametric alternative to the one-way ANOVA.

This test is typically utilized in situations where preliminary analysis reveals that the assumption of normality is violated, or when the measurement scale of the response variable is strictly ordinal. By using the ranks of the data rather than the raw scores themselves, the Kruskal-Wallis procedure effectively sidesteps the need for specific distributional shapes.

Crucially, the Kruskal-Wallis test does not assume that the underlying population distributions are normal, making it far less sensitive to extreme observations, or outliers, compared to the one-way ANOVA, which can be severely impacted by these data points.

Real-World Scenarios for the Kruskal-Wallis Test

To fully grasp the application of this method, considering practical examples is essential. The Kruskal-Wallis test is ideal whenever a researcher needs to compare the central tendency of multiple groups where the data distribution is uncertain or known to be non-normal.

Example 1: Comparing Study Techniques

Imagine a scenario where a researcher seeks to evaluate the effectiveness of three distinct study techniques on exam performance. Ninety students are randomly divided into three groups of 30, with each group employing a specific technique for a month-long preparation period. Upon taking the same comprehensive exam, the researcher collects the scores. If prior evidence suggests that exam scores resulting from these techniques are not normally distributed--perhaps due to ceiling or floor effects--a Kruskal-Wallis test is the appropriate choice. This test will efficiently determine if a statistically significant difference exists among the group medians, thereby indicating whether the study technique has a genuine impact on exam results.

Example 2: Comparing Sunlight Exposure on Plant Growth

A botanist wishes to understand how varying levels of sunlight exposure affect the growth of a specific plant species. Seeds are planted in four different environmental conditions: high sunlight, medium sunlight, low sunlight, and no sunlight. After a month, the height of each plant is measured. If it is known that the distribution of plant heights for this species is highly susceptible to skewness, is non-normally distributed, and is prone to outliers (e.g., a few unusually tall or stunted plants), a parametric test would be inappropriate. The researcher would therefore conduct a Kruskal-Wallis test to determine if the median height significantly differs across the four sunlight exposure groups.

Essential Assumptions for Kruskal-Wallis

While the Kruskal-Wallis test is non-parametric and therefore less restrictive than its parametric counterpart, it still requires that certain conditions are met to ensure the validity and reliability of the results. Before proceeding with the analysis, researchers must confirm the adherence to the following assumptions:

Ordinal or Continuous Response Variable: The variable being measured (the dependent variable) must be either an ordinal variable or a continuous variable. For example, an ordinal variable could be the result of a survey using a Likert scale (e.g., rating agreement on a 5-point scale), while a continuous variable might be a physical measurement, such as weight or height.

Independence of Observations: The data points within each group must be independent of one another, and the groups themselves must be independent. This typically implies that the observations are collected from randomly sampled individuals or units, ensuring that the measurement from one individual does not influence the measurement from another. A well-designed, randomized study setup usually satisfies this condition.

Distributions Have Similar Shapes: Although the test does not assume normality, for the null hypothesis to truly test the equality of population medians, the distributions of the scores in each group must possess a similar shape. If the shapes of the distributions are drastically different across groups, the Kruskal-Wallis test becomes a test of stochastic dominance rather than strictly a test of medians.

Case Study: Evaluating Drug Effects on Pain Rating

To illustrate the computational steps of the Kruskal-Wallis test, consider a study investigating the comparative efficacy of three different drugs on reducing knee pain. A researcher enrolls 30 individuals experiencing similar chronic knee pain and randomly allocates them into three equal groups (N=10), with each group receiving either Drug 1, Drug 2, or Drug 3.

Following one month of treatment, each participant is asked to rate their current knee pain level on a standardized scale ranging from 1 to 100, where 100 signifies the most severe pain imaginable. Since pain ratings on such scales often produce non-normally distributed data, the Kruskal-Wallis test is selected for analysis. The collected pain ratings for all 30 individuals are presented in the table below:

Drug 1	Drug 2	Drug 3
78	71	57
65	66	88
63	56	58
44	40	78
50	55	65
78	31	61
70	45	62
61	66	44
50	47	48
44	42	77

Step-by-Step Execution and Hypothesis Testing

The researcher aims to determine if there is a statistically significant difference among the population medians of knee pain ratings across the three drug groups, using a standard significance level ($alpha$) of 0.05. The analysis proceeds through defining the hypotheses and calculating the test statistic.

Step 1. State the Hypotheses.

The first crucial step in any statistical analysis is to formally define the null hypothesis (H0) and the alternative hypothesis (Ha):

The null hypothesis (H0): The median knee-pain ratings across the three drug groups are equal (i.e., the drugs have no differential effect).

The alternative hypothesis (Ha): At least one of the population median knee-pain ratings is statistically different from the others (i.e., at least one drug has a unique effect).

Step 2. Perform the Kruskal-Wallis Test.

The Kruskal-Wallis Test involves ranking all observations from all groups combined, calculating the sum of the ranks for each group, and using these values to compute the H statistic. For demonstration purposes, we utilize a statistical calculator by entering the raw values provided above:

This Kruskal-Wallis Test calculator compares the medians of three or more independent samples. It is the nonparametric version of the One-Way ANOVA.

Simply enter the values for up to five samples into the cells below, then press the "Calculate" button.

Sample 1

78, 65, 63, 44, 50, 78, 70, 61, 50, 44

Sample 2

71, 66, 56, 40, 55, 31, 45, 66, 47, 42

Sample 3

57, 88, 58, 78, 65, 61, 62, 44, 48, 77

After inputting the data, the 'Calculate' function is initiated to compute the test statistic and the associated P-value:

CALCULATE

H Statistic: 3.08903

p-value: 0.21342

Analyzing the P-Value and Drawing Conclusions

Step 3. Interpret the Results.

The output from the statistical calculation yields a P-value of **0.21342**. To make a decision regarding the null hypothesis, we compare this P-value to the predetermined significance level ($alpha$) of 0.05. The fundamental rule of hypothesis testing states that if the P-value is less than $alpha$, we reject H0; otherwise, we fail to reject H0.

In this scenario, the P-value (0.21342) is significantly greater than the alpha level (0.05). Therefore, we **fail to reject the null hypothesis**. This outcome signifies that, based on the collected data and the Kruskal-Wallis analysis, there is insufficient statistical evidence to conclude that a significant difference exists between the median knee pain ratings resulting from the three tested drugs.

Statistical Software Tutorials

Performing the Kruskal-Wallis test is straightforward across various statistical software packages. The following resources provide detailed guides on implementing this non-parametric procedure using popular tools:

How to Perform a Kruskal-Wallis Test in Python
How to Perform a Kruskal-Wallis Test in SPSS

How to Perform a Kruskal-Wallis Test in SAS