

# What is the purpose and methodology of conducting Dunn's test for multiple comparisons?

Authored by  
**stats writer**

April 19, 2024

## RECOMMENDED CITATION

stats writer (2024). *What is the purpose and methodology of conducting Dunn's test for multiple comparisons?*. PSYCHOLOGICAL SCALES. Retrieved from <https://scales.arabpsychology.com/?p=137199>

Dunn's test is a statistical method used for conducting multiple comparisons between multiple groups or treatments. Its purpose is to determine if there are significant differences between the groups, and if so, which specific groups are significantly different from each other.

The methodology of Dunn's test involves comparing the ranks of the data within each group, rather than the actual values. This helps to reduce the impact of outliers and non-normality in the data. The test also takes into account the overall variability of the data, making it suitable for comparing groups with unequal sample sizes.

To conduct Dunn's test, the data is first ranked across all groups. Then, the test compares the mean ranks of each group to determine if there are significant differences. If a significant difference is found, post-hoc tests are performed to determine which specific groups are significantly different from each other.

Overall, the purpose of Dunn's test is to provide a reliable and robust method for conducting multiple comparisons, while controlling for the potential of false positive results. This makes it a valuable tool for researchers and statisticians in various fields, including medicine, social sciences, and business.

## Dunn's Test for Multiple Comparisons

**A Kruskal-Wallis test is used to determine whether or not there is a statistically significant difference between the medians of three or more independent groups. It is considered to be the non-parametric equivalent of the One-Way ANOVA.**

**If the results of a Kruskal-Wallis test are statistically significant, then it's appropriate to conduct Dunn's Test to determine exactly which groups are different.**

**Dunn's Test performs pairwise comparisons between**

each independent group and tells you which groups are statistically significantly different at some level of  $\alpha$ .

For example, suppose a researcher wants to know whether three different drugs have different effects on back pain. He recruits 30 subjects for the study and randomly assigns them to use Drug A, Drug B, or Drug C for one month and then measures their back pain at the end of the month.

The researcher can perform a Kruskal-Wallis test to determine if the median back pain is equal among the three drugs. If the p-value of the Kruskal-Wallis test is below a certain threshold, it can be said that the three drugs produce different effects.

Following this, the researcher could then perform Dunn's Test to determine *which* drugs produce statistically significant effects.

Dunn's Test: The Formula

You will likely never have to perform Dunn's Test by hand since it can be performed using statistical software (like R, Python, Stata, SPSS, etc.) but the formula to calculate the z-test statistic for the difference

between two groups is:

$$z_i = y_i / \sigma_i$$

where  $i$  is one of the 1 to  $m$  comparisons,  $y_i = W_A - W_B$  (where  $W_A$  is the average of the sum of the ranks for the  $i$ th group) and  $\sigma_i$  is calculated as:

$$\sigma_i = \sqrt{((N(N+1)/12) - (\sum T^2_s - T_s^2)/(12(N-1))) / ((1/n_A) + (1/n_B))}$$

where  $N$  is the total number of observations across all groups,  $r$  is the number of tied ranks, and  $T_s$  is the number of observations tied at the  $s$ th specific tied value.

How to Control the Family-wise Error Rate

Whenever we make multiple comparisons at once, it's important that we control the family-wise error rate. One way to do so is to adjust the p-values that results from the multiple comparisons.

There are several ways to adjust the p-values, but the two most common adjustment methods are:

## 1. The Bonferroni Adjustment

**Adjusted p-value =  $p \cdot m$**

**p:** The original p-value  
**m:** The total number of comparisons being made

## **2. The Sidak Adjustment**

**Adjusted p-value =  $1 - (1-p)^m$**

**where:**

**p:** The original p-value  
**m:** The total number of comparisons being made

**By using one of these p-value adjustments, we can dramatically reduce the probability of committing a type I error among the set of multiple comparisons.**

**[How to Perform Dunn's Test in R](#)**

**[How to Perform Dunn's Test in Python](#)**