

Which test should you use: Tukey, Bonferroni, or Scheffe?

Authored by
stats writer

April 23, 2024

RECOMMENDED CITATION

stats writer (2024). *Which test should you use: Tukey, Bonferroni, or Scheffe?*.

PSYCHOLOGICAL SCALES. Retrieved from <https://scales.arabpsychology.com/?p=138582>

When conducting statistical analysis, it is important to choose the appropriate test to accurately interpret the data and draw meaningful conclusions. When comparing multiple groups, three commonly used tests include Tukey, Bonferroni, and Scheffe.

Tukey test is suitable for comparing groups with equal sample sizes and equal variances, and it allows for all possible pairwise comparisons between groups. Bonferroni test is a more conservative approach that is appropriate for unequal sample sizes and unequal variances, and it controls for the overall type I error rate. Scheffe test is the most versatile and can be used for both equal and unequal sample sizes, as well as unequal variances, but it is less powerful than the other two tests.

Ultimately, the choice between these tests depends on the specific characteristics of the data and the research question at hand. It is important to carefully consider these factors and select the most appropriate test to ensure accurate and reliable results.

Tukey vs. Bonferroni vs. Scheffe: Which Test Should You Use?

A one-way ANOVA is used to determine whether or not there is a statistically significant difference between the means of three or more independent groups.

If the overall p-value from the ANOVA table is less than some significance level, then we have sufficient evidence to say that at least one of the means of the groups is different from the others.

However, this doesn't tell us *which* groups are different from each other. It simply tells us that not all of the group means are equal.

In order to find out exactly which groups are different from each other, we must conduct a post-hoc test that is capable of controlling the family-wise error rate.

Three of the most commonly used post-hoc tests include:

The Tukey Method The Scheffe Method The Bonferroni Method

This tutorial provides an overview of each method along with instructions on which post-hoc test to use depending on the situation.

The Tukey Method

The Tukey post-hoc test should be used when you would like to make pairwise comparisons between group means when the sample sizes for each group are equal.

If the sample sizes are *not* equal, you can use a modified version of the test known as the Tukey-Kramer test.

The term "pairwise" means we only want to compare

two group means at a time.

For example, suppose we have three groups - A, B, C.

The Tukey post-hoc test would allow us to make the following pairwise comparisons:

$$\mu A = \mu B \quad \mu A = \mu C \quad \mu B = \mu C$$

Note that for k groups, there are a total of $k(k-1)/2$ possible pairwise comparisons.

The Scheffe Method

The Scheffe post-hoc test should be used when you would like to make *all* possible contrasts between group means. This test allows you to compare more than just two means at once, unlike the Tukey post-hoc test.

The Scheffe post-hoc test would allow us to make complex comparisons such as:

$$\mu A - \mu B = \mu C - \mu D \quad \mu A + \mu D = \mu B + \mu C$$

While the Scheffe post-hoc test is the most flexible, it is also the most conservative and produces the widest

confidence intervals. This means it has the lowest statistical power and the lowest ability to detect true differences between the groups.

Note that the Scheffe post-hoc test can be used whether or not the group sample sizes are equal.

The Bonferroni Method

The Bonferroni post-hoc test should be used when you have a set of *planned* comparisons you would like to make beforehand.

For example, suppose we have three groups - A, B, C - and we know ahead of time that we're only interested in the following comparisons:

$$\mu_A = \mu_B \quad \mu_B = \mu_C$$

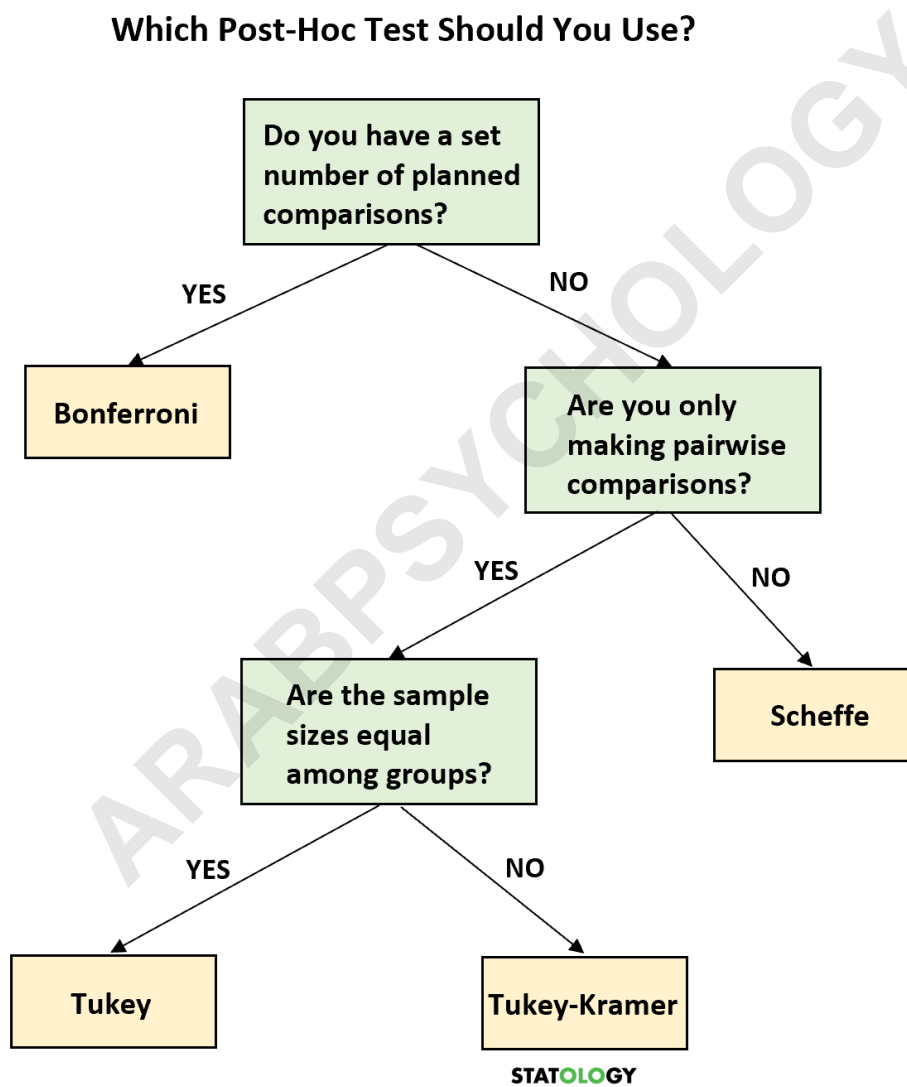
When we have a specific set of planned comparisons we'd like to make ahead of time like this, the Bonferroni post-hoc test produces the most narrow confidence intervals, which means it has the greatest ability to detect true difference between the groups of interest.

Note that the Bonferroni post-hoc test can also be used

whether or not the group sample sizes are equal.

Which Method Should You Use?

The following decision tree helps you decide which post-hoc test you should use depending on the situation:



Closing Thoughts

No matter which post-hoc test you choose to use, you should decide on it *before* conducting the experiment.

This prevents you from choosing a post-hoc test after the experiment that may be more likely to produce significant results, which is considered a dishonest practice in research.

In any case, most statistical software is capable of performing these post-hoc tests so you will rarely have to compute them by hand.