

11.1: Why ANOVA?

We've done all that we can do with t -tests; it's now time to learn a new test statistic! Introducing ANOVA (ANalysis Of VAriance).

Type I Error

You may be wondering why we do not just conduct a t -test to test our hypotheses about three or more groups. After all, we are still just looking at group mean differences. The answer is that our t -statistic formula can only handle up to two groups, one minus the other. In order to use t -tests to compare three or more means, we would have to run a series of individual group comparisons. For only three groups, we would have three t -tests: group 1 vs group 2, group 1 vs group 3, and group 2 vs group 3. This may not sound like a lot, especially with the advances in technology that have made running an analysis very fast, but it quickly scales up. With just one additional group, bringing our total to four, we would have six comparisons: group 1 vs group 2, group 1 vs group 3, group 1 vs group 4, group 2 vs group 3, group 2 vs group 4, and group 3 vs group 4. This makes for a logistical and computation nightmare for five or more groups.

So, why is that a bad thing? Statistical software could run that in a jiffy. The real issue is our probability of committing a Type I Error. Remember [Type I and Type II Errors](#)? Briefly, a Type I error is a false positive; the chance of committing a Type I error is equal to our significance level, α . This is true if we are only running a single analysis (such as a t -test with only two groups) on a single dataset. However, when we start running multiple analyses on the same dataset, we have the same Type I error rate for each analysis. The Type I error rate is cumulative, it is added together with each each new analysis. The more calculations we run, the more our Type I error rate increases. This raises the probability that we are capitalizing on random chance and rejecting a null hypothesis when we should not. ANOVA, by comparing all groups simultaneously with a single analysis, averts this issue and keeps our error rate at the α we set.

Variability

The ANOVA is a cool analysis for other reasons, as well. As the name suggests, it looks at variances (meaning, the squared average difference between each score and the mean), but it looks at more than just the variance for each group (IV level).

What's also cool about ANOVAs is that we can use them to analyze mean differences on a DV when we have more than one IV, not just three or more groups (or levels) of one IV! We'll get to that later. For now, let's learn more about what ANOVAs do with variability.

This page titled [11.1: Why ANOVA?](#) is shared under a [CC BY-NC-SA 4.0](#) license and was authored, remixed, and/or curated by [Michelle Oja](#).

- [11.4: ANOVA and Type I Error](#) by [Foster et al.](#) is licensed [CC BY-NC-SA 4.0](#). Original source: <https://irl.umsl.edu/oer/4>.