

9.5: When to NOT use the Independent Samples t-test

Let's say that you have two groups, that are independent, and you measured a quantitative variable. Sounds like the perfect time to use an independent samples t-test, right?

This is not a trick question. Yes, when you have a quantitative variable for two independent groups, the independent samples t-test is the analysis to complete.

But what if it's not?

When NOT to Use the Independent Samples t-test

When There Are More Than Two Groups

If you wanted to compare more than two groups, you *could* compare pairs of each group by conducting multiple t-tests. Other than being tedious, what do you think that we don't do that? Maybe review [Type I errors](#) to help figure out why we don't do multiple t-tests?

After reminding yourself about Type I errors, you now can see that we increase our chances of rejecting a true null hypothesis (a true null hypothesis says that the means are similar in the population) when we conduct multiple tests on the same data. But don't fret, we have another kind of statistical analysis for when we have more than two groups: An ANOVA! We'll learn about those a little later.

When the Groups are Dependent

Independent is in its name, so it might be a bit obvious, but we wouldn't conduct an independent t-test when the two samples are somehow related. This actually happens a lot, like:

- When we compare the same person Before some experience and then After that experience. This is called a pretest/posttest design.
- When we compare two people who are linked in some way. Often, this is based on a relationship (twins, a parent and child, a romantic couple), but it could also be when two participants are matched on some characteristics (say, GPA), and then the participants are randomly assigned into two conditions (one person with a high GPA is assigned to the tutoring condition and one person with a low GPA is assigned to a no-tutoring condition).

Regardless of the reason why the groups are related, if we have pairs of participants, then an Independent Samples t-test is not the appropriate test for us. But don't fret, we have another kind of statistical test for dependent samples: the Dependent Samples t-test. We'll learn about that in the next chapter!

When the Distribution Doesn't Fit the t-test Assumptions

In statistics, an assumption is some characteristic that we assume is true about our data, and our ability to use our inferential statistics accurately and correctly relies on these assumptions being true. If these assumptions are not true, then our analyses are at best ineffective (e.g. low power to detect effects) and at worst inappropriate (e.g. too many Type I errors). A detailed coverage of assumptions is beyond the scope of this course, but it is important to know that they exist for all analyses.

When the Two Standard Deviations are Very Different

Using the pooled variance to calculate the test statistic relies on an assumption known as homogeneity of variance. This is fancy statistical talk for the idea that the true standard deviation for each group is similar to the other group, and that any differences in the samples' standard deviations is due to random chance (if this sounds eerily similar to the idea of testing the null hypothesis that the true population means are equal, that's because it is exactly the same!) This notion allows us to compute a single pooled variance that uses our easily calculated degrees of freedom. If the assumption is shown to not be true, then we have to use a very complicated formula to estimate the proper degrees of freedom. There are formal tests to assess whether or not this assumption is met, but we will not discuss them here.

Many statistical programs incorporate the test of homogeneity of variance automatically and can report the results of the analysis assuming it is true or assuming it has been violated. You can easily tell which is which by the degrees of freedom: the corrected degrees of freedom (which is used when the assumption of homogeneity of variance is violated) will have decimal places. Fortunately, the independent samples *t*-test is very robust to violations of this assumption (an analysis is "robust" if it works well

even when its assumptions are not met), which is why we do not bother going through the tedious work of testing and estimating new degrees of freedom by hand.

Although it hasn't been highlighted, we've been talking about something called the Student's t-test (there's a fun story about beer and privacy about why it's called that), and that is the most common t-test and is all of the statistics textbooks. However, Welch's t-test is probably a better statistical analysis of two independent groups because it doesn't require that the two standard deviations are similar. The thing is, since no one learns about Welch's t-test, no one uses it. And since no one uses it, if you try to use it, then your supervisor or publisher or professor will be like, "Why are you using this crazy kind of t-test?" which translates to "I don't understand what you're doing, so you can't do it." If you stay in the social sciences, you are encouraged to read a little about Welch's t-test so that you can convince the supervisor/publisher/professor in your life that it's a better statistical analysis than the Student's t-tests. Standard statistical software will run it, and the interpretation is similar to the Student's t-test, so it's really about explaining how cool it is to others to be able to use it.

When the Distribution is Not Normally Distributed

Similar to the assumption of homogeneity of variance, an assumption of t-tests is that the distributions are normally distributed. With the examples that we've been using with small sample sizes, this has probably not been the case, so we are lucky that the t-test is robust.

As we will learn in the next section, when we are worried about the distribution not being normally distributed, we can use non-parametric alternatives.

This page titled [9.5: When to NOT use the Independent Samples t-test](#) is shared under a [CC BY-NC-SA 4.0](#) license and was authored, remixed, and/or curated by [Michelle Oja](#).