Calculating a Dependent Samples T-Test

At this point, you've learned three different hypothesis tests: the one-sample z-test, the one-sample t-test, and the independent-samples t-test. Well done! Now, we're going to learn the dependent-samples t-test. The main difference here is that, instead of having just one group of people measured once (the one-sample z- and t-tests) or having two groups of people being compared against each other (the independent-samples t-test), we have one group of people measured twice. This is why we call this a "dependent" test: Since we're measuring the same group of people twice, their "post-treatment" scores are *dependent on* their "baseline" scores. Let's look at an example to illustrate these ideas and to practice calculating a dependent-samples t-test.

Okay, so let's say you want to know whether watching an episode of the TV show "The Office" makes people happier. An important thing to understand about research and statistics is that you can investigate the same research question in a variety of ways. You could, for example, do an independent samples t-test sort of design here in which half of the people in your study watch "The Office" and the other half are assigned to a control condition (e.g., in which they watch "Parks and Recreation" or watch nothing at all). Then you could measure the happiness of the two groups and see if there is a significant difference using the independent-samples t-test.

Instead, however, we're going to conduct a study that uses a dependent-samples t-test. Here's how this study will work. We'll recruit a sample of, say, 10 participants. First, we'll give them a "happiness scale" to measure their baseline happiness. This represents their happiness before any psychological manipulation. Next, we'll have participants watch an episode of "The Office." Finally, we'll measure the happiness of those same 10 participants to get their post-treatment levels of happiness. Within each participant, we can see if we made a difference. This is what makes the dependent-samples t-test so powerful: In a dependent-samples t-test, *each participant acts as their own control*. Both you and your former self, 20 minutes ago are the same age, gender, race, and so on, making this a perfect comparison to truly isolate the effect of our manipulation (i.e., watching an episode of "The Office").

Okay, that's enough build up. Let's see some data! We'll conduct the study described above, using a "happiness" scale with values that range from 1 ("extremely unhappy") to 50 ("extremely happy"). Remember that, when I show you the data, you're seeing measurements on the same group of people twice. In each row, you have each participant's pair of observed values. This is why the dependent-samples t-test is sometimes called the "paired t-test" or "matched-pairs test."

Okay, here's the data:

Baseline	Post_Treatment
37	41
33	34
29	33
21	21
23	26
24	27
24	25
36	38
38	41
28	41

Now, let's take a look at the formula for the dependent-samples *t*-test and decide where to begin:

Dependent samples t-test

$$t_{\bar{x}_d} = \frac{\bar{x}_d}{s_{\bar{x}_d}}$$

Estimated standard error of the mean difference

$$s_{\bar{x}_d} = \sqrt{\frac{s_d^2}{n}}$$

Notice here that all of our subscripts have "d" in them. Here, "d" stands for "difference." Specifically, it stands for the difference between each participant's post-treatment and baseline scores, as these difference scores represent whether or not your manipulation made a difference (i.e., whether watching an episode of "The Office" actually increased participants' happiness).

So, the first step in any dependent-samples t-test problem is to calculate a new column of difference scores by subtracting each participant's post-treatment score by his or her baseline score. (Note: You can also subtract the baseline scores minus the post-treatment scores; you'll end up with the same t-test statistic with the opposite sign; this is less common but will result in the same p-value, so it doesn't matter too much.)

After doing this, you'll need to calculate the mean and the variance of the difference scores. For any dependent-samples t-test problem, you can essentially forget about the original data after computing the difference scores; everything you need for your formulas you can calculate from the difference score column.

Since we expect happiness to *increase* after watching the TV show episode, we'll subtract `Post_Treatment` from `Baseline`. This means if we get positive values, it indicates happiness increased. Let's calculate the difference scores now:

Baseline	Post_Treatment	Differences
37	41	4
33	34	1
29	33	4
21	21	0
23	26	3
24	27	3
24	25	1
36	38	2
38	41	3
28	41	13

You'll notice that most people's happiness scores increased after our manipulation. If we would've subtracted `Baseline` from `Post_Treatment`, we would end up with negative sign values for the difference score. Again, these will be equivalent, but for ease of interpretation in this example, we subtracted `Post_Treatment` from `Baseline`.

Okay, so let's calculate the mean ($\bar{x}_d$) and the sample variance (s_d^2) of the difference scores. For now, I'll just tell you that the mean of the difference scores is $\bar{x}_d = 3.40$ and the sample variance is $s_d^2 = 13.16$. (I'll show you my work for calculating sample variance as usual on the last page of this *Guide*).

Now that we have these values, let's plug them first into the formula for standard error:

$$s = \sqrt{\frac{s_d^2}{n}} = \sqrt{\frac{13.16}{10}} = 1.15$$

So, our standard error is $s_d^2 = 1.15$. Now, we can plug this, along with the mean, $\bar{x}_d$, into the formula for the dependent-samples *t*-test:

$$t_{\bar{x}_d} = \frac{\bar{x}_d}{s_{\bar{x}_d}} = \frac{3.40}{1.15} = 2.96$$

And that's it! We have our *t*-test statistic for the dependent-samples *t*-test. I'll go ahead and tell you that this *t*-test statistic is extreme enough to yield a *p*-value less than .05, meaning this is a statistically significant difference. It seems like watching an episode of "The Office" does, in fact, make people happier. :)

Calculating The Variance Of The Difference Scores

Okay, let's calculate the variance of the difference scores using are trusty formula for sample variance:

$$s^2 = \frac{\sum (x_i - \bar{x})^2}{n - 1}$$

First, we'll calculate deviations of the difference scores from the mean of the difference scores, and then we'll square those values.

Baseline	Post_Treatment	Differences	deviations	squares
37	41	4	0.6	0.36
33	34	1	-2.4	5.76
29	33	4	0.6	0.36
21	21	0	-3.4	11.56
23	26	3	-0.4	0.16
24	27	3	-0.4	0.16
24	25	1	-2.4	5.76
36	38	2	-1.4	1.96
38	41	3	-0.4	0.16
28	41	13	9.6	92.16

Next, we'll add up the values in the final column to find the Sum of Squares, $\sum (x_i - \bar{x})^2$, which comes out to 118.4.

Finally, we'll divide the Sum of Squares by $n - 1$ to find our sample variance. Dividing 118 by 9 results in a sample variance of 13.16. This is the sample variance of the difference scores, s_d^2.