Calculating an independent samples *t*-test

You now have two hypothesis tests under your belt: the one-sample *z*-test and the one-sample *t*-test. But what do you do when you have two different samples of people that you want to compare? For example, what if you want to know whether men and women differ in their average levels of anxiety? Or whether people with dogs are friendlier than people with cats? These questions cannot be answered using a one-sample *z*- or *t*-test. Instead, we'll need to use a test that is designed to compare two different groups of people. The test we use is known as an independent-samples *t*-test. Let's go through a practice problem to illustrate how the independent-samples *t*-test works.

Let's say that you are a researcher who is interested in investigating whether playing video games makes you less anxious. To test this, you design an experiment in which you recruit a sample of 30 participants. You randomly assign 15 of these participants to an experimental condition in which they are required to play 1 hour of a video game and the remaining 15 participants to a control condition in which they watch two episodes of a TV show instead. Afterwards, each participant is required to report their anxiety via a self-report scale in which 1 = "*not anxious at all*" and 14 = "*very anxious*." Let's take a look at some data below:

Games	TV
6	7
7	8
8	9
5	8
4	5
6	4
7	8
8	7
9	6
11	7
13	8
2	14
4	13
5	9
8	8

Now, let's take a look at the formula for an independent-samples *t*-test to determine how to proceed:

Independent samples t-test

$$t_{(\bar{x}_1-\bar{x}_2)} = \frac{(\bar{x}_1 - \bar{x}_2)}{s_{(\bar{x}_1-\bar{x}_2)}}$$

where:

Estimated standard error of the mean difference

$$s_{(\bar{x}_1-\bar{x}_2)} = \sqrt{\left[\frac{(n_1-1)s_1^2 + (n_2-1)s_2^2}{n_1+n_2-2}\right]\left[\frac{n_1+n_2}{n_1 n_2}\right]}$$

Not the prettiest formula, but notice that you only need a few simple values to complete it:

$$\bar{x}_1 \qquad \bar{x}_2 \qquad n_1 \qquad n_2 \qquad s_1^2 \qquad s_2^2$$

That's right! The only values that go into this formula are sample means, sample sizes, and sample variances–all of which you already know how to calculate! To complete this type of problem, I recommend ignoring the formulas until the very end and instead focusing on finding the 6 values listed above.

Let's start by finding the sample means ($\bar{x}_1$ and $\bar{x}_2$). As usual, these should be pretty easy to calculate by adding up the values in each group and dividing by the number of values in that group. For these data, the sample mean for people in the video game condition is $\bar{x}_1 = 6.87$ and the sample mean for people in the TV condition is $\bar{x}_2 = 8.07$.

Next, let's take care of the sample sizes. In this case, we have 15 participants in each condition, so $n_1 = 15$ and $n_2 = 15$.

Finally, we need to calculate the sample variances of each group, s_1^2 and s_2^2. As usual, I'll show my work for calculating these sample variances on the final page of this document, but for now I'll tell you that $s_1^2 = 7.98$ and $s_2^2 = 6.78$.

Now we have the 6 values we need to plug into the formula for the independent-samples *t*-test:

$$\bar{x}_1 = 6.87 \qquad \bar{x}_2 = 8.07 \qquad n_1 = 15 \qquad n_2 = 15 \qquad s_1^2 = 7.98 \qquad s_2^2 = 6.78$$

Now, let's plug into our formulas, starting with the standard error (i.e., the denominator of the formula for the independent-samples *t*-test, which has its own formula; see above).

$$s_{(\bar{x}_1-\bar{x}_2)} = \sqrt{\left[\frac{(n_1-1)s_1^2+(n_2-1)s_2^2}{n_1+n_2-2}\right]\left[\frac{n_1+n_2}{n_1n_2}\right]}$$

$$s_{(\bar{x}_1-\bar{x}_2)} = \sqrt{\left[\frac{(15-1)(7.98)+(15-1)(6.78)}{15+15-2}\right]\left[\frac{15+15}{(15)(15)}\right]}$$

$$s_{(\bar{x}_1-\bar{x}_2)} = \sqrt{\left[\frac{111.72+94.92}{28}\right]\left[\frac{15+15}{(15)(15)}\right]}$$

$$s_{(\bar{x}_1-\bar{x}_2)} = \sqrt{\left[\frac{111.72+94.92}{28}\right]\left[\frac{30}{225}\right]}$$

$$s_{(\bar{x}_1-\bar{x}_2)} = \sqrt{(7.38)(0.133)}$$

$$s_{(\bar{x}_1-\bar{x}_2)} = 0.99$$

Okay, so now that we've determined that $s_{\bar{x}_1-\bar{x}_2} = 0.99$, we can plug this value into the denominator of the overall independent-samples *t*-test formula:

$$t_{(\bar{x}_1-\bar{x}_2)} = \frac{(\bar{x}_1-\bar{x}_2)}{s_{(\bar{x}_1-\bar{x}_2)}}$$

$$t_{(\bar{x}_1-\bar{x}_2)} = \frac{6.87-8.07}{0.99}$$

$$t_{(\bar{x}_1-\bar{x}_2)} = -1.21$$

And there we have it! $t_{(\bar{x}_1-\bar{x}_2)} = -1.21$. Normally, you'd use statistical software to determine your *p*-value from this *t*-test statistic. I'll tell you that the *p*-value would come out to $p = 0.24$, which is not less than our alpha of 0.05. As a result, we would not have sufficient evidence to say that playing video games affects anxiety more than watching TV.

Calculating sample variances

To remind you, we'll be using the following formula to calculate the sample variance of each group:

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

Let's start with the first group (i.e., the "*Games*" condition). As always, we'll need to first calculate the Sum of Squares by finding the deviations of each value from the mean, squaring those deviations, and then adding up those squares:

Games	deviations	squared_deviations
6	-0.867	0.751
7	0.133	0.018
8	1.133	1.284
5	-1.867	3.484
4	-2.867	8.218
6	-0.867	0.751
7	0.133	0.018
8	1.133	1.284
9	2.133	4.551
11	4.133	17.084
13	6.133	37.618
2	-4.867	23.684
4	-2.867	8.218
5	-1.867	3.484
8	1.133	1.284

Next, we need to add the values in the last column to calculate *SS* (i.e., $\sum(X_i - \bar{x})^2$), which results in 111.733.

Finally, we need to divide the Sum of Squares by $n - 1$, which gives us $\frac{111.733}{14} = 7.98$. So, our variance for the *Games* group is $s_1^2 = 7.98$.

Next, let's calculate the sample variance of the *TV* group by following the same steps as above:

TV	deviations	squared_deviations
7	-1.067	1.138
8	-0.067	0.004
9	0.933	0.871
8	-0.067	0.004
5	-3.067	9.404
4	-4.067	16.538
8	-0.067	0.004
7	-1.067	1.138
6	-2.067	4.271
7	-1.067	1.138
8	-0.067	0.004
14	5.933	35.204
13	4.933	24.338
9	0.933	0.871
8	-0.067	0.004

By adding the values in the final column, we can determine that the Sum of Squares is 94.933. By dividing this value by $n - 1$, we find that our sample variance is 6.78.