

10.7: Effect Sizes and Confidence Intervals

We have seen in previous chapters that even a statistically significant effect needs to be interpreted along with an effect size to see if it is practically meaningful. We have also seen that our sample means, as a point estimate, are not perfect and would be better represented by a range of values that we call a confidence interval. As with all other topics, this is also true of our independent samples t -tests.

Our effect size for the independent samples t -test is still Cohen's d , and it is still just our observed effect divided by the standard deviation. Remember that standard deviation is just the square root of the variance, and because we work with pooled variance in our test statistic, we will use the square root of the pooled variance as our denominator in the formula for Cohen's d . This gives us:

$$d = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{s_p^2}} \quad (10.7.1)$$

For our example above, we can calculate the effect size to be:

$$d = \frac{24.00 - 16.50}{\sqrt{144.48}} = \frac{7.50}{12.02} = 0.62$$

We interpret this using the same guidelines as before, so we would consider this a moderate or moderately large effect.

Our confidence intervals also take on the same form and interpretation as they have in the past. The value we are interested in is the difference between the two means, so our point estimate is the value of one mean minus the other, or \bar{x}_1 minus \bar{x}_2 . Just like before, this is our observed effect and is the same value as the one we place in the numerator of our test statistic. We calculate this value then place the margin of error – still our critical value times our standard error – above and below it. That is:

$$\text{Confidence Interval} = (\bar{X}_1 - \bar{X}_2) \pm t^* \left(s_{\bar{X}_1 - \bar{X}_2} \right) \quad (10.7.2)$$

Because our hypothesis testing example used a one-tailed test, it would be inappropriate to calculate a confidence interval on those data (remember that we can only calculate a confidence interval for a two-tailed test because the interval extends in both directions). Let's say we find summary statistics on the average life satisfaction of people from two different towns and want to create a confidence interval to see if the difference between the two might actually be zero.

Our sample data are $\bar{X}_1 = 28.65$ $s_1 = 12.40$ $n_1 = 40$ and $\bar{X}_2 = 25.40$ $s_2 = 15.68$ $n_2 = 42$. At face value, it looks like the people from the first town have higher life satisfaction (28.65 vs. 25.40), but it will take a confidence interval (or complete hypothesis testing process) to see if that is true or just due to random chance. First, we want to calculate the difference between our sample means, which is $28.65 - 25.40 = 3.25$. Next, we need a critical value from our t -table. If we want to test at the normal 95% level of confidence, then our sample sizes will yield degrees of freedom equal to $40 + 42 - 2 = 80$. From our table, that gives us a critical value of $t^* = 1.990$. Finally, we need our standard error. Recall that our standard error for an independent samples t -test uses pooled variance, which requires the Sum of Squares and degrees of freedom. Up to this point, we have calculated the Sum of Squares using raw data, but in this situation, we do not have access to it. So, what are we to do?

If we have summary data like standard deviation and sample size, it is very easy to calculate the pooled variance, and the key lies in rearranging the formulas to work backwards through them. We need the Sum of Squares and degrees of freedom to calculate our pooled variance. Degrees of freedom is very simple: we just take the sample size minus 1.00 for each group. Getting the Sum of Squares is also easy: remember that variance is standard deviation squared and is the Sum of Squares divided by the degrees of freedom. That is:

$$s^2 = (s)^2 = \frac{SS}{df} \quad (10.7.3)$$

To get the Sum of Squares, we just multiply both sides of the above equation to get:

$$s^2 * df = SS \quad (10.7.4)$$

Which is the squared standard deviation multiplied by the degrees of freedom ($n - 1$) equals the Sum of Squares.

Using our example data:

$$\begin{aligned}(s_1)^2 * df_1 &= SS_1 \\ (12.40)^2 * (40 - 1) &= 5996.64 \\ (s_2)^2 * df_2 &= SS_2 \\ (15.68)^2 * (42 - 1) &= 10080.36\end{aligned}$$

And thus our pooled variance equals:

$$s_p^2 = \frac{SS_1 + SS_2}{df_1 + df_2} = \frac{5996.64 + 10080.36}{39 + 41} = \frac{16077}{80} = 200.96$$

And our standard error equals:

$$s_{\bar{x}_1 - \bar{x}_2} = \sqrt{\frac{s_p^2}{n_1} + \frac{s_p^2}{n_2}} = \sqrt{\frac{200.96}{40} + \frac{200.96}{42}} = \sqrt{5.02 + 4.78} = \sqrt{9.89} = 3.14$$

All of these steps are just slightly different ways of using the same formulae, numbers, and ideas we have worked with up to this point. Once we get out standard error, it's time to build our confidence interval.

$$\begin{aligned}95\%CI &= 3.25 \pm 1.990(3.14) \\ \text{Upper Bound} &= 3.25 + 1.990(3.14) \\ UB &= 3.25 + 6.25 \\ UB &= 9.50 \\ \text{Lower Bound} &= 3.25 - 1.990(3.14) \\ LB &= 3.25 - 6.25 \\ LB &= -3.00 \\ 95\%CI &= (-3.00, 9.50)\end{aligned}$$

Our confidence interval, as always, represents a range of values that would be considered reasonable or plausible based on our observed data. In this instance, our interval (-3.00, 9.50) does contain zero. Thus, even though the means look a little bit different, it may very well be the case that the life satisfaction in both of these towns is the same. Proving otherwise would require more data.

This page titled [10.7: Effect Sizes and Confidence Intervals](#) is shared under a [CC BY-NC-SA 4.0](#) license and was authored, remixed, and/or curated by [Foster et al. \(University of Missouri's Affordable and Open Access Educational Resources Initiative\)](#) via [source content](#) that was edited to the style and standards of the LibreTexts platform.