

One-Sample T-Test

Dr. Tom Pierce
Department of Psychology
Radford University

Let's say that an investigator is interested in the effects of caring for a person with Alzheimer's disease on physical and mental health. In particular, the researcher wants to know whether Alzheimer's caregivers suffer from clinically elevated levels of depression. They administer the Beck Depression Inventory (BDI) to 25 Alzheimer's caregivers. Scores on the BDI range from zero to 40. Scores at or above 25 are considered to reflect clinically significant levels of depression. The mean depression score for the 25 caregivers is 31. The standard deviation for this sample is 3.0. Do these data support the hypothesis that the mean depression score for Alzheimer's caregivers is significantly higher than the cutoff score for depression of 25?

The null hypothesis for this question is that Alzheimer's caregivers do not have BDI scores that are significantly higher than the cutoff score for a clinically elevated level of depression. The alternative hypothesis for this question is that Alzheimer's caregivers have BDI scores that are significantly higher than the cutoff score for a clinically elevated level of depression. Let's say that the researcher decides to use an alpha level of .05 to test this null hypothesis.

This situation sounds a lot like the scenario we had for our Z-test in the last chapter. There, we were comparing the mean I.Q. scores of a sample of seventh graders to a comparison value of 100, the mean I.Q. score for the population. In other words, we conducted our test to see if the sample mean was significantly greater than a comparison value. Here, we're trying to see if the mean of our sample is significantly greater than the comparison value of 25.

So what's different?

I mean, could you answer the question by doing a Z-test? Think for a second. What do you need to do a Z-test? To compute the observed value for Z you take the mean of the sample and then subtract the comparison value. Often, this comparison value is the mean of a population, but it could be any meaningful number. This gives you the difference between the sample mean and this comparison value. We can do that here. Thirty-one minus 25 equals 6.0. Now you'd take this difference of four points and divide it by the standard error of the mean to get a standard score.

$$Z_{\bar{X}} = \frac{\bar{X} - \mu}{\sigma_{\bar{X}}}$$

But how did you get the standard error of the mean? You took the standard deviation of the population and divided it by the square root of the sample size.

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{N}}$$

Can you do that here? What problem do you run into? Well, to calculate the standard error of the mean you have to take the standard deviation of the population and divide it by the square root of N. But we don't have σ . **We don't have the standard deviation of the population.** If we don't have that we can't calculate σ_M , the standard error of the mean. If you don't have σ_M you can't calculate Z! If you can't calculate Z, you can't do a Z-test. Bummer.

So what do you do!

Clearly, we can't do a Z-test. But that doesn't mean we just have to give up. After all, we don't have the standard deviation of the population, but we do have the next best thing. We've got an *estimate* of the standard deviation of the population. We've got the standard deviation of the *sample*. If we take the standard deviation of the sample and divide it by the square root of the sample size, we won't get σ_M but we'll have an estimate of σ_M . The symbol for an estimate of the standard error of the mean is $S_{\bar{x}}$.

$$S_{\bar{x}} = \frac{S}{\sqrt{N}}$$

Using the data provided above we get...

$$S_{\bar{x}} = \frac{S}{\sqrt{N}} = \frac{3}{\sqrt{25}} = \frac{3}{5} = .6$$

Okay, so now we've got an *estimate* of how far off, on average, sample means are going to be off by when they're trying to estimate that comparison value of 25. Now we can take our *one* difference from the comparison value, $31 - 25 = 6$, and divide it by the *average* difference of sample means from this comparison value, 0.6. This will give us an estimate of the number of standard deviations that our sample mean above the comparison value.

$$\text{Estimated standard score for the sample mean} = \frac{31 - 25}{0.6} = \frac{6.0}{0.6} = 10$$

Now we've got to figure out what to call this standard score. We can't call it a value for Z because the number we've just calculated isn't necessarily the real thing. It's only an estimate. This estimate might be a little bit larger or a little bit smaller than it really should be. *Statisticians have decided to refer to an estimate for a value of Z as a value for t .* In situations where we have the standard deviation of the sample, but not the standard deviation of the population, the best we can do is to calculate a value for t . So this is a situation where we need to do a t -test, rather than a Z -test.

Now, you may have noticed that the number crunching to calculate t is exactly the same as the process to calculate a value for Z . The only thing that's different is that some of the symbols are different. But calculating an observed value for t is only half of the story. We need something to compare this observed value for t to. Is a value of +10.0 large enough to be confident that the null hypothesis is false? If we were able to perform a Z -test and we were doing a directional test, we'd use a critical value of +1.645. That critical value is a value for Z , but we're not doing a Z -test. We can't compare an observed value for t to a critical value for Z . We need to compare apples to apples. We've got to figure out what the critical value for t ought to be.

Table ___ provides critical values for t . In order to look up the critical value we need, you need to know three things. To know which column to look in you've got to know (1) whether you're doing a directional (1-tailed) or a non-directional (2-tailed) test and (2) whether you're using an alpha level of .05 or .01. We're doing a one-tailed test with an alpha level of .05, so we need to look in the column that starts at the top with 6.314. To know which row to look in you've got to know the number of degrees of freedom that are associated with the test. The number of degrees of freedom for this test is equal to the number of people in the sample minus one ($N - 1$). We've got 25 participants, so we've got 24 degrees of freedom. When you look at the value in the 24 degrees of freedom row you get 1.71. The critical value we need to use in this t -test is +1.71. This makes the decision rule for this test...

If $t_{\bar{x}} \geq +1.71$, reject H_0

Our observed value for t is 10.0 so our decision is to reject the null hypothesis and accept the alternative hypothesis. The conclusion sentence that might appear in an APA paper is...

Alzheimer's caregivers had scores on the Beck Depression Inventory that were significantly greater than the cutoff score for depression of 25 ($M = 31.0$, $SD = 3.0$), $t(24) = 10.0$, $p < .05$.

Why critical values for t are larger than critical values for Z

Take another look at the critical values for t table. Go to the column where we found the critical value for t for the example above. It starts at the top with 6.314. What's the critical value when you've only got 10 people in the sample? 1.833. Okay, what's the

critical value when there are 25 participants in the sample? 1.711, of course. What's critical value when you've got 121 participants in the sample? 1.658. Now, what if, hypothetically, you had an infinite number of participants in your sample? Look at that! The critical value at the bottom of the column is 1.645 – the identical number as the critical value we would have used if we'd been able to do a Z-test. How did that happen? *Our t-test ends up looking more and more like a Z-test as the sample size gets larger.* Why is this? In a Z-test, changing the sample size didn't do anything to the critical value. *Why does the critical value for t get closer and closer to the critical value for Z as the sample size gets larger?*

To answer this question, let's start with what you know. You know that t is an estimate of Z. You know that t is only an estimate of Z because we only have an estimate of the standard error of the mean. We only have an estimate of the standard error of the mean because we only have an estimate of the standard deviation of the population. We have S instead of σ .

Well so what! Why should that make a difference? You're making my head hurt!

Just relax. This is easy.

You have to think about where the critical value for t comes from. Obviously, the critical value for t corresponds to a point of the scale of possible values for t (See Figure XX). We picked that number because of some choices we made about how the null hypothesis we wanted to test and about how unlikely the null hypothesis would have to be before we just couldn't believe it anymore. In this case we decided to do a directional (one-tailed) test and to use an alpha level of .05. The critical value for t means the same thing as a critical value for Z. It's how far above a standard score of zero we have to go before we hit the start of the outer 5% of values that make up the curve that we're working with. If our observed value for t is greater than or equal to this number (1.711), we know that the odds are less than 5% that it was obtained when the null hypothesis is true.

The critical value for t is based on is a graph of a frequency distribution. The numbers that make up the curve are sample means that have all been converted to values for t. For this reason, it's called a **t curve**.

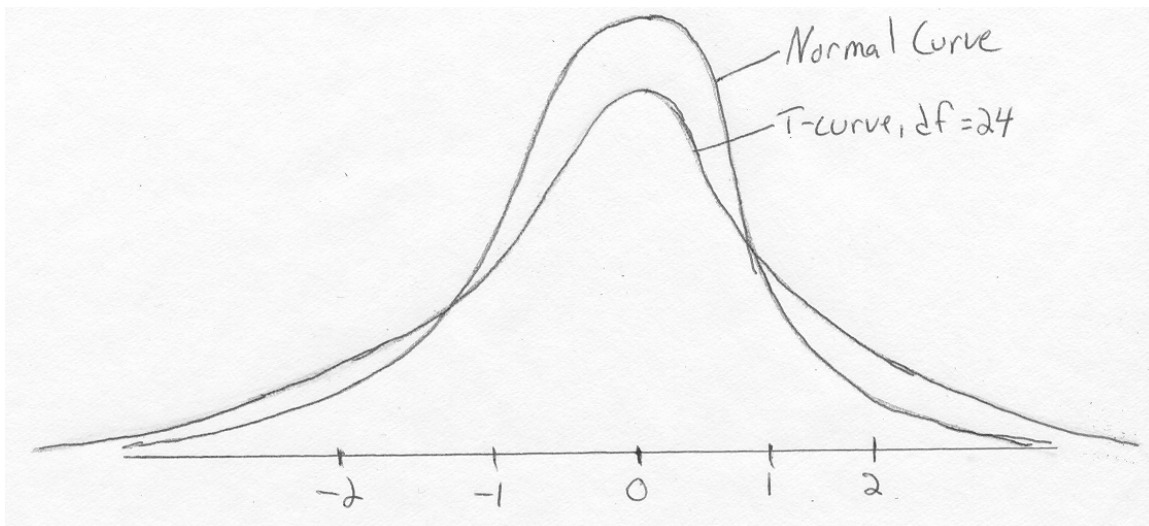
Okay, so why is the critical value for t (1.711) a little bit further away from the center of the curve than the critical value for Z (1.645)? It has to do with the fact that when we convert sample mean to values for t we're only dealing with estimates for these standard scores, not the real things.

Let's say that you could know for sure that the null hypothesis is true and that the mean of a population of people we're working with is 25. You draw a sample of 25 scores from that population. What we'd really like to be able to do is to say exactly how many standard deviations our sample mean is above the comparison value of 25. We'd like to be able to say exactly where that sample mean belongs on the scale of possible standard scores. But we can't do that here. The best we can do to make a best guess about where

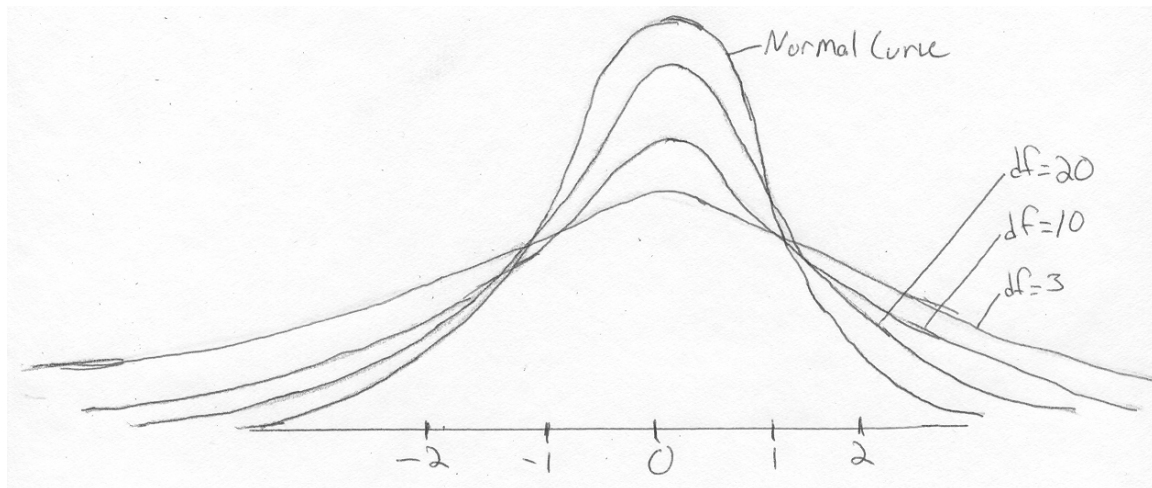
it goes. And we just have to be willing to live with the fact that this best guess might be a little bit high or a little bit low.

Now think about where the critical value for t must come from. Remember, we're trying to compare one sample mean to a bunch of other sample means to see if it belongs with them or not. Except the units we're dealing with here are standard score units. In other words, we're taking all the sample means we could hypothetically collect that are estimates of that comparison value of 25 and converting them all to standard score units – in this case, values for t .

That's the same thing we did in the Z -test example; except that here there is an extra element of uncertainty in placing each value for t where it belongs on the scale. With a Z test, whenever we imagined getting a sample mean and converting it to a standard score we knew exactly where it went on the scale. The normal curve is shape of a frequency distribution made up of thousands and thousands of numbers when you know *exactly* where they go on the scale. Now we're dealing with a bunch of *estimates* of Z – numbers that could always be a little bit too high or a little bit too low. What do you think the distribution of a bunch of estimates is going to look like? Are these estimates going to be as close, on average, to the number they're trying to estimate? NO! They're going to be more spread out! This makes the distribution of values for t flatter and more spread out than the distribution of values for Z . A comparison of the shapes of a normal distribution and the shape of the t -curve that we're working with (24 degrees of freedom) is shown below.



The fewer the number of people in each sample, the less accurate each value for t is going to be when estimating a value for Z . The less accurate the collection of estimates, the more spread out this collection of estimates is going to be. The figure below shows how the shape of the t -curve gets flatter and flatter as the sample size drops.



Each row in the *Critical Values for t* table gives you the critical values for t for a t-curve with its own unique shape. Instead of dealing with one curve (the normal curve) in a Z-test, a t-test has a slightly differently shaped curve for every possible sample size. And some poor underpaid slob has already figured out the critical values for every one of those curves and put them in the t table for you!

The reason the critical values for t gets closer and closer to the critical value for Z as the number of degrees of freedom gets larger is that larger sample sizes give you more accurate estimates. With a small sample size, like $N=6$, values for t don't give you very accurate estimates of values for Z and the shape of the curve is quite a bit flatter than the normal curve. This pushes the start of the outer 5% of the curve further away from the center of the curve. Larger sample sizes put values for t closer to the values for Z that they're trying to estimate.

Now, hypothetically, let's say that you could have the largest sample size you could possibly imagine. You've got an infinite number of people in each sample. What kind of estimate are you going to get if you've got an infinite number of people in your sample? PERFECT! If you could use samples with an infinite number of people in them, every value for t would be a perfect estimate of a value for Z. And the distribution of a bunch of perfect estimates of the real thing has exactly the same shape as a distribution of the real thing. When the sample size is infinite the shape of the t-curve comes to look exactly like the shape of the Z-curve (normal curve). If the shapes of the curves are the same you're going to have to go exactly the same number of standard deviations above a standard score of z to get to the start of the outer 5% of the values that make up the curve. The critical value for t will be the same number as the critical value for Z.