One sided tests

So far all of our tests have been two sided. While this may be a bit easier to understand, this is often not the best way to do a hypothesis test. One simple thing that we can do to get a more powerful test is to do a one sided test. One sided tests are also called *one tailed* or *directional*.

An example of a two sided alternative is what we've been using for our two sample tests:

$$H_1 : \mu_1 \neq \mu_2$$

Notice that this uses the $\neq$ symbol. We reject if we think μ_1 is bigger than μ_2 *or* if we think μ_1 is smaller than μ_2. A one sided test uses only one of these alternatives, and our alternative hypothesis becomes:

$$H_1 : \mu_1 > \mu_2 \quad \textit{or} \quad H_1 : \mu_1 < \mu_2 \quad \text{(one or the other, not both!)}$$

How do we know when to do a one sided test? There are two ways that we can decide to do a one sided test:

1. We know something about the subject matter. Suppose we try to figure out if there is a difference in height between men and women. If $\mu_1 =$ "the true average height of women" and $\mu_2 =$ "the true average height of men, what outcome do we expect? We expect men to be taller. So we can choose our alternative hypothesis as follows:

   $$H_1 : \mu_1 < \mu_2$$

2. Logic tells us something about the problem at hand. Suppose we test a new medication to reduce blood pressure. If If $\mu_1 =$ "the true average systolic blood pressure without medication" and $\mu_2 =$ "the true average systolic blood pressure with medication, what do we expect to happen? We expect blood pressure to drop after taking medication, so we would choose our alternative hypothesis as follows:

   $$H_1 : \mu_1 > \mu_2$$

We use subject matter knowledge or logic to decide if we should do a one sided test.

Note that we *never (never!)* decide to use a one sided test by comparing $\bar{y_1}$ with $\bar{y_2}$:

This is WRONG: if $\bar{y_1} < \bar{y_2}$ then do $H_1 : \mu_1 < \mu_2$

The reason we don't do this is that a one sided test gives you more power. You are only entitled to this extra power if you can decide *ahead of time* to do a one sided test. We'll look more closely at power below.

Now that we know a little about how to decide whether or not to do a one sided test, we need to figure out how to actually carry out a one sided test. The good news is that almost all of our math that we've learned so far stays the same. A one sided two sample T test, for

©2019 Arndt F. Laemmerzahl

example, uses the same calculations as the two sided test. However, there are two things we need to watch out for:

We need to make sure our data agree with the alternative hypothesis.

We need to pull out the correct value from the tables.

Let's use a two sample T test as an example and outline this (the "$\rightarrow$" below indicates the parts you need to do differently):

1. Set up your hypotheses:

   $H_0 : \mu_1 = \mu_2$

   $\rightarrow H_1 : \mu_1 > \mu_2 \quad$ (or $H_1 : \mu_1 < \mu_2$)

2. Pick your value for α

3. $\rightarrow$ *Verify that your data agree with your alternative hypothesis*:

   For example, if your alternative hypothesis is $H_1 : \mu_1 < \mu_2$, you need to make sure that $\bar{y_1} < \bar{y_2}$ (or vice versa).

   If this is not true, *STOP*:

   Don't do anything else. Don't continue your calculations. Your results *contradict* your alternative hypothesis.

   You automatically *fail to reject*. There might also be something screwy with your data.

   If you forget this step, you could make very bad mistakes!

4. Calculate t^* in the usual way. Remember, there are several different versions of the T test; just do everything the same as usual. Don't forget the assumptions of your test (normal distribution, etc.).

5. Make your comparison in the usual way:

   $\rightarrow$ Make sure you use the *one sided* value for t_{table} from the tables.

   If $|t^*| \geq t_{table}$ reject H_0, otherwise fail to reject H_0.

This is not the traditional way of doing one sided tests T tests. Usually you take note of the sign ($+$ or $-$) of t^*, and then compare that to the tabulated value, which may be negative (you may need to change the sign of the table value). Although this does get rid of the requirement of having to check if your data agree with H_1, it is also rather more confusing.

Let's do an example. In general, the larger the animal, the slower the heart rate. For example, the average heart rate of an elephant (about 30) is *much* lower than that for

©2019 Arndt F. Laemmerzahl

humans (about 70).
You want to compare the heart rates between rats and mice. Based on what you just read, who should have the higher heart rate (what do we expect to happen)?

If we let μ_1 =the true average hear rates of rats and μ_2 =the true average heart rate of mice, what do we expect to happen? This is what we put in our alternative hypothesis:

$H_0 : \mu_1 = \mu_2$

$H_1 : \mu_1 < \mu_2$ $\leftarrow$ this is what we *expect* to happen.

So now let's proceed. We'll pick $\alpha = 0.10$. Here are the data:

						$\bar{y}$	s	SE
Heart rate in rats:	478	490	563	617	371	503.8	93.29	41.72
Heart rate in mice:	788	619	365	814	655	648.2	178.98	80.04

Let's make sure our data agree with our alternative hypothesis:

We note that $\bar{y}_1 = 503.8$, and that $\bar{y}_2 = 648.2$.

Our alternative hypothesis is that $\mu_1 < \mu_2$, and since $\bar{y}_1 < \bar{y}_2$ our data agree with the alternative so we continue.

Let's calculate t^*:

$$t^* = \frac{508.8 - 648.2}{\sqrt{\frac{93.29^2}{5} + \frac{178.98^2}{5}}} = -1.5998$$

$$\implies |t^*| = 1.5998$$

And then figure out our degrees of freedom:

$$d.f. = \nu = \frac{(41.72^2 + 80.04^2)^2}{\frac{41.72^4}{4} + \frac{80.04^4}{4}} = 6.042$$

So we use $d.f.$ =6 (we round down).

Now we go to our T tables and look up the *one sided* value:

$t_{table} = t_{0.10,6} = 1.440$

And since $|t^*| = 1.5998 \geq t_{0.10,6} = 1.440$, we reject our null hypothesis and conclude that the heart rates in rats are slower than those in mice.

©2019 Arndt F. Laemmerzahl

So we've learned how to do a one sided or directional test. Why can't we just stick with two sided tests? Because one sided tests have more power. Let's do our example again, but this time as a two sided test:

$H_0 : \mu_1 = \mu_2$

$H_1 : \mu_1 \neq \mu_2$

We pick the same α, so $\alpha = 0.10$.

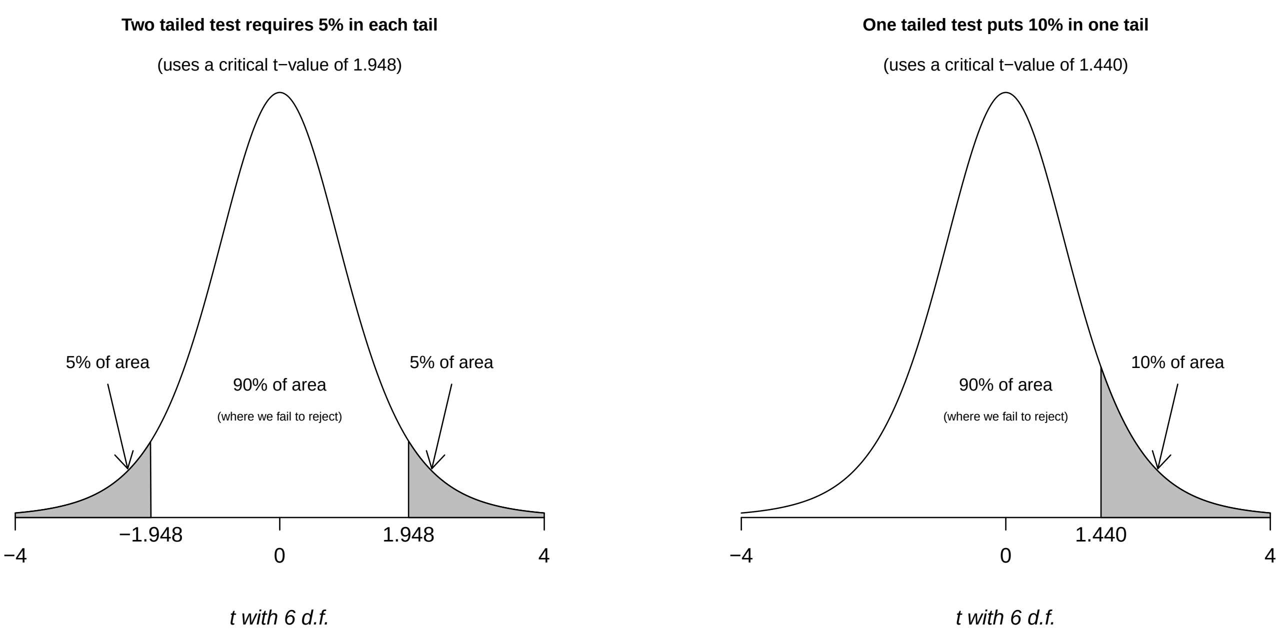
We don't have to make sure our data agree with anything, so we can proceed directly to calculating t^*. We don't actually have to calculate this since we already did the calculation for both t^* and $d.f.$ above.

So we look up our t_{table} value, but since we're now doing a two sided test, we pick:

$t_{table} = t_{0.10,6} = 1.948$

If we now do our comparison, we find that $|t^*| = 1.5998 < t_{0.10,6} = 1.948$, and we *fail to reject* our null hypothesis.

If we do a one sided test, we can see a difference in heart rates, if we do a two sided test, we can *not* find a difference. We lost power by doing a two sided test. Let's look at a picture to see what's going on:



©2019 Arndt F. Laemmerzahl

A one sided test lets us put all of the area on one side, and this lets us use a smaller value of t_{table}, so it makes it easier to reject (and gives us more power).

While it isn't true for all tests, many other tests can be done using a one sided alternative hypothesis. We'll learn some of these in the next few chapters. In the meantime, let's take a quick look at how to do a Mann-Whitney U test that is one sided:

Develop your hypotheses (just as you would for the one-tailed T test).

Decide on α.

Calculate U^*.

You still need to verify that your data agree with H_1. Since the MWU test doesn't always use means in it's hypotheses, what do you do?

> You use K_1 and K_2.
>
> For example, if $H_1 : \mu_1 < \mu_2$, then you make sure that $K_1 < K_2$.
>
> The same thing applies if you're doing $H_1 : D_1 < D_2$; make sure that $K_1 < K_2$.

Now look up your U_{table} value, but make sure you use the values for one sided probabilities. Remember that the MWU tables are laid out so that *each* table is for a single value of α, so make sure you use the one sided values.

Your comparison is the same as always. if $U^* \geq U_{table}$ reject the null hypothesis.

Finally, some comments on one sided tests:

If you can do a one sided test you should. It gives you more power, and as a statistician you should always use the best (i.e., most powerful) test available. But, remember, you should have a reason for choosing a one sided test before you carry out your test. Preferably you should actually *design* your experiment with a one sided test in mind if at all possible.

©2019 Arndt F. Laemmerzahl