

13.2: The One-sample t-test

After some thought, I decided that it might not be safe to assume that the psychology student grades necessarily have the same standard deviation as the other students in Dr Zeppo's class. After all, if I'm entertaining the hypothesis that they don't have the same mean, then why should I believe that they absolutely have the same standard deviation? In view of this, I should really stop assuming that I know the true value of σ . This violates the assumptions of my z-test, so in one sense I'm back to square one. However, it's not like I'm completely bereft of options. After all, I've still got my raw data, and those raw data give me an *estimate* of the population standard deviation:

```
sd( grades )
```

```
## [1] 9.520615
```

In other words, while I can't say that I know that $\sigma=9.5$, I *can* say that $\hat{\sigma}=9.52$.

Okay, cool. The obvious thing that you might think to do is run a z-test, but using the estimated standard deviation of 9.52 instead of relying on my assumption that the true standard deviation is 9.5. So, we could just type this new number into R and out would come the answer. And you probably wouldn't be surprised to hear that this would still give us a significant result. This approach is close, but it's not *quite* correct. Because we are now relying on an *estimate* of the population standard deviation, we need to make some adjustment for the fact that we have some uncertainty about what the true population standard deviation actually is. Maybe our data are just a fluke ... maybe the true population standard deviation is 11, for instance. But if that were actually true, and we ran the z-test assuming $\sigma=11$, then the result would end up being *non-significant*. That's a problem, and it's one we're going to have to address.

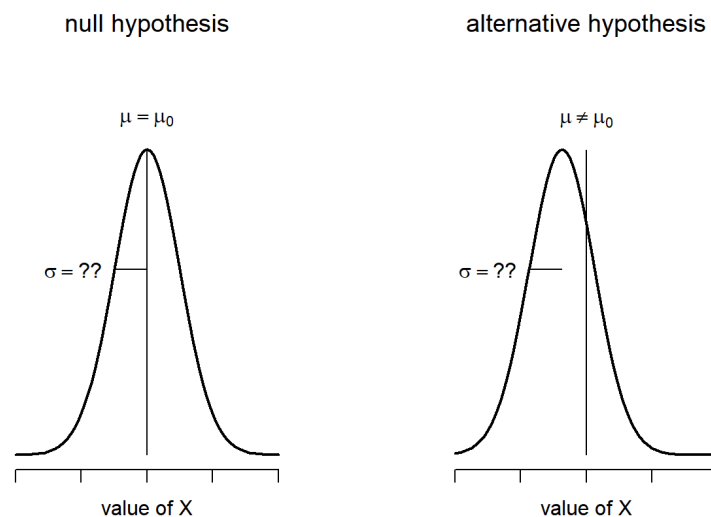


Figure 13.4: Graphical illustration of the null and alternative hypotheses assumed by the (two sided) one sample t-test. Note the similarity to the z-test. The null hypothesis is that the population mean μ is equal to some specified value μ_0 , and the alternative hypothesis is that it is not. Like the z-test, we assume that the data are normally distributed; but we do not assume that the population standard deviation σ is known in advance.

13.2.1 Introducing the t-test

This ambiguity is annoying, and it was resolved in 1908 by a guy called William Sealy Gosset (Student 1908), who was working as a chemist for the Guinness brewery at the time (see Box 1987). Because Guinness took a dim view of its employees publishing statistical analysis (apparently they felt it was a trade secret), he published the work under the pseudonym "A Student", and to this day, the full name of the t-test is actually **Student's t-test**. The key thing that Gosset figured out is how we should accommodate the fact that we aren't completely sure what the true standard deviation is.¹⁸⁷ The answer is that it subtly changes the sampling distribution. In the t-test, our test statistic (now called a t-statistic) is calculated in exactly the same way I mentioned above. If our null hypothesis is that the true mean is μ , but our sample has mean \bar{X} and our estimate of the population standard deviation is $\hat{\sigma}$, then our t statistic is:

$$t = \frac{\bar{X} - \mu}{\hat{\sigma} / \sqrt{N}}$$

The only thing that has changed in the equation is that instead of using the known true value σ , we use the estimate $\hat{\sigma}$. And if this estimate has been constructed from N observations, then the sampling distribution turns into a t -distribution with $N-1$ **degrees of freedom** (df). The t distribution is very similar to the normal distribution, but has “heavier” tails, as discussed earlier in Section 9.6 and illustrated in Figure 13.5. Notice, though, that as df gets larger, the t -distribution starts to look identical to the standard normal distribution. This is as it should be: if you have a sample size of $N=70,000,000$ then your “estimate” of the standard deviation would be pretty much perfect, right? So, you should expect that for large N , the t -test would behave exactly the same way as a z -test. And that’s exactly what happens!

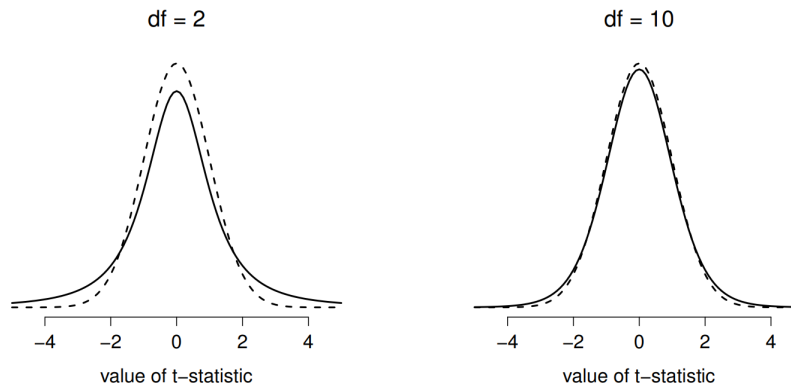


Figure 13.5: The t distribution with 2 degrees of freedom (left) and 10 degrees of freedom (right), with a standard normal distribution (i.e., mean 0 and std dev 1) plotted as dotted lines for comparison purposes. Notice that the t distribution has heavier tails (higher kurtosis) than the normal distribution; this effect is quite exaggerated when the degrees of freedom are very small, but negligible for larger values. In other words, for large df the t distribution is essentially identical to a normal distribution.

13.2.2 Doing the test in R

As you might expect, the mechanics of the t -test are almost identical to the mechanics of the z -test. So there’s not much point in going through the tedious exercise of showing you how to do the calculations using low level commands: it’s pretty much identical to the calculations that we did earlier, except that we use the estimated standard deviation (i.e., something like `se.est <- sd(grades)`), and then we test our hypothesis using the t distribution rather than the normal distribution (i.e. we use `pt()` rather than `pnorm()`). And so instead of going through the calculations in tedious detail for a second time, I’ll jump straight to showing you how t -tests are actually done in practice.

The situation with t -tests is very similar to the one we encountered with chi-squared tests in Chapter 12. R comes with one function called `t.test()` that is very flexible (it can run lots of different kinds of t -tests) and is somewhat terse (the output is quite compressed). Later on in the chapter I’ll show you how to use the `t.test()` function (Section 13.7), but to start out with I’m going to rely on some simpler functions in the `lsr` package. Just like last time, what I’ve done is written a few simpler functions, each of which does only one thing. So, if you want to run a one-sample t -test, use the `oneSampleTTest()` function! It’s pretty straightforward to use: all you need to do is specify `x`, the variable containing the data, and `mu`, the true population mean according to the null hypothesis. All you need to type is this:

```
library(lsr)

oneSampleTTest( x=grades, mu=67.5 )
```

```
##
##   One sample t-test
##
## Data variable:  grades
##
## Descriptive statistics:
##           grades
##   mean      72.300
##   std dev.   9.521
##
## Hypotheses:
##   null:      population mean equals 67.5
##   alternative: population mean not equal to 67.5
##
## Test results:
##   t-statistic:  2.255
##   degrees of freedom:  19
##   p-value:    0.036
##
## Other information:
##   two-sided 95% confidence interval:  [67.844, 76.756]
##   estimated effect size (Cohen's d):  0.504
```

Easy enough. Now let's go through the output. Just like we saw in the last chapter, I've written the functions so that the output is pretty verbose. It tries to describe in a lot of detail what it's actually doing:

```
One sample t-test

Data variable:  grades

Descriptive statistics:
           grades
   mean      72.300
   std dev.   9.521

Hypotheses:
   null:      population mean equals 67.5
   alternative: population mean not equal to 67.5

Test results:
   t-statistic:  2.255
   degrees of freedom:  19
   p-value:    0.036

Other information:
   two-sided 95% confidence interval:  [67.844, 76.756]
   estimated effect size (Cohen's d):  0.504
```

Reading this output from top to bottom, you can see it's trying to lead you through the data analysis process. The first two lines tell you what kind of test was run and what data were used. It then gives you some basic information about the sample: specifically, the sample mean and standard deviation of the data. It then moves towards the inferential statistics part. It starts by telling you what the null and alternative hypotheses were, and then it reports the results of the test: the t-statistic, the degrees of freedom, and the p-

value. Finally, it reports two other things you might care about: the confidence interval for the mean, and a measure of effect size (we'll talk more about effect sizes later).

So that seems straightforward enough. Now what do we *do* with this output? Well, since we're pretending that we actually care about my toy example, we're overjoyed to discover that the result is statistically significant (i.e. p value below 0.05). We could report the result by saying something like this:

With a mean grade of 72.3, the psychology students scored slightly higher than the average grade of 67.5 ($t(19)=2.25$, $p<.05$); the 95% confidence interval is [67.8, 76.8].

where $t(19)$ is shorthand notation for a t -statistic that has 19 degrees of freedom. That said, it's often the case that people don't report the confidence interval, or do so using a much more compressed form than I've done here. For instance, it's not uncommon to see the confidence interval included as part of the stat block, like this:

$t(19)=2.25$, $p<.05$, $CI95=[67.8,76.8]$

With that much jargon crammed into half a line, you know it must be really smart.¹⁸⁸

13.2.3 Assumptions of the one sample t-test

Okay, so what assumptions does the one-sample t -test make? Well, since the t -test is basically a z -test with the assumption of known standard deviation removed, you shouldn't be surprised to see that it makes the same assumptions as the z -test, minus the one about the known standard deviation. That is

- **Normality.** We're still assuming that the the population distribution is normal.[^][A technical comment... in the same way that we can weaken the assumptions of the z -test so that we're only talking about the sampling distribution, we *can* weaken the t test assumptions so that we don't have to assume normality of the population. However, for the t -test, it's trickier to do this. As before, we can replace the assumption of population normality with an assumption that the sampling distribution of \bar{X} is normal. However, remember that we're also relying on a sample estimate of the standard deviation; and so we also require the sampling distribution of $\hat{\sigma}$ to be chi-square. That makes things nastier, and this version is rarely used in practice: fortunately, if the population is normal, then both of these two assumptions are met., and as noted earlier, there are standard tools that you can use to check to see if this assumption is met (Section 13.9), and other tests you can do in it's place if this assumption is violated (Section 13.10).
- **Independence.** Once again, we have to assume that the observations in our sample are generated independently of one another. See the earlier discussion about the z -test for specifics (Section 13.1.4).

Overall, these two assumptions aren't terribly unreasonable, and as a consequence the one-sample t -test is pretty widely used in practice as a way of comparing a sample mean against a hypothesised population mean.

This page titled [13.2: The One-sample t-test](#) is shared under a [CC BY-SA 4.0](#) license and was authored, remixed, and/or curated by [Danielle Navarro](#) via [source content](#) that was edited to the style and standards of the LibreTexts platform.