The Student t Distribution and its Use

James H. Steiger

Department of Psychology and Human Development
Vanderbilt University

The Student t Distribution and its Use

Introduction

- In several previous lectures, we studied, in detail, the characteristics of the Z-statistic for comparing a mean with a null-hypothesized value.
- In the process, we learned a number of general principles about hypothesis testing, power, and the factors affecting power and the sample size n needed to achieve it.
- However, it turns out the Z statistic itself is virtually never used in practice.
- Why? Because the population standard deviation σ is not known, anymore than the population mean is.
- So what do we do?

Introduction

- One obvious strategy is to substitute an estimate of σ in place of σ. The likely candidate is s, the sample standard deviation we studied earlier in the course.
- This would yield a modified Z statistic

$$Z_{modified} = \frac{M - \mu_0}{s/\sqrt{n}} \tag{1}$$

- What do you think will happen if we do that?

The Modified Z Statistic
Thinking Intuitively

- The original Z statistic modified M by subtracting a constant, then dividing by a constant.
- The only thing in the Z statistic that would vary over repeated samples is M, the sample mean.
- This means that the distribution of Z has to be the same shape as the distribution of M.

The Modified Z Statistic

Thinking Intuitively

- The modified Z statistic has a sample quantity in its denominator that varies over repeated samples along with M.
- So now, instead of only one thing varying, you have two.
- It turns out that, as n gets larger and larger, this matters less and less, because s starts acting more and more like the constant that it is estimating.
- In fact, it was known back around 1900 that, as n goes to infinity, the modified Z statistic's distribution got closer and closer to the distribution of the original Z statistic.
- What people didn't know was precisely how to characterize the performance of the modified Z statistic at small sample sizes.

W.S. Gossett and the "Student" t

Some History

- W.S. Gossett was a statistician working for the Guinness brewery when he derived the exact distribution of $Z_{modified}$ under some specific conditions.
- This was seen as something of a landmark development by the statistical community.
- Due to some issues regarding confidentiality and conflict of interest, Gossett was writing under the pen name of "Student" when he published his work, and so the modified Z statistic became known as "Student's t statistic" in his honor.
- The distribution of the statistic became known as "Student's t distribution," and has many applications beyond the simple 1-sample test we are reviewing here.

Basic Facts about the t Distribution

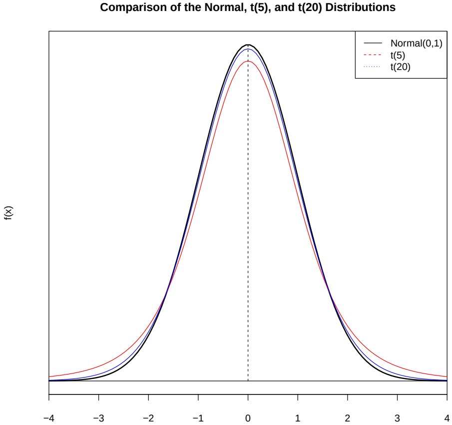
- What are some of the basic facts about the Student t distribution?
- To begin with, it has a single parameter, called the *degrees of freedom*, which we shall abbreviate as df.
- The t distribution is symmetric around a mean of zero.
- For the 1-sample test, $df = n - 1$. Later, we will see a more general formula for df.
- At small df, the t distribution has a shape much like the standard normal, but with larger variability.
- As df increases, the t distribution gets closer and closer to the standard normal distribution in shape.
- Consequently, critical values are somewhat larger than for the normal distribution.
- For example, suppose n is only 10, so $df = 9$. The 0.975 quantile is 2.262, as compared with 1.96 for the normal distribution.
- On the other hand, if $n = 100$, and $df = 99$, the 0.975 quantile is 1.984, only slightly larger than the normal distribution value.

Basic Facts about the t Distribution

- On the next slide, there is a picture comparing the probability density of a standard normal distribution with that of Student's t distributions with 5 (red) and with 20 (blue) degrees of freedom.
- Comparison of this slide with Figure 9.1 from the Gravetter-Walnau textbook on the following slide shows that the Gravetter-Walnau figure is not an accurate representation of the actual densities. Figure 9.1 seems to show equal differences in the heights of the Z, t_5, and t_{20} distributions at 0, when in fact the density of the t_{20} is much closer to the Z than to the t_5 distribution.

Basic Facts about the t Distribution

Comparison of Normal, t(5), and t(20) Distributions

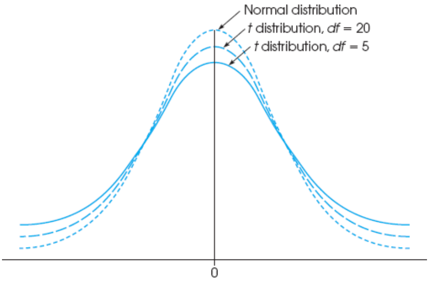


Basic Facts about the t Distribution

GW Figure 9.1

Figure 9.1

Distributions of the t statistic for different values of degrees of freedom are compared to a normal Z-score distribution. Like the normal distribution, t distributions are bell-shaped and symmetrical and have a mean of zero. However, t distributions have more variability, indicated by the flatter and more spread-out shape. The larger the value of df is, the more closely the t distribution approximates a normal distribution.



Critical Values of the t Distribution

- When using the t distribution for statistical testing, we need *critical values* (rejection points), just as with the normal distribution.
- With the normal distribution, there are two parameters (μ and σ), but all normal distributions have the same shape, so critical values for any normal distribution can be computed from the critical values for the standard normal.
- For example, the 0.975 quantile of the standard normal is 1.96, and the 0.975 quantile for any other normal distribution may be found by multiplying 1.96 by that distribution's σ and then adding that distribution's μ.
- On the other hand, the t distribution changes its shape as the *df* parameter changes.
- Consequently, t distribution tables give only a few key quantiles as a function of degrees of freedom.
- Gravetter & Walnau, Table 9.1 is an example of a section from a typical t distribution table.

Critical Values of the t Distribution
From Tables of the t Distribution (GW Table 9.1)

Table 9.1

A portion of the t-distribution table. The numbers in the table are the values of t that separate the tail from the main body of the distribution. Proportions for one or two tails are listed at the top of the table, and df values for t are listed in the first column.

Proportion in One Tail						
	0.25	**0.10**	**0.05**	**0.025**	**0.01**	**0.005**
Proportion in Two Tails Combined						
df	**0.50**	**0.20**	**0.10**	**0.05**	**0.02**	**0.01**
1	1.000	3.078	6.314	12.706	31.821	63.657
2	0.816	1.886	2.920	4.303	6.965	9.925
3	0.765	1.638	2.353	3.182	4.541	5.841
4	0.741	1.533	2.132	2.776	3.747	4.604
5	0.727	1.476	2.015	2.571	3.365	4.032
6	0.718	1.440	1.943	2.447	3.143	3.707

Critical Values of the t Distribution
Using the R Functions

- A superior alternative to the use of tables is to use the R functions for the t distribution.
- In keeping with standard R nomenclature, the key functions are `pt` and `qt`.
- For example, Table 9.1 on the preceding slide gives the upper tail probability of 0.01 for a t value of 3.365 with 5 degrees of freedom.
- An upper tail probability of 0.01 corresponds to the 0.99 quantile of the distribution, which is easily calculated with `qt` below.

```
> qt(0.99,5)

[1] 3.36493
```

Critical Values of the t Distribution

Let's Practice!

- Let's practice computing some critical values.
- Remember the rule. If the test is 2-tailed, half the α probability goes in each tail, and so if $\alpha = 0.05$, 2-tailed, the upper critical value is at the 0.975 quantile.
- If the test is 1-tailed, all the α goes in the rejection region, so if the rejection region is in the upper tail, and $\alpha = 0.05$, then the critical value would be at the 0.95 quantile.
- You also need to remember that the degrees of freedom for the 1-sample t statistic are $df = n - 1$.

Critical Values of the t Distribution

Let's Practice!

Example (Critical Value Calculation: Example 01)

Suppose you want to test the null hypothesis that $\mu = 100$ with a sample of size $n = 25$, and an α of 0.05. What will the critical value(s) for the t statistic be?

(Answer on next slide . . .)

Critical Values of the t Distribution

Let's Practice!

Example (Critical Value Calculation: Example 01 . . . continued)

Suppose you want to test the null hypothesis that $\mu = 100$ with a sample of size $n = 25$, and an α of 0.05. What will the critical value(s) for the t statistic be?

Answer. This is a 2-tailed test, so half the α is in each tail. With 0.025 probability in the upper tail, the upper critical value will be at the 0.975 quantile. The lower critical value will be its symmetric opposite at the 0.025 quantile. Degrees of freedom are $df = n - 1 = 25 - 1 = 24$. Using R, we get

```
> upper.critical.value <- qt(0.975,24)
> lower.critical.value <- qt(0.025,24)
> upper.critical.value
[1] 2.063899
> lower.critical.value
[1] -2.063899
```

Table B.2 in the textbook gives the upper critical value as 2.064.

Critical Values of the t Distribution

Let's Practice!

Example (Critical Value Calculation: Example 02)

Suppose you want to test the null hypothesis that $\mu \leq 100$ with a sample of size $n = 60$, and an α of 0.01. What will the critical value(s) for the t statistic be?

(Answer on next slide . . .)

Critical Values of the t Distribution
Let's Practice!

Example (Critical Value Calculation: Example 02 . . . continued)

Suppose you want to test the null hypothesis that $\mu \leq 100$ with a sample of size $n = 60$, and an α of 0.01. What will the critical value(s) for the t statistic be?

Answer. This is a 1-tailed test, with a rejection region in the upper tail, so all the α is in the upper tail. With 0.01 probability in the upper tail, the upper critical value will be at the 0.99 quantile. Degrees of freedom are $df = n - 1 = 60 - 1 = 59$. Using R, we get

```
> critical.value <- qt(0.99,59)
> critical.value
[1] 2.391229
```

The 1-Sample t Test

- The 1-Sample t statistic is

$$t_{n-1} = \frac{M - \mu_0}{s/\sqrt{n}} \tag{2}$$

- The notation t_{n-1} reminds us that the t statistic has $n - 1$ degrees of freedom.
- To perform the t test, we simply compute the statistic and see if it "beats" its critical value.
- A couple of examples should suffice.

The 1-Sample t Test
Let's Practice!

Example (The 1-Sample t Test: Example 01)

Suppose you want to test the null hypothesis that $\mu = 100$ with a sample of size $n = 25$, and an α of 0.05. You observe a sample mean of 107.23 and a sample standard deviation of 14.87. Perform the 1-sample t test.

(Answer on next slide . . .)

The 1-Sample t Test

Let's Practice!

Example (The 1-Sample t Test: Example 01 …continued)

Suppose you want to test the null hypothesis that $\mu = 100$ with a sample of size $n = 25$, and an α of 0.05. You observe a sample mean of 107.23 and a sample standard deviation of 14.87. Perform the 1-sample t test.

Answer. This is a 2-tailed test, and we already calculated the critical values to be ± 2.064. The test statistic itself can be easily computed in R as

```
> t.observed <- (107.23 - 100)/(14.87/sqrt(25))
> t.observed
[1] 2.431069
```

Since the observed value of t exceeds the positive critical value, the null hypothesis is "rejected at the 0.05 significance level, 2-tailed."

The 1-Sample t Test

Let's Practice!

Example (The 1-Sample t Test: Example 02)

Suppose you want to test the null hypothesis that $\mu \leq 100$ with a sample of size $n = 101$, and an α of 0.01. You observe a sample mean of 104.11 and a sample standard deviation of 16.04. Perform the 1-sample t test.

(Answer on next slide . . .)

The 1-Sample t Test

Let's Practice!

Example (The 1-Sample t Test: Example 02 . . . continued)

Suppose you want to test the null hypothesis that $\mu \leq 100$ with a sample of size $n = 101$, and an α of 0.01. You observe a sample mean of 104.11 and a sample standard deviation of 16.04. Perform the 1-sample t test.

Answer. This is a 1-tailed test, and the critical value is in the upper tail at the 0.99 quantile. The degrees of freedom are $df = 101 - 1 = 100$. The critical value is

```
> critical.value <- qt(0.99,100)
> critical.value
[1] 2.364217
```

(Continued on next slide . . .)

The 1-Sample t Test

Let's Practice!

Example (The 1-Sample t Test: Example 02 ...continued)

The test statistic itself can be easily computed in R as

```
> t.observed <- (104.11 - 100)/(16.04/sqrt(101))
> t.observed
[1] 2.575124
```

Since the observed value of t exceeds the positive critical value, the null hypothesis is "rejected at the 0.01 significance level, 1-tailed."

Introduction

In the context of the 1-sample t, Gravetter & Walnau mention two measures of effect size, the estimated standardized effect size, and r^2, the proportion of variance accounted for by knowing the sample mean.

We'll review each of these briefly in the next sections.

Estimating E_s

- Recall that the standardized effect size, E_s, is defined as

$$E_s = \frac{\mu - \mu_0}{\sigma} \tag{3}$$

and is the amount by which the null hypothesis is wrong, in standard deviation units.

Estimating E_s

- When σ is not known, we estimate E_s from our data as

$$\hat{E}_s = \hat{d} = \frac{M - \mu_0}{s} \tag{4}$$

The "hat" in the above equations means "estimate of."

- Note that, since

$$\begin{aligned} t &= \frac{M - \mu_0}{s/\sqrt{n}} \\ &= \sqrt{n}\frac{M - \mu_0}{s} \\ &= \sqrt{n}\hat{E}_s \end{aligned}$$

it immediately follows that the estimate of E_s may be directly calculated from the t statistic and the sample size as

$$\hat{E}_s = \hat{d} = \frac{t}{\sqrt{n}} = \frac{t}{\sqrt{df + 1}} \tag{5}$$

Estimating r^2
the Proportion of Variance Accounted For

- Another well known measure of effect size is r^2.
- r^2 has a very general meaning in statistics when nested models are compared — it is the proportional reduction in the sum of squared errors of a model made by adding complexity to the model.
- Suppose we had a statistical model that the mean of the population from which a sample of 5 scores was taken is 0. That's all we know about the 5 scores. Our model, in other words, is that $\mu = \mu_0 = 0$.
- Statistical theory tells us that, if we were to have to "estimate" these 5 scores from our model, prior to seeing them, the best we could do, in the long run, would be to use the population mean. So suppose we do that.

Estimating r^2
the Proportion of Variance Accounted For

- Suppose our 5 scores were 1,2,3,4,5.
- Remember that the null hypotheis is $\mu = \mu_0 = 0$.
- If the null hypothesis is true, our best estimate, in the long run, is to estimate each of the 5 scores as 0.
- In that case, what would be the sum of squared errors?
- Let's let R compute that for us.

Estimating r^2
the Proportion of Variance Accounted For

Below, we see that the sum of squared errors is 55 using $\mu = 0$ as our model.

```
> X <- 1:5
> mu_0 <- 0
> X.hat <- rep(mu_0,5)
> E <- X - X.hat
> E.squared <- E^2
> demo_0 <- data.frame(X,X.hat,E,E.squared)
> demo_0

  X X.hat E E.squared
1 1     0 1         1
2 2     0 2         4
3 3     0 3         9
4 4     0 4        16
5 5     0 5        25

> SS_0 <- sum(E.squared)
> SS_0

[1] 55
```

Estimating r^2
the Proportion of Variance Accounted For

- Now suppose we consider that perhaps μ isn't 0, and the null hypothesis is false.
- Using that as our model, we use the data to guess just what the value of μ actually is. We use the sample mean. Since the scores are 1,2,3,4,5, the sample mean is 3.
- Suppose we now use $\mu = 3$ as our model, and "estimate" each score as the sample mean.
- In that case, what would be the sum of squared errors?
- Again, let's let R do it for us.

Estimating r^2
the Proportion of Variance Accounted For

By foresaking the null hypothesis and letting the data speak, we reduce the sum of squared errors from 55 to 10.

```
> X <- 1:5
> mu <- mean(X)
> X.hat <- rep(mu,5)
> E <- X - X.hat
> E.squared <- E^2
> demo_1 <- data.frame(X,X.hat,E,E.squared)
> demo_1

  X X.hat  E E.squared
1 1     3 -2         4
2 2     3 -1         1
3 3     3  0         0
4 4     3  1         1
5 5     3  2         4

> SS_1 <- sum(E.squared)
> SS_1

[1] 10
```

Estimating r^2

the Proportion of Variance Accounted For

- With H_0 as our "model" our sum of squared errors was 55.
- With H_1 as our "model" our sum of squared errors is only 10.
- The proportional reduction in the errors is

$$r^2 = \frac{55 - 10}{55} = \frac{45}{55} = 0.81818\ldots \tag{6}$$

- As the sample mean moves away from μ_0, this value will approach 1, because as μ_0 becomes more false, the more we can gain by estimating μ and using that estimate.

Calculating r^2 Directly from the t Statistic

- Exploiting some well known algebraic relationships, it is possible to prove that

$$r^2 = \frac{t^2}{t^2 + df} \tag{7}$$

- Thus, if given a 1-sample t statistic and the df, we can compute r^2 directly.

Calculating r^2 Directly from the t Statistic

- We can demonstrate that with the data set we've been using

```
> X

[1] 1 2 3 4 5

> t <- sqrt(5) * mean(X) / sd(X)
> t

[1] 4.242641

> df <- 4
> t^2/(t^2 + df)

[1] 0.8181818
```