EDP308: STATISTICAL LITERACY

The University of Texas at Austin, Fall 2020

RAZ: Rebecca A. Zárate, MA

Overview

- Hypothesis Testing – One Sample t-test
  - σ Unknown
- Translating Numbers, Again
  - Intervals Around the Null
- One Sample t-test Examples
- One Sample t-tests in R
  - Summary Data
  - Full Datasets

One Sample t-test

Hypothesis Testing – One Sample t-test

- We've actually done a bit of this already with the z and the t tests from the previous section, but now we focus on just the t-test (σ unknown)
  - We were using them as examples for hypothesis testing

$$t_{stat} = \frac{\bar{x} - \mu}{s/\sqrt{n}}$$

Purpose of a One Sample t-test

- A one-sample t-test compares the mean of ONE sample mean ($\bar{x}$) to an assumed population mean (μ).
- To determine if the sample is truly different from an assumed population mean
  - Ex. A group of students that received statistics tutoring vs the normal population of those that do not receive tutoring
- A "Statistically Significant Difference"

One Sample t-test Formula

- The formula one sample t-test statistic should appear familiar:

$$t_{stat} = \frac{\bar{x} - \mu}{s/\sqrt{n}}$$

- Numerator $(\bar{x} - \mu)$ the difference between the sample mean and the assumed population mean
- Denominator $(s/\sqrt{n})$ is the standard error
- t-statistic (t) is the standardized (translated raw mean) score we can compare to a critical value

Try it. Freshman Fifteen

You've heard that the average amount of weight *change* for the first year of college is +15 pounds. To test this claim, you take a sample of 25 UT sophomores and ask them to report the amount their weight has *changed* during the previous year. You calculate summary statistics for your sample as:

$$\bar{x} = 13\ lbs,$$
$$s = 4.0$$

Test the claim of the "Freshman 15" using $\alpha = .05$.

Use the steps of hypothesis testing!

What are the hypotheses? One or two tail test?

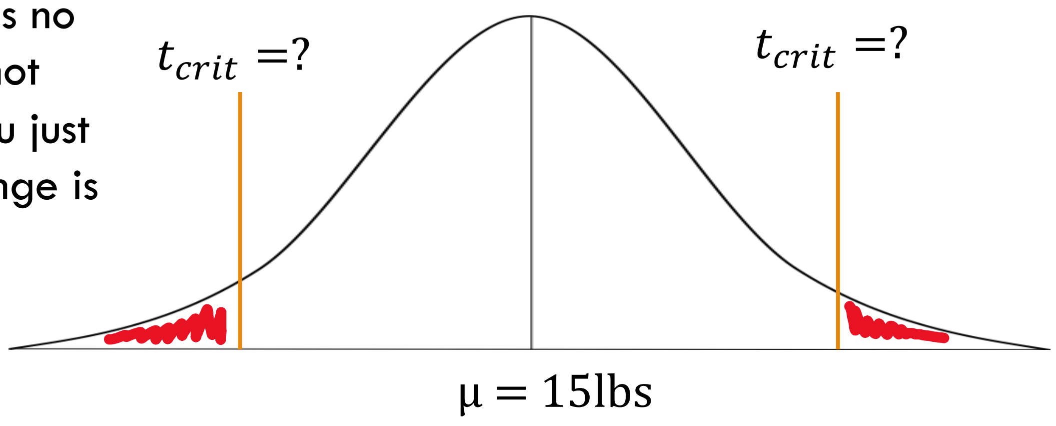
Try it. Freshman Fifteen

Step 1: State the Hypotheses - and draw the picture!

H_0: Freshmen gain an average of 15 pounds during their first year of college.
H_1: Freshmen do not gain an average of 15 pounds during their first year of college.

$$H_0: \mu = 15$$
$$H_1: \mu \neq 15$$

This is a two-tailed test.
(Freshmen may gain less than 15 or more than 15, hence two tailed. There is no assertion of direction. You are not assuming one way or the other, you just want to know if the amount of change is equal to 15 or not.)



Try it. Freshman Fifteen

Step 2: Level of Significance

$$\alpha = .05$$

Step 3: Statistical Test = One sample t-test

$$t_{stat} = \frac{\bar{x} - \mu}{s/\sqrt{n}}$$

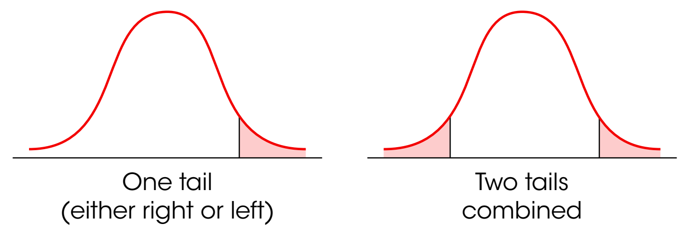
Critical t

- $\alpha = .05$
- n = 25

What is our critical value?

TABLE B.2 **The *t* Distribution**

Table entries are values of *t* corresponding to proportions in one tail or in two tails combined.



	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
7	0.711	1.415	1.895	2.365	2.998	3.499
8	0.706	1.397	1.860	2.306	2.896	3.355
9	0.703	1.383	1.833	2.262	2.821	3.250
10	0.700	1.372	1.812	2.228	2.764	3.169
11	0.697	1.363	1.796	2.201	2.718	3.106
12	0.695	1.356	1.782	2.179	2.681	3.055
13	0.694	1.350	1.771	2.160	2.650	3.012
14	0.692	1.345	1.761	2.145	2.624	2.977
15	0.691	1.341	1.753	2.131	2.602	2.947
16	0.690	1.337	1.746	2.120	2.583	2.921
17	0.689	1.333	1.740	2.110	2.567	2.898
18	0.688	1.330	1.734	2.101	2.552	2.878
19	0.688	1.328	1.729	2.093	2.539	2.861
20	0.687	1.325	1.725	2.086	2.528	2.845
21	0.686	1.323	1.721	2.080	2.518	2.831
22	0.686	1.321	1.717	2.074	2.508	2.819
23	0.685	1.319	1.714	2.069	2.500	2.807
24	0.685	1.318	1.711	2.064	2.492	2.797
25	0.684	1.316	1.708	2.060	2.485	2.787
26	0.684	1.315	1.706	2.056	2.479	2.779
27	0.684	1.314	1.703	2.052	2.473	2.771
28	0.683	1.313	1.701	2.048	2.467	2.763
29	0.683	1.311	1.699	2.045	2.462	2.756
30	0.683	1.310	1.697	2.042	2.457	2.750
40	0.681	1.303	1.684	2.021	2.423	2.704
60	0.679	1.296	1.671	2.000	2.390	2.660
120	0.677	1.289	1.658	1.980	2.358	2.617
∞	0.674	1.282	1.645	1.960	2.326	2.576

Critical t

- $\alpha = .05$
- n = 25
- df = 24
- t-critical =
  - ± 2.064

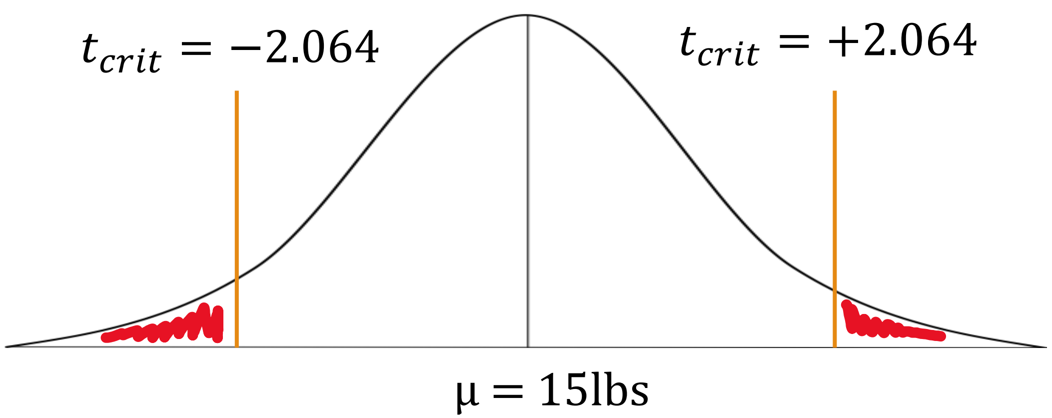
df	Proportion in Two Tails Combined 0.50	0.20	0.10	0.05
1	1.000	3.078	6.314	12.706
2	0.816	1.886	2.920	4.303
3	0.765	1.638	2.353	3.182
4	0.741	1.533	2.132	2.776
5	0.727	1.476	2.015	2.571
6	0.718	1.440	1.943	2.447
7	0.711	1.415	1.895	2.365
8	0.706	1.397	1.860	2.306
9	0.703	1.383	1.833	2.262
10	0.700	1.372	1.812	2.228
11	0.697	1.363	1.796	2.201
12	0.695	1.356	1.782	2.179
13	0.694	1.350	1.771	2.160
14	0.692	1.345	1.761	2.145
15	0.691	1.341	1.753	2.131
16	0.690	1.337	1.746	2.120
17	0.689	1.333	1.740	2.110
18	0.688	1.330	1.734	2.101
19	0.688	1.328	1.729	2.093
20	0.687	1.325	1.725	2.086
21	0.686	1.323	1.721	2.080
22	0.686	1.321	1.717	2.074
23	0.685	1.319	1.714	2.069
24	0.685	1.318	1.711	2.064
25	0.684	1.316	1.708	2.060
26	0.684	1.315	1.706	2.056
27	0.684	1.314	1.703	2.052
28	0.683	1.313	1.701	2.048
29	0.683	1.311	1.699	2.045
30	0.683	1.310	1.697	2.042
40	0.681	1.303	1.684	2.021
60	0.679	1.296	1.671	2.000
120	0.677	1.289	1.658	1.980
∞	0.674	1.282	1.645	1.960

Try it. Freshman Fifteen

Step 4: Find the Critical Value(s)

Finally, we know that our test is a two-tailed test, meaning that our two critical values should cut off tails with probabilities of .025 each (½ of .05 is .025).

The t-critical values that satisfies $df = 24$ and $t_{.025}$ are $t_{crit} = \pm 2.064$.



Try it. Freshman Fifteen

Step 5: Calculating the test statistic

Substituting our sample statistics into the formula below, we have:

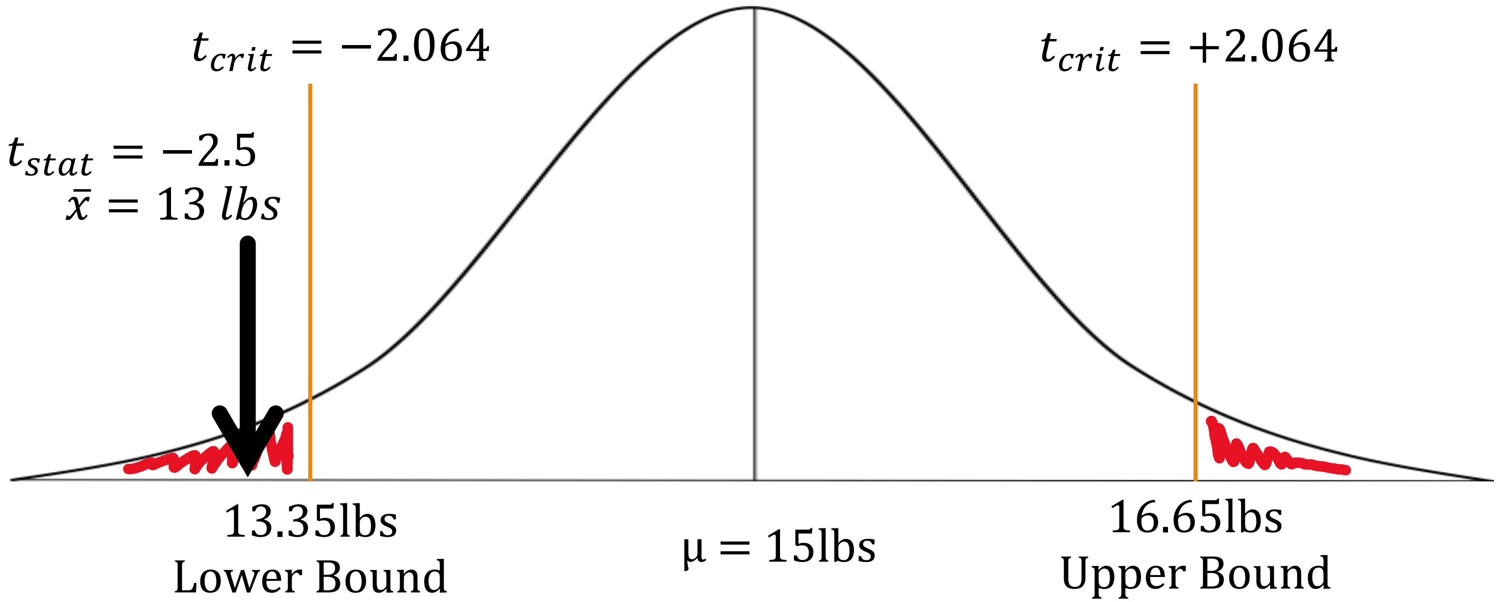
$$t_{stat} = \frac{\bar{x} - \mu}{s/\sqrt{n}} = \frac{13 - 15}{4/\sqrt{25}} = \frac{-2}{.8} = -2.5$$

So, $t_{stat} = -2.5$

Where does our t-statistic fall on the distribution?

Try it. Freshman Fifteen

- This is what it looks like on our curve, assuming that the NULL is true, i.e. +15lbs is the average weight change
  - I've "translated" the critical value into pounds, what do you think our results will be?

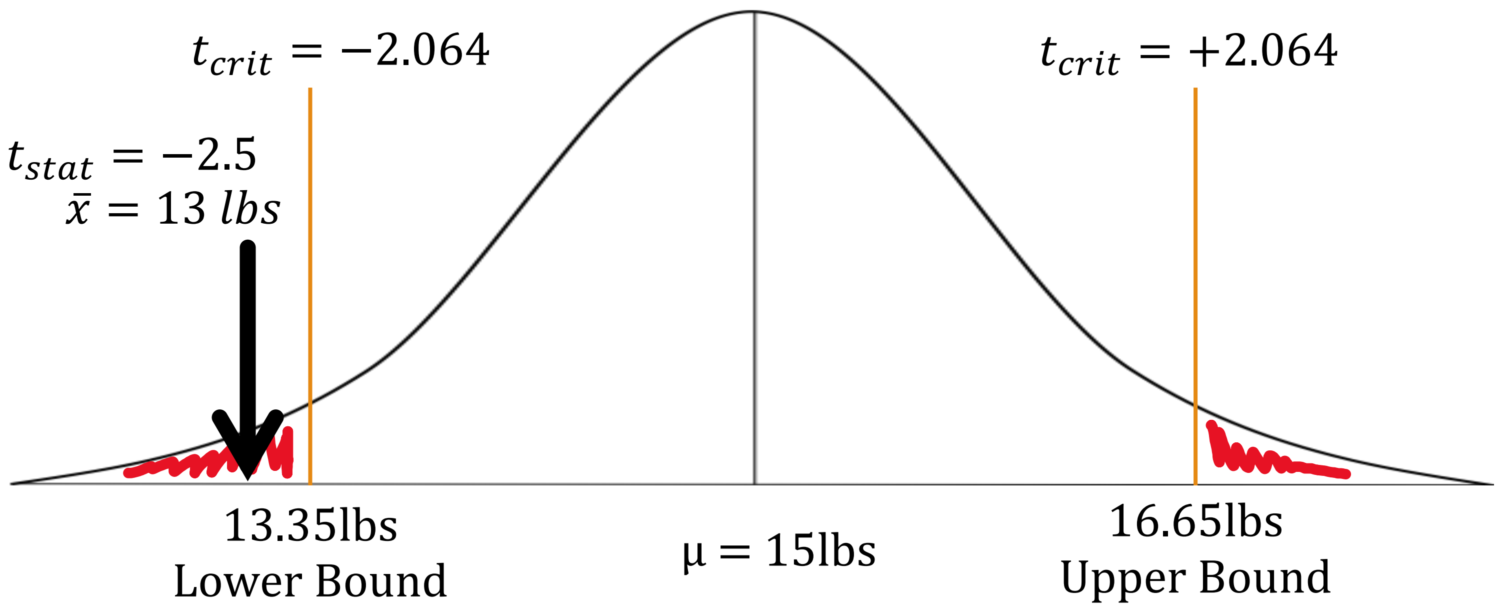


Try it. Freshman Fifteen

Step 6: Make a Conclusion

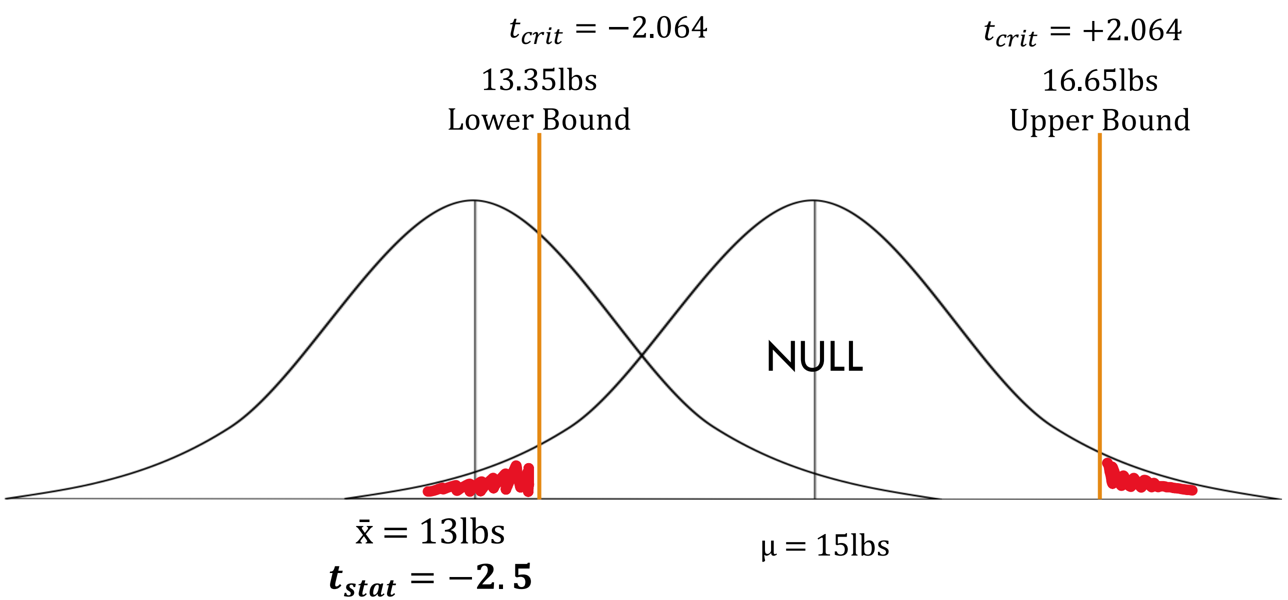
Our $t_{stat} = -2.5$, and our $t_{crit} = \pm 2.064$

Our t_{stat} is past our t_{crit}, so we reject H_0



Think Critically…

- Our results were “statistically significant”, but what does that mean practically? What could it mean in regards to our sampled group?
  - Maybe UT students don’t gain as much weight
    - Maybe Austin is a healthier city
    - Maybe you sampled some athletes and your sample was not representative
  - Also, it would not have been significant if it was just .35lbs more…



Translating Numbers, Again

Translating Between Scales

- I have been including numbers in the standardized t-scale and numbers in the original scale of the question
  - This is not always common practice, but I find it helps makes the idea more tangle
- You can make the same conclusions as your t-test, but the scientific community prefers standardized numbers (because we aren't always familiar with the original scale)
  - You still have to report you t-test statistic, but if you want to double check your work, here's how…

Translating Critical Values to Raw Scores

We don't actually create CI for Nulls, this is just for the example and for translating.

Margin of Error:
The amount of error based on SE and a desired level of confidence in the original number scale

$$\text{“}CI_{Null}\text{”} = \mu \pm t_{critcal} * \frac{\sigma}{\sqrt{n}}$$

Point Estimate:
The **<u>NULL's mean</u>**, in the original number scale

t-critical:
The level of significance I want (ex. $\alpha = .05$, t = refer to chart) in standardized scale

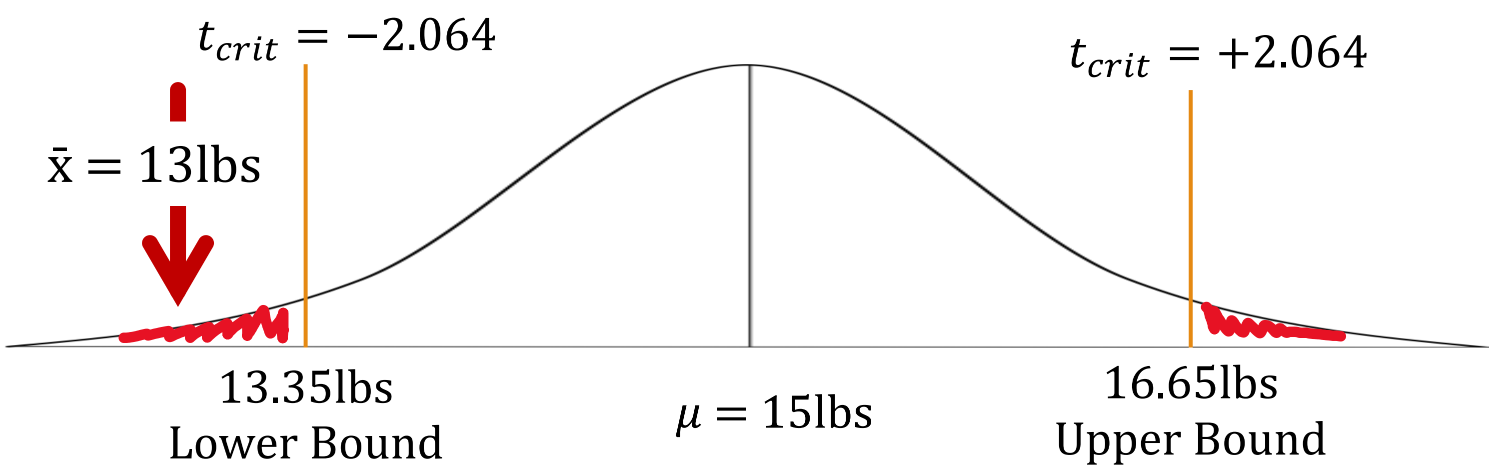
Standard Error (SE):
Standard deviation of the sampling distribution

Translating Between Scales

$$\text{“}CI_{Null}\text{”} = \mu \pm t_{critcal} * \frac{\sigma}{\sqrt{n}}$$

$$\text{“}CI_{Null}\text{”} = 15\text{lb} \pm 2.064 * \frac{4}{\sqrt{25}}$$

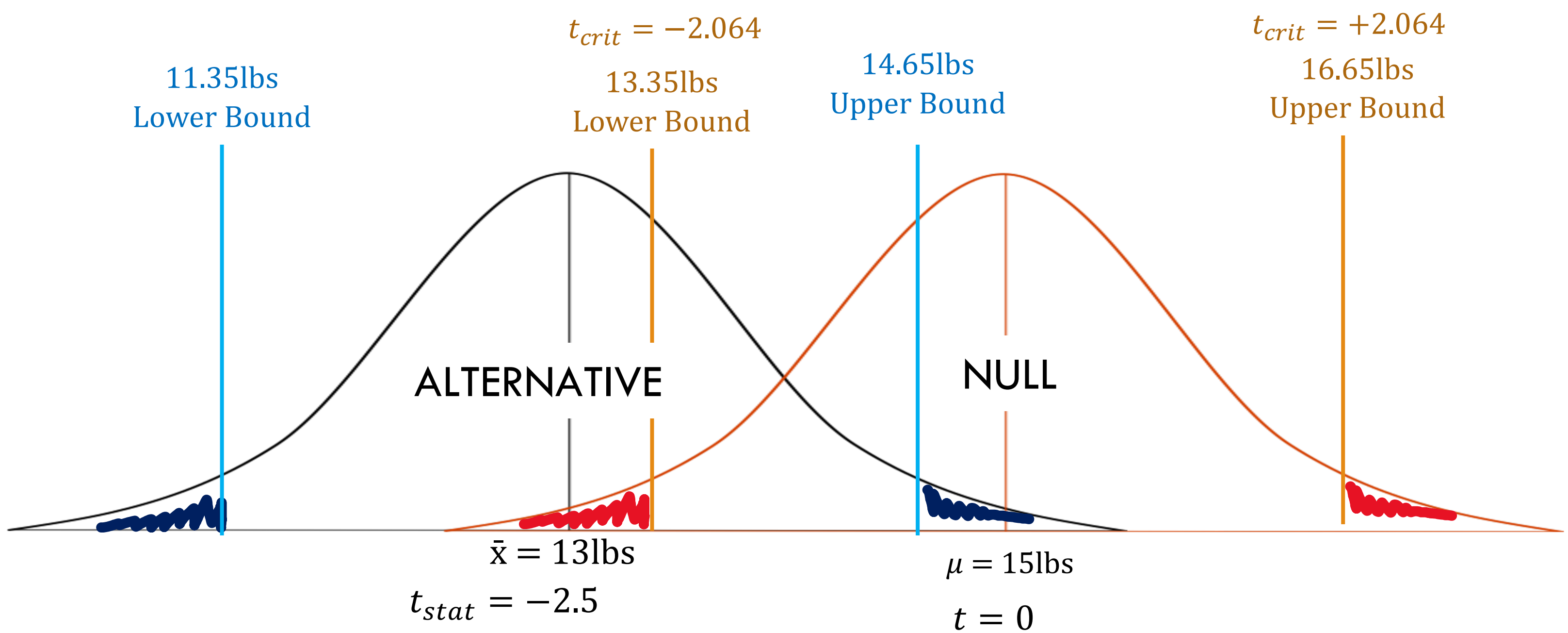
$$[13.35, 16.65] = 15\text{lbs} \pm 1.65$$



Confidence Interval for Our Sample Mean

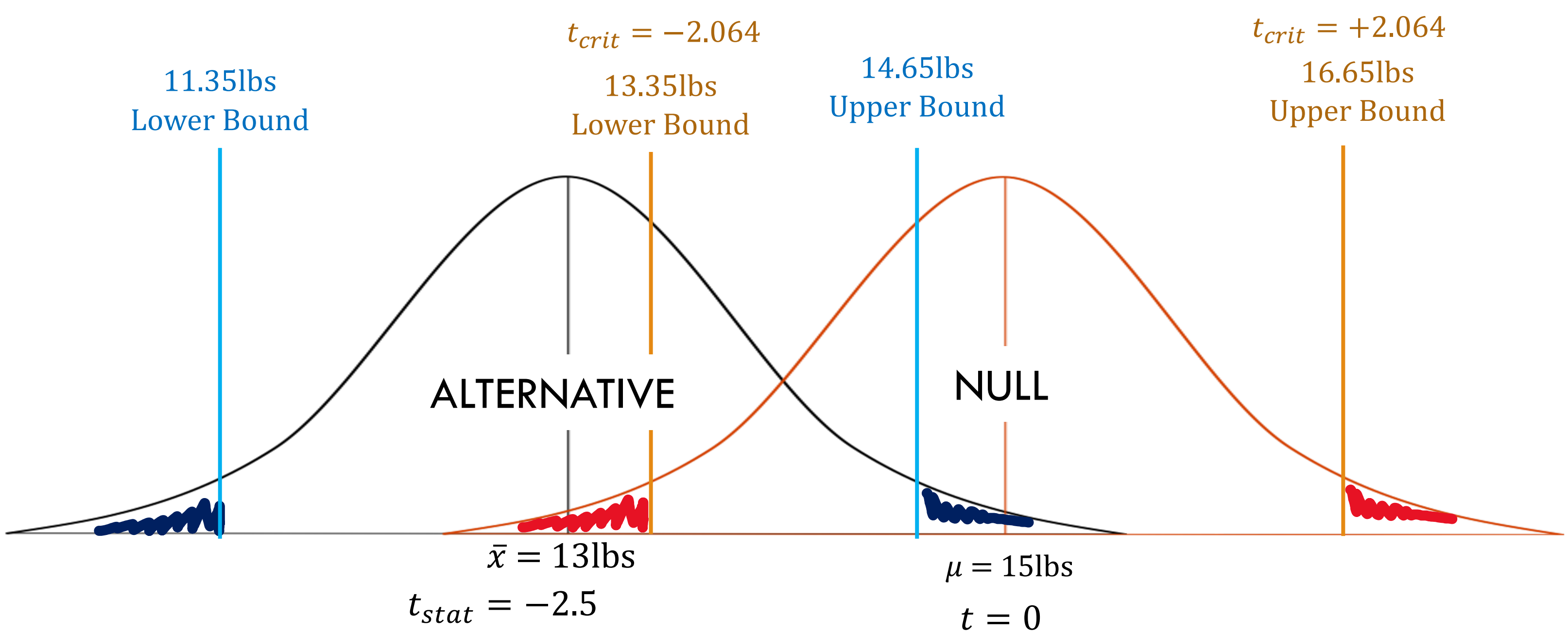
$$CI = 13\text{lb} \pm 2.064 * \frac{4}{\sqrt{25}}$$

$$[11.35, 14.65] = 13\text{lbs} \pm 1.65$$



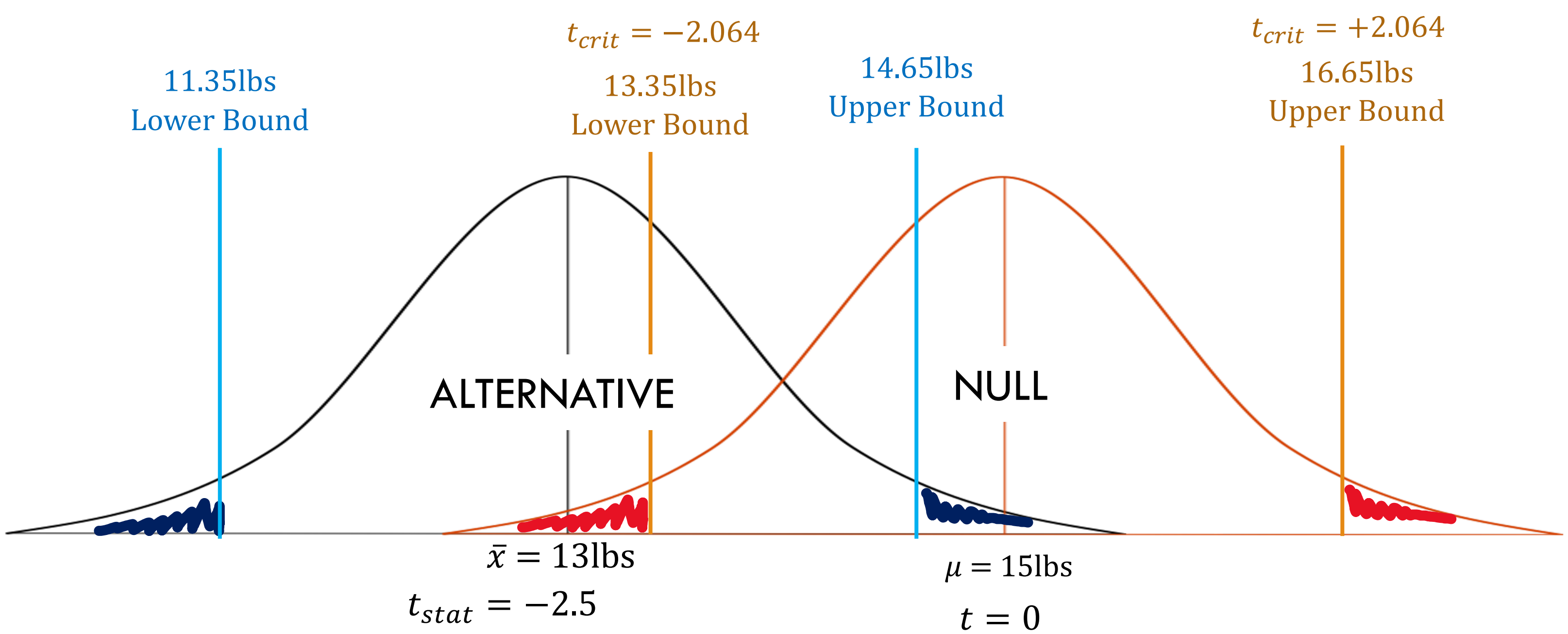
Comparing

- Notice how the confidence interval for our sample statistic ($\bar{x} = 13$) does *not* include 15 lbs (μ), the population parameter that we were testing against. Similarly, the "confidence interval" around the null 15 lbs (μ) does not contain the sample mean of 13 lbs.



Conclusions and Chances

- Considering the sample size and standard error, the sample mean of $\bar{x} = 13$ is not a reasonable value in the null world. It is extreme enough for us to conclude that the probability of seeing a $\bar{x} = 13$ if the null were true is less than .05, less than a 5% chance.



One Sample t-tests are for…

- Situations when:
  - You are comparing your sample mean to a **known population mean**
    - Ex. Did this school district score lower than the national average?
    - Ex. Are UT students more or less likely to gain the national average of 15lbs their Freshman year?
    - Ex. Does tracking behavior increase the number of minutes/week of exercise compared to the average of those who don't track their behavior (30 minutes/week)?
  - σ **unknown**
    - If it was known, you'd use a z-test

Up Next…

- One Sample t-tests are for one sample to compare against some known population mean, but what if we want to compare two groups that we sampled?
- Then we need an…

Independent Samples t-test

But first a quick detour to Effect Sizes (Cohen's d)

One Sample t-tests in R

One Sample t-tests in R: Freshman 15

Using summary statistics.

```
#########################################
##### t-test One sample, TWO Tailed #####
#########################################

# First fill in the information that is know: Freshman 15

sample_mean <- 13
pop_mean <- 15
sd <- 4
n <- 25
df <- n-1

# Find the critical values: 95% Critical t-values for 95% Confidence, significance = 0.05
two_tail_crit_t_95 <- qt(p = c(.025, .975), df = df) # critical ts = -2.064, 2.2064

# Calculate the t-statistic
t_stat_one_sample <- (sample_mean - pop_mean)/(sd/sqrt(n)) # t-statistic = -2.5

# Our t-statistic (-2.5) is further out than the critcal t values (-2.064, 2.2064)
# We reject the null hypothesis.
```

Confidence Interval: Freshman 15

Using summary statistics.

```
########################################
#### Confidence Intervals for One Sample ###
############# t -test ###################
########################################

# Calculate the Margin of Error
# This will be two numbers, the +MOE and the -MOE
moe <- two_tail_crit_t_95 * sd/sqrt(n)

# You only have to add the MOE to the average to get back the upper and lower bound
confidence_interval <- sample_mean + moe

# The 95% CI [11.35, 14.65]
```

One Sample t-tests in R with Data

Using data.

```
# This is how to do a One-Sample t-test with data (rather than summary statistics)

# R has some datasets included, we can use the "InsectSprays" data set for the example
data <- InsectSprays

# R will automatically calculate the means number of bug bites when you put "data" in
# Then you tell R what the population mean you are testing against is
# Let's say the average number of bug bites for someone not wear any bug spray is mu = 14
# We want to do a two-tailed test to see if there is any difference wear bug spray

t.test(x = data$count, mu = 14, alternative = "two.sided")

# We could also run a one-tailed test, either left or right, by specifying
# alternative = "less" or alternative = "greater" respectively
# Here is would make more sense to do a left sided, LESS than, test bc it looks
# like people who wear bug spray get less bug bites

t.test(x = data$count, mu = 14, alternative = "less")

# Both of these tests were significant. R also gives you the t-statistic, degrees of freedom,
# a p-value, a confidence interval, and the mean difference between the sample and mu
```

Data Source: R Datasets “InsectSpray”
Beall, G., (1942) The Transformation of data from entomological field experiments,*Biometrika*, **29**, 243–262.

One Sample t-tests in R with Data Output

```
> t.test(x = data$count, mu = 14, alternative = "two.sided")

        One Sample t-test

data:  data$count
t = -5.3009, df = 71, p-value = 1.241e-06
alternative hypothesis: true mean is not equal to 14
95 percent confidence interval:
  7.807311 11.192689
sample estimates:
mean of x
      9.5
```

Two Tailed

```
> t.test(x = data$count, mu = 14, alternative = "less")

        One Sample t-test

data:  data$count
t = -5.3009, df = 71, p-value = 6.204e-07
alternative hypothesis: true mean is less than 14
95 percent confidence interval:
    -Inf 10.9148
sample estimates:
mean of x
      9.5
```

One Tailed (Left)

Less Than