Not much left!

Unit 8 – Hypothesis Testing Day 2

Your Realizing-summer-is-almost-over
Professor Colton

Unit 8 Day 2- Outline

<u>Unit 8 – Hypothesis Testing</u>

Hypothesis Testing for Population Means with KNOWN SD

- Define Parameter and State Hypotheses
- Rejection Region
- Test Statistic, P-value Method and Conclusion

Hypothesis Testing for Population Means with UNKNOWN SD

- Similarities and Differences
- Calc work for Test Statistic and P-value Method

Decisions in Hypothesis Tests

- Describing Type 1 and 2 Errors

Hypothesis Tests for Means with KNOWN $\boldsymbol{\sigma}$!

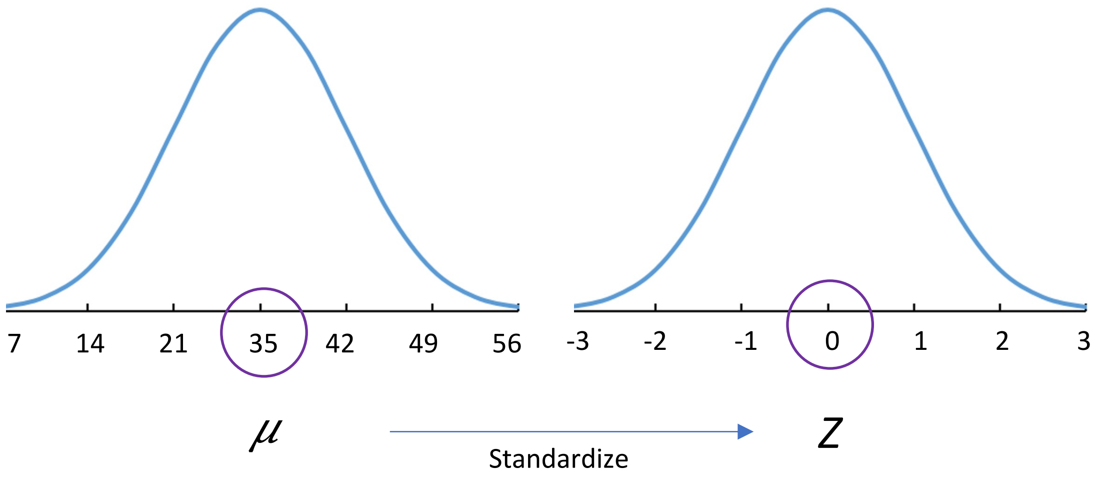
- All of the previous Hypothesis tests overview applies, now we are just going to apply it specifically to a Means Test!

- And going back to the Confidence Interval unit, we have a known population standard deviation! So the same logic and implications of that apply here as well!

- We will be doing a **Z-Test**!

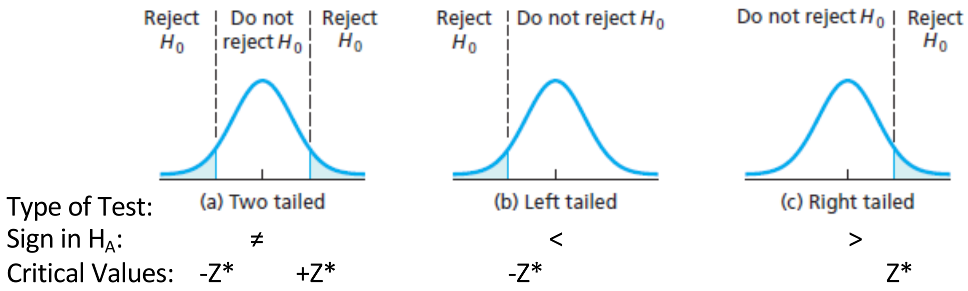
The Hypothesis Statements - Review

1. State the Hypotheses
- Define parameter + context.



<u>Define Parameter</u>

- Always define your **parameter** at the start!
- Think about the variable / quantity of interest!
  - Quantitative (numeric) → population mean μ



<u>Null Hypothesis</u> H_0

- This is the <u>status quo</u>, typically a *known* value of the parameter (μ_0)
  - **When written symbolically ALWAYS =**

<u>Alternative Hypothesis</u> H_A

- May be left-tailed (<), right-tailed (>), or two-tailed (≠).
  - Uses the <u>same value</u> of the parameter as in the Null hypothesis H_0

<u>In general</u>

- H_0: $\mu = \mu_0$ and H_A: μ ≠,<,> μ_0

Examples:

'Research from previous studies suggests the average number of people

- Equal to → H_0: $\mu = 7$

'The owner believes his average monthly profit is more than $50,000'

- In this case, greater than → H_A: $\mu > 50,000$

LCQ – Hypotheses

Problem: (1) Define the parameter of interest and (2) State the Null and Alternative Hypotheses and the directionality of the test (two-tailed, left-tailed or right-tailed) for the following scenarios:

a) A random sample of 15 human body temperatures were obtained. Assume that human body temperatures are normally distributed. Is there sufficient evidence to conclude that the true mean human body temperature differs from 98.6ºF?

b) In 2012, a large number of foreclosed homes in Washington, D.C. were sold. Real estate experts say the standard deviation for sales the past 10 years was $190,000. In one community, a random sample of 30 foreclosed homes sold for an average of $443,705. A prospective home-buyer wants to know if prices have decreased from the 2002 average of $450,000.

c) Test 1 grades on the most fun class you've ever taken averaged 80.76 with standard deviation 13.34 points. From a random sample of 19 Test 2 grades, there was a mean of 83.39. Your super cool instructor wants to know if the Test 2 grades improved.

LCQ – Hypotheses

Problem: (1) Define the parameter of interest and (2) State the Null and Alternative Hypotheses and the directionality of the test (two-tailed, left-tailed or right-tailed) for the following scenarios:

a) A random sample of 15 human body temperatures were obtained. Assume that human body temperatures are normally distributed. Is there sufficient evidence to conclude that the true mean human body temperature **differs** from 98.6°F?

Let μ = the true mean of human body temperature → VERY GOOD!

First try:
$H_0 = 98.6$ and $H_A \neq 98.6$ → INCORRECT! MISSING μ!!!

H_0: $\mu = 98.6$ and H_A: $\mu \neq 98.6$ → two-tailed → NOW IT'S PERFECT!

b) In 2012, a large number of foreclosed homes in Washington, D.C. were sold. Real estate experts say the standard deviation for sales the past 10 years was $190,000. In one community, a random sample of 30 foreclosed homes sold for an average of $443,705. A prospective home-buyer wants to know if prices have **decreased** from the 2002 average of $450,000.

First try:
H_0: $p = 450{,}000$ and H_A: $p < 450{,}000$ → INCORRECT!! NOT TALKING ABOUT PROPORTIONS!!! SHOULD BE NO Ps

Let μ = the true mean of the prices of foreclosed homes → PERFECT

H_0: $\mu = 443{,}705$ and H_A: $\mu < 450{,}000$ → INCORRECT! Remember both numbers in the H_0 and H_A need to be the same! So which one is correct?????
Hypotheses are NEVER based on SAMPLE information!!!! H_0: $\mu = 450{,}000$ and H_A: $\mu < 450{,}000$ left-tailed → NOW CORRECT!

c) Test 1 grades on the most fun class you've ever taken averaged 80.76 with standard deviation 13.34 points. From a random sample of 19 Test 2 grades, there was a mean of 83.39 Your super cool instructor wants to know if the Test 2 grades **improved**.

Both correct (I won't be too picky with the context, technically we are looking at our Test 2 grade average and comparing that to a value from the Test 1 (the null):
Let μ be the true mean of the test 2 grades.
Let μ = represent the true mean of test scores

H_0: $\mu = 80.76$ and H_A: $\mu > 8.076$ right-tailed → CORRECT!

Rejection Region for Means with Known $\boldsymbol{\sigma}$

2. Sketch Rejection Region based of Significance Level

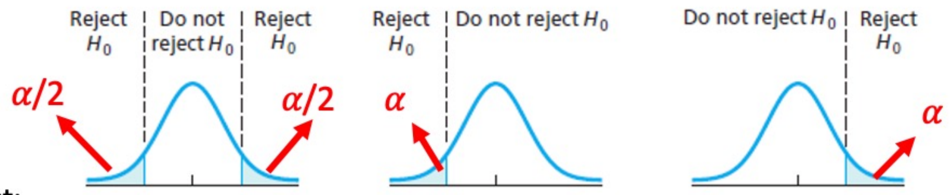
<u>Rejection Region for Means Test with KNOWN $\boldsymbol{\sigma}$</u>

- This is the SAME as for a Proportions Test!

<u>Review</u>

- We have to determine the the when <u>there is or is not enough evidence</u> against the Null.

- Our Rejection Region (RR) is based on whether we are doing a one or two tailed test (this is the direction from the H_A)!



Type of Test:	(a) Two tailed	(b) Left tailed	(c) Right tailed
Sign in H_A:	$\neq$	$<$	$>$
Rejection Region:	Both sides	Left side	Right side
	Z	Z	Z

Using Calc - Test Statistic and P-Value for Means with Known σ

3. **Compute P-value** (and Test Statistic).

(Original) Setup
Scientists discovered a new mountain range under the sea. Lets assume the sea mountain heights are normally distribution with known standard deviation of 5000 ft.

From a random sample of 13 peaks, there was an average height of 12,000 ft. Is there enough evidence to conclude the average heights of these new sea mountains is different than the Rocky Mountains, which average 14,400 ft? Use α = 0.12

Formula for Z_{stat} by hand:

Test for	H_0	Test statistic
Pop. mean μ	$\mu = \mu_0$	$z = \frac{\bar{x} - \mu_0}{\frac{\sigma}{\sqrt{n}}}$

GOAL: Conduct a Hypothesis Test!

1. Z-Test
   a) Input = Stats
   b) μ_0 = the Null mean
   c) σ = population SD
   d) $\bar{x}$ = sample mean
   e) n = sample size
   f) μ: Alternative hypothesis

   Calculate or Draw

New Scenario
Three more sea mountains were discovered. The new sample mean is equal to 15,000 ft. Is there enough evidence to say the sea mountains are taller than the Rockies. Use α = 0.10

- Run another ZTest

Using Calc - Test Statistic and P-Value for Means with Known σ

3. **Compute P-value** (and Test Statistic).

(Original) Setup
Scientists discovered a new mountain range under the sea. Lets assume the sea mountain heights are normally distribution with known standard deviation of 5000 ft.

From a random sample of 13 peaks, there was an average height of 12,000 ft. Is there enough evidence to conclude the average heights of these new sea mountains is different than the Rocky Mountains, which average 14,400 ft? Use α = 0.12

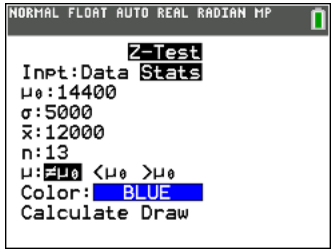
Formula for Z_{stat} by hand:

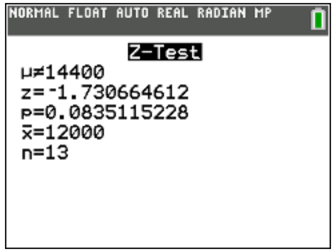
Test for	H_0	Test statistic
Pop. mean μ	$\mu = \mu_0$	$z = \dfrac{\bar{x} - \mu_0}{\frac{\sigma}{\sqrt{n}}}$

GOAL: Conduct a Hypothesis Test!

1. Z-Test
   a) Input = Stats
   b) μ_0 = the Null mean
   c) σ = population SD
   d) $\bar{x}$ = sample mean
   e) n = sample size
   f) µ: Alternative hypothesis

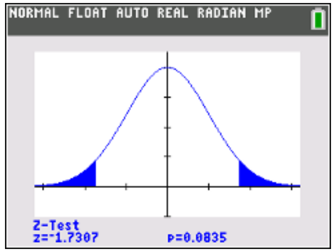
   Calculate or Draw





H_0: μ = 14,400
H_A: μ ≠ 14,400

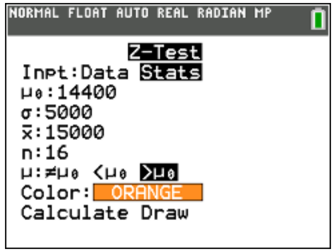
Calculate Output
µ = Alternative hypothesis
z = Z_{stat}
p = p-value
$\bar{x}$ = sample proportion
n = sample size

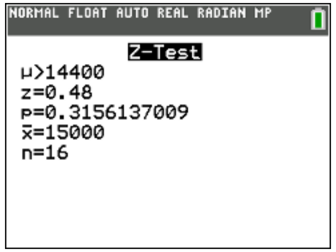


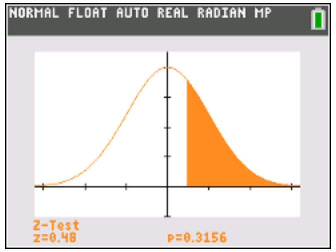
Draw Output
Plot (and displays values) of p = p-value and z = Z_{stat} on the standard normal curve

New Scenario
Three more sea mountains were discovered. The new sample mean is equal to 15,000 ft. Is there enough evidence to say the sea mountains are **taller** than the Rockies. Use α = 0.10
- Run another ZTest

H_0: μ = 14,400
H_A: μ > 14,400







LCQ – Conclusions and Interpretations

4. Conclude and Interpret

- State whether you reject H_0 or fail to reject H_0 AND WHY!
- Interpret your results in the context of the problem

Problem: Write the conclusions and interpretations for the previous scenarios using our results.

a) Original Setup: Scientists discovered a new mountain range under the sea. Lets assume the sea mountain heights are normally distribution with known standard deviation of 5000 ft. From a random sample of 13 peaks, there was an average height of 12000 ft. Is there enough evidence to conclude the average heights of these new sea mountains is different than the Rocky Mountains, which average 14,400 ft? Use α = 0.12

b) New Scenario: Three more sea mountains were discovered. The new sample mean is equal to 15,000 ft. Is there enough evidence to say the sea mountains are taller than the Rockies. Use α = 0.10

LCQ – Conclusions and Interpretations

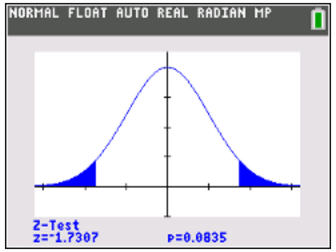
4. Conclude and Interpret
- State whether you reject H_0 or fail to reject H_0 AND WHY!
- Interpret your results in the context of the problem

Problem: Use the specified method to write the conclusions and interpretations for the previous scenarios using our results.

a) Original Setup: Scientists discovered a new mountain range under the sea. Lets assume the sea mountain heights are normally distribution with known standard deviation of 5000 ft. From a random sample of 13 peaks, there was an average height of 12000 ft. Is there enough evidence to conclude the average heights of these new sea mountains is different than the Rocky Mountains, which average 14,400 ft? Use α = 0.12

Need these…

$H_0: \mu = 14,400$
$H_A: \mu \neq 14,400$
$\alpha = 0.12$



P-Value

p-value = ZTest(μ_0 = 14400, σ = 5000, $\bar{x}$ = 12000, n = 13, $\mu \neq \mu_0$) = 0.0835
Comparing p-value to alpha → Decision on H_0: p-value = 0.0835 < 0.12 = α → Reject H_0!

Conclusion / Interpretation **has TWO sentences**
Sentence 1: Put comparison in words: *P-value = 0.08 is less than alpha = 0.12, so we reject the null hypothesis*

Sentence 2: There (is / is not) enough evidence to conclude (alternative hypothesis is words)
There IS enough evidence to conclude ($\mu \neq 14,400$) *the* true mean height of sea mountains is not equal to 14,400 ft

LCQ – Conclusions and Interpretations Cont…

4. **Conclude** and **Interpret**

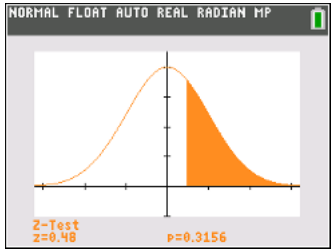
- State whether you reject H_0 or fail to reject H_0 AND WHY!
- Interpret your results in the context of the problem

Problem: Use the specified method to write the conclusions and interpretations for the previous scenarios using our results.

b) New Scenario: Three more sea mountains were discovered. The new sample mean is equal to 15,000 ft. Is there enough evidence to say the sea mountains are taller than the Rockies. Use α = 0.10

Need these (new) …

$H_0: \mu = 14,400$
$H_A: \mu > 14,400$
$\alpha = 0.10$



P-Value

p-value = ZTest(μ_0 = 14400, $\boldsymbol{\sigma}$ = 5000, $\bar{x}$ = 15000, n = 16, $\mu > \mu_0$) = 0.3156
p-value = 0.3156 < 0.10 = $\alpha \rightarrow$ Fail to Reject H_0

Conclusion and Interpretation

Because our p-value = 0.3156 is greater than the significance level 0.10, we fail to reject the Null hypothesis. There is NOT sufficient evidence to conclude that the true mean height of the sea mountains are greater than 14,400 ft, which is the average height of the Rocky Mountains.

Hypothesis Tests for Means with UNKNOWN **σ**!

- All of the previous Hypothesis tests overview and for Means applies, now we are just going to look at the situation when we have an unknown **σ** (unlike before)!

- And going back to the Confidence Interval unit, we have now have an unknown population standard deviation! So the same logic and implications of that apply here as well!

- We will be doing a **T-Test**!

T Test vs Z-Test - Step Similarities

1. **State the Hypotheses**
   - **Define parameter + context.**

2. **Sketch** Rejection Region based of Significance Level

3. **Compute** P-value (and Test Statistic).

4. **Conclude** and **Interpret**
   - State whether you reject H_0 or fail to reject H_0 AND WHY!
   - Interpret your results in the context of the problem

Samesies

- EXACT same overall process as for testing a Mean with known **σ**!
- BUT NOW, the population standard deviation **σ** is UNKNOWN and we only have the sample standard deviation *s*…
- And the Conclusion / Interpretation is the same as well!

Different

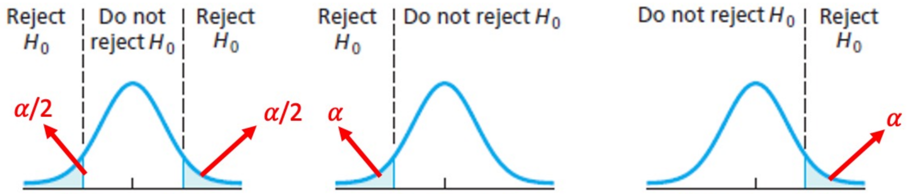
- The ONLY steps that involve something sightly different are the Rejection Region and the Test Statistic / P-Value

Rejection Region for Means with Unknown **σ**

2. Sketch Rejection Region based of Significance Level

Rejection Region for Means Test with UNKNOWN **σ**

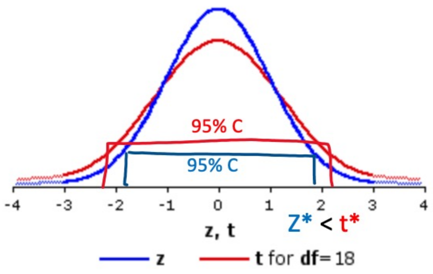
- It is now based on the ***t*-distribution** rather than the standard normal distribution Z (just like it was for the Confidence Intervals).
  - So the curve we draw is a T-curve!
  - (All of the CVs will be **t*s** with the **correct degrees of freedom** *df)*



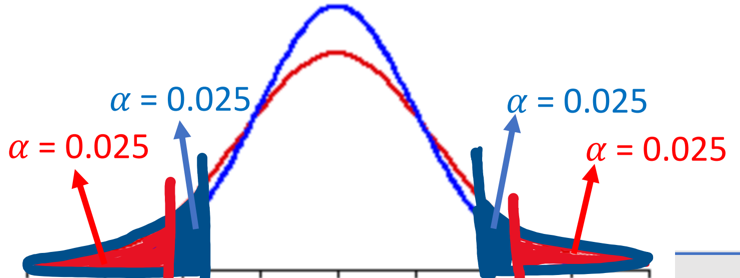
Type of Test:	(a) Two tailed	(b) Left tailed	(c) Right tailed
Critical Value(s):	-t* +t*	-t*	+t*
Rejection Region:	Both sides	Left side	Right side

Effect of Z vs T on Hypothesis Tests

- Recall from CI that because we have to **estimate σ** with ***s***, there is inherently more variability (uncertainty) which produces wider *t*-intervals compared to Z-intervals.



- In Hypothesis Tests, this translates to **Rejections Regions** being further away from the center for the same α!
- Which means we need to have more extreme results in order to reject when switching from Z to T Tests



Using Calc - Test Statistic and P-Value for Means with Unknown σ

3. **Compute P-value** (and Test Statistic).

Formula for t_{stat} by hand:

Test for	H_0	Test statistic
Pop. mean μ	$\mu = \mu_0$	$t = \frac{\bar{x} - \mu_0}{\frac{s}{\sqrt{n}}}$

(ALMOST SAME Original) Setup
Scientists discovered a new mountain range under the sea. Lets assume the sea mountain heights are normally distribution.

From a random sample of 13 peaks, there was an average height of 11,308 ft and standard deviation of 5,287 ft. Is there enough evidence to conclude the average heights of these new sea mountains is different than the Rocky Mountains, which average 14,400 ft? Use α = 0.12

*** Only a sample standard deviation is provided, so σ is unknown → T-Test*

GOAL: Conduct a Hypothesis Test!

1. T-Test

- Option 1) Input = Stats

a) μ_0 = the Null mean

b) $\bar{x}$ = sample mean

c) Sx = sample SD

d) n = sample size

e) μ: Alternative hypothesis

Calculate or Draw

- Option 2) Input = Data
  - Enter raw data in L_1

a) μ_0 = the Null mean

b) List = L1

c) Freq = 1

d) μ: Alternative hypothesis

Calculate or Draw

Height
7700
11500
16000
5800
9000
5500
12400
22100
14200
19300
8300
9100
6100

Mean = 11308
SD = 5287

Using Calc - Test Statistic and P-Value for Means with Unknown σ

3. **Compute P-value** (and Test Statistic).

Formula for t_{stat} by hand:

Test for	H_0	Test statistic
Pop. mean μ	$\mu = \mu_0$	$t = \frac{\bar{x} - \mu_0}{\frac{s}{\sqrt{n}}}$

(ALMOST SAME Original) Setup
Scientists discovered a new mountain range under the sea. Lets assume the sea mountain heights are normally distribution.

From a random sample of 13 peaks, there was an average height of 11,308 ft and standard deviation of 5,287 ft. Is there enough evidence to conclude the average heights of these new sea mountains is different than the Rocky Mountains, which average 14,400 ft? Use α = 0.12

*** Only a sample standard deviation is provided, so σ is unknown → T-Test*

GOAL: Conduct a Hypothesis Test!

1. T-Test

Height
7700
11500
16000
5800
9000
5500
12400
22100
14200
19300
8300
9100
6100
Mean = 11308
SD = 5287

- Option 1) Input = Stats

a) μ_0 = the Null mean

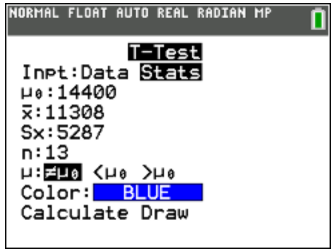
b) $\bar{x}$ = sample mean

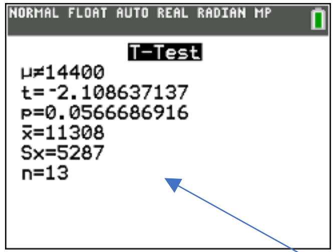
c) Sx = sample SD

d) n = sample size

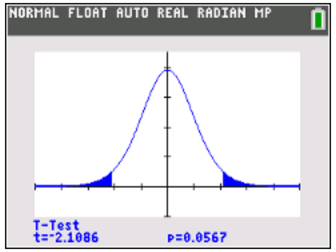
e) μ: Alternative hypothesis

Calculate or Draw





Calculate Output
μ = Alternative hypothesis
t = t_{stat}
p = p-value
$\bar{x}$ = sample proportion
Sx = sample SD
n = sample size



Draw Output
Plot (and displays values) of p = p-value and t = t_{stat} on the t curve with df = n -1

H_0: μ = 14,400

H_A: μ ≠ 14,400

Same results! (maybe a little roundoff error from $\bar{x}$ *or s)*

T-Test(input = Stats, μ0 = 14,400, x-bar = 11308, Sx = 5287, n = 13, μ ≠ μ0)

- Option 2) Input = Data
  - Enter raw data in L_1

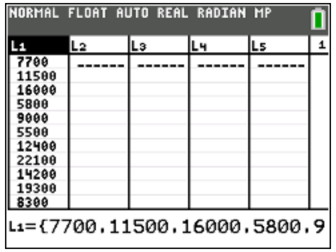
a) μ_0 = the Null mean

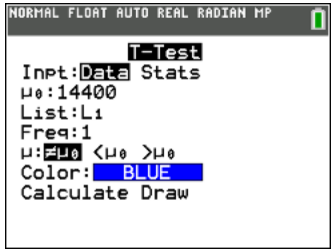
b) List = L1

c) Freq = 1

d) μ: Alternative hypothesis

Calculate or Draw

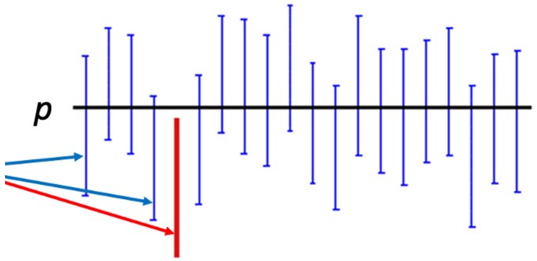




T-Test(input = Data, μ0 = 14,400, List = L1, Freq = 1, μ ≠ μ0)

Decisions in Hypothesis Tests

- Recall from Confidence Intervals that it's not a guarantee that our interval captures the true population parameter!
    - As researchers we can do things to minimize the chances this happens such as having a high confidence level and a large sample size.
    - But we are working with real data when we take samples, and this is what your intervals are based on. There is always the possibility that our sample data leads us astray resulting in an interval that misses the population parameter ☹



A 95% confidence interval indicates that 19 out of 20 samples (95%) from the same population will produce confidence intervals that contain the population parameter.

- Of course, we never actually know if we capture or don't capture (because we don't know the truth). But it's something that we have to keep in mind when interpreting and making decisions based on our results.
- **This same dilemma is present in Hypothesis Tests as well!**
- There is the real possibility that we are making the WRONG conclusion to either Reject or Fail to Reject the Null hypothesis.
    - These are called **Type 1 and Type 2 Errors**!

Incorrect Decisions Example

Null Hypothesis: Not pregnant
~~Alternative Hypothesis: Pregnant~~

WRONG CONCLUSION!

Telling someone they <u>ARE pregnant</u> when in reality they are NOT.

ANOTHER WRONG CONCLUSION!

Telling someone they are <u>NOT pregnant</u> when in reality they ARE.

Type I and Type II Error

Type I Error

- This occurs when we incorrectly **reject** a TRUE Null hypothesis
  - False Positive → The test says you have COVID, but you actually don't
- *(Reject the Null Hypothesis and)* ***Conclude the Alternative***, *when in reality the Null Hypothesis is actually correct.*
  - Recall α is the probability of rejecting the Null hypothesis. This also means that …
  - Probability of committing Type I error = α

Type II error

- This occurs when we **fail to (2) reject** a FALSE Null hypothesis
  - False Negative → The test says you don't have COVID, but you actually do
- *(Fail to (2) Reject the Null Hypothesis and)* ***Conclude the Null***, *when in reality the the alternative is actually correct.*

** If I ask you to describe a Type 1 or Type 2 error, these are the structures you should use + CONTEXT!!

There's two layers here:

1. The TRUTH (which we don't actually know)
2. Our DECISION (which he hope is correct)

Given the Null Hypothesis Is

Your Decision Based On a Random Sample	True	False
Reject	Type I Error	Correct Decision
Do Not Reject	Correct Decision	Type II Error

Two Types of Errors in Decision Making

https://home.ubalt.edu/ntsbarsh/Business-stat/opre504.htm

Errors Example

Your Decision Based On a Random Sample		Given the Null Hypothesis Is: True	Given the Null Hypothesis Is: False
	Reject	Type I Error	Correct Decision
	Do Not Reject	Correct Decision	Type II Error

Two Types of Errors in Decision Making

Setup: All commercial elevators must pass yearly inspections. An inspector has to choose between certifying an elevator as safe (no repairs needed) or saying that the elevator is not safe (repairs are needed). There are two hypotheses:

H_0: The elevator is not safe (repairs are needed)

H_A: The elevator is safe (no repairs needed)

MAYBE ADD EXAMPLE WHERE WE ACTUALLY USE NUMBERS FROM A HYPOTHESIS TEST AND THEN DETERMINE WHICH TYPE OF ERROR WAS MADE BY (example from Connect HW below)

a) Describe Type I error.

b)Describe Type II error.

Determine whether the outcome is a Type I error, a Type II error, or a correct decision.

A test is made of $H_0: \mu = 30$ versus $H_1: \mu \neq 30$.

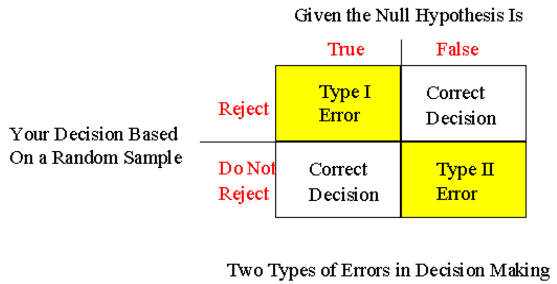
The true value of μ is 30, and H_0 is rejected.

The outcome of the test is a (Choose one) ▼.

- Type I error
- correct decision
- Type II error

Errors Example - Solution

https://home.ubalt.edu/ntsbarsh/Business-stat/opre504.htm



H_0: The elevator is not safe (repairs are needed)

H_A: The elevator is safe (no repairs needed)

a) Describe Type I error → *(Reject the Null Hypothesis and)* ***Conclude the Alternative***, *when in reality the Null Hypothesis is actually correct*

- *Description → A Type I error is wrongly concluding that the elevator is safe and no repairs are needed,* ***when actually the elevator is not safe.***
    - *We add the phrase "wrongly concluding" because we are committing an error, so of course our conclusion is incorrect*

b) Describe Type II error → *(Fail to (2) Reject the Null Hypothesis and)* ***Conclude the Null***, *when in reality the the alternative is actually correct*

- *Description → A Type II error is wrongly concluding that the elevator is not safe and repairs are needed, when actually the elevator is safe.*

TIP: When trying to decide which type of error has been committed in a problem, use these facts:

- *A Type 1 error can ONLY happen when we REJECT the null*
- *A Type 2 error can ONLY happen when we FAIL TO REJECT the null*

LCQ – Type 1 and Type 2 Errors

Given the Null Hypothesis Is

Your Decision Based On a Random Sample		True	False
	Reject	Type I Error	Correct Decision
	Do Not Reject	Correct Decision	Type II Error

Two Types of Errors in Decision Making

Setup: A restaurant got their supply of food (meat, veggies, etc.) a few days late and are trying to decide if they are still able to serve it or not. There are two hypotheses:

H_0: The food is still fresh

H_A: The food has gone bad

a) Describe Type I error

b) Describe Type II error

LCQ – Type 1 and Type 2 Errors

		True	False
Your Decision Based On a Random Sample	Reject	Type I Error	Correct Decision
	Do Not Reject	Correct Decision	Type II Error

Two Types of Errors in Decision Making

Setup: A restaurant got their supply of food (meat, veggies, etc.) a few days late and are trying to decide if they are still able to serve it or not. There are two hypotheses:

H_0: The food is still fresh

H_A: The food has gone bad

a) Describe Type I error→ (*Reject the Null Hypothesis and) **Conclude the Alternative**, <u>when in reality the Null Hypothesis is actually correct</u>*

<u>Descriptions</u> - Options

1) The food still would be thought as fresh, when the food is not fresh → INCORRECT! This would mean concluding the Null is TRUE but in reality the Alternative is correct (this is actually Type 2!)

2) Type I error is wrongly concluding that the food is not fresh → ALMOST, this has the first part perfect! But MISSING the 'in reality the Null is TRUE'. So NEED to ADD 'but in reality it's good'

3) We wrongly conclude that the food has gone bad, when it is still good → Very good! Stating we wrongly conclude alternative in context but the Null is actually true context!

b) Describe Type II error → *(Fail to (2) Reject the Null Hypothesis and) **Conclude the Null**, <u>when in reality the the alternative is actually correct</u>*

<u>Descriptions</u> - Options

1) We wrongly conclude that the food is fresh → ALMOST again! MISSING the second key part of committing an error, the truth being different than our conclusion! NEED to HAVE 'when it actually it has gone bad'

2) We wrongly conclude that the food is still good when it has actually gone bad → PERFECT! Conclude H_A in context but actually H_0 context is true!

To describe these errors, it is really as simple as plugging and chugging into the general description! Here's what I mean:

- *Conclude the Null, when in reality the the alternative is actually correct*
- ***Wrongly** Concluding the food is still fresh, when in reality the food has gone bad → Can add 'wrongly' at the start further indicating that we are making the incorrect decision!*