Almost there!!

Unit 8 – Hypothesis Testing

Your Scientist Professor Colton

Unit 8 - Outline

Unit 8 – Hypothesis Testing

Intro

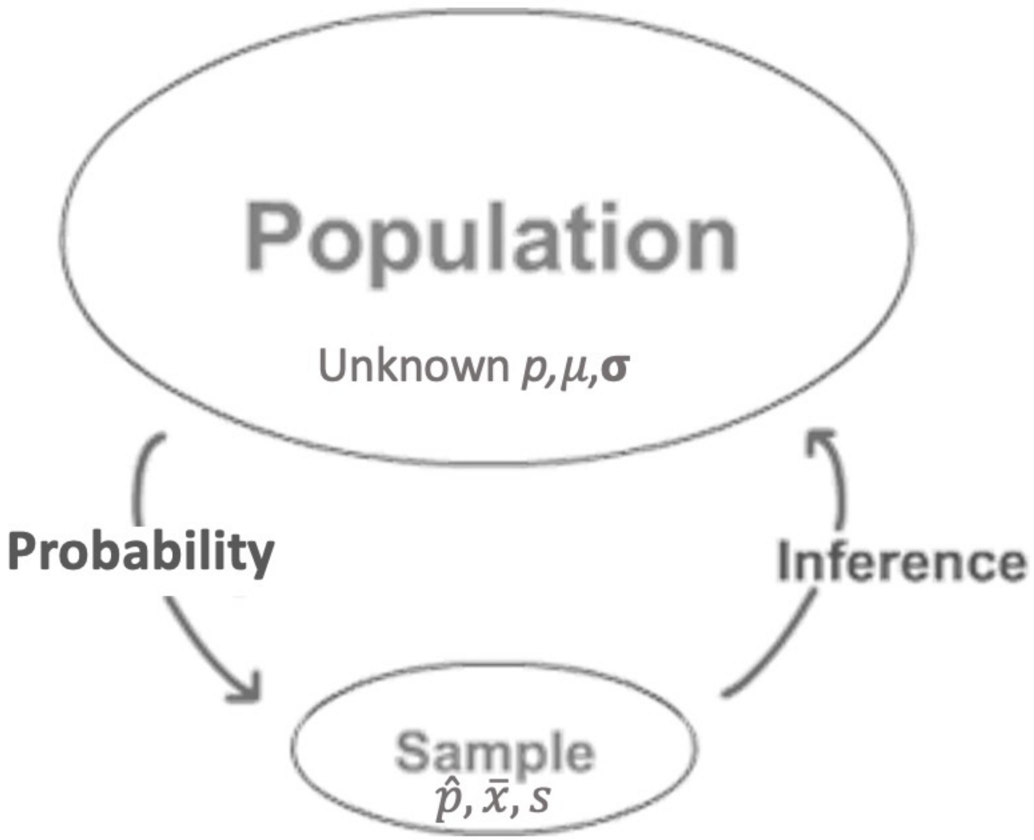
Hypothesis Testing Overview

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

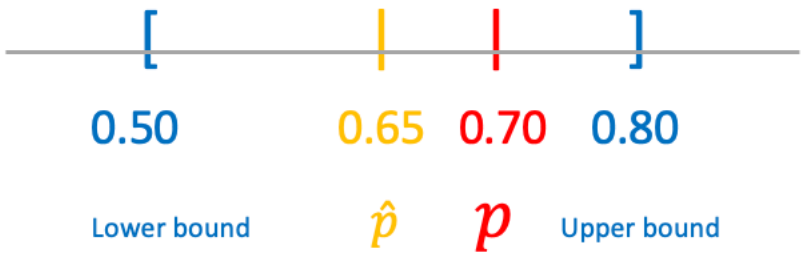
Hypothesis Testing for Population Proportions

- Calculator Work and Conclusions
- Full Example

Inference! Our Second Look



Confidence Intervals



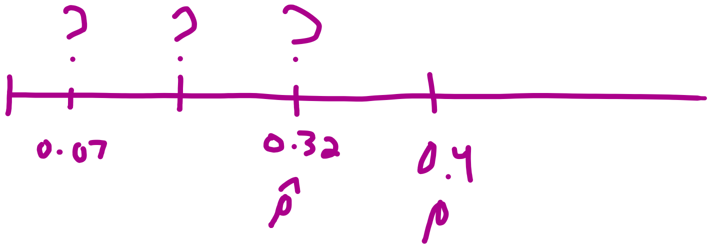
What's next??

Intro to Hypothesis Tests

Testing a Claim

A company is criticized because only 13 of 43 people in executive-level positions are women. The company explains that this is not surprising given that only 40% of their employees are women. What do you think?

- This is a big claim!
- Is there actually enough evidence to back this up?
  - Is the sample proportion 13/43 = 0.32 strange enough????
  - What if it was 30/43?? Or 3/43???



Compare this type of question to where we have been:

Estimating a Parameter

Estimate the proportion of female employees at this company with a confidence interval.

Logic Behind Hypothesis Testing

- We believe something about a population
  - **The Null Hypothesis (H_0)**
  - Ex) 40% of the company employees are women, so 40% of the executives should also be women

- We want to determine if something else is true
  - **The Alternative Hypothesis (H_A)**
  - Ex) Less than 40% of the executives are women

- We use sample data to make a decision / conclusion
  - **Compute the Test Statistic (TS)**
  - Ex) 13 out of the 43 executives are women, use this to find the TS
  - And then determine if we will continue to believe the null hypothesis or will reject it in favor of the alternative hypothesis.

- Enough evidence?
  - How far must the sample statistic be from the hypothesized parameter?
  - This is set before running the test and depends on the the significance level α

Full Problem

Here is an <u>entire hypothesis problem</u> worked out <u>perfectly</u> to show us where we are going!
- <u>Then we will break it down piece by piece!</u>

Setup: The campus bookstore is determining if they need to increase their marketing budget. They would like at least 65% of students to buy their textbooks directly from them rather than off-campus stores. In order to check this, they took a random sample of 137 students in which 81 students said they buy their books at the campus bookstore.

Is there sufficient evidence to conclude that the proportion of students who purchase their books is less than 0.65? Test this at the 10% significance level.

<u>Solution</u>

Hypotheses:

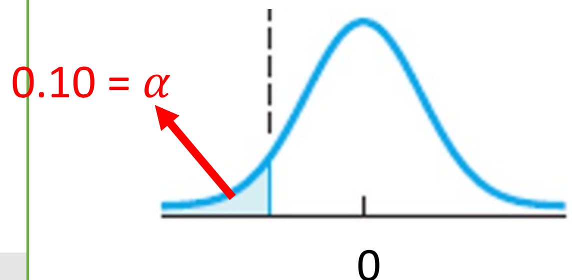
Let p = true proportion of students who purchase textbooks at the campus bookstore

$H_0: p = 0.65$
$H_A: p < 0.65$

Rejection Region:

α = 0.1

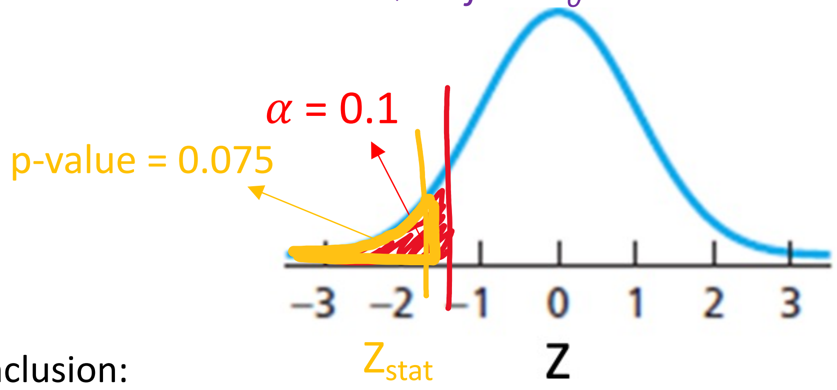
RR = Lower 10% of standard normal curve



P-Value:

p-value = 1-PropZTest(p_0 = 0.65, x = 81, n = 137, prop < p_0) = 0.075

p-value = 0.075 < α = 0.1 → Reject H_0!



Conclusion:

Because our p-value = 0.075 is less than the significance level 0.1, we reject the Null hypothesis.

We have sufficient evidence to conclude that the true proportion of students who buy their textbooks at the campus bookstore is less than 0.65.

- *The marketing team should increase their budget to reach their*

Hypothesis Test Steps – This is Your Life Now…

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

The Hypothesis Statements

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

Full Example

Is there sufficient evidence to conclude that the proportion of students who purchase their books is less than 0.65?

(Parameter):
Let p = true proportion of students who purchase textbooks at the campus bookstore

Define Parameter

- Always define your **parameter** at the start!
  - This helps orient us with the context of the problem and sets us up to correctly write our hypothesis statements
- Think about the variable / quantity of interest!
  - Categorical (“success” and “failure”) → population proportion (p)
  - Quantitative (numeric) → population mean μ

This is similar to how we talked about the ***parameter + context*** *in the confidence interval interpretation!*

More Examples:
- Let p = population proportion of students is statistics courses
- Let μ = true mean height of oak trees in meters

The Hypothesis Statements

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

<u>Null Hypothesis</u> H_0

- The Null hypothesis is an equation that includes the population parameter
  - **When written symbolically ALWAYS =**
- This is the <u>status quo</u>, typically a *known* value of the parameter (p_0 or μ_0)
- Said another way, it's what we are starting with as TRUE
  - In doing so, we are placing the hypothesized value at the <u>center of our distribution</u>!

In general

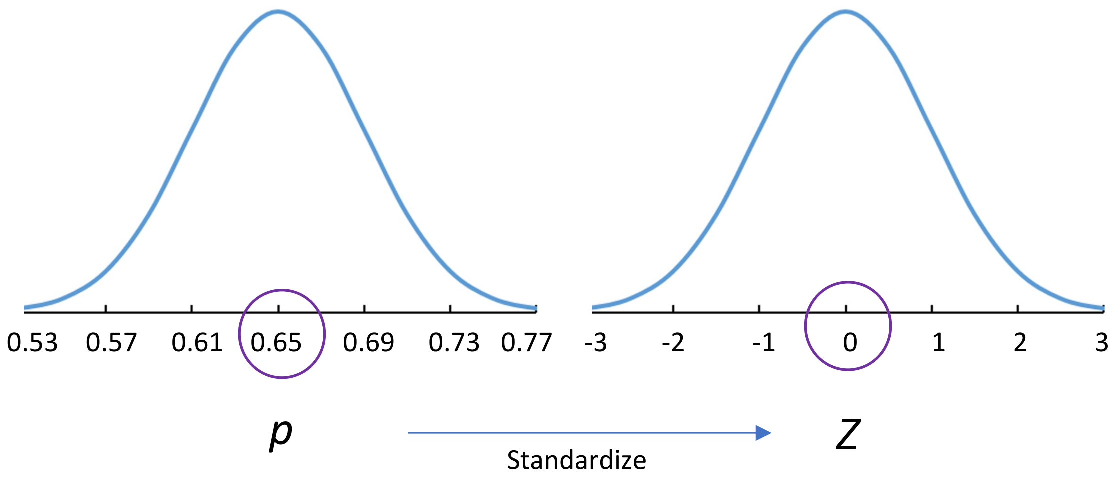
- H_0: $p = p_0$
- H_0: $\mu = \mu_0$

<u>Full Example</u>

Is there sufficient evidence to conclude that the proportion of students who purchase their books is less than 0.65?

Hypotheses:
Let p = true proportion of students who purchase textbooks at the campus bookstore

$H_0: p = 0.65$



More Examples:
H_0: $p = 0.3$, p = population proportion of students is statistics courses
H_0: $\mu = 70$, μ = true mean height of oak trees in meters

'Research from previous studies suggests the average number of people is 7'
- Equal to → H_0: $\mu = 7$

The Hypothesis Statements

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

<u>Alternative Hypothesis</u> H_A

- The Alternative hypothesis is another equation that includes the population parameter
  - Uses the <u>same value</u> of the parameter as in the Null hypothesis H_0
- May be left-tailed (<), right-tailed (>), or two-tailed (≠).
  - Depends on if we are interested in simply different than (≠), or a directional difference (< or >)
- This is the research interest, what <u>we want to prove</u>

In general

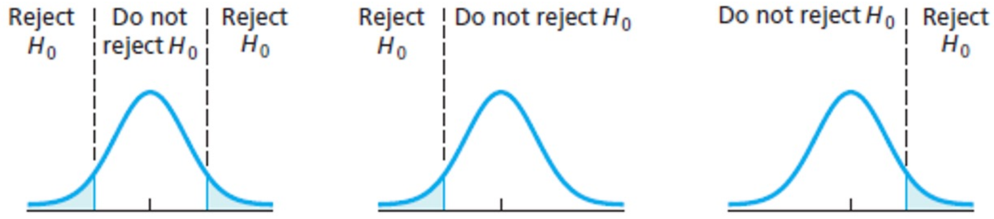
- $H_0: p = p_0$ and $H_A: p \neq, <, > p_0$
- $H_0: \mu = \mu_0$ and $H_A: \mu \neq, <, > \mu_0$

<u>Full Example</u>

Is there sufficient evidence to conclude that the proportion of students who purchase their books is less than 0.65?

Hypotheses:
Let p = true proportion of students who purchase textbooks at the campus bookstore

$H_0: p = 0.65$
$H_A: p < 0.65$



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

More Examples:
H_0: p = 0.3, p = population proportion of students is statistics courses
H_A: p > 0.3

H_0: μ = 70, μ = true mean height of oak trees in meters
H_A: $\mu \neq 70$

‘The owner believes his average monthly profit is more than $50,000’
- In this case, greater than → H_A: μ > 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 company reports that last year 40% of their reports in accounting were on time. From a random sample this year, they want to know if that proportion has **changed**.

b) Last year, 42% of the employees enrolled in at least one wellness class at the company's site. Using a survey from randomly selected employees, they want to know if a **greater** percentage is planning to take a wellness class this year.

c) There are two political candidates, and one wants to know from the recent polls if she is going to win a majority of votes in next week's election.

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 company reports that last year 40% of their reports in accounting were on time. From a random sample this year, they want to know if that proportion has changed.

Let p = the true proportion of reports that were on time
H_0: $p = 0.4$ and H_A: $p \neq 0.4$ → two tailed! PERFECTO!!

b) Last year, 42% of the employees enrolled in at least one wellness class at the company's site. Using a survey from randomly selected employees, they want to know if a greater percentage is planning to take a wellness class this year.

Options
1) Let p = the true ~~number~~ PROPORTION of employees enrolled (+ more context) → Have to say PROPORTION (not 'number', 'amount', etc.)
2) P = TRUE proportion of employees enrolled in at least one wellness class → Need to make sure to say TRUE or POPULATION proportion

Options
1) $H_0 > 0.42$ and $H_A < 0.42$ → several things to correct: (1) 'H_0' is NOT our PARAMETER, p is (have to include p); (2) equals sign = always goes in the Null hypothesis; (3) from the wording of the problem ('if a greater percentage is planning') indicates greater than for the Alternative and have to include our PARAMETER, less than 0.42, p
2) H_0: $p = 0.42$ and H_A: $p > 0.42$ → RIGHT (greater than) tailed! YES!!
H_A: $p \geq 0.42$ → INCORRECT, should never include any equality in the alternative hypothesis

c) There are two political candidates, and one wants to know from the recent polls if she is going to win a majority of votes in next week's election.

Let p = the true proportion of votes she will receive
H_0: $p = 0.5$ → *We use 0.5 because there was no prior information about how she had been polling, so we just start with assuming they are tied (evenly 50/50)*

Options
H_A = wins election → *INCORRECT! Yes this is the correct context, but what does this mean for the population proportion??? Have to have an INEQUALITY with our PARAMETER*
H_A: $p > \cancel{0.51}$ → *Alternative hypothesis is INCORRECT! We have to have the SAME VALUE in the Null and Alternative; the strictly greater sign take care of getting 'the majority' because anything more than 0.5 is technically the majority (even 0.50001)*
H_A: $p > 0.5$ RIGHT tailed → *CORRECT!*

Rejection Region

2. Sketch Rejection Region based of Significance Level

Rejection Region (RR)

- We have to determine the the when there is or is not enough evidence against the Null.
    - In other words, which tail do we make the conclusion of reject and how large is the area!
- 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)!

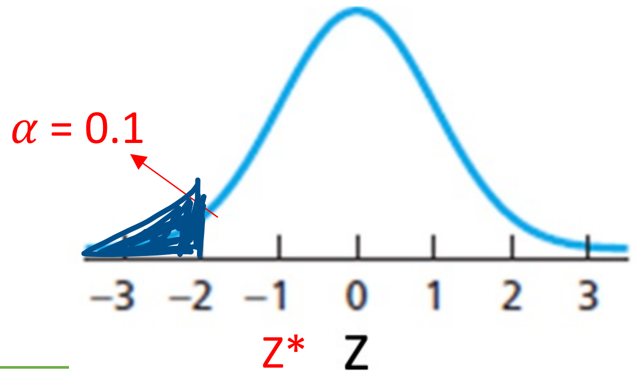
Significance Level

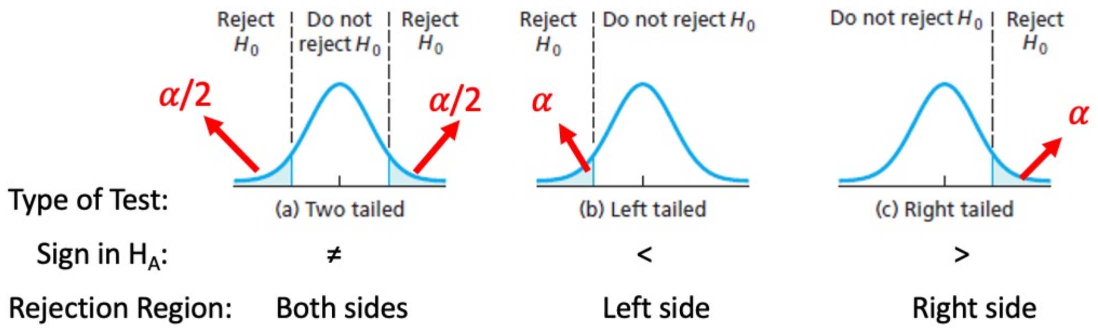
- **The significance level** (α) of the test is what determines how large the Rejection Region is!
    - It is chosen *before* running the test. Setups will say something similar to: ”Determine if there is enough evidence at the 5% significance level.”
- It represents the probability of rejecting the Null Hypothesis.
- Visually, it is the area under the curve in the direction of the Alternative hypothesis
- For a two-tailed test, area for α is split equally between the upper and lower tails!

Full Example

Is there sufficient evidence to conclude that the proportion of students who purchase their books is less than 0.65? Test this at the 10% significance level.

Rejection Region:
α = 0.1
RR = Lower 10% of standard normal curve

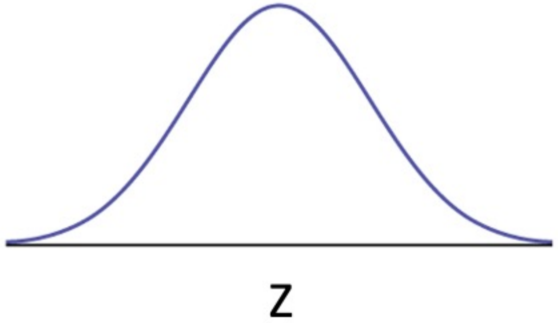




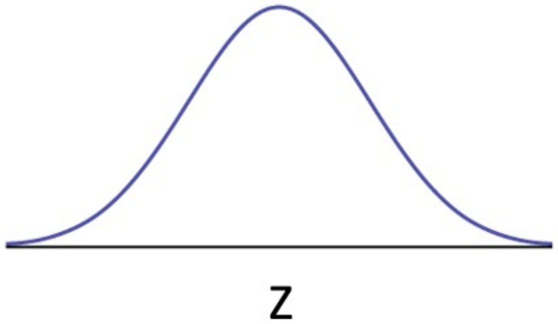
Mini LCQ

Problem: Sketch the Rejection Region for the following Alternative hypotheses and significance levels:

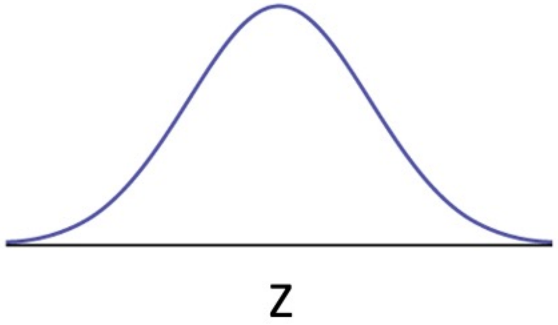
a) H_A: $p > 0.7$, $\alpha = 0.08$



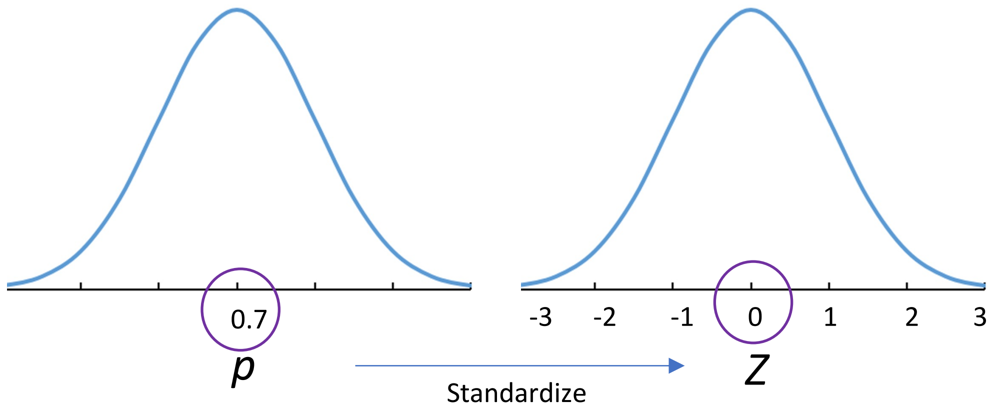
b) H_A: $p \neq 0.7$, $\alpha = 0.15$



c) H_A: $p < 0.7$, $\alpha = 4\%$



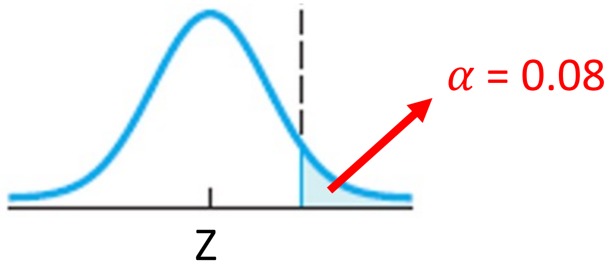
Mini LCQ Solution



Note: We can draw these RR on t
standardized. We will use the Z c
and applies to all problems

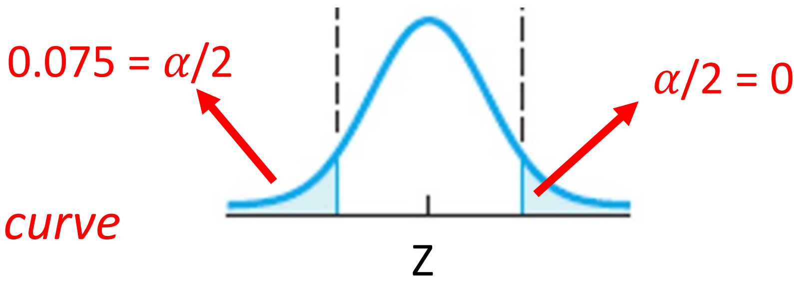
Problem: Sketch the Rejection Region for the following Alternative hypotheses and significance levels:

a) H_A: p > 0.7, α = 0.08



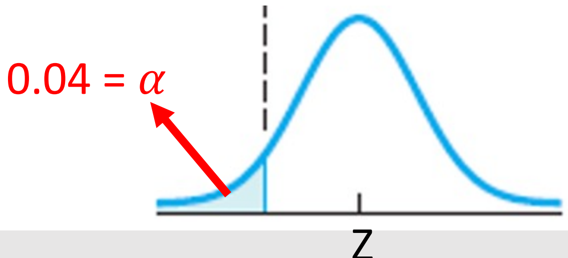
Right tailed → RR = upper 0.08 probability on the standard normal curve

b) H_A: p ≠ 0.7, α = 0.15



Two tailed → RR = lower and upper 0.075 probability on the standard normal curve
RR are always to the outside, so never the middle probability

c) H_A: p < 0.7, α = 4%



Left tailed → RR = lower 0.04 probability on the standard normal curve

Test Statistic

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

Test Statistic (TS)

****** NOTE!!* We are not going to use the Test Statistic to make conclusions, but it is helpful to know what it is!*

- **Test Statistic (TS)** is calculated using our sample statistics (or directly from the data)
- Technically, the TS is just the standardized (Z) score of our sample statistic based on the corresponding sampling distribution of that statistic under the Null hypothesis
- We could plot the TS to see if it is in the rejection region!

*** These are based on the* ***CLT formulas*** *proportions and means*

Test for	H_0	Test statistic
Pop. mean μ	$\mu = \mu_0$	$z = \dfrac{\bar{x} - \mu_0}{\frac{\sigma}{\sqrt{n}}}$
Pop. prop. p	$p = p_0$	$z = \dfrac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}}$

Full Example

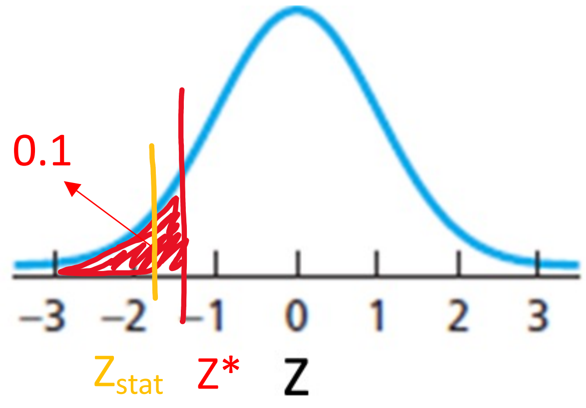
Is there sufficient evidence to conclude that the proportion of students who purchase their books is less than 0.65? Test this at the 10% significance level.

Test Statistic:

$TS = Z_{stat}$ = 1-PropZTest(p_0 = 0.65, x = 81, n = 137, prop < p_0) = -1.44

$$\hat{p} = \frac{81}{137} \approx 0.59 \rightarrow -1.44 = Z_{stat}$$

Z_{stat} is in the Rejection Region → Reject H_0!



- We are *not* going to calculate these by hand because we've had lots of practice before
- Instead, we will learn new calculator functions shortly!
  - 1-PropZTest and Z-Test

P-Value

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

P-Value

- The **p-value** is the probability of getting a result as, or more extreme than, the result obtained from the sample given (assuming) the Null Hypothesis (H_0) is TRUE.
  - In other terms, "The P-Value is a measure of how plausible the data are, given our null hypothesis."

P-Value Method

- This way just involves comparing the p-value to the significance level.

Decision RULE:

If: p-value $\leq \alpha$ Reject H_0! *This means the TS in in the RR!*

If: p-value $> \alpha$ Fail to Reject H_0!

- If the p-value is less than the significance level, this is the implication:
  - Our statistic or anything more extreme (smaller or larger; depends on alternative) isn't likely with the Null Hypothesis being True!! And thus we REJECT!

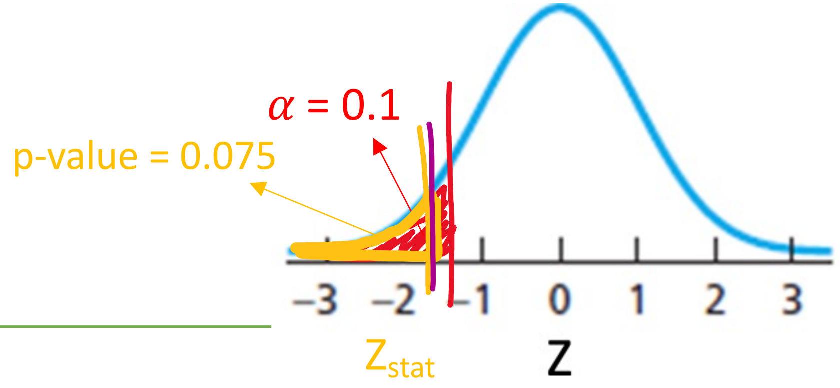
Full Example

Is there sufficient evidence to conclude that the proportion of students who purchase their books is less than 0.65? Test this at the 10% significance level.

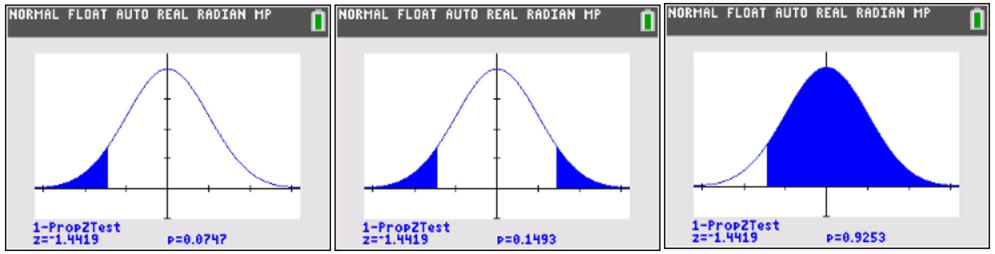
P-Value:

p-value = 1-PropZTest(p_0 = 0.65, x = 81, n = 137, prop < p_0) = 0.075

p-value = 0.075 < α = 0.1 → Reject H_0!

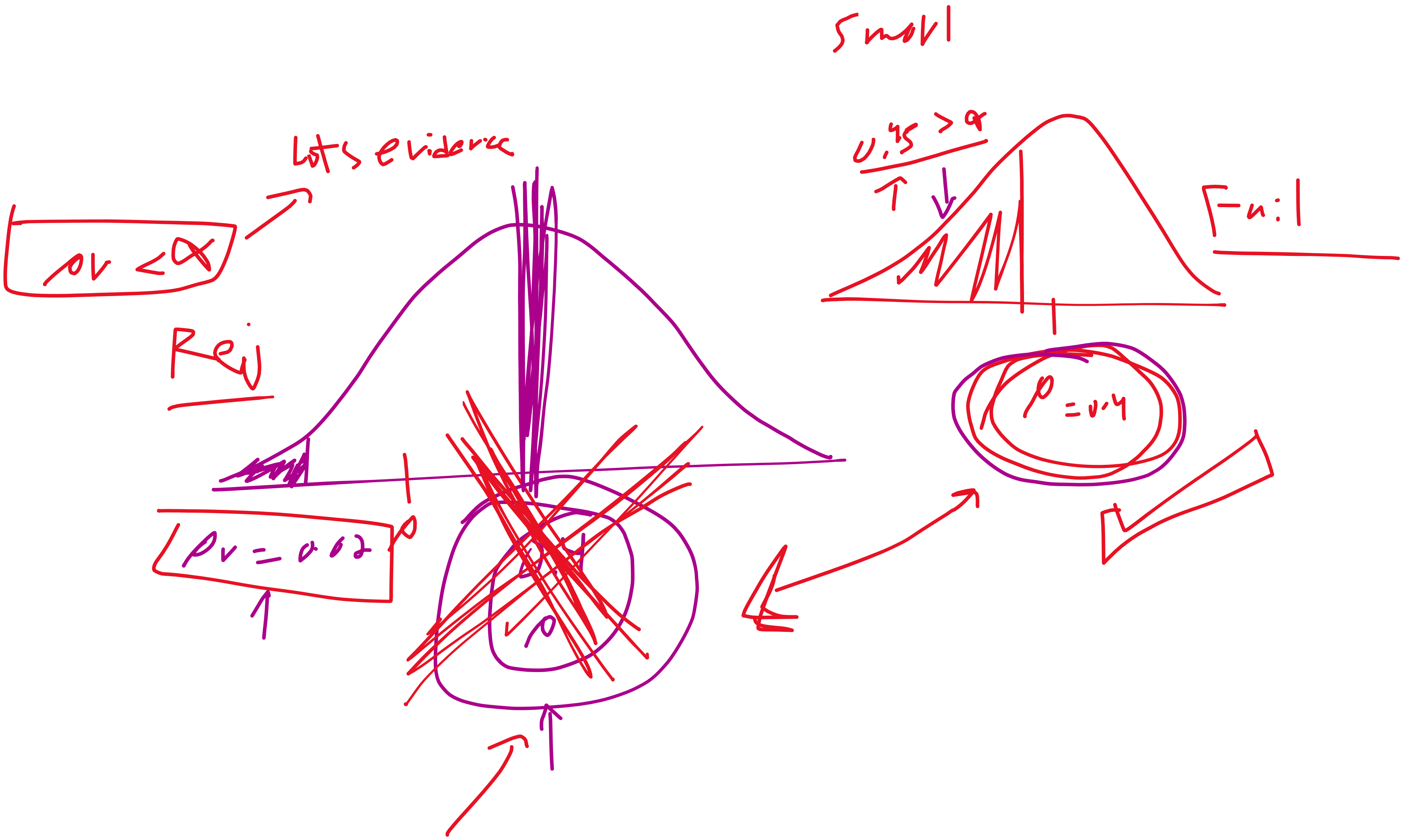


- Our calculator draws the p-value.
- P-Value depends on the alternative hypothesis!



H_A: $p < 0.65$ $p \neq 0.65$ $p > 0.65$

* With two-tailed tests, the p-value is double that of the one-tailed


Lots evidence
pv < α
Rej
pv = 0.02
p
Small
0.45 > α
Fail
p = 0.4

Conclude and Interpret

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

This is where we formally (nicely) say the results of our hypothesis test!

How to Write Our Conclusion

(General structure)

Reject or Rail to reject NULL

First part → Because our P-Value (**INSERT P-Value**) is (**LESS** or **GREATER**) than our significance level (**INSERT SIGNIFICANCE LEVEL**), we (**REJECT** or **FAIL TO REJECT**) the Null Hypothesis.

Second part → There (**IS** or **IS NOT**) sufficient evidence to conclude (**THE ALTERNATIVE HYPOTHESIS**).

DECISION RULE!!
- If P-value $\leq \alpha$, **REJECT** (use the **Red Text**)
- If P-value $> \alpha$, **FAIL TO REJECT** (use the **Blue Text**)

*** Always talk about the results in terms of the the problem! Don't just write '… conclude the alternative hypothesis'

Full Example

Is there sufficient evidence to conclude that the proportion of students who purchase their books is less than 0.65? Test this at the 10% significance level.

Conclusion:
Because our p-value = 0.075 is less than the significance level 0.1, we reject the Null hypothesis.

We have sufficient evidence to conclude that the true proportion of students who buy their textbooks at the campus bookstore is less than 0.65.

*** **NEVER SAY WE ACCEPT THE NULL HYPOTHESIS!!!!!!**
We aren't saying that the Null is true, just that we can't say it's wrong! Think NOT GUILTY (not "Innocent")

Hypothesis Tests for Proportions!

- Everything above applies, now we are just going to apply it specifically to a Proportion Test!

Using Calc - Test Statistic and P-Value for Proportions

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

Formula for Z_{stat} by hand:

Test for	H_0	Test statistic
Pop. prop. p	$p = p_0$	$z = \dfrac{\hat{p} - p_0}{\sqrt{\dfrac{p_0(1-p_0)}{n}}}$

(Original) Setup
In 2020, A NatGeo Poll interviewed 1200 hiking enthusiasts and asked "Are you more afraid of spiders or snakes???" Out of the 1200 people, 768 responded "Ewww, snakes….". Is there enough evidence to conclude the proportion of people who are more afraid of snakes is **different** than the 2010 proportion of 0.65. Use α = 0.1

GOAL: Conduct a Hypothesis Test!

1. 1-PropZTest
   a) p_0 = the Null proportion
   b) x = number of successes
   c) n = sample size
   d) prop: Alternative hypothesis

   Calculate or Draw

New Scenario
Now lets say the 2020 sample proportion is equal to 0.62 (same sample size) AND we want to know if the proportion has decreased from 2010. Use α = 0.05

- Run another 1-PropZTest

Using Calc - Test Statistic and P-Value for Proportions

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

Formula for Z_{stat} by hand:

Test for	H_0	Test statistic
Pop. prop. p	$p = p_0$	$z = \dfrac{\hat{p} - p_0}{\sqrt{\dfrac{p_0(1-p_0)}{n}}}$

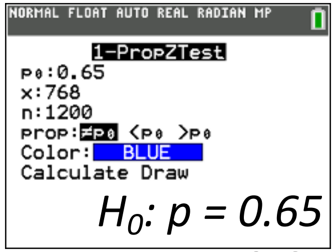
(Original) Setup

In 2020, A NatGeo Poll interviewed 1200 hiking enthusiasts and asked "Are you more afraid of spiders or snakes???" Out of the 1200 people, 768 responded "Ewww, snakes….". Is there enough evidence to conclude the proportion of people who are more afraid of snakes is different than the 2010 proportion of 0.65. Use $\alpha = 0.1$

GOAL: Conduct a Hypothesis Test!

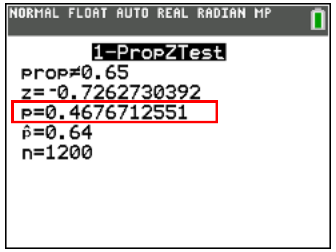
1. 1-PropZTest
   a) p_0 = the Null proportion
   b) x = number of successes
   c) n = sample size
   d) prop: Alternative hypothesis

   Calculate or Draw

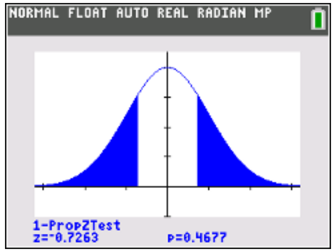


H_0: $p = 0.65$

H_A: $p \neq 0.65$



Calculate Output
prop = Alternative hypothesis
z = Z_{stat}
p = p-value
$\hat{p}$= sample proportion
n = sample size



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

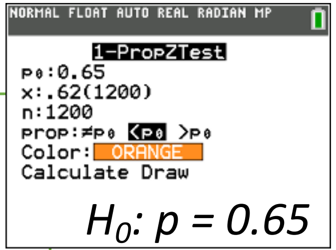
Show work: 1-PropZTest(p_0 = 0.65, x = 768, n = 1200, prop ≠ p)

Then we are looking for the p-value!

New Scenario

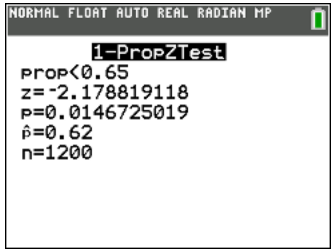
Now lets say the 2020 sample proportion is equal to 0.62 (same sample size) AND we want to know if the proportion has decreased from 2010. Use $\alpha = 0.05$

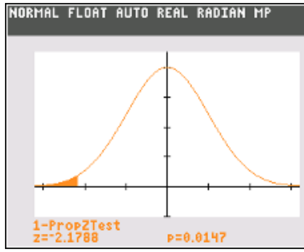
- Run another 1-PropZTest



H_0: $p = 0.65$

H_A: $p < 0.65$





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: In 2020, A NatGeo Poll interviewed 1200 hiking enthusiasts and asked "Are you more afraid of spiders or snakes???" Out of the 1200 people, 768 responded "Ewww, snakes….". Is there enough evidence to conclude the proportion of people who are more afraid of snakes is different than the 2010 proportion of 0.65. Use α = 0.1

b) New Scenario: Now lets say the 2020 sample proportion is equal to 0.62 (same sample size) AND we want to know if the proportion has decreased from 2010. Use α = 0.05

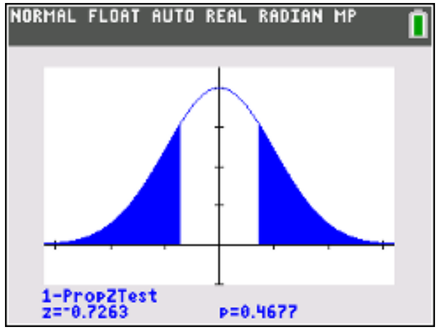
LCQ – Conclusions and Interpretations

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: In 2020, A NatGeo Poll interviewed 1200 hiking enthusiasts and asked “Are you more afraid of spiders or snakes???” Out of the 1200 people, 768 responded “Ewww, snakes….”. Is there enough evidence to conclude the proportion of people who are more afraid of snakes is different than the 2010 proportion of 0.65. Use α = 0.1



Need these… H_0: $p = 0.65$ *Let p = true proportion of hikers who are more afraid of snakes*

H_A: $p \neq 0.65$

$\alpha = 0.1$

P-Value

p-value = 1-PropZTest(p_0 = 0.65, x = 768, n = 1200, prop ≠ p_0) = 0.4667

p-value = 0.4667 < 0.10 = α *→ Fail to Reject H_0! → The first part of the conclusion is really just writing this comparison in words*

Conclusion and Interpretation

First part → *Decision (comparison of p-value and significance level):*

*Because our p-value = 0.4667 is **greater** than the significance level 0.10, we fail to reject the Null hypothesis → Perfect for the first part*

Second part → *Second part (then what does the decision mean):*

*First attempt at second part → There is **not** enough evidence to conclude the true proportion is different than the null hypothesis (0.65) → The ending part can be improved! Need context!!*

Improvement

*There is **not** enough evidence to conclude to that the true proportion of hikers who are more afraid of snakes is different than 0.65*

*** the underlined portion above is where we talk about the alternative hypothesis in CONTEXT! We can really just directly put into words what each part of the H_A represents!*

Here’s what I mean: There is NOT enough evidence that p (the true proportion of hikers that are more afraid of snakes) is ≠ (different than) 0.65! So easy!

Entire Process LCQ

Problem: 24% of the state of Ohio speaks a second language. The local community center took a random of 96 community members in which 26 speak a second language.

Is there sufficient evidence that the proportion of local community members that speak a second language is higher than 0.24? Use α = 0.15.

Solution:

Hypotheses and Parameter:

Rejection Region:

Test Statistic and P-value:

Conclusion and Interpretation:

Entire Process LCQ

Problem: 24% of the state of Ohio speaks a second language. The local community center took a random of 96 community members in which 26 speak a second language.

Is there sufficient evidence that the proportion of local community members that speak a second language is higher than 0.24? Use α = 0.15.

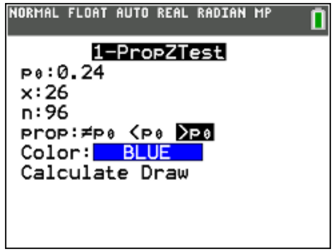
Solution:

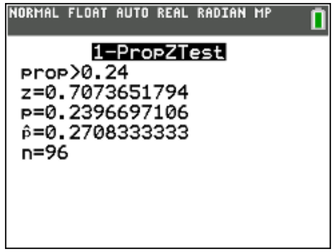
Hypotheses and Parameter:

Let p = true proportion of community members who speak a second language
$H_0: p = 0.24$
$H_A: p < 0.24$

Rejection Region:

Test Statistic and P-value:
1-PropZTest(p_0 = 0.24, x = 26, n = 96, prop > p_0)
p-value = 0.293 < 0.15 = α → Fail to Reject H_0!





Conclusion and Interpretation:
Because the p-value = 0.293 is greater than the significance level 0.15, we fail to reject the null hypothesis.
There is not sufficient evidence to conclude that the true proportion of community members to speak a second language is greater than 0.24