Hypothesis Testing

Sara Esfahani

Economics 30330

Spring 2017

Introduction

- Hypothesis testing is an extension of confidence interval estimation
- In this chapter we will learn how hypothesis testing will help us to determine if we believe the true population parameter exists within an interval that we created
- We start by making some assumptions about the population parameter (mean or proportion)
  - The Null Hypothesis H_0
  - The Alternative Hypothesis H_a
    - If multiple alternative hypotheses exist, denote them by H_1, H_2, …

What is Hypothesis Testing?

1. State the **null & alternative** hypotheses

2. Take a **sample**

3. Use the sample results to **reject** or **do not reject** the Null Hypothesis

4. Draw **conclusions**

What is a Null Hypothesis?

H_0 is an assumption about μ or P

H_a is counter to H_0

- Say we would like to determine if the light bulbs from a particular manufactory burns on an average of 1050 hours

$$H_0\colon \mu = 1050$$

$$H_a\colon \mu \neq 1050$$

What is a Null Hypothesis?

- Let's say we take a sample of n=50 light bulbs and estimate $\bar{x} = 1055$

- Does this mean that the true mean is NOT equal to 1050 hours?

- Or is it close enough to the true mean for us to believe the mean might be 1050?

- What if the $\bar{x}$ is 1070?

What is a Null Hypothesis?

- We know the closer the sample mean is to the population mean of 1050, we are less likely to reject the null hypothesis (more likely to fail to reject the null hypothesis)

- The opposite is true as well:

  – The further the sample mean is from 1050, the more likely we are to reject the null hypothesis

- How close is close enough?

Developing a Null Hypothesis

- Now if the question of research is whether the number of burn hours is at least 1050.

$$H_0:\ \mu \geq 1050$$

$$H_a:\ \mu < 1050$$

Forms of Hypothesis Tests

$$\begin{cases} H_0\colon \mu \geq \mu_0 \\ H_a\colon \mu < \mu_0 \end{cases}$$

One-tailed

(Lower-tail)

$$\begin{cases} H_0\colon \mu \leq \mu_0 \\ H_a\colon \mu > \mu_0 \end{cases}$$

One-tailed

(Upper-tail)

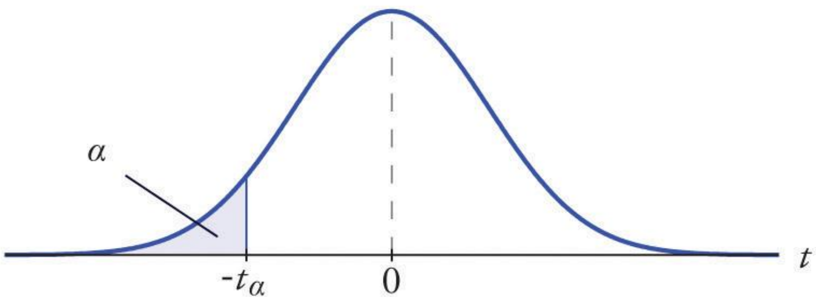
$$\begin{cases} H_0\colon \mu = \mu_0 \\ H_a\colon \mu \neq \mu_0 \end{cases}$$

Two-tailed

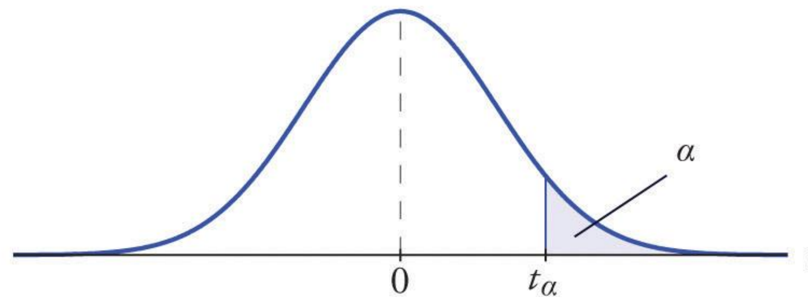
- Note the equality part of the expression always appears in the null hypothesis
- Keep in mind that the alternative hypothesis is often what the test is attempting to establish
- Alternative hypothesis is also called research hypothesis

Forms of Hypothesis Tests

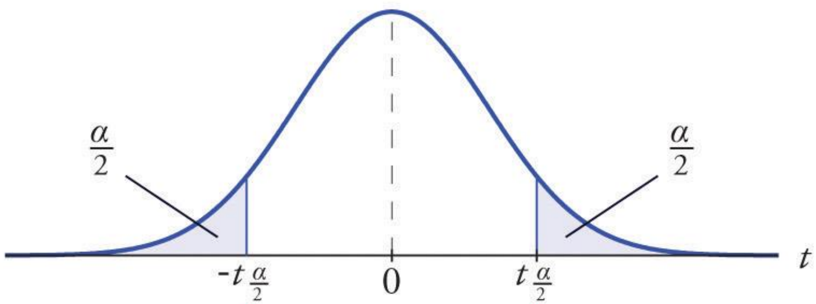
$H_a : \mu < \mu_0$



$H_a : \mu > \mu_0$



$H_a : \mu \neq \mu_0$



Hypothesis Testing Process

- After we form the null and alternative hypotheses
  - Take Sample
    - Look for Evidence
      - One of two conclusions:
        1. Reject the null hypothesis
        2. Fail to reject the null hypothesis

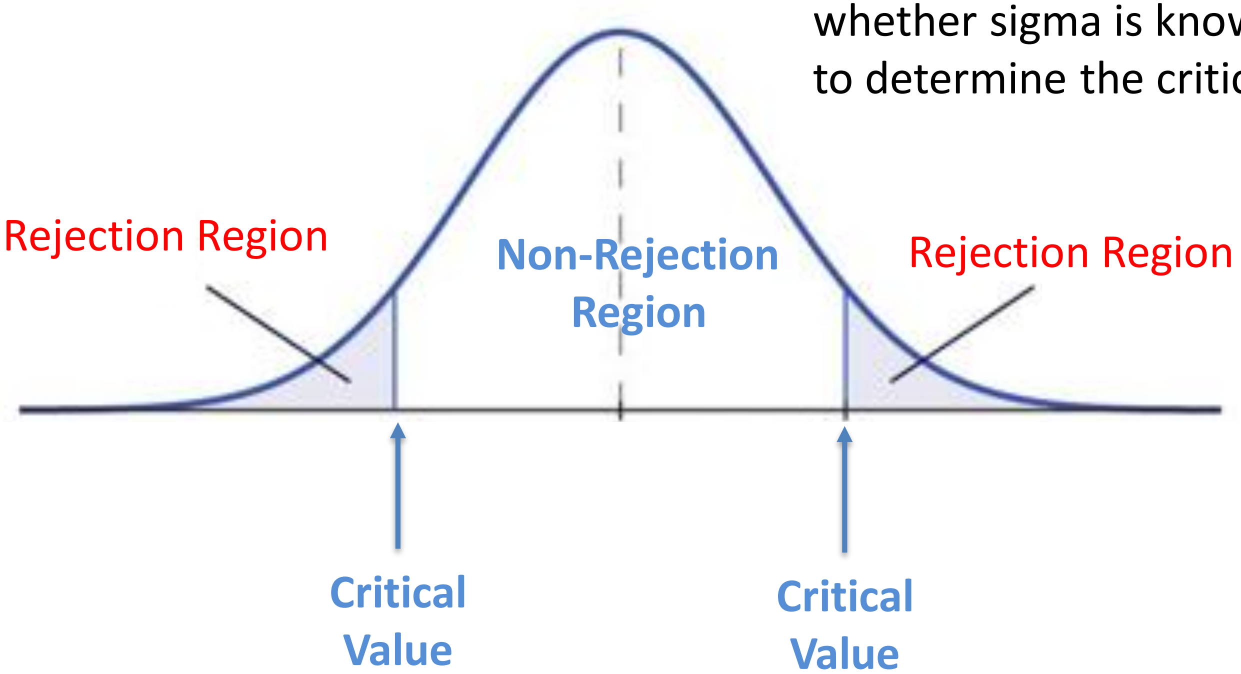
Critical Value

If we draw a sample and get a too high or too low sample statistics, we reject the null

Q: What determines the critical values?

A: α or level of significance

We use z-table or t-table (depending in whether sigma is known or unknown) to determine the critical values



Type I and Type II Errors

- We don’t always draw the correct conclusion
- Sometimes we get a bad sample, which leads us to draw a wrong conclusion

Type I and Type II Errors

		Population Condition	
		H_0 True	H_a True
Conclusion	Do not reject H_0	Correct Conclusion	**Type II Error**
	Reject H_0	**Type I Error**	Correct Conclusion

- Type I Error happens α percent of the time (depends on the level of significance)
- Type II Error = β

Let's start with a two-tailed test

- From the light bulbs example

$$H_0:\ \mu = 1050$$

$$H_a:\ \mu \neq 1050$$

- We have a sample of size n = 50 and we know σ = 65, and we get $\bar{X}$= 1045

- Calculate test statistics: $Z = \frac{\bar{X}-\mu_0}{\frac{\sigma}{\sqrt{n}}} = \frac{1045-1050}{\frac{65}{\sqrt{50}}}$=-0.5441

Critical Value Approach

- To look up the critical value, we need to know α

- Say α = 0.05, so $\frac{\alpha}{2}$ = 0.025

- From z-table we get the critical values of -1.96 and 1.96

- Our test statistics is -0.5441

- In non-rejection region: Do not reject the null hypothesis

- There is no evidence to conclude that the true mean is different from 1050 hours

p-Value Approach

- We have the test statistics from before = -0.5441

- To find p-value we calculate $P(z \leq -0.5441) = ?$

  From z-table: $P(z \leq -0.5441) = 0.2946$

- Since we have two tails, we double the p-value

  $0.2946 \times 2 = 0.5892$

p-Value Approach

- Rule for p-value approach:
  - Reject the null if p-value $\leq \alpha$
  - Do not reject the null if p-value $> \alpha$

- We have

$$\text{p-value} = 0.5892 \text{ and } \alpha = 0.05$$

  - That is $0.5892 > 0.05$
  - Therefore, DO NOT REJECT the null hypothesis
  - Notice we got the same conclusion as when we used critical value approach

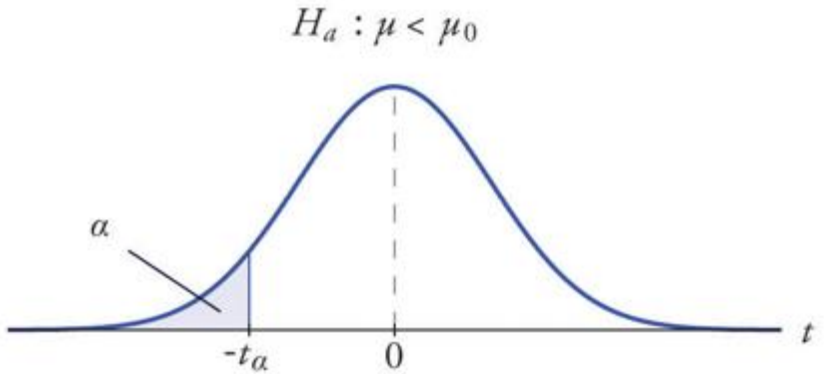
Now let's look at a one-tailed test

- Example: The department claims the average grade on an exam is at least 65. Is there evidence that the true mean is lower than claimed?
- We have a sample of size n = 75. We know σ = 15 & we get $\bar{X}$= 60
- Let's write the null and alternative hypotheses

$$H_0\colon \mu \geq 65$$

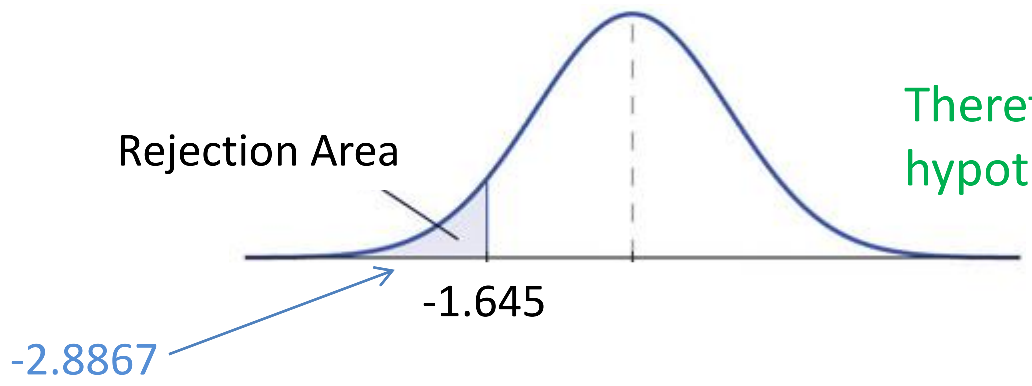
$$H_a\colon \mu < 65$$

- Remember the alternative hypothesis is also called research hypothesis



Hypothesis Testing

- Calculate test statistics: $Z = \frac{\bar{X}-\mu_0}{\frac{\sigma}{\sqrt{n}}} = \frac{60-65}{\frac{15}{\sqrt{75}}} = -2.8867$
- Critical value for α = 0.05
  - Row number -1.6
  - Column number: between 0.04 and 0.05
  - So, -1.645
- Compare and come to statistical conclusion



Therefore, REJECT the null hypothesis

p-Value Approach

- Rule for p-value approach:
  - Reject the null if p-value $\leq \alpha$
  - Do not reject the null if p-value $> \alpha$

- We have

$$\text{p-value} = P(z \leq -2.8867) = 0.0019$$

  - That is $0.0019 < 0.05$
  - Therefore, REJECT the null hypothesis
  - Notice we got the same conclusion as when we used critical value approach

For Sigma Unknown

- What if sigma is unknown?
  - Use s as a point estimate for σ
  - Use t table instead of z table
  - Everything else is the same as before

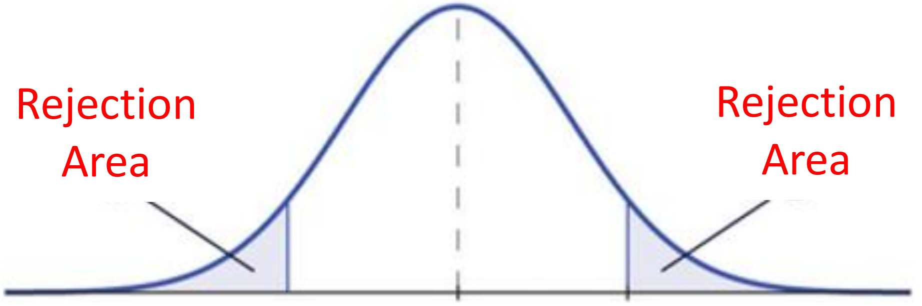
  - Test statistics here is called t,

$$t = \frac{\bar{X} - \mu_0}{\frac{s}{\sqrt{n}}}$$

Example of Two-tailed Test

- We want to know if the mean fill of cereal boxes is 16 ounces?

$$H_0: \mu = 16$$
$$H_a: \mu \neq 16$$



- ✓ Take Sample
- ✓ Calculate Test Statistics
- ✓ Look up a Critical Value
- ✓ Come to Statistical Conclusion

Example of Two-tailed Test

- ✓ Take Sample

n = 30 $\bar{x}$ = 15.3 s = 0.5

- ✓ Calculate Test Statistics

$$t = \frac{\bar{X} - \mu_0}{\frac{s}{\sqrt{n}}} = \frac{15.3 - 16}{\frac{0.5}{\sqrt{30}}} = -7.675$$

- ✓ Look up a Critical Value

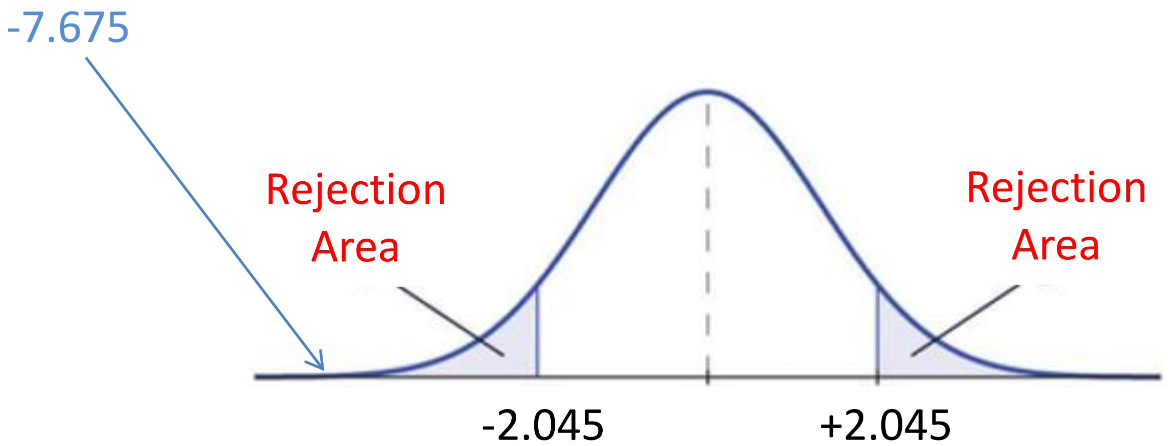
For $\alpha = 0.05$, we get $\frac{\alpha}{2} = 0.025$ and df=n-1=29

Critical value = -2.045 and +2.045

- ✓ Come to Statistical Conclusion

Conclusion

- Compare the test statistics with the critical values



- Therefore, REJECT the null hypothesis
- Conclude that the mean fill is not 16 ounces

p-Value Approach

- Look up the test statistics -7.675 in t-table
- With df = 29, p-value has to be a very small number (close to zero)
- Obviously p-value ≤ α
- Therefore REJECT the null
- Conclude there is evidence that the mean is not equal to 16 ounces

Example of One-tailed Test

- Testing whether the mean oil change time is less than 10 minutes
- We have

$$H_0\text{: } \mu \leq 10$$

$$H_a\text{: } \mu > 10$$

- We take a sample

n = 25 $\bar{x}$ = 11.5 s = 4

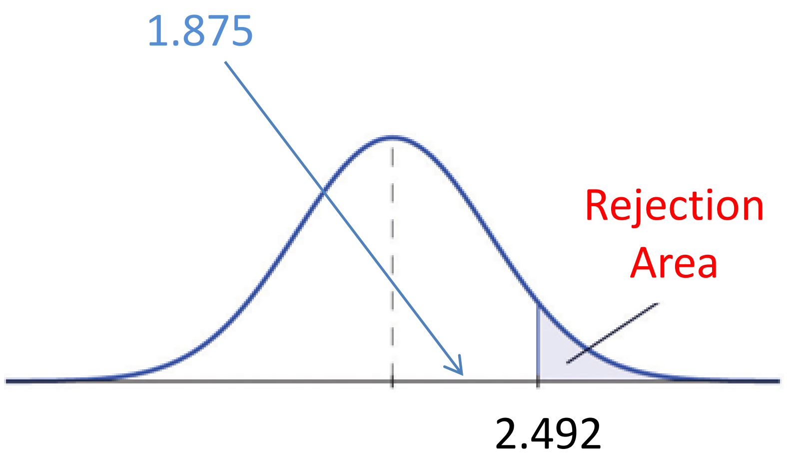
- Test statistics

$$t = \frac{\bar{X} - \mu_0}{\frac{s}{\sqrt{n}}} = \frac{11.5 - 10}{\frac{4}{\sqrt{25}}} = 1.875$$

Example of One-tailed Test

For $\alpha = 0.01$, and df=n-1=24

Critical value = 2.492



- Do Not Reject the null
- We don’t find evidence that the true mean is greater than 10 minutes

p-Value Approach

- Look up the test statistics 1.875 in t-table

- With df = 24, p-value
  - Row number 24
  - Column number: between 0.05 and 0.025
  - Both not less than $\alpha = 0.01$

- Therefore Do Not Reject the null

- Conclude there is no evidence that the mean is greater than 10 minutes

Hypothesis Testing
Population Proportion

- Just as we had for means, there are three forms of hypothesis testing for population proportion

$H_0: p \geq p_0$	$H_0: p \leq p_0$	$H_0: p = p_0$
$H_a: p < p_0$	$H_a: p > p_0$	$H_a: p \neq p_0$
One-tailed	One-tailed	Two-tailed
(Lower-tail)	(Upper-tail)	

Pop Quiz*! True or False…

1. When testing a hypothesis with proportions, you should use a *t value* to calculate the margin of error.

2. When testing a hypothesis with proportions, you should use the *sample proportion* to calculate the standard error.

**Not for credit.*

Let’s start with a two-tailed example

- A retailer believes 80% of his customers use credit card when marking a purchase. A sample of 200 customers showed 155 used a credit card to make their purchase. At the 0.05 level of significance, is there evidence that the true population proportion differs from 80%?

- First: What is the hypothesis?

$$H_0: p = 0.8$$

$$H_a: p \neq 0.8$$

Test Statistics

- Then calculate test statistics:

$$Z = \frac{\bar{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}}$$

- From the question:

$$\bar{p} = \frac{155}{200} = 0.775$$

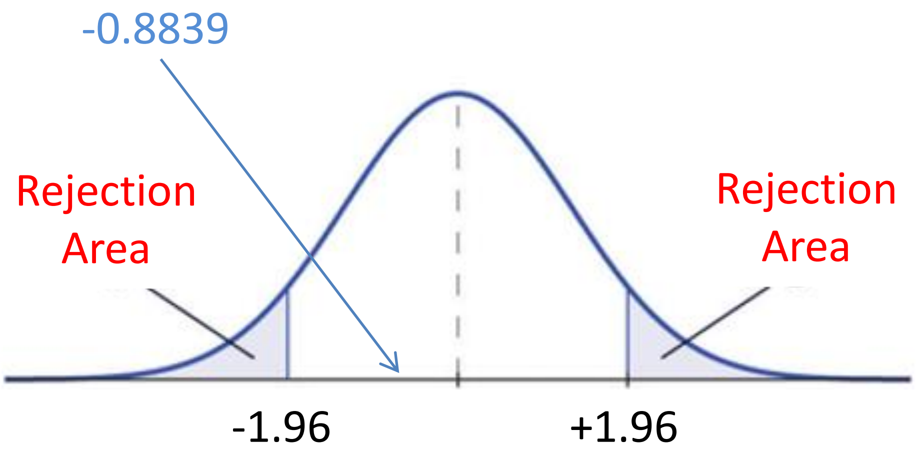
- Then we can calculate the test statistics:

$$Z = \frac{\bar{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}} = \frac{0.775 - 0.8}{\sqrt{\frac{0.8(1-0.8)}{200}}} = -0.8839$$

Critical Value

- For $\alpha = 0.05$, we look up $\frac{\alpha}{2} = 0.025$ in z-table:

Critical values are -1.96 and +1.96



- Do NOT REJECT the null hypothesis

Conclusion

- There is no evidence that the proportion of customers that use credit cards differ from 80%

p-Value Approach

- Look up the test statics in the z table
- $z = 0.8839$ or 0.88
- From the table p-value = 0.1894
- Because it is two-tailed test, double p-value first
- So $0.1894 \times 2 = 0.3788$
- We reject the null if p-value $\leq \alpha$
- We have $0.3788 < 0.05$
- Therefore, DO NOT REJECT the null
- There is no evidence that the proportion is different from 80%

Example of a one-tailed test

- Research Question: Are more than 80% of American's right handed?

- In a sample of 100 Americans, 87 were right handed.

Check Necessary Assumptions

- Can we use normal approximation method?

- Yes because $np = 100 * 0.80 = 80$ and $n(1-p) = 100 * 0.20 = 20$ are both greater than 5

- What are the null and alternative hypotheses?

$$H_0: p \leq 0.8$$

$$H_a: p > 0.8$$

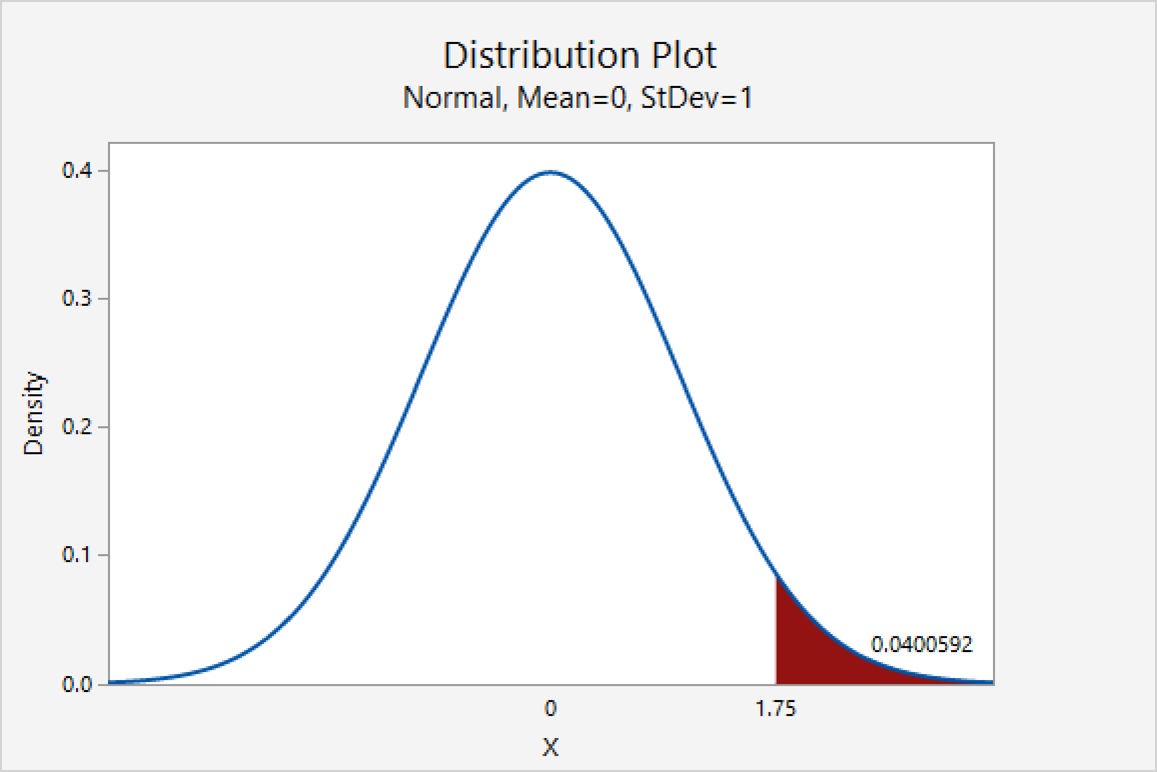
  - This is a right-tailed test because we want to know if the proportion is greater than .80

Calculate the appropriate test statistics

$$Z = \frac{\bar{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}} = \frac{0.87 - 0.80}{\sqrt{\frac{0.8(1-0.8)}{100}}} = 1.75$$

- According to the table $P(z < 1.75) = 0.9599$
- Therefore $P(z \geq 1.75) = 1 - 0.9599 = 0.0401$

Conclusion?



- Reject the null hypothesis.
- There is statistical evidence to state that more than 80% of American's are right handed.

More Examples to Review Chapter 9

Example 1

Suppose we know with certainty that 35% of graduate students smoke. In a survey of 31 ND grad students, 29.03% say they smoke (9 students)

- Are ND grad students different from the average graduate students?

Example 2

A subsample of hourly wages for 115 teen (16 – 19) workers are drawn from a large survey.

$$\bar{x} = \$8.45 \qquad s = 3.20$$

We know that the federal minimum wage is $7.25.

- Are teens working at minimum wage?

Example 3: Pulse Rate

- A research study measured the pulse rates of 57 college men and found a mean pulse rate of 70.4211 beats per minute with a standard deviation of 9.9480 beats per minute.
- Researchers want to know if the mean pulse rate for all college men is different from the current standard of 72 beats per minute.

Solution to Example 3: Pulse Rate

$$H_0: \mu = 72$$

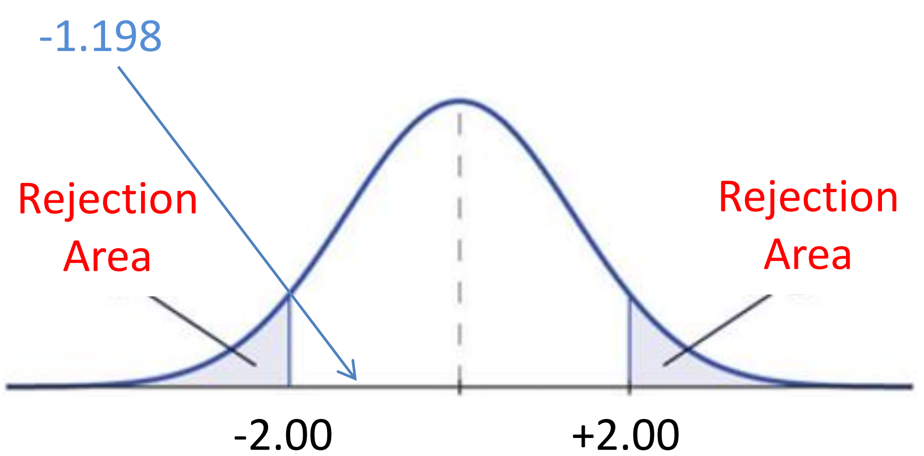
$$H_a: \mu \neq 72$$

$$t = \frac{\bar{X} - \mu_0}{\frac{s}{\sqrt{n}}} = \frac{70.4211 - 72}{\frac{9.9480}{\sqrt{57}}} = -1.198$$

$$df = n - 1 = 57 - 1 = 56$$

Solution to Example 3: Pulse Rate

- Compare the test statistics with the critical values



- Therefore, FAIL TO REJECT the null hypothesis
- Conclude there is not sufficient evidence to state that the mean pulse of college men is different from 72. It is reasonably possible that this sample came from a population with mean of 72.