USC School of Business

Data Driven Decision Making: Hypothesis Testing

GSBA 545, Fall 2021

Professor Dawn Porter

Hypothesis Testing

- Null and Alternative Hypotheses in Testing
- Large Sample Tests about a Mean:
  - One-Sided Alternatives Hypotheses
  - Two-Sided Alternative Hypotheses
- *p*-Value approach
- Small Sample Tests about a Population Mean
- Proportion Hypothesis Testing

Null & Alternative Hypotheses

Null hypothesis, denoted H_0, is a statement of the basic proposition being tested.

- The statement generally represents the status quo and is not rejected unless there is convincing sample evidence that it is false.

Alternative or **research hypothesis**, denoted H_a, is an alternative (to the null hypothesis) statement that will be accepted *only if* there is convincing sample evidence that it is true.

Types of Hypotheses

One-Sided, Greater Than

$H_0: \mu \leq 50$	$H_a: \mu > 50$	(Trash bag breaking strength)
$H_0: \mu \leq 12$	$H_a: \mu > 12$	(*Metro EMS* response times)

One-Sided, Less Than

$H_0: \mu \geq 19.5$	$H_a: \mu < 19.5$	(Accounts receivable days)
$H_0: \mu = 183$	$H_a: \mu < 183$	(Cholesterol change with black tea consumption)

Two-Sided, Not Equal To

$H_0: \mu = 4.5$	$H_a: \mu \neq 4.5$	(Camshaft bolt diameters)
$H_0: \mu = 6$	$H_a: \mu \neq 6$	(*Glow Toothpaste* tube wt)

Type I & Type II Errors

Type I Error: Rejecting H_0 when it is true
Type II Error: Failing to reject H_0 when it is false

Conclusion \ *Truth*	H_0 **True**	H_0 **False**
Accept H_0	Correct Conclusion	Type II Error
Reject H_0	Type I Error	Correct Conclusion

Hypotheses: SF Rental Prices

San Francisco Rental Prices

A firm managing rental properties is assessing an expansion into an area of San Francisco. To cover its costs, the firm needs rents in this area to average more than $2000/month.

The director of the firm wants to formulate a hypothesis test that uses a sample of rental properties to determine whether the rental price goal of more than $2000 is feasible.

Are rental prices in that area in San Francisco high enough to warrant expansion?

Null & Alternative: SF Rental Prices

Null Hypothesis: $H_0: \mu \leq \$2000$

Rental prices in the area are not high enough for the firm's expansion; no action taken.

Alternative Hypothesis: $H_a: \mu > \$2000$

Rental prices in the area are high enough for the expansion; firm plans to expand.

* μ = mean rental price for San Francisco properties in a particular area.

Type I & Type II Errors: SF Rental Prices

Type I Error:	Incorrectly deciding the area IS profitable enough
Type II Error:	Incorrectly deciding the area is NOT profitable enough

Conclusion	*Truth*: H_0 True: $\mu \leq \$2000$	*Truth*: H_0 False: $\mu > \$2000$
Do not reject H_0	Correct Conclusion	Type II Error
Reject H_0	Type I Error	Correct Conclusion

Null & Alternative: TV Watching

TV Watching

Past research showed urban preschool children aged 3-5 watch an average of 22.6 hours of television per week. A market research firm believes that the historical mean is too low.

A random sample of 60 urban preschool children was taken and its mean was 24.5 hours of television watching with a standard deviation of 6.1 hours. They plan to test the hypotheses at a significance level of $\alpha = 0.01$.

What is your conclusion about the hours of TV watching?

Null & Alternative: TV Watching

Null Hypothesis: H_0: $\mu \leq 22.6$ hours

The weekly time preschool children watch TV has not significantly increased.

Alternative Hypothesis: H_a: $\mu > 22.6$ hours

The weekly time preschool children watch TV has significantly increased.

* μ = amount of time per week preschool children watch TV.

Type I & Type II Errors: TV Watching

USC School of Business

Type I Error:	Incorrectly deciding children DO watch more TV
Type II Error:	Incorrectly deciding the children no NOT watch more TV

Conclusion \ Truth	H_0 True: $\mu \leq 22.6$	H_0 False: $\mu > 22.6$
Do not reject H_0	Correct Conclusion	Type II Error
Reject H_0	Type I Error	Correct Conclusion

Large Sample, One-Sided Tests

If $n \geq 30$, calculate the ***Z* test statistic** as:

$$z = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}}$$

A **critical value** is the corresponding Z value from a standard normal table corresponding to a desired level of significance α.

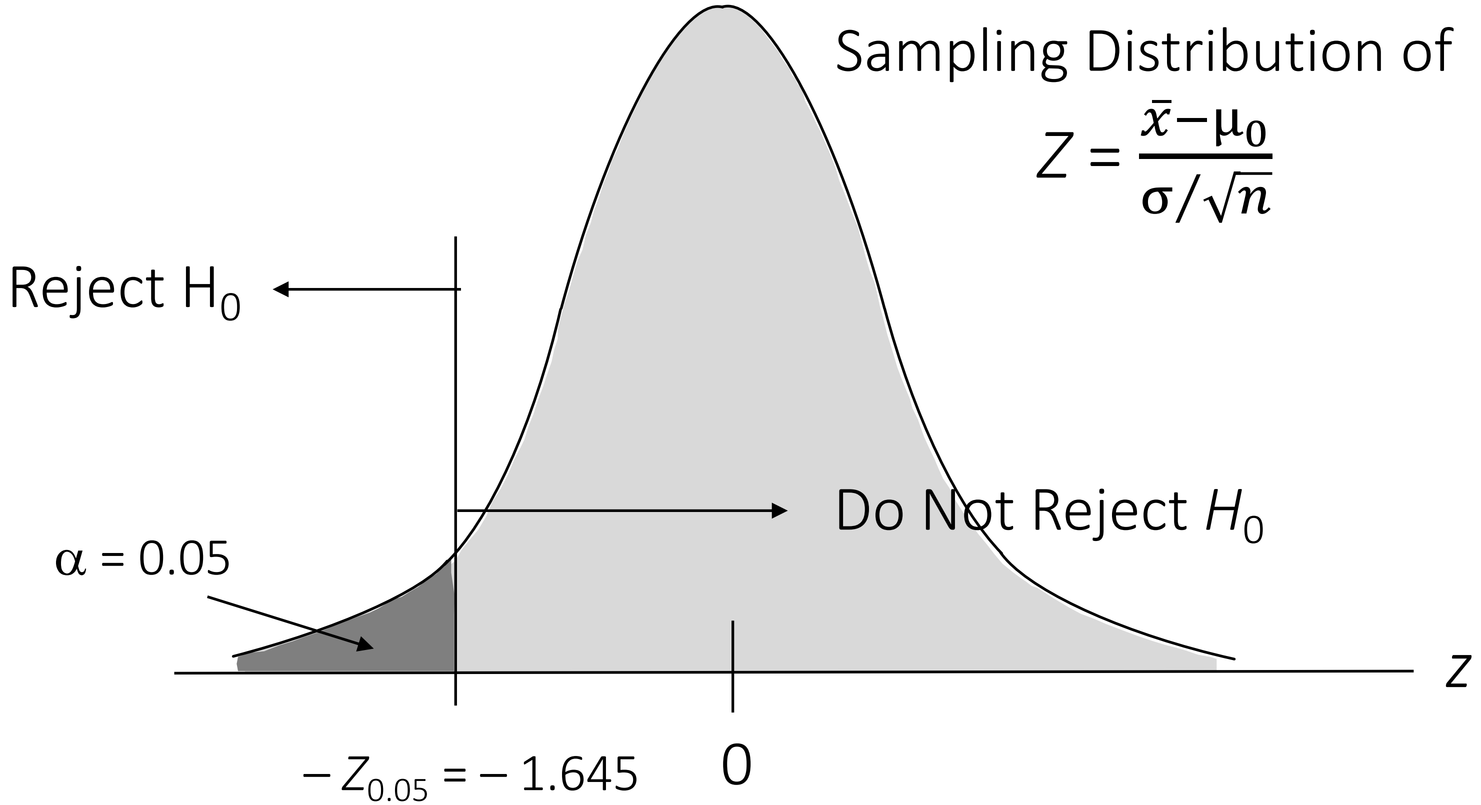
The **rejection rules** for H_0 are:

- Lower tail: Reject H_0 if $z < -z_\alpha$
- Upper tail: Reject H_0 if $z > z_\alpha$

* If σ is unknown and n is large ($n \geq 30$), you can estimate σ by s.

Rejection Regions: Lower Tail

H_0: $\mu \geq \mu_0$ vs H_a: $\mu < \mu_0$ at $\alpha = 0.05$

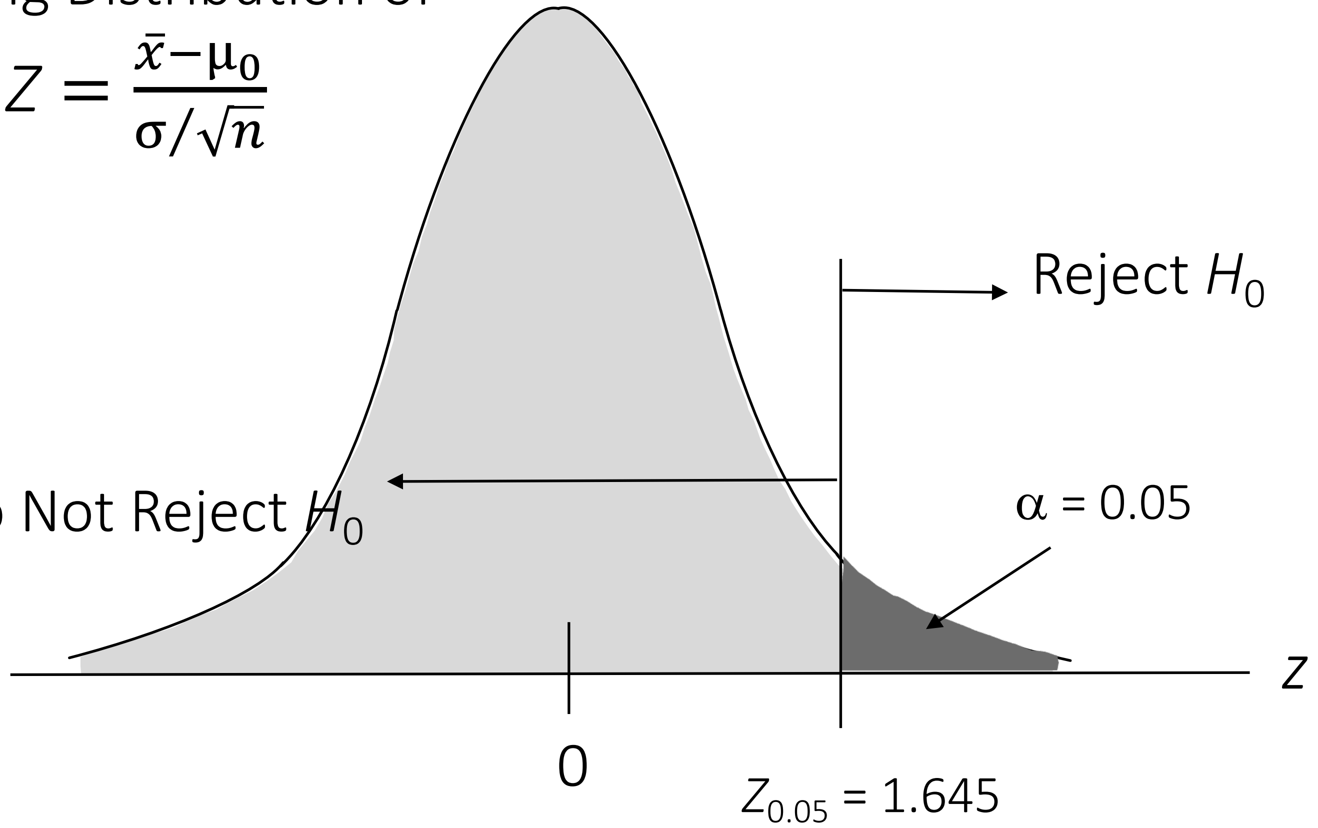


Rejection Regions: Upper Tail

H_0: $\mu \le \mu_0$ vs H_a: $\mu > \mu_0$ at $\alpha = 0.05$

Sampling Distribution of

$$Z = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}}$$



USC School of Business

Hypothesis Testing Steps

1. Determine **null and alternative hypotheses**.
2. Specify level of **significance α**.
3. Calculate **test statistic** value.
4. Determine **critical value(s)** or calculate the ***p*-value**.
5. Decide whether to **reject** H_0 and interpret the statistical result in (real-world) managerial terms.

Hypotheses: SF Rental Prices

San Francisco Rental Prices

Recall a firm managing rental properties is assessing an expansion into an area of San Francisco and needs the rents in this area to average more than \$2000/month. The firm's director will test this at a significance level of α = 0.01.

Rental prices for a sample of 115 properties in San Francisco were found to have an average of $\bar{x} = \$2157$ and the standard deviation s = \$581.

Are rental prices in that area in San Francisco high enough to warrant expansion?

Hypothesis Testing: SF Rental Prices

1. Determine **null and alternative hypotheses.**
2. Specify level of **significance α**.
3. Calculate **test statistic** value.
4. Determine **critical value(s)**.
5. Decide whether to **reject H_0** and interpret the statistical result.

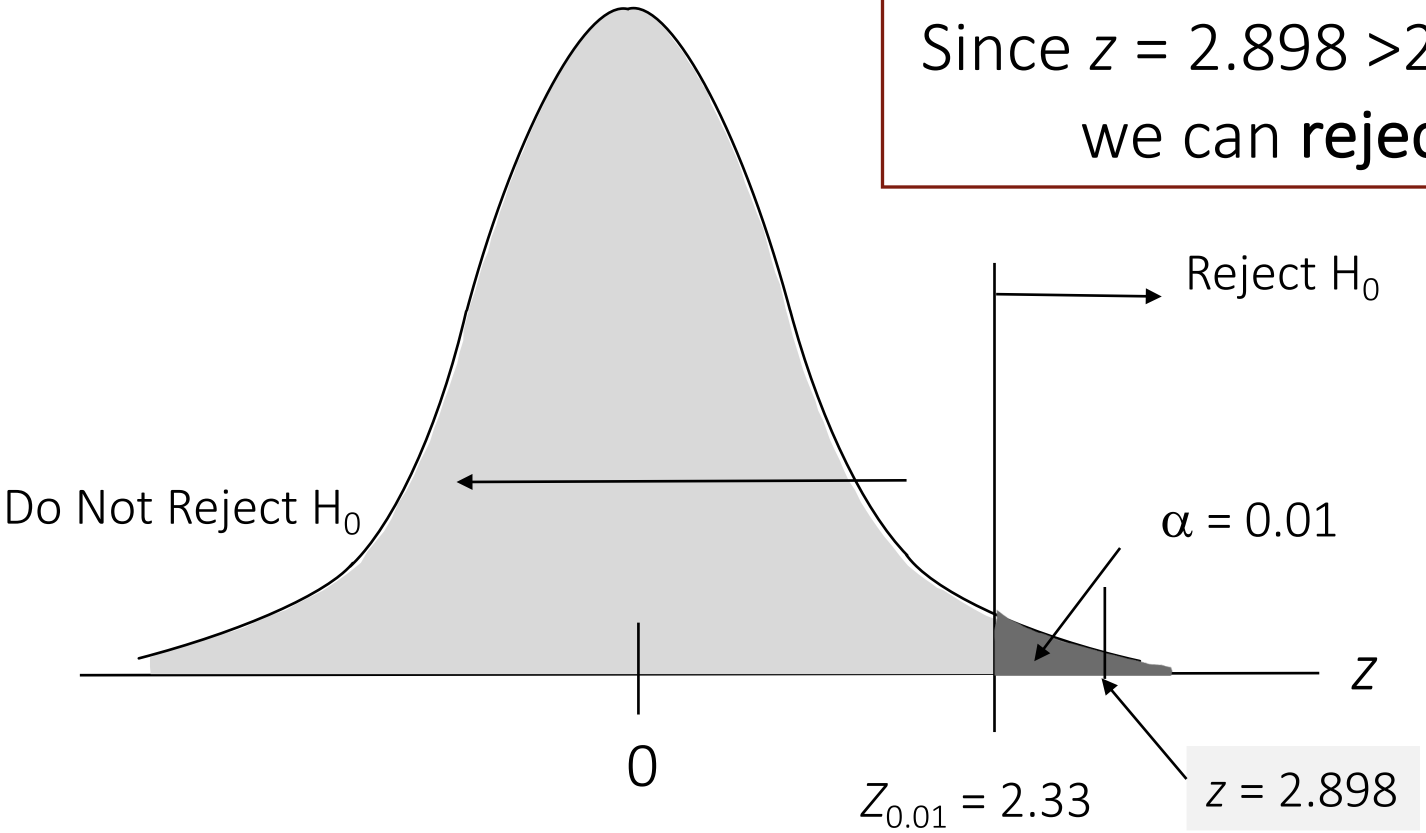
1. H_0: $\mu \le \$2000$ vs H_a: $\mu > \$2000$
2. $\alpha = 0.01$
3. $z = \frac{\bar{x} - \mu_0}{s/\sqrt{n}} = \frac{2157 - 2000}{581/\sqrt{115}} = 2.898$
4. Reject H_0 if $z > z_{0.01} = 2.33$
5. Since $z = 2.898 > 2.33$, we can **reject H_0** and conclude the rental prices are high enough.

Hypothesis Testing: SF Rental Prices

We are testing at $\alpha = 0.01$

H_0: $\mu \le \$2000$ vs H_a: $\mu > \$2000$

Since $z = 2.898 > 2.33 = z_{0.01}$, we can **reject** H_0.



Hypotheses: TV Watching

TV Watching

Recall that past research showed urban preschool children aged 3-5 watch an average of 22.6 hours of television per week and a market research firm believes that the historical mean is too low.

A random sample of 60 urban preschool children was taken and its mean was 24.5 hours of television watching with a standard deviation of 6.1 hours. They plan to test the hypotheses at a significance level of $\alpha = 0.01$.

What is your conclusion about the hours of TV watching?

Hypothesis Testing: TV Watching

1. Determine **null and alternative hypotheses.**
2. Specify level of **significance α.**
3. Calculate **test statistic** value.
4. Determine **critical value(s).**
5. Decide whether to **reject H_0** and interpret the statistical result.

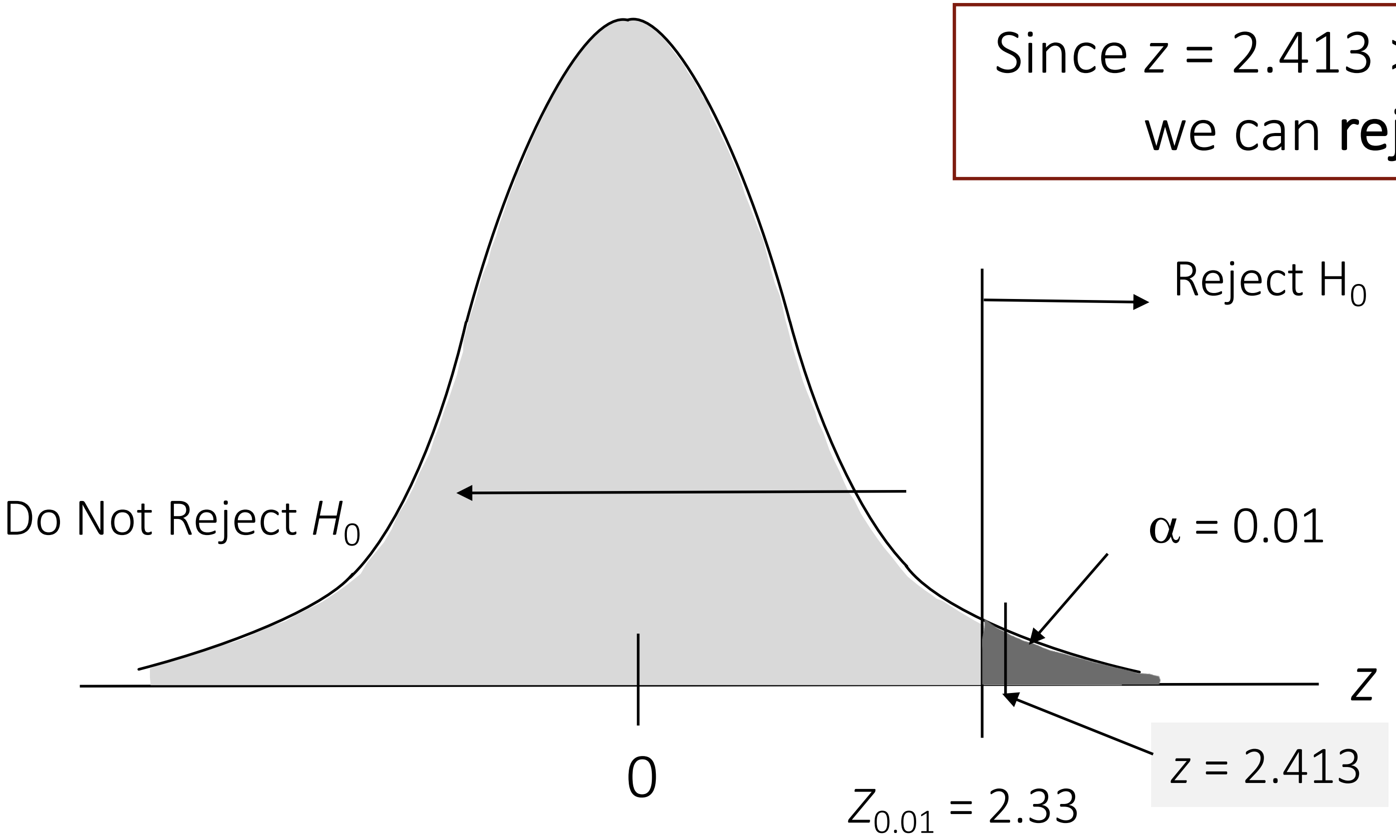
1. H_0: $\mu = 22.6$ vs H_a: $\mu > 22.6$
2. $\alpha = 0.01$
3. $z = \frac{\bar{x}-\mu_0}{s/\sqrt{n}} = \frac{24.5-22.6}{6.1/\sqrt{60}} = 2.413$
4. Reject H_0 if $z > z_{0.01} = 2.33$
5. Since $z = 2.413 > 2.33$, we can **reject H_0** and conclude the viewing times have increased.

Hypothesis Testing: TV Watching

We are testing
at $\alpha = 0.01$

H_0: $\mu = 22.6$ vs H_a: $\mu > 22.6$

Since $z = 2.413 > 2.33 = z_{0.01}$,
we can **reject H_0**.



Large Sample Testing: 2-Sided

A **two-sided** hypothesis test is appropriate when the context of the situation either

(1) does NOT indicate any expected direction results, or
(2) indicates BOTH directions are to be assessed.

If $n \geq 30$, the Z **test statistic** is again calculated as: $z = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}}$

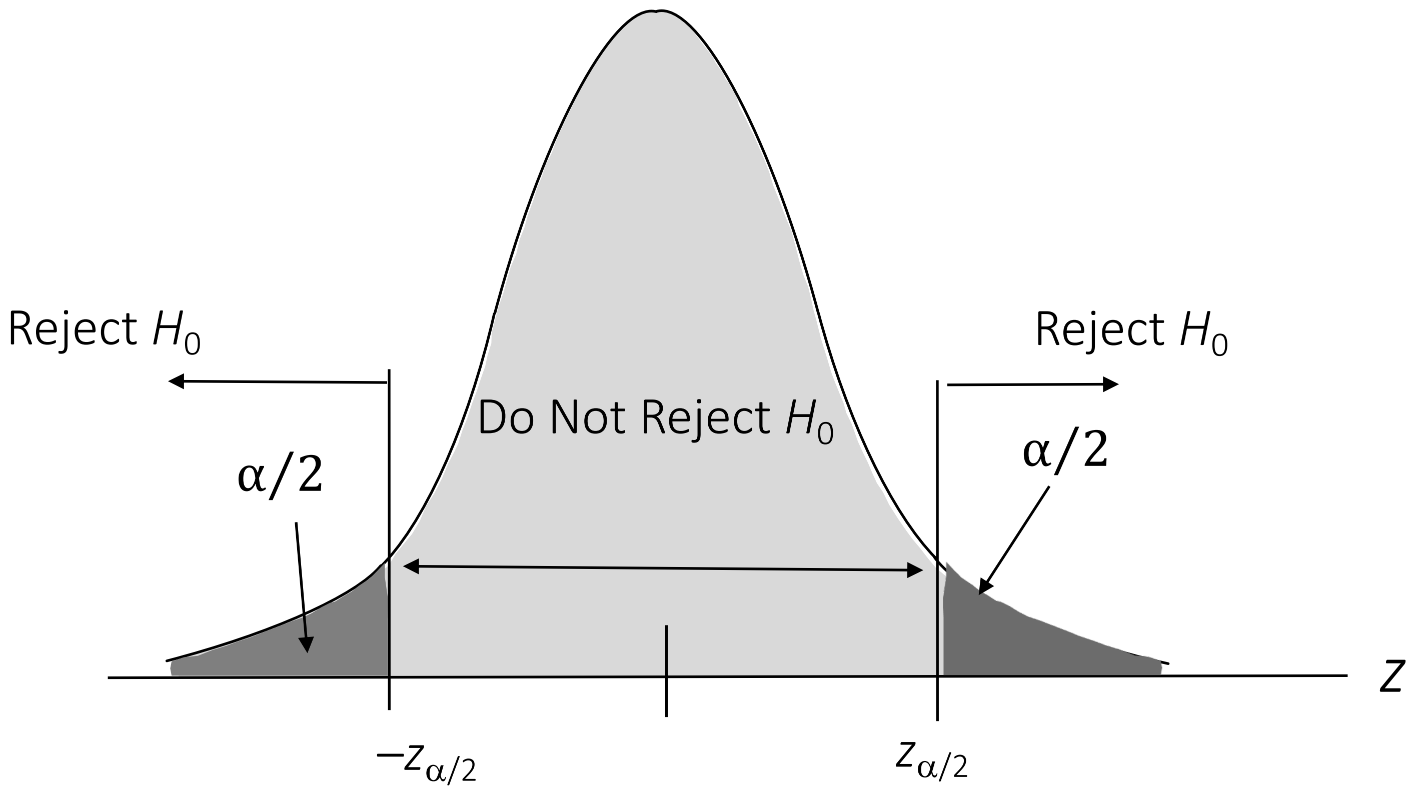
Critical values are the corresponding $\pm Z$ values from a standard normal table corresponding to a desired level of significance $\alpha/2$.

The **rejection rules** for H_0 are: Reject H_0 if $z < -z_{\alpha/2}$ or if $z > z_{\alpha/2}$.

* If σ is unknown and n is large ($n \geq 30$), you can estimate σ by s.

Large Sample Testing: 2-Sided

Assume we are testing H_0: $\mu = \mu_0$ vs H_a: $\mu \neq \mu_0$



Hypotheses: Glow Toothpaste

Glow Toothpaste

The production line for Glow toothpaste is designed to fill tubes with a mean weight of 6 oz. Periodically, a sample of 30 tubes will be selected in order to check the filling process at α = 0.01.

Quality assurance procedures indicate the filling process will continue if the sample results are consistent with the mean filling weight of 6 oz.; otherwise the process will be adjusted.

A sample of 30 toothpaste tubes gives $\bar{x} = 6.1$ oz with a standard deviation of 0.2 oz.

Are the toothpaste tubes being filled, on average, to a weight of 6 oz?

Hypothesis Testing: Glow Toothpaste

1. Determine **null and alternative hypotheses.**
2. Specify level of **significance α**.
3. Calculate **test statistic** value.
4. Determine **critical value(s).**
5. Decide whether to **reject H_0** and interpret the statistical result.

1. H_0: $\mu = 6$ oz vs H_a: $\mu \neq 6$ oz

2. $\alpha = 0.01$

3. $z = \frac{\bar{x} - \mu_0}{s/\sqrt{n}} = \frac{6.1 - 6}{0.20/\sqrt{30}} = 2.74$

4. Reject H_0 if $|z| > z_{0.005} = 2.575$

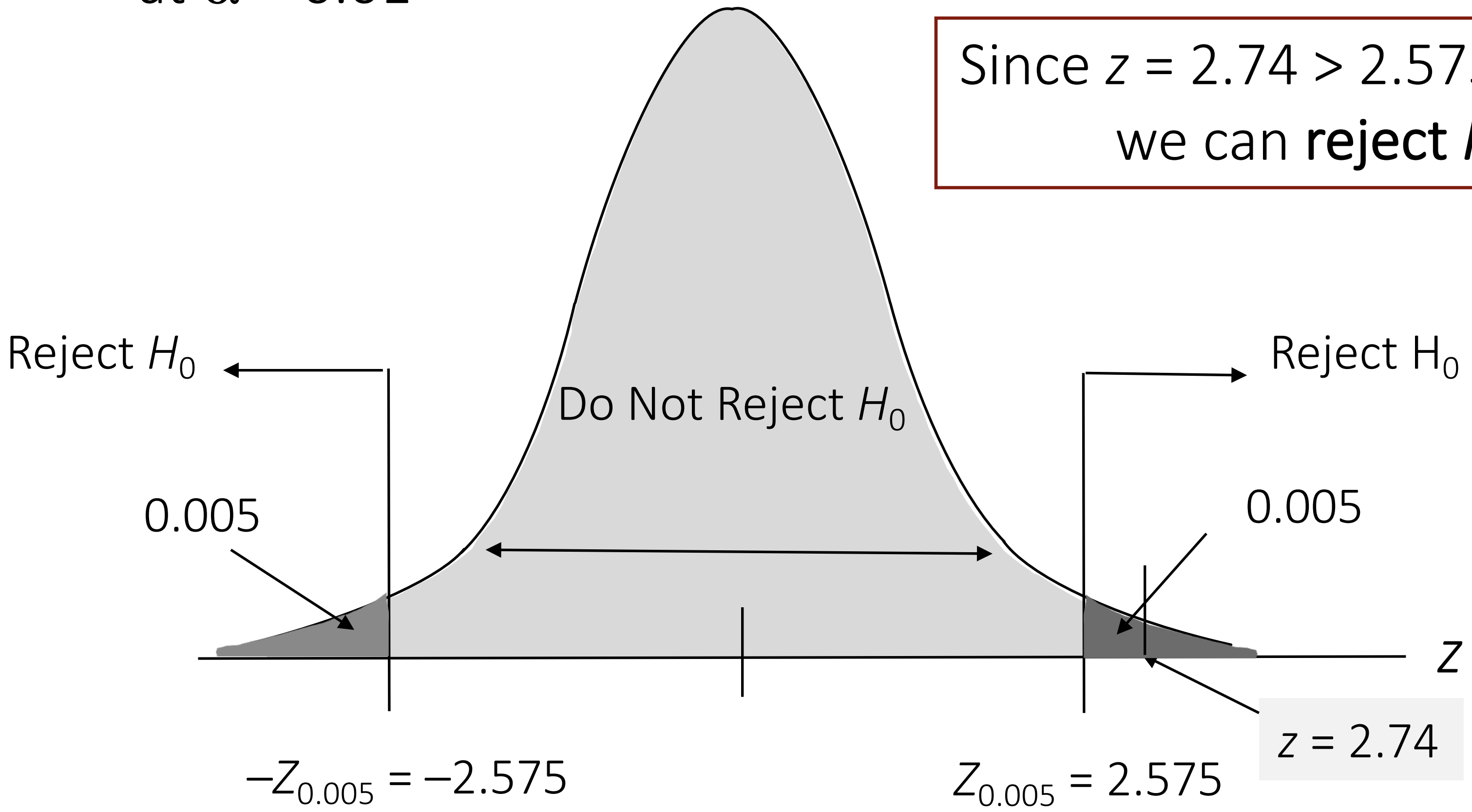
5. Since $z = 2.74 > 2.575$, we can **reject H_0** and conclude the mean filling weight is NOT 6 oz.

Hypothesis Testing: Glow Toothpaste

We are testing at α = 0.01

H_0: $\mu = 6$ oz vs H_a: $\mu \neq 6$ oz

Since $z = 2.74 > 2.575 = z_{0.005}$, we can **reject H_0.**



2-Sided Hypothesis Testing: CI Approach

A **two-sided hypothesis** can also be tested with a confidence interval calculation.

Create a confidence interval for the population mean, μ, in the usual way from the sample statistics.

The **rejection rule** for H_0 is:

- If the CI does contain μ_0, **do NOT reject H_0.**
- If the CI does not contain μ_0, **reject H_0.**

CI Approach: Glow Toothpaste

Create a confidence interval for the population mean, μ, in the usual way from the sample statistics at $\alpha = 0.01$:

$$\boxed{\bar{x} \pm z_{\alpha/2}\frac{s}{\sqrt{n}} = 6.1 \pm 2.575\frac{0.20}{\sqrt{30}} = 6.1 \pm 0.094 \rightarrow [6.006, \quad 6.194]}$$

Since the confidence interval does NOT contain H_0: $\mu = 6$ oz, we should **reject the null hypothesis** and conclude the mean filling weight is NOT 6 oz.

*Note: this is the same result we obtained using the two-tailed hypothesis test approach.

Hypothesis Testing: p-Values

A **p–value** is :

- The smallest level of a for which the null hypothesis can be rejected.
- The area under the curve beyond the calculated test statistic, in the direction of the alternative hypothesis.
- The level of support for the null hypothesis being true.

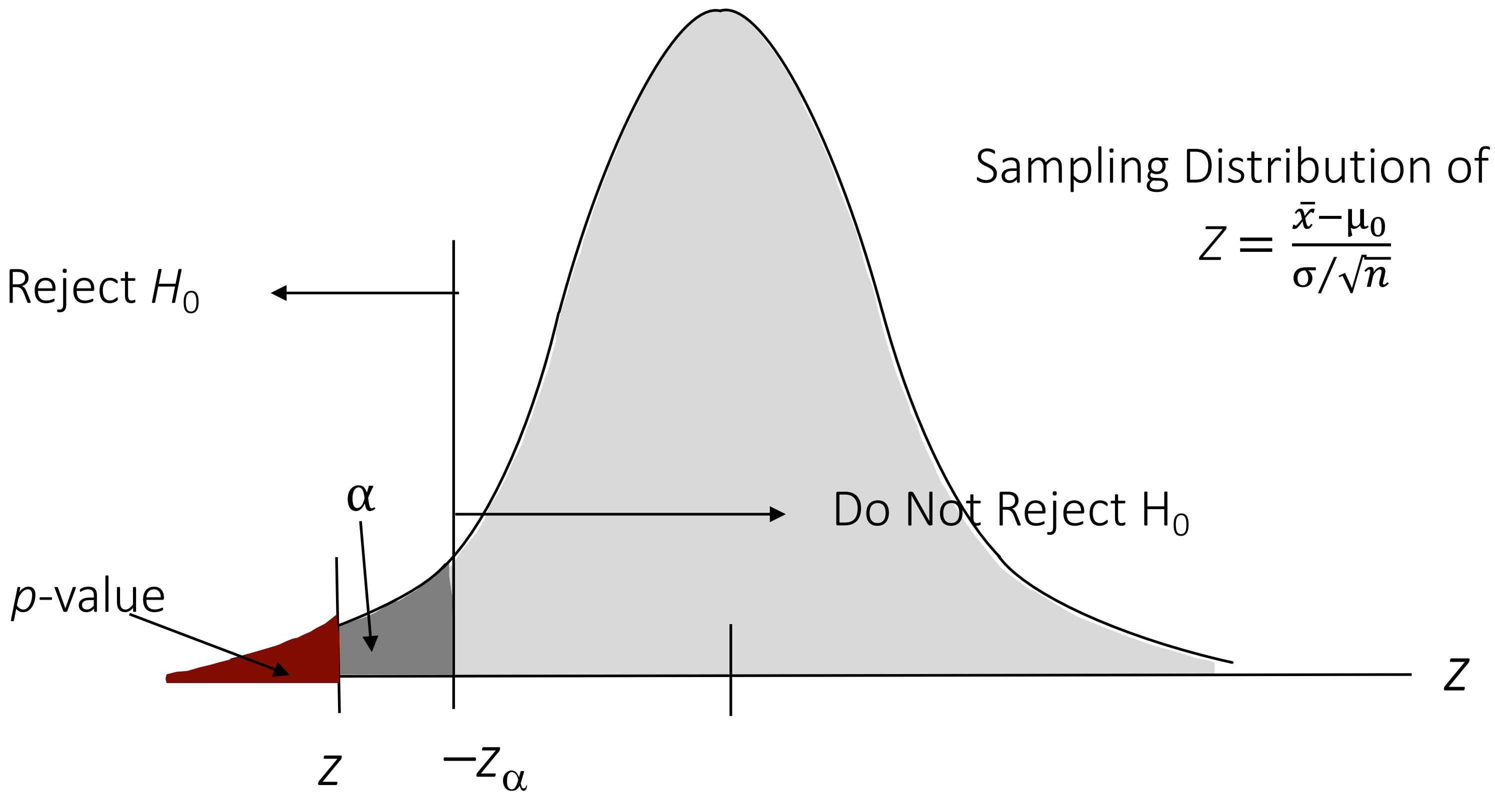
Rule: Reject H_0 if the ***p*-value** < α

*Note: when conducting a two-sided test, the area under the curve needs to be doubled.

Hypothesis Testing: p-Values

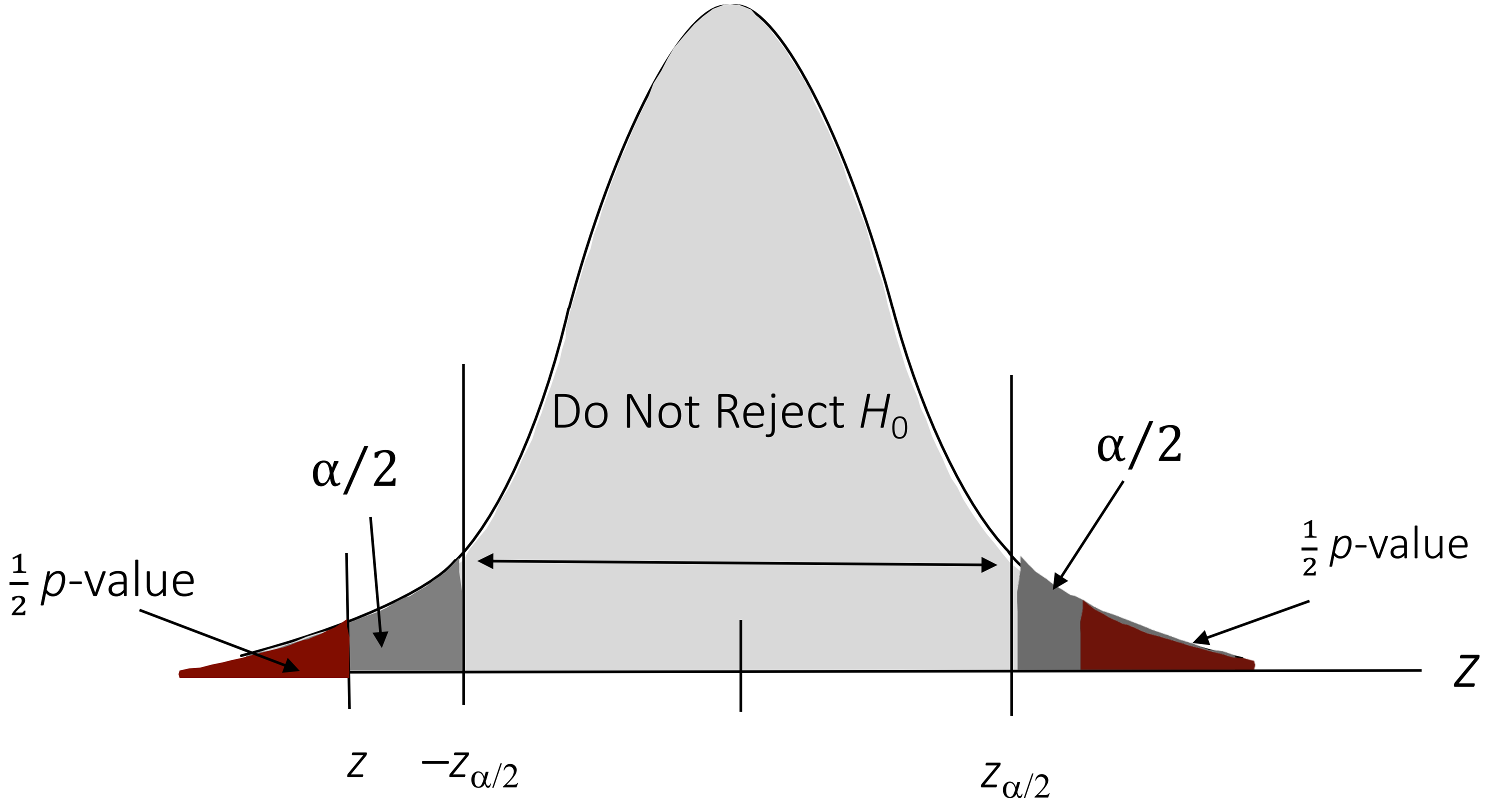
Assume we are testing

H_0: $\mu \geq \mu_0$ vs H_a: $\mu < \mu_0$



2-Sided Hypothesis Testing: *p*-Values

Assume we are testing $H_0\!: \mu = \mu_0$ vs $H_a\!: \mu \neq \mu_0$



p-Values: SF Rental Prices

USC School of Business

San Francisco Rental Prices

Recall a firm managing rental properties is assessing an expansion into an area of San Francisco but the rents need to average more than \$2000/month. The firm's director will test this at $\alpha = 0.01$.

Rental prices for a sample of 115 properties in San Francisco had an average of $\bar{x} = \$2157$ and a standard deviation $s = \$581$.

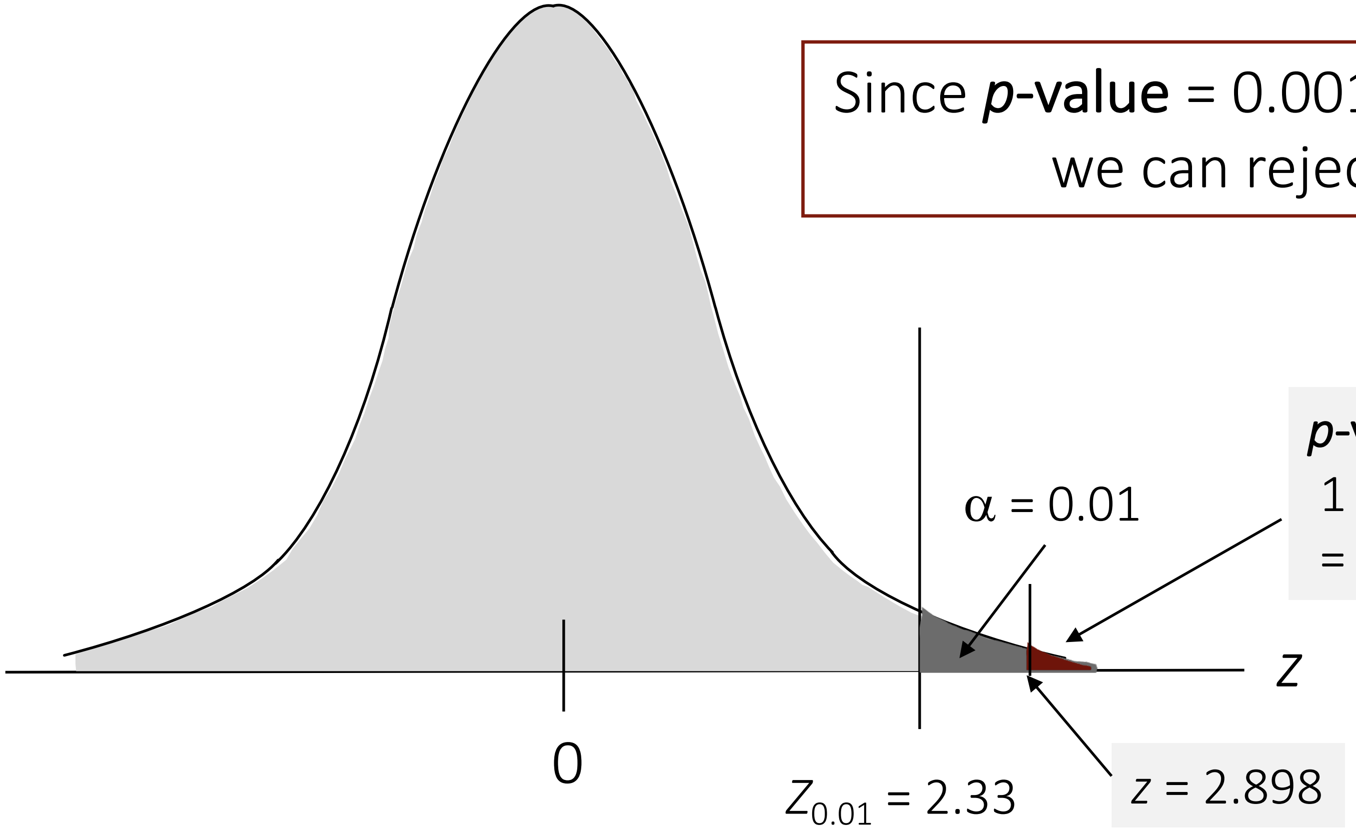
1. H_0: $\mu \le \$2000$ vs H_a: $\mu > \$2000$
2. $\alpha = 0.01$
3. $z = \frac{\bar{x} - \mu_0}{s/\sqrt{n}} = \frac{2157 - 2000}{581/\sqrt{115}} = 2.898$
4. Reject H_0 if $z > z_{0.01} = 2.33$
5. Since the *p*-value = 0.0019 < 0.01 = α, we can **reject** H_0 and conclude the rental prices are high enough. (Remember also that $z = 2.898 > 2.33 \rightarrow$ **reject** H_0.)

p–Values: SF Rental Prices

We are testing at $\alpha = 0.01$

H_0: $\mu \leq \$2000$ vs H_a: $\mu > \$2000$

Since ***p*-value** $= 0.0019 < \alpha = 0.01$, we can reject H_0.



p-Values: TV Watching

TV Watching

Recall that past research showed urban preschool children aged 3-5 watch an average of 22.6 hrs of TV per week. A market research firm feels this mean is too low.

A sample of 60 children was taken, with a mean 24.5 hrs and a std dev of 6.1 hrs of TV watching. They will test the hypotheses at a $\alpha = 0.01$.

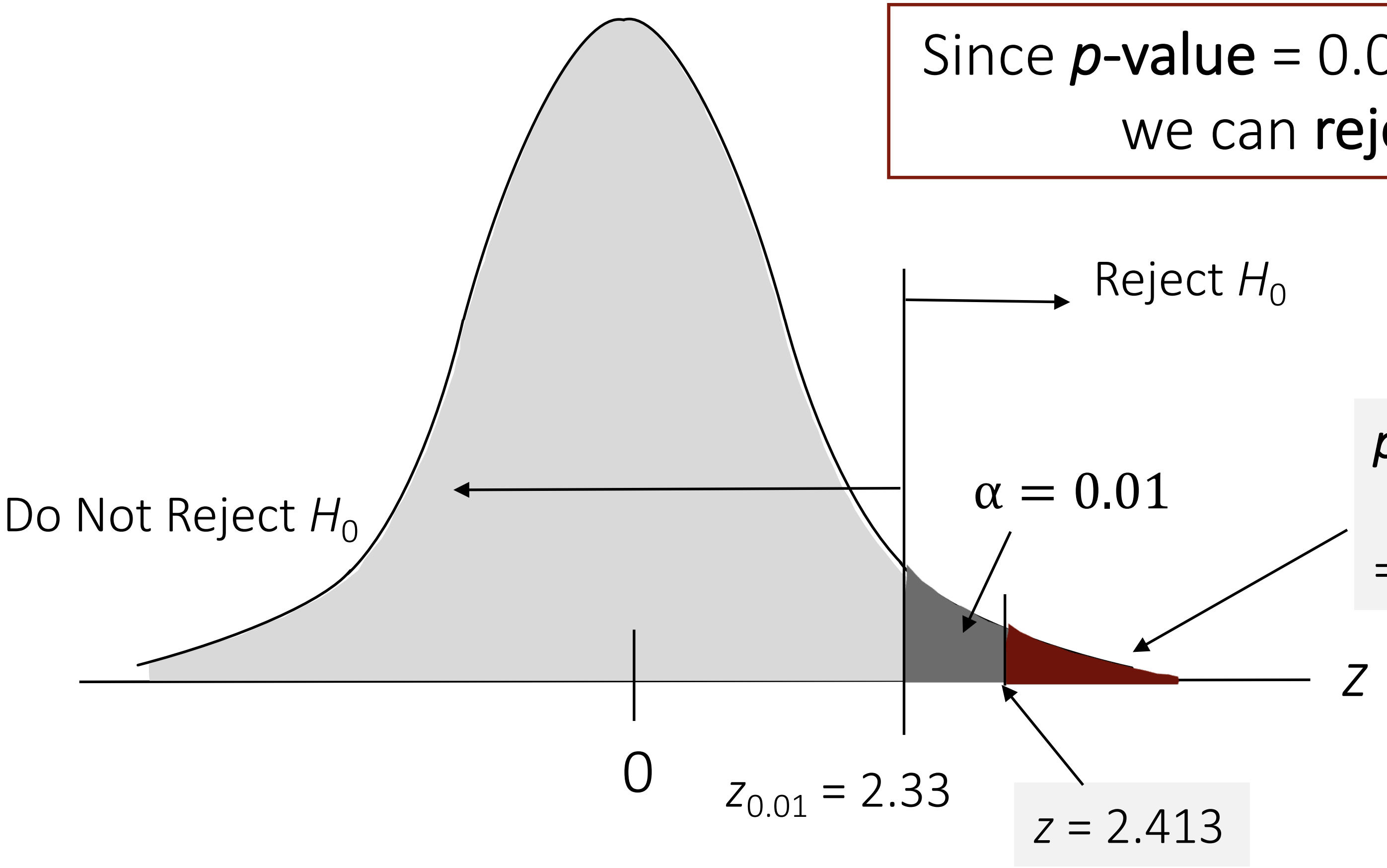
1. H_0: $\mu = 22.6$ vs H_a: $\mu > 22.6$
2. $\alpha = 0.01$
3. $z = \frac{\bar{x}-\mu_0}{s/\sqrt{n}} = \frac{24.5-22.6}{6.1/\sqrt{60}} = 2.413$
4. Reject H_0 if $z > z_{0.01} = 2.33$
5. Since the p-value = 0.008 < 0.01 = α, we can **reject H_0** and conclude the viewing times have increased. (Remember also that $z = 2.413 > 2.33 \rightarrow$ **reject H_0**.)

p-Values: TV Watching

We are testing at $\alpha = 0.01$

H_0: $\mu = 22.6$ vs H_a: $\mu > 22.6$

Since ***p*-value** $= 0.008 < 0.01 = \alpha$, we can **reject** H_0.



p-Values: Glow Toothpaste

Glow Toothpaste

Recall Glow toothpaste's production should fill tubes to a mean weight of 6 oz. 30-tube samples periodically selected to check the filling process at $\alpha = 0.01$.

A sample of 30 toothpaste tubes gives $\bar{x} = 6.1$ oz with a standard deviation of 0.2 oz.

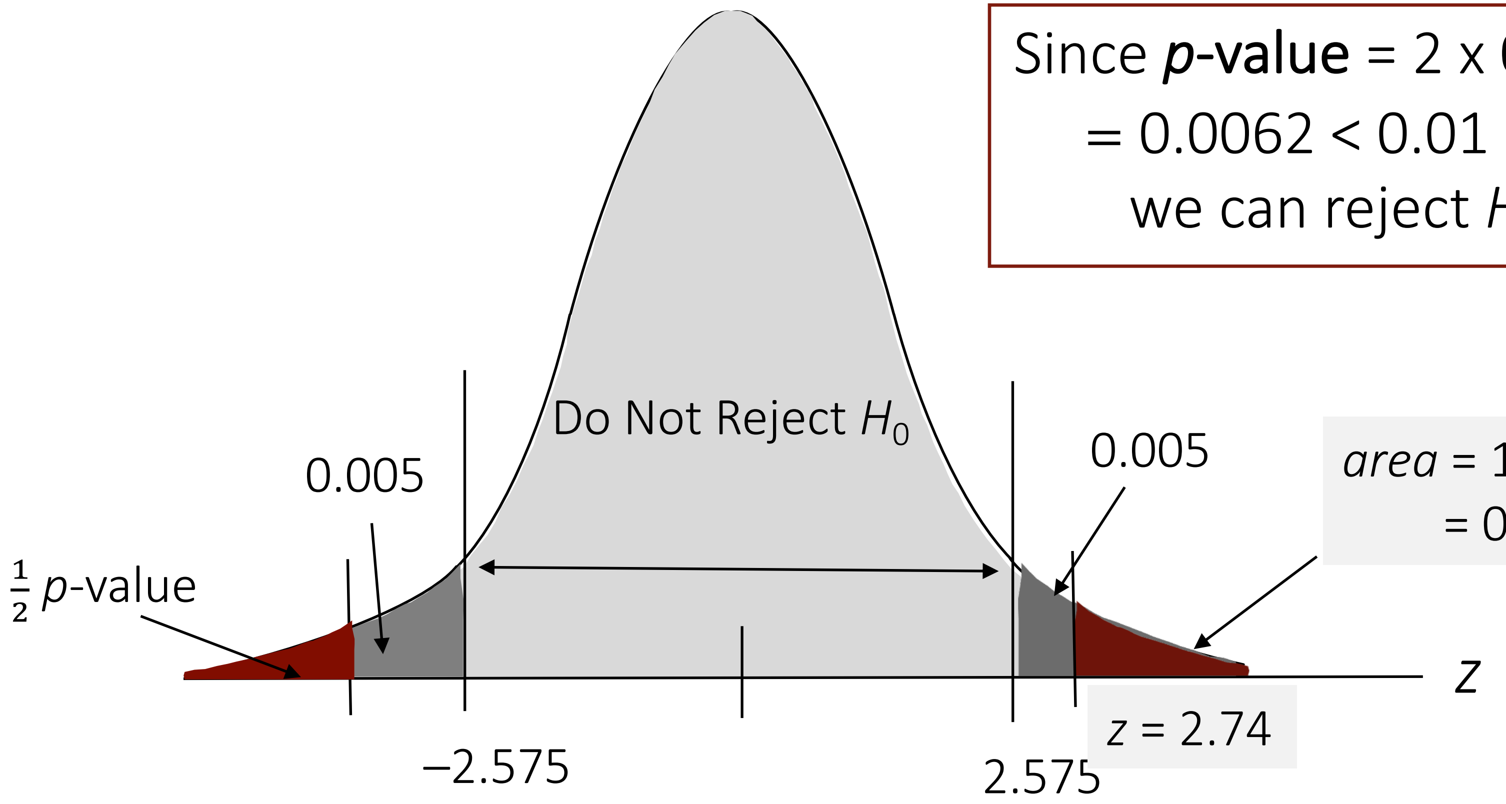
1. H_0: $\mu = 6$ oz vs H_a: $\mu \neq 6$ oz
2. $\alpha = 0.01$
3. $z = \frac{\bar{x} - \mu_0}{s/\sqrt{n}} = \frac{6.1 - 6}{0.20/\sqrt{30}} = 2.74$
4. Reject H_0 if $|z| > z_{0.005} = 2.575$
5. Since the *p*-value = 2 x 0.0031 = 0.0062 < 0.01 = α, we can **reject** H_0 and conclude the mean filling weight is NOT 6 oz. (Remember also that $z = 2.74 > 2.575 \rightarrow$ **reject** H_0.)

p-Values: Glow Toothpaste

We are testing at α = 0.01

H_0: μ = 6 oz vs H_a: $\mu \neq$ 6 oz

Since ***p*-value** = 2 x 0.0031 = 0.0062 < 0.01 = α, we can reject H_0.



Small Sample Hypothesis Testing

When $n < 30$, calculate the ***t* test statistic** as:

$$t = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}}$$

A **critical value** is the corresponding t value, with $df = n - 1$,from the T-distribution table corresponding to significance level α.

The **rejection rule** for H_0 is:

- Lower tail: Reject H_0 if $t < -t_{\alpha,n-1}$
- Upper tail: Reject H_0 if $t > t_{\alpha,n-1}$
- Two-sided: Reject H_0 if $|t| > t_{\alpha/2,n-1}$

* If σ is unknown, you can estimate σ by s.

Small Sample Test: Highway MPH

Highway Patrol Speed Traps

A State Highway Patrol periodically samples vehicle speeds at various locations on a particular roadway. A sample of vehicle speeds is used to test the hypothesis H_0: $\mu \leq 65$ at $\alpha = 0.05$.

The locations where H_0 is rejected are deemed the best locations for radar traps. At one particular location, a sample of 24 vehicles shows a mean speed of 66.2mph with a standard deviation of 4.2mph.

Statistically, are people driving faster than the speed limit at that location?

Small Sample Test: Highway MPH

1. Determine **null and alternative hypotheses.**
2. Specify level of **significance α.**
3. Calculate **test statistic** value.
4. Determine **critical value(s).**
5. Decide whether to **reject H_0** and interpret the statistical result.

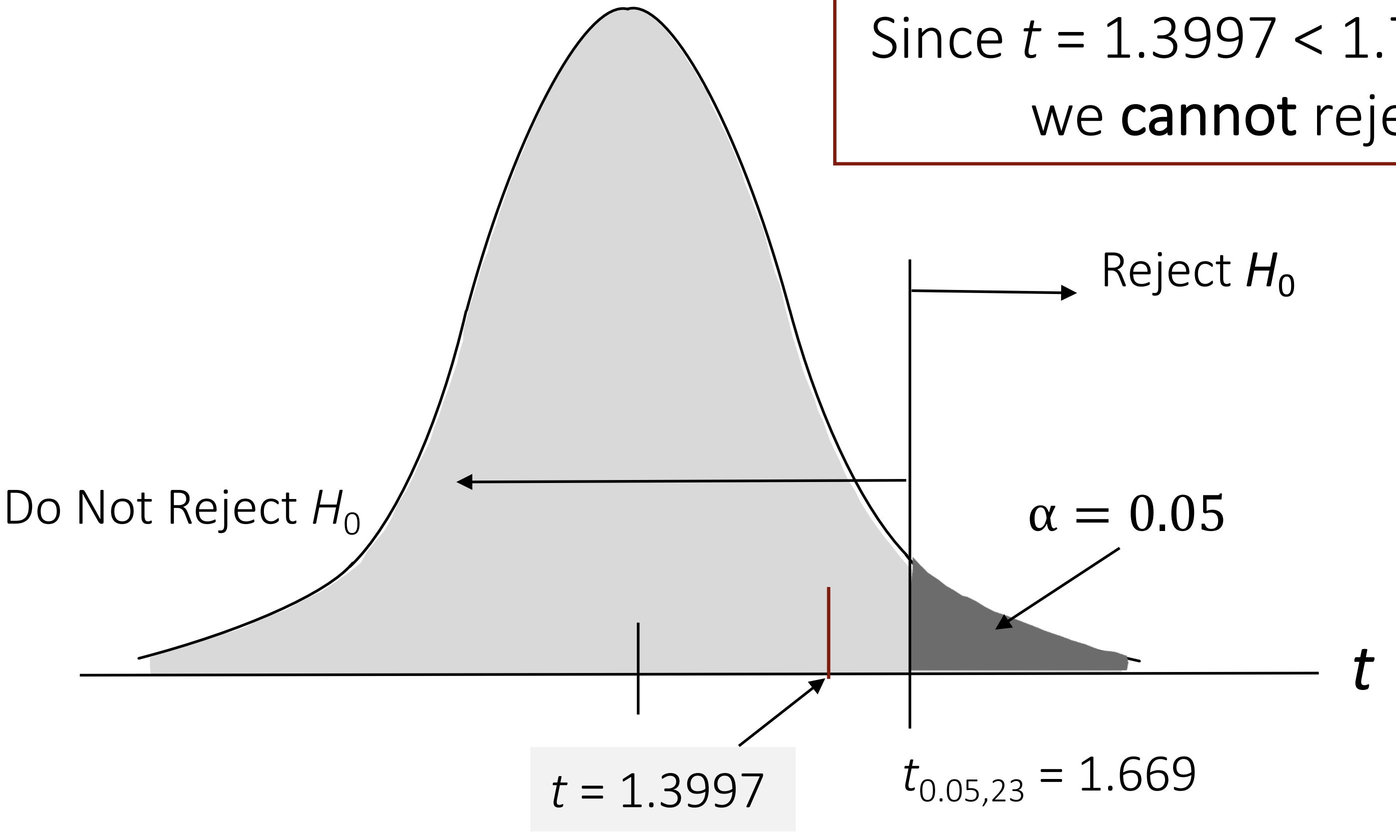
1. H_0: $\mu \leq 65$ vs H_a: $\mu > 65$
2. $\alpha = 0.05$
3. $t = \frac{\bar{x} - \mu_0}{s/\sqrt{n}} = \frac{66.2 - 65}{4.2/\sqrt{24}} = 1.3997$
4. Reject H_0 if $t > t_{0.05,23} = 1.714$
5. Since $t = 1.3997 < 1.714$, we **cannot reject H_0** and conclude the cars are NOT driving faster than 65mph on average.

Small Sample Test: Highway MPH

We are testing
at $\alpha = 0.05$

H_0: $\mu \leq 65$ vs H_a: $\mu > 65$

Since $t = 1.3997 < 1.714 = t_{0.05,23}$,
we **cannot** reject H_0.

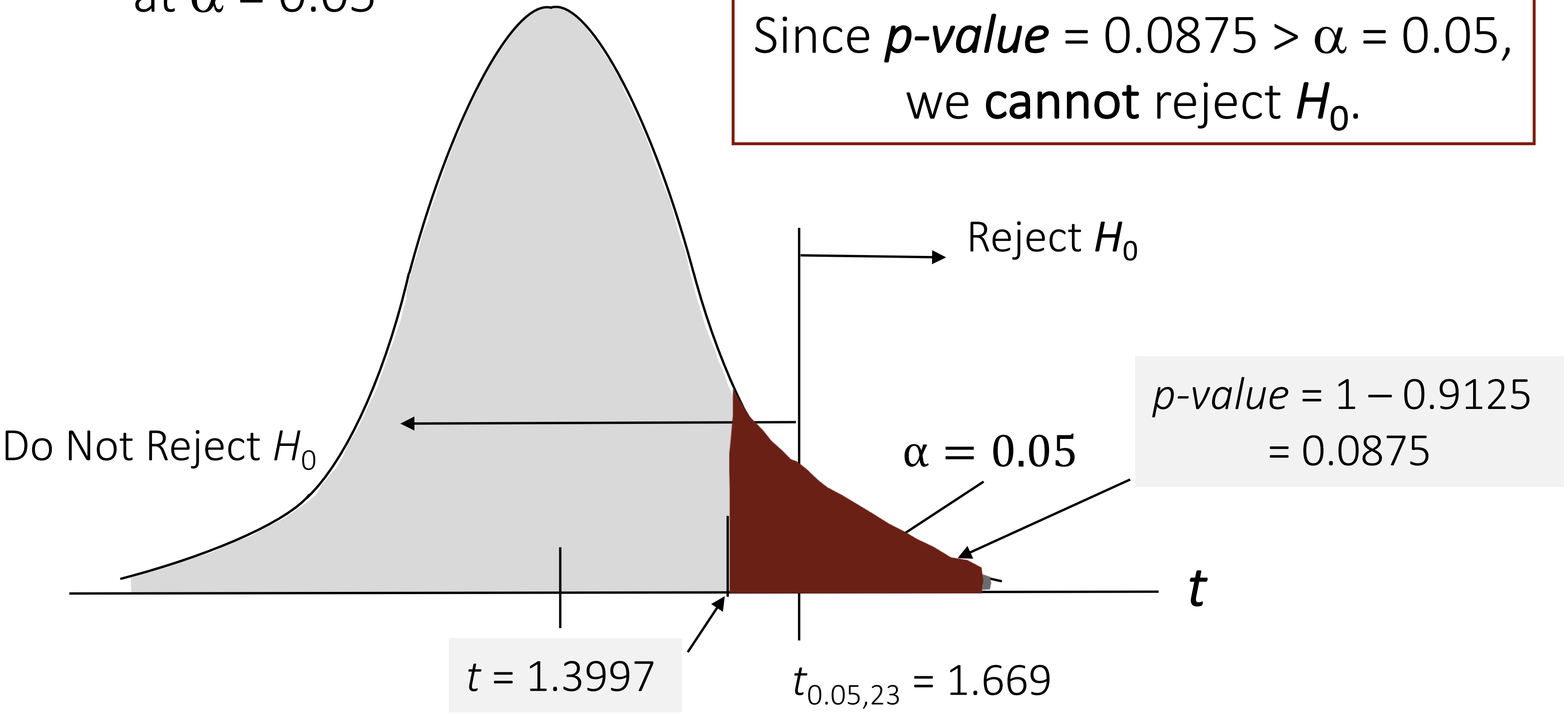


p-Value: Highway MPH

We are testing at $\alpha = 0.05$

H_0: $\mu \le 65$ vs H_a: $\mu > 65$

Since ***p-value*** = 0.0875 > α = 0.05, we **cannot** reject $\boldsymbol{H_0}$.



* The p-value here needs to be found using software (Excel, Minitab, JMP)

Small Sample Test: CC Interest Rates

Credit Card Interest Rates

In the past, the average interest rate on major credit card purchases was 18.8%. Due to changes in economic conditions and payment methods, it is of interest to see if this rate has significantly decreased from the past value.

A random sample of 15 credit cards was taken and the average rate was 17.9% with a standard deviation of 1.09%. They plan to test the hypotheses at a significance level of $\alpha = 0.05$.

Have the credit card interest rates significantly decreased?

Credit Card Interest Rates: Test Statistic

1. Determine **null and alternative hypotheses.**
2. Specify level of **significance α.**
3. Calculate **test statistic** value.
4. Determine **critical value(s).**
5. Decide whether to **reject H_0** and interpret the statistical result.

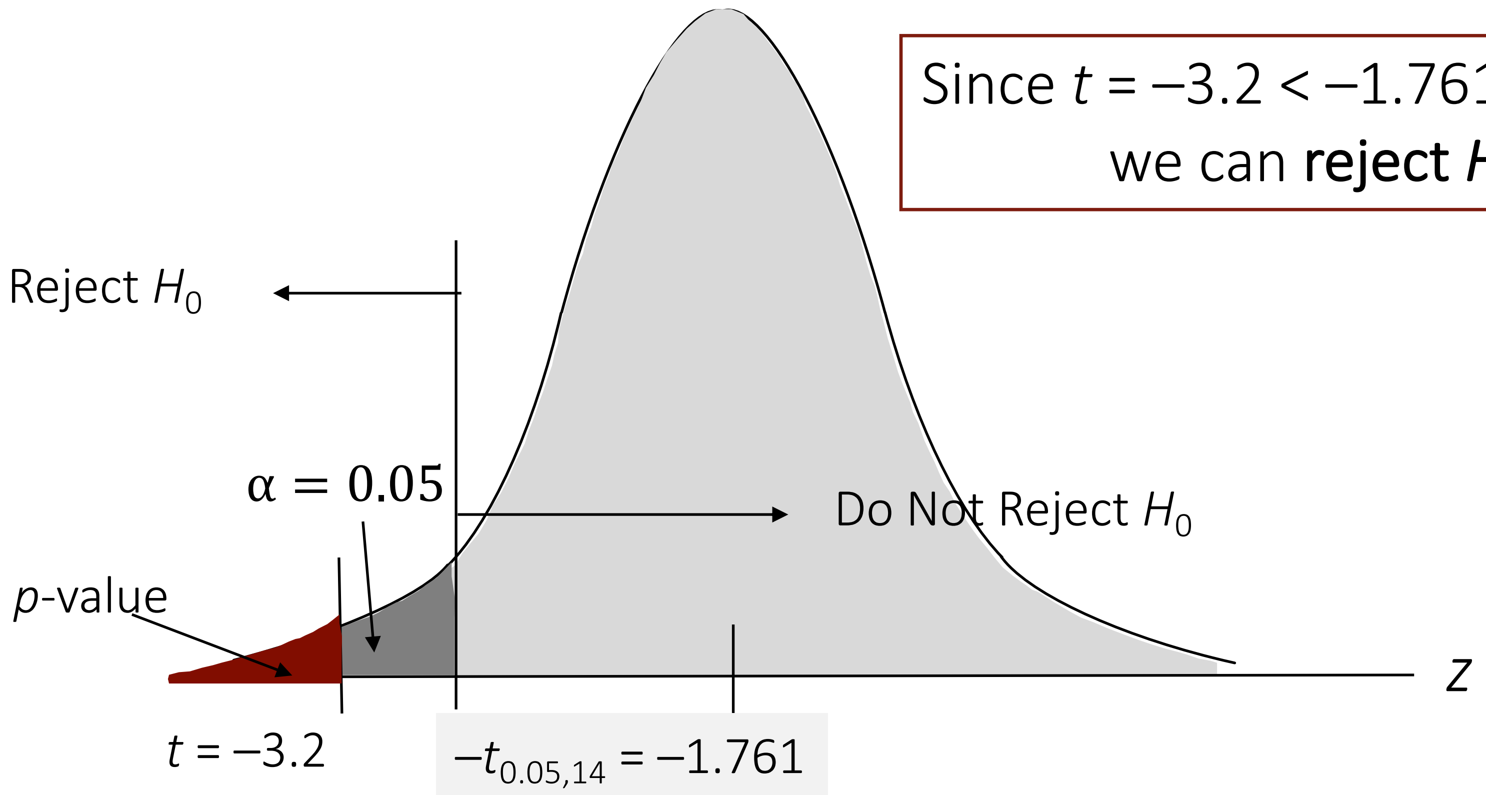
1. H_0: $\mu = 18.8$ vs H_a: $\mu < 18.8$
2. $\alpha = 0.05$
3. $t = \frac{\bar{x} - \mu_0}{s/\sqrt{n}} = \frac{17.9 - 18.8}{1.09/\sqrt{15}} = -3.2$
4. Reject H_0 if $t < t_{0.05,14} = -1.761$
5. Since $t = -3.2 < -1.761$, we can **reject** H_0 and conclude the credit card rates have decreased.

Credit Card Interest Rates: Test Statistic

We are testing at $\alpha = 0.05$

H_0: $\mu = 18.8$ vs H_a: $\mu < 18.8$

Since $t = -3.2 < -1.761 = t_{0.05,14}$, we can **reject H_0**.

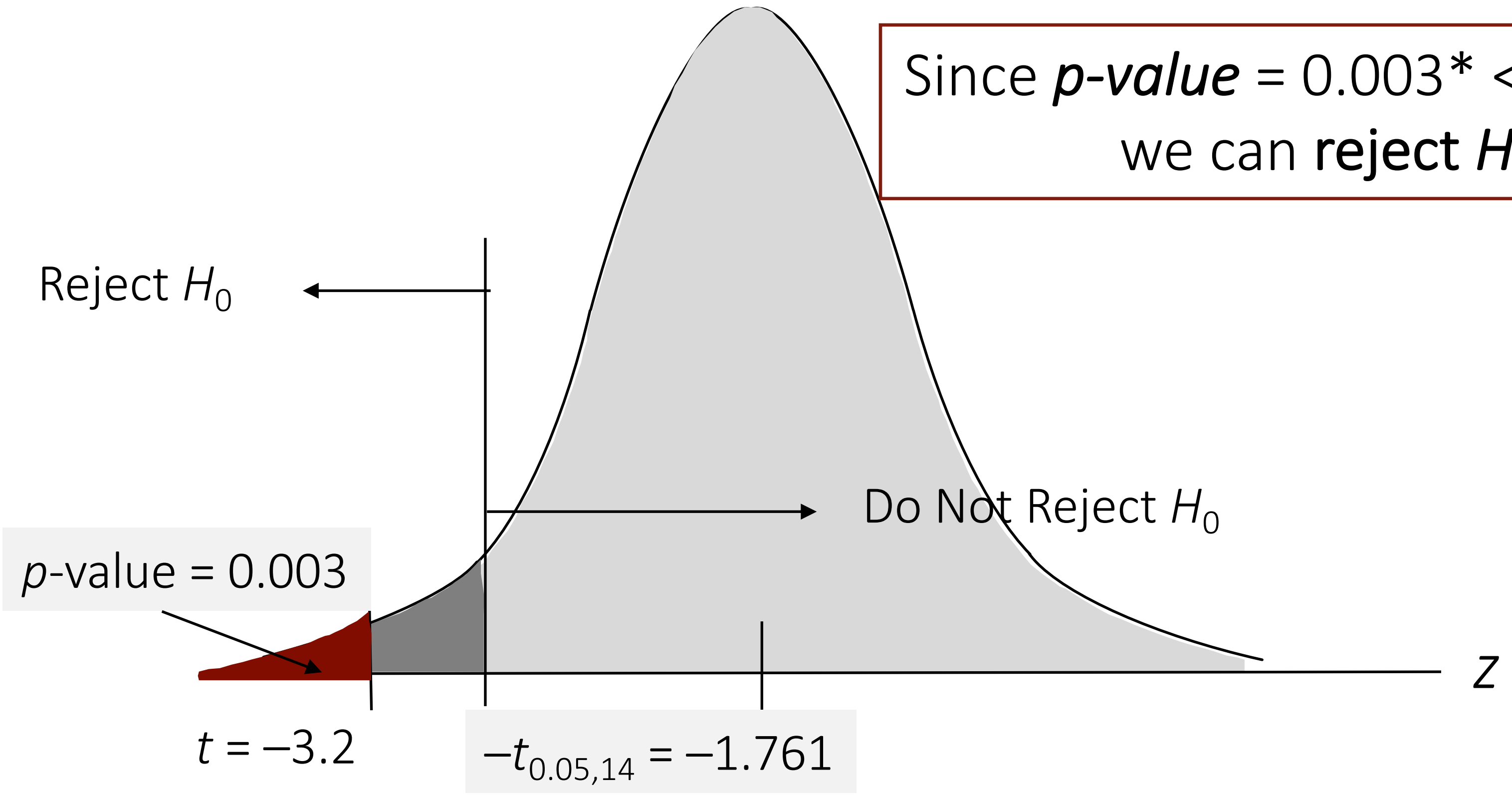


Credit Card Interest Rates: *p*-Value

We are testing at $\alpha = 0.05$

H_0: $\mu = 18.8$ vs H_a: $\mu < 18.8$

Since ***p-value*** $= 0.003^* < \alpha = 0.05$, we can **reject** $\boldsymbol{H_0}$.



* The p-value here needs to be found using software (Excel, Minitab, JMP)

Hypothesis Testing: Proportions

Calculate the *Z* **test statistic*** as:

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

A **critical value** is the corresponding *Z* value from a standard normal table corresponding to a desired level of significance α.

The **rejection rules** for H_0 are:

- **Lower tail:** Reject H_0 if $z < -z_\alpha$
- **Upper tail:** Reject H_0 if $z > z_\alpha$
- **Two-sided:** Reject H_0 if $z > z_{\alpha/2}$ or if $z < -z_{\alpha/2}$
- ***p*-value:** Reject H_0 if *p*-value $< \alpha$

*Note: technically we must have $np \geq 5$ and $n(1-p) \geq 5$.

Hypothesis Testing: Proportions

One-Sided, Greater Than

$H_0: p \leq 0.05$ $H_a: p > 0.05$ (Burger King Ad)

One-Sided, Less Than

$H_0: p \geq 0.20$ $H_a: p < 0.20$ (Spam filter percentage)

Two-Sided, Not Equal To

$H_0: p = 0.50$ $H_a: p \neq 0.50$ (National Safety Council)

Proportion Testing: Spam Filtering

Spam Filtering

A firm is considering buying software that claims to filter spam email better than what the firm currently uses, which results in only 24% of incoming emails being spam. The office manager will use a trial version of the software to test the system, and she has determined that it will be worth its cost if it can reduce the spam level to less than 20%.

A sample of 100 employees showed an average spam rate of 12% when using the new trial software. Use these data to test the spam filter's rate with $\alpha = 0.05$.

Is the spam filtering software valuable enough for the firm to purchase it?

USC School of Business

Proportion Testing: Spam Filtering

1. Determine **null and alternative hypotheses.**
2. Specify level of **significance α.**
3. Calculate **test statistic** value.
4. Determine **critical value(s).**
5. Decide whether to **reject H_0** and interpret the statistical result.

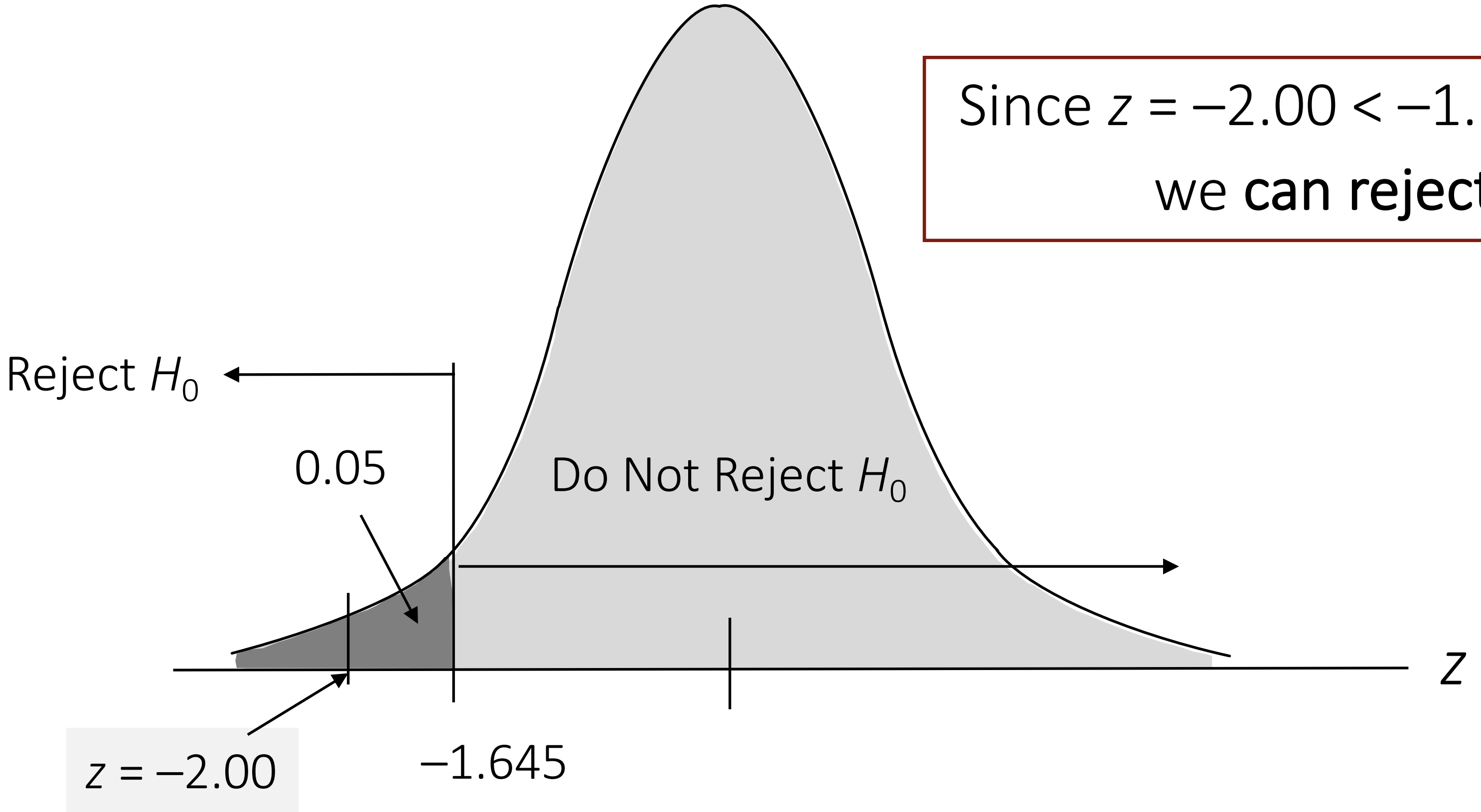
1. H_0: $p \geq 0.2$ vs H_a: $p < 0.2$
2. $\alpha = 0.05$
3. $z = \dfrac{\hat{p}-p_0}{\sqrt{\dfrac{p_0(1-p_0)}{n}}} = \dfrac{0.12-0.2}{\sqrt{\dfrac{0.2(1-0.2)}{100}}} = -2.00$
4. Reject H_0 if $z < -z_{0.05} = -1.645$
5. Since $z = -2.00 < -1.645$, we **can reject H_0** and should conclude the spam filtering software is valuable enough to purchase.

Proportion Testing: Spam Filtering

We are testing at $\alpha = 0.05$

H_0: $p \geq 0.2$ vs H_a: $p < 0.2$

Since $z = -2.00 < -1.645 = -z_{0.05}$, we **can reject** H_0.



p-Value: Spam Filtering

We are testing at $\alpha = 0.05$

H_0: $p \geq 0.2$ vs H_a: $p < 0.2$

Since ***p*-value** = 0.0228 < 0.05 = α, we **can reject** H_0.