IE 340/440

PROCESS IMPROVEMENT THROUGH PLANNED EXPERIMENTATION

Fundamentals of Hypothesis Testing: One-Sample Tests

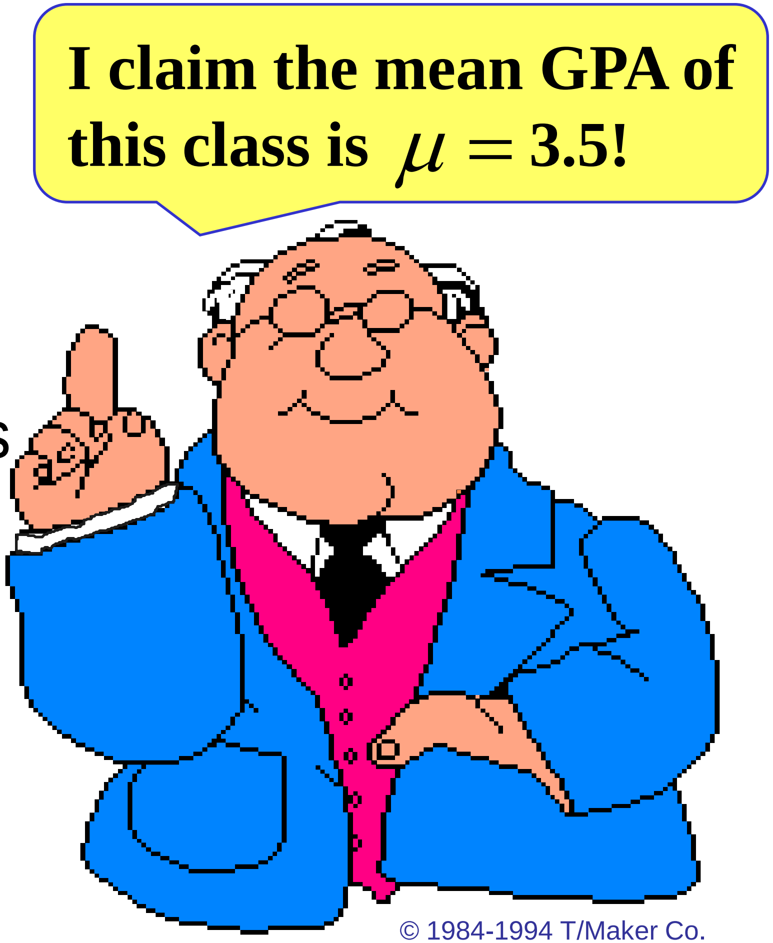
Dr. Xueping Li
University of Tennessee

Chapter Topics

- Hypothesis Testing Methodology
- Z Test for the Mean (σ Known)
- p-Value Approach to Hypothesis Testing
- Connection to Confidence Interval Estimation
- One-Tail Tests
- t Test for the Mean (σ Unknown)
- Z Test for the Proportion
- Potential Hypothesis-Testing Pitfalls and Ethical Issues

What is a Hypothesis?

- A Hypothesis is a Claim (Assumption) about the Population Parameter
  - Examples of parameters are population mean or proportion
  - The parameter must be identified before analysis



The Null Hypothesis, H_0

- States the Assumption (Numerical) to be Tested
  - E.g., The mean GPA is 3.5
  - $H_0 : \mu = 3.5$
- Null Hypothesis is Always about a Population Parameter ($H_0 : \mu = 3.5$), Not about a Sample Statistic ($H_0 : \bar{X} = 3.5$)
- Is the Hypothesis a Researcher Tries to Reject

The Null Hypothesis, H_0 *(continued)*

- Begin with the Assumption that the Null Hypothesis is True
  - Similar to the notion of innocent until proven guilty

- Refer to the Status Quo
- Always Contains the "=" Sign
- The Null Hypothesis May or May Not be Rejected

The Alternative Hypothesis, H_1

- Is the Opposite of the Null Hypothesis
  - E.g., The mean GPA is NOT 3.5 ($H_1 : \mu \neq 3.5$)
- Challenges the Status Quo
- Never Contains the "=" Sign
- The Alternative Hypothesis May or May Not Be Accepted (i.e., The Null Hypothesis May or May Not Be Rejected)
- Is Generally the Hypothesis that the Researcher Claims

Error in Making Decisions

- Type I Error
  - Reject a true null hypothesis
    - When the null hypothesis is rejected, we can say that "We have shown the null hypothesis to be false (with some 'slight' probability, i.e. α, of making a wrong decision)
  - Has serious consequences
  - Probability of Type I Error is α
    - Called level of significance
    - Set by researcher

Error in Making Decisions

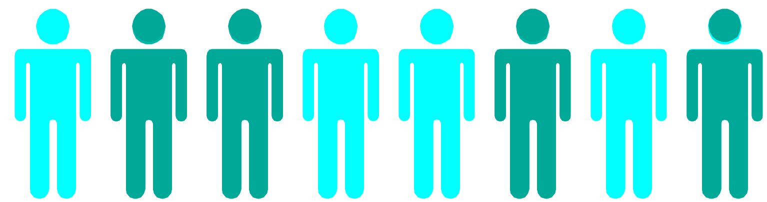
(continued)

- Type II Error
  - Fail to reject a false null hypothesis
  - Probability of Type II Error is β
  - The power of the test is $(1-\beta)$
- Probability of Not Making Type I Error
  - $(1-\alpha)$
  - Called the Confidence Coefficient

Hypothesis Testing Process

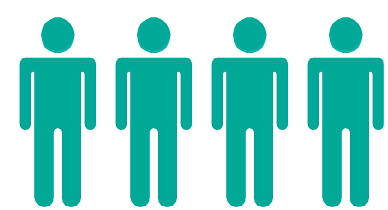
Assume the population mean GPA is 3.5

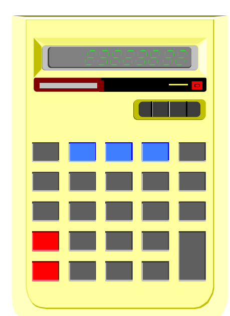
$(H_0 : \mu = 3.5)$



Identify the Population

Take a Sample





$(\bar{X} = 2.4)$

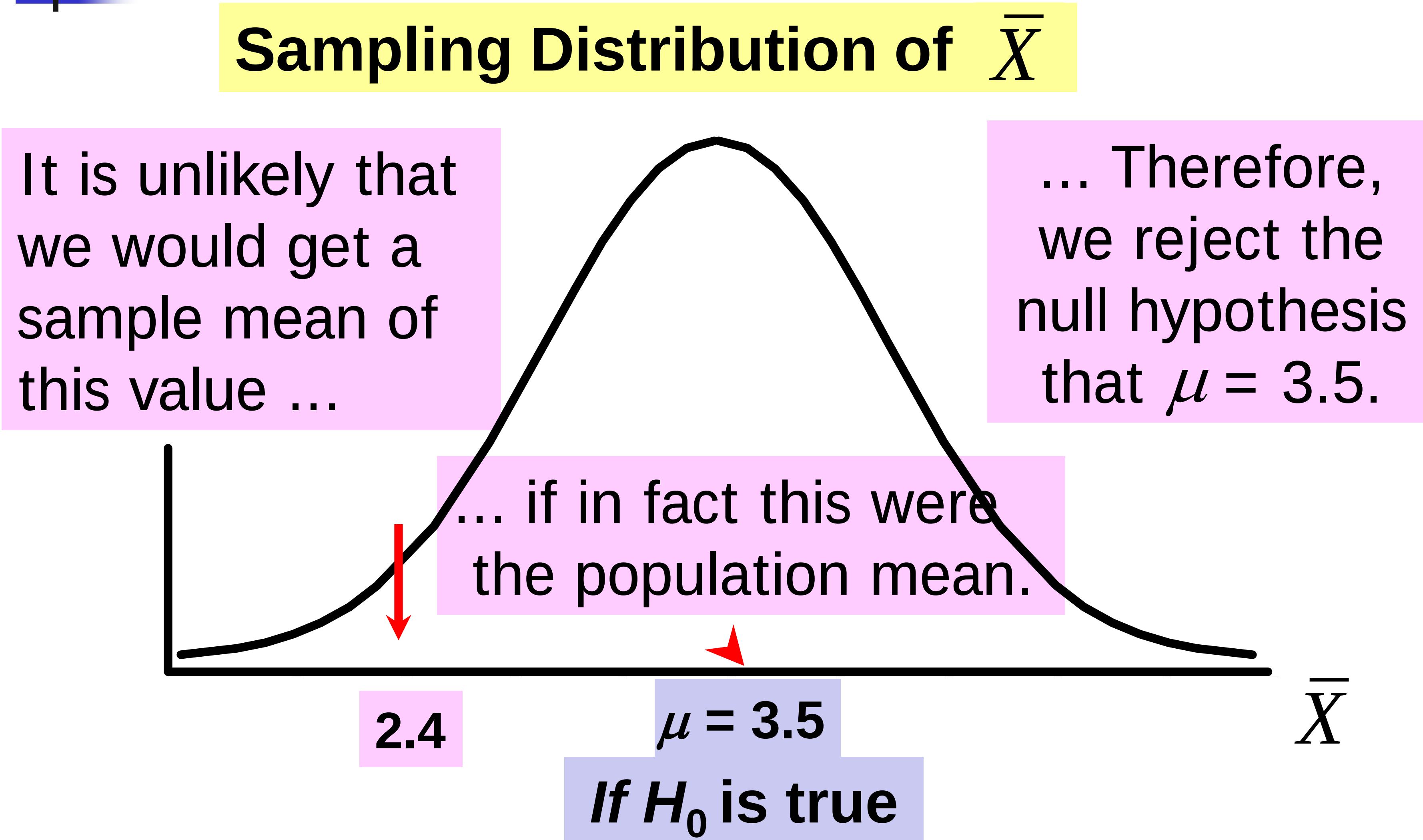
Is $\overline{\mathrm{X}} = 2.4$ likely if $\mu = 3.5$?

No, not likely!

REJECT

Null Hypothesis

Reason for Rejecting H_0



Level of Significance, α

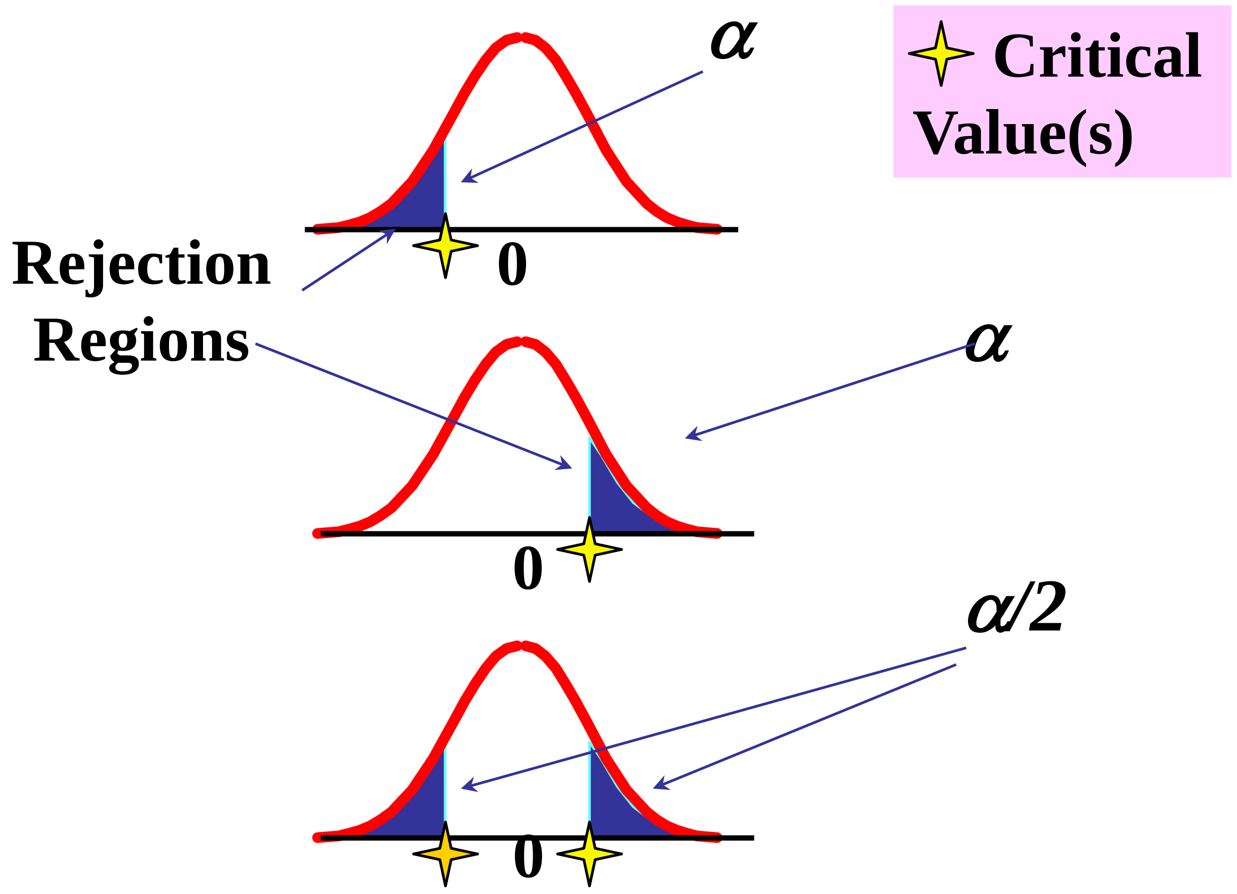
- Defines Unlikely Values of Sample Statistic if Null Hypothesis is True
  - Called rejection region of the sampling distribution
- Designated by α , (level of significance)
  - Typical values are .01, .05, .10
- Selected by the Researcher at the Beginning
- Controls the Probability of Committing a Type I Error
- Provides the Critical Value(s) of the Test

Level of Significance and the Rejection Region

H_0: $\mu \geq 3.5$
H_1: $\mu < 3.5$

H_0: $\mu \leq 3.5$
H_1: $\mu > 3.5$

H_0: $\mu = 3.5$
H_1: $\mu \neq 3.5$



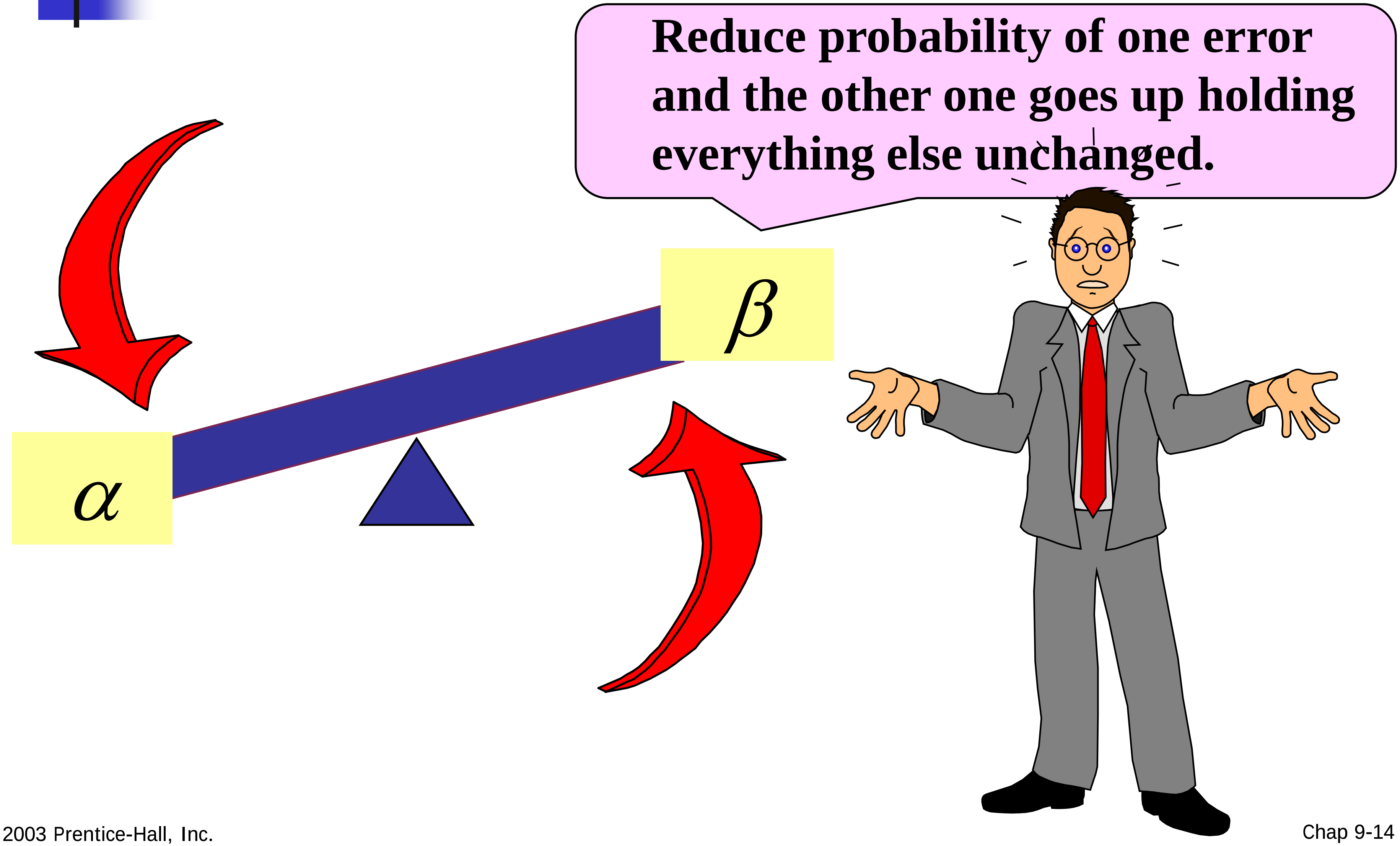
Result Probabilities

H_0: Innocent

Jury Trial		
	The Truth	
Verdict	Innocent	Guilty
Innocent	**Correct**	**Error**
Guilty	**Error**	**Correct**

Hypothesis Test		
	The Truth	
Decision	H_0 True	H_0 False
Do Not Reject **H_0**	**1 - α**	**Type II Error (β)**
Reject **H_0**	**Type I Error (α)**	**Power (1 - β)**

Type I & II Errors Have an Inverse Relationship



Factors Affecting Type II Error

- True Value of Population Parameter
  - β increases when the difference between the hypothesized parameter and its true value decrease
- Significance Level
  - β increases when α decreases

  $\beta \uparrow \quad \alpha \downarrow$
- Population Standard Deviation
  - β increases when σ increases

  $\beta \uparrow \quad \sigma \uparrow$
- Sample Size
  - β increases when n decreases

  $\beta \uparrow \quad n \downarrow$

How to Choose between Type I and Type II Errors

- Choice Depends on the Cost of the Errors
- Choose Smaller Type I Error When the Cost of Rejecting the Maintained Hypothesis is High
  - A criminal trial: convicting an innocent person
  - The Exxon Valdez: causing an oil tanker to sink
- Choose Larger Type I Error When You Have an Interest in Changing the Status Quo
  - A decision in a startup company about a new piece of software
  - A decision about unequal pay for a covered group

Critical Values Approach to Testing

- Convert Sample Statistic (e.g., $\bar{X}$) to Test Statistic (e.g., *Z*, *t* or *F* –statistic)
- Obtain Critical Value(s) for a Specified α from a Table or Computer
  - If the test statistic falls in the critical region, reject H_0
  - Otherwise, do not reject H_0

p-Value Approach to Testing

- Convert Sample Statistic (e.g., $\bar{X}$) to Test Statistic (e.g., *Z, t* or *F* –statistic)
- Obtain the p-value from a table or computer
  - p-value: probability of obtaining a test statistic as extreme or more extreme ($\leq$ or $\geq$) than the observed sample value given H_0 is true
  - Called observed level of significance
  - Smallest value of α that an H_0 can be rejected
- Compare the p-value with α
  - If p-value $\geq \alpha$, do not reject H_0
  - If p-value $< \alpha$, reject H_0

General Steps in Hypothesis Testing

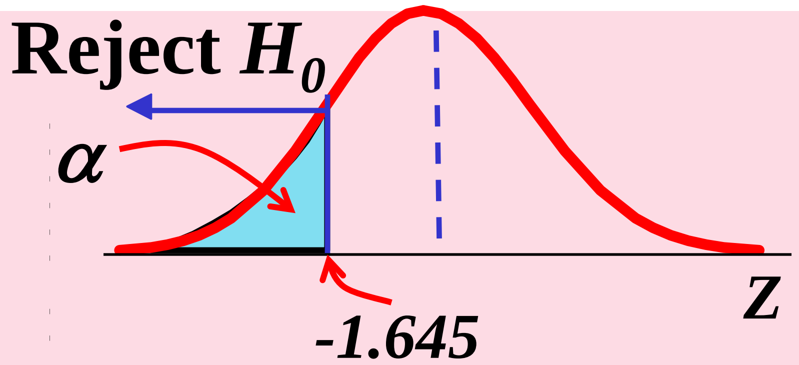
E.g., Test the Assumption that the True Mean # of TV Sets in U.S. Homes is at Least 3 (σ Known)

1. State the H_0	$H_0 : \mu \geq 3$
2. State the H_1	$H_1 : \mu < 3$
3. Choose α	$\alpha = .05$
4. Choose *n*	$n = 100$
5. Choose Test	Z *test*

General Steps in Hypothesis Testing

(continued)

6. Set up critical value(s)



7. Collect data

100 households surveyed

8. Compute test statistic and p-value

Computed test stat =-2,

p-value = .0228

9. Make statistical decision

Reject null hypothesis

10. Express conclusion

The true mean # TV set is less than 3

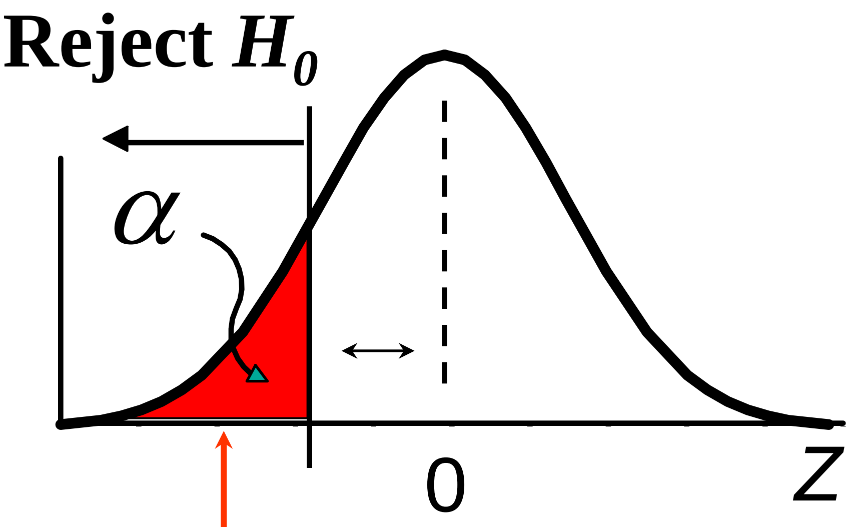
One-Tail Z Test for Mean (σ Known)

- Assumptions
  - Population is normally distributed
  - If not normal, requires large samples
  - Null hypothesis has $\leq$ or $\geq$ sign only
  - σ is known
- Z Test Statistic
  - $$Z = \frac{\bar{X} - \mu_{\bar{X}}}{\sigma_{\bar{X}}} = \frac{\bar{X} - \mu}{\sigma / \sqrt{n}}$$

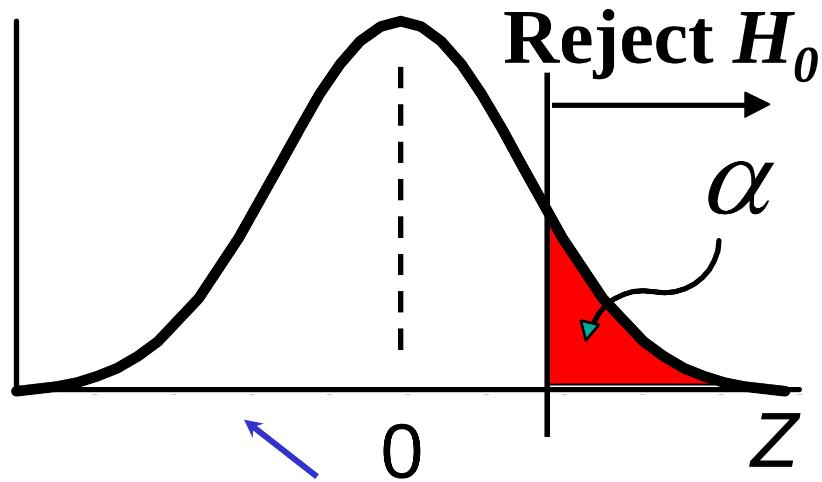
Rejection Region

$$H_0: \mu \geq \mu_0$$
$$H_1: \mu < \mu_0$$

$$H_0: \mu \leq \mu_0$$
$$H_1: \mu > \mu_0$$



Z must be *significantly* below 0 to reject H_0



Small values of Z don't contradict H_0 ; don't reject H_0 !

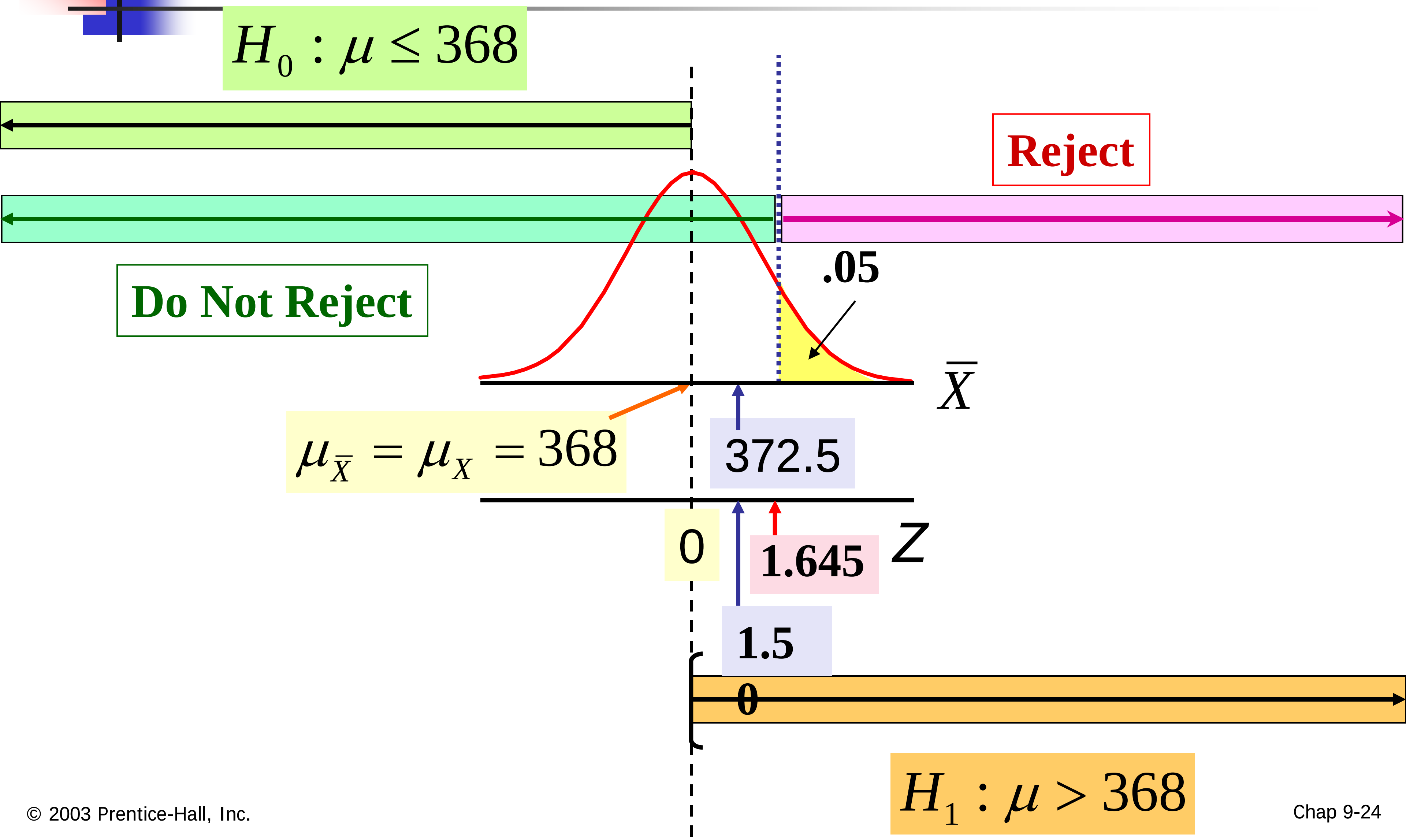
Example: One-Tail Test

Does an average box of cereal contain more than 368 grams of cereal? A random sample of 25 boxes showed $\bar{X}$ = 372.5. The company has specified σ to be 15 grams. Test at the α = 0.05 level.

H_0: $\mu \leq 368$

H_1: $\mu > 368$

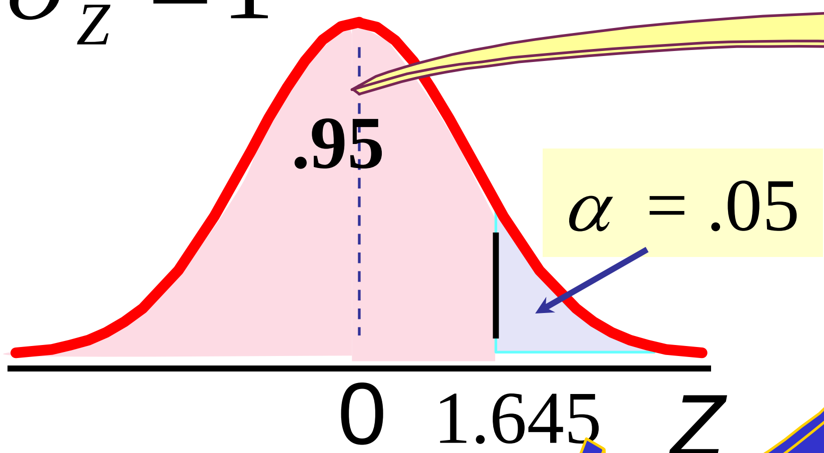
Reject and Do Not Reject Regions



Finding Critical Value: One-Tail

What is Z given $\alpha = 0.05$?

$$\sigma_Z = 1$$



Critical Value = 1.645

Standardized Cumulative Normal Distribution Table (Portion)

Z	.04	.05	.06
1.6	**.9495**	**.9505**	.9515
1.7	.9591	.9599	.9608
1.8	.9671	.9678	.9686
1.9	.9738	.9744	.9750

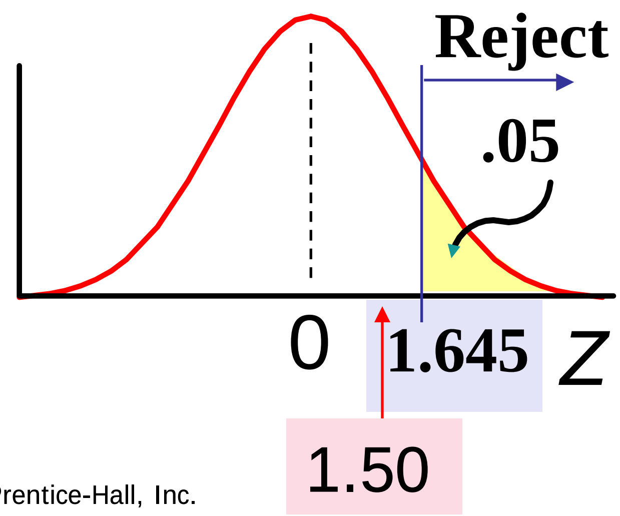
Example Solution: One-Tail Test

H_0: $\mu \le 368$
H_1: $\mu > 368$

$\alpha = 0.05$

$n = 25$

Critical Value: 1.645



Test Statistic:

$$Z = \frac{\bar{X} - \mu}{\sigma / \sqrt{n}} = 1.50$$

Decision:

Do Not Reject at α = .05.

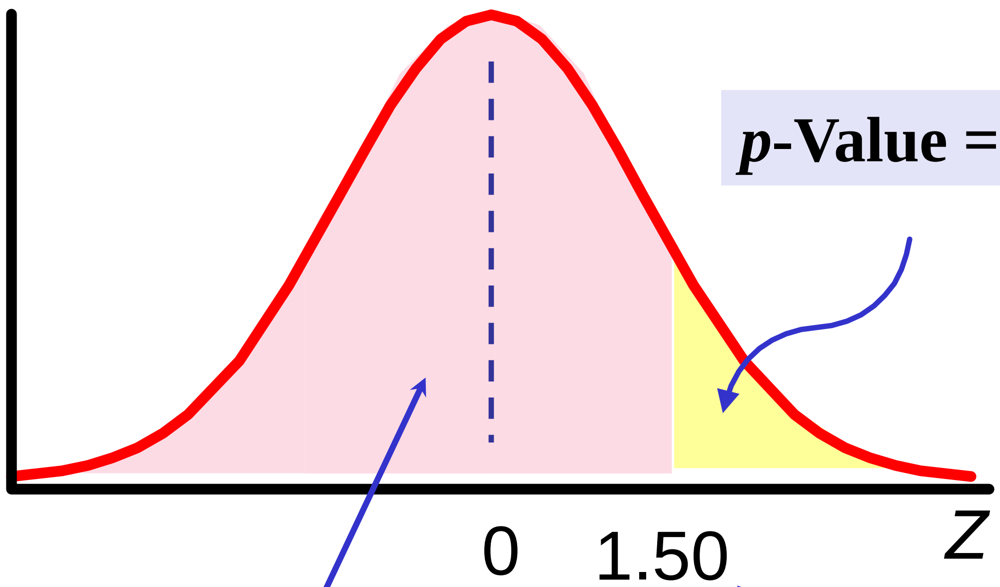
Conclusion:

Insufficient Evidence that True Mean is More Than 368.

p -Value Solution

p-Value is $P(Z \geq 1.50) = 0.0668$

Use the alternative hypothesis to find the direction of the rejection region.



p-Value =.0668

$$\begin{array}{r} 1.0000 \\ -\ .9332 \\ \hline .0668 \end{array}$$

From Z Table: Lookup 1.50 to Obtain .9332

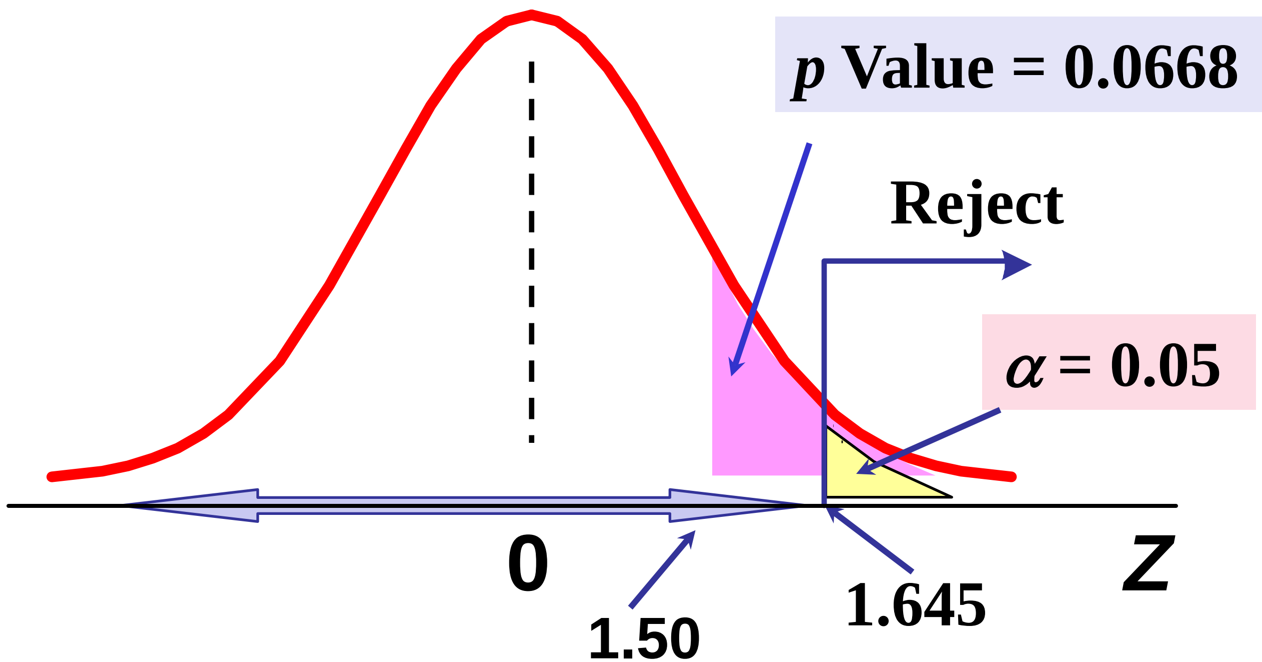
Z Value of Sample Statistic

p -Value Solution

(continued)

$(p\text{-Value} = 0.0668) \geq (\alpha = 0.05)$

Do Not Reject.



Test Statistic 1.50 is in the Do Not Reject Region

One-Tail Z Test for Mean (σ Known) in PHStat

- PHStat | One-Sample Tests | Z Test for the Mean, Sigma Known ...
- Example in Excel Spreadsheet

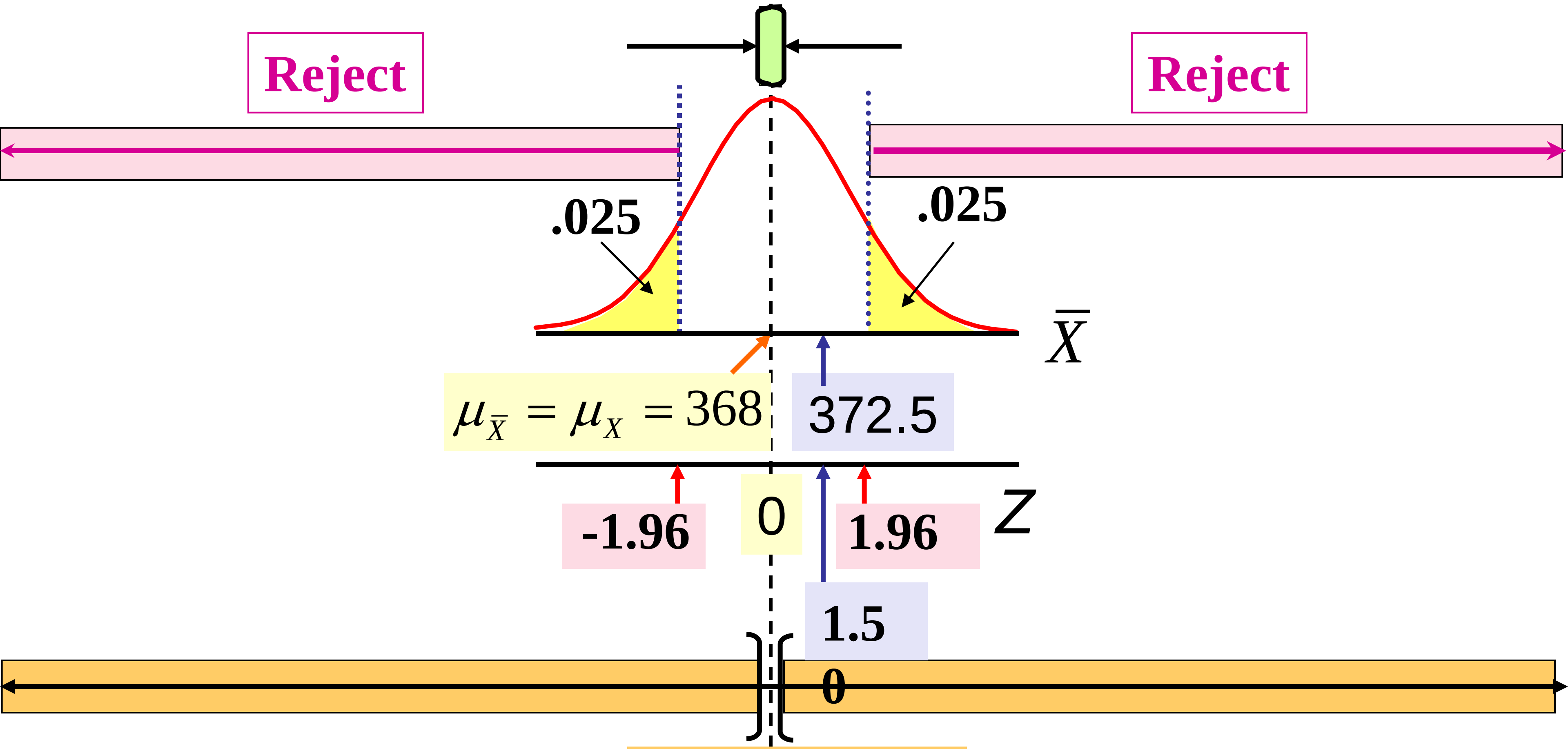
Example: Two-Tail Test

Does an average box of cereal contain 368 grams of cereal? A random sample of 25 boxes showed $\bar{X}$ = 372.5. The company has specified σ to be 15 grams and the distribution to be normal. Test at the $\alpha =$ 0.05 level.

$$H_0: \mu = 368$$
$$H_1: \mu \neq 368$$

Reject and Do Not Reject Regions

$H_0 : \mu = 368$



$H_1 : \mu \neq 368$

Example Solution: Two-Tail Test

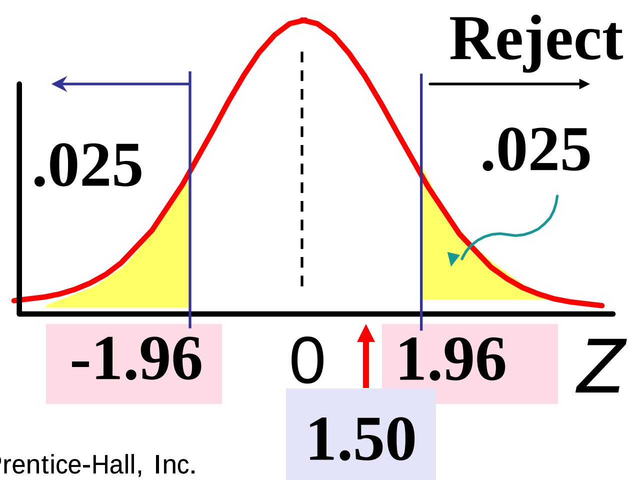
H_0: $\mu = 368$

H_1: $\mu \neq 368$

$\alpha = 0.05$

$n = 25$

Critical Value: ± 1.96



Test Statistic:

$$Z = \frac{\bar{X} - \mu}{\sigma / \sqrt{n}} = \frac{372.5 - 368}{15 / \sqrt{25}} = 1.50$$

Decision:

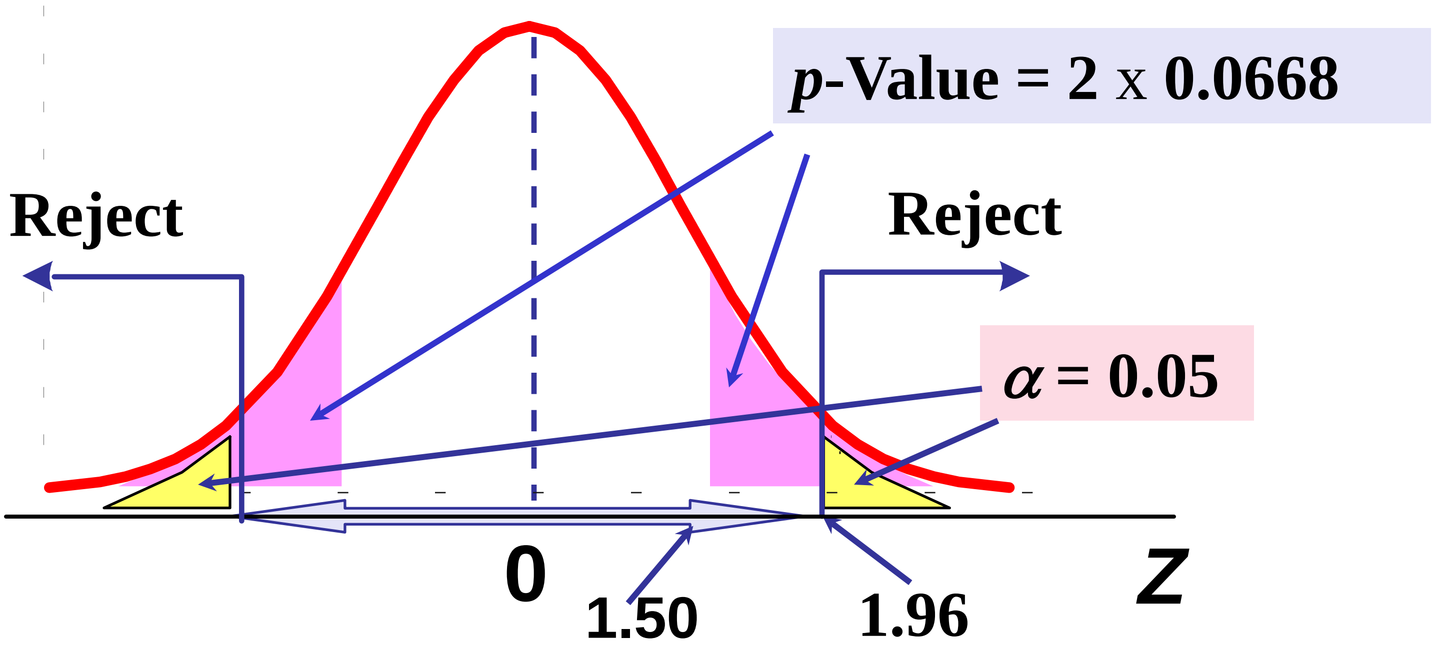
Do Not Reject at $\alpha = .05$.

Conclusion:

Insufficient Evidence that True Mean is Not 368.

p-Value Solution

(p-Value = 0.1336) $\geq$ (α = 0.05)
Do Not Reject.



Test Statistic 1.50 is in the Do Not Reject Region

Two-Tail Z Test for Mean (σ Known) in PHStat

- PHStat | One-Sample Tests | Z Test for the Mean, Sigma Known ...
- Example in Excel Spreadsheet

Connection to Confidence Intervals

For $\bar{X} = 372.5, \ \sigma = 15$ and $n = 25$,

the 95% confidence interval is:

$$372.5 - (1.96)15/\sqrt{25} \le \mu \le 372.5 + (1.96)15/\sqrt{25}$$

or

$$366.62 \le \mu \le 378.38$$

We are 95% confident that the population mean is between 366.62 and 378.38.

If this interval contains the hypothesized mean (368), we do not reject the null hypothesis.

It does. Do not reject.

t Test: σ Unknown

- Assumption
  - Population is normally distributed
  - If not normal, requires a large sample
  - σ is unknown
- *t* Test Statistic with *n-1* Degrees of Freedom
  - $$t = \frac{\bar{X} - \mu}{S / \sqrt{n}}$$

Example: One-Tail t Test

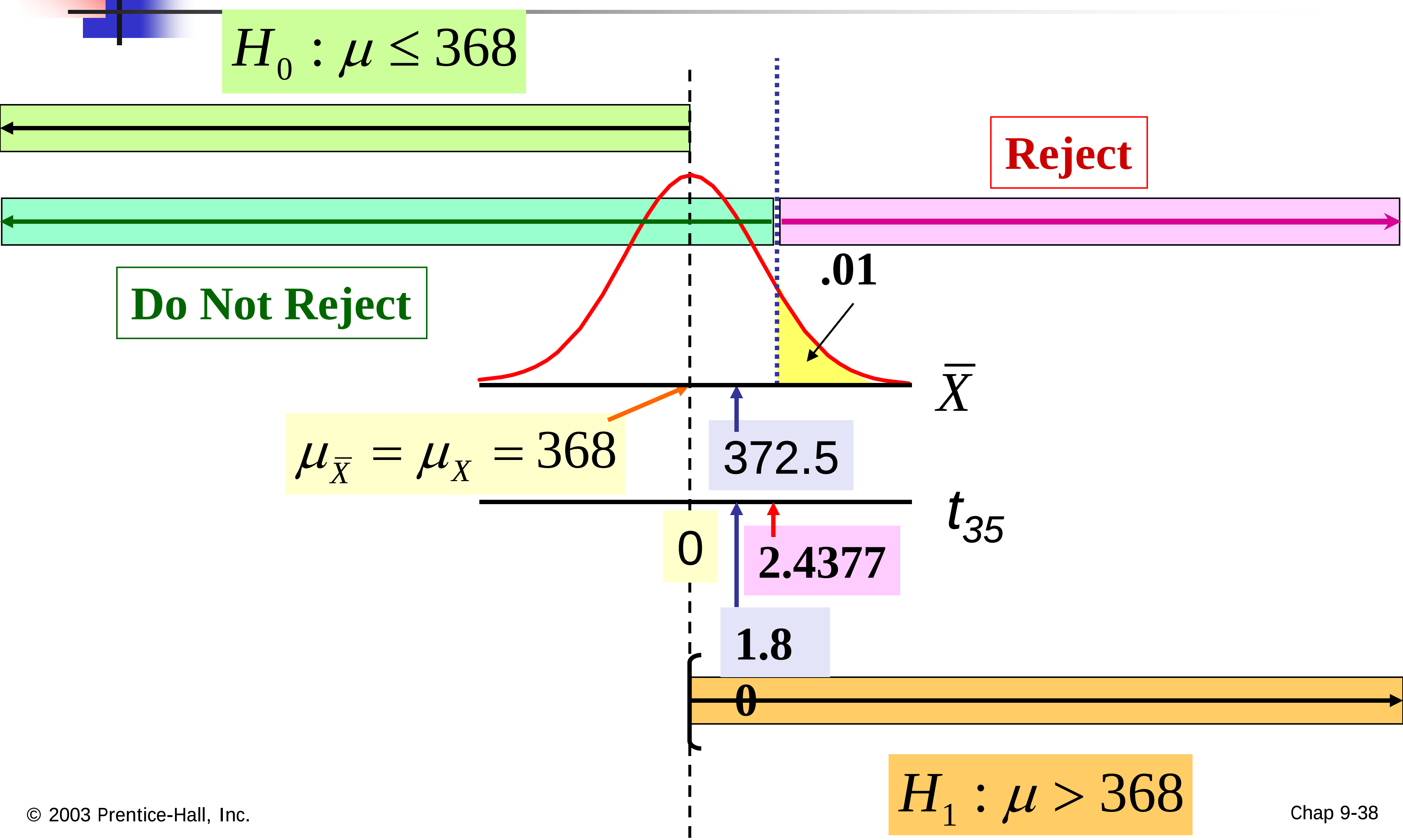
Does an average box of cereal contain more than 368 grams of cereal? A random sample of 36 boxes showed $\bar{X} = 372.5$, and $s = 15$. Test at the $\alpha = 0.01$ level.

σ is not given

$H_0: \mu \le 368$

$H_1: \mu > 368$

Reject and Do Not Reject Regions



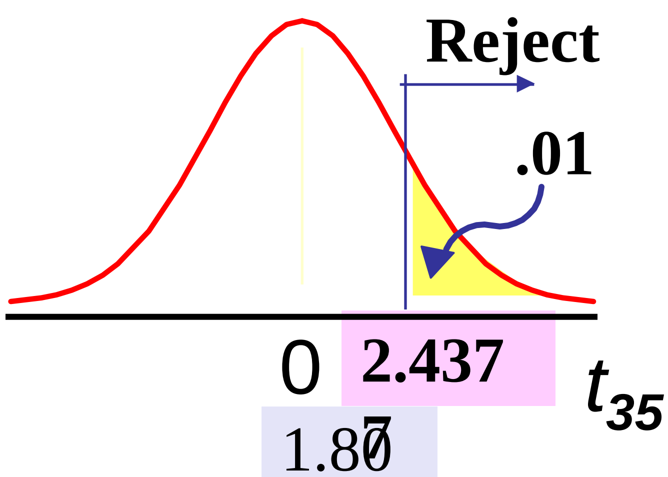
Example Solution: One-Tail

$H_0: \mu \leq 368$
$H_1: \mu > 368$

$\alpha = 0.01$

$n = 36$, df $= 35$

Critical Value: 2.4377



Test Statistic:

$$t = \frac{\bar{X} - \mu}{S / \sqrt{n}} = \frac{372.5 - 368}{15 / \sqrt{36}} = 1.80$$

Decision:

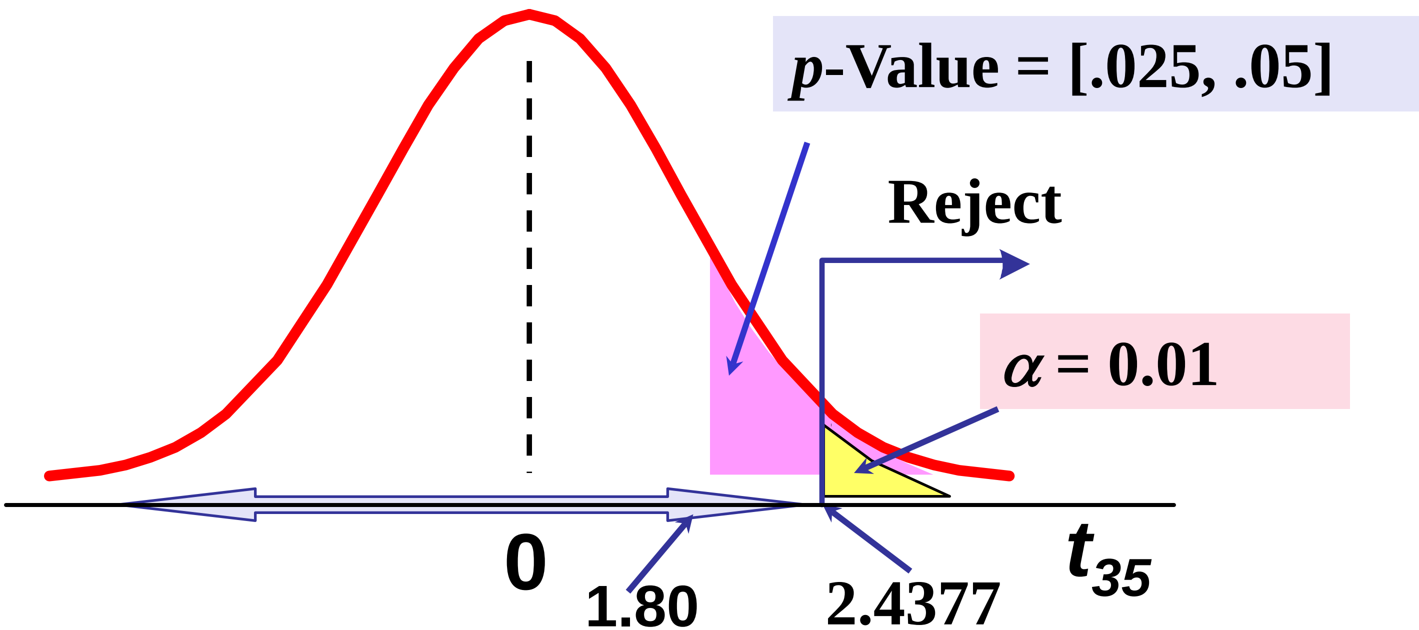
Do Not Reject at $a = .01$.

Conclusion:

Insufficient Evidence that True Mean is More Than 368.

p -Value Solution

(p-Value is between .025 and .05) $\geq$ (α = 0.01)
Do Not Reject.



Test Statistic 1.80 is in the Do Not Reject Region

t Test: σ Unknown in PHStat

- PHStat | One-Sample Tests | *t* Test for the Mean, Sigma Known ...
- Example in Excel Spreadsheet

Proportion

- Involves Categorical Variables
- Two Possible Outcomes
  - "Success" (possesses a certain characteristic) and "Failure" (does not possess a certain characteristic)
- Fraction or Proportion of Population in the "Success" Category is Denoted by p

Proportion

(continued)

- Sample Proportion in the Success Category is Denoted by p_S
  - $$p_s = \frac{X}{n} = \frac{\text{Number of Successes}}{\text{Sample Size}}$$
- When Both np and $n(1-p)$ are at Least 5, p_S Can Be Approximated by a Normal Distribution with Mean and Standard Deviation
  - $$\mu_{p_s} = p \qquad \sigma_{p_s} = \sqrt{\frac{p(1-p)}{n}}$$

Example: *Z* Test for Proportion

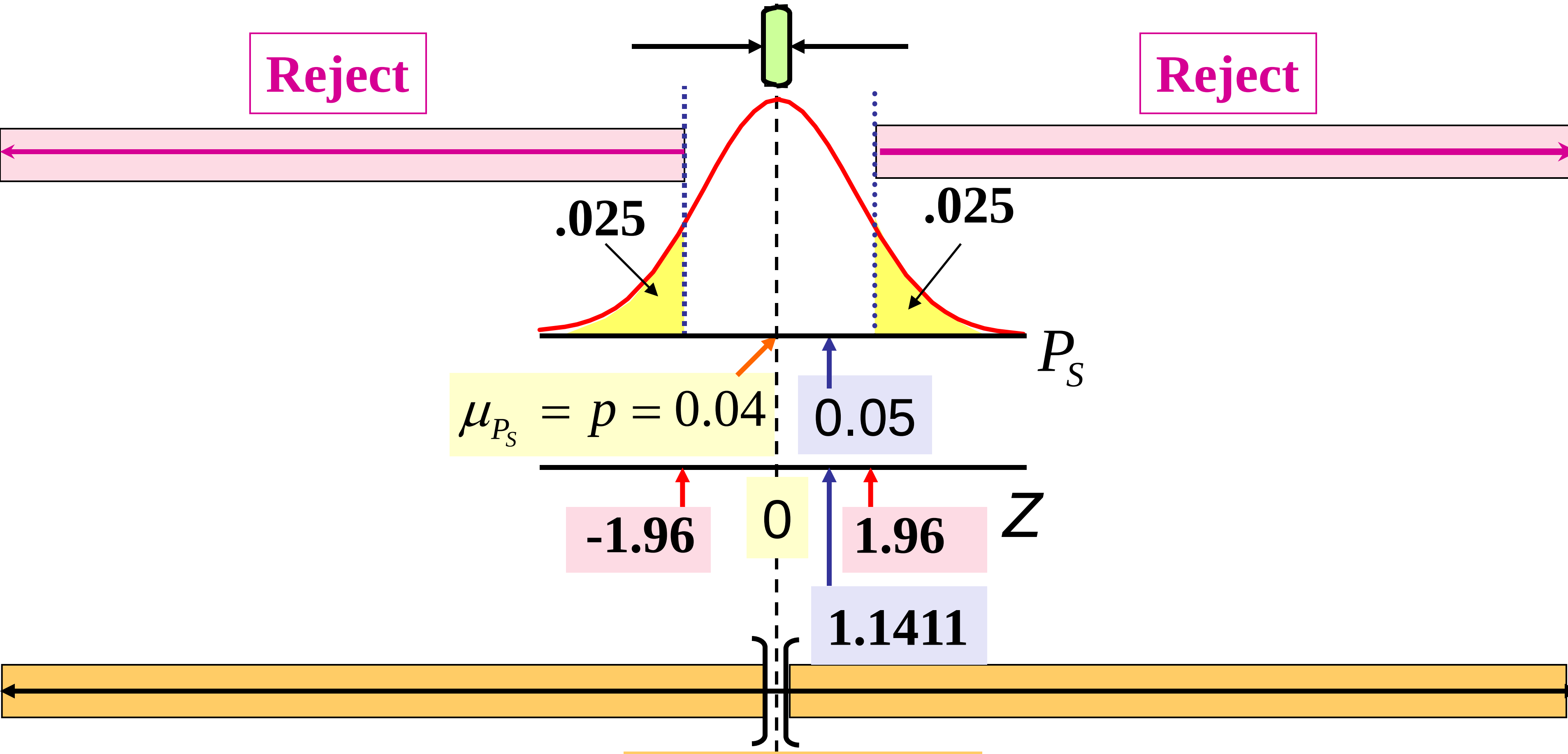
A marketing company claims that a survey will have a 4% response rate. To test this claim, a random sample of 500 were surveyed with 25 responses. Test at the α = .05 significance level.

Check:

$$np = 500(.04) = 20$$
$$\geq 5$$
$$n(1-p) = 500(1-.04)$$
$$= 480 \geq 5$$

Reject and Do Not Reject Regions

$H_0 : p = 0.04$



$H_1 : p \neq 0.04$

Z Test for Proportion: Solution

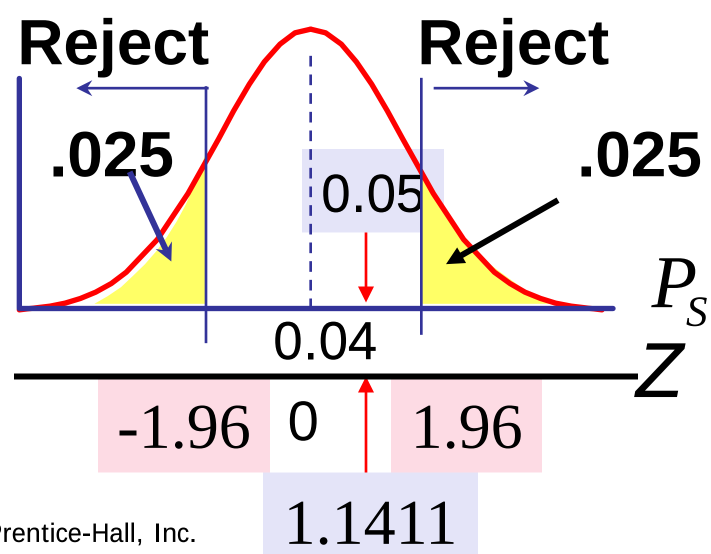
$H_0: p = .04$

$H_1: p \neq .04$

$\alpha = .05$

$n = 500$

Critical Values: ± 1.96



Test Statistic:

$$Z \cong \frac{p_S - p}{\sqrt{\frac{p(1-p)}{n}}} = \frac{.05 - .04}{\sqrt{\frac{.04(1-.04)}{500}}} = 1.1411$$

Decision:

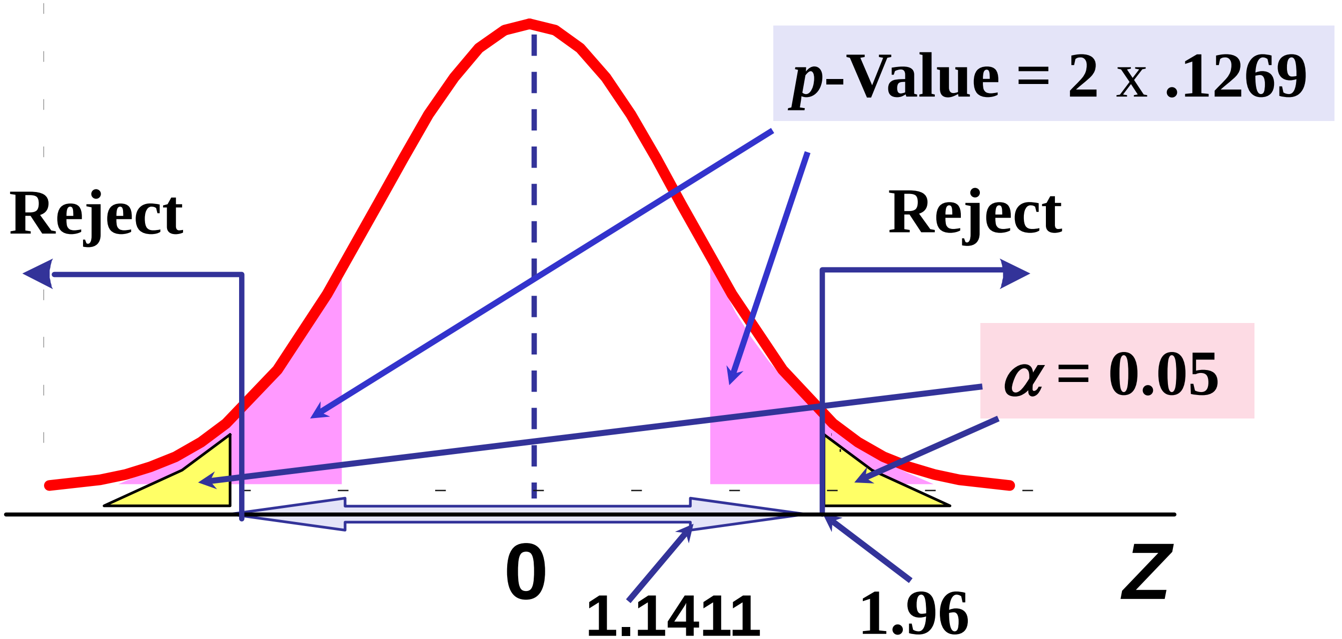
Do not reject at $\alpha = .05$.

Conclusion:

We do not have sufficient evidence to reject the company's claim of 4% response rate.

p -Value Solution

(p-Value = 0.2538) ≥ (α = 0.05)
Do Not Reject.



Test Statistic 1.1411 is in the Do Not Reject Region

Z Test for Proportion in PHStat

- PHStat | One-Sample Tests | Z Test for the Proportion ...
- Example in Excel Spreadsheet

Potential Pitfalls and Ethical Issues

- Data Collection Method is Not Randomized to Reduce Selection Biases
- Treatment of Human Subjects are Manipulated Without Informed Consent
- Data Snooping is Used to Choose between One-Tail and Two-Tail Tests, and to Determine the Level of Significance

Potential Pitfalls and Ethical Issues

(continued)

- Data Cleansing is Practiced to Hide Observations that do not Support a Stated Hypothesis
- Fail to Report Pertinent Findings

Chapter Summary

- Addressed Hypothesis Testing Methodology
- Performed Z Test for the Mean (σ Known)
- Discussed p–Value Approach to Hypothesis Testing
- Made Connection to Confidence Interval Estimation

Chapter Summary

(continued)

- Performed One-Tail and Two-Tail Tests
- Performed t Test for the Mean (σ Unknown)
- Performed Z Test for the Proportion
- Discussed Potential Pitfalls and Ethical Issues