

# Fundamentals of Hypothesis Testing: One-Sample Tests

Prof. dr. Siswanto Agus Wilopo, M.Sc., Sc.D.  
Department of Biostatistics, Epidemiology  
and Population Health  
Faculty of Medicine  
Universitas Gadjah Mada

# Learning Objectives

## In this lecture, you learn:

- The basic principles of hypothesis testing
- How to use hypothesis testing to test a mean or proportion
- The assumptions of each hypothesis-testing procedure, how to evaluate them, and the consequences if they are seriously violated
- How to avoid the pitfalls involved in hypothesis testing
- The ethical issues involved in hypothesis testing

# What is a Hypothesis?

- A hypothesis is a claim (assumption) about a population parameter:
  - population mean

**Example: The mean of systolic blood pressure among adults of this city is  $\mu = 120$  mmHg**

- population proportion

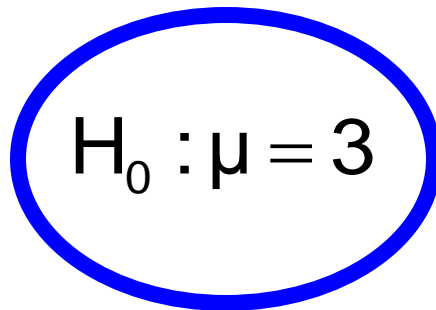
**Example: The proportion of adults in this city with hypertension is  $\pi = 0.118$**

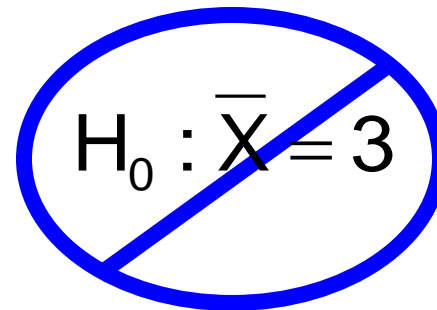
# The Null Hypothesis, $H_0$

- States the claim or assertion to be tested

**Example:** The average number of children in NTT woman is equal to three ( $H_0 : \mu = 3$ )

- Is always about a population parameter, not about a sample statistic

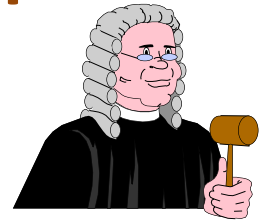

$$H_0 : \mu = 3$$


$$H_0 : \bar{X} = 3$$

# 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
- Refers to the status quo
- Always contains “=”, “ $\leq$ ” or “ $\geq$ ” sign
- May or may not be rejected

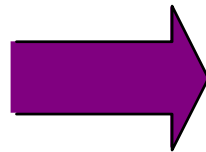


# The Alternative Hypothesis, $H_1$

- Is the opposite of the null hypothesis
  - e.g., The average number of children in NTT woman is not equal to 3 (  $H_1: \mu \neq 3$  )
- Challenges the status quo
- Contains the “=”, “ $\leq$ ” or “ $\geq$ ” sign
- May or may not be proven
- Is generally the hypothesis that the researcher is trying to prove

# Hypothesis Testing Process

**Claim:** the population mean age is 50.  
(Null Hypothesis:  
 $H_0: \mu = 50$ )

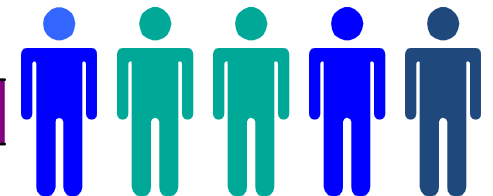


**Population**



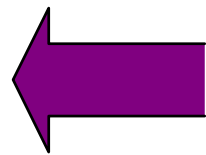
Now select a random sample

Is  $\bar{X}=20$  likely if  $\mu = 50$ ?



**Sample**

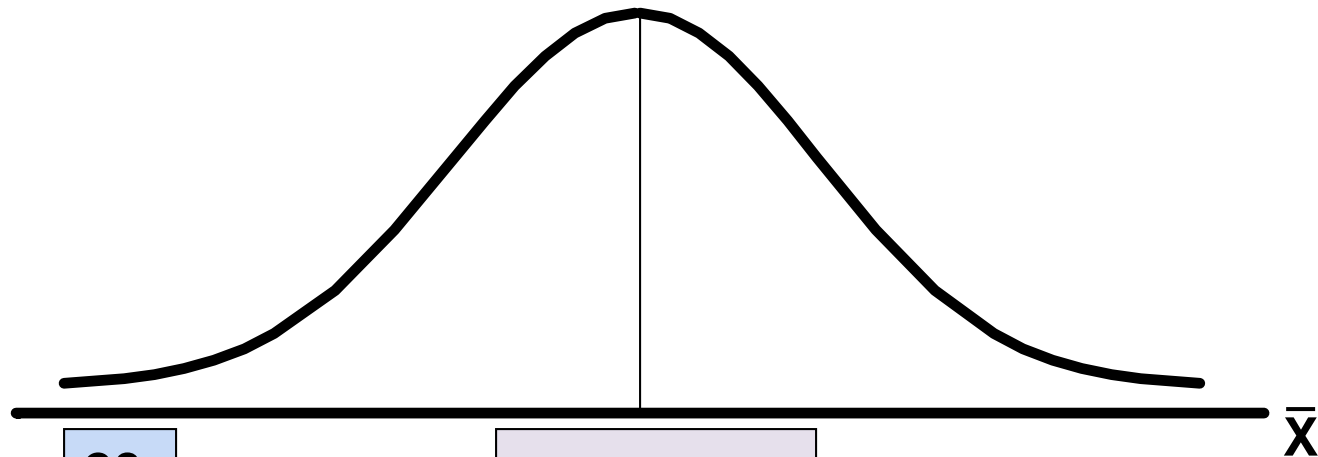
Suppose the sample mean age is 20:  $\bar{X} = 20$



If not likely,  
**REJECT**  
Null Hypothesis

# Reason for Rejecting $H_0$

Sampling Distribution of  $\bar{X}$



20

$\mu = 50$   
If  $H_0$  is true

$\bar{X}$

If it is unlikely that we would get a sample mean of this value ...

... if in fact this were the population mean...

... then we reject the null hypothesis that  $\mu = 50$ .



# Level of Significance, $\alpha$

- Defines the unlikely values of the sample statistic if the null hypothesis is true
  - Defines rejection region of the sampling distribution
- Is designated by  $\alpha$ , (level of significance)
  - Typical values are 0.01, 0.05, or 0.10
- Is selected by the researcher at the beginning
- Provides the critical value(s) of the test

# Level of Significance and the Rejection Region

Level of significance =  $\alpha$

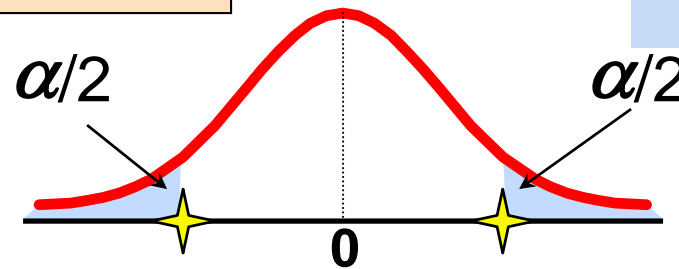
✦ Represents critical value

Rejection region is shaded

$$H_0: \mu = 3$$

$$H_1: \mu \neq 3$$

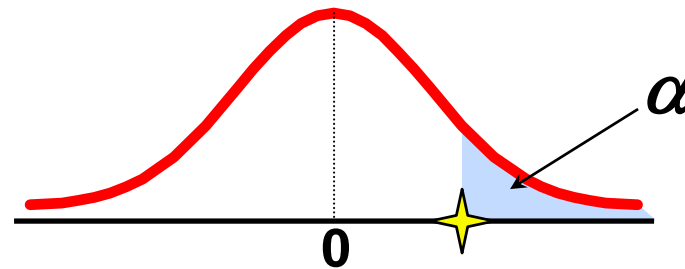
Two-tail test



$$H_0: \mu \leq 3$$

$$H_1: \mu > 3$$

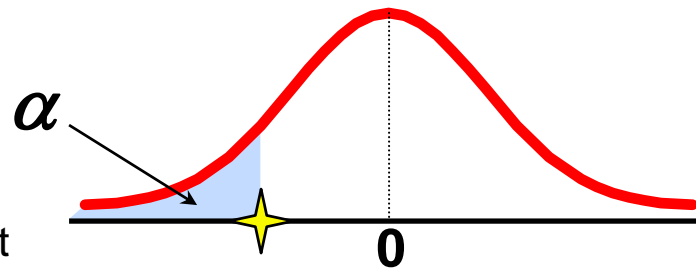
Upper-tail test



$$H_0: \mu \geq 3$$

$$H_1: \mu < 3$$

Lower-tail test



# Errors in Making Decisions

- **Type I Error**

- Reject a true null hypothesis
- Considered a serious type of error

The probability of Type I Error is  $\alpha$

- Called **level of significance** of the test
- Set by the researcher in advance

# Errors in Making Decisions

*(continued)*

- **Type II Error**
  - **Fail to reject a false null hypothesis**

The probability of Type II Error is  $\beta$

# Outcomes and Probabilities

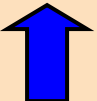

## Possible Hypothesis Test Outcomes

	Actual situation	
Decision	$H_0$ True	$H_0$ False
Do Not Reject $H_0$	No error ( $1 - \alpha$ )	Type II Error ( $\beta$ )
Reject $H_0$	Type I Error ( $\alpha$ )	No Error ( $1 - \beta$ )

**Key:**  
**Outcome**  
**(Probability)**

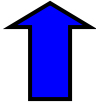

# Type I & II Error Relationship

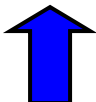

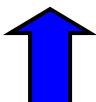



- Type I and Type II errors cannot happen at the same time
  - Type I error can only occur if  $H_0$  is true
  - Type II error can only occur if  $H_0$  is false

If Type I error probability (  $\alpha$  ) , then  
Type II error probability (  $\beta$  ) 

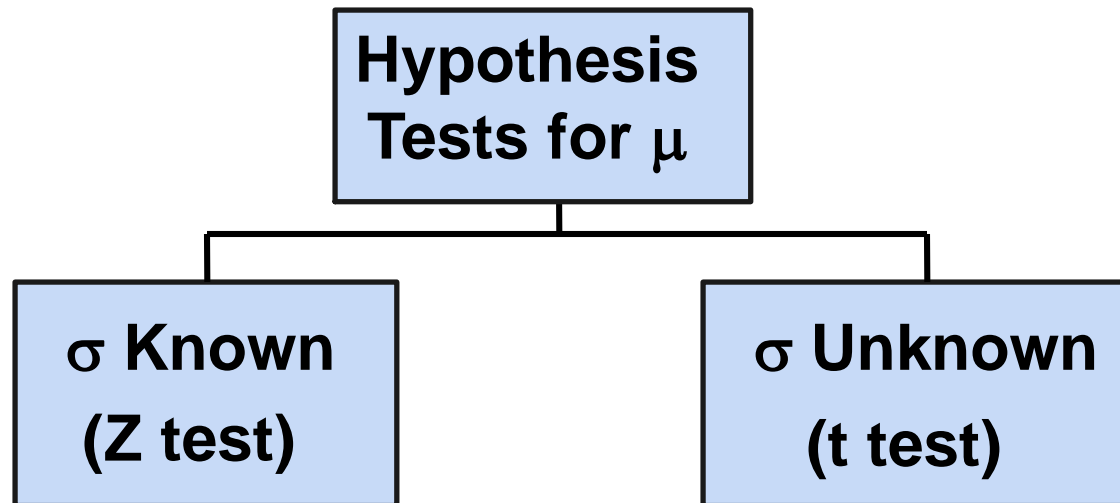
# Factors Affecting Type II Error

- All else equal,

- $\beta$   when the difference between hypothesized parameter and its true value 

- $\beta$   when  $\alpha$  
- $\beta$   when  $\sigma$  
- $\beta$   when  $n$  

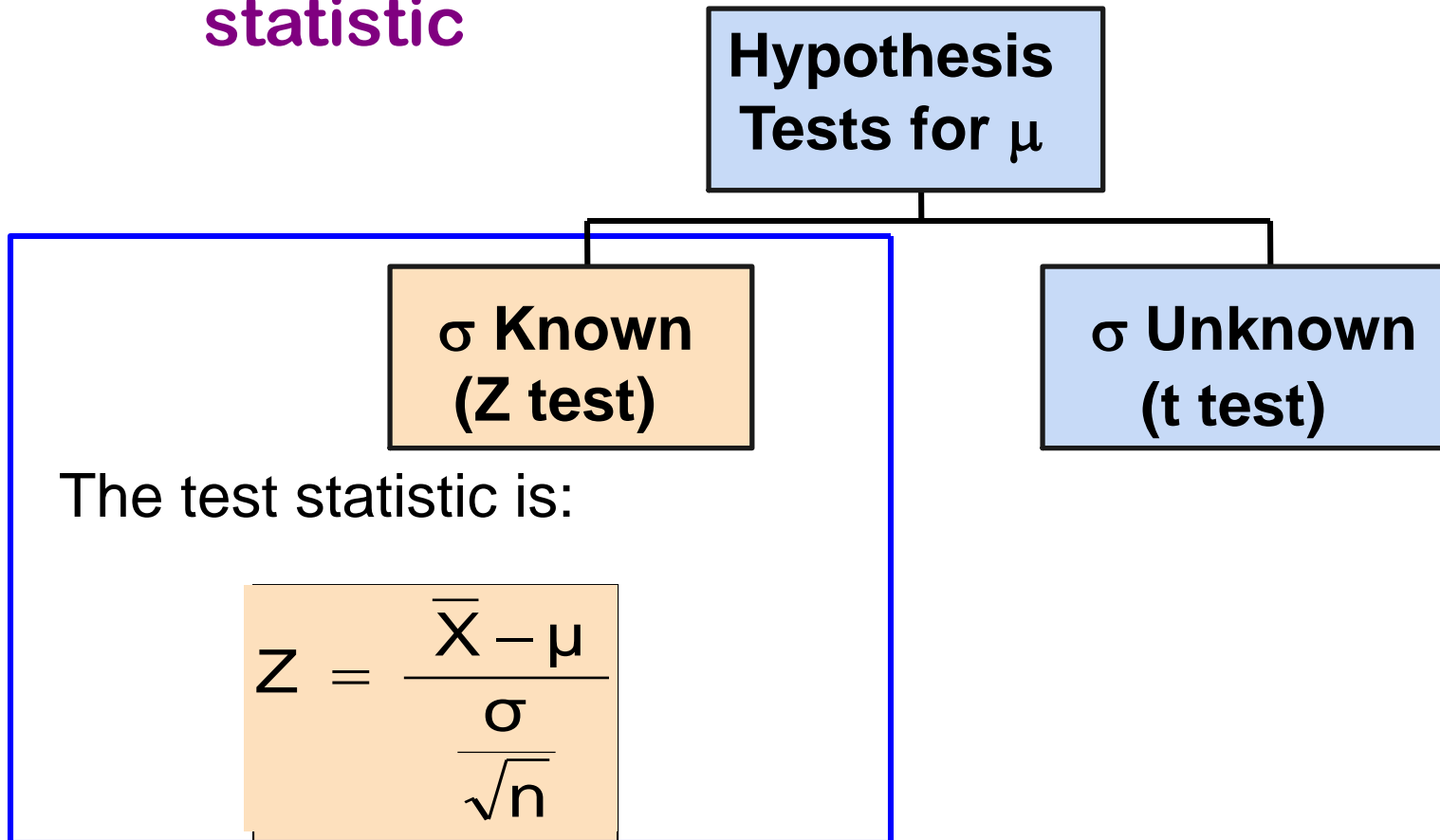
# Hypothesis Tests for the Mean





# Z Test of Hypothesis for the Mean ( $\sigma$ Known)

- Convert sample statistic ( $\bar{X}$ ) to a Z test statistic

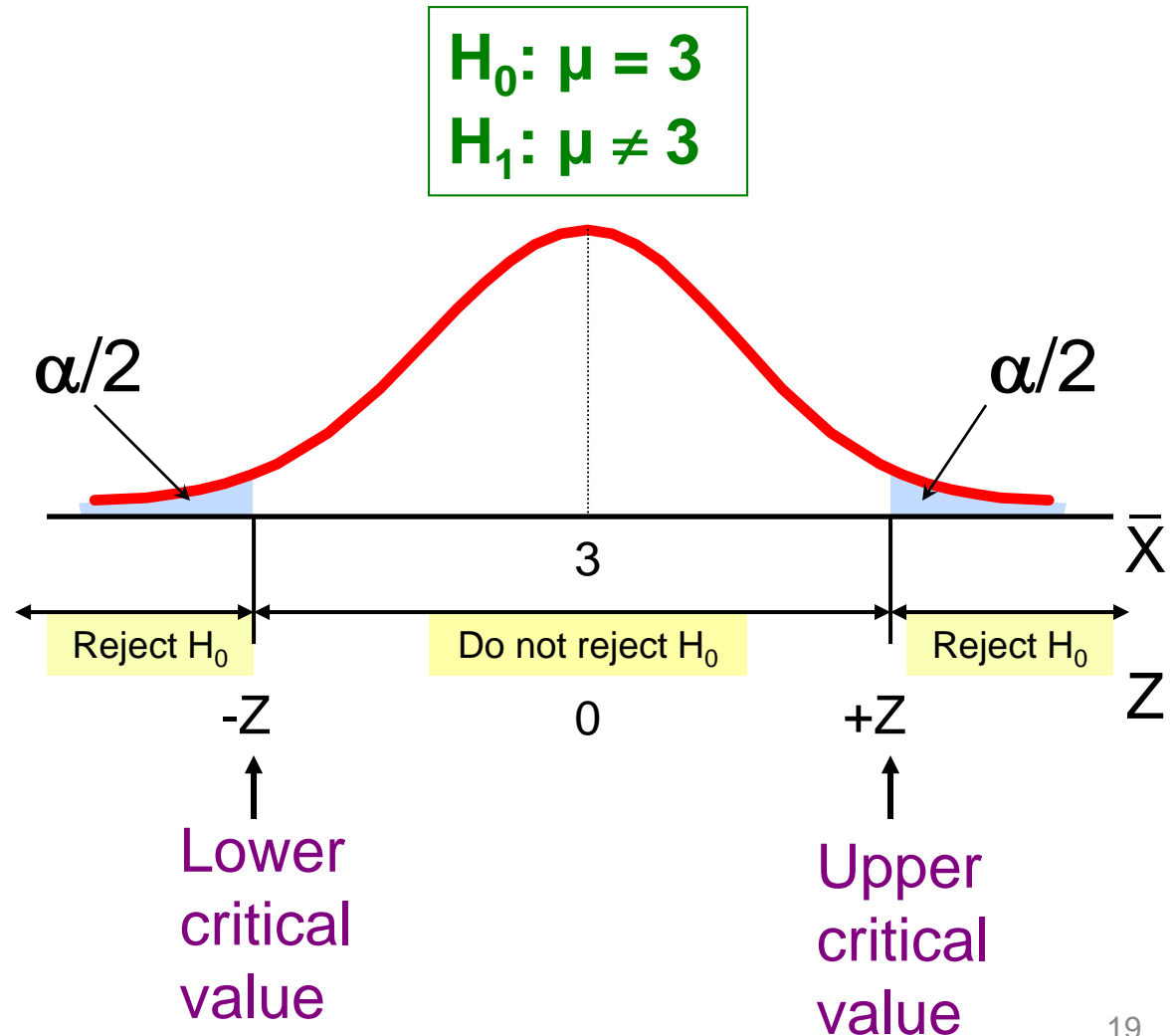


# Critical Value Approach to Testing

- For a two-tail test for the mean,  $\sigma$  known:
- Convert sample statistic ( $\bar{X}$ ) to test statistic (Z statistic)
- Determine the critical Z values for a specified level of significance  $\alpha$  from a table or computer
- **Decision Rule:** If the test statistic falls in the rejection region, reject  $H_0$ ; otherwise do not reject  $H_0$

# Two-Tail Tests

- There are two cutoff values (critical values), defining the regions of rejection



# 6 Steps in Hypothesis Testing

1. State the null hypothesis,  $H_0$  and the alternative hypothesis,  $H_1$
2. Choose the level of significance,  $\alpha$ , and the sample size,  $n$
3. Determine the appropriate test statistic and sampling distribution
4. Determine the critical values that divide the rejection and nonrejection regions

# 6 Steps in Hypothesis Testing

*(continued)*

5. Collect data and compute the value of the test statistic
6. Make the statistical decision and state the managerial conclusion.
  - If the test statistic falls into the non-rejection region, do not reject the null hypothesis  $H_0$ .
  - If the test statistic falls into the rejection region, reject the null hypothesis.
  - Express the managerial conclusion in the context of the problem

# Hypothesis Testing Example

Test the claim that the true mean of children in NTT woman is # equal to 3.  
(Assume  $\sigma = 0.8$ )

1. State the appropriate null and alternative hypotheses
  - $H_0: \mu = 3$      $H_1: \mu \neq 3$  (This is a two-tail test)
2. Specify the desired level of significance and the sample size
  - Suppose that  $\alpha = 0.05$  and  $n = 100$  are chosen for this test

# Hypothesis Testing Example

*(continued)*

3. Determine the appropriate technique
  - $\sigma$  is known so this is a Z test.
4. Determine the critical values
  - For  $\alpha = 0.05$  the critical Z values are  $\pm 1.96$
5. Collect the data and compute the test statistic
  - Suppose the sample results are  
 $n = 100$ ,  $\bar{X} = 2.84$  ( $\sigma = 0.8$  is assumed known)

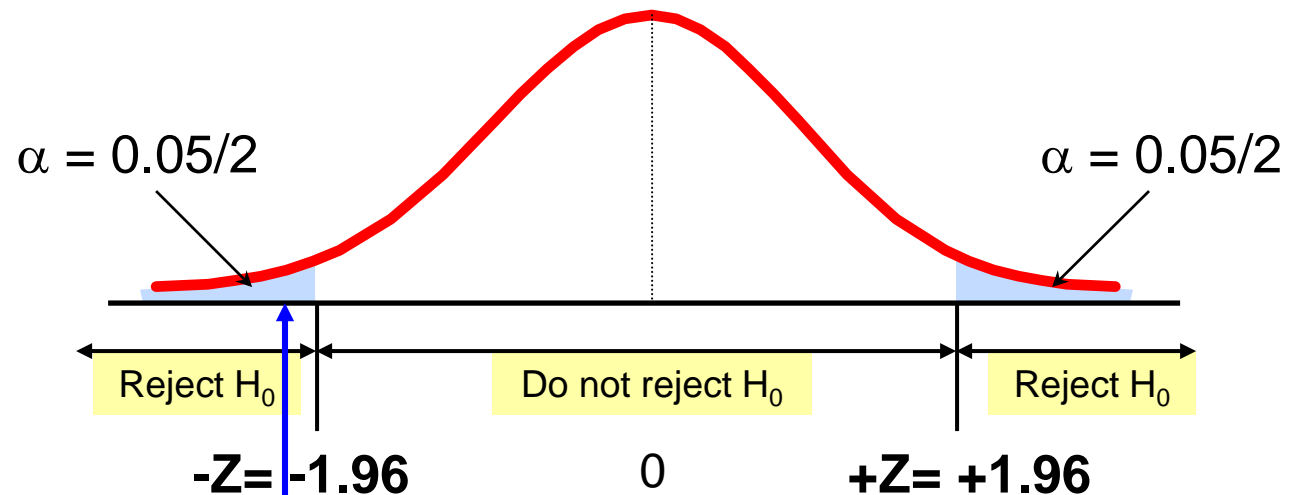
So the test statistic is:

$$Z = \frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}} = \frac{2.84 - 3}{\frac{0.8}{\sqrt{100}}} = \frac{-.16}{.08} = -2.0$$

# Hypothesis Testing Example

(continued)

## 6. Is the test statistic in the rejection region?



Reject  $H_0$  if  
 $Z < -1.96$  or  
 $Z > 1.96$ ;  
otherwise  
do not  
reject  $H_0$

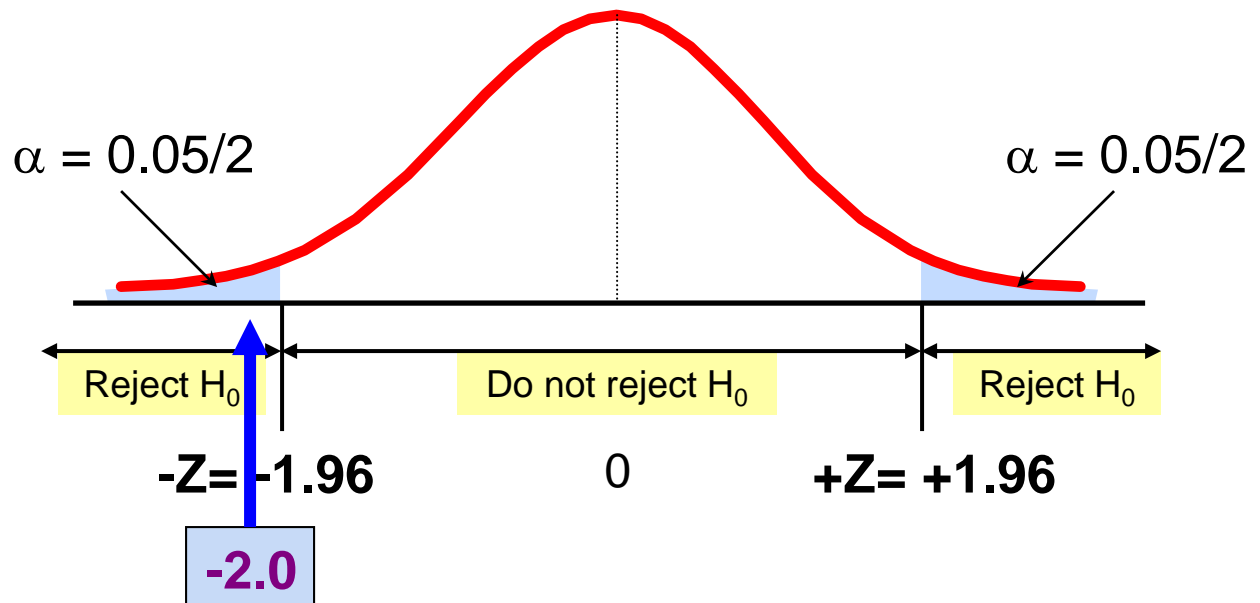
Here,  $Z = -2.0 < -1.96$ , so the  
test statistic is in the rejection  
region



# Hypothesis Testing Example

(continued)

6(continued). Reach a decision and interpret the result



Since  $Z = -2.0 < -1.96$ , we reject the null hypothesis and conclude that there is sufficient evidence that the mean number of children of NTT woman is not equal to 3

# p-Value Approach to Testing

- p-value: Probability of obtaining a test statistic more extreme ( $\leq$  or  $\geq$ ) than the observed sample value given  $H_0$  is true
  - Also called observed level of significance
  - Smallest value of  $\alpha$  for which  $H_0$  can be rejected

# p-Value Approach to Testing

- Convert Sample Statistic (e.g.,  $\bar{X}$ ) to Test Statistic (e.g., Z statistic ) (continued)
- Obtain the p-value from a table or computer
- Compare the p-value with  $\alpha$ 
  - If p-value  $< \alpha$  , reject  $H_0$
  - If p-value  $\geq \alpha$  , do not reject  $H_0$

# p-Value Example

- Example:** How likely is it to see a sample mean of 2.84 (or something further from the mean, in either direction) if the true mean is  $\mu = 3.0$ ?

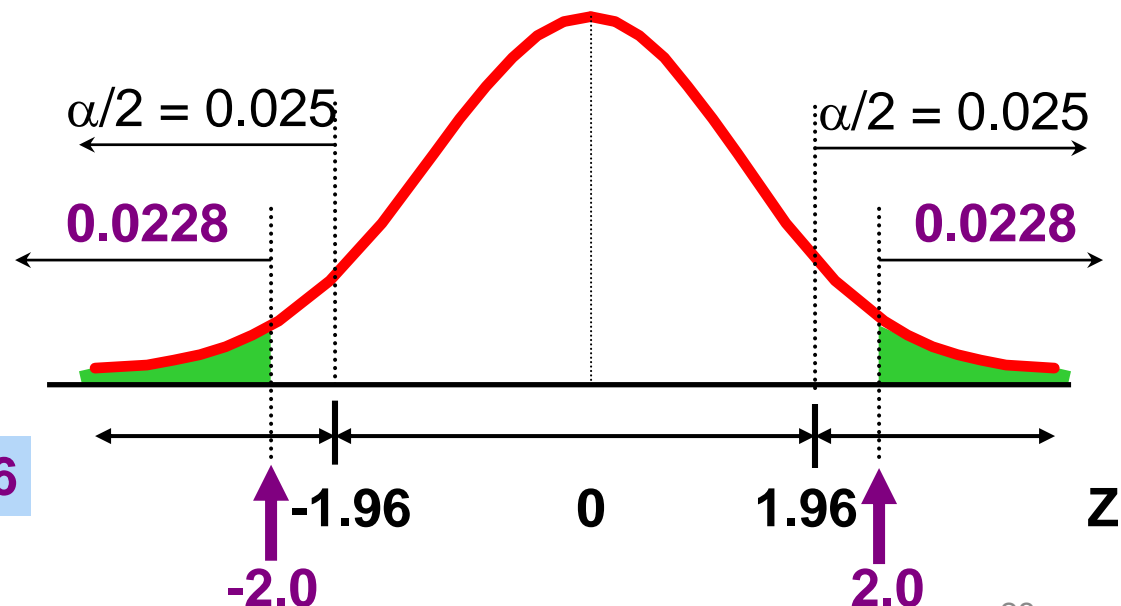
$\bar{X} = 2.84$  is translated to a Z score of  $Z = -2.0$

$$P(Z < -2.0) = 0.0228$$

$$P(Z > 2.0) = 0.0228$$

**p-value**

$$= 0.0228 + 0.0228 = 0.0456$$



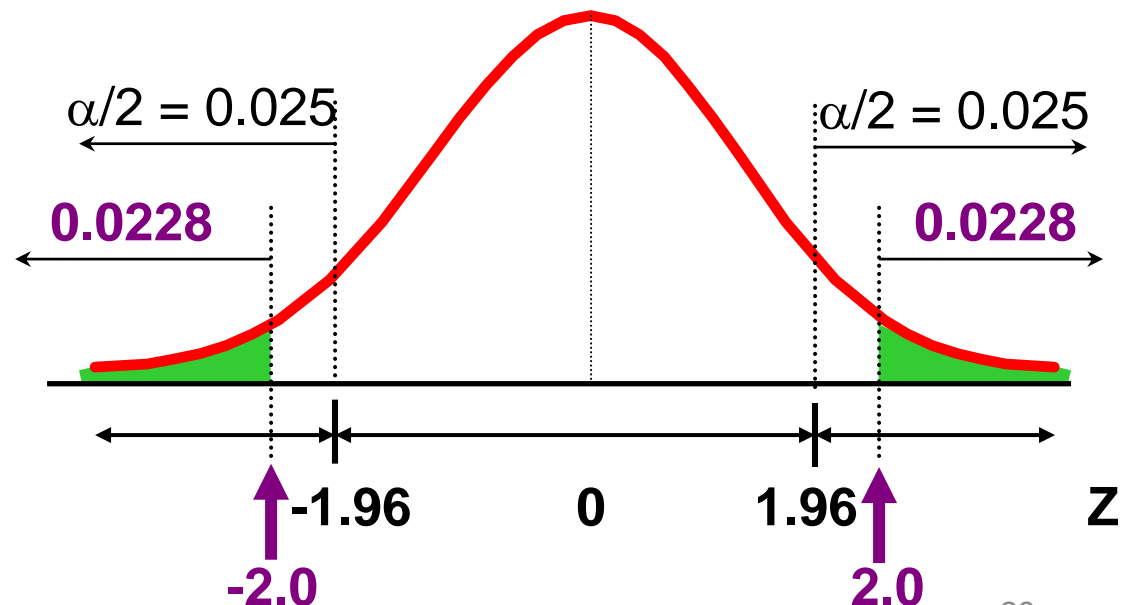
# p-Value Example

(continued)

- Compare the p-value with  $\alpha$ 
  - If p-value  $< \alpha$ , reject  $H_0$
  - If p-value  $\geq \alpha$ , do not reject  $H_0$

Here: p-value = 0.0456  
 $\alpha = 0.05$

Since  $0.0456 < 0.05$ ,  
we reject the null  
hypothesis



# Connection to Confidence Intervals

- For  $\bar{X} = 2.84$ ,  $\sigma = 0.8$  and  $n = 100$ , the 95% confidence interval is:

$$2.84 - (1.96) \frac{0.8}{\sqrt{100}} \quad \text{to} \quad 2.84 + (1.96) \frac{0.8}{\sqrt{100}}$$

$$2.6832 \leq \mu \leq 2.9968$$

- Since this interval does not contain the hypothesized mean (3.0), we reject the null hypothesis at  $\alpha = 0.05$

# One-Tail Tests

- In many cases, the alternative hypothesis focuses on a particular direction

$$\begin{array}{l} H_0: \mu \geq 3 \\ H_1: \mu < 3 \end{array}$$



This is a **lower**-tail test since the alternative hypothesis is focused on the lower tail below the mean of 3

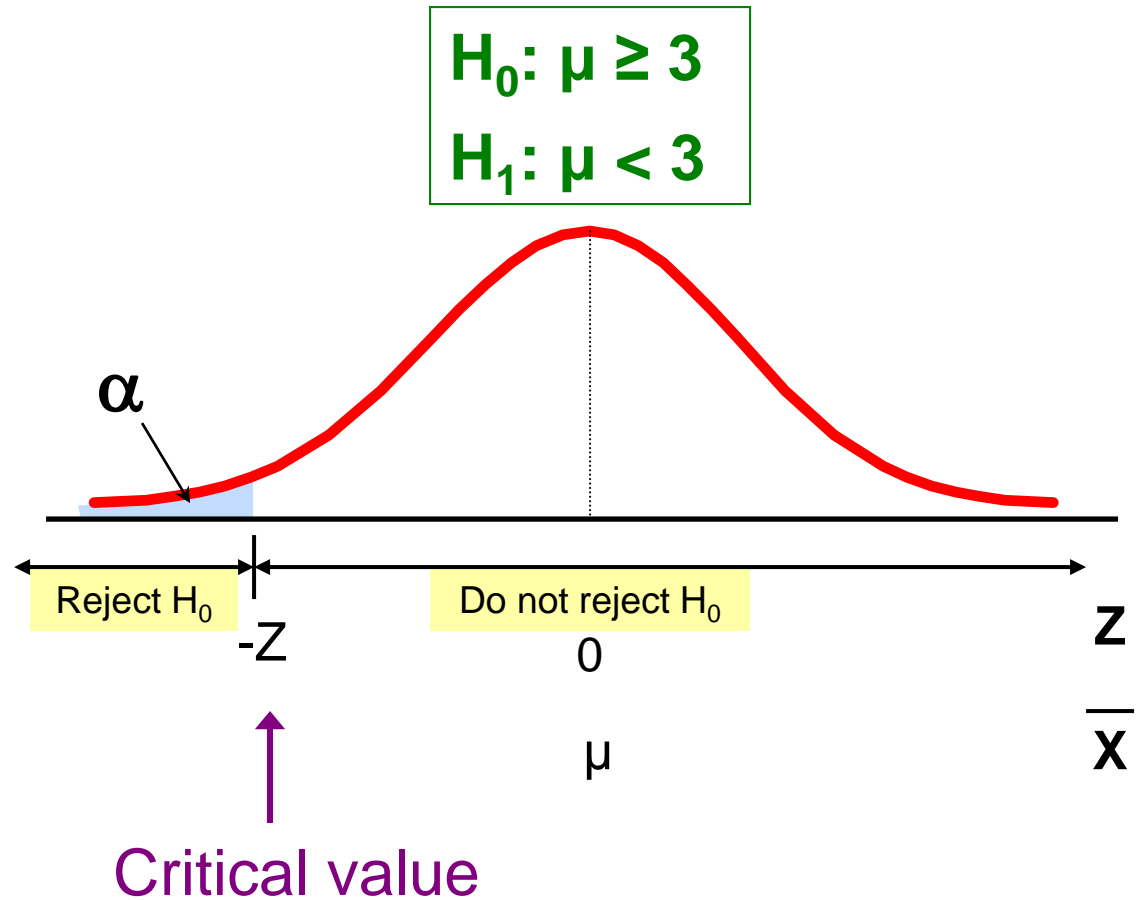
$$\begin{array}{l} H_0: \mu \leq 3 \\ H_1: \mu > 3 \end{array}$$



This is an **upper**-tail test since the alternative hypothesis is focused on the upper tail above the mean of 3

# Lower-Tail Tests

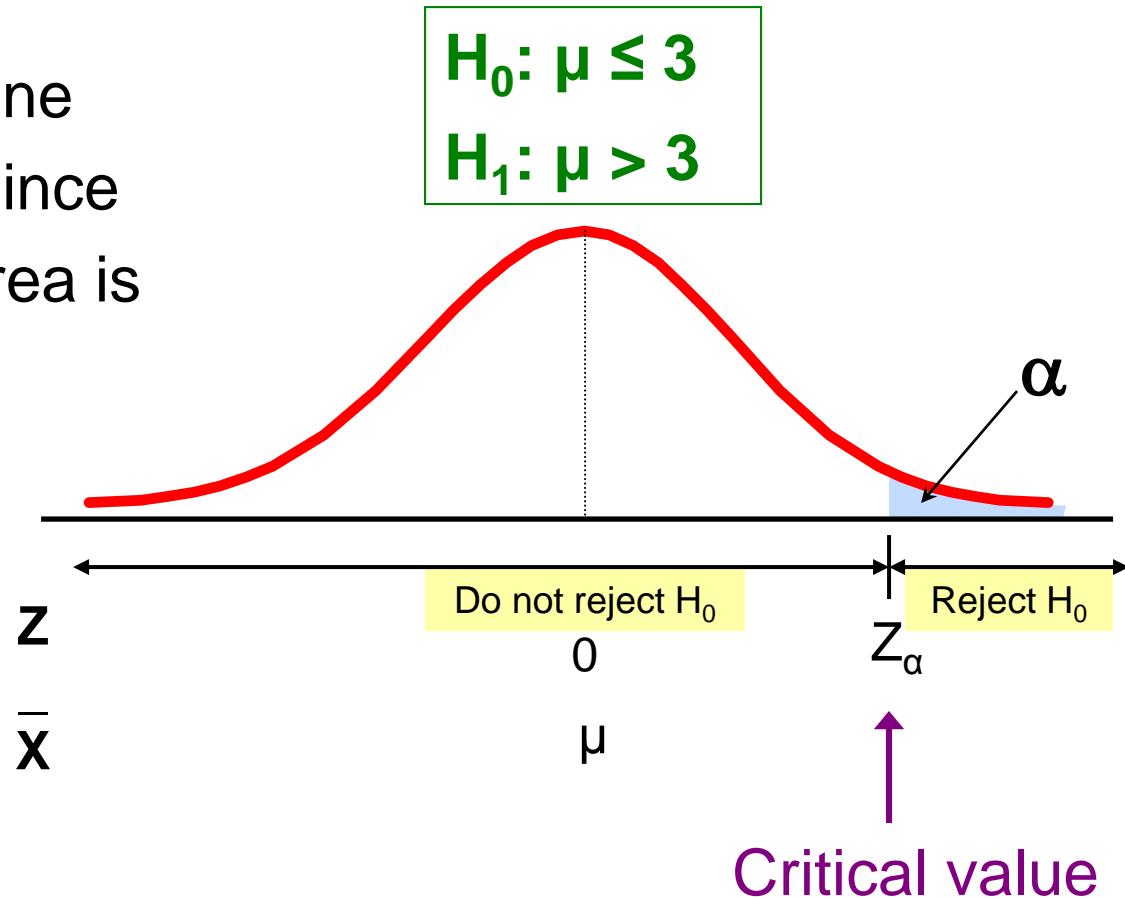
- There is only one critical value, since the rejection area is in only one tail





# Upper-Tail Tests

- There is only one critical value, since the rejection area is in only one tail



# Example: Upper-Tail Z Test for Mean ( $\sigma$ Known)

An investigator thinks that annual health expenditure have increased, and now average over \$52 per year. The investigator wishes to test this claim. (Assume  $\sigma = 10$  is known)

Form hypothesis test:

$H_0: \mu \leq 52$     the average is not over \$52 per month

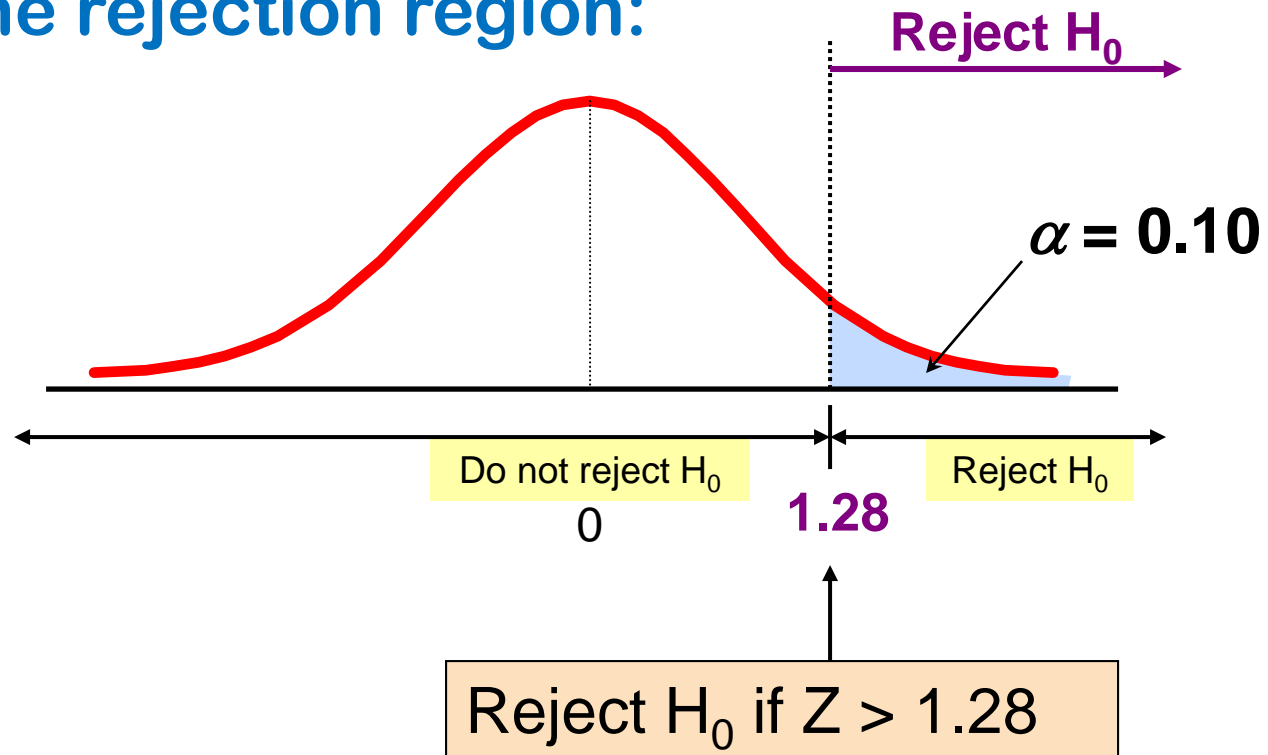
$H_1: \mu > 52$     the average **is** greater than \$52 per year  
(i.e., sufficient evidence exists to support the  
invitagotor's claim)

# Example: Find Rejection Region

(continued)

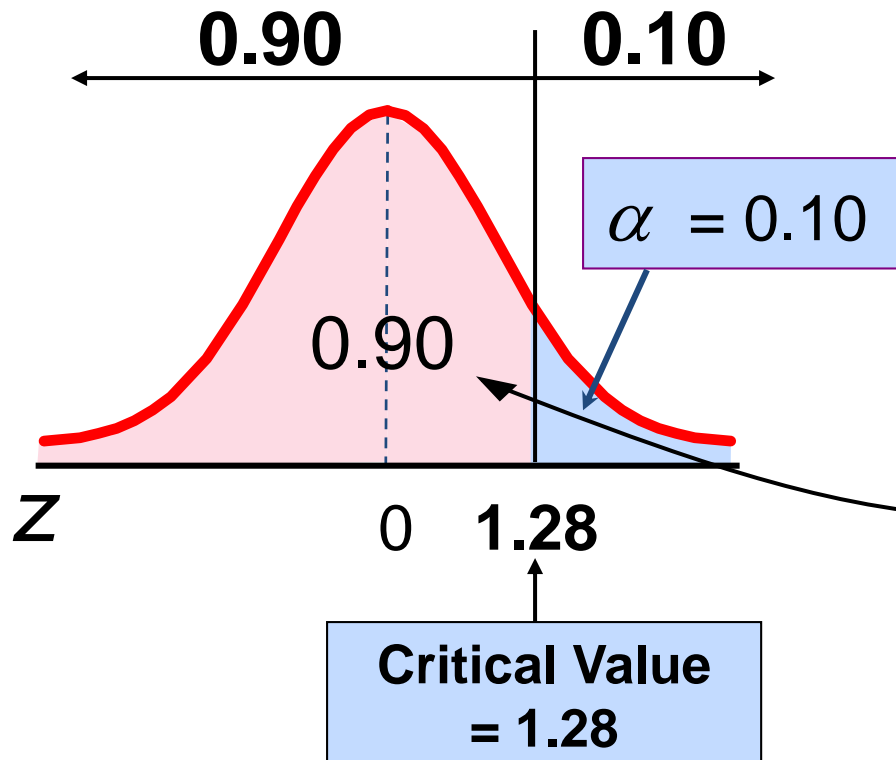
- Suppose that  $\alpha = 0.10$  is chosen for this test

Find the rejection region:



# Review: One-Tail Critical Value

What is Z given  $\alpha = 0.10$ ?



Standardized Normal  
Distribution Table (Portion)

Z	.07	<b>.08</b>	.09
1.1	.8790	.8810	.8830
<b>1.2</b>	.8980	<b>.8997</b>	.9015
1.3	.9147	.9162	.9177

# Example: Test Statistic

*(continued)*

Obtain sample and compute the test statistic

Suppose a sample is taken with the following results:  $n = 64$ ,  $\bar{X} = 53.1$  ( $\sigma=10$  was assumed known)

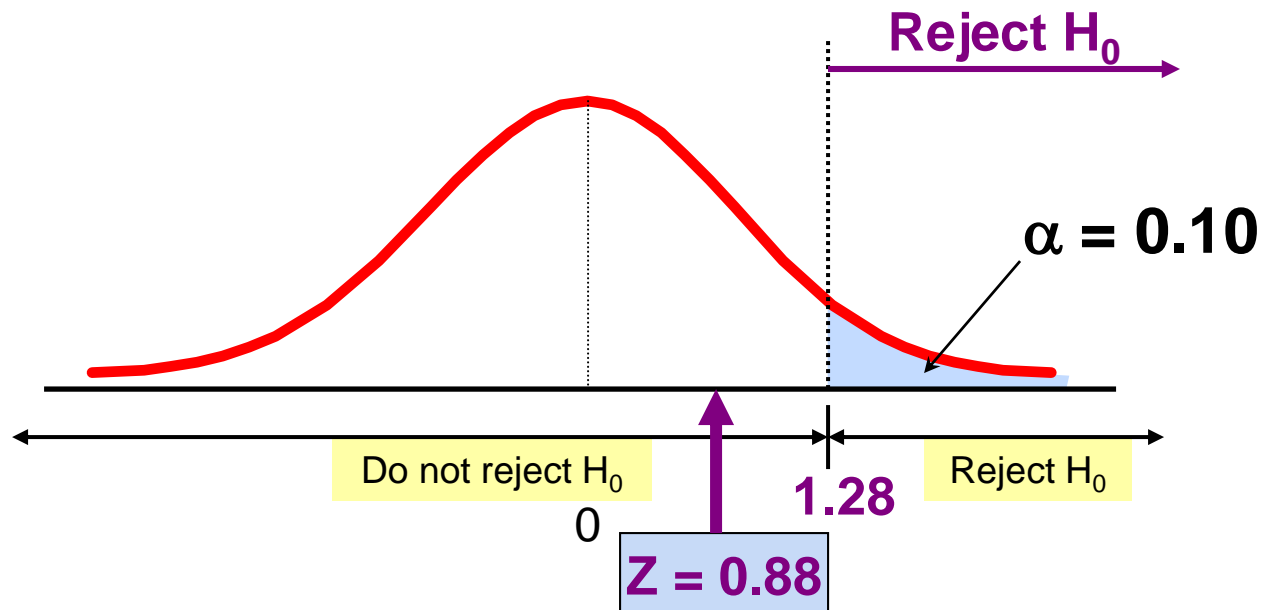
– Then the test statistic is:

$$Z = \frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}} = \frac{53.1 - 52}{\frac{10}{\sqrt{64}}} = 0.88$$

# Example: Decision

(continued)

Reach a decision and interpret the result:



**Do not reject  $H_0$  since  $Z = 0.88 \leq 1.28$**

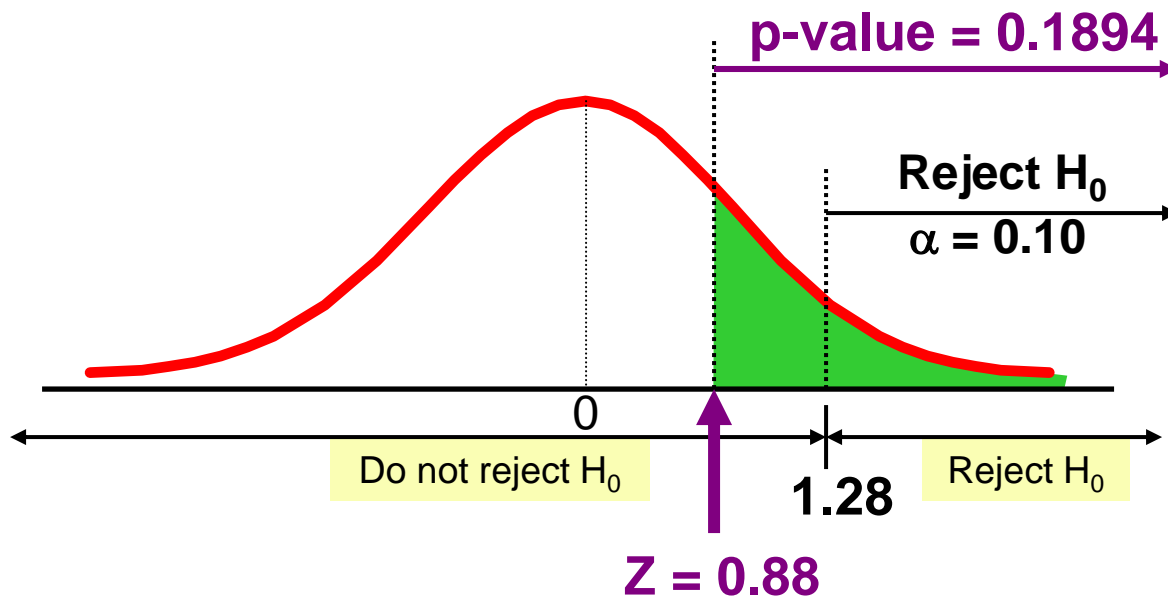
i.e.: there is not sufficient evidence that the mean bill is over \$52

# p-Value Solution

(continued)

Calculate the p-value and compare to  $\alpha$

(assuming that  $\mu = 52.0$ )



$$P(\bar{X} \geq 53.1)$$

$$= P\left(Z \geq \frac{53.1 - 52.0}{10/\sqrt{64}}\right)$$

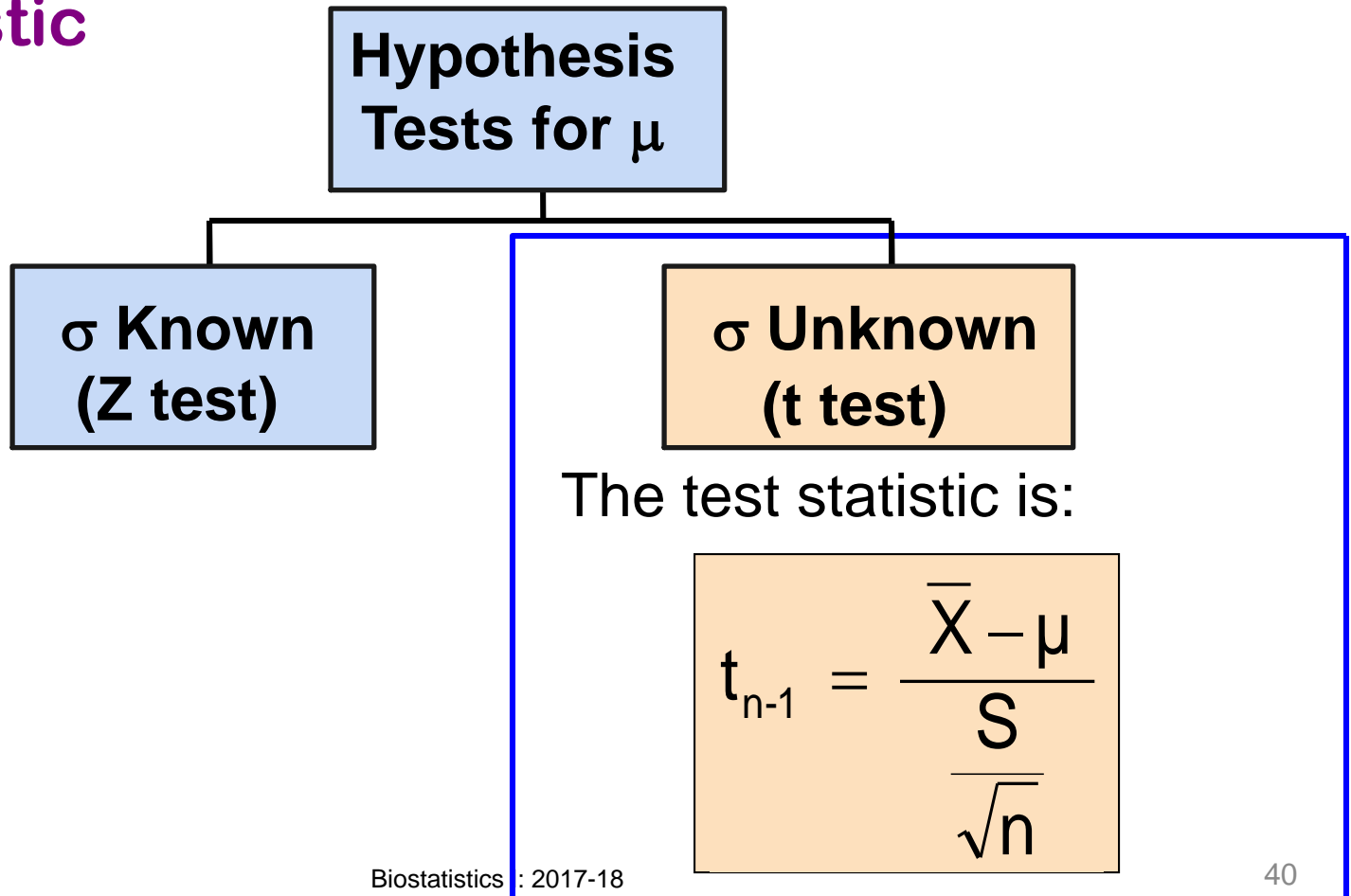
$$= P(Z \geq 0.88) = 1 - 0.8106$$

$$= 0.1894$$

**Do not reject  $H_0$  since p-value = 0.1894 >  $\alpha = 0.10$**

# t Test of Hypothesis for the Mean ( $\sigma$ Unknown)

- Convert sample statistic ( $\bar{X}$ ) to a t test statistic





# Example: Two-Tail Test ( $\sigma$ Unknown)

The average height of male in Indonesia is said to be 168 cm. A random sample of 25 male resulted in

$$\bar{X} = 172.50 \text{ cm and}$$

$$S = 15.40 \text{ cm.}$$

Test at the  $\alpha = 0.05$  level.

(Assume the population distribution is normal)

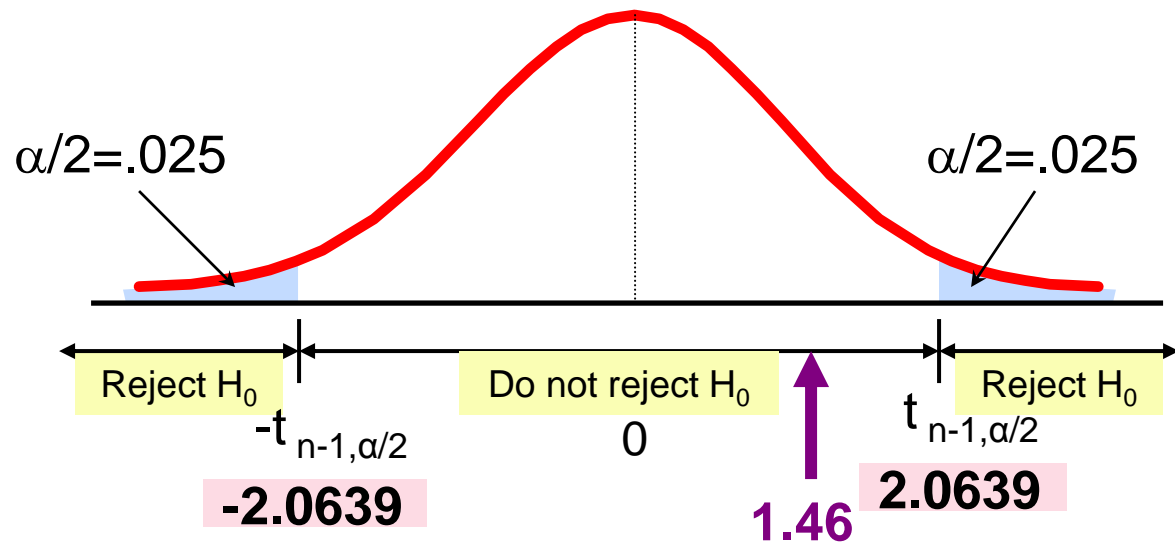
$$H_0: \mu = 168$$

$$H_1: \mu \neq 168$$

# Example Solution: Two-Tail Test

$$H_0: \mu = 168$$
$$H_1: \mu \neq 168$$

- $\alpha = 0.05$
- $n = 25$
- $\sigma$  is unknown, so use a t statistic
- Critical Value:



$$t_{n-1} = \frac{\bar{X} - \mu}{\frac{S}{\sqrt{n}}} = \frac{172.50 - 168}{\frac{15.40}{\sqrt{25}}} = 1.46$$

**Do not reject  $H_0$ :** not sufficient evidence that true mean height is different than 168 cm

# Connection to Confidence Intervals

- For  $\bar{X} = 172.5$ ,  $S = 15.40$  and  $n = 25$ , the 95% confidence interval is:

$$172.5 - (2.0639) 15.4/\sqrt{25} \quad \text{to} \quad 172.5 + (2.0639) 15.4/\sqrt{25}$$

$$166.14 \leq \mu \leq 178.86$$

- Since this interval contains the Hypothesized mean (168), we do not reject the null hypothesis at  $\alpha = 0.05$

# Hypothesis Tests for Proportions

- Involves categorical variables
- Two possible outcomes
  - “Success” (possesses a certain characteristic)
  - “Failure” (does not possess that characteristic)
- Fraction or proportion of the population in the “success” category is denoted by  $\pi$

# Proportions

(continued)

- Sample proportion in the success category is denoted by  $p$

$$p = \frac{X}{n} = \frac{\text{number of successes in sample}}{\text{sample size}}$$

- When both  $n\pi$  and  $n(1-\pi)$  are at least 5,  $p$  can be approximated by a normal distribution with mean and standard deviation

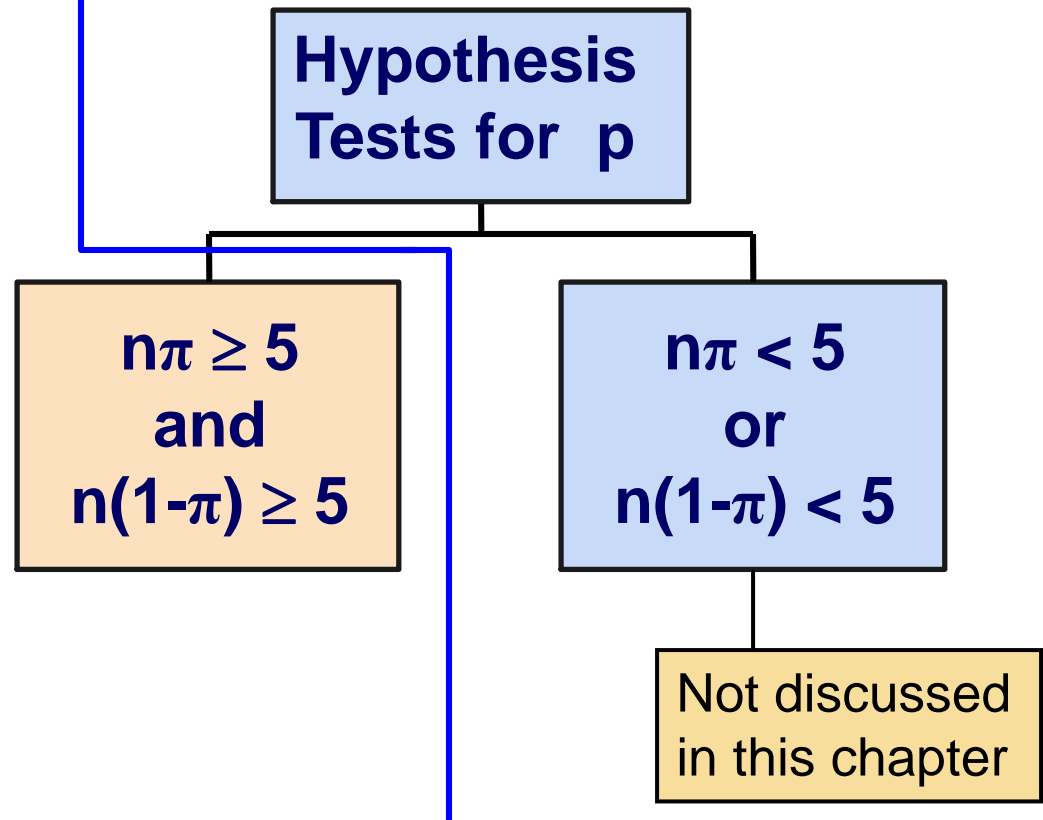
$$\mu_p = \pi$$

$$\sigma_p = \sqrt{\frac{\pi(1-\pi)}{n}}$$

# Hypothesis Tests for Proportions

- The sampling distribution of  $\hat{p}$  is approximately normal, so the test statistic is a Z value:

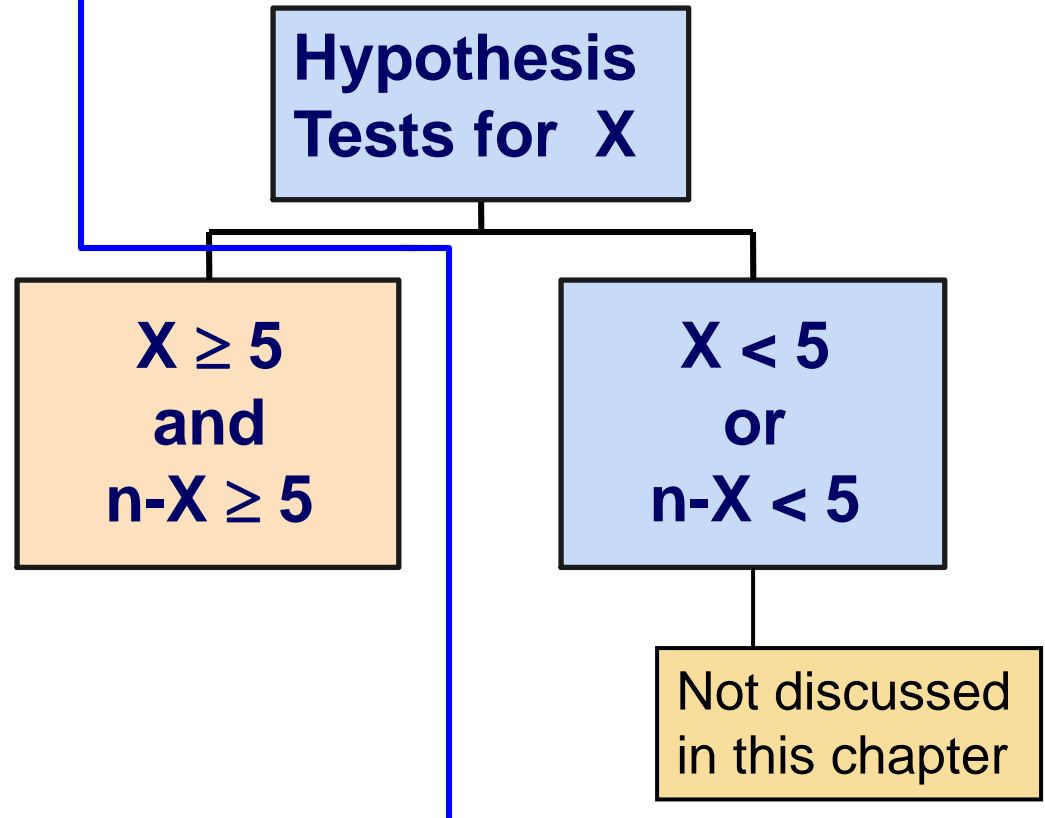
$$Z = \frac{\hat{p} - \pi}{\sqrt{\frac{\pi(1-\pi)}{n}}}$$



# Z Test for Proportion in Terms of Number of Successes

- An equivalent form to the last slide, but in terms of the number of successes,  $X$ :

$$Z = \frac{X - n\pi}{\sqrt{n\pi(1-\pi)}}$$



# Example: Z Test for Proportion

A researcher claims that it receives 8% responses from its mailing. To test this claim, a random sample of 500 were surveyed with 25 responses. Test at the  $\alpha = 0.05$  significance level.

Check:

$$n\pi = (500)(.08) = 40$$

$$n(1-\pi) = (500)(.92) = 460$$





# Z Test for Proportion: Solution

$$H_0: \pi = 0.08$$

$$H_1: \pi \neq 0.08$$

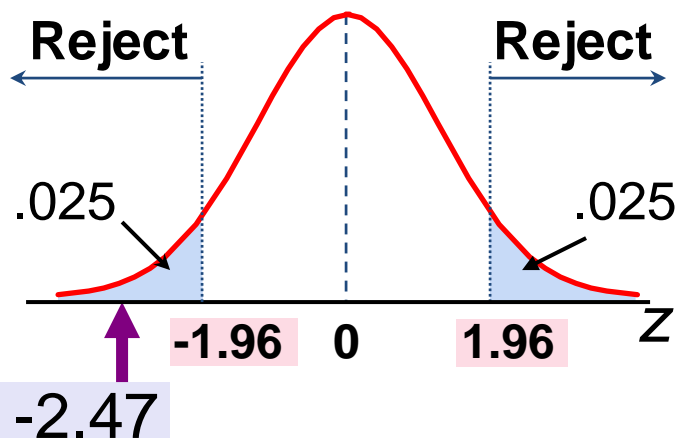
$$\alpha = 0.05$$

$$n = 500, p = 0.05$$

**Test Statistic:**

$$Z = \frac{p - \pi}{\sqrt{\frac{\pi(1 - \pi)}{n}}} = \frac{.05 - .08}{\sqrt{\frac{.08(1 - .08)}{500}}} = -2.47$$

**Critical Values:  $\pm 1.96$**



**Decision:**

Reject  $H_0$  at  $\alpha = 0.05$

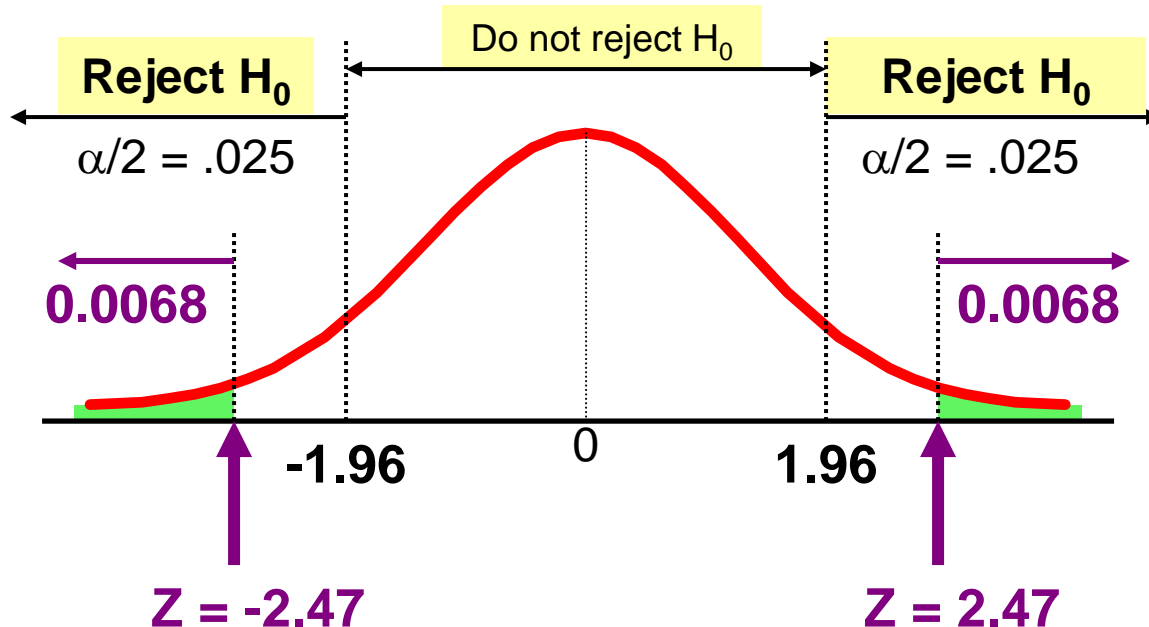
**Conclusion:**

There is sufficient evidence to reject the researcher claim of 8% response rate.

# p-Value Solution

(continued)

Calculate the p-value and compare to  $\alpha$   
(For a two-tail test the p-value is always two-tail)



**p-value = 0.0136:**

$$P(Z \leq -2.47) + P(Z \geq 2.47) \\ = 2(0.0068) = 0.0136$$

**Reject  $H_0$  since p-value = 0.0136 <  $\alpha$  = 0.05**

# Potential Pitfalls and Ethical Considerations

- Use randomly collected data to reduce selection biases
- Do not use human subjects without informed consent
- Choose the level of significance,  $\alpha$ , and the type of test (one-tail or two-tail) before data collection
- Do not employ “data snooping” to choose between one-tail and two-tail test, or to determine the level of significance
- Do not practice “data cleansing” to hide observations that do not support a stated hypothesis
- Report all pertinent findings

# Chapter Summary

- Addressed hypothesis testing methodology
- Performed Z Test for the mean ( $\sigma$  known)
- Discussed critical value and p-value approaches to hypothesis testing
- Performed one-tail and two-tail tests

# Chapter Summary

*(continued)*

- Performed t test for the mean ( $\sigma$  unknown)
- Performed Z test for the proportion
- Discussed pitfalls and ethical issues