Business Statistic

Week 10

Fundamentals of Hypothesis Testing: One-Sample Test

Agenda

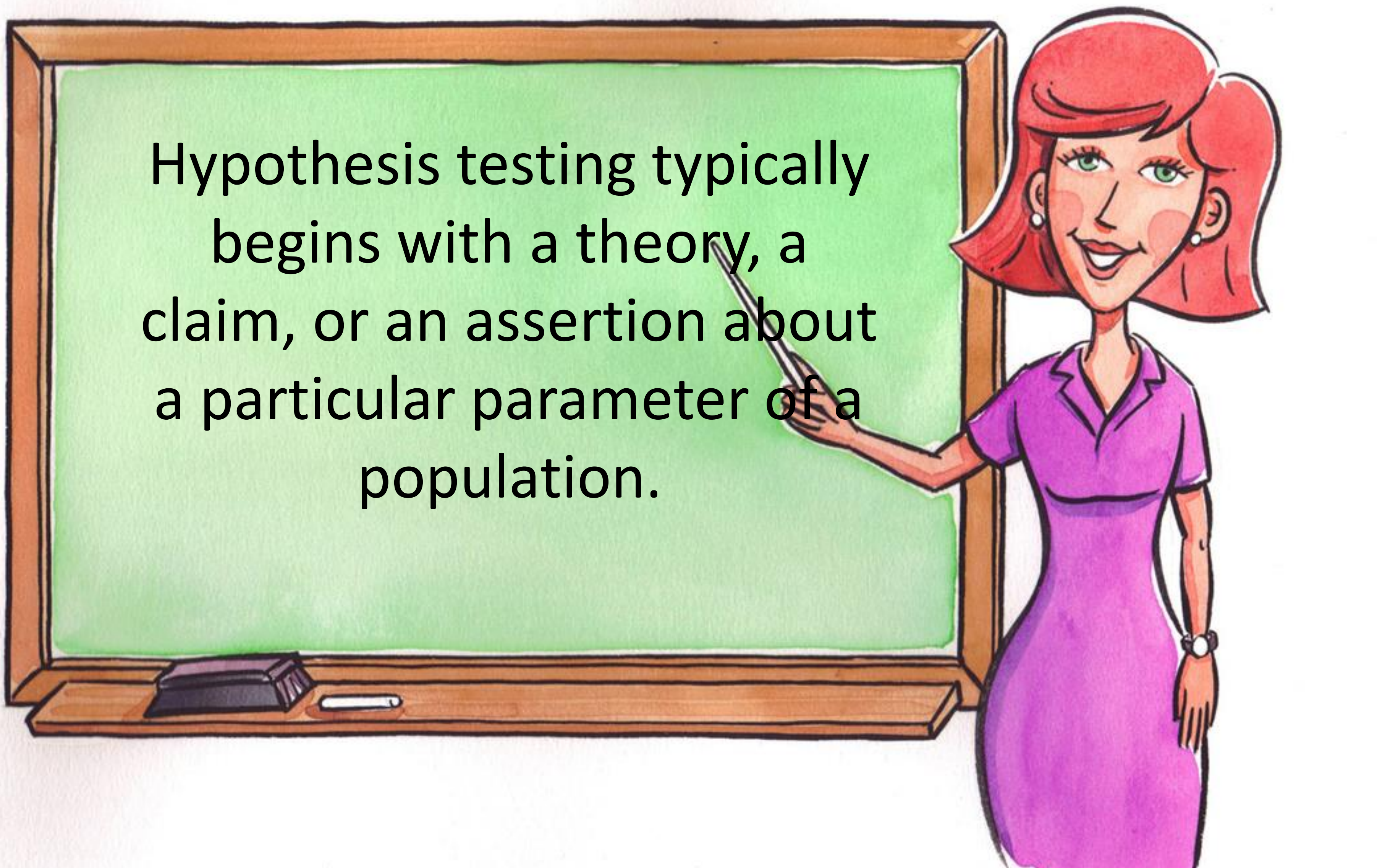
Time	Activity
20 minutes	Fundamentals of Hypothesis-Testing Methodology
40 minutes	z Test of Hypothesis for The Mean (σ Known)
40 minutes	t Test of Hypothesis for The Mean (σ Unknown)
40 minutes	Z Test of Hypothesis for Proportions
60 minutes	Exercise

Learning Objectives

In this chapter, you learn:

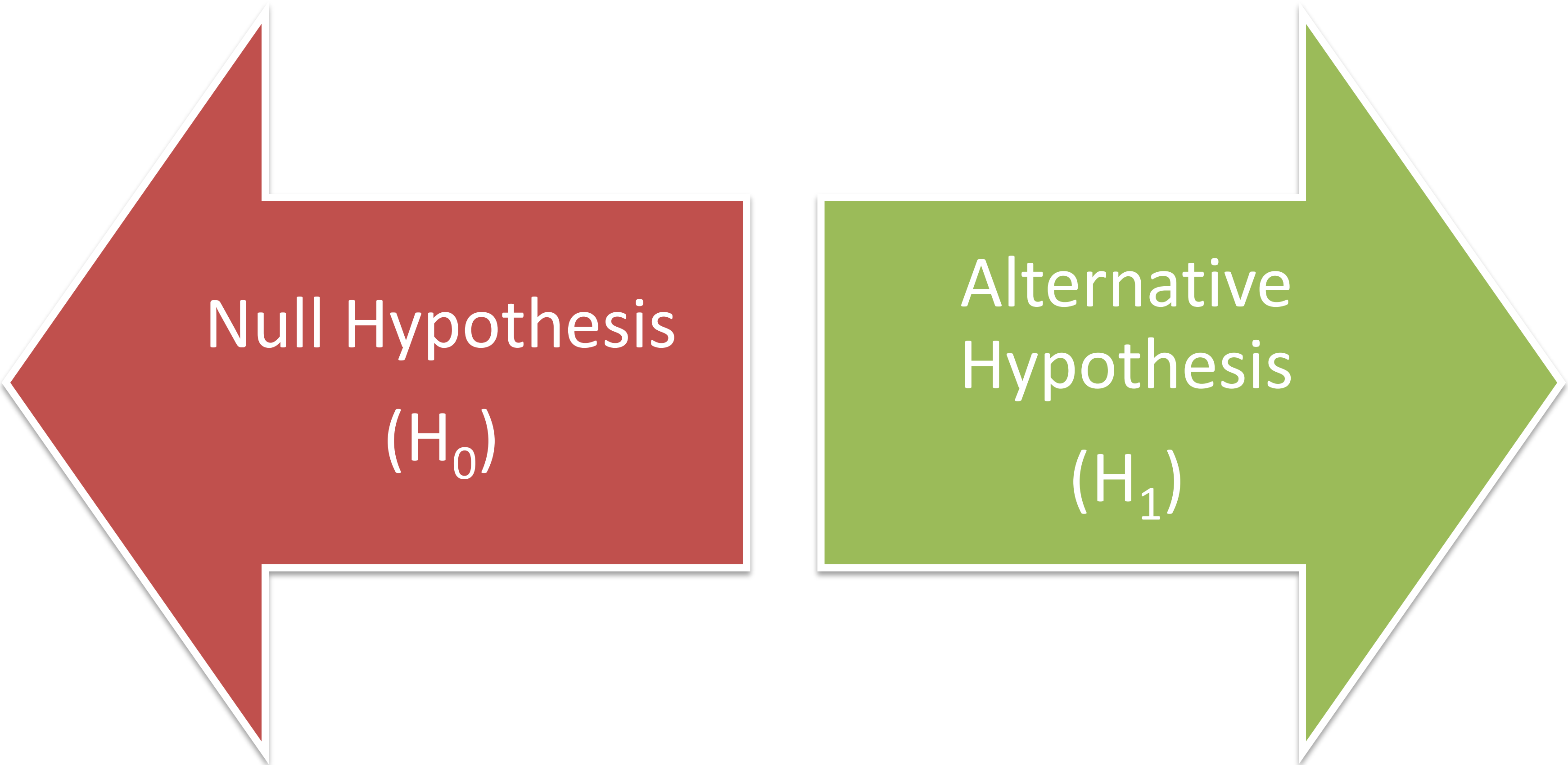
- The basic principles of hypothesis testing
- How to use hypothesis testing to test a mean or proportion
- The assumptions of each hypothesis-testing procedure, how to evaluate them, and the consequences if they are seriously violated
- How to avoid the pitfalls involved in hypothesis testing
- Ethical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

Hypothesis



Hypothesis testing typically begins with a theory, a claim, or an assertion about a particular parameter of a population.

Hypotheses



The Null Hypothesis, H_0

- Begin with the assumption that the null hypothesis is true
  - Similar to the notion of innocent until proven guilty

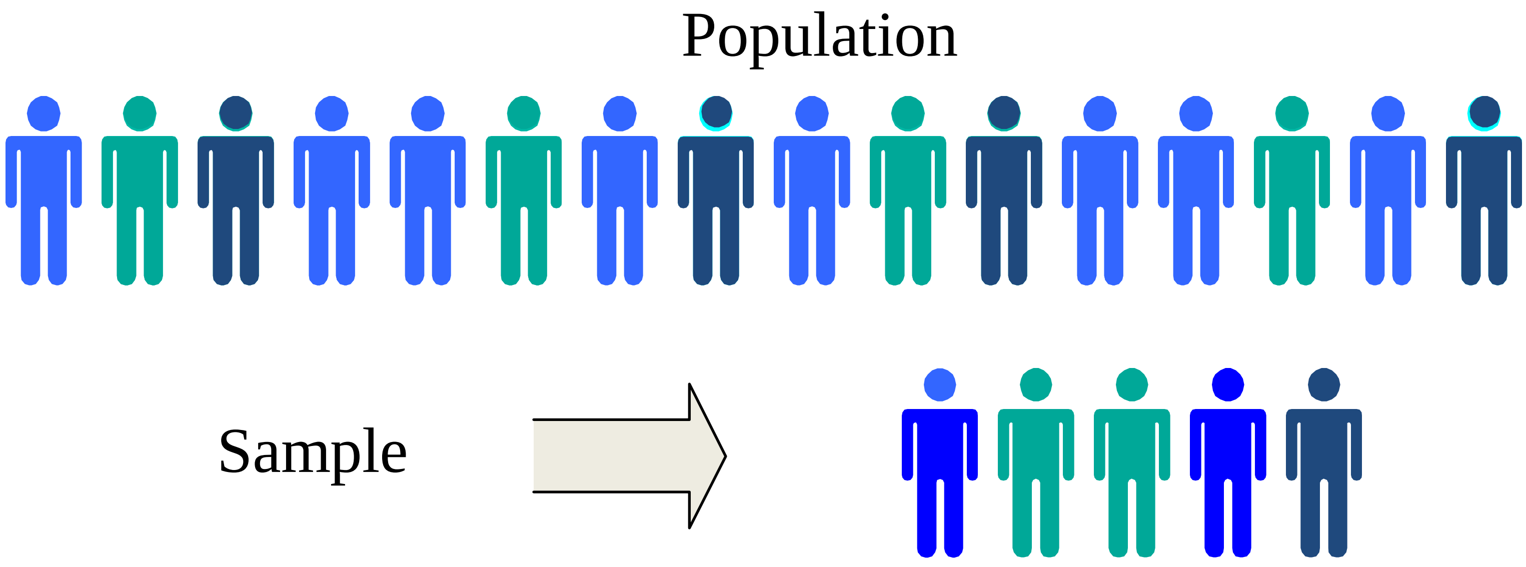
- Refers to the status quo or historical value
- Always contains "=" , "≤" or "≥" sign
- May or may not be rejected

The Alternative Hypothesis, H_1

- Is the opposite of the null hypothesis
  - e.g., The average number of TV sets in U.S. homes is not equal to 3 ($H_1: \mu \neq 3$)
- Challenges the status quo
- Never contains the "=" , "≤" or "≥" sign
- May or may not be proven
- Is generally the hypothesis that the researcher is trying to prove

The Hypothesis Testing Process

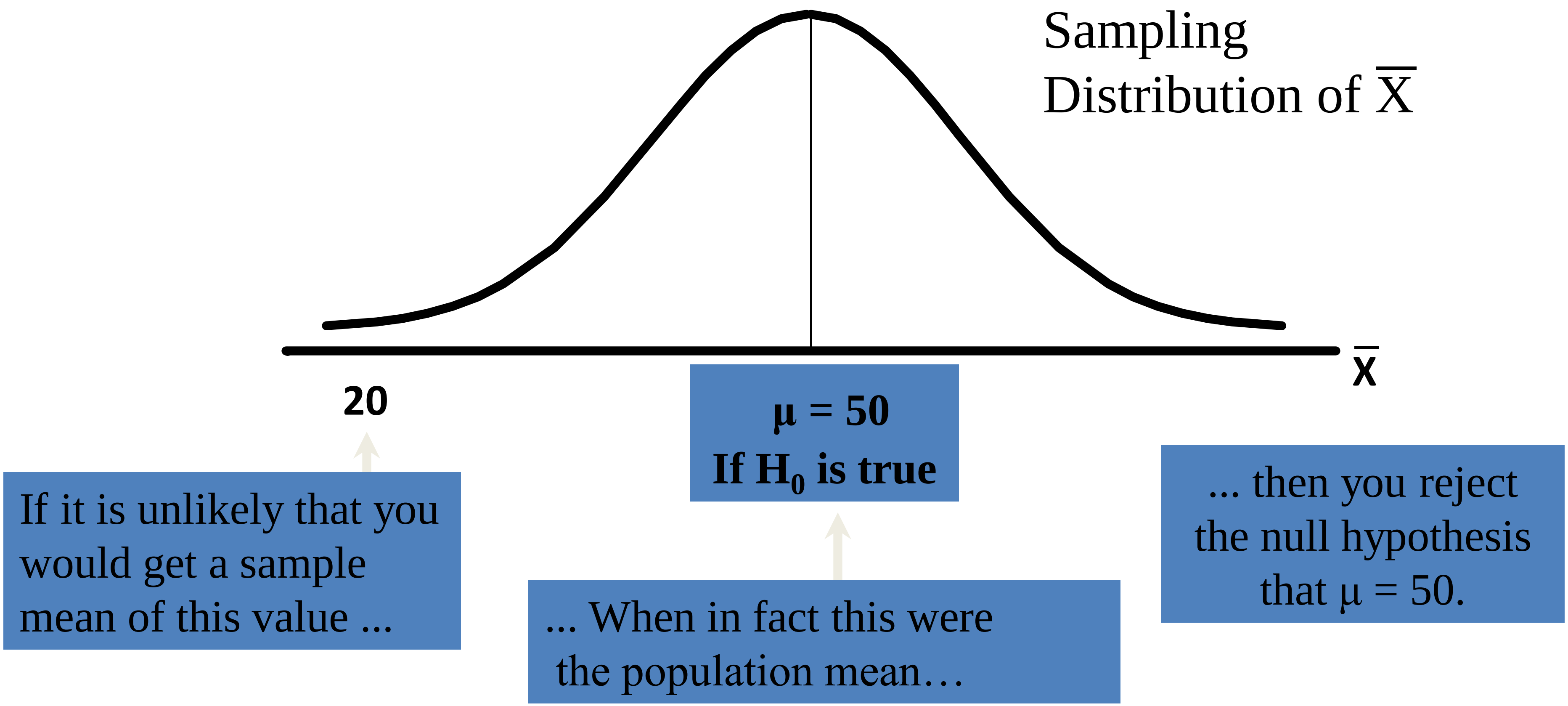
- Claim: The population mean age is 50.
  - H_0: $\mu = 50$, H_1: $\mu \neq 50$
- Sample the population and find sample mean.



The Hypothesis Testing Process

- Suppose the sample mean age was $\overline{X} = 20$.
- This is significantly lower than the claimed mean population age of 50.
- If the null hypothesis were true, the probability of getting such a different sample mean would be very small, so you reject the null hypothesis .
- In other words, getting a sample mean of 20 is so unlikely if the population mean was 50, you conclude that the population mean must not be 50.

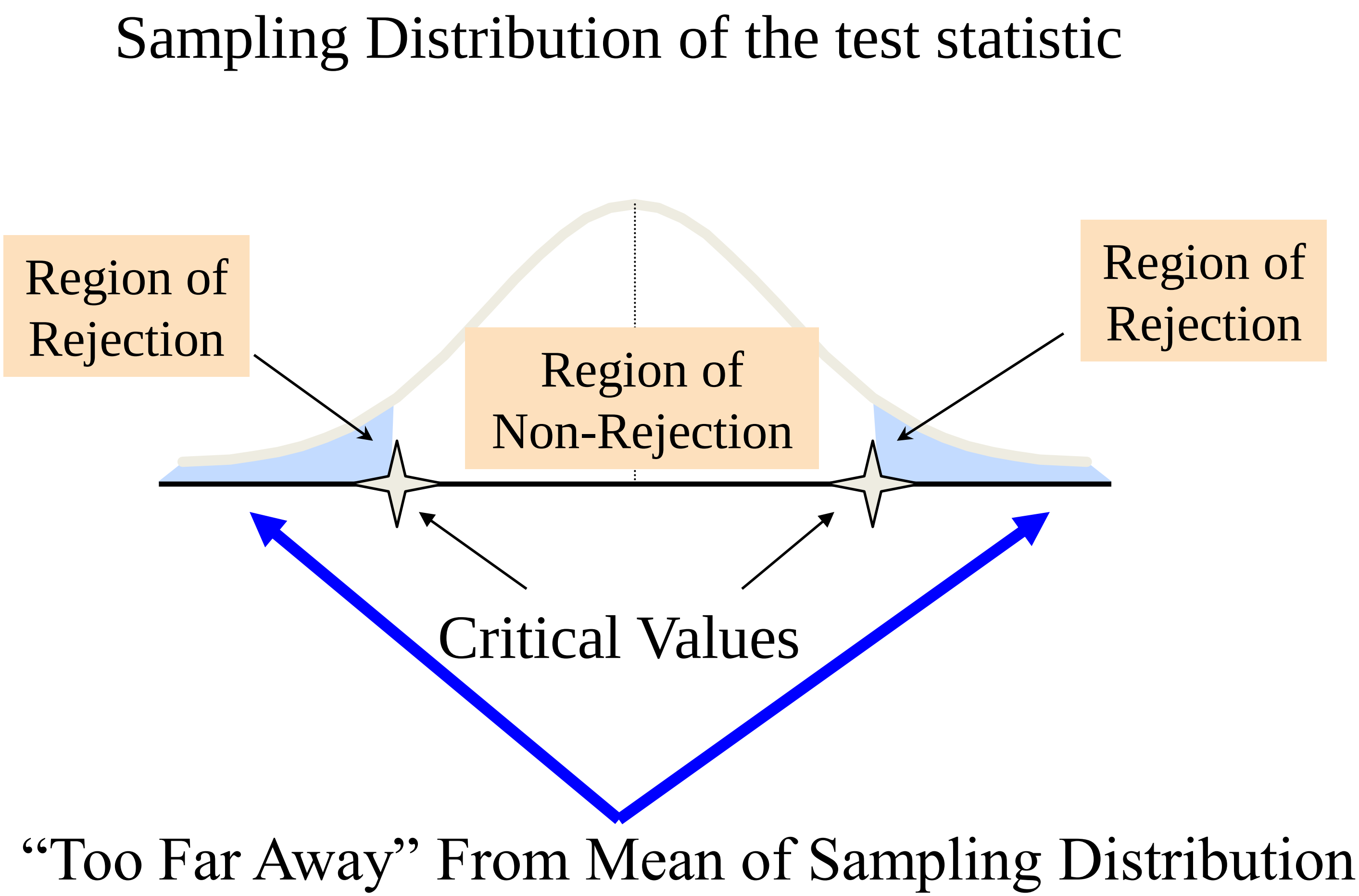
The Hypothesis Testing Process



The Test Statistic and Critical Values

- If the sample mean is close to the assumed population mean, the null hypothesis is not rejected.
- If the sample mean is far from the assumed population mean, the null hypothesis is rejected.
- How far is “far enough” to reject H_0?
- The critical value of a test statistic creates a “line in the sand” for decision making -- it answers the question of how far is far enough.

The Test Statistic and Critical Values



Possible Errors in Hypothesis Test Decision Making

Possible Hypothesis Test Outcomes		
	Actual Situation	
Decision	H_0 True	H_0 False
Do Not Reject H_0	No Error Probability 1 - α	Type II Error Probability β
Reject H_0	Type I Error Probability α	No Error Probability 1 - β

Possible Results in Hypothesis Test Decision Making

- The confidence coefficient $(1-\alpha)$ is the probability of not rejecting H_0 when it is true.
- The confidence level of a hypothesis test is $(1-\alpha)*100\%$.
- The power of a statistical test $(1-\beta)$ is the probability of rejecting H_0 when it is false.

Type I & II Error Relationship

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

If Type I error probability (α) ⬆ , then

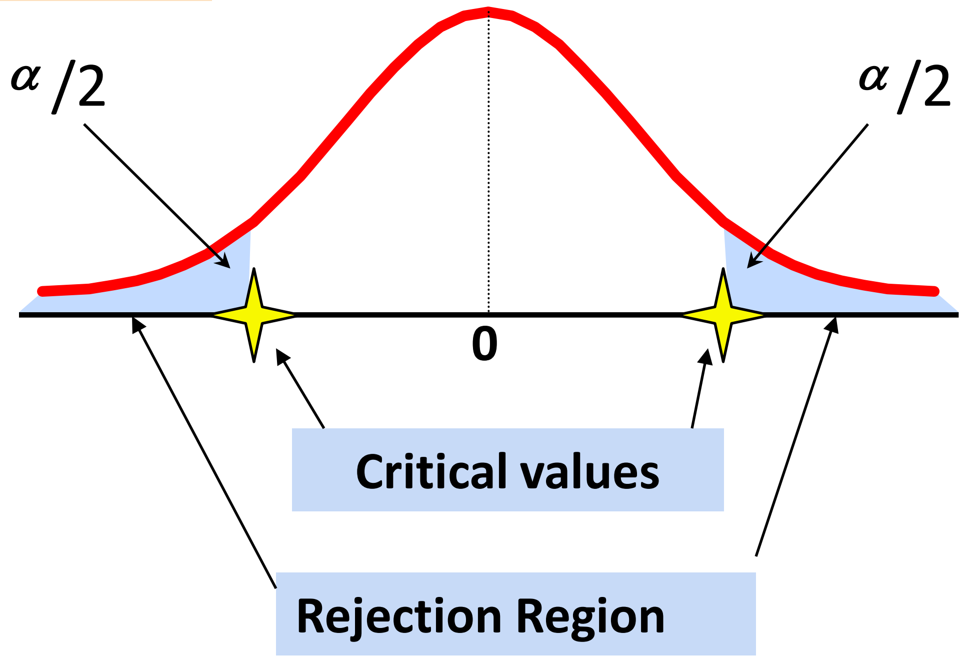
Type II error probability (β) ⬇

Level of Significance and the Rejection Region

H_0: $\mu = 3$

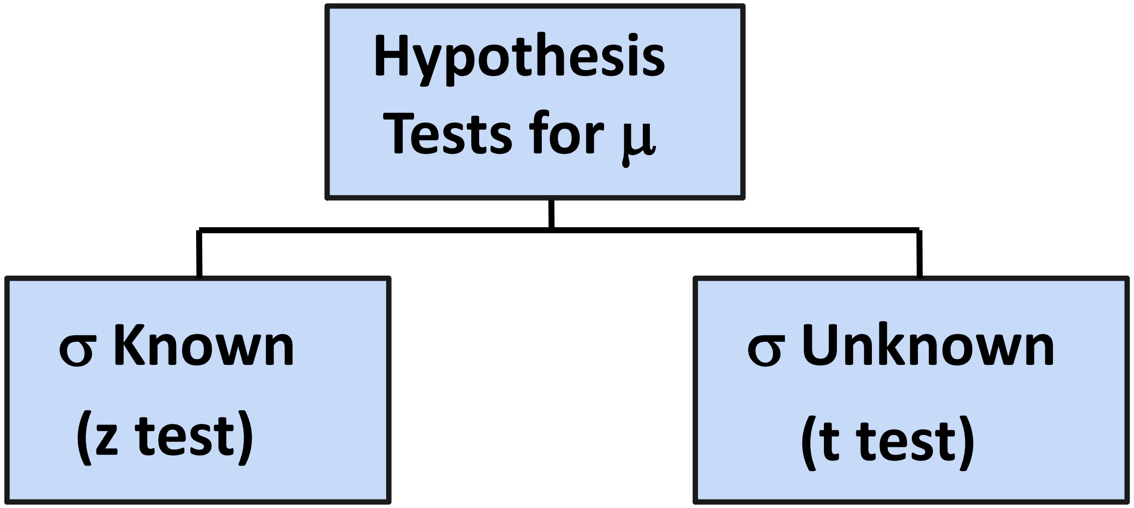
H_1: $\mu \neq 3$

Level of significance = α



This is a two-tail test because there is a rejection region in both tails

Hypothesis Tests for the Mean



z TEST OF HYPOTHESIS FOR THE MEAN (σ KNOWN)

Z Test of Hypothesis for the Mean (σ Known)

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

Hypothesis Tests for μ

σ Known (z test)

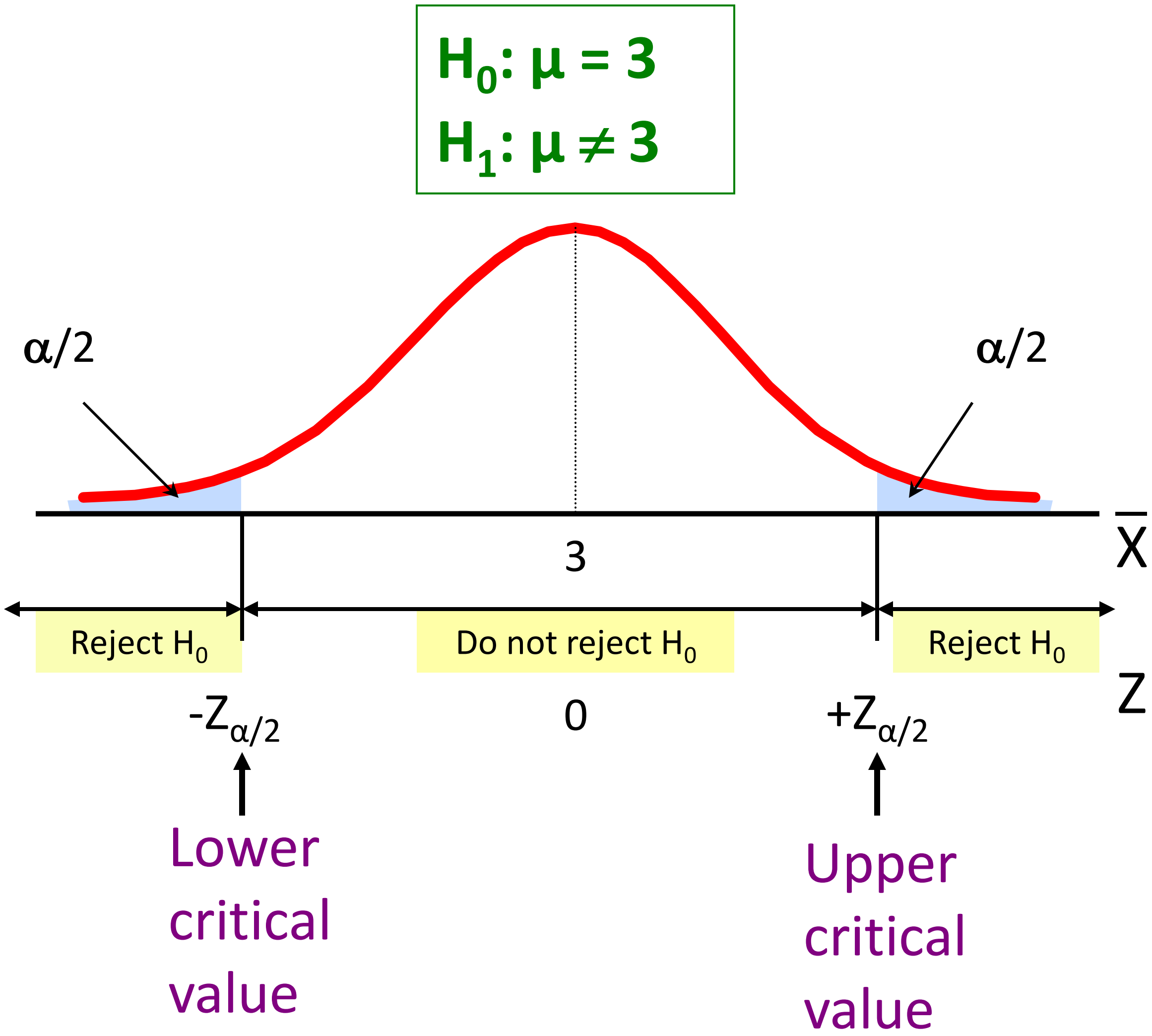
σ Unknown (t test)

The test statistic is:

$$Z_{STAT} = \frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}}$$

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, α, 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
5. Collect data and compute the value of the test statistic

Statistical Decision

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.

p-Value Approach to Testing

- p-value: Probability of obtaining a test statistic equal to or more extreme than the observed sample value given H_0 is true
  - The p-value is also called the observed level of significance
  - It is the smallest value of α for which H_0 can be rejected

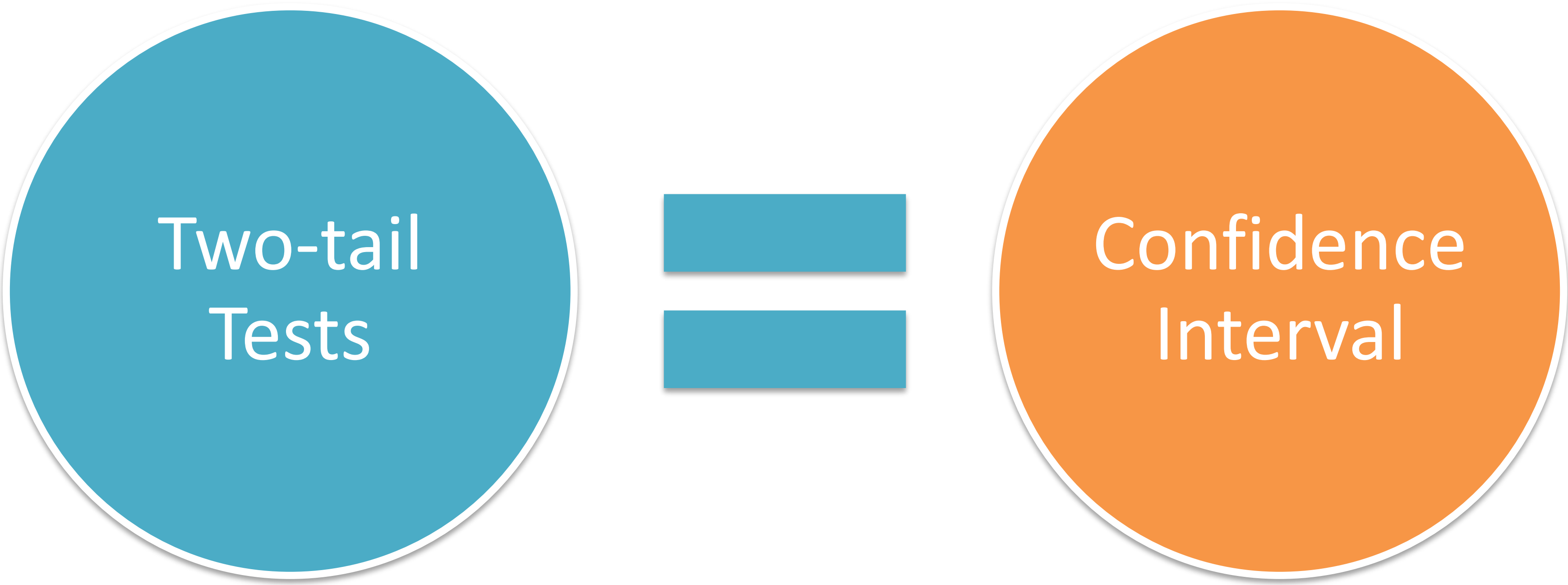
The 5 Step p-value approach to Hypothesis Testing

1. State the null hypothesis, H_0 and the alternative hypothesis, H_1
2. Choose the level of significance, α, and the sample size, n
3. Determine the appropriate test statistic and sampling distribution
4. Collect data and compute the value of the test statistic and the p-value
5. Make the statistical decision and state the managerial conclusion.

Statistical Decision

If the p-value is $< \alpha$ then reject H_0, otherwise do not reject H_0.

Connection Between Two Tail Tests and Confidence Intervals



EXERCISE

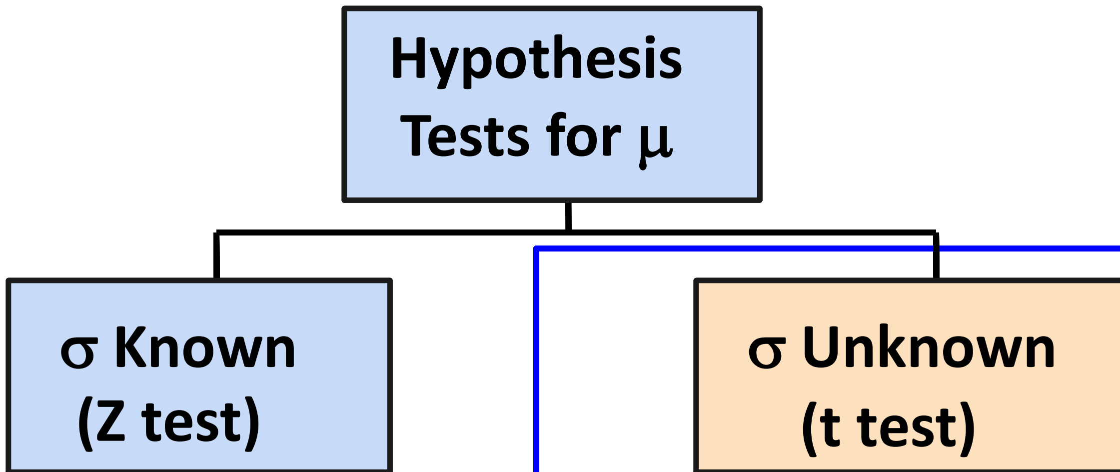
Fast-food Restaurant

You are the manager of a fast-food restaurant. The business problem is to determine whether the population mean waiting time to place an order has changed in the past month from its previous population mean value of 4.5 minutes. From past experience, you can assume that the population is normally distributed, with a population standard deviation of 1.2 minutes. You select a sample of 25 orders during a one-hour period. The sample mean is 5.1 minutes. Determine whether there is evidence at the 0.05 level of significance that the population mean waiting time to place an order has changed in the past month from its previous population mean value of 4.5 minutes.

t TEST OF HYPOTHESIS FOR THE MEAN (σ UNKNOWN)

t Test of Hypothesis for the Mean (σ Unknown)

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



The test statistic is:

$$t_{STAT} = \frac{\bar{X} - \mu}{\frac{S}{\sqrt{n}}}$$

Example Two-Tail t Test Using A p-value from Excel

- Since this is a t-test we cannot calculate the p-value without some calculation aid.
- The Excel output below does this:

t Test for the Hypothesis of the Mean

Data		
Null Hypothesis μ=	**$**	**168.00**
Level of Significance		**0.05**
Sample Size		**25**
Sample Mean	**$**	**172.50**
Sample Standard Deviation	**$**	**15.40**

Intermediate Calculations			
Standard Error of the Mean	$	3.08	=B8/SQRT(B6)
Degrees of Freedom		24	=B6-1
t test statistic		**1.46**	=(B7-B4)/B11

Two-Tail Test		
Lower Critical Value	-2.0639	=-TINV(B5,B12)
Upper Critical Value	2.0639	=TINV(B5,B12)
p-value	0.157	=TDIST(ABS(B13),B12,2)
Do Not Reject Null Hypothesis		=IF(B18<B5, "Reject null hypothesis", "Do not reject null hypothesis")

p-value > α
So do not reject H_0

EXAMPLE

Saxon Home Improvement

Saxon Home Improvement distributes home improvement supplies in the northeastern United States. The business objective is to determine whether the mean amount per sales invoice is unchanged from the $120 of the past five years. As an accountant for the company, you need to determine whether this amount changes. You collect the data from a sample of n = 12 sales invoices. You decide to use $\alpha = 0.05$. You organize and store the data from a random sample of 12 sales invoices, and got a *sample mean of* $112.85 and *standard deviation of* $20.80.

ONE-TAIL TESTS

One-Tail Tests

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

H_0: $\mu \geq 3$

H_1: $\mu < 3$

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

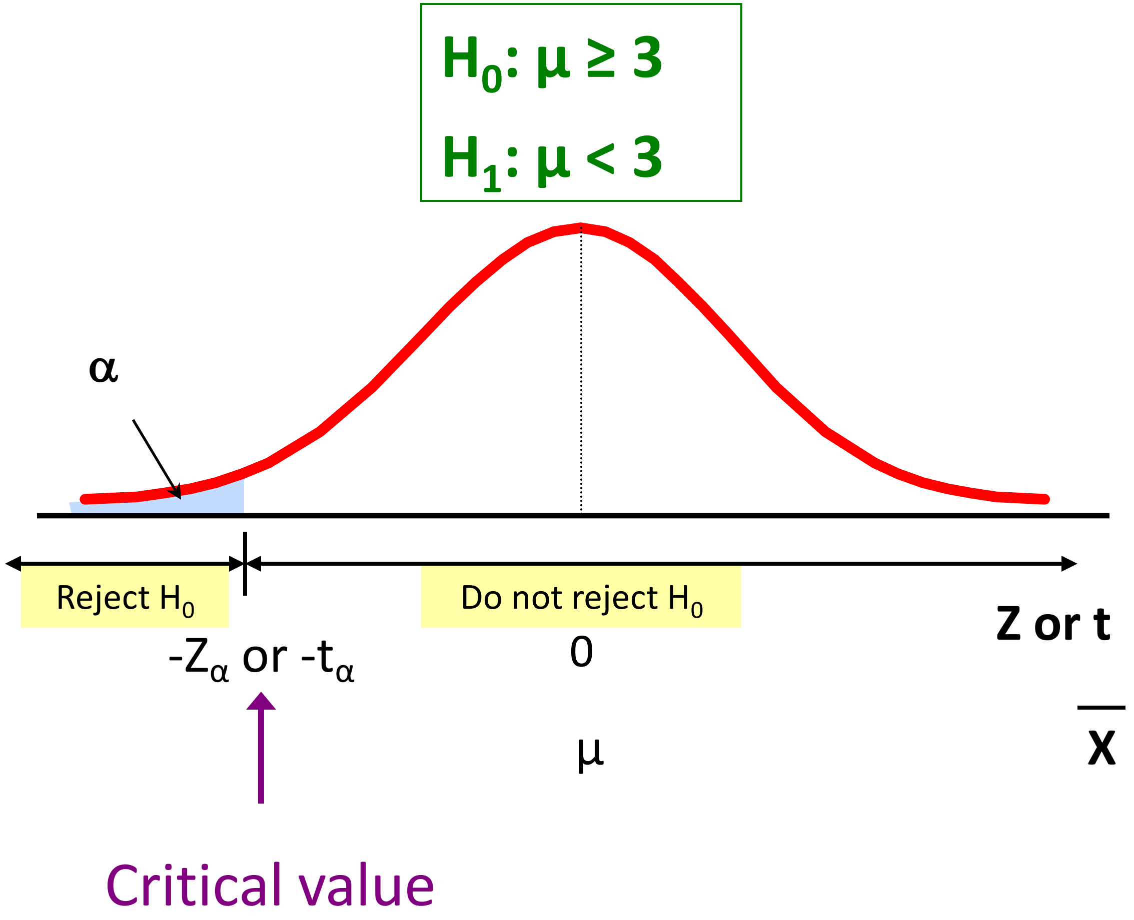
H_0: $\mu \leq 3$

H_1: $\mu > 3$

→ 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



McDonald’s (cont’d)

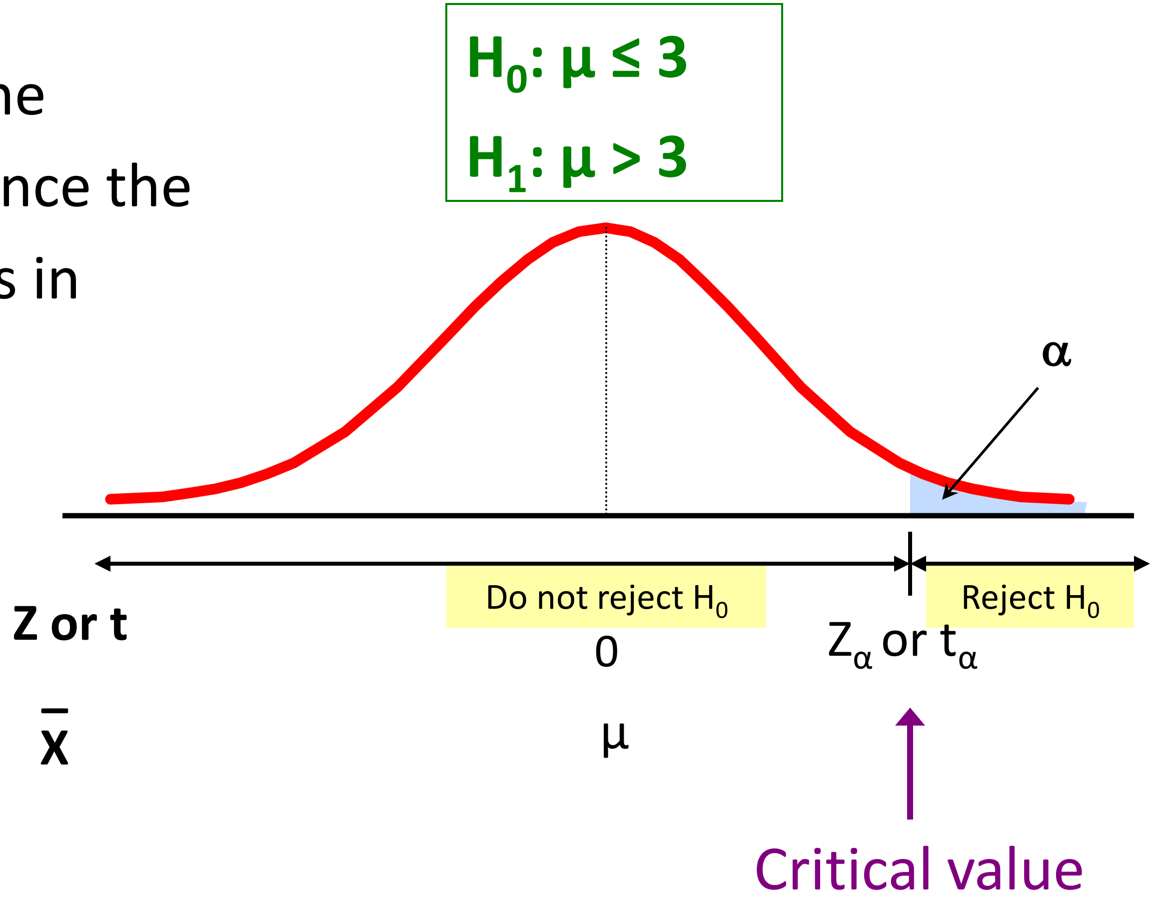
The speed with which customers are served is of critical importance to the success of the service. In one past study, McDonald’s had a mean service time of 174.22 seconds, which was only ninth best in the industry. Suppose that McDonald’s began a quality improvement effort to reduce the service time by deploying an improved drive-through service process in a sample of 25 stores.

McDonald’s

You wish to determine whether the new drive-through process has a mean that is less than 174.22 seconds. From the sample of 25 stores you selected, you find that the sample mean service time at the drive-through equals 162.96 seconds and the sample standard deviation equals 20.2 seconds. (Use a = 0.05)

Upper-Tail Tests

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



Chocolate Bars

A company that manufactures chocolate bars is particularly concerned that the mean weight of a chocolate bar is not greater than 6.03 ounces. A sample of 50 chocolate bars is selected; the sample mean is 6.034 ounces, and the sample standard deviation is 0.02 ounces. Using the α = 0.01 level of significance, is there evidence that the population mean weight of the chocolate bars is greater than 6.03 ounces?

Example: Upper-Tail t Test for Mean (σ unknown)

A phone industry manager thinks that customer monthly cell phone bills have increased, and now average over \$52 per month. The company wishes to test this claim. (Assume a normal population)

Form hypothesis test:

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

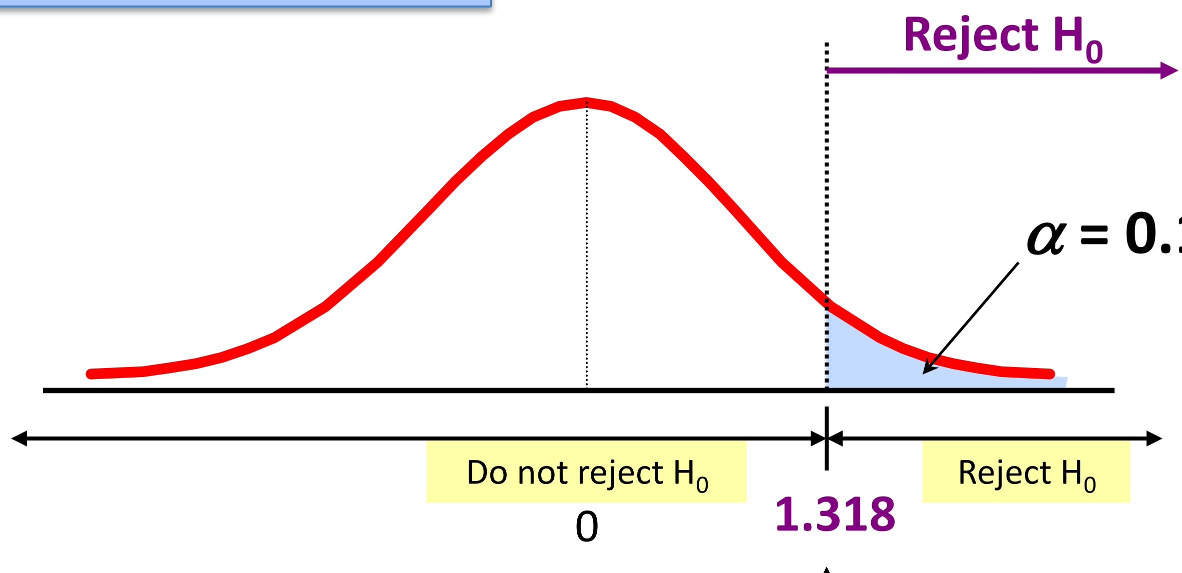
H_1: $\mu > 52$ the average **is** greater than \$52 per month
(i.e., sufficient evidence exists to support the manager's claim)

Example: Find Rejection Region

(continued)

- Suppose that α = 0.10 is chosen for this test and n = 25.

Find the rejection region:



Reject H_0 if t_{STAT} > 1.318

Example: Test Statistic

(continued)

Obtain sample and compute the test statistic

Suppose a sample is taken with the following results: n = 25, $\bar{X}$ = 53.1, and S = 10

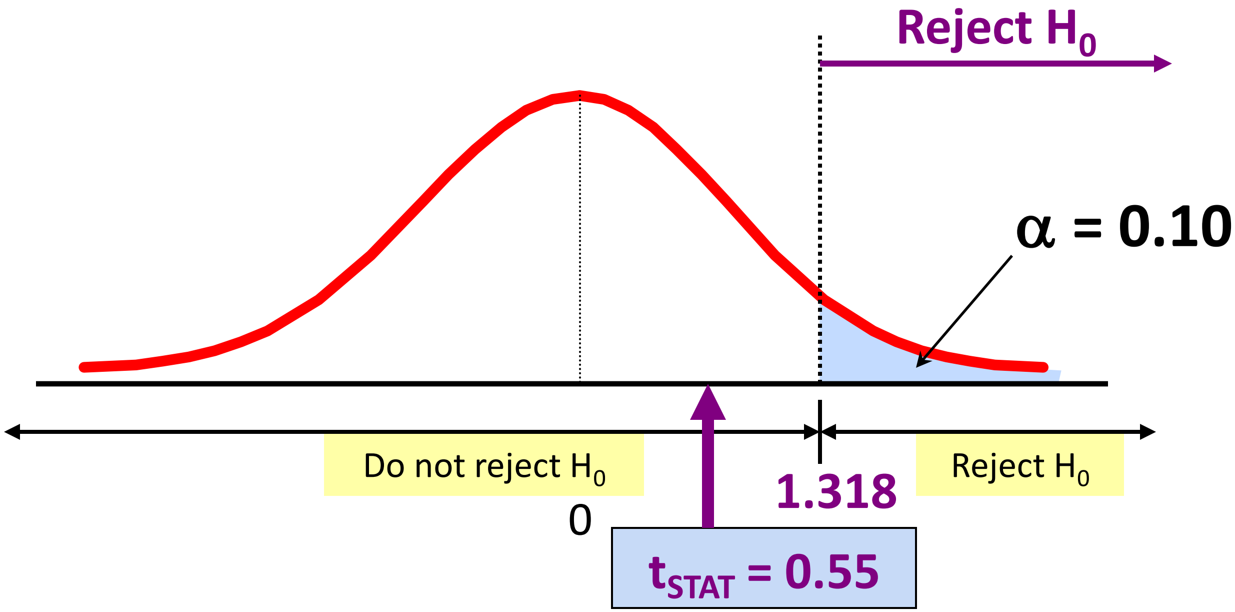
– Then the test statistic is:

$$t_{STAT} = \frac{\bar{X} - \mu}{\frac{S}{\sqrt{n}}} = \frac{53.1 - 52}{\frac{10}{\sqrt{25}}} = 0.55$$

Example: Decision

(continued)

Reach a decision and interpret the result:

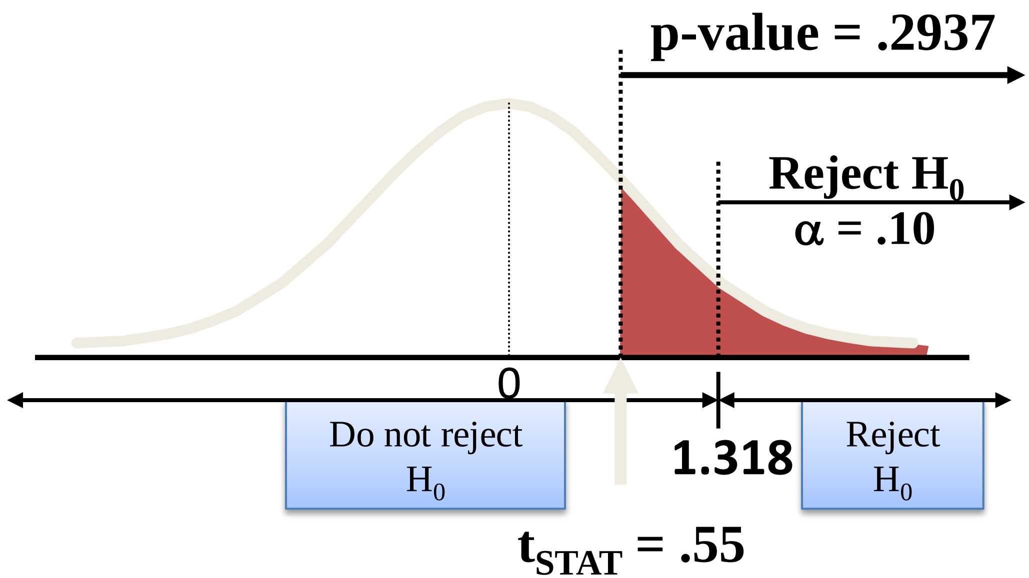


Do not reject H_0 since $t_{STAT} = 0.55 \leq 1.318$

there is not sufficient evidence that the mean bill is over $52

Example: Utilizing The p-value for The Test

- Calculate the p-value and compare to α (p-value below calculated using excel spreadsheet on next page)



Do not reject H_0 since p-value = .2937 > α = .10

Z TEST OF HYPOTHESIS FOR PROPORTIONS

Hypothesis Tests for Proportions

- Involves categorical variables
- Two possible outcomes
  - Possesses characteristic of interest
  - Does not possess characteristic of interest
- Fraction or proportion of the population in the category of interest is denoted by π

Proportions

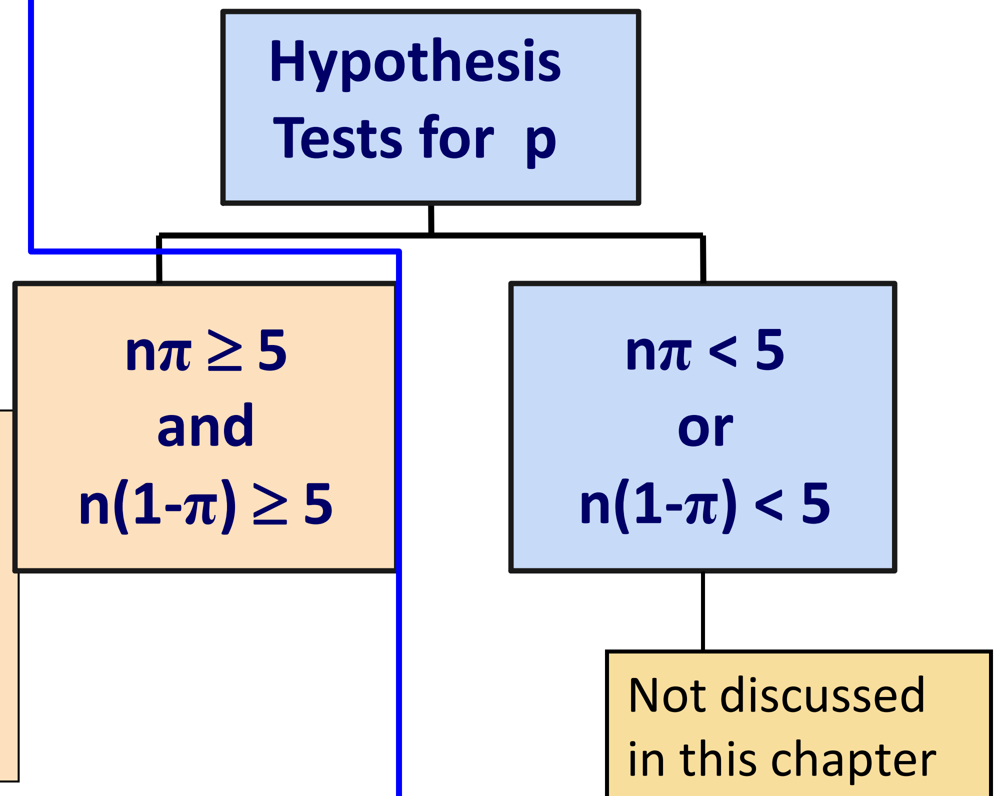
(continued)

- Sample proportion in the category of interest is denoted by p
  - $$p = \frac{X}{n} = \frac{\text{number in category of interest in sample}}{\text{sample size}}$$
- 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
  - $$\mu_p = \pi \qquad \sigma_p = \sqrt{\frac{\pi(1-\pi)}{n}}$$

Hypothesis Tests for Proportions

- The sampling distribution of p is approximately normal, so the test statistic is a Z_{STAT} value:

$$Z_{STAT} = \frac{p - \pi}{\sqrt{\frac{\pi(1-\pi)}{n}}}$$



Z Test for Proportion in Terms of Number in Category of Interest

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

Hypothesis Tests for X

$X \geq 5$ and $n\text{-}X \geq 5$

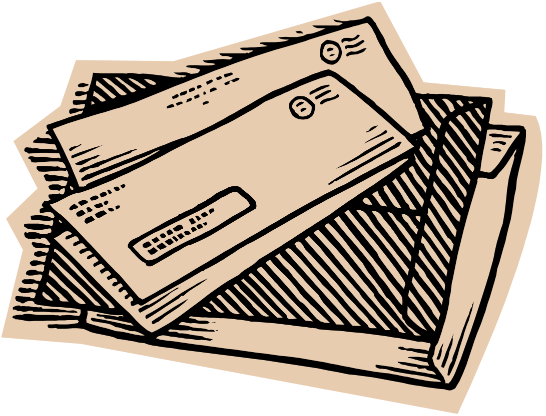
$$Z_{STAT} = \frac{X - n\pi}{\sqrt{n\pi(1-\pi)}}$$

$X < 5$ or $n\text{-}X < 5$

Not discussed in this chapter

Example: Z Test for Proportion

A marketing company 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 α = 0.05 significance level.



Check:

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

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

✓

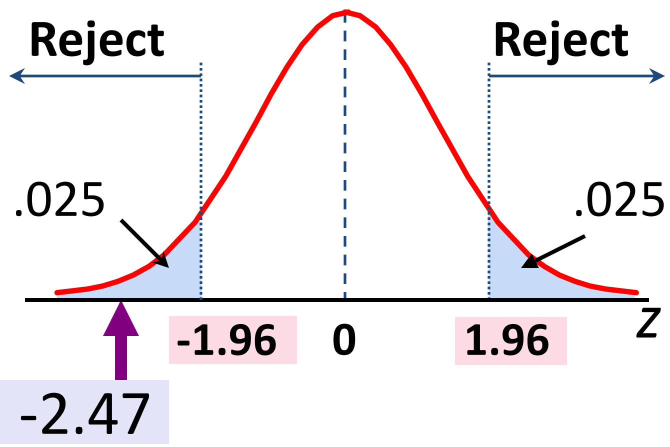
Z Test for Proportion: Solution

H_0: π = 0.08 H_1: $\pi \neq$ 0.08

α = 0.05

n = 500, p = 0.05

Critical Values: ± 1.96



Test Statistic:

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

Decision:

Reject H_0 at α = 0.05

Conclusion:

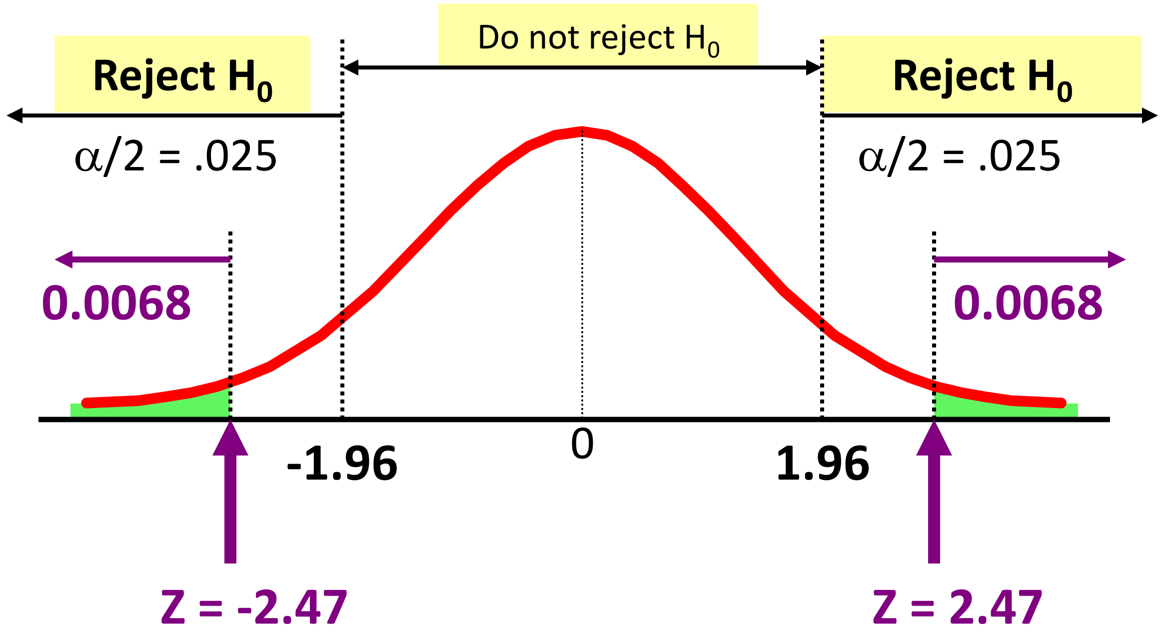
There is sufficient evidence to reject the company's claim of 8% response rate.

p-Value Solution

(continued)

Calculate the p-value and compare to α

(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 < α = 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, α, and the type of test (one-tail or two-tail) before data collection
- Report all pertinent findings including both statistical significance and practical importance

EXERCISE

9.16

The manager of a paint supply store wants to determine whether the mean amount of paint contained in 1-gallon cans purchased from a nationally known manufacturer is actually 1 gallon. You know from the manufacturer's specifications that the standard deviation of the amount of paint is 0.02 gallon. You select a random sample of 50 cans, and the mean amount of paint per 1-gallon can is 0.995 gallon. Is there evidence that the mean amount is different from 1.0 gallon? (Use $\alpha = 0.01$)

9.26

A stationery store wants to estimate the mean retail value of greeting cards that it has in its inventory. A random sample of 100 greeting cards indicates a mean value of \$2.55 and a standard deviation of \$0.44. Is there evidence that the population mean retail value of the greeting cards is different from \$2.50? (Use a 0.05 level of significance.)

8.46

In a recent year, the Federal Communications Commission reported that the mean wait for repairs for Verizon customers was 36.5 hours. In an effort to improve this service, suppose that a new repair service process was developed. This new process, used for a sample of 100 repairs, resulted in a sample mean of 34.5 hours and a sample standard deviation of 11.7 hours. Is there evidence that the population mean amount is less than 36.5 hours? (Use a 0.05 level of significance.)

9.50 (cont'd)

The per-store daily customer count (i.e., the mean number of customers in a store in one day) for a nationwide convenience store chain that operates nearly 10,000 stores has been steady, at 900, for some time. To increase the customer count, the chain is considering cutting prices for coffee beverages by approximately half. The small size will now be \$0.59 instead of \$0.99, and the medium size will be \$0.69 instead of \$1.19. Even with this reduction in price, the chain will have a 40% gross margin on coffee.

9.50

To test the new initiative, the chain has reduced coffee prices in a sample of 34 stores, where customer counts have been running almost exactly at the national average of 900. After four weeks, the sample stores stabilize at a mean customer count of 974 and a standard deviation of 96. This increase seems like a substantial amount to you, but it also seems like a pretty small sample. Do you think reducing coffee prices is a good strategy for increasing the mean customer count? At the 0.01 level of significance, is there evidence that reducing coffee prices is a good strategy for increasing the mean customer count?

9.58

Of 1,000 respondents aged 24 to 35, 65% reported that they preferred to “look for a job in a place where I would like to live” rather than “look for the best job I can find, the place where I live is secondary.” At the 0.05 level of significance, is there evidence that the proportion of all young jobseekers aged 24 to 35 who preferred to “look for a job in a place where I would like to live” rather than “look for the best job I can find, the place where I live is secondary” is different from 60%?

9.59 (cont’d)

The telephone company wants to investigate the desirability of beginning a marketing campaign that would offer customers the right to purchase an additional telephone line at a substantially reduced installation cost. The campaign will be initiated if there is evidence that more than 20% of the customers would consider purchasing an additional telephone line if it were made available at a substantially reduced installation cost. A random sample of 500 households is selected. The results indicate that 135 of the households would purchase the additional telephone line at a reduced installation cost.

9.59

a. At the 0.05 level of significance, is there evidence that more than 20% of the customers would purchase the additional telephone line?
b. How would the manager in charge of promotional programs concerning residential customers use the results in (a)?

9.72 (cont'd)

The owner of a gasoline station wants to study gasoline purchasing habits of motorists at his station. He selects a random sample of 60 motorists during a certain week, with the following results:

- The amount purchased was $\bar{X}$= 11.3 gallons, S = 3.1 gallons.
- Eleven motorists purchased premium-grade gasoline.

9.72

a. At the 0.05 level of significance, is there evidence that the population mean purchase was different from 10 gallons?
b. At the 0.05 level of significance, is there evidence that less than 20% of all the motorists at the station purchased premium-grade gasoline?
c. What is your answer to (a) if the sample mean equals 10.3 gallons?
d. What is your answer to (b) if 7 motorists purchased premium-grade gasoline?

THANK YOU