Statistical Analysis in Fin Mkts

MSF 502

Li Cai

CHAPTER 9 Hypothesis Testing

Chapter 9 Learning Objectives (LOs)

LO 9.1: **Define the null hypothesis and the alternative hypothesis.**

LO 9.2: **Distinguish between Type I and Type II errors.**

LO 9.3: **Explain the steps of a hypothesis test using the *p*-value approach.**

LO 9.4: **Explain the steps of a hypothesis test using the critical value approach.**

LO 9.5: **Differentiate between the test statistics for the population mean.**

LO 9.6: **Specify the test statistic for the population proportion.**

Undergraduate Study Habits

- Are today's college students studying hard or hardly studying?
- A recent study asserts that over the past five decades the number of hours that the average college student studies each week has been steadily dropping (*The Boston Globe,* July 4, 2010).
- In 1961, students invested 24 hours per week in their academic pursuits, whereas today's students study an average of 14 hours per week.

Undergraduate Study Habits

- As dean of a large university in California, Susan Knight wonders if the study trend is reflective of students at her university.
- Susan randomly selected 35 students to ask about their average study time per week. Using these results, Susan wants to
  1. Determine if the mean study time of students at her university is below the 1961 national average of 24 hours per week.
  2. Determine if the mean study time of students at her university differs from today's national average of 14 hours per week.

9.1 Point Estimators and Their Properties

LO 9.1 Define the null hypothesis and the alternative hypothesis.

- **Hypothesis tests resolve conflicts between two competing opinions (hypotheses).**
- **In a hypothesis test, define**
  - **H_0, the null hypothesis, the presumed default state of nature or status quo.**
  - **H_A, the alternative hypothesis, a contradiction of the default state of nature or status quo.**

LO 9.1

9.1 Point Estimators and Their Properties

- In statistics we use sample information to make inferences regarding the unknown population parameters of interest.
- We conduct hypothesis tests to determine if sample evidence contradicts H_0.
- On the basis of sample information, we either
  - "Reject the null hypothesis"
    - Sample evidence *is* inconsistent with H_0.
  - "Do not reject the null hypothesis"
    - Sample evidence *is not* inconsistent with H_0.
    - We do not have enough evidence to "accept" H_0.

LO 9.1

9.1 Point Estimators and Their Properties

- **Defining the Null Hypothesis and Alternative Hypothesis**

- **General guidelines:**
  - **Null hypothesis, H_0, states the status quo.**
  - **Alternative hypothesis, H_A, states whatever we wish to establish (i.e., contests the status quo).**
  - **Use the following signs in hypothesis tests**

H_0	$=$	$\geq$	$\leq$
H_A	$\neq$	$<$	$>$

← **specify the status quo,**

← **contradict H_0.**

  - **Note that H_0 always contains the "equality."**

LO 9.1

9.1 Point Estimators and Their Properties

- **One-Tailed versus Two-Tailed Hypothesis Tests**
  - **Two-Tailed Test**
    - **Reject H_0 on either side of the hypothesized value of the population parameter.**
    - **For example:**

$$H_0: \mu = \mu_0 \text{ versus } H_A: \mu \neq \mu_0$$
$$H_0: p = p_0 \text{ versus } H_A: p \neq p_0$$

  - **The "≠" symbol in H_A indicates that both tail areas of the distribution will be used to make the decision regarding the rejection of H_0.**

LO 9.1

9.1 Point Estimators and Their Properties

- **One-Tailed versus Two-Tailed Hypothesis Tests**

  - **One-Tailed Test**
    - **Reject H_0 only on one side of the hypothesized value of the population parameter.**
    - **For example:**

      $H_0: \mu \leq \mu_0$ **versus** $H_A: \mu > \mu_0$ **(right-tail test)**

      $H_0: \mu \geq \mu_0$ **versus** $H_A: \mu < \mu_0$ **(left-tail test)**

  - **Note that the inequality in H_A determines which tail area will be used to make the decision regarding the rejection of H_0.**

LO 9.1

9.1 Point Estimators and Their Properties

- **Three Steps to Formulate Hypotheses**

1. **Identify the relevant population parameter of interest (e.g., μ or p).**
2. **Determine whether it is a one- or a two-tailed test.**

H_0	H_A	Test Type
$=$	$\neq$	Two-tail
$\geq$	$<$	One-tail, Left-tail
$\leq$	$>$	One-tail, Right-tail

3. **Include some form of the equality sign in H_0 and use H_A to establish a claim.**

LO 9.1

9.1 Point Estimators and Their Properties

- **Example: A trade group predicts that back-to-school spending will average \$606.40 per family this year. A different economic model is needed if the prediction is wrong.**
  1. **Parameter of interest is μ since we are interested in the average back-to-school spending.**
  2. **Since we want to determine if the population mean differs from \$606.4 (i.e, ≠), it is a two-tail test.**
  3. **H_0: $\mu = 606.4$**
     H_A: $\mu \neq 606.4$

LO 9.1

9.1 Point Estimators and Their Properties

- **Example: A television research analyst wishes to test a claim that more than 50% of the households will tune in for a TV episode. Specify the null and the alternative hypotheses to test the claim.**

1. **Parameter of interest is p since we are interested in the proportion of households.**
2. **Since the analyst wants to determine whether $p > 0.50$, it is a one-tail test.**
3. $H_0: p \leq 0.50$
   $H_A: p > 0.50$

9.1 Point Estimators and Their Properties

LO 9.2 Distinguish between Type I and Type II errors.

- **Type I and Type II Errors**
  - **Type I Error: Committed when we reject H_0 when H_0 is actually true.**
    - **Occurs with probability α. α is chosen *a priori*.**
  - **Type II Error: Committed when we do not reject H_0 and H_0 is actually false.**
    - **Occurs with probability β. Power of the test = $1-\beta$**
  - **For a given sample size *n*, a decrease in α will increase β and vice versa.**
  - **Both α and β decrease as *n* increases.**

LO 9.2

9.1 Point Estimators and Their Properties

- **This table illustrates the decisions that may be made when hypothesis testing:**

Decision	Null hypothesis is true	Null hypothesis is false
Reject the null hypothesis	Type I error	Correct decision
Do not reject the null hypothesis	Correct decision	Type II error

- **Correct Decisions:**
  - **Reject H_0 when H_0 is false.**
  - **Do not reject H_0 when H_0 is true.**
- **Incorrect Decisions:**
  - **Reject H_0 when H_0 is true (Type I Error).**
  - **Do not reject H_0 when H_0 is false (Type II Error).**

LO 9.2

9.1 Point Estimators and Their Properties

- **Example: Consider the following competing hypotheses that relate to the court of law.**
  - ***H_0:* An accused person is innocent**
    ***H_A:* An accused person is guilty**
- **Consequences of Type I and Type II errors:**
  - **Type I error: Conclude that the accused is guilty when in reality, she is innocent.**
  - **Type II error: Conclude that the accused is innocent when in reality, she is guilty.**

9.2 Hypothesis Test of the Population Mean When σ Is Known

LO 9.3 Explain the steps of a hypothesis test using the *p*-value approach.

- **Hypothesis testing enables us to determine whether the sample evidence is inconsistent with what is hypothesized under the null hypothesis (H_0).**
- **Basic principle: First assume that H_0 is true and then determine if sample evidence contradicts this assumption.**
- **Two approaches to hypothesis testing:**
  - **The *p*-value approach.**
  - **The critical value approach.**

LO 9.3

9.2 Hypothesis Test of the Population Mean When σ Is Known

- **The *p*-value Approach**
  - **The value of the test statistic for the hypothesis test of the population mean μ when the population standard deviation σ is known is computed as**

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

    where μ_0 is the hypothesized mean value.
  - ***p*-value: the likelihood of obtaining a sample mean that is at least as extreme as the one derived from the given sample, under the assumption that the null hypothesis is true**

LO 9.3

9.2 Hypothesis Test of the Population Mean When σ Is Known

- **The *p*-value Approach**
  - **Under the assumption that $\mu = \mu_0$, the *p*-value is the likelihood of observing a sample mean that is at least as extreme as the one derived from the given sample.**
  - **The calculation of the *p*-value depends on the**

Alternative Hypothesis	*p*-value
H_A: $\mu > \mu_0$	Right-tail probability: $P(Z \geq z)$
H_A: $\mu < \mu_0$	Left-tail probability: $P(Z \leq z)$
H_A: $\mu \neq \mu_0$	Two-tail probability: $2P(Z \geq z)$ if $z > 0$ or $2P(Z \leq z)$ if $z < 0$

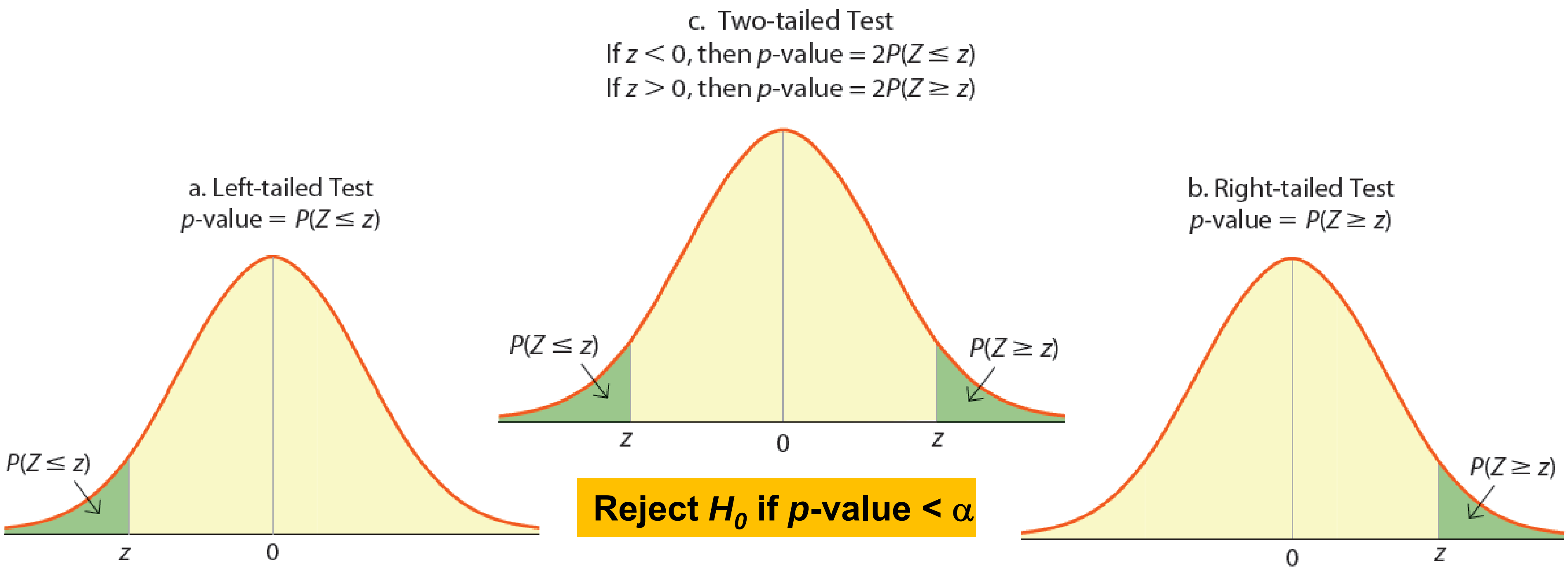
  - **Decision rule: Reject H_0 if *p*-value < α.**

LO 9.3

9.2 Hypothesis Test of the Population Mean When σ Is Known

The *p*-value Approach

- **Determining the *p*-value depending on the specification of the competing hypotheses.**



Reject H_0 if p-value $< \alpha$

LO 9.3

9.2 Hypothesis Test of the Population Mean When σ Is Known

- **Four Step Procedure Using The *p*-value Approach**
  - **Step 1. Specify the null and the alternative hypotheses.**
  - **Step 2. Specify the test statistic and compute its value.**
  - **Step 3. Calculate the p-value.**
  - **Step 4. State the conclusion and interpret the results.**

LO 9.3

9.2 Hypothesis Test of the Population Mean When σ Is Known

Example: The *p*-value Approach

- **Consider the following:** $n = 25,\ \bar{x} = 71,\ \sigma = 9$
- **Step 1. State the hypotheses:** $H_0: \mu \leq 67$

  $H_A: \mu > 67$

  Thus, $\mu_0 = 67$
- **Step 2. Given that the population is normally distributed with a known standard deviation,**

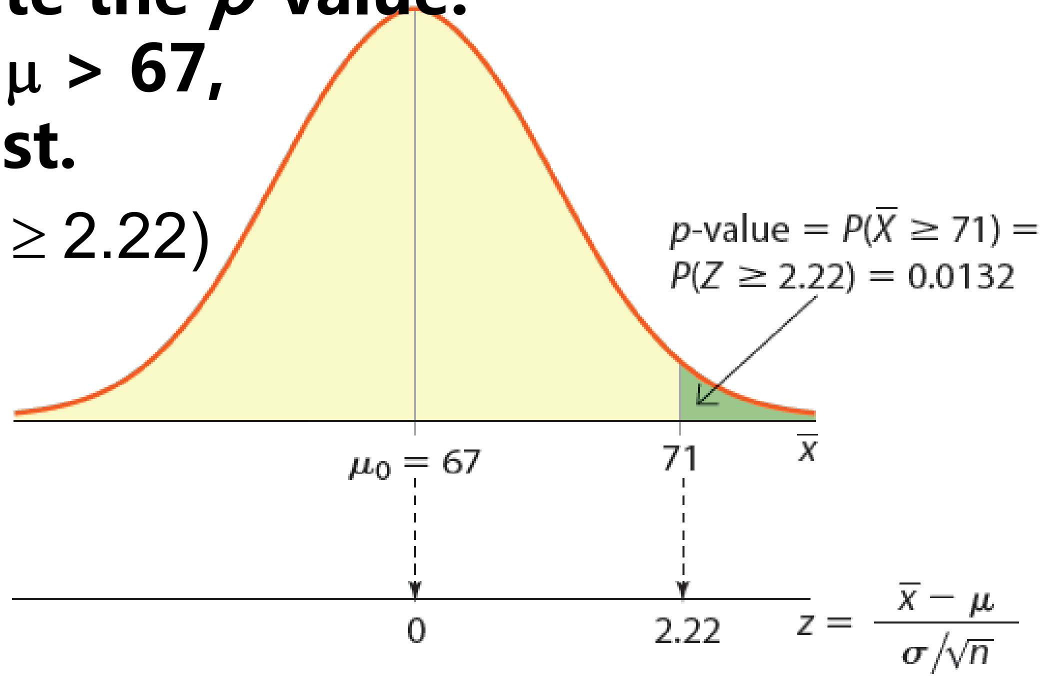
  $\sigma = 9$, we c … **f the test statistic as**

$$z = \frac{\bar{x} - \mu_0}{\sigma / \sqrt{n}} = \frac{71 - 67}{9 / \sqrt{25}} = 2.22$$

LO 9.3

9.2 Hypothesis Test of the Population Mean When σ Is Known

- **Example: The *p*-value Approach**
  - **Unstandardized Normal Distribution:** $\bar{x} = 71 \qquad \mu_0 = 67$
    Standardized Normal Distribution: $z = 2.22 \qquad \mu = 0$
  - **Step 3. Now compute the *p*-value:**
    Note that since H_A: μ > 67,
    this is a right-tail test.
  - **Thus,** $P(\bar{X} \geq 71) = P(Z \geq 2.22)$
    $= 1 - 0.9868$
    $= 0.0132$
  - ***p*-value = 0.0132**
    or 1.32%



LO 9.3

9.2 Hypothesis Test of the Population Mean When σ Is Known

- **Example: The *p*-value Approach**
  - ***p*-value = 0.0132 or 1.32%**
  - **Typically, before implementing a hypothesis test, we choose a value for α = 0.01, 0.05, or 0.1 and reject H_0 when the *p*-value < α.**
  - **Let's say, before conducting the study, we chose α = 0.05.**
  - **Step 4. Since *p*-value = 0.0132 < α = 0.05, we reject H_0 and conclude that the sample data support the alternative claim that μ > 67.**

9.2 Hypothesis Test of the Population Mean When σ Is Known

LO 9.4 Explain the steps of a hypothesis test using the critical value approach.

- **The Critical Value Approach**
  - **Rejection region: a region of values such that if the test statistic falls into this region, then we reject H_0.**
    - **The location of this region is determined by H_A.**
  - **Critical value: a point that separates the rejection region from the nonrejection region.**

LO 9.4

9.2 Hypothesis Test of the Population Mean When σ Is Known

The Critical Value Approach

- The critical value approach specifies a region such that if the value of the test statistic falls into the region, the null hypothesis is rejected.
- The critical value depends on the alternative

Alternative Hypothesis	Critical Value
$H_A: \mu > \mu_0$	Right-tailed critical value is z_α, where $P(Z \geq z_\alpha) = \alpha$.
$H_A: \mu < \mu_0$	Left-tailed critical value is $-z_\alpha$, where $P(Z \leq -z_\alpha) = \alpha$.
$H_A: \mu \neq \mu_0$	Two-tailed critical values $-z_{\alpha/2}$ and $z_{\alpha/2}$, where $P(Z \geq z_{\alpha/2}) = \alpha/2$.

- Decision Rule: Reject H_0 if:

  $z > z_\alpha$ for a right-tailed test

  $z < -z_\alpha$ for a left-tailed test

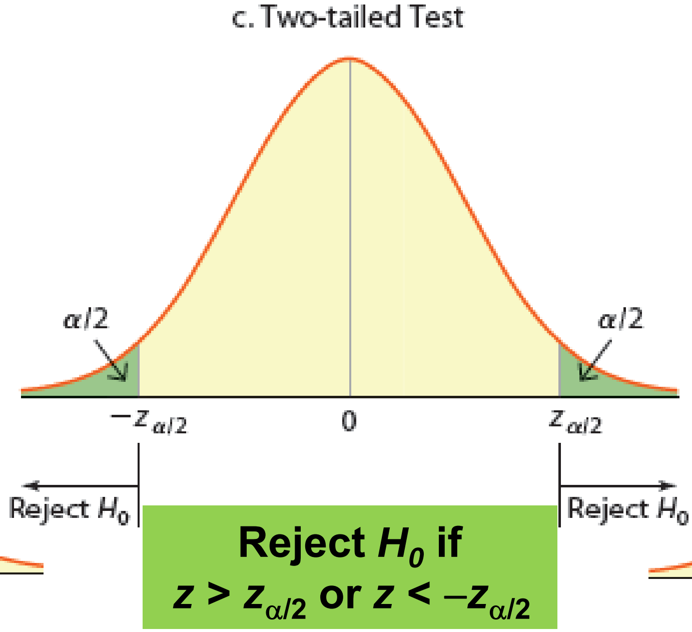
  $z > z_{\alpha/2}$ or $z < -z_{\alpha/2}$ for a two-tailed test

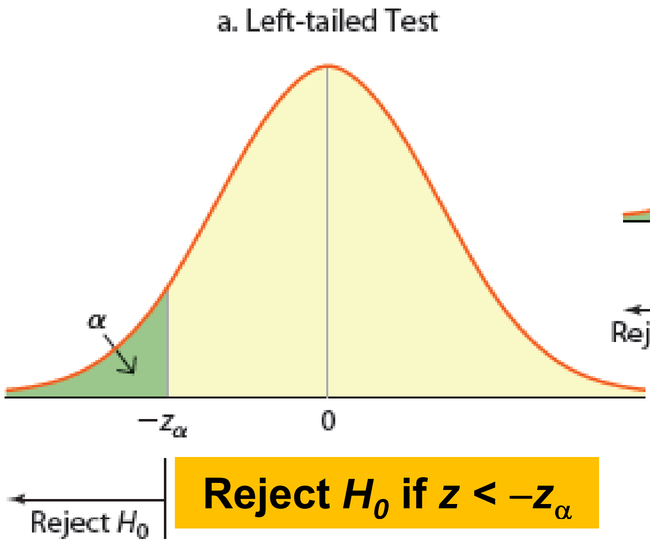
LO 9.4

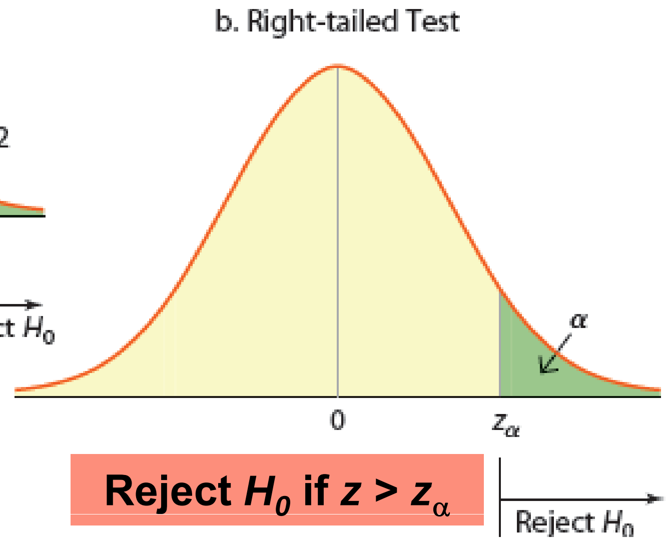
9.2 Hypothesis Test of the Population Mean When σ Is Known

The Critical Value Approach

- Determining the critical value(s) depending on the specification of the competing hypothese







LO 9.4

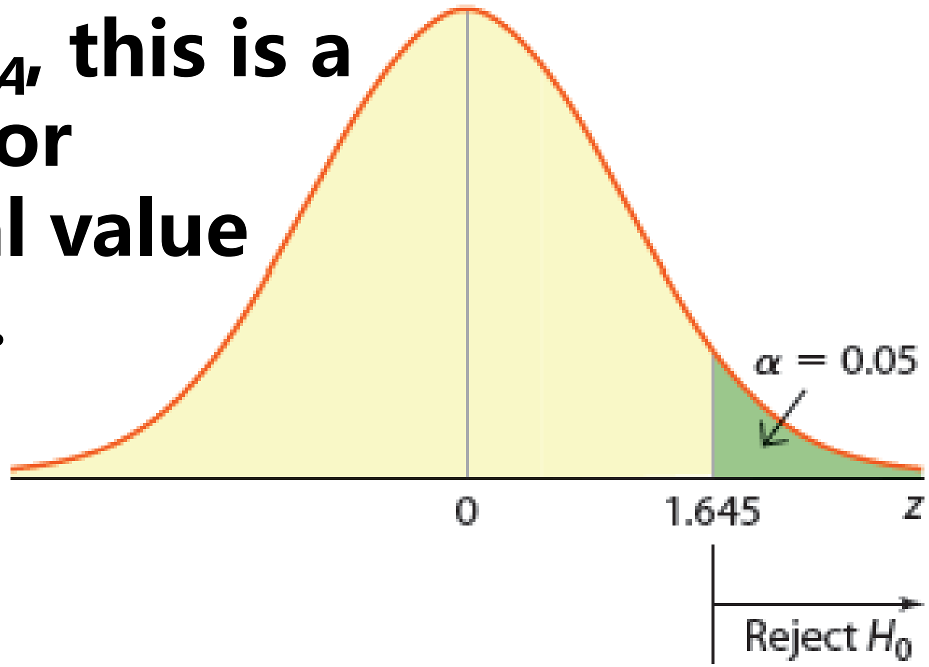
9.2 Hypothesis Test of the Population Mean When σ Is Known

- **Four Step Procedure Using the Critical Value Approach**
  - **Step 1. Specify the null and the alternative hypotheses.**
  - **Step 2. Specify the test statistic and compute its value.**
  - **Step 3. Find the critical value *or* values.**
  - **Step 4. State the conclusion and interpret the results.**

LO 9.4

9.2 Hypothesis Test of the Population Mean When σ Is Known

- **Example: The Critical Value Approach**
  - **Step 1.** H_0: $\mu \leq 67$, H_A: $\mu > 67$
  - **Step 2. From previous example,** $z = 2.22$
  - **Step 3. Based on** H_A**, this is a right-tail test and for** $\alpha = 0.05$**, the critical value is** $z_\alpha = z_{0.05} = 1.645$**.**

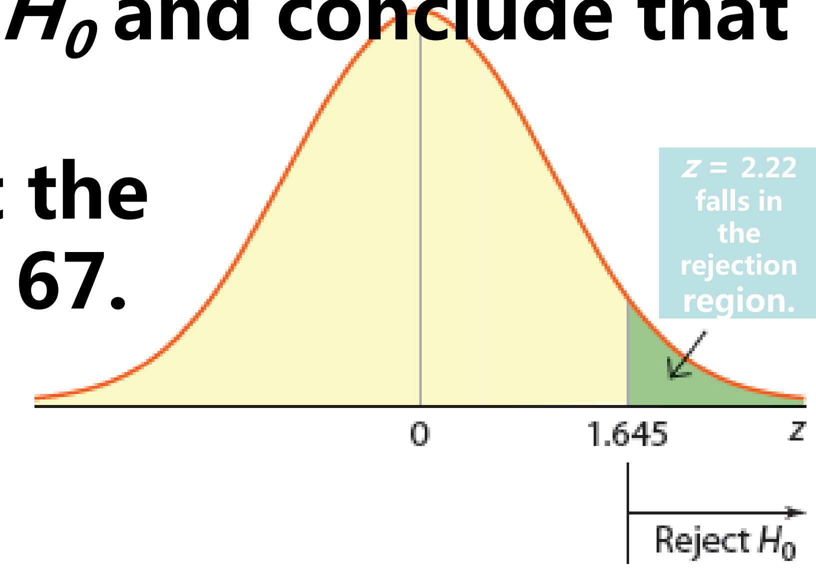


LO 9.4

9.2 Hypothesis Test of the Population Mean When σ Is Known

Example: The Critical Value Approach

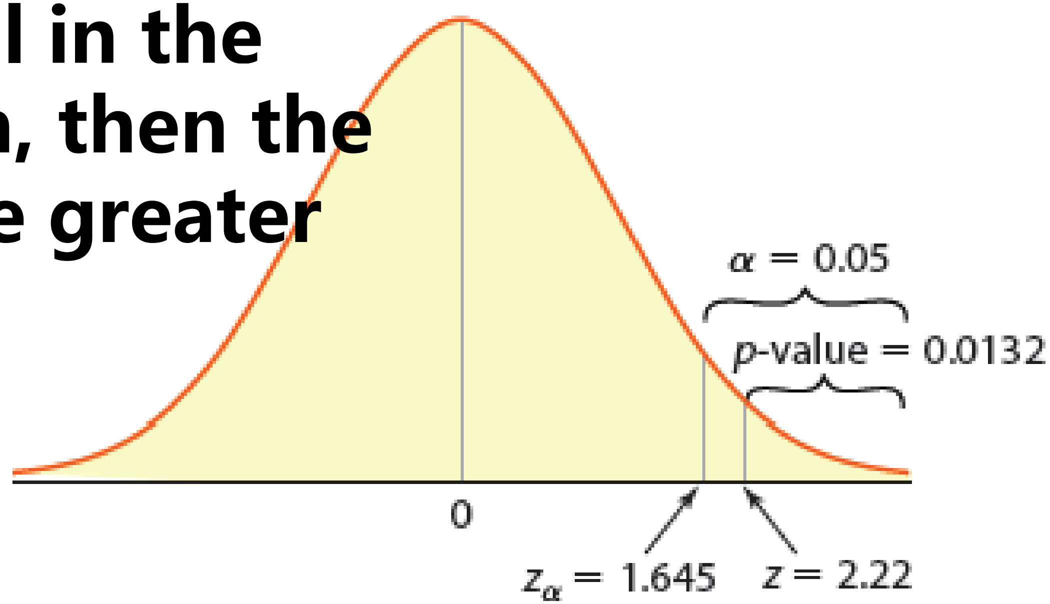
- **Step 4. Reject H_0 if $z > 1.645$.**
  - **Since $z = 2.22 > z_\alpha = 1.645$, the test statistic falls in the rejection region. Therefore, we reject H_0 and conclude that the**
    sample data support the alternative claim $\mu > 67$.
- **This conclusion is the same as that from the *p-value* approach.**



LO 9.4

9.2 Hypothesis Test of the Population Mean When σ Is Known

- **Example: The Critical Value Approach**
  - **If *z* falls in the rejection region, then the *p*-value must be less than α.**
  - **If *z* does not fall in the rejection region, then the *p*-value must be greater than α.**



LO 9.4

9.2 Hypothesis Test of the Population Mean When σ Is Known

- **Confidence Intervals and Two-Tailed Hypothesis Tests**
  - **Given the significance level α, we can use the sample data to construct a $100(1 - \alpha)\%$ confidence interval for the population mean μ.**
  - **Decision Rule**
    - **Reject H_0 if the confidence interval *does not* contain the value of the hypothesized mean μ_0.**
    - **Do not reject H_0 if the confidence interval *does* contain the value of the hypothesized**

LO 9.4

9.2 Hypothesis Test of the Population Mean When σ Is Known

- **Implementing a Two-Tailed Test Using a Confidence Interval**
  - **The general specification for a 100(1 – α)% confidence interval of the population mean μ when the population standard deviation σ is known is computed as**

$$\bar{x} \pm z_{\alpha/2}\,\sigma/\sqrt{n}$$

or

$$\left[\bar{x} - z_{\alpha/2}\,\sigma/\sqrt{n},\ \bar{x} + z_{\alpha/2}\,\sigma/\sqrt{n}\right]$$

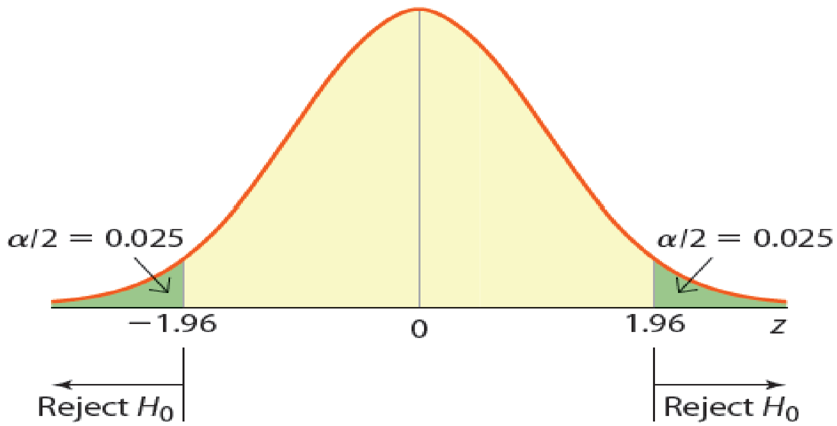
  - **Decision rule: Reject H_0 if** $\mu_0 < \bar{x} - z_{\alpha/2}\,\sigma/\sqrt{n}$

or if $\mu_0 > \bar{x} + z_{\alpha/2}\,\sigma/\sqrt{n}$

LO 9.4

9.2 Hypothesis Test of the Population Mean When σ Is Known

- **Example: Recall that a research analyst wishes to determine if average back-to-school spending differs from $606.40.**
  - **Out of 30 randomly drawn households from a normally distributed population, the standard deviation is $65 and sample mean is $622.85.**
  - **Step 1.** $H_0: \mu = 606.4$, $H_A: \mu \neq 606.4$
  - **Step 2.** $z = 1.39$
  - **Step 3. Based on** H_A**, this is a two-tail test and for** $\alpha = 0.05$**, the critical value is** $z_{\alpha/2} = z_{0.025} = \pm 1.96$**.**



9.3 Hypothesis Test of the Population Mean When σ Is Unknown

LO 9.5 Differentiate between the test statistics for the population

- **Test Statistic for μ When σ is Unknown**
  - **When the population standard deviation σ is unknown, the test statistic for testing the population mean μ is assumed to follow the t_{df} distribution with (n – 1) degrees of freedom (df).**
  - **The value of t_{df} is computed as** $t_{df} = \dfrac{\bar{x} - \mu_0}{s/\sqrt{n}}$

LO 9.5

9.3 Hypothesis Test of the Population Mean When σ Is Unknown

- **Example**
  - **Consider the following:** $n = 35,\ \bar{x} = 16.37,\ s = 7.22$
  - **Step 1. State the hypotheses:** $H_0 : \mu \geq 24$

    $H_A : \mu < 24$

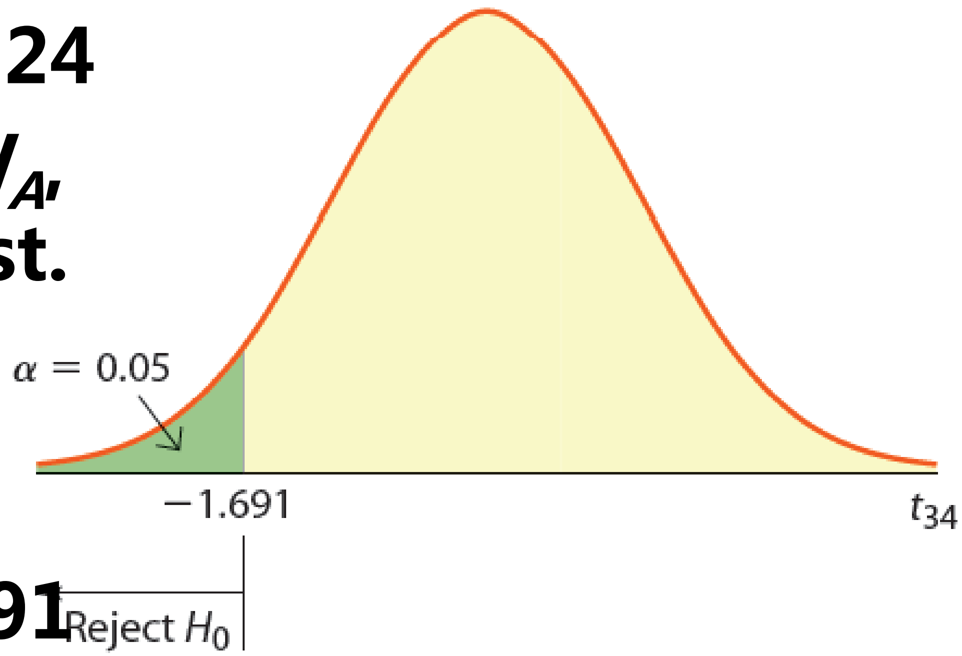
    Thus, μ_0 = 24
  - **Step 2. Because *n* = 35 (i.e, n > 30), we can assume that the sample mean is normally distributed and thus compute the value of the test statistic:**

$$t_{34} = \frac{\bar{x} - \mu_0}{s/\sqrt{n}} = \frac{16.37 - 24}{7.22/\sqrt{35}} = -6.25$$

LO 9.5

9.3 Hypothesis Test of the Population Mean When σ Is Unknown

- **Example: The Critical Value Approach**
  - $n = 35,\ \bar{x} = 16.37,\ s = 7.22,\ t_{34} = -6.25$
  - H_0: $\mu \geq 24$, H_A: $\mu < 24$
  - **Step 3. Based on H_A, this is a left-tail test. For α = 0.05 and n–1 = 34 df, the critical value is $t_{\alpha,df} = t_{0.05,34} = 1.691$ (–1.691 due to symmetry).**



LO 9.5

9.3 Hypothesis Test of the Population Mean When σ Is Unknown

- **Example: The Critical Value Approach**
  - **Step 4. State the conclusion and interpret the results.**
    - **Reject H_0 if $t_{34} < -t_{0.05,34} = -1.691$.**
    - **Since $t_{34} = -6.25$ is less than $t_{0.05,34} = -1.691$, we reject H_0 and conclude that the sample data support the alternative claim that $\mu < 24$.**

LO 9.5

9.3 Hypothesis Test of the Population Mean When σ Is Unknown

- **Example: The *p*-value Approach**
  - $n = 35, \ \bar{x} = 16.37, \ s = 7.22$
  - **Step 1. H_0: μ = 14, H_A: μ ≠ 14**
  - **Step 2. Compute the value of the test statistic.**

$$t_{34} = \frac{\bar{x} - \mu_0}{s/\sqrt{n}} = \frac{16.37 - 14}{7.22/\sqrt{35}} = 1.94$$

  - **Step 3. Compute the *p*-value.**
    - **Since t_{34}=1.94 > 0, the *p*-value for a two-tailed test is 2P($T_{34} \geq t_{34}$). Referencing the t_{df} table for *df* = 34, we find that the exact probability P($T_{34} \geq$ 1.94) cannot be determined.**

LO 9.5

9.3 Hypothesis Test of the Population Mean When σ Is Unknown

- **Example: The *p*-value Approach**
  - **Step 3. Compute the *p*-value (continued).**
    - **Look up t_{34} = 1.94 in the t-table to find the p-**

	Area in Upper Tail, α					
df	0.20	0.10	0.05	0.025	0.01	0.005
1	1.376	3.078	6.341	12.706	31.821	63.657
⋮	⋮	⋮	⋮	⋮	⋮	⋮
34	0.852	1.307	1.691	2.032	2.441	2.728

    - **Note that t_{34} = 1.94 lies between 1.691 and 2.032.**
    - **Thus, 0.025 < $P(T_{34} \geq 1.94)$ < 0.05. However, because this is a two-tail test, we multiply by two to get 0.05 < *p*-value < 0.10.**

LO 9.5

9.3 Hypothesis Test of the Population Mean When σ Is Unknown

- **Example: The *p*-value Approach**
  - **0.05 < *p*-value < 0.10**
  - α **= 0.05.**
  - **Step 4. State the conclusion and interpret the results.**
    - **Since the *p*-value satisfies 0.05 < *p*-value < 0.10, it must be greater than α = 0.05.**
    - **Thus, we do not reject H_0 and conclude that the mean study time of students at the university is not statistically different from today's national average of 14 hours per week.**

9.4 Hypothesis Test of the Population Proportion

LO 9.6 Specify the test statistic for the population

- **Test Statistic for *p*.**
  - $\bar{P}$ **can be approximated by a normal distribution if $np \geq 5$ and $n(1-p) \geq 5$.**
  - **Test statistic for the hypothesis test of the population proportion *p* is assumed to follow the *z* distribution:**

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

where $\bar{p} = x/n$ and p_0 is the hypothesized value of the population proportion.

LO 9.6

9.4 Hypothesis Test of the Population Proportion

- **Example:**
  - $n = 180,\ x = 67,\ p_0 = 0.4$
  - **Step 1. H_0: $p \geq 0.4$, H_A: $p < 0.4$**
  - **Step 2. Compute the value of the test statistic.**
    - **First verify that the sample is large enough:**

      $np_0 = 67 \times 0.4 = 26.8 > 5$

      $n(1 - p_0) = 67 \times 0.6 = 40.2 > 5$

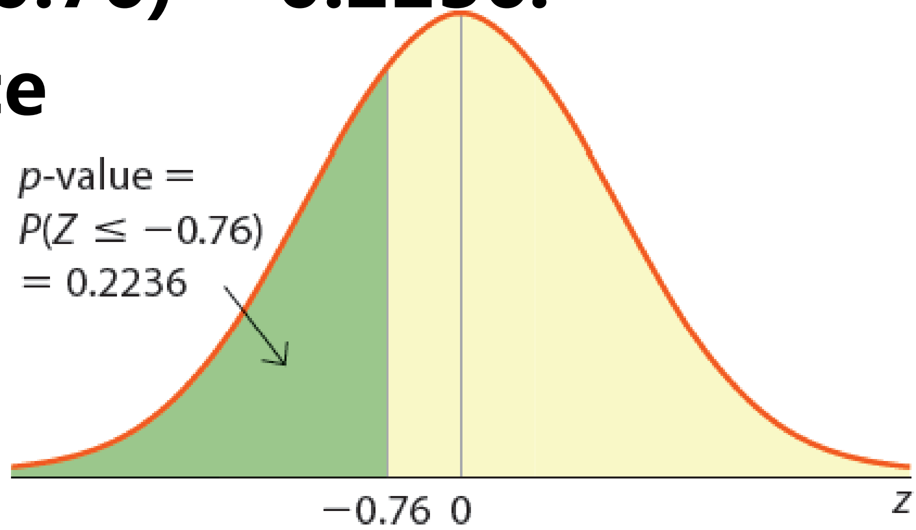
      $\bar{p}$
    - **Compu 0.3722**

$$z = \frac{\bar{p} - p_0}{\sqrt{p_0(1-p_0)/n}} = \frac{0.3722 - 0.4}{\sqrt{0.4(1-0.4)/180}} = -0.76$$

LO 9.6

9.4 Hypothesis Test of the Population Proportion

- **Example:**
  - **Step 3. Compute the *p*-value.**
    - **Based on H_A: $p < 0.4$, this is a left-tailed test. Compute the *p*-value as: $P(Z \leq z) = P(Z \leq -0.76) = 0.2236$.**
    - **Let the significance level $\alpha = 0.10$.**



LO 9.6

9.4 Hypothesis Test of the Population Proportion

- **Example:**
  - **Step 4. State the conclusion and interpret the results.**
    - **p-value = 0.2236 > α = 0.10.**
    - **Do not reject H_0: $p \geq 0.4$ and conclude H_A: $p < 0.4$.**
    - **Thus, the magazine's claim that fewer than 40% of households in the United States have changed their lifestyles because of escalating gas prices is not justified by the sample data.**

End of Chapter