Chapter 9 Hypothesis Testing

Overview

Hypothesis testing is a statistical procedure used to provide evidence in favor of some statement (called a *hypothesis*). For instance, hypothesis testing might be used to assess whether a population parameter, such as a population mean, differs from a specified standard or previous value. In this chapter we discuss testing hypotheses about population means and proportions.

In order to illustrate how hypothesis testing works, we revisit several cases introduced in previous chapters and also introduce some new cases:

The e-billing Case: The consulting firm uses hypothesis testing to provide strong evidence that the new electronic billing system has reduced the mean payment time by more than 50 percent.

The Cheese Spread Case: The cheese spread producer uses hypothesis testing to supply extremely strong evidence that fewer than 10 percent of all current purchasers would stop buying the cheese spread if the new spout were used.

The Commercial Loan Case: The bank uses hypothesis testing to provide strong evidence that the mean debt-to-equity ratio for its portfolio of commercial loans is less than 1.5.

The Trash Bag Case: A marketer of trash bags uses hypothesis testing to support its claim that the mean breaking strength of its new trash bag is greater than 50 pounds. As a result, a television network approves use of this claim in a commercial.

The Valentine's Day Chocolate Case: A candy company projects that this year's sales of its special valentine box of assorted chocolates will be 10 percent higher than last year. The candy company uses hypothesis testing to assess whether it is reasonable to plan for a 10 percent increase in sales of the valentine box.

Overview

9.1 The Null and Alternative Hypotheses and Errors in Testing

9.2 z Tests about a Population Mean: σ Known

9.3 t Tests about a Population Mean: σ Unknown

9.4 z Tests about a Population Proportion

9.5 Type II Error Probabilities and Sample Size Determination (Optional)

9.6 The Chi-Square Distribution (Optional)

9.7 Statistical Inference for a Population Variance (Optional)

9.1 The Null and Alternative Hypotheses and Errors in Hypothesis Testing

The Null Hypothesis and the Alternative Hypothesis

In hypothesis testing:

1. The **null hypothesis,** denoted H_0, is the statement being tested. Often, the null hypothesis is a statement of "no difference" or "no effect." The null hypothesis is not rejected unless there is convincing sample evidence that it is false.

2. The **alternative,** or **research, hypothesis,** denoted H_a, is a statement that will be accepted only if there is convincing sample evidence that it is true.

What is a Hypothesis?

- In statistics, a ***hypothesis*** is an assumption (claim) about the **population parameter.**
- Population Parameter: such as mean, variance, proportion (percentage of a targeted group), etc.
- Population mean μ . Example 1: The average number of TV sets in US homes is equal to 3.
- Population proportion p. Example 2: 68% of adults in this city have cell phones.

Philosophy behind Hypothesis Testing

- If a person said he never scolded someone. Can he prove it?
- Show the audio and video recordings of every moment he grew up, all the written things, etc., and prove that these physical evidences are complete, true, and uninterrupted.
- **This is impossible!**
- Even if he finds some witnesses, such as his classmates, family members, and colleagues, it can *only prove* that for some moments when those witnesses were present, he was not heard scolding.

Philosophy behind hypothesis testing

- Conversely, if it is easy to prove that this person has scolded someone, it is enough to be caught once.
- It seems that trying to affirm something is difficult, but negating is much easier.
- This is the philosophy behind hypothesis testing.
- **Proof by Contradiction**

Hypothesis Testing

The Null Hypothesis
H_0

VS

What is Hypothesis testing?

The Alternative Hypothesis
H_a

Hypothesis Testing

- **Hypothesis test** provides *a formal procedure* for comparing observed data (sample) with a claim about a population
- The test is designed to assess the strength of the evidence against the **null hypothesis** (H_0)
- Typically, H_0 asserts that there is nothing unusual, nothing unexpected, no effect, no difference, etc.
- **Alternative hypothesis** H_a is specified as another choice

Null hypothesis can *never be proved* by any data!!

Null Hypothesis--H_0

- Begin with the assumption that the null hypothesis is true
  - e.g., The person is presumed to not guilty at the outset of the court trial .
- Is always about population parameter
- Always contains "=" , "≤" or "≥" sign
- May or may not be rejected

Alternative Hypothesis--H_a

- Is the opposite of the null hypothesis
  - e.g., The person is guilty in the court trial.
- Challenges the *existing state*
- Never contains the "=" , "≤" or "≥" sign
- May or may not be proven
- **Is generally the hypothesis that the researcher is trying to prove**

Example 9.1 Trash Bag Case

- Tests show that the current trash bag has a mean breaking strength μ close to but not exceeding 50 lbs
- The null hypothesis H_0 is that the new bag has a mean breaking strength that is 50 lbs or less
- The new bag's mean breaking strength is not known and is in question, but it is hoped that the new bag is stronger than the current one
- The alternative hypothesis H_a is that the new bag has a mean breaking strength that exceeds 50 lbs
- H_0: $\mu \leq 50, H_a: \mu > 50$

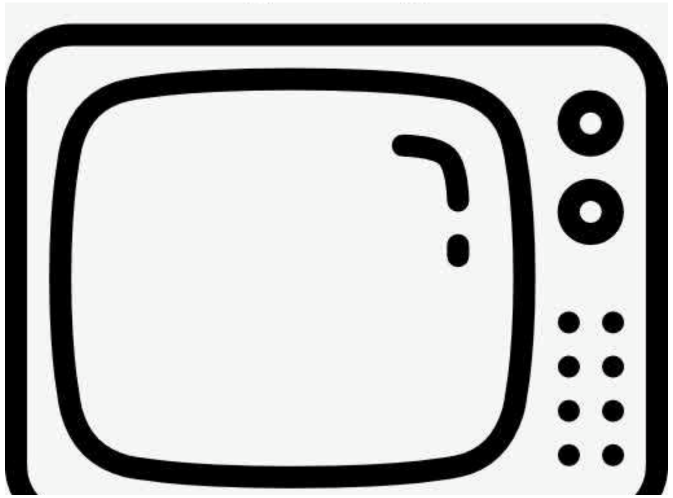
Example 9.2 Payment Time Case

- With a new billing system, the mean bill paying time μ is hoped to be less than 19.5 days
- The alternative hypothesis H_a is that the new billing system has a mean payment time that is less than 19.5 days
- With the old billing system, the mean bill paying time μ was close to but not less than 39 days
- The null hypothesis H_0 is that the new billing system has a mean payment time close to but not less than 19.5 days
- H_0: $\mu \geq 19.5, H_a < 19.5$

Example

- Check if the average number of TV sets in US homes is equal to 3. A random sample of 100 homes showed $\bar{X}$= 2.84. Assume that σ = 0.8 is known.

- H_0: $\mu = 3$, H_a: $\mu \neq 3$



Null and Alternative Hypotheses

- Right-Sided, Upper-Tailed, “Greater Than” Alternative

$$H_0: \mu \leq \mu_0 \quad \text{vs.} \quad H_a: \mu > \mu_0$$

- Left-Sided, lower-Tailed, “Less Than” Alternative

$$H_0 : \mu \geq \mu_0 \quad \text{vs.} \quad H_a : \mu < \mu_0$$

- Two-Sided, Two-Tailed, “Not Equal To” Alternative

$$H_0 : \mu = \mu_0 \quad \text{vs.} \quad H_a : \mu \neq \mu_0$$

where μ_0 is a given constant value (with the appropriate units) that is a comparative value

Types of Decisions

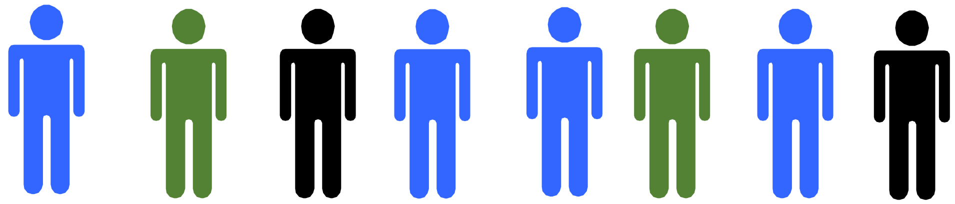
- As a result of testing H_0 vs. H_a, will have to decide either of the following decisions for the null hypothesis H_0:

Do not reject H_0 OR Reject H_0

Hypothesis Testing Process

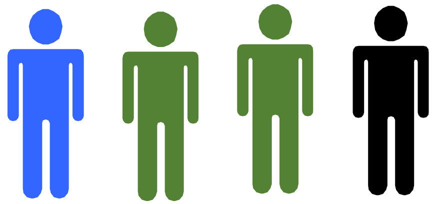
Claim: the mean
Number of TV sets
In US homes is 3.

H_0: μ = 3



Population

Now select a
random sample



Sample

Suppose the
sample mean
Is 2.84

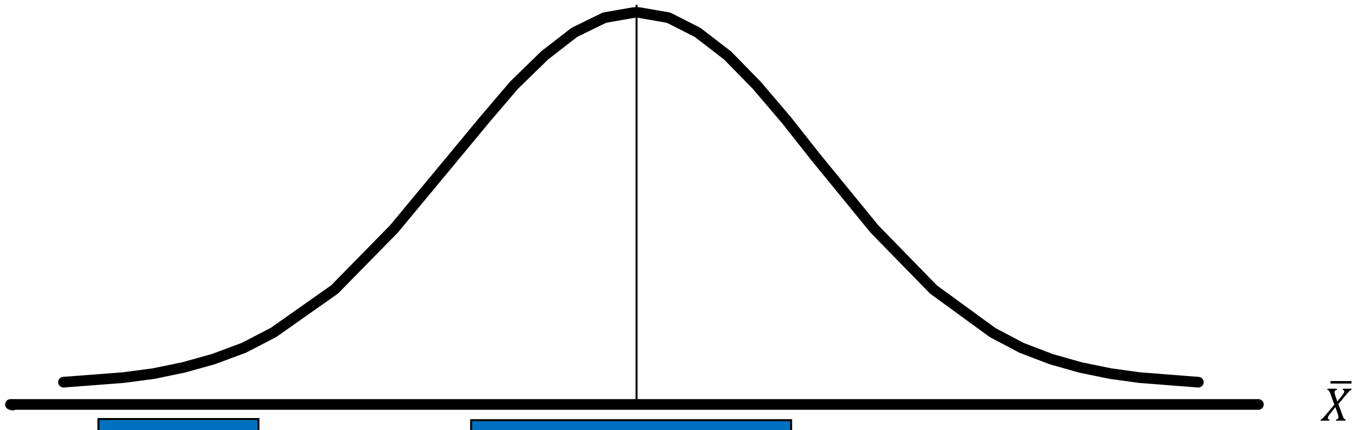
Is $\bar{X} = 2.84$ likely if μ = 3?

If not likely,

Reject H_0

Reason for Rejecting H_0

Sampling Distribution of $\bar{X}$



2.84

$\mu = 3$
If H_0 is true

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 $\mu = 3$.

Level of Significance--α

- Defines the unlikely values of the sample statistic if the null hypothesis is true
  - Defines rejection region of the sampling distribution
- Is designated by α , (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

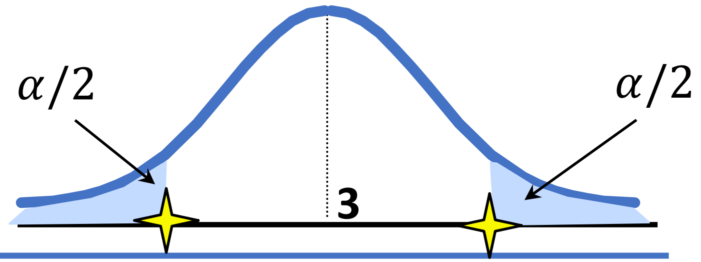
Level of Significance and the Rejection Region

Level of significance = α

H_0: $\mu = 3$

H_a: $\mu \neq 3$

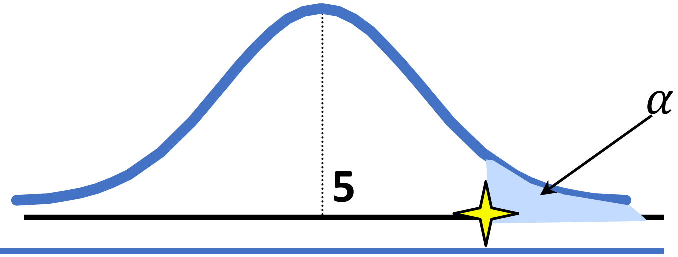
Two-tailed test



H_0: $\mu \leq 5$

H_a: $\mu > 5$

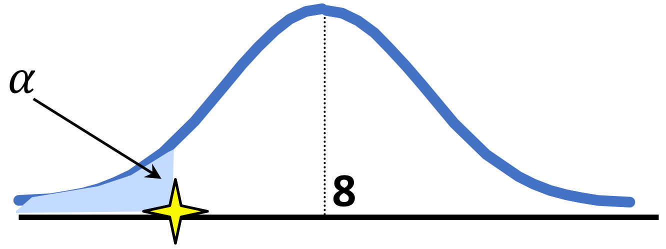
Right-tailed test



H_0: $\mu \geq 8$

H_a: $\mu < 8$

Left-tailed test



Represents critical value

Rejection region is shaded

Errors in Making Decisions

- **Type I Error**
  - Rejecting null hypothesis H_0 when it is true
  - Considered a serious type of error
- The probability of Type I Error is α
  - Called **level of significance** of the test
  - Set by the researcher in advance

Errors in Making Decisions

- **Type II Error**
  - Failing to reject the null hypothesis H_a when it is false
- Denote that the probability of Type II Error is β

Hypothesis Tests for a Population Mean μ

- σ Known------ z-test
- σ Unknown------ t-test

9.2 *z* Tests about a Population mean μ (σ Known)

- Convert sample statistic ($\overline{X}$) to a ***Z*-test statistic**
- Test statistic: $Z = \frac{\bar{X}-\mu}{\sigma/\sqrt{n}}$
- σ is known, by Central Limit Theorem, $\bar{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right)$ and

$$Z = \frac{\bar{X}-\mu}{\sigma/\sqrt{n}} \sim N(0,1)$$

Steps in Testing a “Greater Than” Alternative

1. State the null and alternative hypotheses
2. Specify the significance level alpha
3. Select the test statistic
4. Determine the critical value rule for rejecting H_0
5. Collect the sample data and calculate the value of the test statistic
6. Decide whether to reject H_0 by using the test statistic and the critical value rule
7. Interpret the statistical results in managerial terms and assess their practical importance

Recall: critical values and rejection region

Level of significance = α

H_0: $\mu = 3$

H_a: $\mu \neq 3$

Two-tailed test

H_0: $\mu \leq 5$

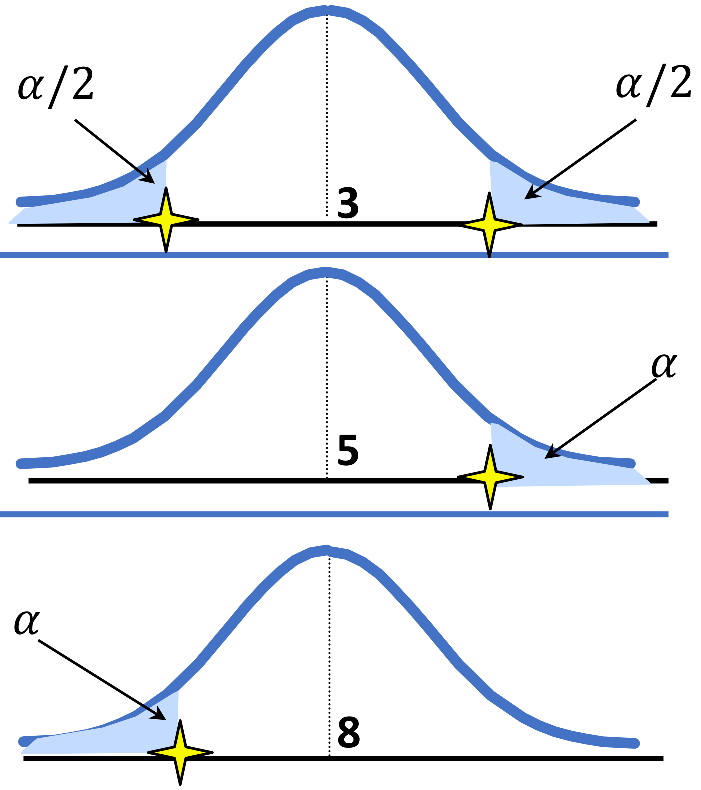
H_a: $\mu > 5$

Right-tailed test

H_0: $\mu \geq 8$

H_a: $\mu < 8$

Left-tailed test



Represents critical value

Rejection region is shaded

Steps in Testing a “Greater Than” Alternative in Trash Bag Case #1

1. State the null and alternative hypotheses

   H_0: $\mu \leq 50$
   H_a: $\mu > 50$

2. Specify the significance level $\alpha = 0.05$

3. Select the test statistic

$$z = \frac{\bar{x} - 50}{\sigma_{\bar{x}}} = \frac{\bar{x} - 50}{\sigma / \sqrt{n}}$$

Steps in Testing a "Greater Than" Alternative in Trash Bag Case #2

4. Determine the critical value rule for deciding whether or not to reject H_0
   - Reject H_0 in favor of H_a if the test statistic z is greater than the rejection point z_α
     - This is the critical value rule
   - In the trash bag case, the critical value rule is to reject H_0 if the calculated test statistic z is > 1.645

Steps in Testing a "Greater Than" Alternative in Trash Bag Case #3

$$z = \frac{\bar{x} - 50}{\sigma/\sqrt{n}} = \frac{50.575 - 50}{1.65/\sqrt{40}} = 2.20$$

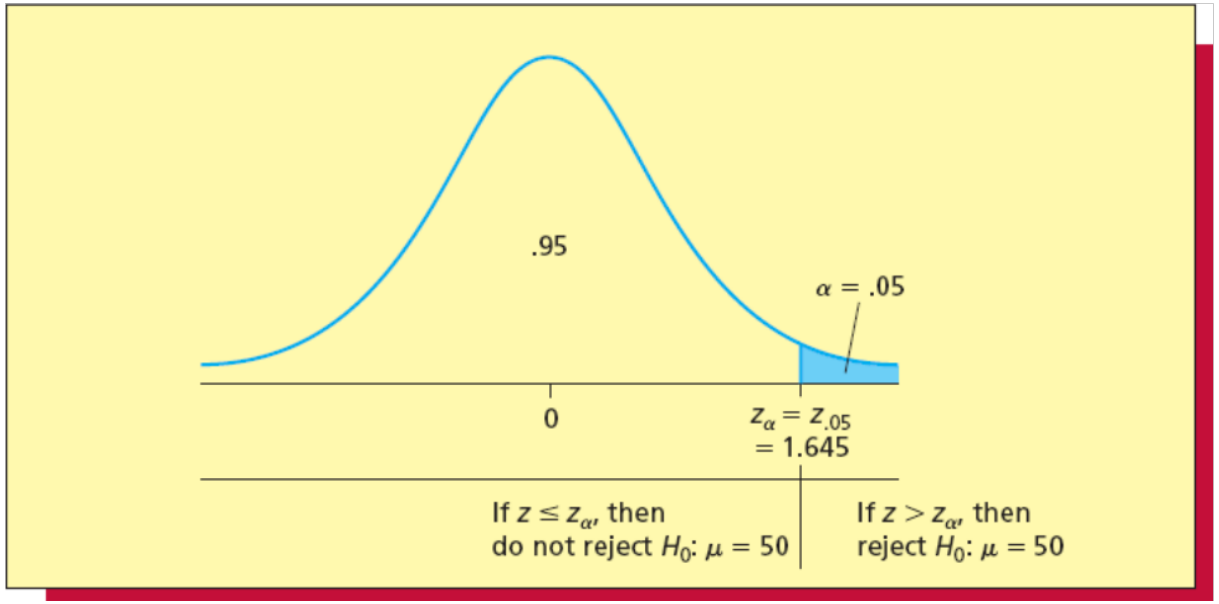


Figure 9.1

Steps in Testing a "Greater Than" Alternative in Trash Bag Case #4

- Decide whether to reject H_0 by using the test statistic and the rejection rule
  - Compare the value of the test statistic to the critical value according to the critical value rule
  - In the trash bag case, $z = 2.20$ is greater than $z_{0.05} = 1.645$
  - Therefore reject H_0: $\mu \leq 50$ in favor of H_a: $\mu > 50$ at the 0.05 significance level
- Interpret the statistical results in managerial terms and assess their practical importance

Effect of α

- At $\alpha = 0.01$, the rejection point is $z_{0.01} = 2.33$
- In the trash example, the test statistic
  $z = 2.20$ is $< z_{0.01} = 2.33$
- Therefore, cannot reject H_0 in favor of H_a at the $\alpha = 0.01$ significance level
  - This is the opposite conclusion reached with $\alpha=0.05$
  - So, the smaller we set α, the larger is the rejection point, and the stronger is the statistical evidence that is required to reject the null hypothesis H_0

The p-Value

- The p-value is the probability of the obtaining the sample results if the null hypothesis H_0 is true
- Sample results that are not likely if H_0 is true have a low p-value and are evidence that H_0 is not true
  - The p-value is the smallest value of α for which we can reject H_0
- The p-value is an alternative to testing with a z test statistic

Steps in Testing a "Less Than" Alternative in Payment Time Case #1

1. State the null and alternative hypotheses
   - H_0: $\mu \geq 19.5$ vs.
   - H_a: $\mu < 19.5$
2. Specify the significance level $\alpha = 0.01$
3. Select the test statistic

$$z = \frac{\bar{x} - 19.5}{\sigma_{\bar{x}}} = \frac{\bar{x} - 19.5}{\sigma / \sqrt{n}}$$

Steps in Testing a "Less Than" Alternative in Payment Time Case #2

4. Determine the rejection rule for deciding whether or not to reject H_0
   - The rejection rule is to reject H_0 if the calculated test statistic –z is less than –2.33
5. Collect the sample data and calculate the value of the test statistic

$$z = \frac{\bar{x} - 19.5}{\sigma / \sqrt{n}} = \frac{18.1077 - 19.5}{4.2 / \sqrt{65}} = -2.67$$

Steps in Testing a "Less Than" Alternative in Payment Time Case #3

- Decide whether to reject H_0 by using the test statistic and the rejection rule
  - In the payment time case, $z = –2.67$ is less than $z_{0.01} = –2.33$
  - Therefore reject H_0: $\mu \geq 19.5$ in favor of H_a: $\mu < 19.5$ at the 0.01 significance level
- Interpret the statistical results in managerial terms and assess their practical importance

Steps Using a p-value to Test a "Less Than" Alternative

4. Collect the sample data, compute the value of the test statistic, and calculate the p-value by corresponding to the test statistic value
5. Reject H_0 if the p-value is less than α

Steps in Testing a "Not Equal To" Alternative in Valentine Day Case #1

1. State null and alternative hypotheses
   - H_0: $\mu = 330$ vs.
   - H_a: $\mu \neq 330$
2. Specify the significance level $\alpha = 0.05$
3. Select the test statistic

$$z = \frac{\bar{x} - 330}{\sigma_{\bar{x}}} = \frac{\bar{x} - 330}{\sigma / \sqrt{n}}$$

Steps in Testing a "Not Equal To" Alternative in Valentine Day Case #2

4. Determine the rejection rule for deciding whether or not to reject H_0
   - Rejection points are z_α=1.96,–z_α=– 1.96
   - Reject H_0 in favor of H_a if the test statistic z satisfies either:
     - z greater than the rejection point $z_{\alpha/2}$, or
     - –z less than the rejection point $-z_{\alpha/2}$

Steps in Testing a "Not Equal To" Alternative in Valentine Day Case #3

5. Collect the sample data and calculate the value of the test statistic

$$z = \frac{\bar{x} - 330}{\sigma/\sqrt{n}} = \frac{326 - 330}{40/\sqrt{100}} = -1.00$$

- Decide whether to reject H_0 by using the test statistic and the rejection rule
- Interpret the statistical results in managerial terms and assess their practical importance

Steps Using a p-value to Test a "Not Equal To" Alternative

4. Collect the sample data and compute the value of the test statistic
    - Calculate the p-value by corresponding to the test statistic value
    - The p-value is $0.1587 \cdot 2 = 0.3174$
5. Reject H_0 if the p-value is less than α

Interpreting the Weight of Evidence Against the Null Hypothesis

- If $p < 0.10$, there is **some** evidence to reject H_0
- If $p < 0.05$, there is **strong** evidence to reject H_0
- If $p < 0.01$, there is **very strong** evidence to reject H_0
- If $p < 0.001$, there is **extremely strong** evidence to reject H_0

Exercise 1

- In the payment time case, we assume that σ is known and σ = 4.2 days. A sample of *n* = 65, $\bar{X}$ = 18.1077 days
- Have we strong evidence that the mean payment time for new billing system is less than 19.5 days at α=0.01 significance level ?

Exercise 2

- 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 **25** grams. Test at the **0.05** level.

p-Value Approach for Hypothesis Testing

- The **p-value** is the probability, computed assuming that the null hypothesis H_0 is true, of observing a value of the test statistic that is at least as contradictory to H_0 and supportive of H_a as the value actually computed from the sample data.
- **p-value:** Probability of obtaining a test statistic more extreme (≤ or ≥) than the observed sample value **given H_0 is true**
  - Also called observed level of significance
  - **Smallest value of α** for which **H_0 can be rejected**

Five Steps in Hypothesis Testing-population mean *--using p-value approach*

1. State the null hypothesis, H_0 and the alternative hypothesis, H_a
2. Specify the significance level α
3. Select the test statistic
4. Collect the sample data, compute the value of the test statistic, and compute the p-value.
5. **Reject H_0 at level of significance a if the p-value is less than α**. Interpret the statistical results.

Example 9.4 Trash Bag Case

- Tests show that the current trash bag has a mean breaking strength μ close to but not exceeding 50 lbs, The new bag's mean breaking strength is not known and is in question, but it is hoped that the new bag is stronger than the current one
- σ=1.65 was assumed known and α = 0.05
- Suppose a sample is taken with the following results: n = 40, $\bar{X} = 50.575$

Example 9.4 Trash Bag Case (p-Value Solution)

1. State the appropriate null and alternative hypotheses
   - H_0: $\mu \leq 50, H_a: \mu > 50$ (This is a right-tailed test)
2. Specify the significance level α
   - α = 0.05 is chosen for this test.
3. Select the test statistic
   - σ is known, by Central Limit Theorem, $\bar{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right)$ and $\frac{\bar{X}-\mu}{\sigma/\sqrt{n}} \sim N(0,1)$

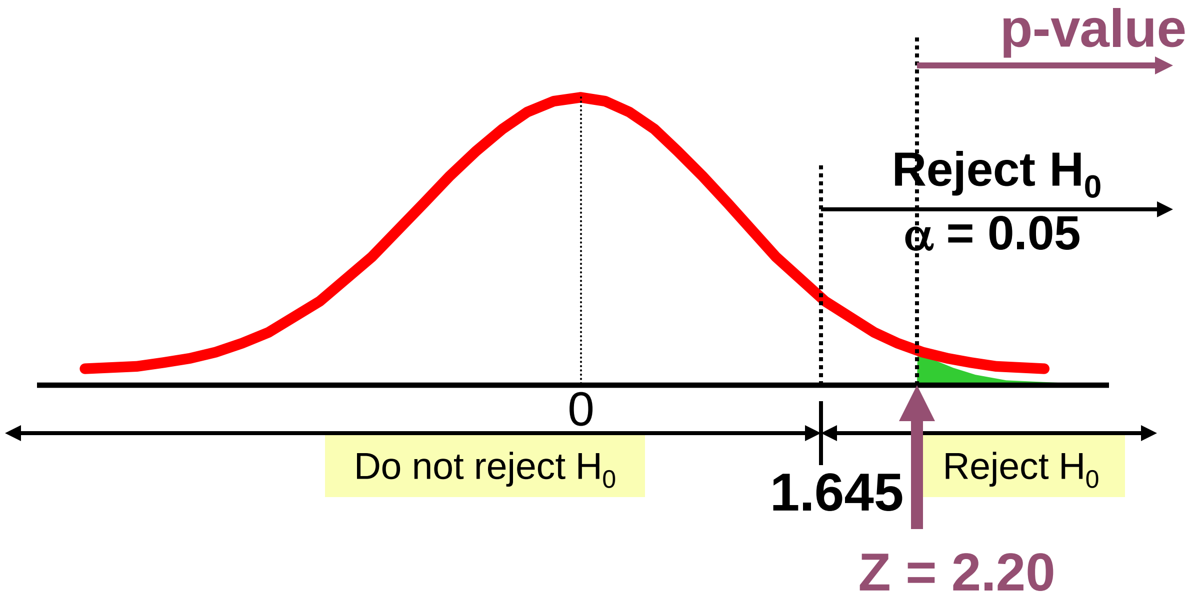
Example 9.4 Trash Bag Case (p-Value Solution)

4. Collect the sample data and calculate the value of the test statistic, and **the p-value**

- n = 40, $\bar{X}$ = 50.575 (σ =1.65 is assumed known)
- So the test statistics is $\frac{\bar{X}-\mu}{\sigma/\sqrt{n}} = \frac{50.575-50}{1.65/\sqrt{40}}$ =2.20

Example 9.4 *Trash Bag Case (p-Value Solution)*

4. Calculate the p-value and compare to α (assuming that $\mu = 50$)



$$P(\overline{X} \geq 50.575)$$

$$= P\left(Z \geq \frac{50.575 - 50}{1.65/\sqrt{40}} \right)$$

$$= P(Z \geq 2.20) = 0.5 - 0.4861$$

$$= 0.0139$$

5. Reject H_0 since p-value = 0.0139 $<$ α = 0.05

Weight of Evidence Against the Null

- Calculate the test statistic and the corresponding p-value
- Rate the strength of the conclusion about the null hypothesis H_0 according to these rules:
  - If $p < 0.10$, then there is **some** evidence to reject H_0
  - If $p < 0.05$, then there is **strong** evidence to reject H_0
  - If $p < 0.01$, then there is **very strong** evidence to reject H_0
  - If $p < 0.001$, then there is **extremely strong** evidence to reject H_0

Example 9.3: p-value approach

- Test the claim that the true mean number of TV sets in US homes is equal to 3 at the significance level 0.05. A random sample of 100 homes showed $\bar{X}$= 2.84. Assume that σ = 0.8 is known.

1. State the appropriate null and alternative hypotheses

   H_0: μ = 3 H_a: μ ≠ 3   (This is a two-tailed test)

2. Specify the significance level α

   α = 0.05 is chosen for this test.

3. Select the test statistic

   σ is known, by Central Limit Theorem, $\bar{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right)$ and $\frac{\bar{X}-\mu}{\sigma/\sqrt{n}} \sim N(0,1)$

Example 9.3: p-value approach

4. Collect the sample data and calculate the value of the test statistic and the p-value

n = 100, $\bar{X}$ = 2.84 (σ = 0.8 is assumed known)

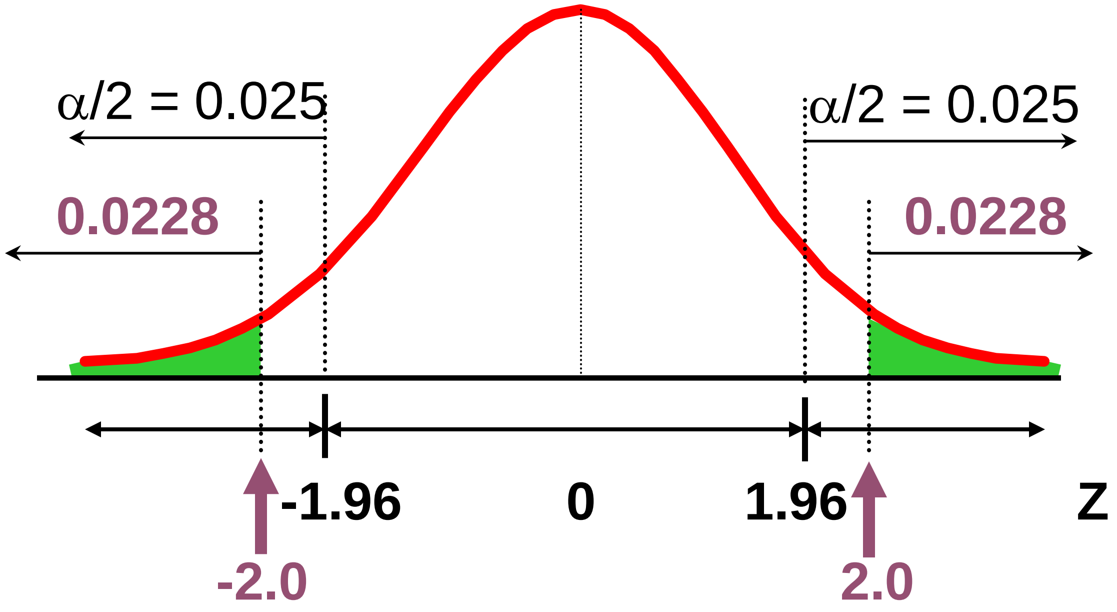
So the test statistics is $\frac{\bar{X}-\mu}{\sigma/\sqrt{n}} = \frac{2.84-3}{0.8/\sqrt{100}} = -2.0$

$P(Z < -2.0) = 0.0228$

$P(Z > 2.0) = 0.0228$

p-value

= 0.0228 + 0.0228 = 0.0456



Example 9.3: p-value approach

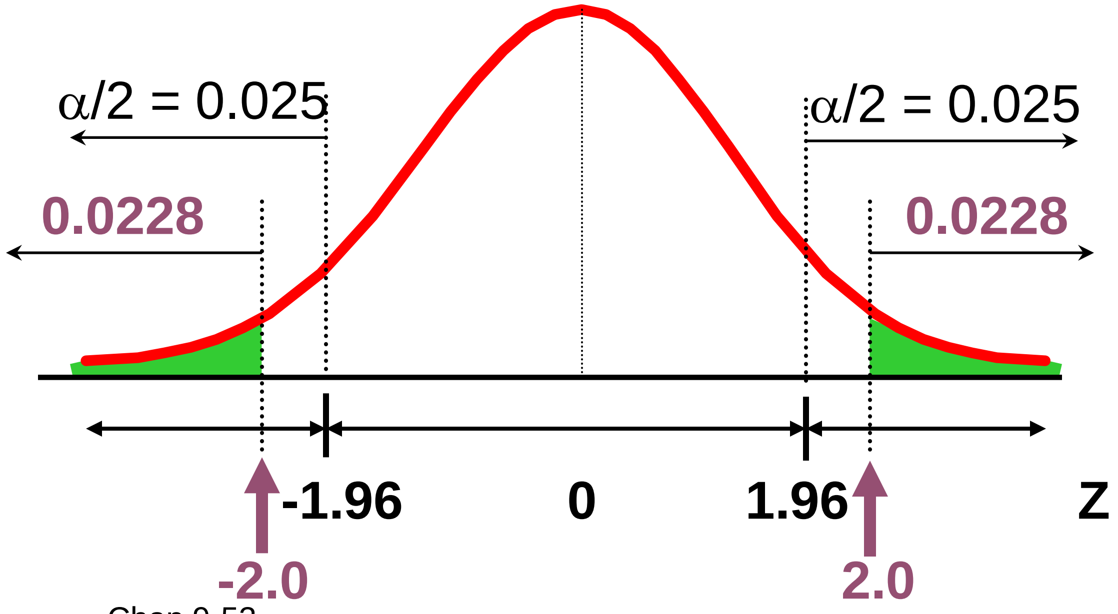
(continued)

5. Compare the p-value with α

- If p-value $<$ α, reject H_0
- If p-value $\geq$ α, do not reject H_0

Here: p-value = 0.0456
α = 0.05

Since 0.0456 < 0.05, we reject the null hypothesis



Connection to Confidence Intervals

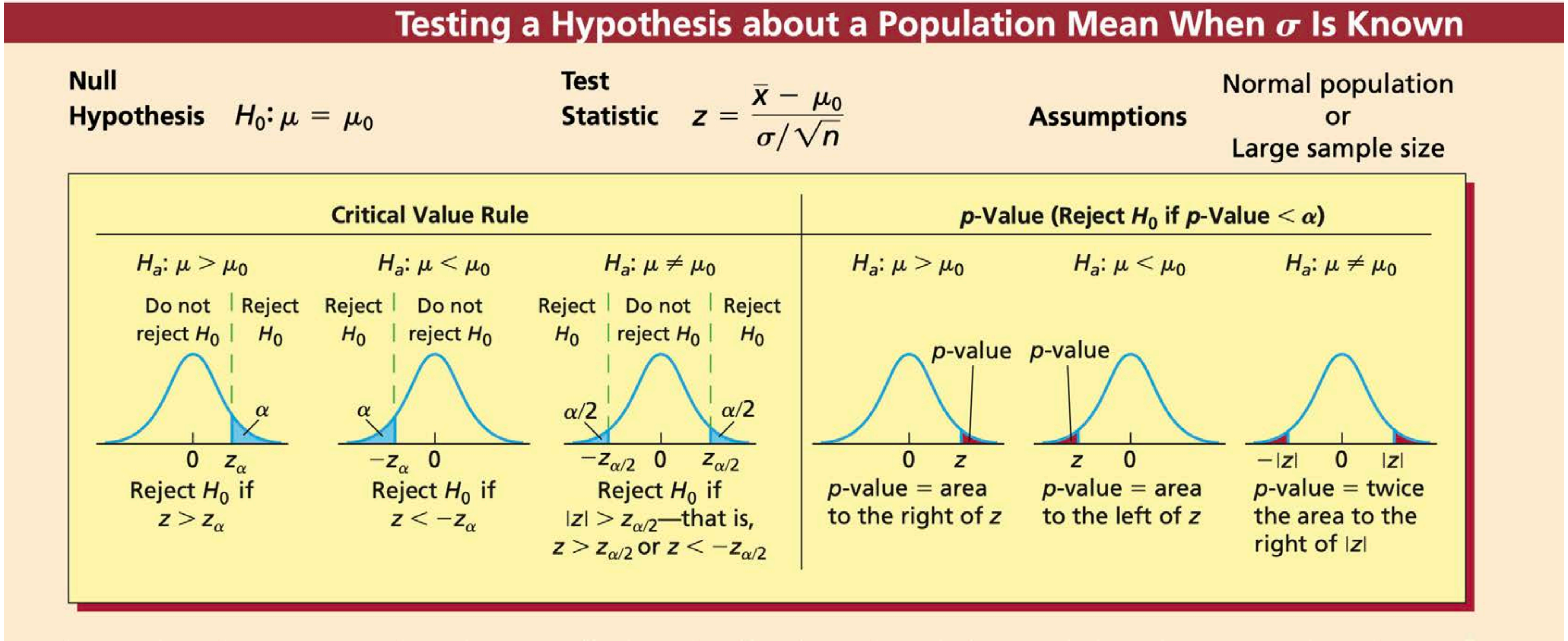
- For $\bar{X} = 2.84$, $\sigma = 0.8$ and $n = 100$, the 95% confidence interval is:

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

$$\boxed{2.6832 \leq \mu \leq 2.9968}$$

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

z Tests about a Population mean (σ Known)



Exercise 1 (p-value approach)

- In the payment time case, we assume that σ is known and σ = 4.2 days. A sample of n = 65, $\bar{X}$ = 18.1077 days
- Have we strong evidence that the mean payment time for new billing system is less than 19.5 days at α=0.01 significance level ?

The Five Steps of Hypothesis Testing

1 State the null hypothesis H_0 and the alternative hypothesis H_a.

2 Specify the level of significance α.

3 Select the test statistic.

Using a critical value rule:

4 Determine the critical value rule for deciding whether to reject H_0.

5 Collect the sample data, compute the value of the test statistic, and decide whether to reject H_0 by using the critical value rule. Interpret the statistical results.

Using a *p*-value:

4 Collect the sample data, compute the value of the test statistic, and compute the *p*-value.

5 Reject H_0 at level of significance α if the *p*-value is less than α. Interpret the statistical results.

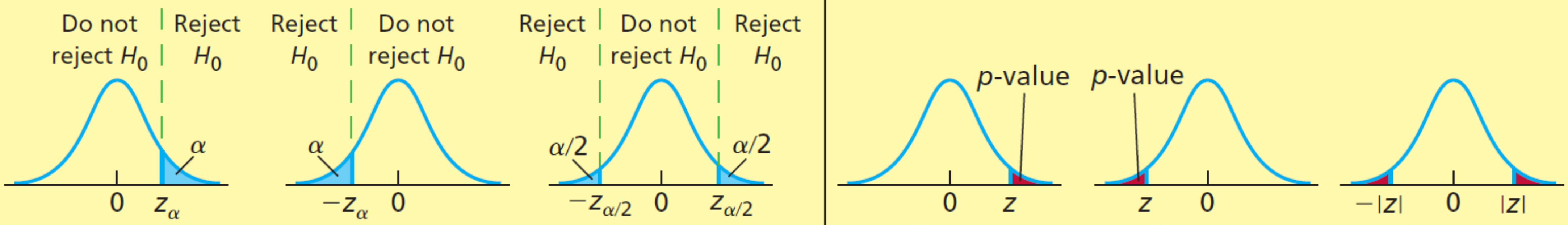
Testing a Hypothesis about a Population Mean When σ Is Known

Null Hypothesis $H_0: \mu = \mu_0$

Test Statistic $z = \dfrac{\bar{x} - \mu_0}{\sigma/\sqrt{n}}$

Assumptions Normal population or Large sample size

Critical Value Rule | **p-Value (Reject H_0 if p-Value $< \alpha$)**



Critical Value Rule			p-Value (Reject H_0 if p-Value $< \alpha$)		
$H_a: \mu > \mu_0$	$H_a: \mu < \mu_0$	$H_a: \mu \neq \mu_0$	$H_a: \mu > \mu_0$	$H_a: \mu < \mu_0$	$H_a: \mu \neq \mu_0$
Reject H_0 if $z > z_\alpha$	Reject H_0 if $z < -z_\alpha$	Reject H_0 if $|z| > z_{\alpha/2}$—that is, $z > z_{\alpha/2}$ or $z < -z_{\alpha/2}$	p-value = area to the right of z	p-value = area to the left of z	p-value = twice the area to the right of $|z|$

9.3 *t* Tests about a Population mean (σ Unknown)

- Convert sample statistic ($\bar{X}$) to a ***t*-test statistic**
- Test statistic: $t = \frac{\bar{X}-\mu}{s/\sqrt{n}}$
- σ is unknown
- If the sampled population is normally distributed (or if the sample size is large—at least 30), then this sampling distribution is exactly (or approximately) a t distribution having n-1 degrees of freedom.

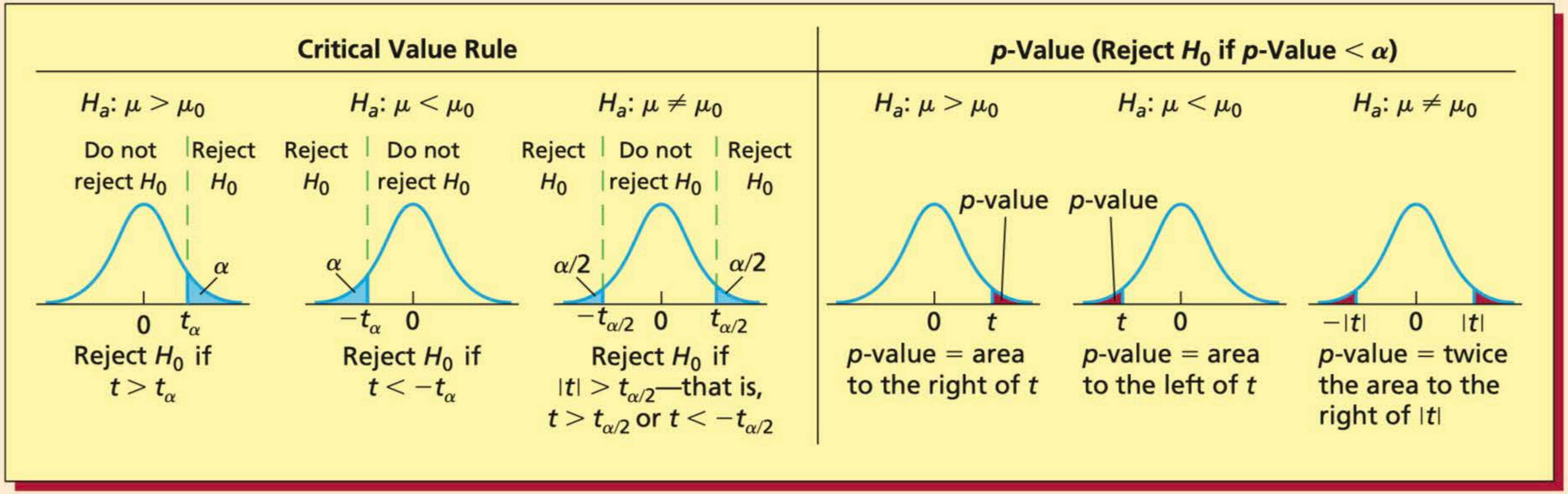
t Tests about a Population mean (σ Unknown)

A t Test about a Population Mean: σ Unknown

Null Hypothesis $H_0: \mu = \mu_0$

Test Statistic $t = \dfrac{\bar{x} - \mu_0}{s/\sqrt{n}}$ $\quad df = n - 1$

Assumptions Normal population or Large sample size



Example: Two-Tail Test (σ Unknown)

The average cost of a hotel room in New York is said to be \$168 per night. A random sample of 25 hotels resulted in $\bar{X}$ = \$172.50 and s = \$15.40. Test at the α = 0.05 level. (Assume the population distribution is normal)

1. State the appropriate null and alternative hypotheses
   - H_0: $\mu = 168$, H_a: $\mu \neq 168$ (This is a two-tailed test)
2. Specify the significance level α
   - α = 0.05 is chosen for this test.
3. Select the test statistic
   - σ is unknown, use t-statistic: $t = \frac{\bar{X}-\mu}{s/\sqrt{n}}$, df=n-1=25-1=24

Example: Two-Tail Test (σ Unknown)

4. Determine the critical values

- For α = 0.05, use the t-table

to obtain the critical values $\pm t_{0.025}$ are $\pm$ 2.0639

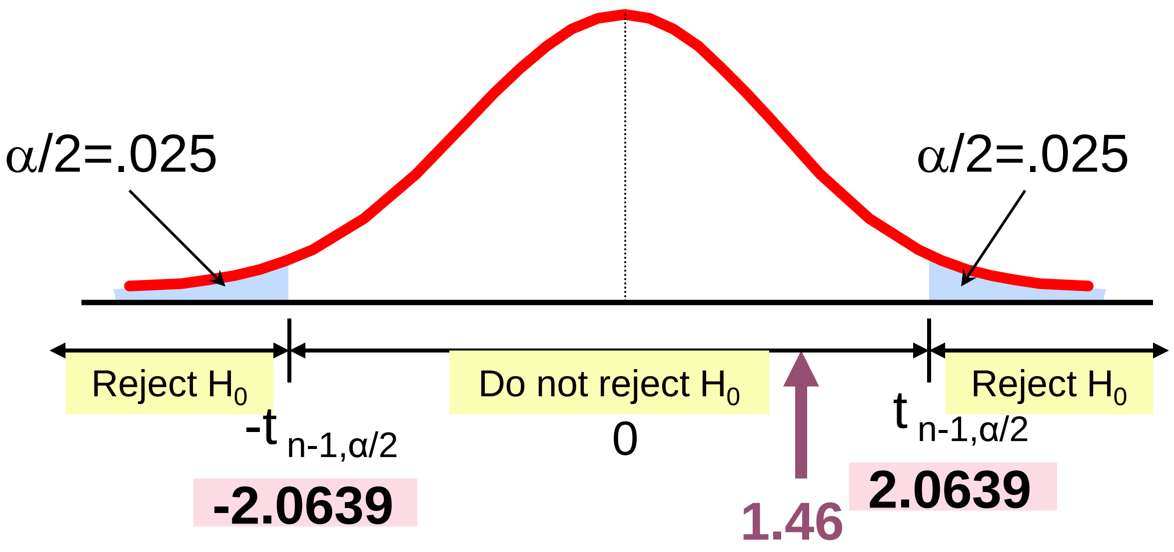
5. Collect the sample data and calculate the value of the test statistic

- n = 25, $\bar{X}$ = 172.50, s = 15.40
- So the test statistics is $\frac{\bar{X}-\mu}{s/\sqrt{n}} = \frac{172.50-168}{15.40/\sqrt{25}} = 1.46$

Example: Two-Tail Test (σ Unknown)

- **α = 0.05**
- **n = 25**
- **σ is unknown, so use a *t* statistic**
- **Critical Value:**

$t_{24} = \pm$ 2.0639



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

Do not reject H_0: not sufficient evidence that true mean cost is different than $168

Connection to Confidence Intervals

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

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

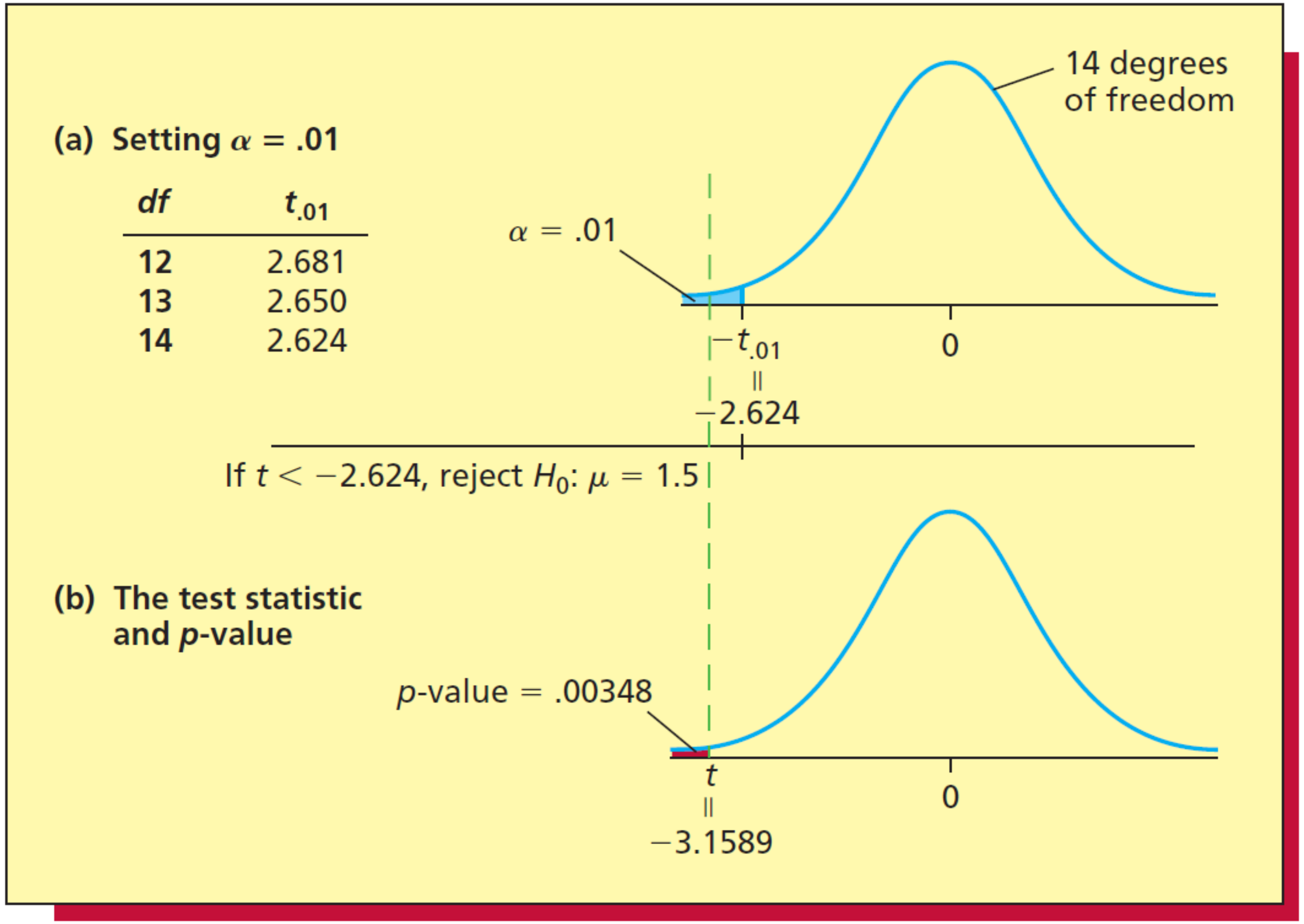
$$\boxed{166.14 \le \mu \le 178.86}$$

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

EXAMPLE The Commercial Loan Case: Mean Debt-to-Equity Ratio

One measure of a company's financial health is its *debt-to-equity ratio.* This quantity is defined to be the ratio of the company's corporate debt to the company's equity. If this ratio is too high, it is one indication of financial instability. For obvious reasons, banks often monitor the financial health of companies to which they have extended commercial loans. Suppose that, in order to reduce risk, a large bank has decided to initiate a policy limiting the mean debt-to-equity ratio for its portfolio of commercial loans to being less than 1.5. In order to assess whether the mean debt-to-equity ratio μ of its (current) commercial loan portfolio is less than 1.5, the bank will test the **null hypothesis H_0: $\mu = 1.5$ versus the alternative hypothesis H_a: $\mu < 1.5$.** In this situation,

FIGURE 9.5 Testing H_0: $\mu = 1.5$ versus H_a: $\mu < 1.5$ by Using a Critical Value and the p-Value



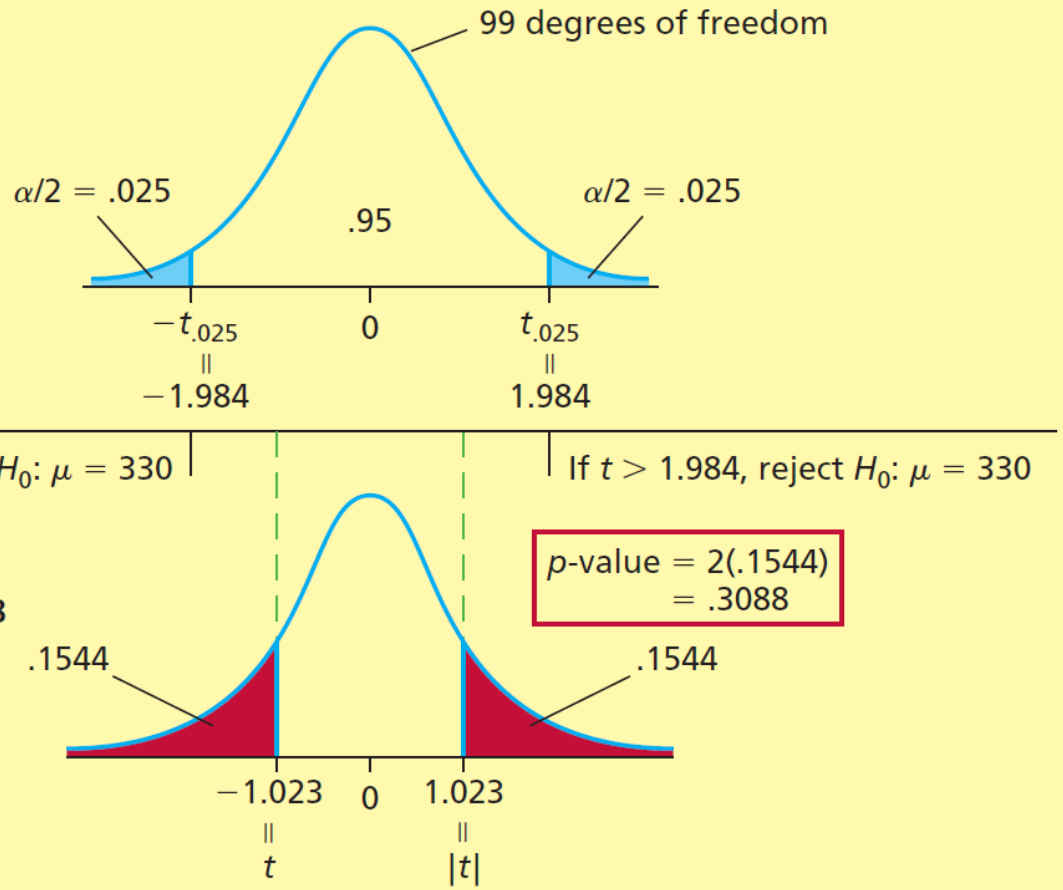

```
Test of mu = 1.5 vs < 1.5

                                              95% Upper
Variable   N     Mean   StDev  SE Mean        Bound      T      P
Ratio     15   1.3433  0.1921   0.0496       1.4307  -3.16  0.003
```

Because $t = -3.1589$ is less than $-t_{.01} = -2.624$, we reject H_0: $\mu = 1.5$ in favor of H_a: $\mu < 1.5$. That is, we conclude (at an α of .01) that the population mean debt-to-equity ratio of the bank's commercial loan portfolio is less than 1.5. This, along with the fact that the sample mean $\bar{x} = 1.3433$ is slightly less than 1.5, implies that it is reasonable for the bank to conclude that the population mean debt-to-equity ratio of its commercial loan portfolio is slightly less than 1.5.

Recall that in three cases discussed in Section 9.2 we tested hypotheses by assuming that the population standard deviation σ is known and by using z tests. If σ is actually not known in these cases (which would probably be true), we should test the hypotheses under consideration by using t tests. Furthermore, recall that in each case the sample size is large (at least 30). **In general, it can be shown that if the sample size is large, the t test is approximately valid even if the sampled population is not normally distributed (or mound shaped).** Therefore, consider the Valentine's Day chocolate case and testing $\boldsymbol{H_0: \mu = 330}$ **versus** $\boldsymbol{H_a: \mu \neq 330}$ at the **.05 level of significance.** To perform the hypothesis test, assume that we will randomly select $n = 100$ large

FIGURE 9.6 Testing H_0: $\mu = 330$ versus H_a: $\mu \neq 330$ by Using Critical Values and the *p*-Value



df	$t_{.025}$
97	1.985
98	1.984
99	1.984

```
Test of mu = 330 vs not = 330

Variable    N     Mean   StDev  SE Mean        95% CI             T       P
Boxes     100   326.00   39.10     3.91  (318.24, 333.76)    -1.02   0.309
```

retail stores and use their anticipated order quantities to calculate the value of the **test statistic t in the summary box**. Then, because the alternative hypothesis H_a: $\mu \neq 330$ implies a two tailed test, we will **reject H_0: $\mu = 330$ if the absolute value of t is greater than $t_{\alpha/2} = t_{.025} = 1.984$ (based on $n - 1 = 99$ degrees of freedom)**—see Figure 9.6(a). Suppose that when the sample is randomly selected, the mean and the standard deviation of the $n = 100$ reported order quantities are calculated to be $\bar{x} = 326$ and $s = 39.1$. The **value of the test statistic** is

$$t = \frac{\bar{x} - 330}{s/\sqrt{n}} = \frac{326 - 330}{39.1/\sqrt{100}} = -1.023$$

Because $|t| = 1.023$ is less than $t_{.025} = 1.984$, we cannot reject H_0: $\mu = 330$ by setting α equal to .05. It follows that we cannot conclude (at an α of .05) that this year's population mean order quantity of the valentine box by large retail stores will differ from 330 boxes. Therefore, the candy company will base its production of valentine boxes on the 10 percent projected sales increase.

9.4 z Tests about a Population Proportion p

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

z Tests about a Population Proportion p

- In one sample, Sample proportion in the success category is denoted by $\hat{p}$
  - $\hat{p} = \dfrac{\text{number of success in one sample}}{\text{sample size}}$
- When both np and $n(1\text{-}p)$ are at least 5, the sampling distribution of $\hat{p}$ can be approximated by a normal distribution with mean and standard deviation
  - $\mu_{\hat{p}} = p,\ \sigma_{\hat{p}} = \sqrt{\dfrac{p(1-p)}{n}}$

z Tests about a Population Proportion *p*

- Convert sample statistic ($\hat{p}$) to a ***z*-test statistic**
- Test statistic: $Z = \frac{\hat{p}-p}{\sqrt{\frac{p(1-p)}{n}}}$
- $np \geq 5, n(1-p) \geq 5, Z = \frac{\hat{p}-p}{\sqrt{\frac{p(1-p)}{n}}} \sim N(0,1)$
- An equivalent form to the test statistic, in terms of the number of successes, X:

$$Z = \frac{X - np}{\sqrt{np(1-p)}}$$

z Tests about a Population Proportion p

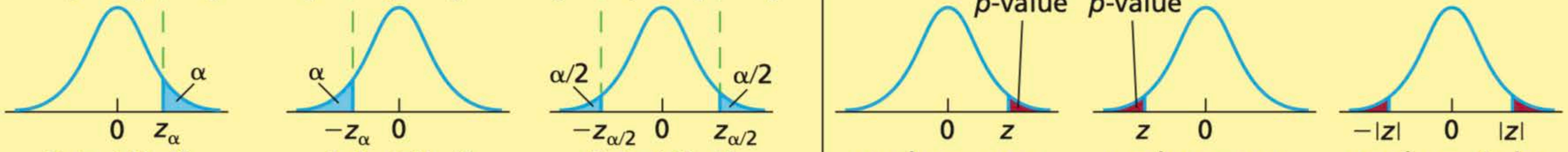
A Large Sample Test about a Population Proportion

Null Hypothesis $H_0: p = p_0$

Test Statistic $z = \dfrac{\hat{p} - p_0}{\sqrt{\dfrac{p_0(1-p_0)}{n}}}$

Assumptions[3] $np_0 \geq 5$ and $n(1-p_0) \geq 5$

Critical Value Rule			***p*-Value (Reject H_0 if *p*-Value $< \alpha$)**		
$H_a: p > p_0$	$H_a: p < p_0$	$H_a: p \neq p_0$	$H_a: p > p_0$	$H_a: p < p_0$	$H_a: p \neq p_0$
Do not reject H_0 | Reject H_0	Reject H_0 | Do not reject H_0	Reject H_0 | Do not reject H_0 | Reject H_0			



Reject H_0 if $z > z_\alpha$	Reject H_0 if $z < -z_\alpha$	Reject H_0 if $|z| > z_{\alpha/2}$—that is, $z > z_{\alpha/2}$ or $z < -z_{\alpha/2}$	*p*-value = area to the right of z	*p*-value = area to the left of z	*p*-value = twice the area to the right of $|z|$

EXAMPLE 9.5 The Cheese Spread Case: Improving Profitability

Recall that the soft cheese spread producer has decided that replacing the current spout with the new spout is profitable only if p, the true proportion of all current purchasers who would stop buying the cheese spread if the new spout were used, is less than .10. The producer feels that it is unwise to change the spout unless it has very strong evidence that p is less than .10. Therefore, the spout will be changed if and only if the null hypothesis H_0: $p = .10$ can be rejected in favor of the alternative hypothesis H_a: $p < .10$ at the .01 level of significance.

$$z = \frac{\hat{p} - p_0}{\sqrt{\dfrac{p_0(1 - p_0)}{n}}} = \frac{.063 - .10}{\sqrt{\dfrac{.10(1 - .10)}{1{,}000}}} = -3.90$$

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 $\alpha = 0.05$ significance level.

1. State the appropriate null and alternative hypotheses
   - $H_0: p = 0.08, H_a: p \neq 0.08$ (This is a two-tailed test)
2. Specify the significance level α
   - $\alpha = 0.05$ is chosen for this test.
3. Select the test statistic
   - $np_0 = 500 * 0.08 = 40 > 5, n(1 - p_0) = 500 * (1 - 0.08) = 460 > 5,$
   - $Z = \frac{\hat{p} - p}{\sqrt{\frac{p(1-p)}{n}}} \sim N(0,1)$

Example: z-Test for Proportion-critical value

4. Determine the critical values
   - For α = 0.05, use the standard normal table

   to obtain the critical values $\pm z_{0.025}$ are ± 1.96

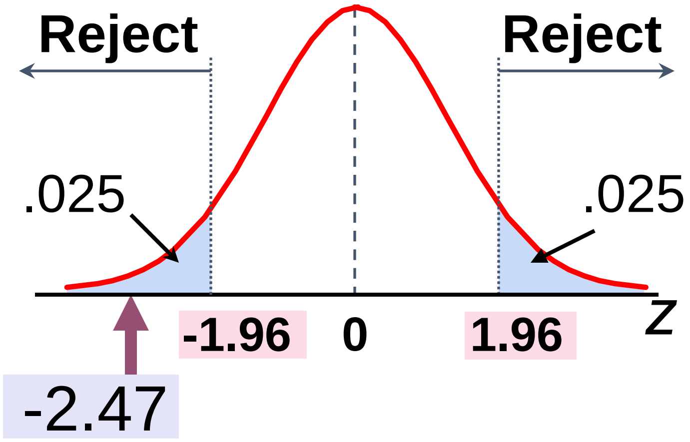
5. Collect the sample data and calculate the value of the test statistic
   - n = 500, X= 25, $\hat{p} = \frac{25}{500} = 0.05$
   - So the test statistics is $Z = \frac{\hat{p}-p}{\sqrt{\frac{p(1-p)}{n}}} = \frac{0.05-0.08}{\sqrt{\frac{0.08(1-0.08)}{500}}} = -2.47$

Example: z Test for Proportion-critical value

Test Statistic:

$$Z = \frac{\hat{p} - p}{\sqrt{\frac{p(1-p)}{n}}} = \frac{0.05 - 0.08}{\sqrt{\frac{0.08(1-0.08)}{500}}} = -2.47$$

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



Decision:

Reject H_0 at $\alpha = 0.05$

Conclusion:

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

Example 9.8: z-Test for Proportion-p value

4. Collect the sample data and calculate the value of the test statistic, and **the p-value**

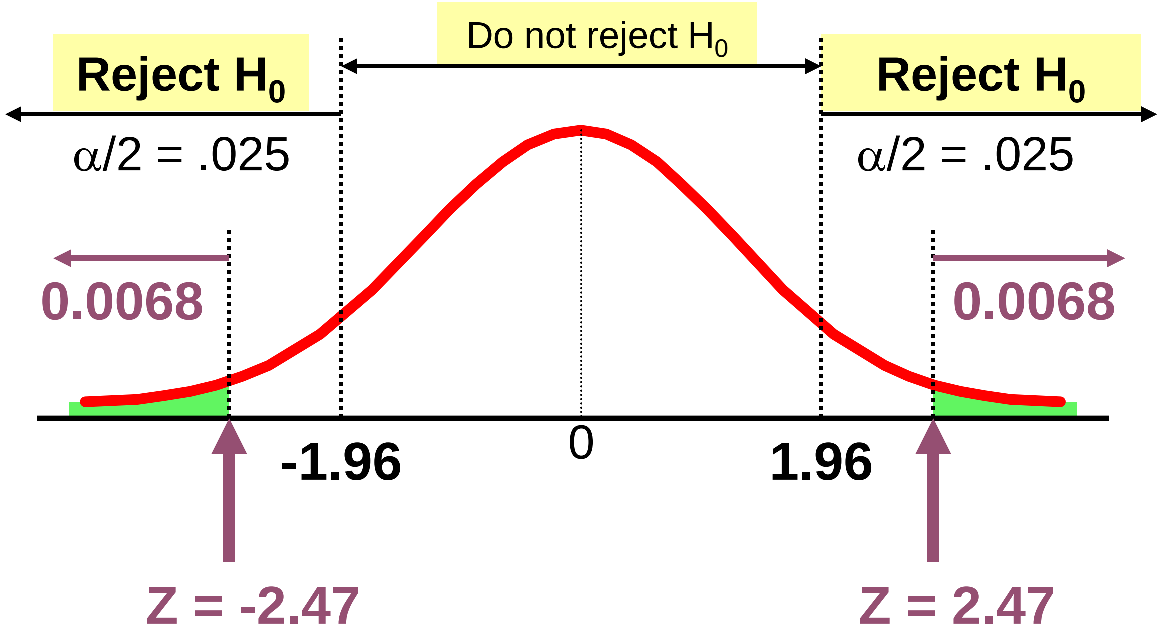
- n = 500, X= 25, $\hat{p} = \frac{25}{500} = 0.05$
- So the test statistics is $Z = \frac{\hat{p}-p}{\sqrt{\frac{p(1-p)}{n}}} = \frac{0.05-0.08}{\sqrt{\frac{0.08(1-0.08)}{500}}} = -2.47$

Example: z-Test for Proportion-p value

(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:

$$P(Z \le -2.47) + P(Z \ge 2.47)$$

$$= 2(0.0068) = 0.0136$$

5. Reject H_0 since p-value = 0.0136 < α = 0.05

EXAMPLE 9.7 The Phantol Case: Testing the Effectiveness of a Drug

Recent medical research has sought to develop drugs that lessen the severity and duration of viral infections. Virol, a relatively new drug, has been shown to provide relief for 70 percent of all patients suffering from viral upper respiratory infections. A major drug company is developing a competing drug called Phantol. The drug company wishes to investigate whether Phantol is more effective than Virol. To do this, the drug company will test a hypothesis about the proportion, p, of all patients whose symptoms would be relieved by Phantol. **The null hypothesis to be tested is H_0: $p = .70$, and the alternative hypothesis is H_a: $p > .70$.** If H_0 can be rejected in favor of H_a at the **.05 level of significance,** the drug company will conclude that Phantol helps more than the 70 percent of patients helped by Virol. To perform the hypothesis test, we will randomly select

EXAMPLE 9.8 The Electronic Article Surveillance Case: False Alarms

A sports equipment discount store is considering installing an electronic article surveillance device and is concerned about the proportion, p, of all consumers who would never shop in the store again if the store subjected them to a false alarm. Suppose that industry data for general discount stores says that 15 percent of all consumers say that they would never shop in a store again if the store subjected them to a false alarm. To determine whether this percentage is different for the sports equipment discount store, the store will test the **null hypothesis H_0: $p = .15$ versus the alternative hypothesis H_a: $p \neq .15$** at the **.05 level of significance.** To perform the hypothesis test, the store will randomly select $n = 500$ consumers, find the proportion $\hat{p}$ of these consumers who say that they would never shop in the store again if the store subjected them to a false alarm, and calculate the value of the **test statistic z in the summary box.** Then, because the alternative

$$z = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1 - p_0)}{n}}} = \frac{.14 - .15}{\sqrt{\frac{.15(1 - .15)}{500}}} = -.63$$

Because $|z| = .63$ is less than $z_{.025} = 1.96$, we cannot reject $H_0: p = .15$ in favor of $H_a: p \neq .15$. That is, we cannot conclude (at an α of .05) that the percentage of all people who would never shop in the sports discount store again if the store subjected them to a false alarm differs from the general discount store percentage of 15 percent.

Exercise 3

- Roder's Discount Store chain issues its own credit card. Lisa, the credit manager, wants to find out if the mean monthly unpaid balance is more than $400. The level of significance is set at .05. A random check of 25 unpaid balances revealed the sample mean to be $407 and the sample standard deviation to be $38. *Should Lisa conclude that the population mean is greater than $400?*

Exercise 4

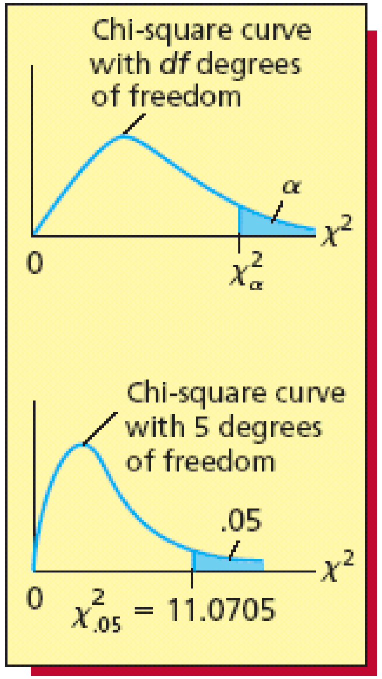
- In the past, 15% of the mail order solicitations for a certain charity resulted in a financial contribution. A new solicitation letter that has been drafted is sent to a sample of 200 people and 45 responded with a contribution. *At the 0.05 significance level can it be concluded that the new letter is more effective?*

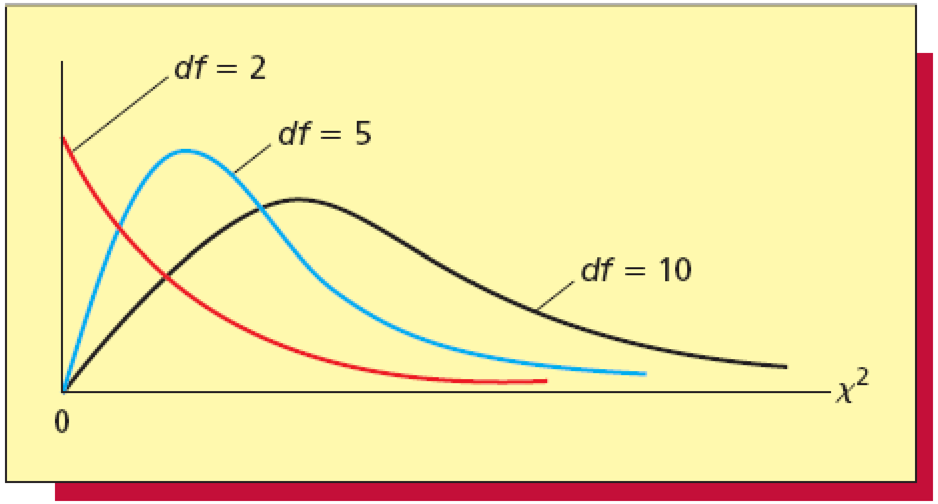
Exercise 5

- A major videocassette rental chain is considering opening a new store in an area that currently does not have any such stores. The chain will open if there is evidence that more than 5,000 of the 20,000 households in the area are equipped with videocassette recorders (VCRs). It conducts a telephone poll of 300 randomly selected households in the area and finds that 96 have VCRs. Find the *p*-value associated with the test statistic in this problem and make decision if the chain open a new store in the area.

9.6 The Chi-Square Distribution

- The chi-square χ^2 distribution depends on the number of degrees of freedom
- A chi-square point χ^2_α is the point under a chi-square distribution that gives right-hand tail area α





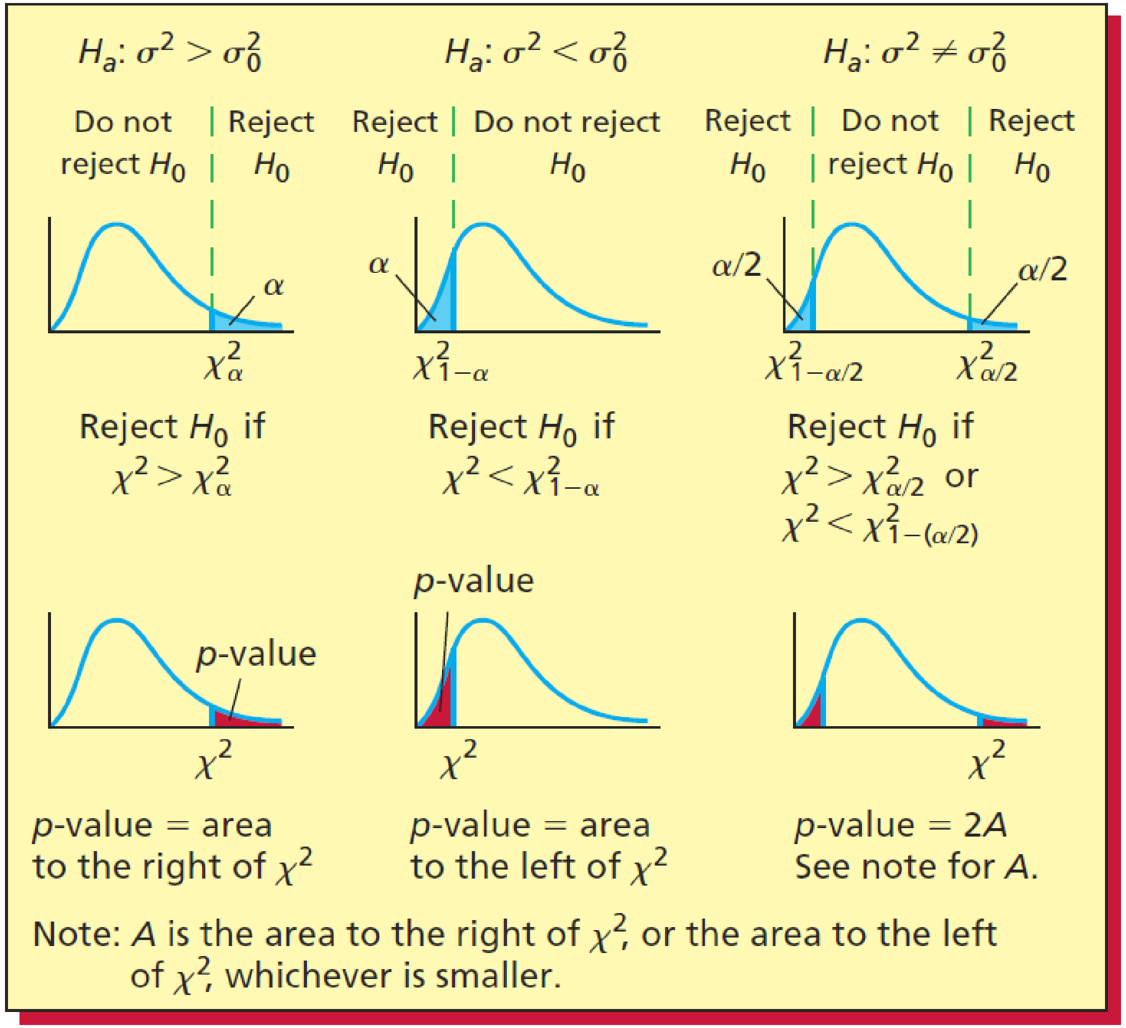
9.7 Statistical Inference for Population Variance (Optional)

- If s^2 is the variance of a random sample of n measurements from a normal population with variance σ^2
- The sampling distribution of the statistic
  $(n - 1)\, s^2 / \sigma^2$ is a chi-square distribution with $(n – 1)$ degrees of freedom
- Can calculate confidence interval and perform hypothesis testing

Confidence Interval for Population Variance

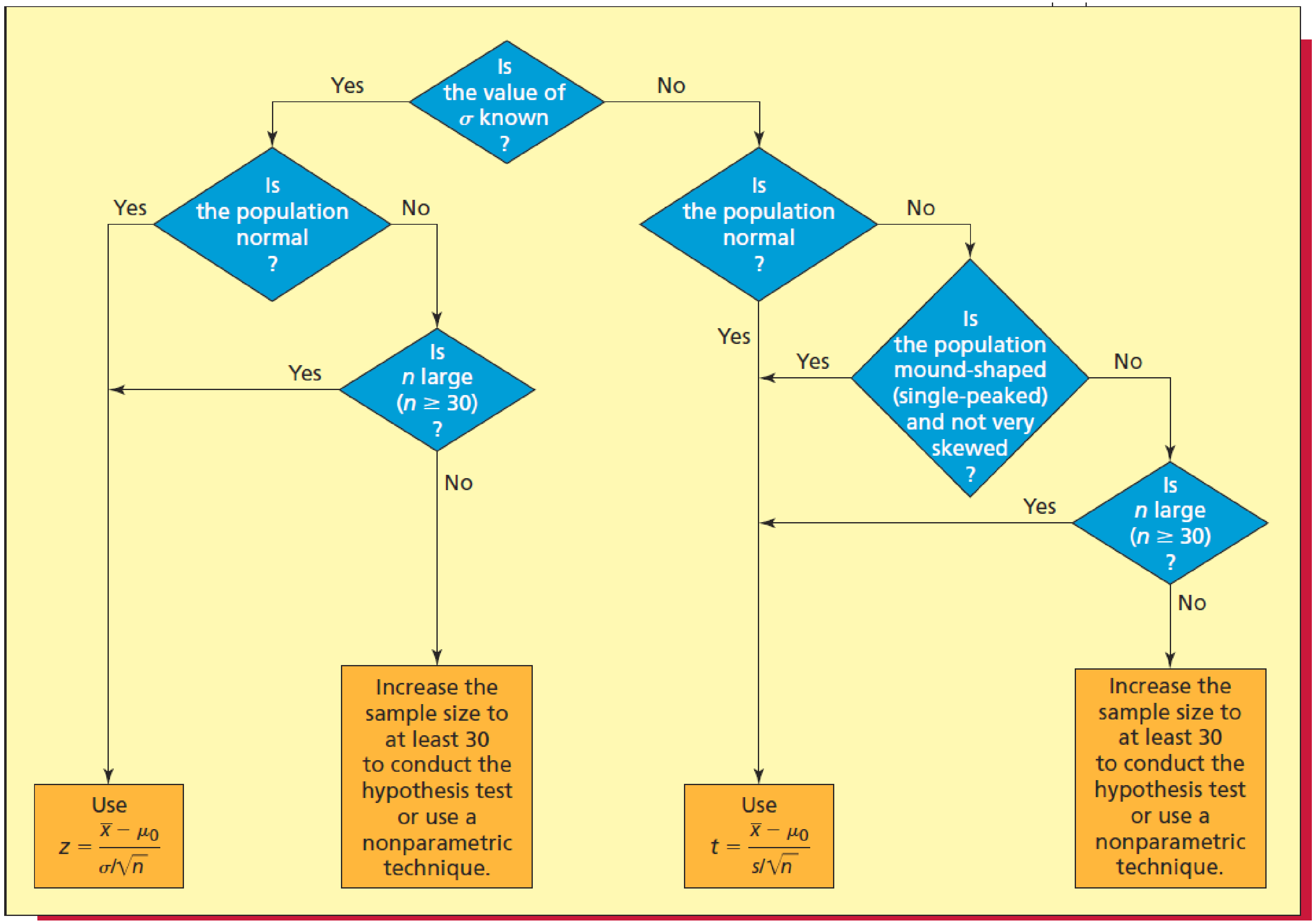
$$\left(\frac{(n-1)s^2}{\chi^2_{\alpha/2}}, \frac{(n-1)s^2}{\chi^2_{1-\alpha/2}} \right)$$

Hypothesis Testing for Population Variance



$$\chi^2 = \frac{(n-1)s^2}{\sigma_0^2}$$

Selecting an Appropriate Test Statistic for a Test about a Population Mean



Summary

Chapter Summary

We began this chapter by learning about the two hypotheses that make up the structure of a hypothesis test. The **null hypothesis** is the statement being tested. The null hypothesis is often a statement of "no difference" or "no effect," and it is not rejected unless there is convincing sample evidence that it is false. The **alternative, or, research, hypothesis** is a statement that is accepted only if there is convincing sample evidence that it is true and that the null hypothesis is false. In some situations, the alternative hypothesis is a statement for which we wish to find supportive evidence. We also learned that two types of errors can be made in a hypothesis test. A **Type I error** occurs when we reject a true null hypothesis, and a **Type II error** occurs when we do not reject a false null hypothesis.

We studied two commonly used ways to conduct a hypothesis test. The first involves comparing the value of a test statistic with what is called a **critical value,** and the second employs what is called a ***p*-value.** The *p*-value measures the weight of evidence against the null hypothesis. The smaller the *p*-value, the more we doubt the null hypothesis.

The specific hypothesis tests we covered in this chapter all dealt with a hypothesis about one population parameter. First, we studied a test about a **population mean** that is based on the assumption that the population standard deviation **σ is known.** This test employs the **normal distribution.** Second, we studied a test about a population mean that assumes that **σ is unknown.** We learned that this test is based on the ***t* distribution.** Figure 9.9 presents a flowchart summarizing how to select an appropriate test statistic to test a hypothesis about a population mean. Then we presented a test about a **population proportion** that is based on the **normal distribution.** We concluded this chapter by studying Type II error probabilities, and we showed how we can find the sample size needed to make both the probability of a Type I error and the probability of a serious Type II error as small as we wish.

Thank you!