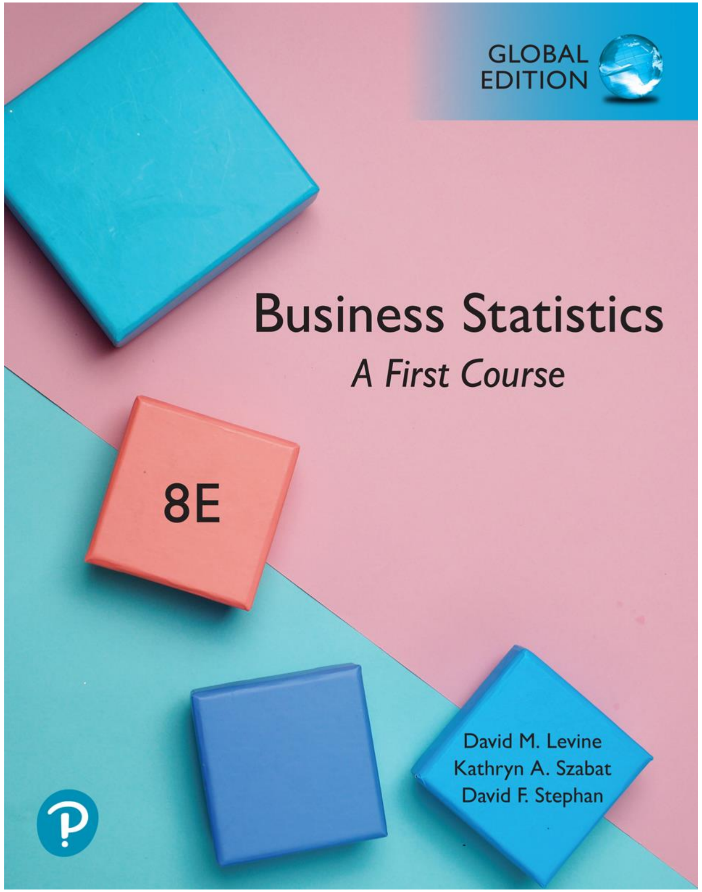


Chapter 9

Fundamentals of Hypothesis Testing: One-Sample Tests

Objectives

In this chapter, you learn:

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

What is a Hypothesis?

DCOVA

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

Example: The mean monthly cell phone bill in this city is μ = $42.

  - population proportion:

Example: The proportion of adults in this city with cell phones is $\pi = 0.88$.

The Null Hypothesis, H_0

DCOVA

- States the claim or assertion to be tested.

Example: The mean diameter of a manufactured bolt is 30mm $(H_0 : \mu = 30)$

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

$$H_0 : \mu = 30$$

$$\cancel{H_0 : \bar{X} = 30}$$

The Null Hypothesis, H_0 (continued)

DCOVA

- Begin with the assumption that the null hypothesis is true.
  - Similar to the notion of innocent until proven guilty.
- Represents the current belief in a situation.
- Always contains "=", or "≤", or "≥" sign.
- May or may not be rejected.

H_0 = ≤ ≥

The Alternative Hypothesis, H_1

DCOVA

- Is the opposite of the null hypothesis.
  - e.g., The mean diameter of a manufactured bolt is not equal to 30mm (H_1: $\mu \neq 30$).
- Challenges the status quo.
- Never contains the "=", or "≤", or "≥" sign.
- May or may not be proven.
- Is generally the hypothesis that the researcher is trying to prove.

The Hypothesis Testing Process

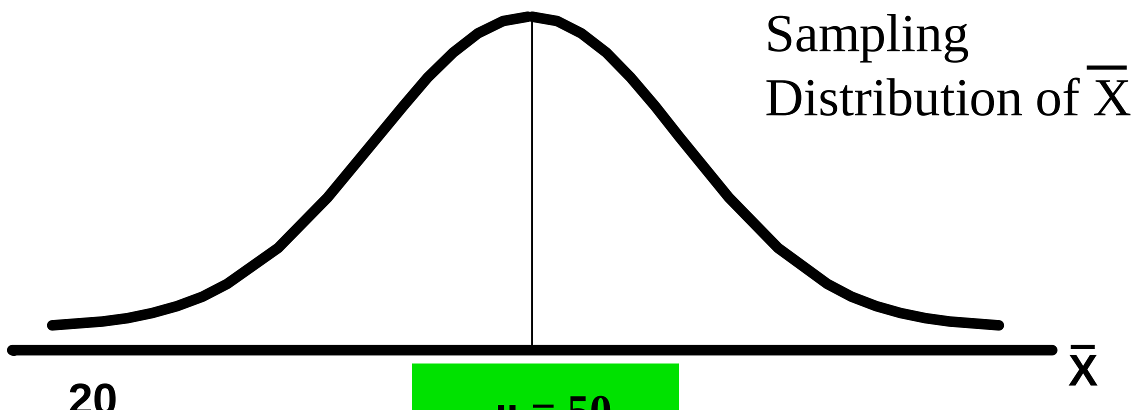
DCOVA

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

The Hypothesis Testing Process (continued)

- 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 (continued)



$\mu = 50$
If H_0 is true

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

... When in fact this was the population mean…

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

The Test Statistic and Critical Values

DCOVA

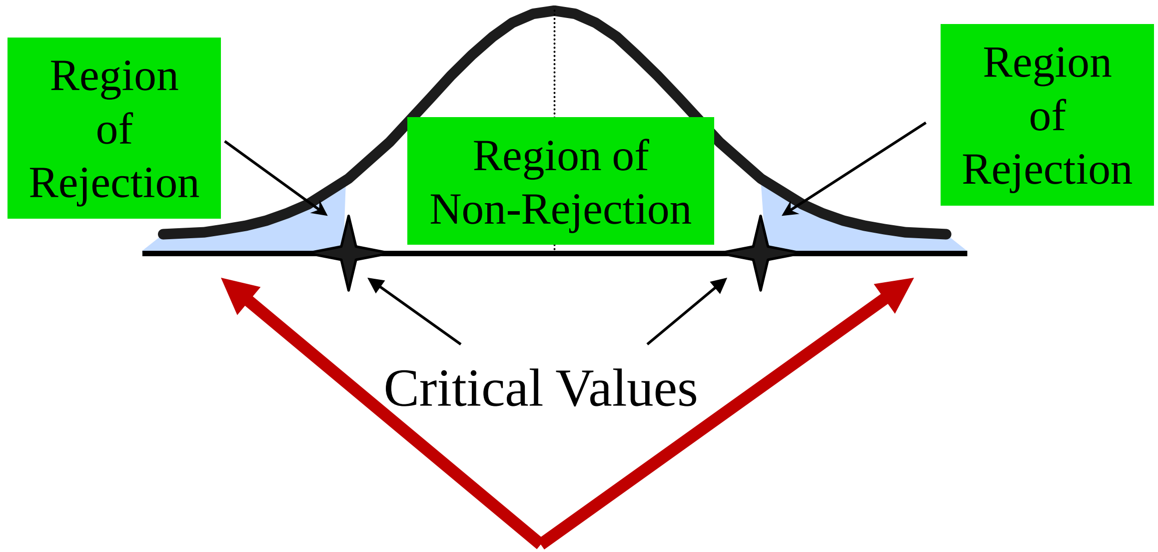
- If the sample mean is close to the stated population mean, the null hypothesis is not rejected.
- If the sample mean is far from the stated population mean, the null hypothesis is rejected.
- How far is "far enough" to reject H_0?
- The critical value of a test statistic creates a "line in the sand"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

The Test Statistic and Critical Values

(continued)

DCOVA

Sampling Distribution of the test statistic



"Too Far Away" From Mean of Sampling Distribution

Risks in Decision Making Using Hypothesis Testing

DCOVA

- **Type I Error:**
  - Reject a true null hypothesis.
  - A Type I error is a “false alarm.”
  - The probability of a Type I Error is $\boldsymbol{\alpha}$**.**
    - Called level of significance of the test.
    - Set by researcher in advance.
- **Type II Error:**
  - Failure to reject a false null hypothesis.
  - Type II error represents a “missed opportunity.”
  - The probability of a Type II Error is $\boldsymbol{\beta}$**.**

Possible Errors in Hypothesis Test Decision Making

(continued)

Possible Hypothesis Test Outcomes		
	Actual Situation	
Decision	H_0 True	H_0 False
Do Not Reject H_0	Correct Decision Confidence = 1 - α	Type II Error P(Type II Error) = β
Reject H_0	Type I Error P(Type I Error) = α	Correct Decision Power = (1 – β)

Possible Errors in Hypothesis Test Decision Making

(continued) DCOVA

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

Type I & II Error Relationship

DCOVA

- 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) ⇩

Factors Affecting Type II Error

DCOVA

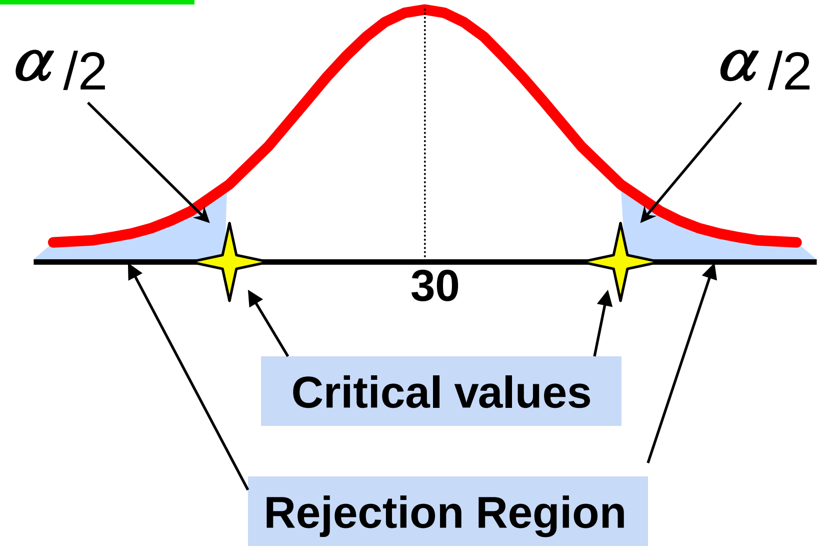
- All else equal,
  - $\beta \uparrow$ when the difference between hypothesized parameter and its true value $\downarrow$.
  - $\beta \uparrow$ when $\alpha \downarrow$.
  - $\beta \uparrow$ when $\sigma \uparrow$.
  - $\beta \uparrow$ when $n \downarrow$.

Level of Significance and the Rejection Region

DCOVA

H_0: μ = 30
H_1: μ ≠ 30

Level of significance = α



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

Hypothesis Tests for the Mean

DCOVA

Hypothesis Tests for μ

- **σ Known** **(Z test)**
- **σ Unknown** **(t test)**

Z Test of Hypothesis for the Mean (σ Known)

DCOVA

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

$\bar{X} \pm Z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$

Hypothesis Tests for μ

σ Known (Z test)

σ Unknown (t test)

The test statistic is:

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

Critical Value Approach to Testing

DCOVA

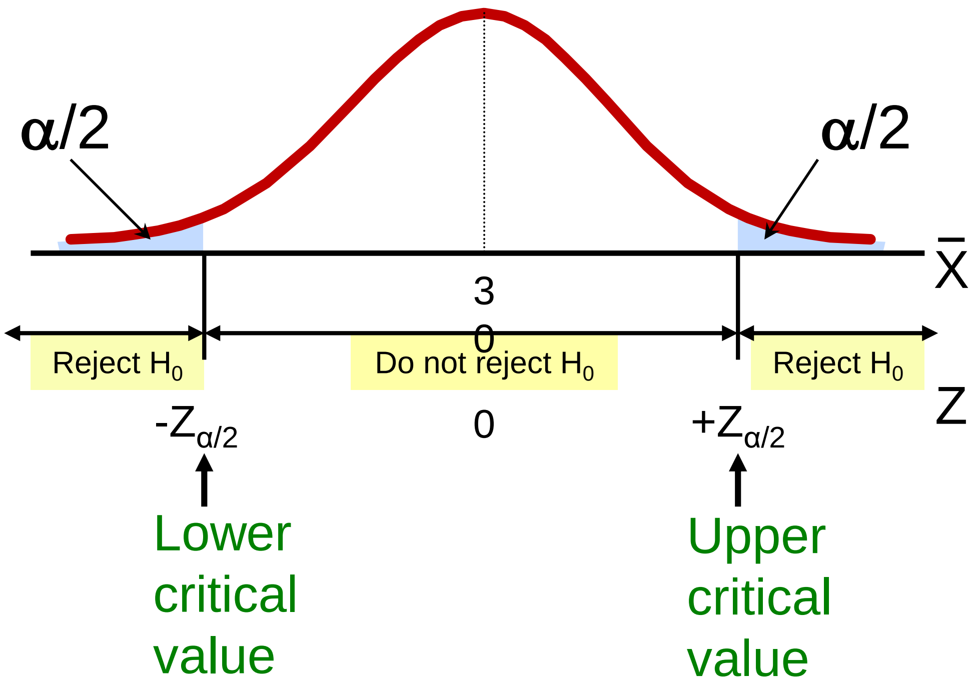
- For a two-tail test for the mean, σ known:
- Convert sample statistic ($\bar{X}$) to test statistic (Z_{STAT}.)
- Determine the critical Z values for a specified level of significance α from a table or by using computer software.
- Decision Rule: If the test statistic falls in the rejection region,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

Two-Tail Tests

DCOVA

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

H_0: $\mu = 30$
H_1: $\mu \neq 30$



Steps in The Critical Value Approach To Hypothesis Testing

DCOVA

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. The level of significance is based on the relative importance of Type I and Type II errors in the situation.
3. Determine the appropriate test statistic and sampling distribution.
4. Determine the critical values that divide the rejection and nonrejection regions.

Steps in The Critical Value Approach To Hypothesis Testing (continued)

DCOVA

5. Collect the sample data, organize the results, and compute the value of the test statistic.
6. Make the statistical decision, determine whether the assumptions are valid, and state the managerial conclusion in the context of the theory, claim, or assertion being tested. If the test statistic falls into the nonrejection region, you do not reject the null hypothesis H_0. If the test statistic falls into the rejection region, reject the null hypothesis.

Hypothesis Testing Example

DCOVA

Test the claim that the true mean diameter of a manufactured bolt is 30mm. (Assume σ = 0.8)

1. State the appropriate null and alternative hypotheses:
   - H_0: μ = 30 H_1: μ ≠ 30 (This is a two-tail test).
2. Specify the desired level of significance and the sample size:
   - Suppose that α = 0.05 and n = 100 are chosen for this test.

Hypothesis Testing Example

(continued)

3. Determine the appropriate technique:
   - σ is assumed known so this is a Z test.
4. Determine the critical values:
   - For $\alpha = 0.05$ the critical Z values are ±1.96.
5. Collect the data and compute the test statistic.
   - Suppose the sample results are:

   $n = 100$, $\bar{X} = 29.84$ ($\sigma = 0.8$ is assumed known).

$\mu = 30$

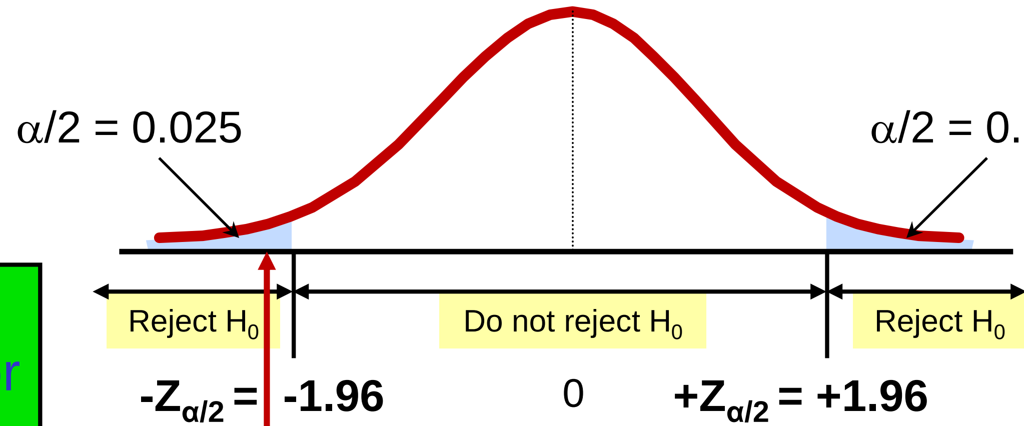
$\bar{X} = 29.84$

So the test statistic is:

$$Z_{STAT} = \frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}} = \frac{29.84 - 30}{\frac{0.8}{\sqrt{100}}} = \frac{-0.16}{0.08} = -2.0$$

Hypothesis Testing Example (continued)

- **6. Is the test statistic in the rejection region?**



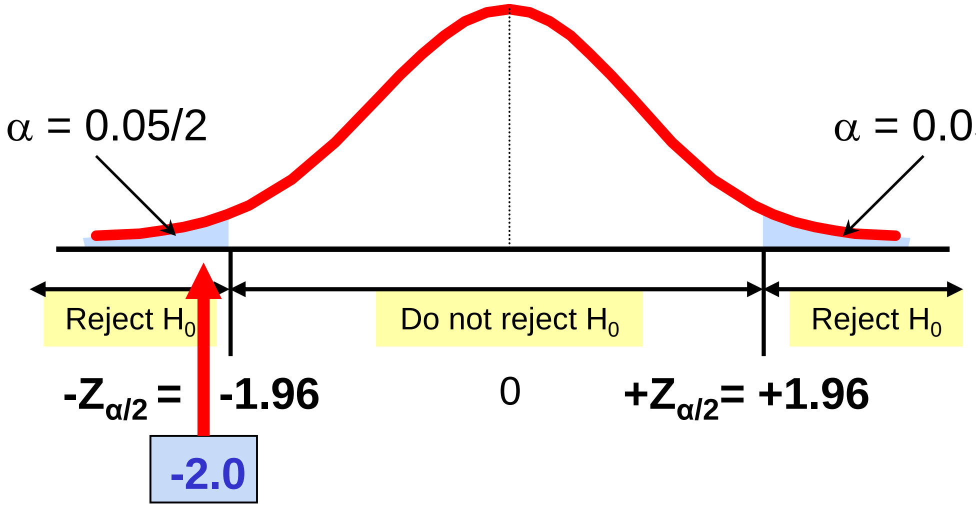
Reject H_0 if $Z_{STAT} < -1.96$ or $Z_{STAT} > 1.96$; otherwise do not reject H_0.

Here, $Z_{STAT} = -2.0 < -1.96$, so the test statistic is in the rejection region.

Hypothesis Testing Example

(continued)

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



Since Z_{STAT} = -2.0 < -1.96, reject the null hypothesis and conclude there is sufficient evidence that the mean diameter of a manufactured bolt is not equal to 30.

p-Value Approach to Testing

DCOVA

- 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.

p-Value Approach to Testing: Interpreting the p-value

DCOVA

- Compare the p-value with α:
  - If p-value $< \alpha$, reject H_0.
  - If p-value $\geq \alpha$, do not reject H_0.
- Remember
  - If the p-value is low then H_0 must go.

The 5 Step p-value approach to Hypothesis Testing

DCOVA

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. The level of significance is based on the relative importance of the risks of a type I and a type II error.**
3. **Determine the appropriate test statistic and sampling distribution.**
4. **Collect the sample data, compute the value of the test statistic and the p-value.**
5. **Make the statistical decision and state the managerial conclusion in the context of the theory, claim, or assertion being tested. If the p-value is $< \alpha$ reject H_0.**

p-value Hypothesis Testing Example

DCOVA

Test the claim that the true mean diameter of a manufactured bolt is 30mm. (Assume $\sigma = 0.8$)

1. State the appropriate null and alternative hypotheses:
   - $H_0: \mu = 30$ $H_1: \mu \neq 30$ (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.

p-value Hypothesis Testing Example (continued)

3. Determine the appropriate technique:
   - σ is assumed known so this is a Z test.
4. Collect the data, compute the test statistic and the p-value.
   - Suppose the sample results are:

     $n = 100, \quad \bar{X} = 29.84$ ($\sigma = 0.8$ is assumed known.)

So the test statistic is:

$$Z_{STAT} = \frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}} = \frac{29.84 - 30}{\frac{0.8}{\sqrt{100}}} = \frac{-0.16}{0.08} = -2.0$$

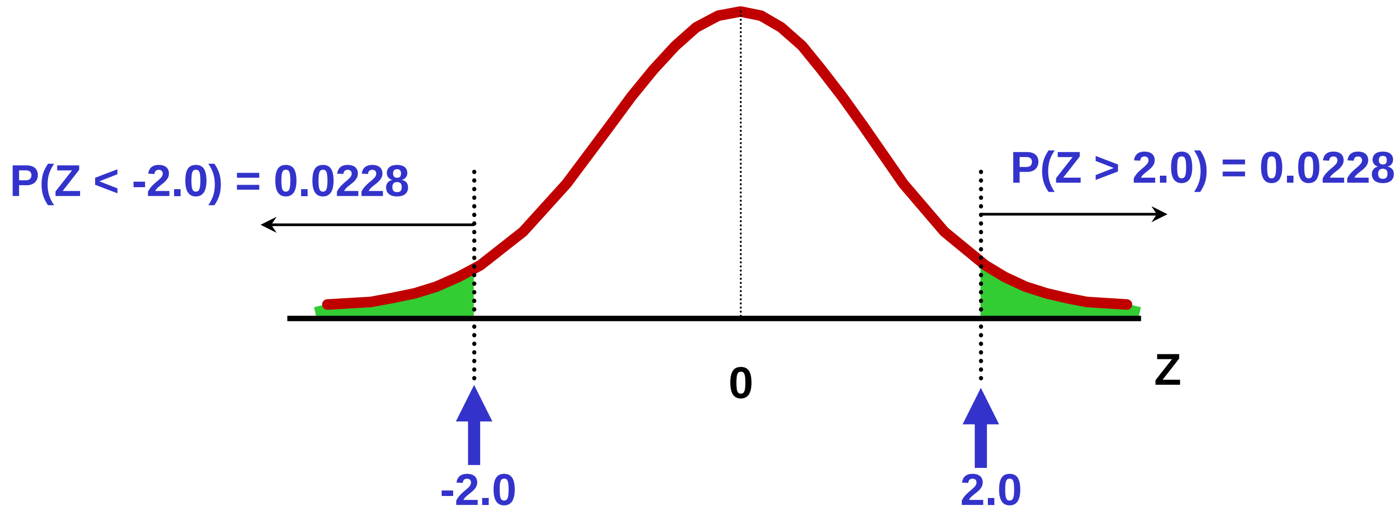
p-Value Hypothesis Testing Example: Calculating the p-value

(continued)

DCOVA

4. (continued) Calculate the p-value.

- How likely is it to get a Z_{STAT} of -2 (or something further from the mean (0), in either direction) if H_0 is true?



p-value = 0.0228 + 0.0228 = 0.0456

p-value Hypothesis Testing Example

(continued)

DCOVA

- 5. Is the p-value < α?
  - Since p-value = 0.0456 < α = 0.05 Reject H_0.
- 5. (continued) State the managerial conclusion in the context of the situation.
  - There is sufficient evidence to conclude the mean diameter of a manufactured bolt is not equal to 30mm.

Connection Between Two Tail Tests and Confidence Intervals

- For $\bar{X} = 29.84$, $\sigma = 0.8$ and $n = 100$, the 95% confidence interval is:

$$\boxed{29.84 - (1.96)\frac{0.8}{\sqrt{100}} \quad \text{to} \quad 29.84 + (1.96)\frac{0.8}{\sqrt{100}}}$$

$$\boxed{29.6832 \le \mu \le 29.9968}$$

- Since this interval does not contain the hypothesized mean (30), we reject the null hypothesis at $\alpha = 0.05$.

Do You Ever Truly Know σ?

DCOVA

- Probably not!
- In virtually all real world business situations, σ is not known.
- If there is a situation where σ is known then μ is also known (since to calculate σ you need to know μ.)
- If you truly know μ there would be no need to gather a sample to estimate it.

Hypothesis Testing for the Mean: σ Unknown

DCOVA

- If the population standard deviation is unknown, you instead use the sample standard deviation S.
- Because of this change, you use the t distribution instead of the Z distribution to test the null hypothesis about the mean.
- When using the t distribution you must assume the population you are sampling from follows a normal distribution.
- All other steps, concepts, and conclusions are the same.

t Test of Hypothesis for the Mean (σ Unknown)

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

Hypothesis Tests for μ

σ Known (Z test)

σ Unknown (t test)

The test statistic is:

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

Example: Two-Tail Test (σ Unknown)

DCOVA

The mean cost of a hotel room in New York is said to be \$168 per night. To determine if this is true, a random sample of 25 hotels is taken and resulted in an $\bar{X}$ of \$172.50 and an S of \$15.40. Test the appropriate hypotheses at $\alpha = 0.05$.

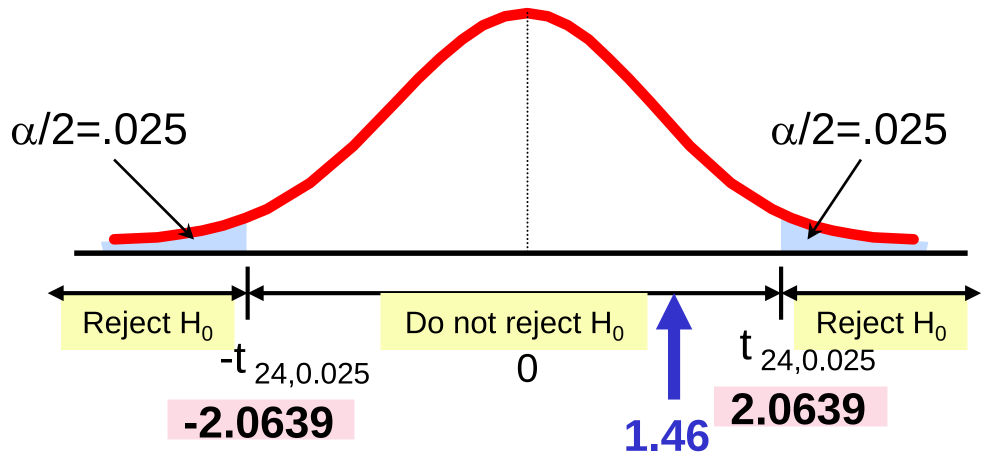
(Assume the population distribution is normal)

H_0: $\mu = 168$

H_1: $\mu \neq 168$

Example Solution: Two-Tail t Test

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



- $\alpha = 0.05$.
- $n = 25$, $df = 25-1=24$.
- σ is unknown, so use a **t statistic.**
- Critical Value:

$\pm t_{24,0.025} = \pm\ 2.0639$.

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

Do not reject H_0: insufficient evidence that true mean cost is different from \$168.

To Use the t-test Must Assume the Population Is Normal

- As long as the sample size is not very small and the population is not very skewed, the t-test can be used.
- To evaluate the normality assumption:
  - Determine how closely sample statistics match the normal distribution's theoretical properties.
  - Construct a histogram or stem-and-leaf display or boxplot or a normal probability plot.
  - Section 6.3 has more details on evaluating normality.

Example Two-Tail t Test Using A p-value from Excel, JMP, & Minitab

- Since this is a t-test we cannot calculate the p-value without some calculation aid.
- The Excel, JMP, & Minitab output below all do this:

	A	B
1	t Test for the Hypothesis of the Mean	
2		
3	Data	
4	Null Hypothesis μ=	120
5	Level of Significance	0.05
6	Sample Size	12
7	Sample Mean	112.85
8	Sample Standard Deviation	20.8
9		
10	Intermediate Calculations	
11	Standard Error of the Mean	6.0044
12	Degrees of Freedom	11
13	*t* Test Statistic	-1.1908
14		
15	Two-Tail Test	
16	Lower Critical Value	-2.2010
17	Upper Critical Value	2.2010
18	*p*-Value	0.2588
19	Do not reject the null hypothesis	

Test Inputs

Hypothesized Mean	120
Sample Average	112.85
Sample Standard Deviation	20.8
Sample Size	12
Significance Level (alpha)	0.05

Test Results

Result	Value
Standard Error of the Mean	6.0044
t-score	-1.1908
t Critical Values	+/- 2.201
Observed Significance (p-value)	0.2588

Fail to Reject Null Hypothesis

One-Sample T

Descriptive Statistics

N	Mean	StDev	SE Mean	95% CI for μ
12	112.85	20.80	6.00	(99.63, 126.07)

μ: mean of Sample

Test

Null hypothesis H_0: $\mu = 120$

Alternative hypothesis H_1: $\mu \neq 120$

T-Value	P-Value
-1.19	0.259

Connection of Two Tail Tests to Confidence Intervals

DCOVA

- For $\bar{X} = 172.5$, $S = 15.40$ and $n = 25$, the 95% confidence interval for μ 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 $\alpha = 0.05$.

One-Tail Tests

DCOVA

- 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: $\mu \leq 3$

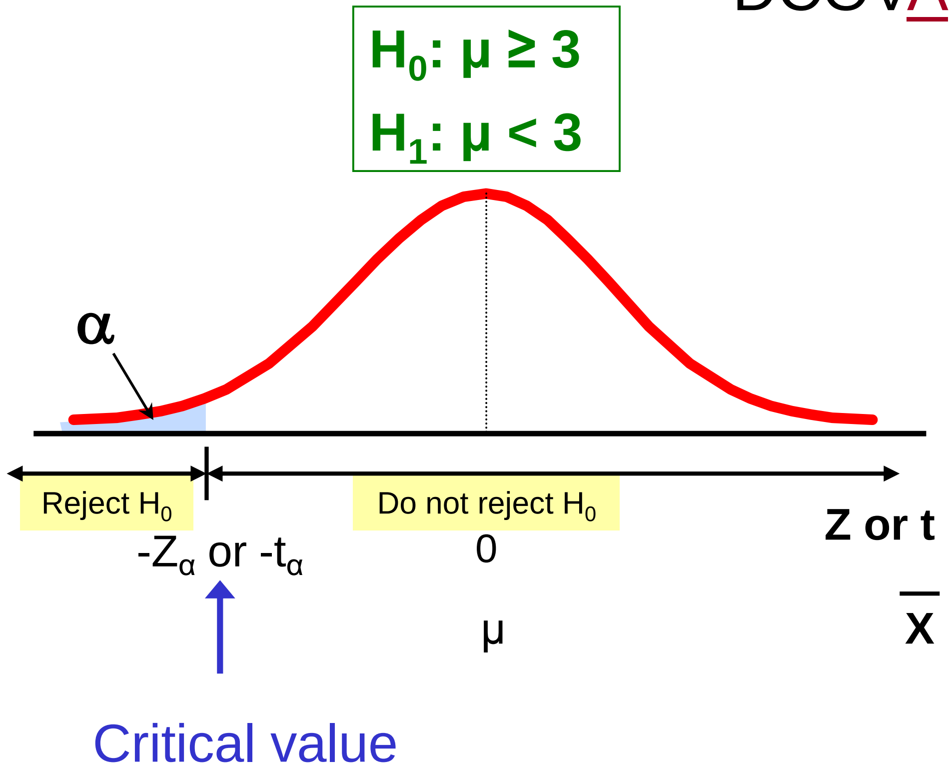
H_1: $\mu > 3$

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

Lower-Tail Tests

DCOVA

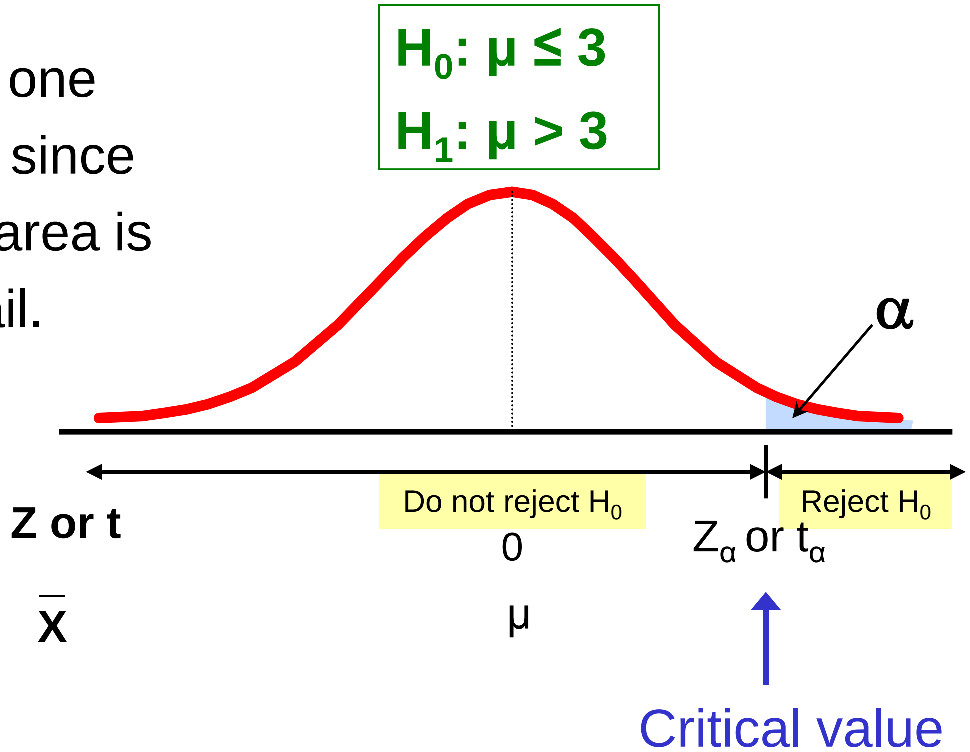
- There is only one critical value, since the rejection area is in only one tail.



Upper-Tail Tests

DCOVA

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



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

DCOVA

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.)

$H_0: \mu \le 52$

$H_1: \mu > 52$

Form hypothesis test:

H_0: μ ≤ 52	the mean is not over $52 per month
H_1: μ > 52	the mean **is** greater than $52 per month (i.e., sufficient evidence exists to support the manager's claim)

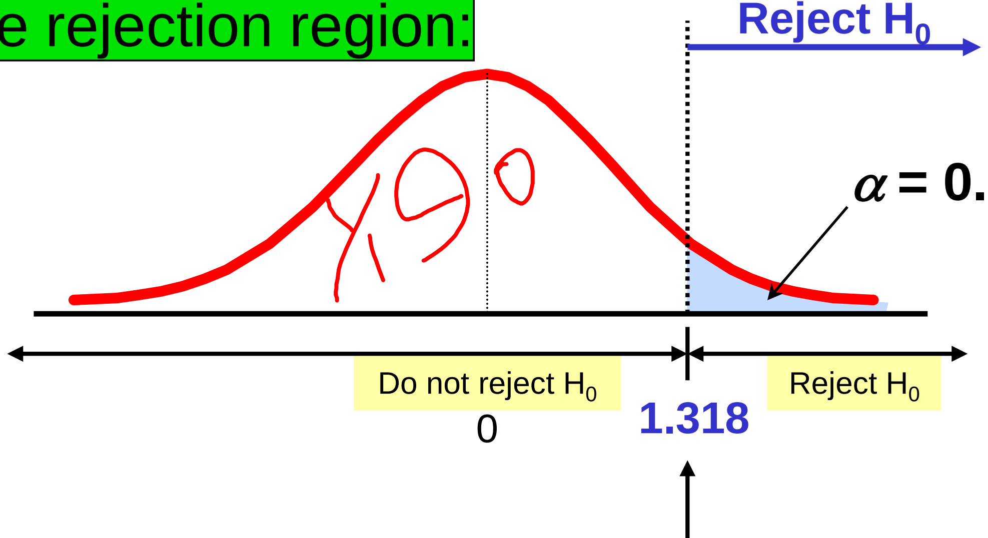
Example: Find Rejection Region

(continued)

DCOVA

- Suppose that $\alpha = 0.10$ is chosen for this test and $n = 25$.

Find the rejection region:



Reject H_0 if $t_{STAT} > 1.318$

Example: Test Statistic

(continued)

DCOVA

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:

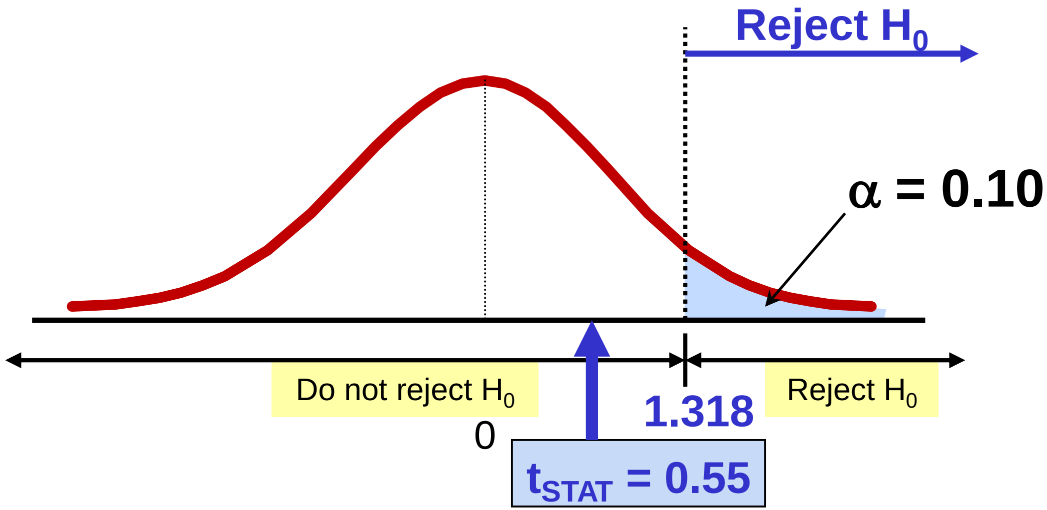
$$t_{STAT} = \frac{\bar{X} - \mu}{\frac{S}{\sqrt{n}}} = \frac{53.1 - 52}{\frac{10}{\sqrt{25}}} = 0.55$$

Example: Decision

(continued)

DCOVA

Reach a decision and interpret the result.



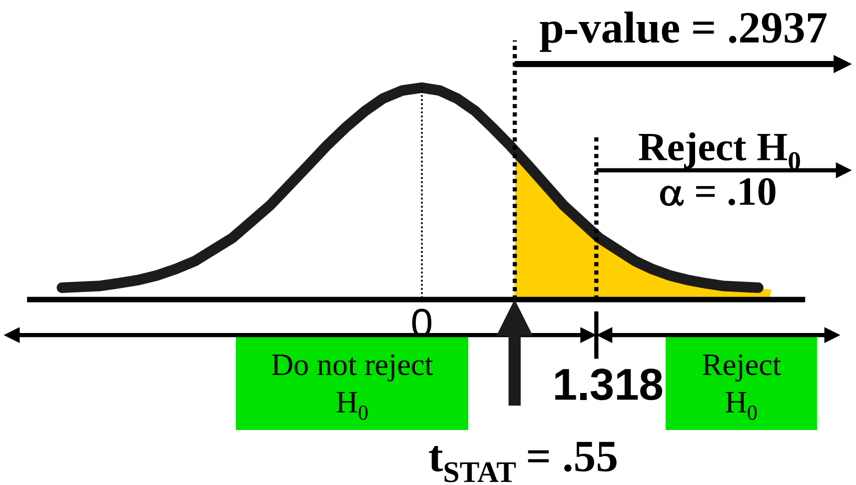
Do not reject H_0 since $t_{STAT} = 0.55 < 1.318$.

There is not sufficient evidence that the mean bill is over $52.

Example: Utilizing The p-value for The Upper Tail t-Test

DCOVA

- Calculate the p-value and compare to α (p-value calculation via Excel, Minitab, & JMP shown on next page)



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

Calculating The p-value for The Upper Tail t-Test In Excel, JMP, & Minitab

DCOVA

	A	B
1	**t Test for the Hypothesis of the Mean**	
2		
3	**Data**	
4	**Null Hypothesis** μ =	**52**
5	**Level of Significance**	**0.10**
6	**Sample Size**	**25**
7	**Sample Mean**	**53.1**
8	**Sample Standarad Deviation**	**10**
9		
10	**Intermediate Calculations**	
11	Standard Error of the Mean	2.00
12	Degrees of Freedom	24
13	**t Test Statistic**	**0.55**
14		
15	**Upper Tail Test**	
16	**Upper Critical Value**	**1.318**
17	**p-value**	**0.2937**
18	**Do not reject the null hypothesis**	

Test Inputs

Hypothesized Mean	52
Sample Average	53.1
Sample Standard Deviation	10
Sample Size	25
Significance Level (alpha)	0.1

Test Results

Result	Value
Standard Error of the Mean	2
t-score	0.55
t Critical Values	1.3178
Observed Significance (p-value)	0.2937

Fail to Reject Null Hypothesis

One-Sample t

Descriptive Statistics

N	Mean	SE Mean	95% CI for μ
25	53.1	2.0	(49.678, 56.522)

μ: mean of sample

Test

Null hypothesis $H_0: \mu = 52$

Alternative hypothesis $H_1: \mu > 52$

t-Value	P-Value
0.55	0.2937

Hypothesis Tests for Proportions

DCOVA

- 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

DCOVA

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

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

Hypothesis Tests for p

$n\pi \geq 5$ and $n(1-\pi) \geq 5$

$n\pi < 5$ or $n(1-\pi) < 5$

Not discussed in this book.

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

DCOVA

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

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

Hypothesis Tests for X

$X \geq 5$ and $n-X \geq 5$

$X < 5$ or $n-X < 5$

Not discussed in this book.

Example: Z Test for Proportion

DCOVA

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.

Check:

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

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

✓

Z Test for Proportion: Solution

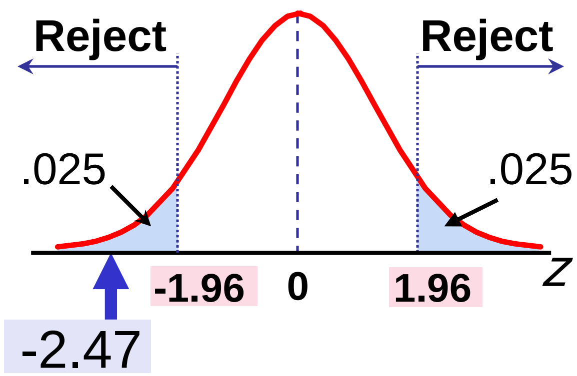
DCOVA

H_0: $\pi = 0.08$
H_1: $\pi \neq 0.08$

$\alpha = 0.05$

$n = 500$, $p = 0.05$

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 $\alpha = 0.05$

Conclusion:

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

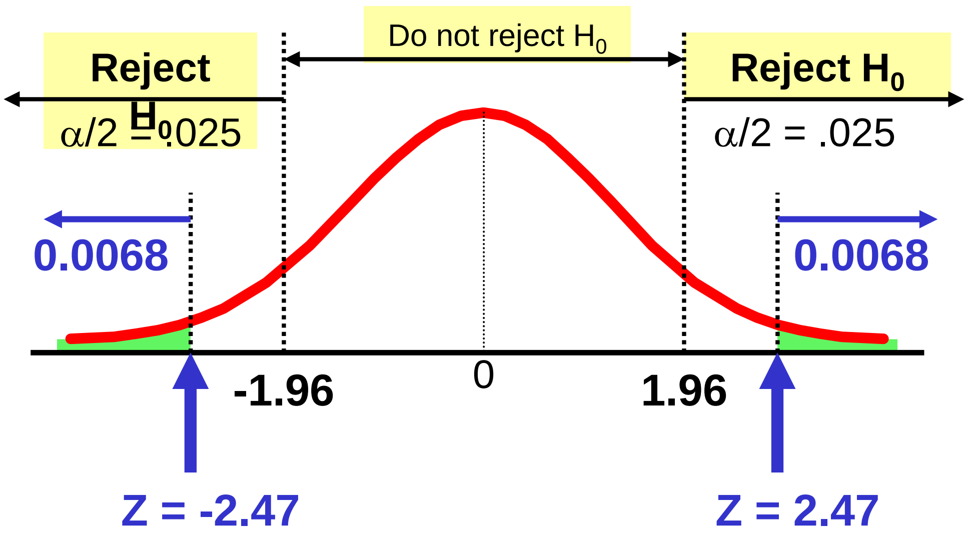
p-Value Solution

(continued)

DCOVA

Calculate the p-value and compare to α

(For a two-tail test the p-value is always two-tail.)



p-value = 0.0136:

$$P(Z \le -2.47) + P(Z \ge 2.47)$$
$$= 2(0.0068) = 0.0136$$

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

Questions To Address In The Planning Stage

- What is the goal of the survey, study, or experiment?
- How can you translate this goal into a null and an alternative hypothesis?
- Is the hypothesis test one or two tailed?
- Can a random sample be selected?
- What types of data will be collected? Numerical? Categorical?
- What level of significance should be used?
- Is the intended sample size large enough to achieve the desired power?
- What statistical test procedure should be used and why?
- What conclusions & interpretations can you reach from the results of the planned hypothesis test?

Failing to consider these questions can lead to bias or incomplete results.

Statistical Significance vs Practical Significance

- Statistically significant results (rejecting the null hypothesis) are not always of practical significance.
  - This is more likely to happen when the sample size gets very large.
- Practically important results might be found to be statistically insignificant (failing to reject the null hypothesis.)
  - This is more likely to happen when the sample size is relatively small.

Reporting Findings & Ethical Issues

- Should document & report both good & bad results.
- Should not just report statistically significant results.
- Reports should distinguish between poor research methodology and unethical behavior.
- Ethical issues can arise in:
  - The use of human subjects.
  - The data collection method.
  - The type of test being used.
  - The level of significance being used.
  - The cleansing and discarding of data.
  - The failure to report pertinent findings.

Chapter Summary

In this chapter we discussed:

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