Hypothesis Testing

Chapter 9 (9.1 - 9.3)

Learning Objectives

In this chapter, you learn:

- The basic principles of hypothesis testing
- How to test population means
- Required assumptions
- How to read and interpret results of hypothesis testing

What is a Hypothesis?

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

  - **population mean**

Example: The mean monthly cell phone bill in this city is μ = $72

  - **population proportion**

Example: The proportion of adults in this city with cell phones is $\pi = 0.78$

The Null Hypothesis, H_0

- 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

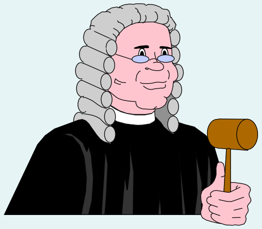
$$H_0 : \mu = 30$$

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

The Null Hypothesis, H_0

(continued)

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



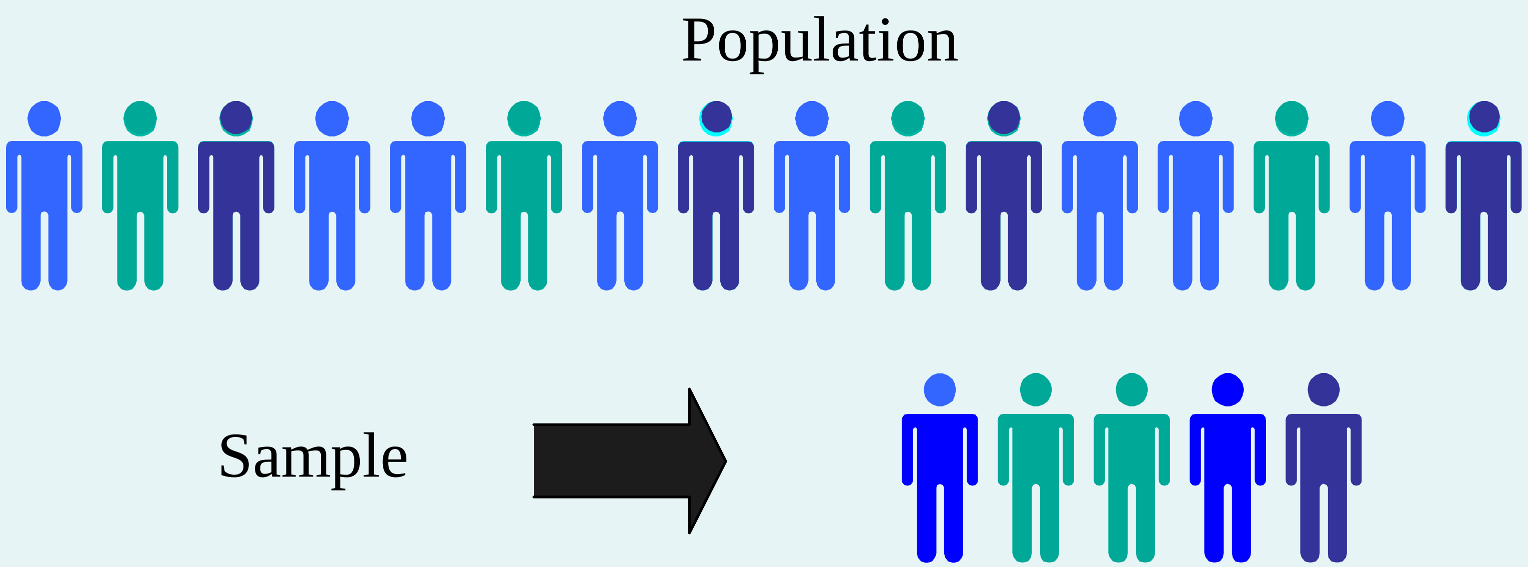
- Refers to the status quo or historical value
- Always contains "=", or "≤", or "≥" sign
- May or may not be rejected, according to the test results

The Alternative Hypothesis, H_1

- Is the opposite of the null hypothesis
  - e.g., The average diameter of a manufactured bolt is not equal to 30mm ($H_1: \mu \neq 30$)
- Challenges the status quo
- Contains "≠", or "<", or ">" sign
- May be supported or not supported by evidence that comes from data

The Hypothesis Testing Process

- Claim: The population mean age is 50.
  - H_0: μ = 50, H_1: μ ≠ 50
- Sample the population and find sample mean.



The Hypothesis Testing Process

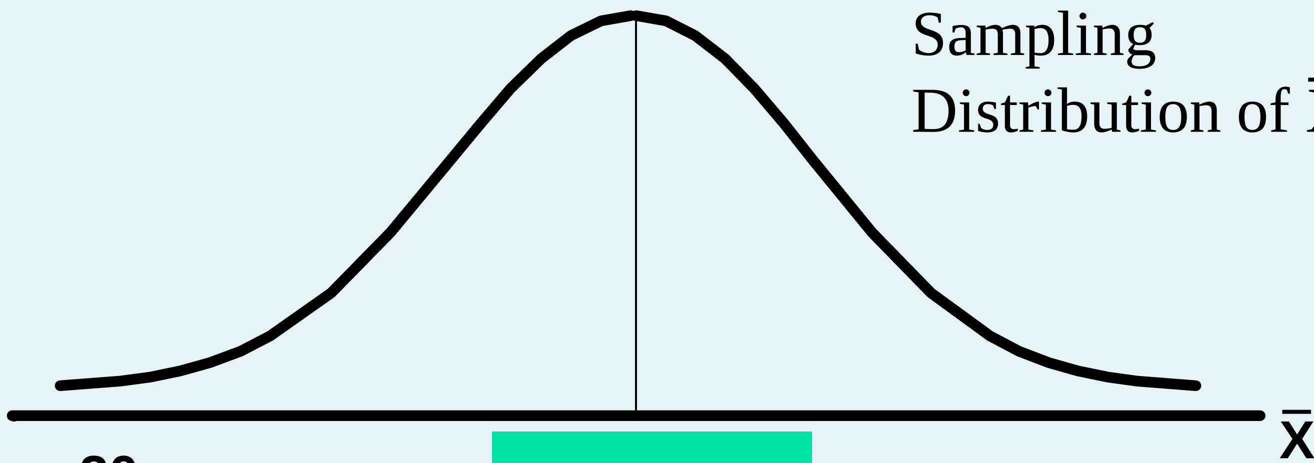
(continued)

- 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)

Sampling Distribution of $\bar{X}$



20

μ = 50
If H_0 is true

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

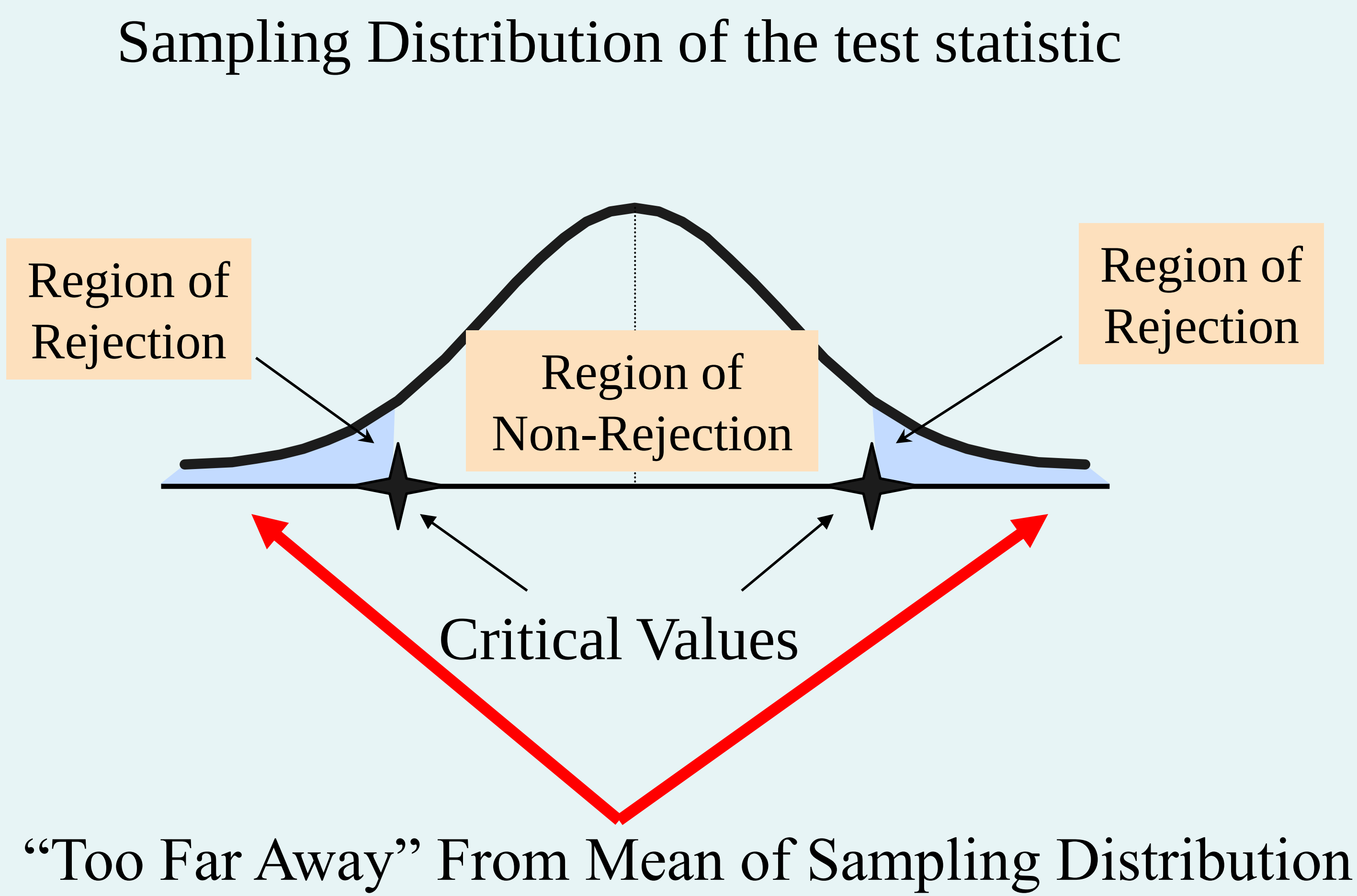
... When in fact this were the population mean…

... then you reject the null hypothesis that μ = 50.

The Test Statistic and Critical Values

- 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



Possible Errors in Hypothesis Test Decision Making

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

Possible Errors in Hypothesis Test Decision Making

- **Type I Error**
  - Reject a true null hypothesis
  - Considered a serious type of error
  - We control the probability of a Type I Error at α
    - Called level of significance of the test
    - Set by researcher in advance
- **Type II Error**
  - Failure to reject a false null hypothesis
  - The probability of a Type II Error is $\boldsymbol{\beta}$
  - We try to minimize $\boldsymbol{\beta}$ while guaranteeing the given α

Possible Errors in Hypothesis Test Decision Making

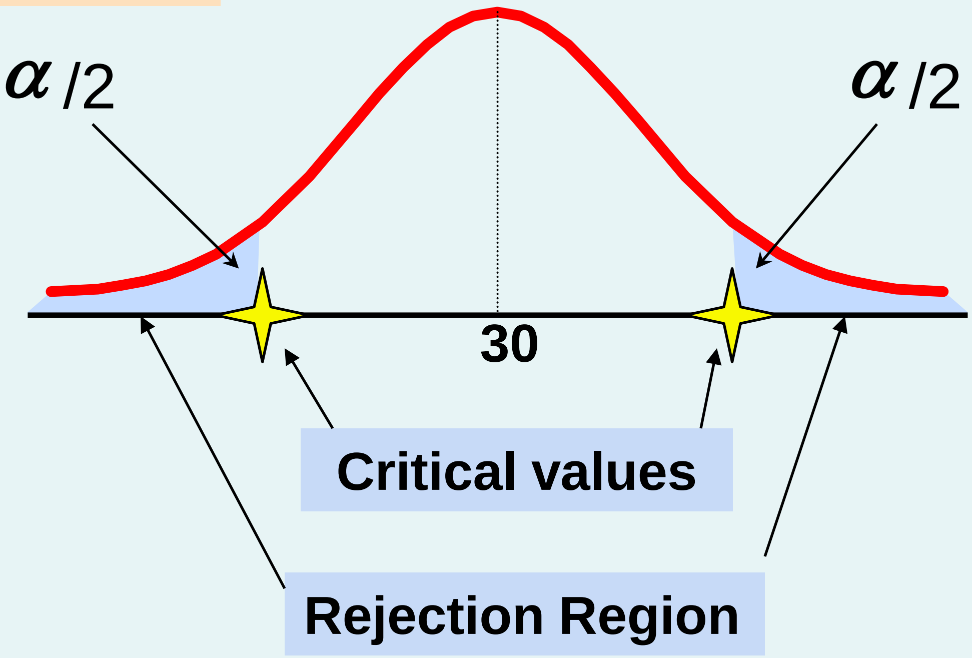
(continued)

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

Level of Significance and the Rejection Region

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

Level of significance = α

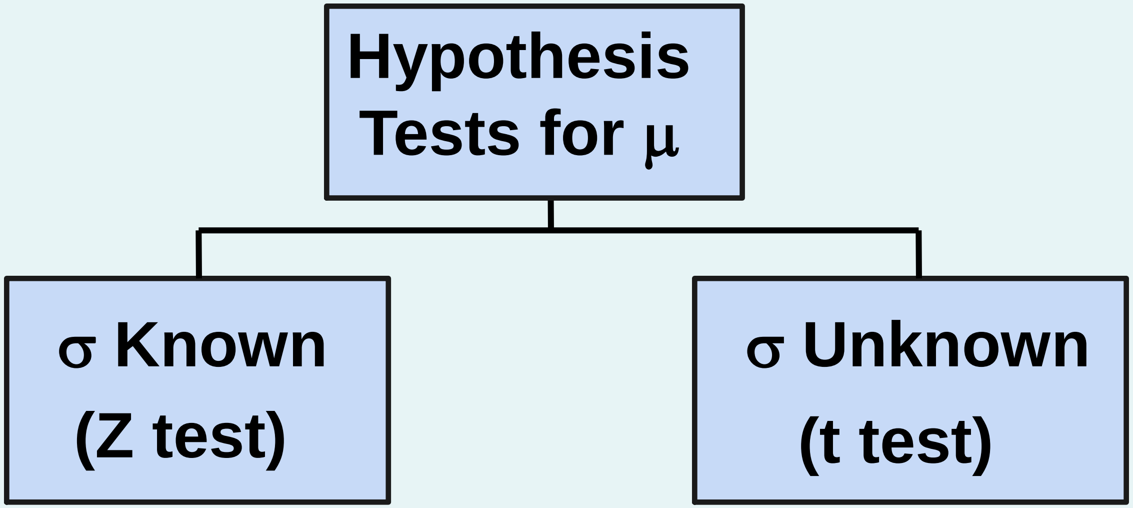


This is a two-tail test because there is a rejection region in both tails

Examples

9.10, 9.11 on p. 317

Hypothesis Tests for the Mean



Z Test of Hypothesis for the Mean (σ Known)

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

Hypothesis Tests for μ

σ Known (Z test)

σ Unknown (t test)

The test statistic is:

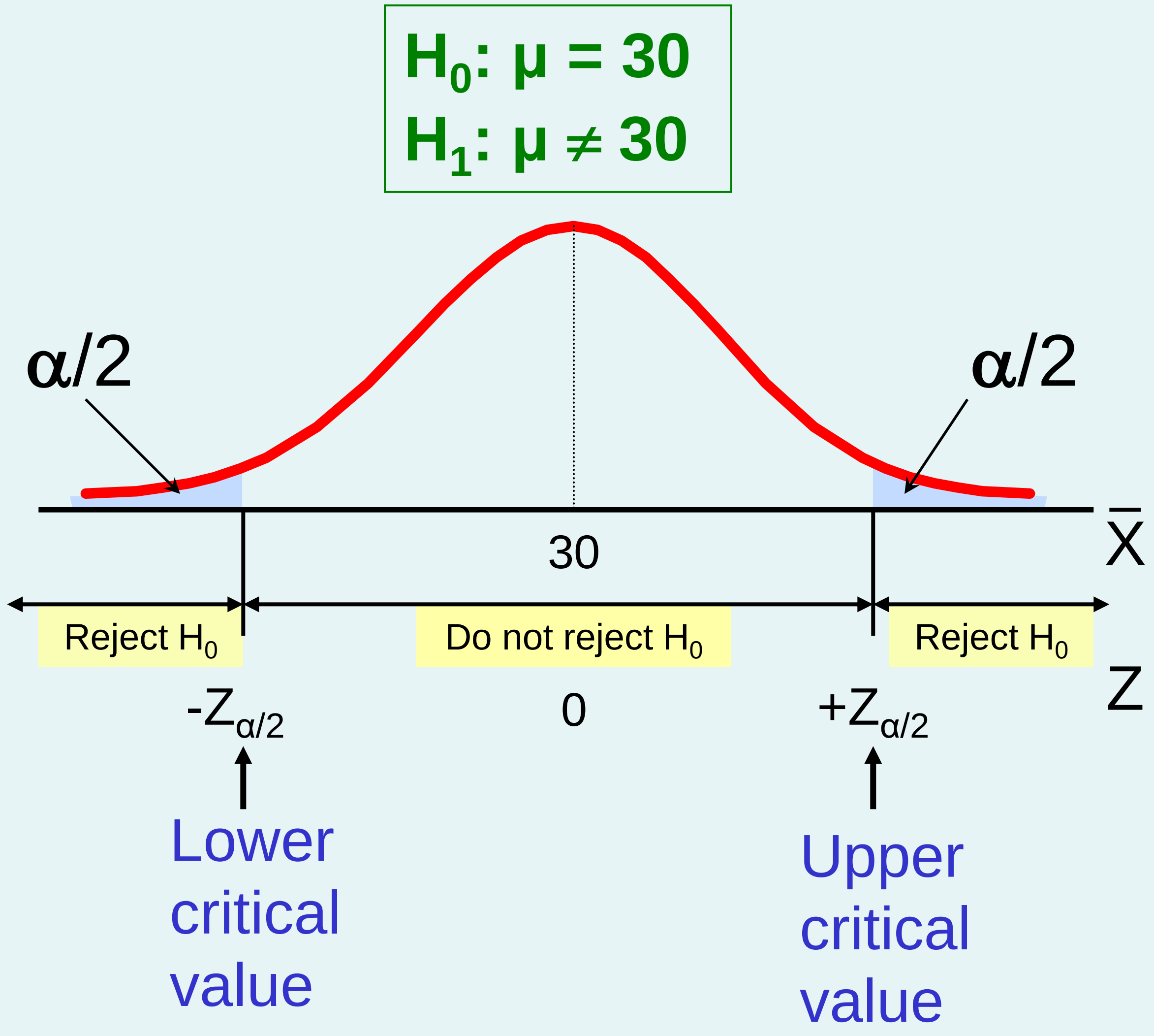
$$Z_{STAT} = \frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}}$$

Critical Value Approach to Testing

- 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 computer
- Decision Rule: If the test statistic falls in the rejection region, reject H_0 ; otherwise do not reject H_0

Two-Tail Tests

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



6 Steps in Hypothesis Testing

1. State the null hypothesis, H_0 and the alternative hypothesis, H_1
2. Choose the level of significance, α, and the sample size, n
3. Determine the appropriate test statistic and sampling distribution
4. Determine the critical values that divide the rejection and non-rejection (acceptance) regions

6 Steps in Hypothesis Testing

(continued)

5. Collect data and compute the value of the test statistic
6. Make the statistical decision and state the managerial conclusion. If the test statistic falls into the non-rejection region, *do not reject* the null hypothesis H_0. If the test statistic falls into the rejection region, *reject* the null hypothesis. It means the data provides significant evidence, at level α, to reject H_0 in favor of H_1. Express the managerial conclusion in the context of the problem

Hypothesis Testing Example

Test the claim that the true mean diameter of a manufactured bolt is 30 mm. (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 level of significance and the sample size
   - Suppose that $\alpha = 0.05$ and $n = 100$ are chosen for this test

Hypothesis Testing Example

(continued)

3. Determine the appropriate technique
   - σ is assumed known so this is a Z test.
4. Determine the critical values
   - For α = 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 (σ = 0.8 is assumed known)

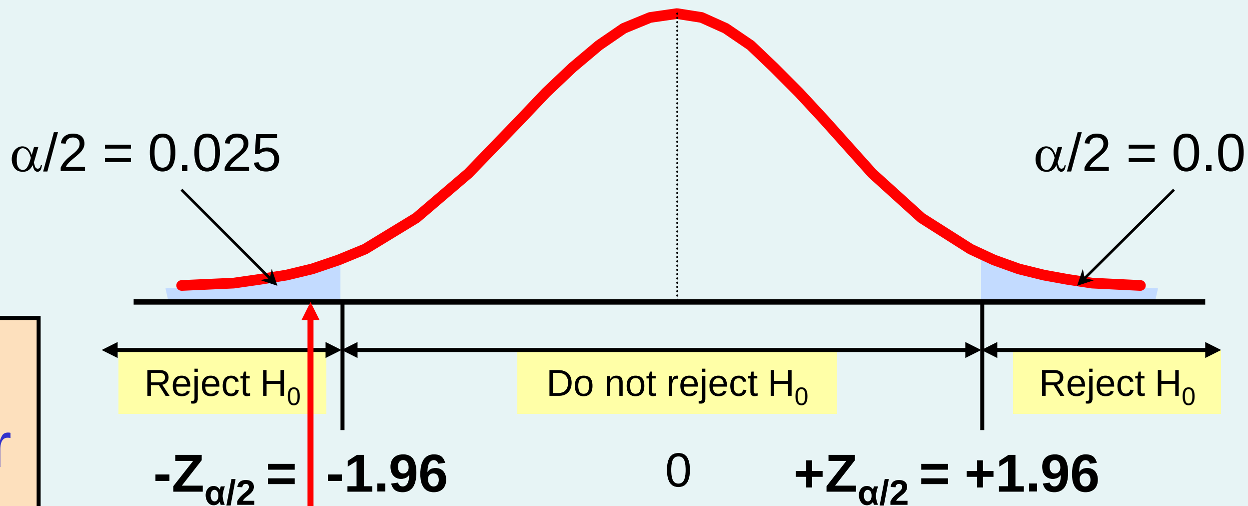
So the test statistic is:

$$Z_{STAT} = \frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}} = \frac{29.84 - 30}{\frac{0.8}{\sqrt{100}}} = \frac{-.16}{0.08} = -2.0$$

Hypothesis Testing Example

(continued)

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



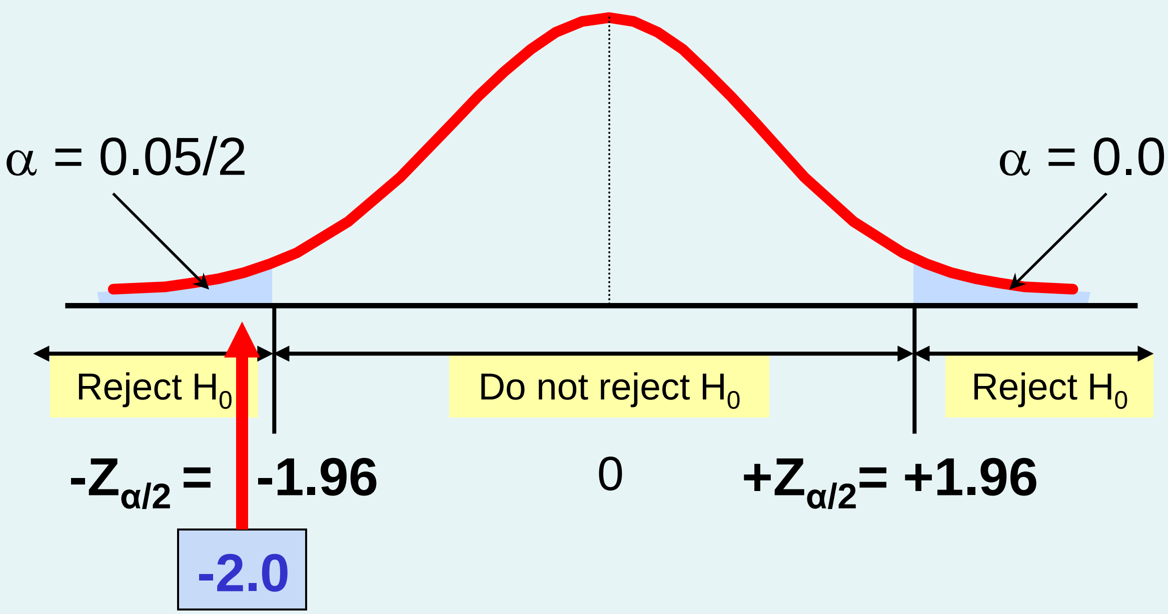
Reject H_0 if $Z_{STAT} < -1.96$ or $Z_{STAT} > 1.96$; otherwise do not reject H_0

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

Examples

9.14a on p. 317

p-Value Approach to Testing

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

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

- 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

1. State the null hypothesis, H_0 and the alternative hypothesis, H_1
2. Choose the level of significance, α, and the sample size, n
3. Determine the appropriate test statistic and sampling distribution
4. Collect data and compute the value of the test statistic and the p-value
5. Make the statistical decision and state the managerial conclusion. If the p-value is $< \alpha$ then reject H_0, otherwise do not reject H_0. State the managerial conclusion in the context of the problem

p-value Hypothesis Testing Example

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

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, $\bar{X}$ = 29.84 (σ = 0.8 is assumed known)

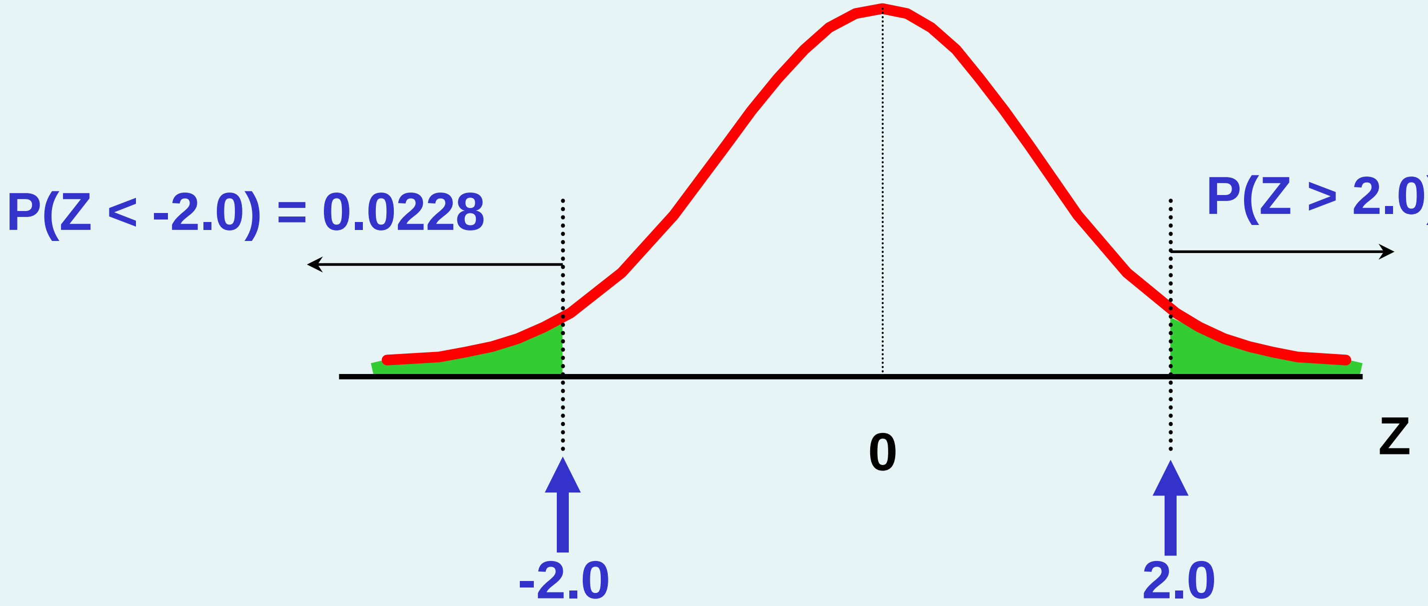
So the test statistic is:

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

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

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)

- 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 average diameter of a manufactured bolt is not equal to 30mm.

Examples

9.14b on p. 317

Example

9.14c on p. 317

Connection Between Two-Tail Tests and Confidence Intervals

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

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

$$29.6832 \le \mu \le 29.9968$$

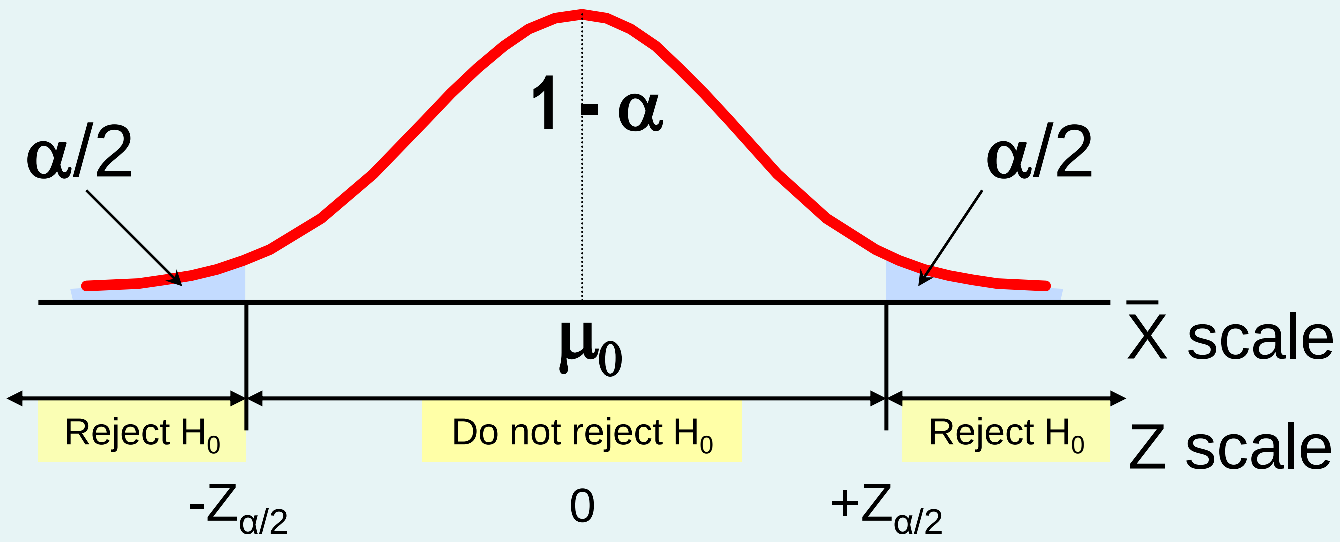
- Since this interval does not contain the hypothesized mean (30), we reject the null hypothesis at $\alpha = 0.05$

Hypothesis Testing: σ Unknown

- 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.

Two-Sided Tests and Confidence Intervals

Reject Ho: $\mu = \mu_0$ at level α in favor of Ho: $\mu \neq \mu_0$
if
the $(1-\alpha)100\%$ confidence interval does not contain μ_0



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)**

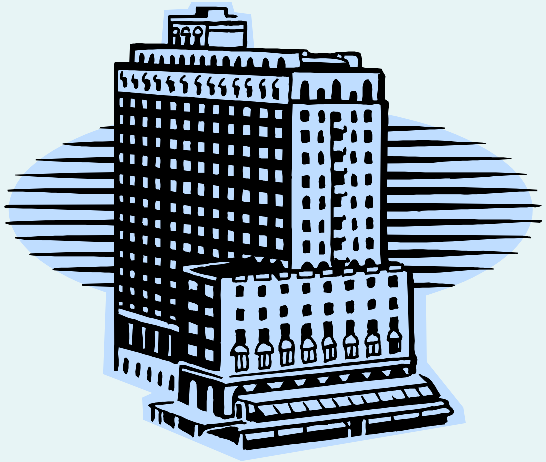
The test statistic is:

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

Example: Two-Tail Test (σ Unknown)

The average 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: $\mu = 168$

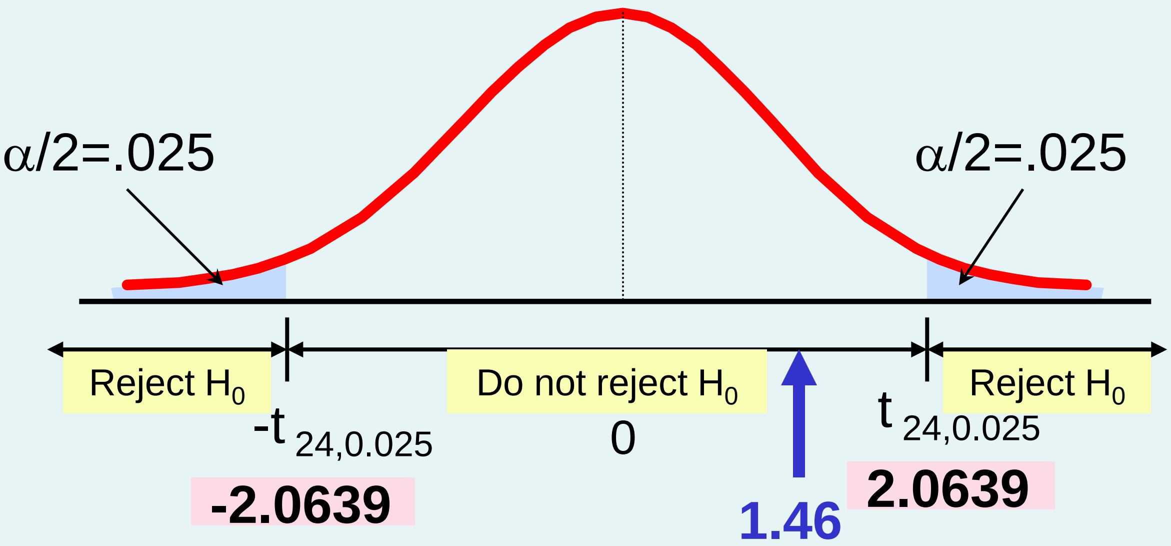
H_1: $\mu \neq 168$

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:

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



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

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

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

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

t Test for the Hypothesis of the Mean

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

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

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

p-value > α
So do not reject H_0

Two-Sided Tests and Confidence Intervals

- 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 $\alpha = 0.05$

Example

9.32 on p. 323

One-Tail Tests

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

H_0: μ ≥ 3

H_1: μ < 3

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

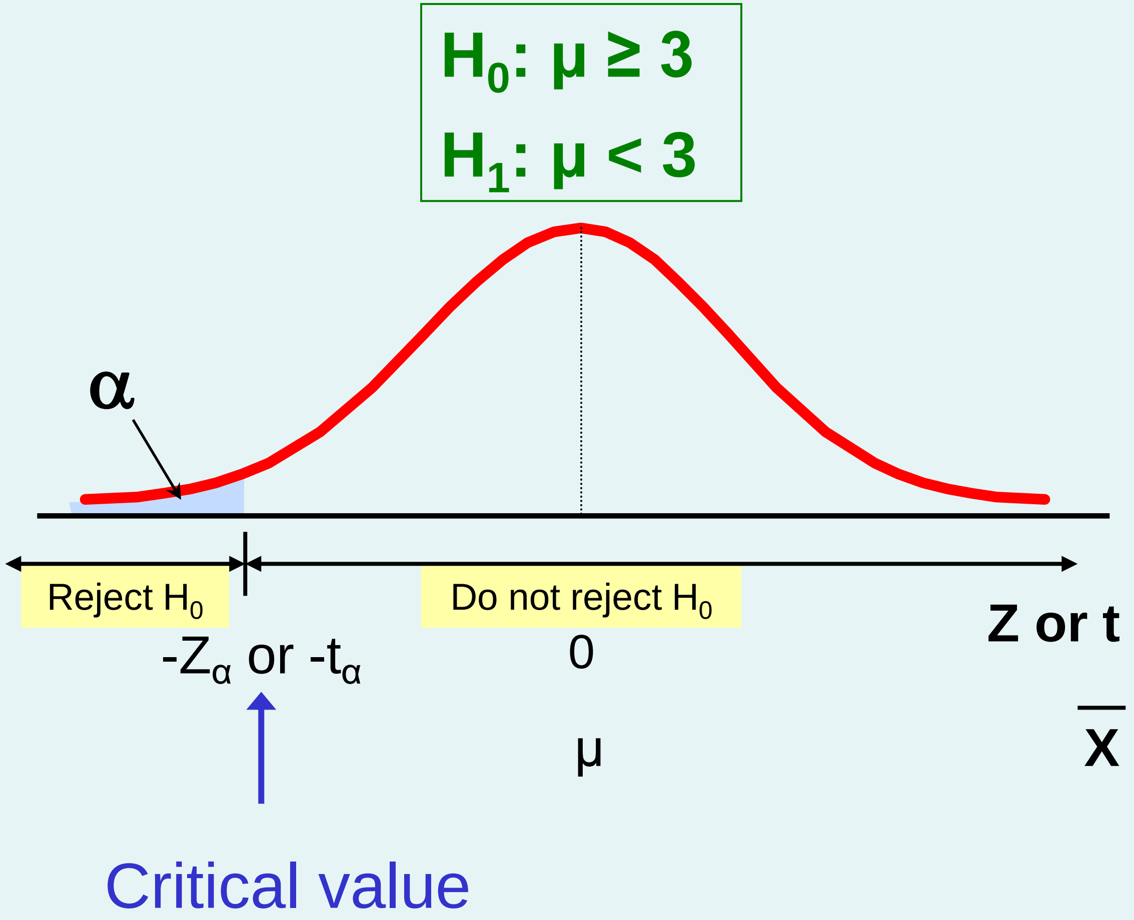
H_0: μ ≤ 3

H_1: μ > 3

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

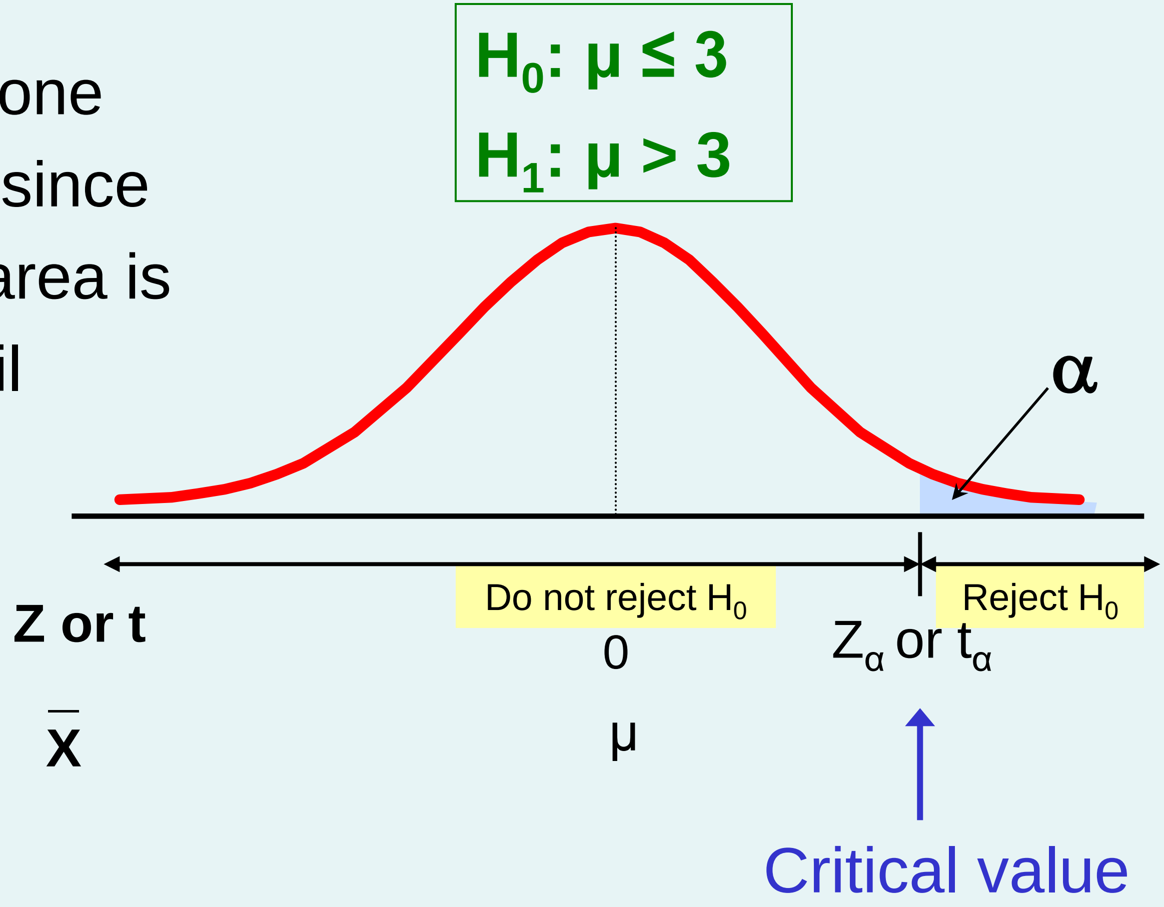
Lower-Tail Tests

- There is only one critical value, since the rejection area is in only one tail
- Use α instead of $\alpha/2$



Upper-Tail Tests

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



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

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

Form hypothesis test:

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

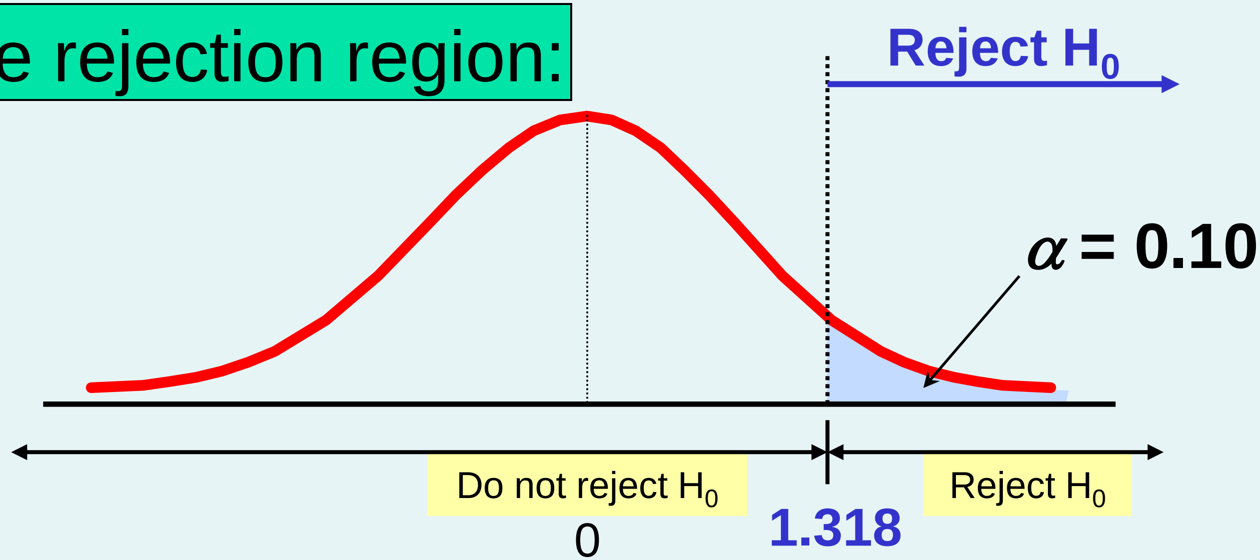
.

Example: Find Rejection Region

(continued)

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

Find the rejection region:



Reject H_0 if $t_{STAT} > 1.318$

Example: Test Statistic

(continued)

Obtain sample and compute the test statistic

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

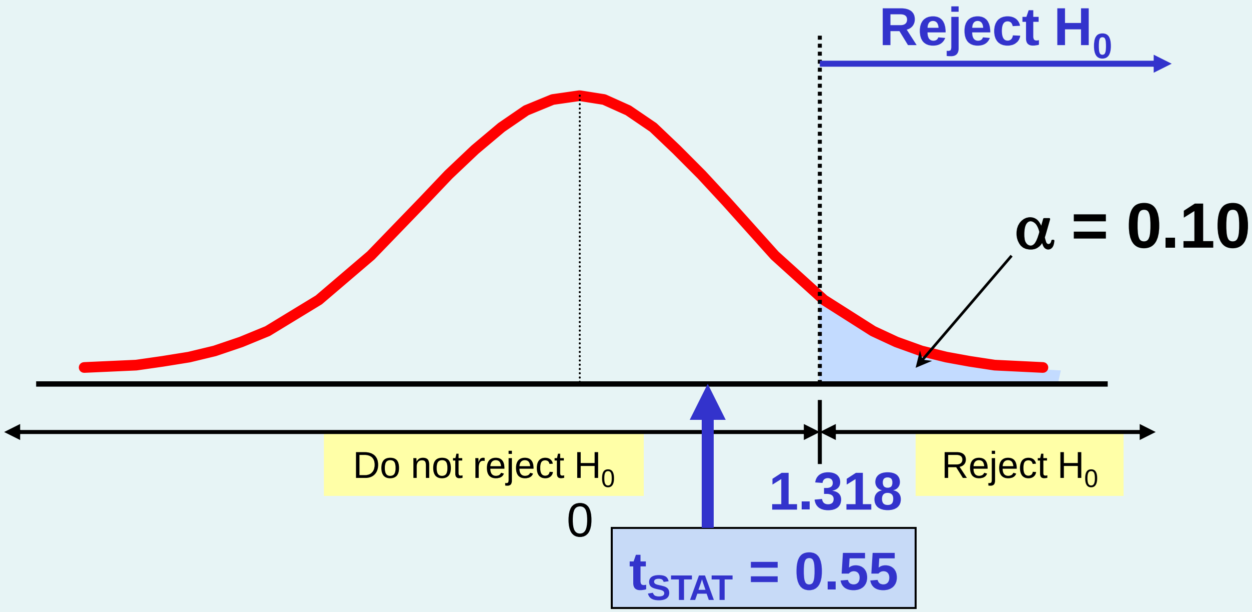
- Then the test statistic is:

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

Example: Decision

(continued)

Reach a decision and interpret the result:

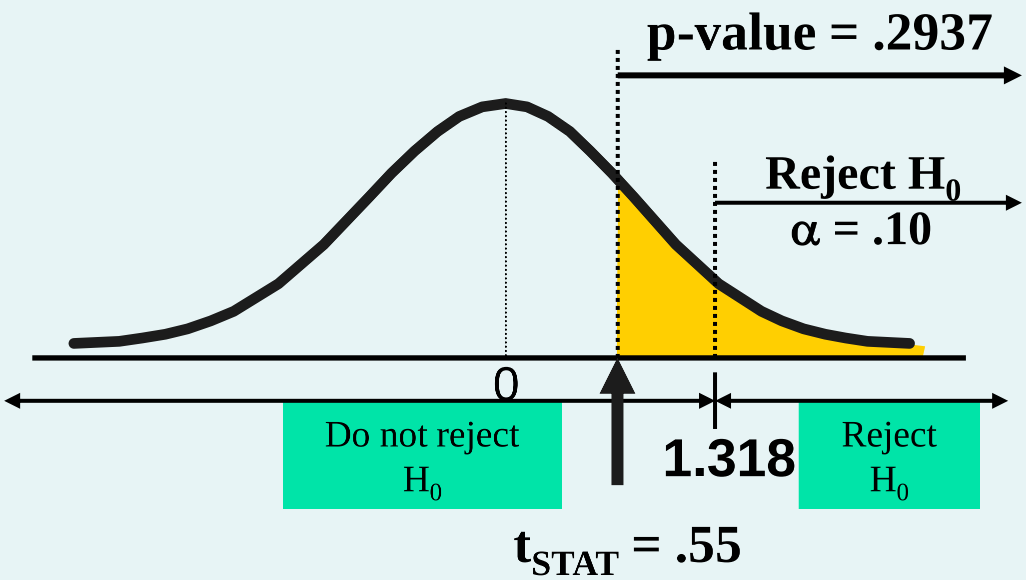


Do not reject H_0 since $t_{STAT} = 0.55 \leq 1.318$

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

Example: Utilizing The p-value for The Test

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



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

Excel Spreadsheet Calculating The p-value for An Upper Tail t Test

	A	B	
1	**t Test for the Hypothesis of the Mean**		
2			
3	**Data**		
4	Null Hypothesis **μ=**	184.2	
5	Level of Significance	0.05	
6	Sample Size	25	
7	Sample Mean	170.8	
8	Sample Standard Deviation	21.3	
9			
10	Intermediate Calculations		
11	Standard Error of the Mean	4.2600	**=B8/SQRT(B6)**
12	Degrees of Freedom	24	**=B6 - 1**
13	*t* Test Statistic	-3.1455	**=(B7 - B4)/B11**
14			
15	**Lower-Tail Test**		
16	Lower Critical Value	-1.7109	**=-T.INV.2T(2 * B5, B12)**
17	*p*-Value	0.0022	**=IF(B13 < 0, E11, E12)**
18	Reject the null hypothesis		**=IF(B17 < B5,"Reject the null hypothesis", "Do not reject the null hypothesis")**

	D	E	
10	One-Tail Calculations		
11	T.DIST.RT value	0.0022	**=T.DIST.RT(ABS(B13), B12)**
12	1-T.DIST.RT value	0.9978	**=1 - E11**

Example

9.50 on p. 328

.