Statistics for business and decision making

Prof. Ilaria Benedetti

Lesson 7. Inferential Statistics: Hypotesis test

What is a Hypothesis?

- A hypothesis is a claim (assumption) 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 $\pi = 0.68$

The Null Hypothesis, H0

- States the claim or assertion to be tested

  Example: The average number of TV sets in U.S. Homes is equal to three $H_0: \mu = 3$

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

$H_0: \mu = 3$ $\qquad$ $H_0: \bar{X} = 3$

The Null Hypothesis, H_0

(continued)

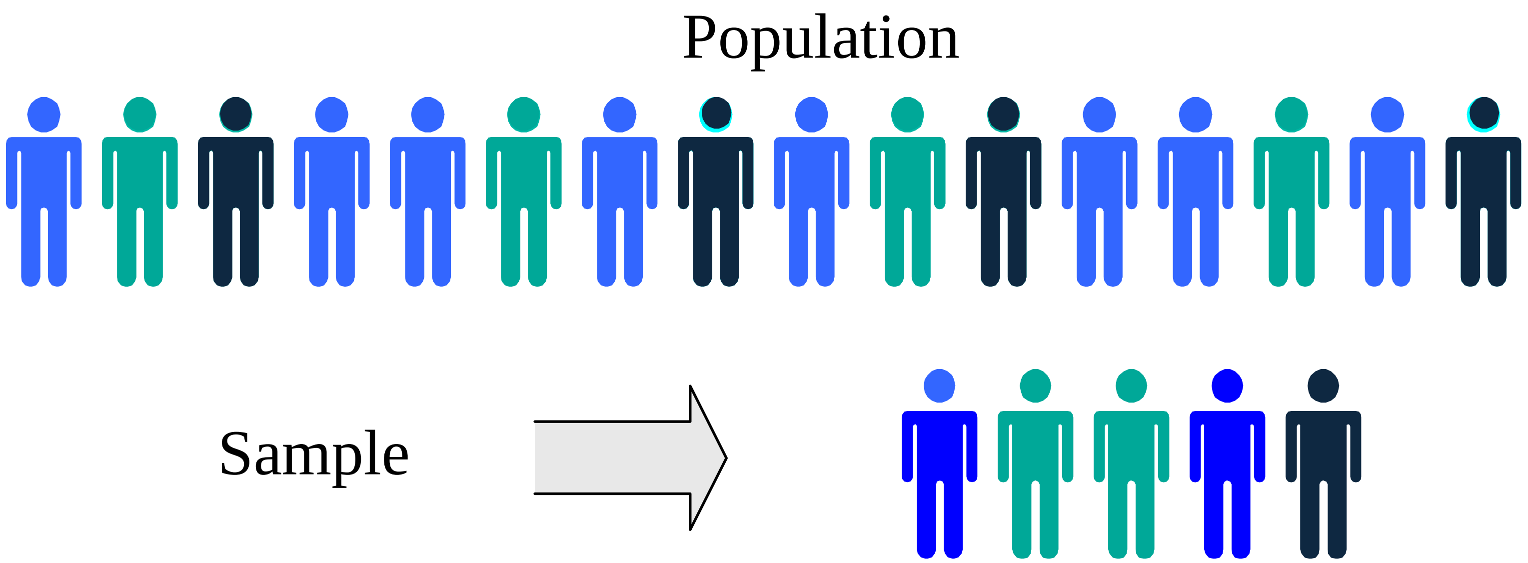
- Begin with the assumption that the null hypothesis is true
- Refers to the status quo or historical value
- Always contains "=" sign
- May or may not be rejected

The Alternative Hypothesis, H_1

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

The Hypothesis Testing Process

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



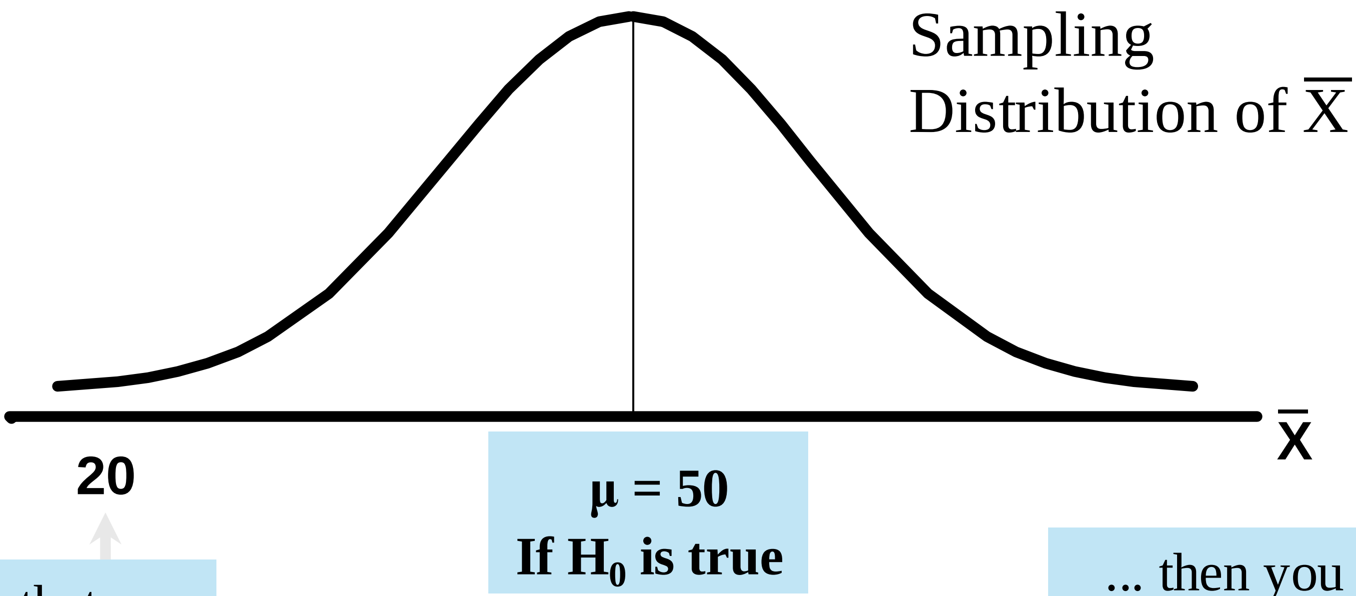
The Hypothesis Testing Process

(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)



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

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

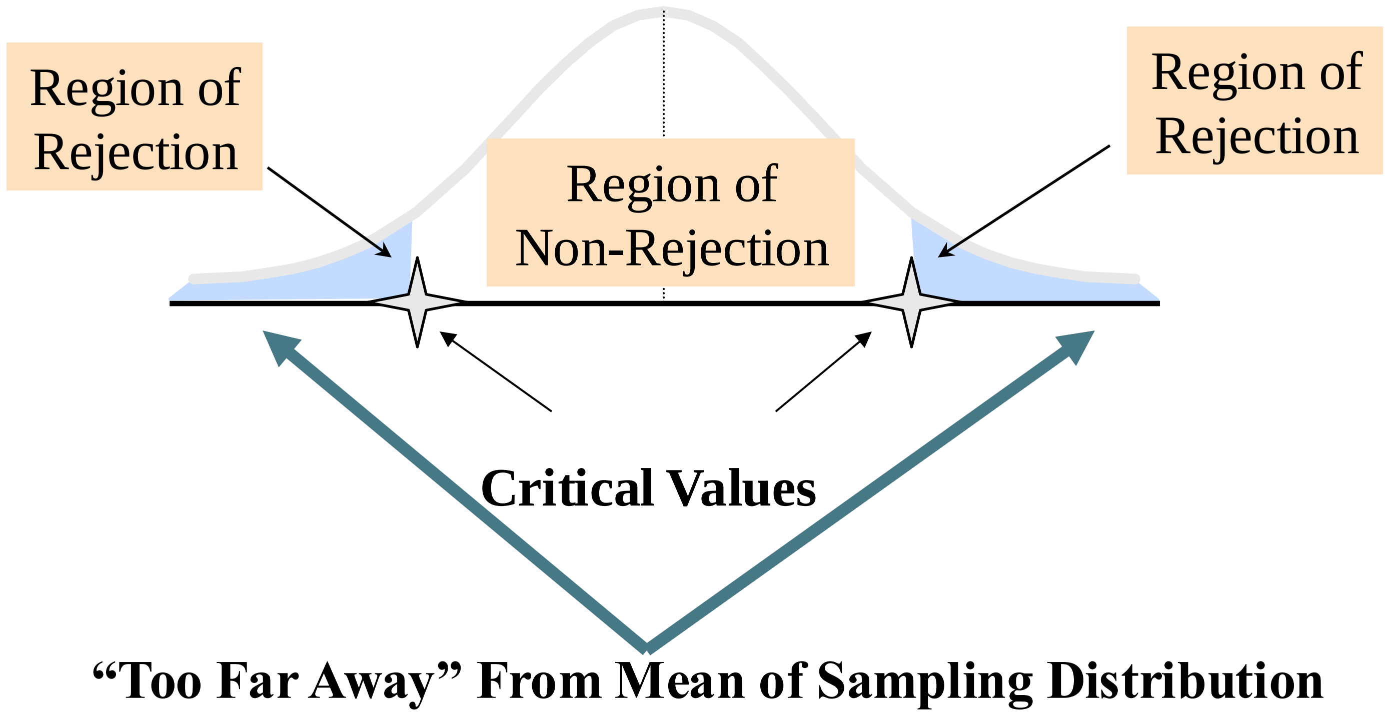
... then you reject the null hypothesis that $\mu = 50$.

The Test Statistic and Critical Values

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

The Test Statistic and Critical Values

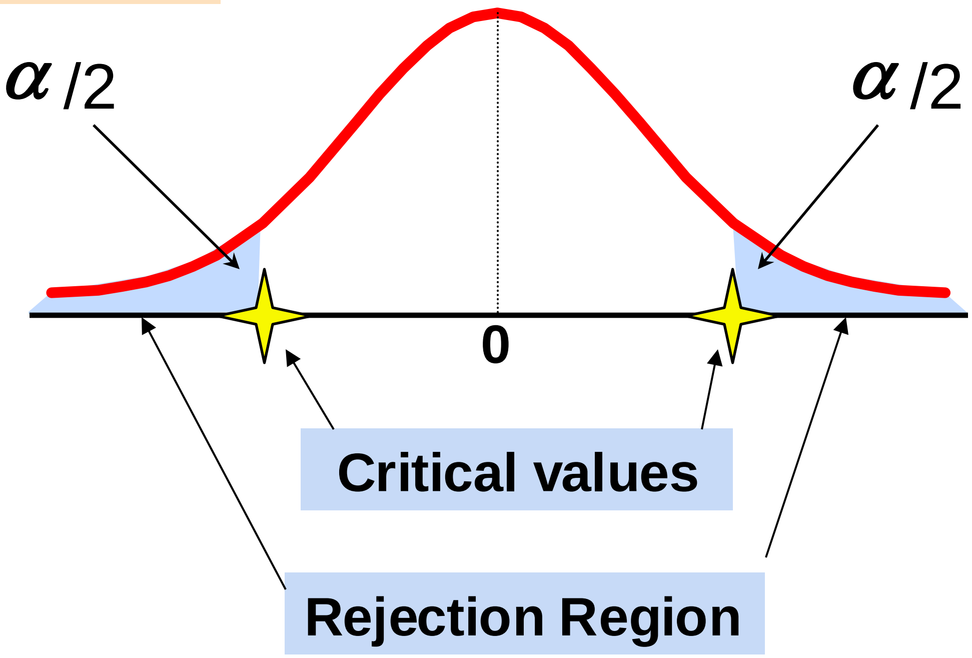
Sampling Distribution of the test statistic



Level of Significance and the Rejection Region

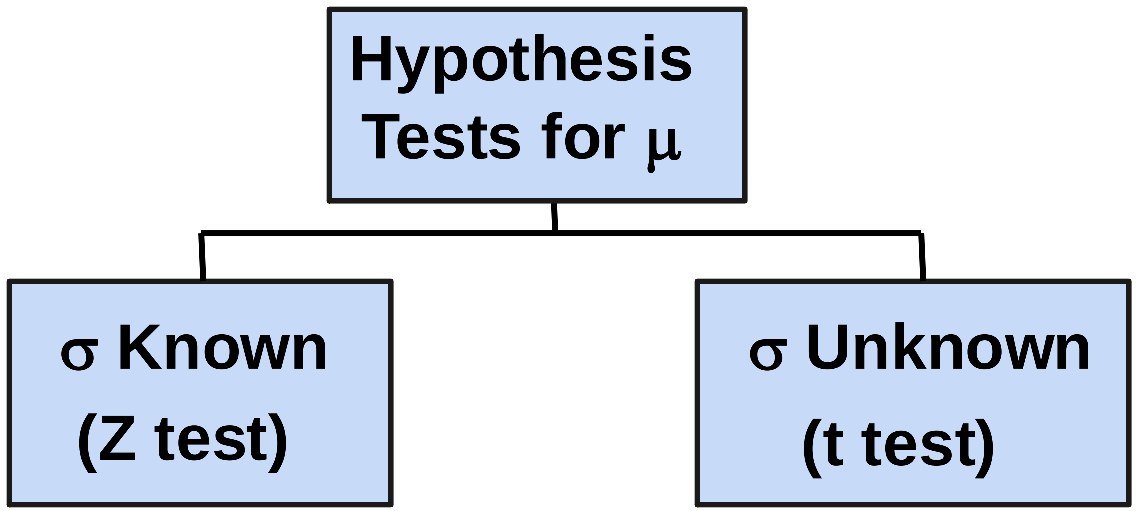
H_0: $\mu = 3$
H_1: $\mu \neq 3$

Level of significance = α



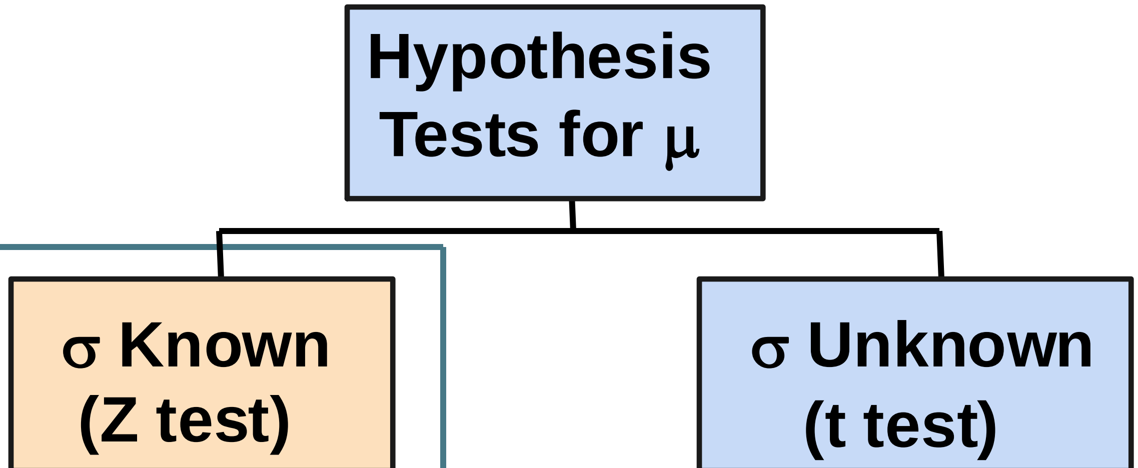
This is a two-tail test because there is a rejection region in both tails

Hypothesis Tests for the Mean



Z Test of Hypothesis for the Mean (σ Known)

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



The test statistic is:

$$Z = \frac{\bar{x} - \mu}{\sigma / \sqrt{n}}$$

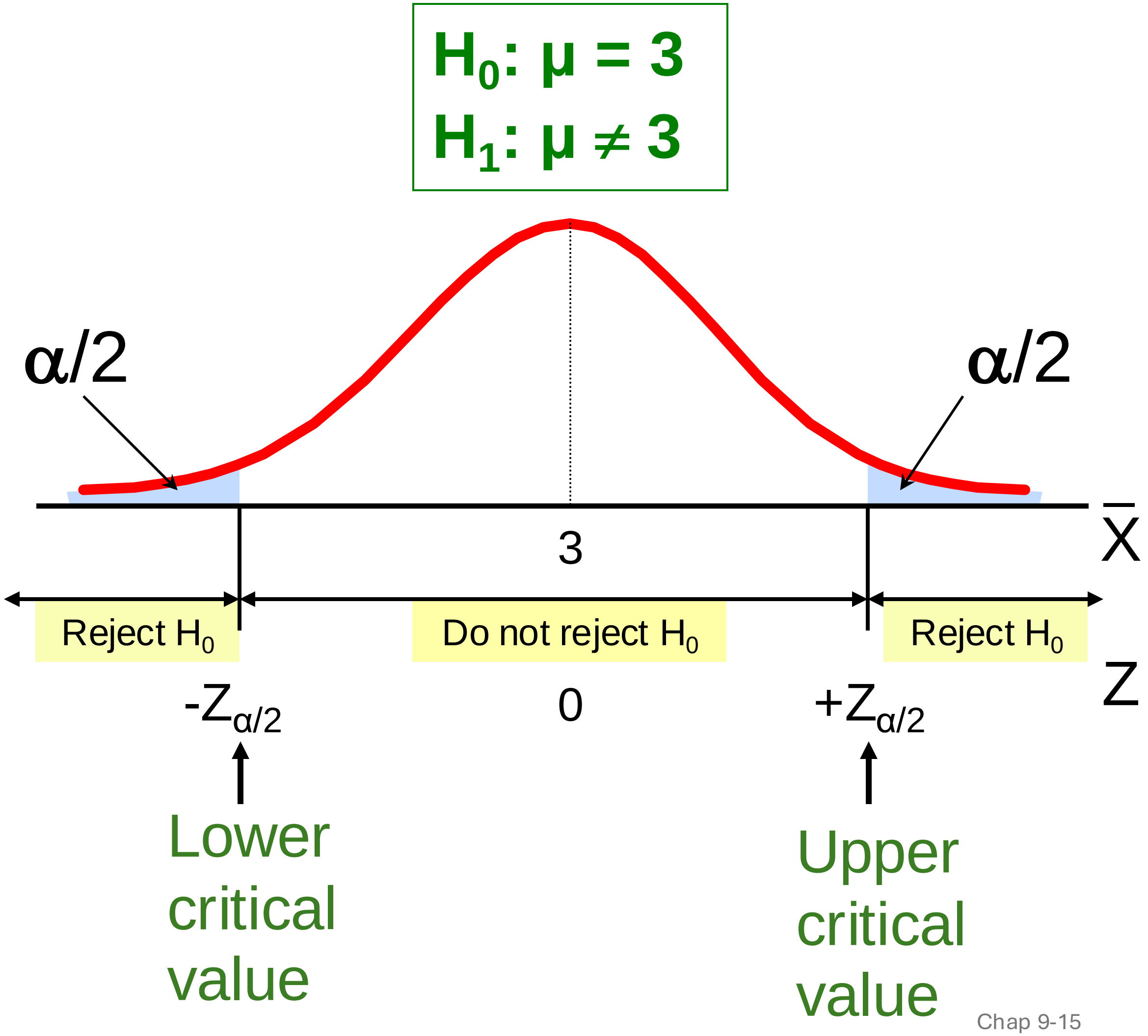
Critical Value Approach to Testing

For a two-tail test for the mean, σ known:

1. Convert sample statistic to test statistic (Z_{STAT})
2. Determine the critical Z values for a specified level of significance α from a table or computer
3. 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 nonrejection 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 nonrejection region, do not reject the null hypothesis H_0. If the test statistic falls into the rejection region, reject the null hypothesis. Express the managerial conclusion in the context of the problem

Hypothesis Testing Example

Test the claim that the true mean # of TV sets in US homes is equal to 3.

(Assume $\sigma = 0.8$)

1. State the appropriate null and alternative hypotheses
   - **H_0: $\mu = 3$ H_1: $\mu \neq 3$ (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**

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, $\bar{X} = 2.84$ ($\sigma = 0.8$ is assumed known)

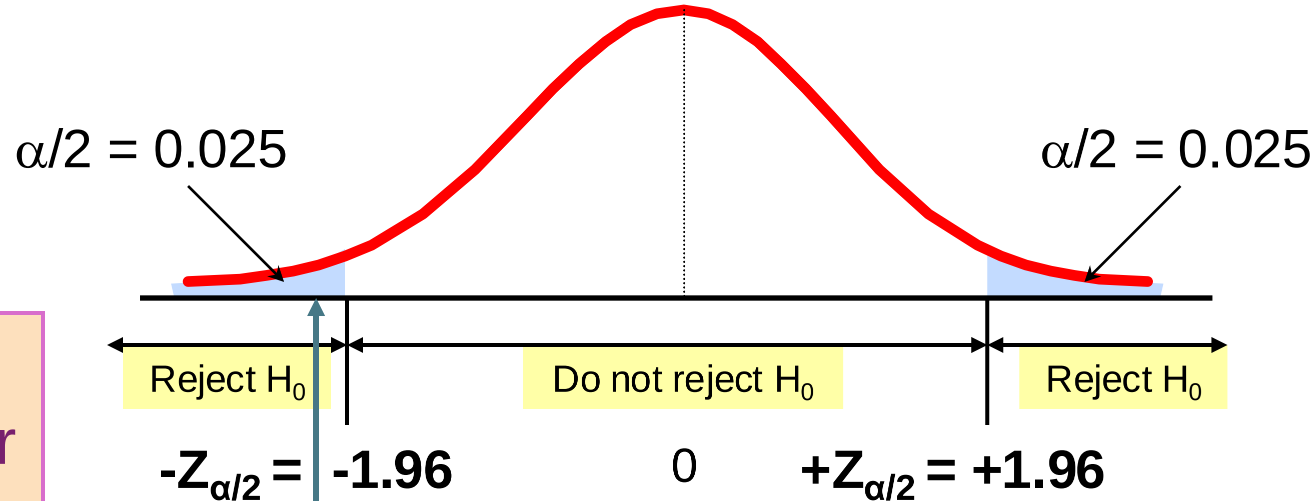
So the test statistic is:

$$Z = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}} \qquad Z = \frac{2.84 - 3}{0.8/\sqrt{100}} = -2$$

Hypothesis Testing Example

(continued)

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

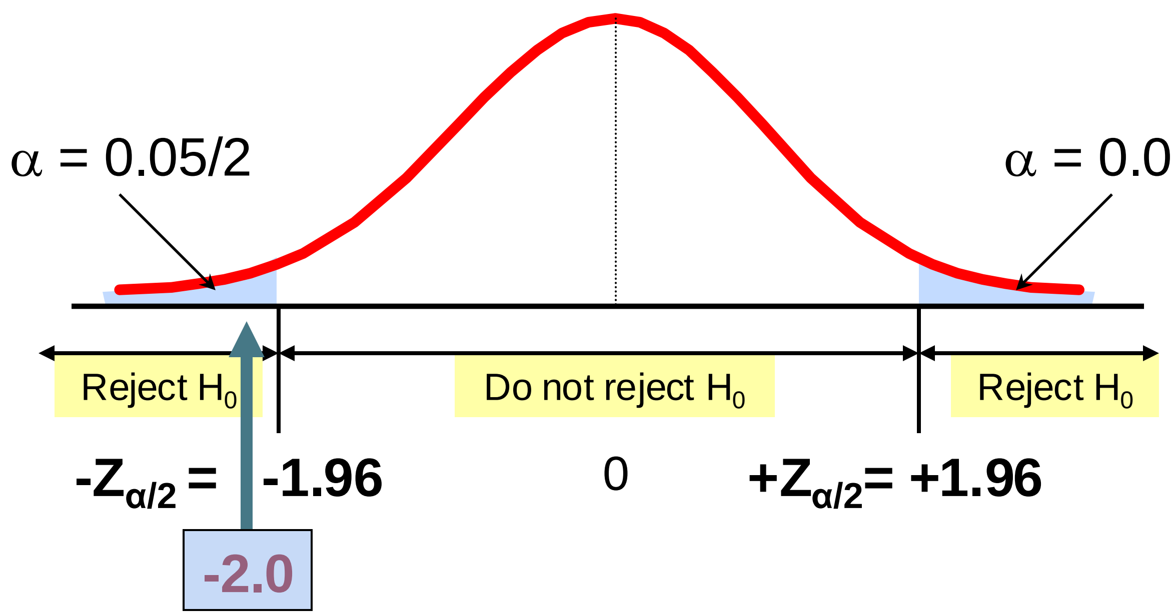


Reject H_0 if $Z_{STAT} < -1.96$ or $Z_{STAT} > 1.96$; otherwise do not reject H_0

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

Hypothesis Testing Example

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



Since Z_{STAT} = -2.0 < -1.96, reject the null hypothesis and conclude there is sufficient evidence that the mean number of TVs in US homes is not equal to 3

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 (Z-calculated)

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 # of TV sets in US homes is equal to 3.
(Assume σ = 0.8)

1. State the appropriate null and alternative hypotheses
   - **H_0: μ = 3 H_1: $\mu \neq$ 3 (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

3. Determine the appropriate technique *(continued)*
   - 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, $\bar{X}$ = 2.84 (σ = 0.8 is assumed known)

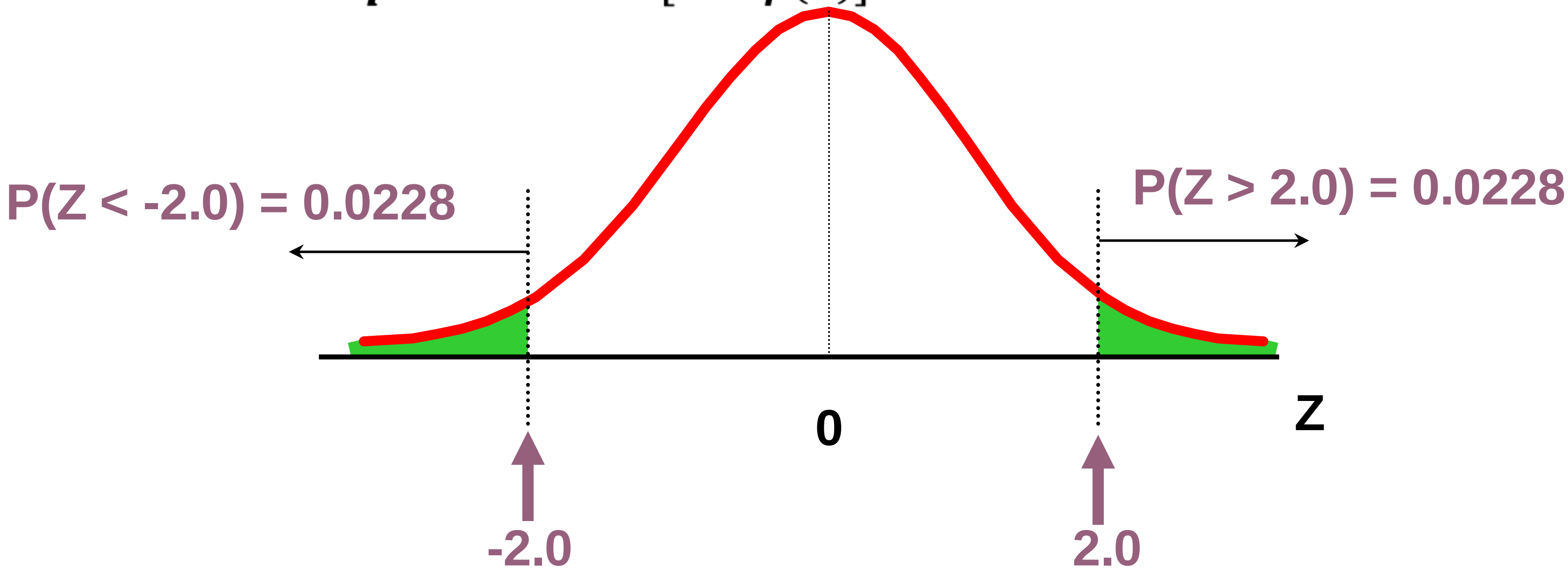
So, the test statistic is:

$$Z = \frac{2.84 - 3}{0.8/\sqrt{100}} = -2$$

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?**
  - $p - value = 2[1 - \phi(Z)]$



p-value = 2*(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 number of TVs in US homes is not equal to 3.

Connection Between Two Tail Tests and Confidence Intervals

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

$$\bar{x} - z_{\frac{\alpha}{2}} * \frac{\sigma}{\sqrt{n}} \le \mu \le \bar{x} + z_{\frac{\alpha}{2}} * \frac{\sigma}{\sqrt{n}}$$

$$2.6832 \le \mu \le 2.9968$$

- Since this interval does not contain the hypothesized mean (3.0), **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.

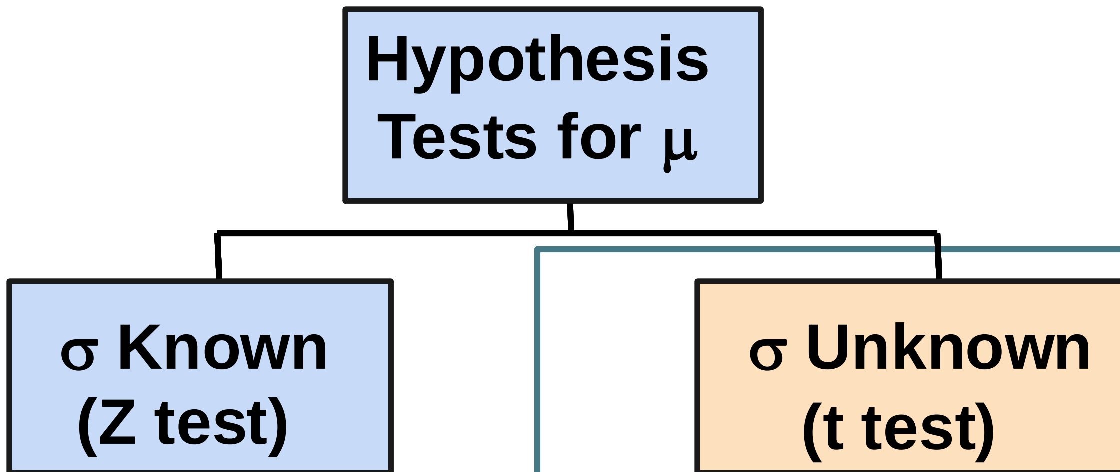
- Because of this change, you use the t distribution instead of the Z distribution to test the null hypothesis about the mean.

- 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.

t Test of Hypothesis for the Mean (σ Unknown)

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



The test statistic is:

$$T = \frac{\bar{x} - \mu}{S/\sqrt{n}}$$

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: $\mu \neq 168$

$$s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2$$

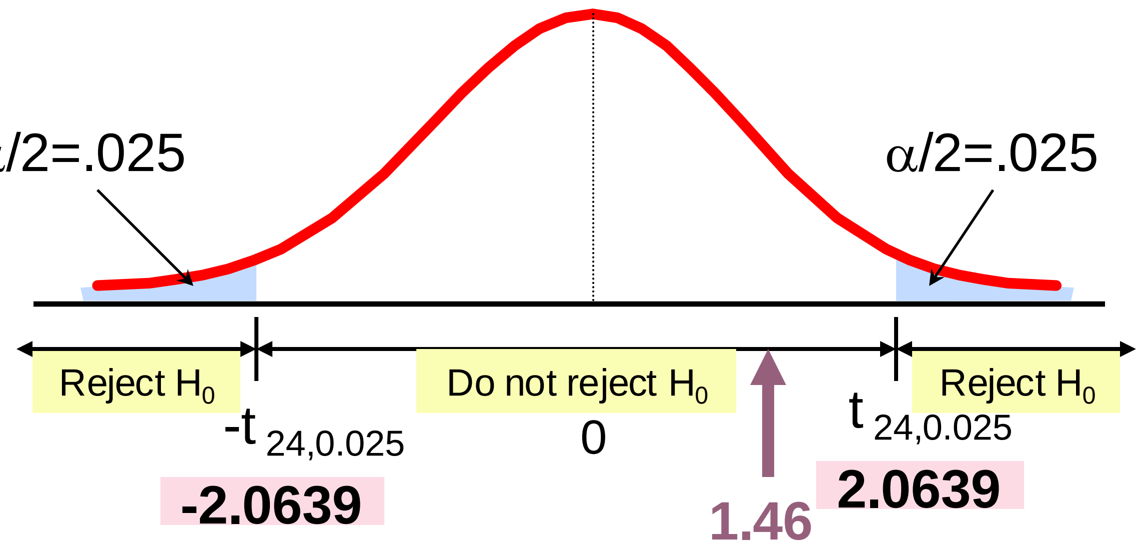
$$s = \sqrt{s^2} = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2}$$

Example Solution: Two-Tail t Test

H_0: μ = 168
H_1: μ ≠ 168

- α = 0.05
- n = 25, df = 25-1=24
- σ is unknown, so use a t statistic
- Critical Value:

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



$$\rightarrow T = \frac{\bar{x} - \mu}{S/\sqrt{n}} \quad T = \frac{172.5 - 168}{15.40/\sqrt{25}} = 1.46$$

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

Using the Test Statistic *t*	
Enter *t* -->	1.46
***df* -->**	24
***p* -value (Lower Tail)**	0.9214
***p* -value (Upper Tail)**	0.0786
***p* -value (Two Tail)**	0.1573

p-value > α
So do not reject H_0

Connection of Two Tail Tests to Confidence Intervals

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

$$\bar{x} - t_{\frac{\alpha}{2};n-1} * \frac{S}{\sqrt{n}} \leq \mu \leq \bar{x} + t_{\frac{\alpha}{2};n-1} * \frac{S}{\sqrt{n}}$$

166.14 ≤ μ ≤ 178.86

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

One-Tail Tests

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

H_0: $\mu = 3$

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

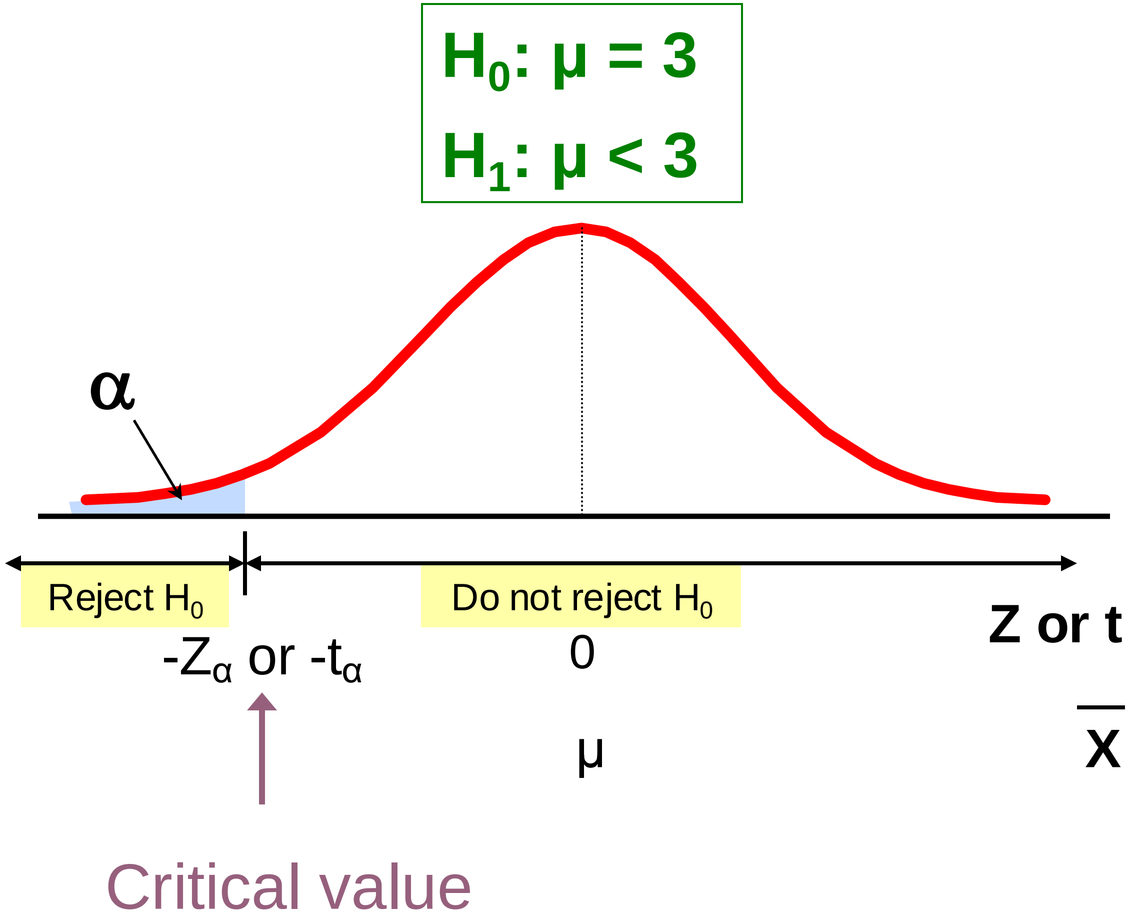
H_0: $\mu = 3$

H_1: $\mu > 3$

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

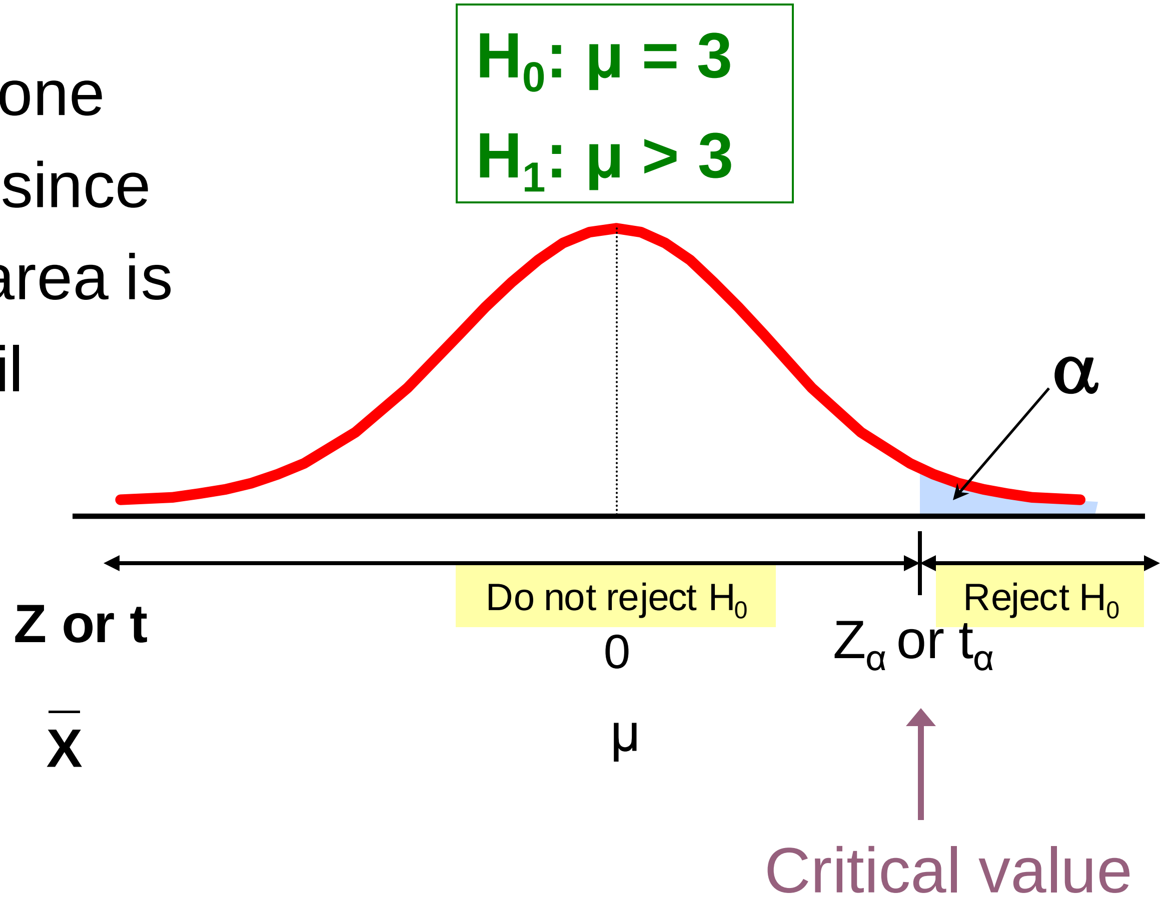
Lower-Tail Tests

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



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: μ = 52 the average is $52 per month

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

Example: Test Statistic

(continued)

Obtain sample and compute the test statistic

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

α = 0.10

- **Then the test statistic is:**

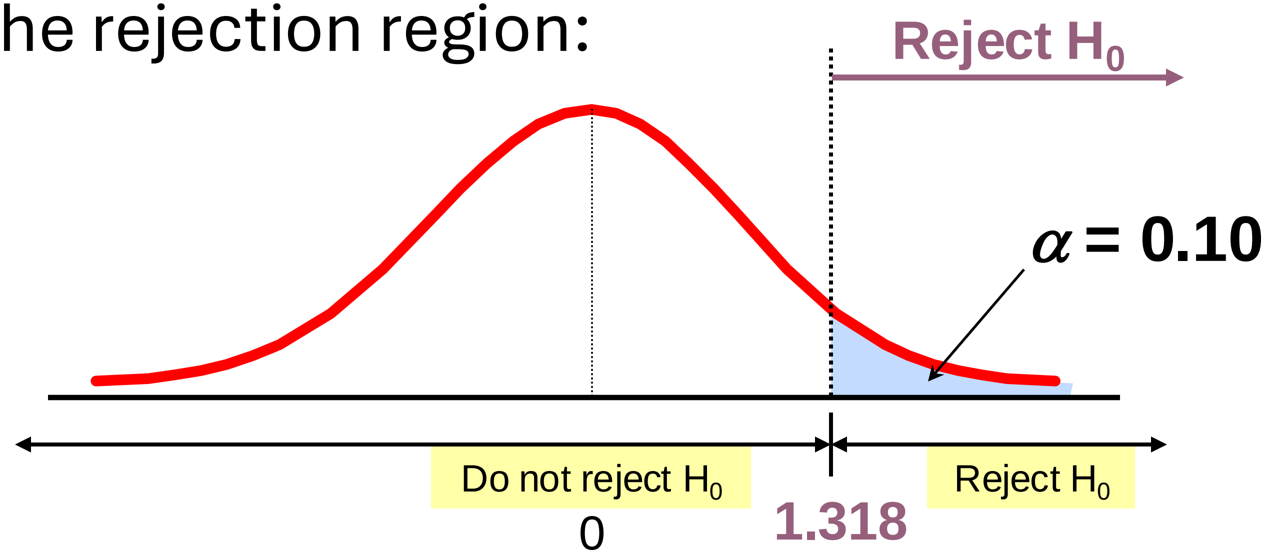
$$T = \frac{\bar{x} - \mu}{S/\sqrt{n}} \quad T = \frac{53.1 - 52}{10/\sqrt{25}} = 0.55$$

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: Final Decision

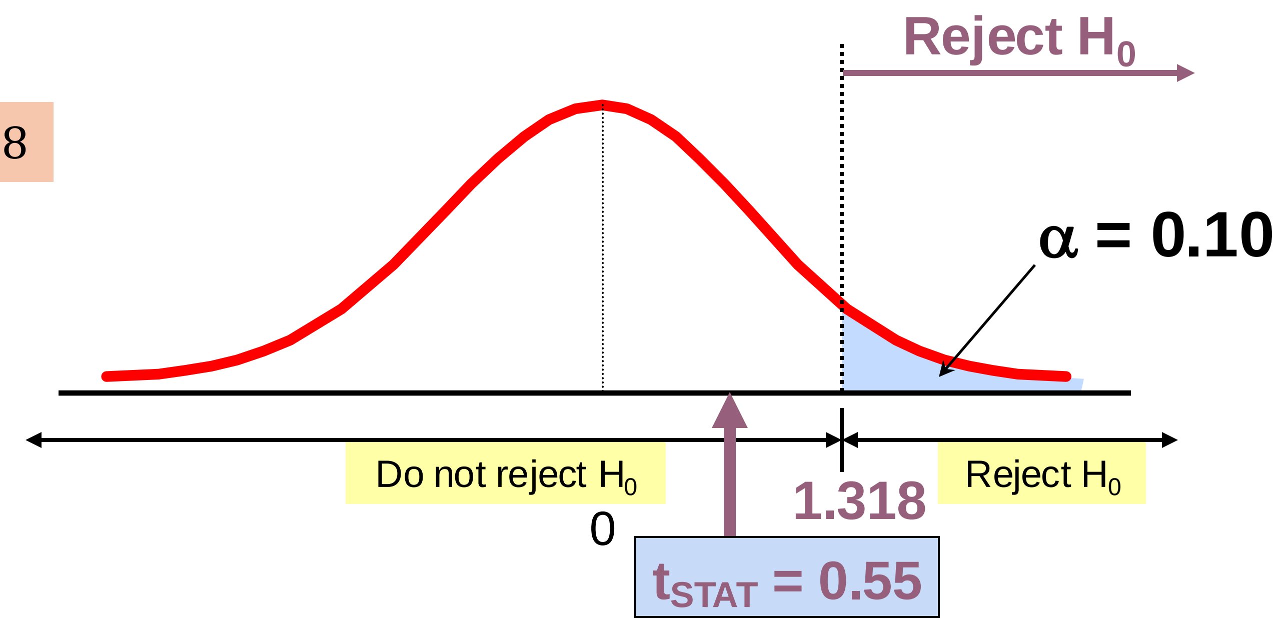
(continued)

Reach a decision and interpret the result:

$\alpha = 0.10$

$t_{0,10;24} = 1.318$

Tstat=0.55

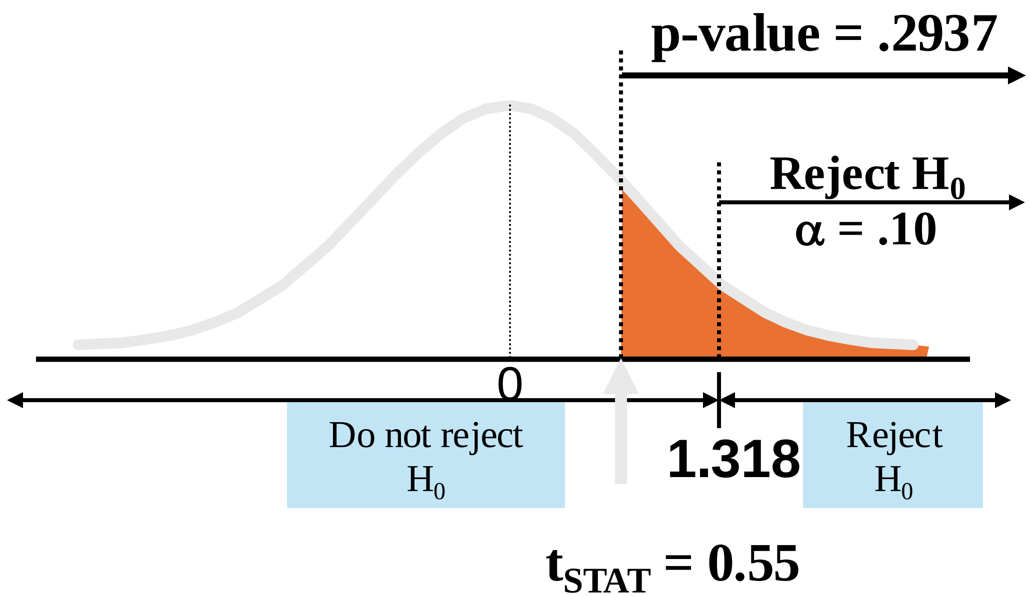


Do not reject H_0 since $t_{STAT} = 0.55 \leq 1.318$

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 = 0.2937 > α = .10

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

Using the Test Statistic *t*	
Enter *t* -->	0.55
***df* -->**	24
***p* -value (Lower Tail)**	0.7063
***p* -value (Upper Tail)**	0.2937
***p* -value (Two Tail)**	0.5874

=SE(E6<0;DISTRIB.T(-E6;E7;1);1-DISTRIB.T(E6;E7;1))

=1-E9

=2*(MIN(E9;E10))

Hypothesis Tests for Proportions

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

Proportions

(continued)

- Sample proportion in the category of interest is denoted by p

- $p = \frac{successful\ cases}{total\ cases}$

- When both nπ and n(1-π) are at least 5, p can be approximated by a normal distribution with mean and standard deviation

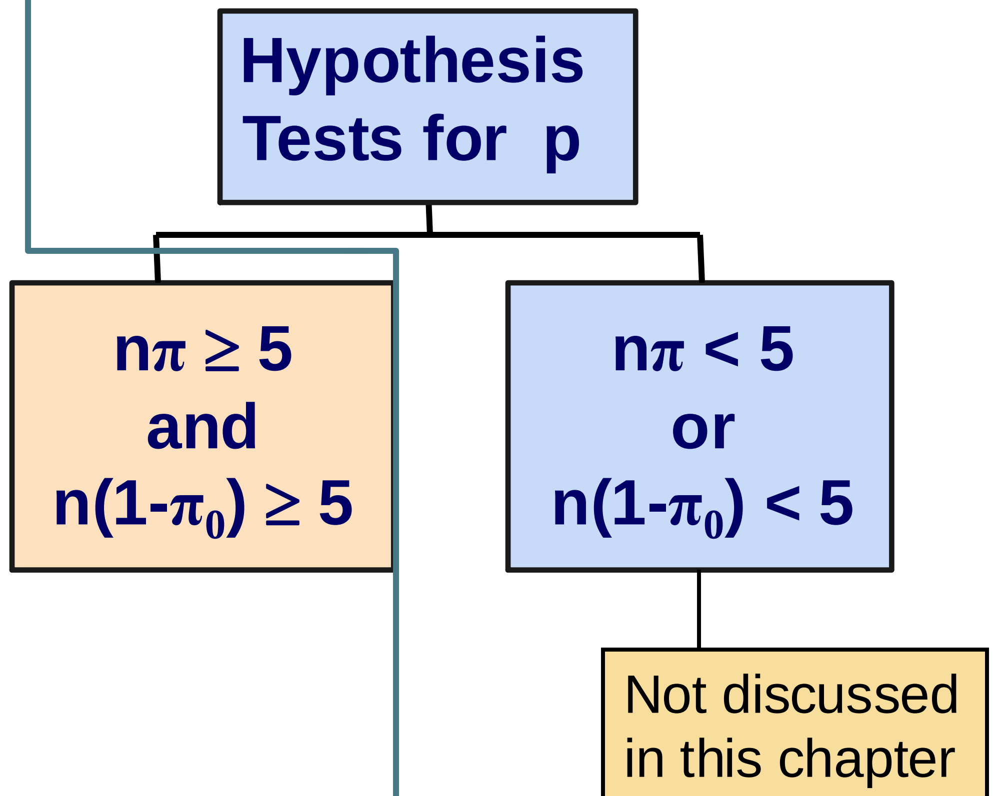
$$\bar{x} = p$$

$$S = \sqrt{\frac{p(1-p)}{n}}$$

Hypothesis Tests for Proportions

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

$$Z = \frac{\overline{X} - \pi_0}{\sqrt{\dfrac{\pi_0(1-\pi_0)}{n}}}$$



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.

Check:

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

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

✓

Example: Z Test for Proportion (solution)

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

$\alpha = 0.05$

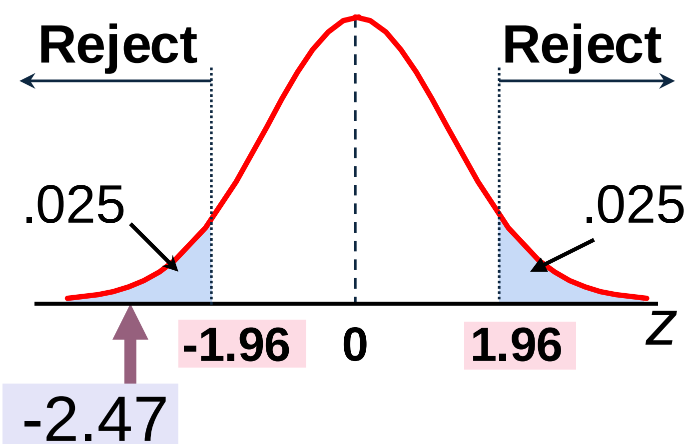
n = 500, p = 0.05

Test Statistic:

$$Z = \frac{\bar{X} - \pi_0}{\sqrt{\frac{\pi_0(1 - \pi_0)}{n}}}$$

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

Critical Values: ± 1.96



Decision:

Reject H_0 at $\alpha = 0.05$

Conclusion:

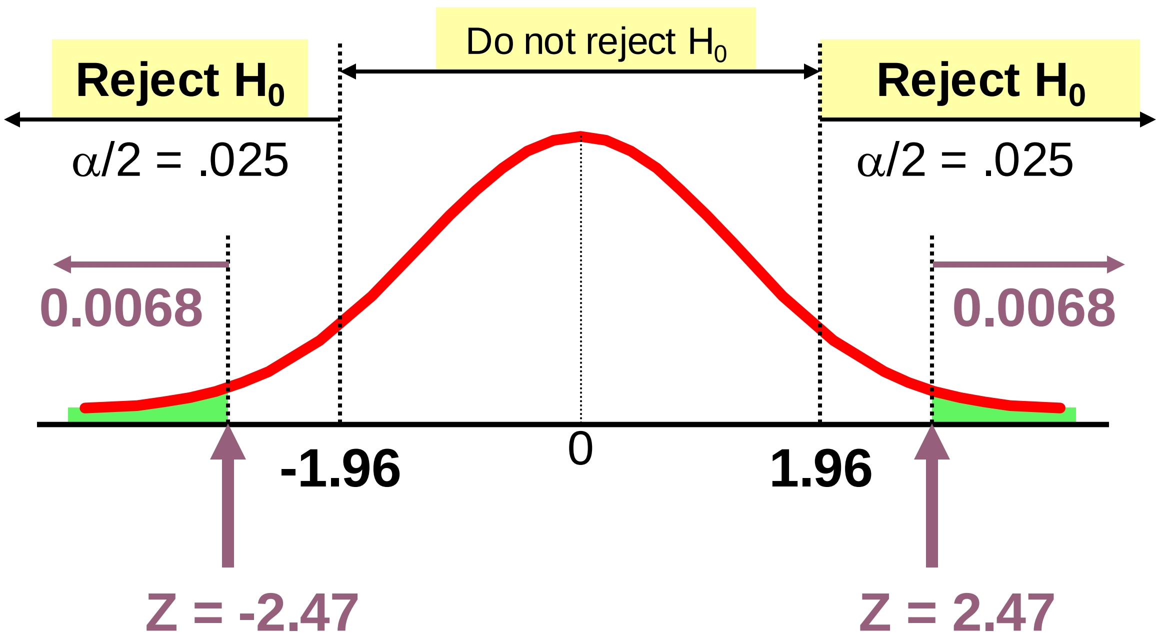
There is sufficient evidence to reject the company's claim of 8% response rate.

p-Value Solution

(continued)

Calculate the p-value and compare to α

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



$$p - value =$$
$$2 * [1 - \phi(z)] =$$
$$2 * [1 - \phi(\ 2.47)]$$

p-value = 0.0136

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

Possible Errors in Hypothesis Test Decision Making

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

- **Type II Error**
  - Failure to reject false null hypothesis
  - The probability of a Type II Error is β

Possible Errors in Hypothesis Test Decision Making

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

Possible Errors in Hypothesis Test Decision Making

- 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

- 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

- All else equal,
  - β ↑ when the difference between hypothesized parameter and its true value ↓
  - β ↑ when α ↓
  - β ↑ when σ ↑
  - β ↑ when n ↓

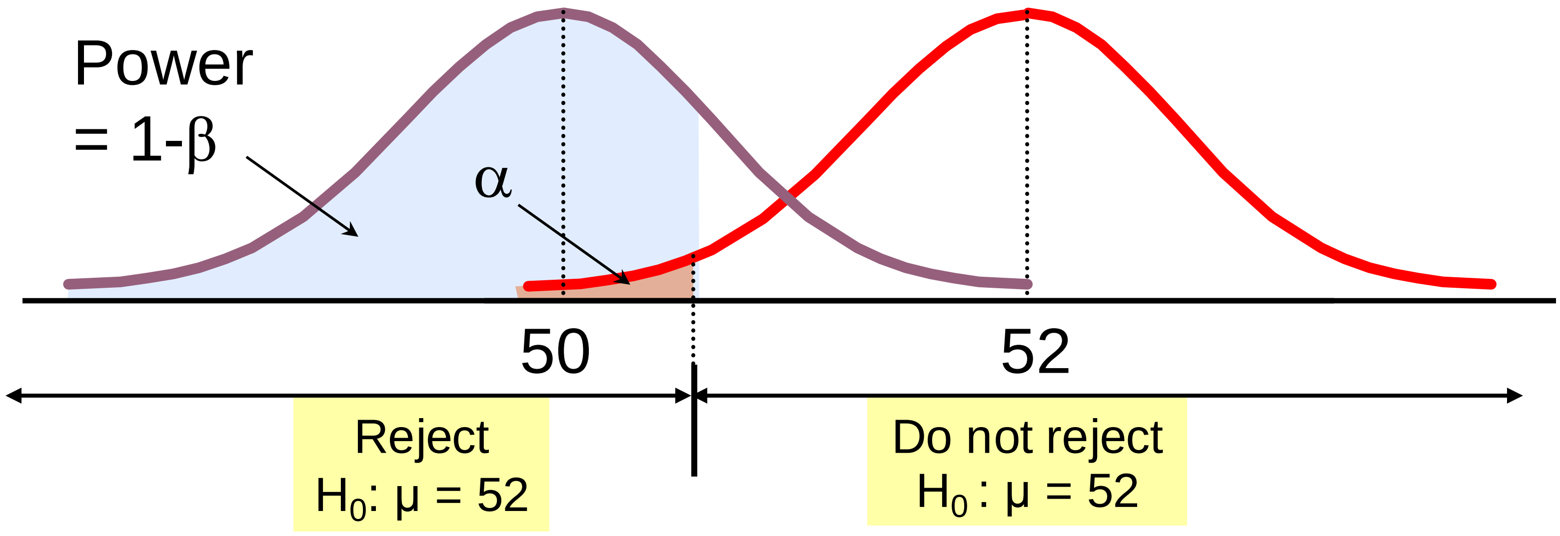
The Power of a Test

The power of the test is the probability of correctly rejecting a false H_0

$$\begin{cases} H_0: \mu = 52 \\ H_1: \mu > 52 \end{cases}$$

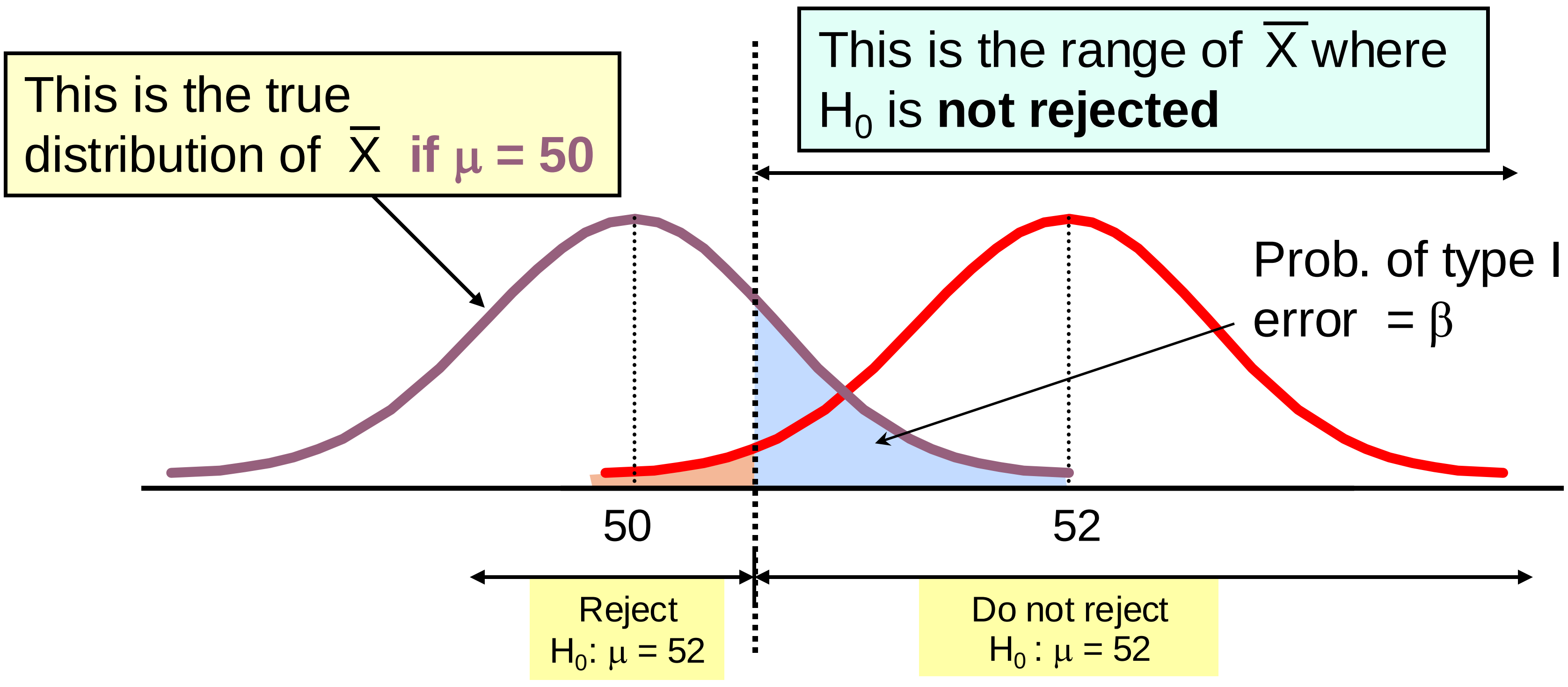
$\alpha = 0.10$

Suppose we correctly reject $H_0: \mu = 52$
when in fact the true mean is $\mu = 50$



Type II Error

Suppose we do not reject H_0: $\mu = 52$ when in fact the true mean is $\mu = 50$

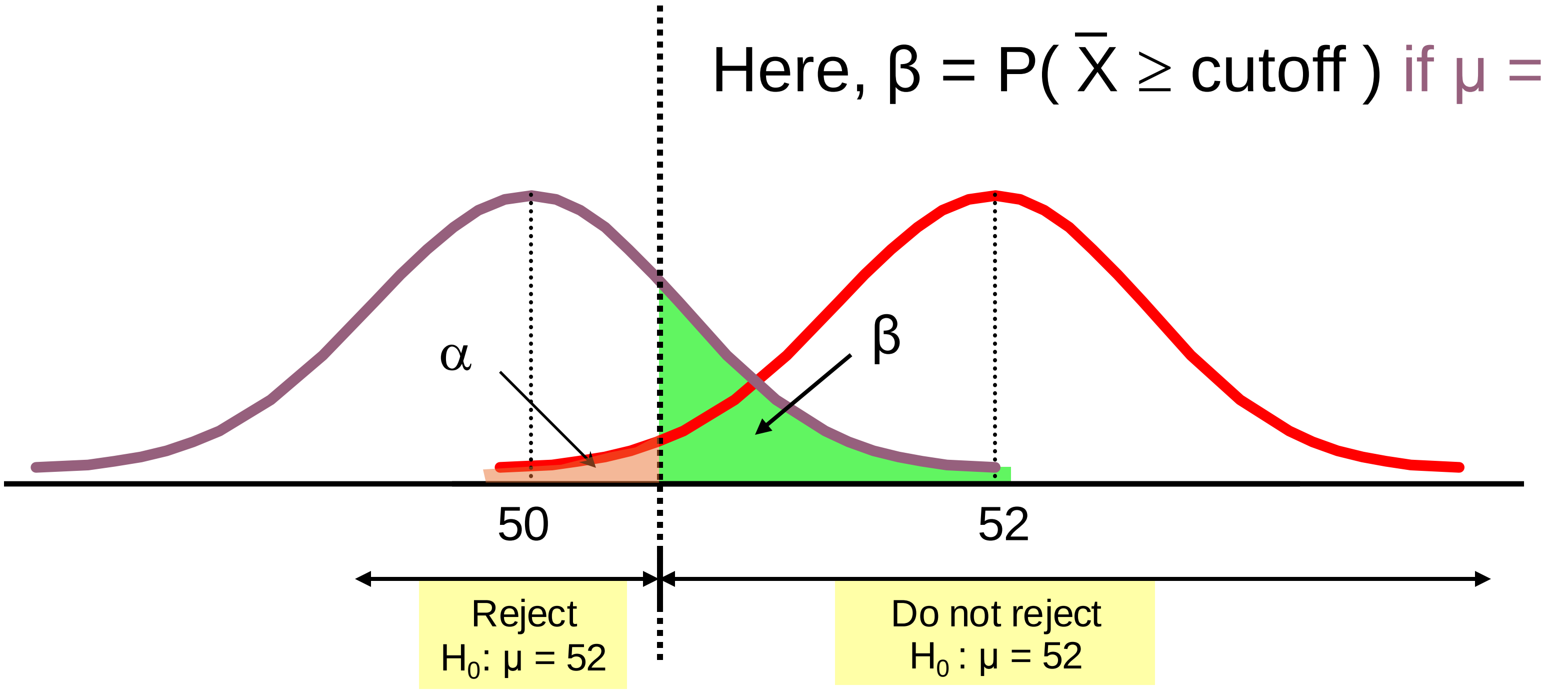


Type II Error

(continued)

- Suppose we do not reject H_0: $\mu = 52$ when in fact the true mean is $\mu = 50$

Here, $\beta = P(\bar{X} \geq \text{cutoff})$ if $\mu = 50$



Power of the Test

- Conclusions regarding the power of the test:

1. A one-tail test is more powerful than a two-tail test
2. An increase in the level of significance (α) results in an increase in power
3. An increase in the sample size results in an increase in power

Potential Pitfalls and Ethical Considerations

- Use *randomly collected data* to reduce selection biases
- Do not use human subjects without informed consent
- Choose the level of significance, α, and the type of test (one-tail or two-tail) before data collection
- Do not employ "data snooping" to choose between one-tail and two-tail test, or to determine the level of significance
- Do not practice "data cleansing" to hide observations that do not support a stated hypothesis
- Report all pertinent findings including both statistical significance and practical importance

Exercises on Hypothesis test

Problem 1 (true variance known)

Suppose the manufacturer claims that the mean lifetime of a light bulb is equal to 10,000 hours. In a sample of 30 light bulbs, it was found that they only last 9,900 hours on average. Assume the population standard deviation is 120 hours. At 0.05 significance level, can we reject the claim by the manufacturer against the alternative that the mean is lower?

Problem 2 (true variance unknown)

Suppose the food label on a cookie bag states that there are equal 2 grams of saturated fat in a single cookie. In a sample of 35 cookies, it is found that the mean amount of saturated fat per cookie is 2.1 grams. Assume that the sample standard deviation is 0.3 gram. At .05 significance level, can we reject the claim on food label against the alternative in which the mean of saturated fat is higher?

Problem 3 (proportion)

Suppose a coin toss turns up 12 heads out of 20 trials. At .05 significance level, can one reject the null hypothesis that the coin toss is fair?