Basic Business Statistics
11^{th} Edition

Chapter 9

Fundamentals of Hypothesis Testing: One-Sample Tests

Learning Objectives

In this chapter, you learn:

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

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 π = 0.68

The Null Hypothesis, H_0

- States the claim or assertion to be tested

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

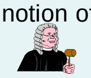
- Is always about a population parameter, not about a sample statistic

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

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



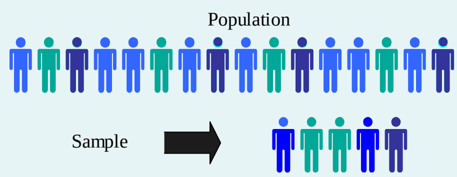
- Refers to the status quo or historical value
- Always contains "=" , "≤" or "≥" sign
- May or may not be rejected

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.



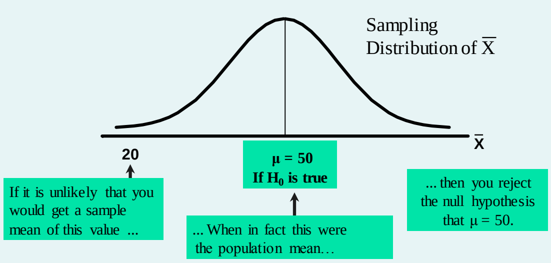
The Hypothesis Testing Process

(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

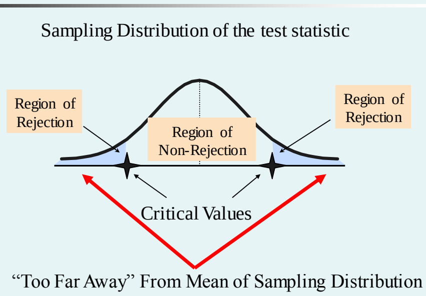
(continued)



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



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 *(continued)*

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 *(continued)*

- 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) $\uparrow$, then Type II error probability (β) $\downarrow$

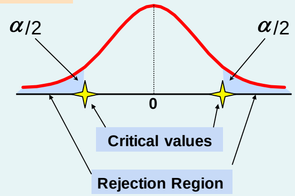
Factors Affecting Type II Error

- All else equal,
  - β $\uparrow$ when the difference between hypothesized parameter and its true value $\downarrow$
  - β $\uparrow$ when α $\downarrow$
  - β $\uparrow$ when σ $\uparrow$
  - β $\uparrow$ when *n* $\downarrow$

Level of Significance and the Rejection Region

H_0: μ = 3
H_1: μ ≠ 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

Hypothesis Tests for μ

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

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:

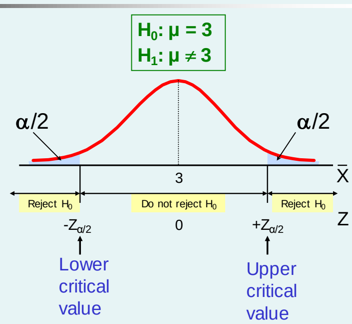
$$Z_{STAT} = \frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}}$$

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 nonrejection regions

6 Steps in Hypothesis Testing

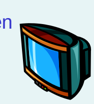
(continued)

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 σ = 0.8)

1. State the appropriate null and alternative hypotheses
   - H_0: μ = 3 H_1: μ ≠ 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



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

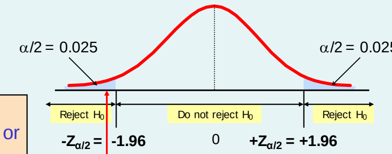
So the test statistic is:

$$Z_{STAT} = \frac{\bar{X}-\mu}{\frac{\sigma}{\sqrt{n}}} = \frac{2.84-3}{\frac{0.8}{\sqrt{100}}} = \frac{-.16}{.08} = -2.0$$

Hypothesis Testing Example

(continued)

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



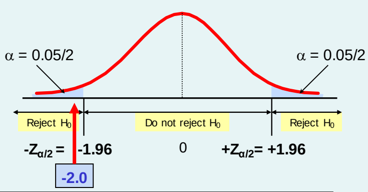
Reject H_0 if Z_{STAT} < -1.96 or Z_{STAT} > 1.96; otherwise do not reject H_0

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 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 < α , reject H_0
  - If p-value ≥ α , 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 < α 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: μ ≠ 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

(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

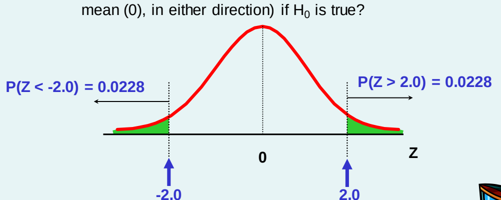
     n = 100, $\bar{X}$ = 2.84 (σ = 0.8 is assumed known)

   So the test statistic is:

$$Z_{STAT} = \frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}} = \frac{2.84 - 3}{\frac{0.8}{\sqrt{100}}} = \frac{-.16}{.08} = -2.0$$

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 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:

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

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

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

Do You Ever Truly Know σ?

- 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: 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

Hypothesis Tests for μ

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

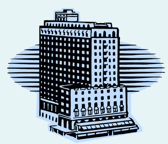
The test statistic is:

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

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)



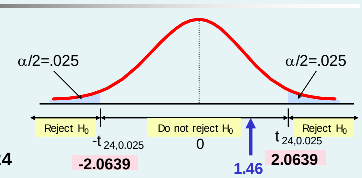
H_0: μ = 168
H_1: μ ≠ 168

Example Solution: Two-Tail t Test

H_0: μ = 168
H_1: μ ≠ 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 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:

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

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

One-Sample T

Test of mu = 168 vs not = 168

N	Mean	StDev	SE Mean	95% CI	T	P
25	172.50	15.40	3.08	(166.14, 178.86)	1.46	0.157

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:

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

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

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

One-Tail Tests

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

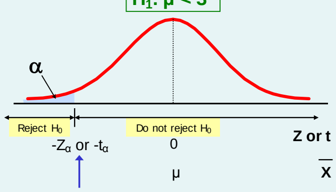
H_0: μ ≥ 3
H_1: μ < 3
→ This is a lower-tail test since the alternative hypothesis is focused on the lower tail below the mean of 3

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

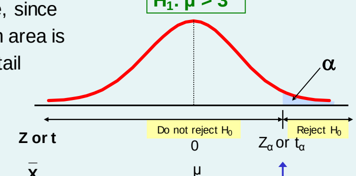
H_0: μ ≥ 3
H_1: μ < 3



Upper-Tail Tests

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

H_0: μ ≤ 3
H_1: μ > 3



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:

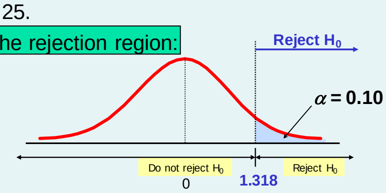
H_0: μ ≤ 52	the average is not over $52 per month
H_1: μ > 52	the average **is** greater than $52 per month (i.e., sufficient evidence exists to support the manager's claim)

Example: Find Rejection Region

(continued)

- Suppose that α = 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 $S = 10$

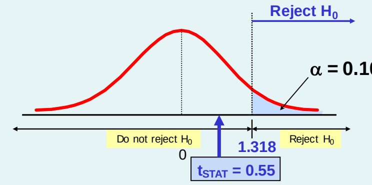
- Then the test statistic is:

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

Example: Decision

(continued)

Reach a decision and interpret the result:

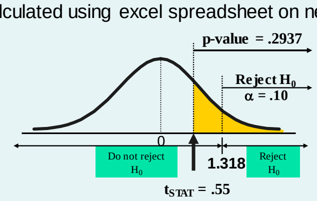


Do not reject H_0 since $t_{STAT} = 0.55 \le 1.318$

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 The Upper Tail t Test

t Test for the Hypothesis of the Mean

Data		
Null Hypothesis μ=	**52.00**	
Level of Significance	**0.1**	
Sample Size	**25**	
Sample Mean	**53.10**	
Sample Standard Deviation	**10.00**	
Intermediate Calculations		
Standard Error of the Mean	2.00	=B8/SQRT(B6)
Degrees of Freedom	24	=B6-1
t test statistic	**0.55**	=(B7-B4)/B11
Upper Tail Test		
Upper Critical Value	1.318	=TINV(2*B5,B12)
p-value	0.2937	=TDIST(ABS(B13),B12,1)
Do Not Reject Null Hypothesis		=IF(B18<B5, "Reject null hypothesis", "Do not reject null hypothesis")

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
  - $p = \frac{X}{n} = \frac{\text{number in category of interest in sample}}{\text{sample size}}$
- 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

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

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

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 chapter

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

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

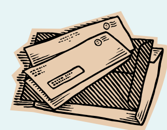
$$Z_{STAT} = \frac{X-n\pi}{\sqrt{n\pi(1-\pi)}}$$

Hypothesis Tests for X

- **$X \geq 5$ and $n-X \geq 5$**
- **$X < 5$ or $n-X < 5$** — Not discussed in this chapter

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:

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

$n(1-\pi) = (500)(.92) = 460$ ✓

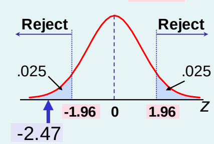
Z Test for Proportion: Solution

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:

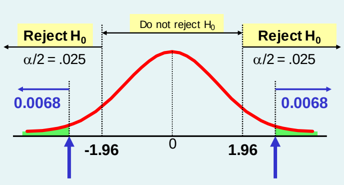
There is sufficient evidence to reject the company's claim of 8% response rate.

p-Value Solution

(continued)

Calculate the p-value and compare to α

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



p-value = 0.0136:

$P(Z \leq -2.47) + P(Z \geq 2.47)$

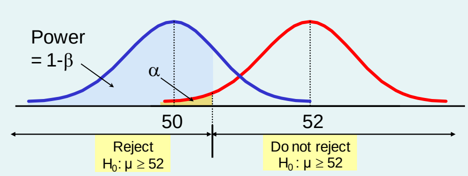
$= 2(0.0068) = 0.0136$

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

The Power of a Test

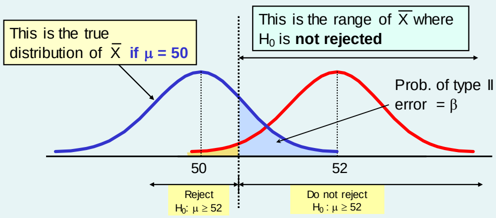
- **The power of the test is the probability of correctly rejecting a false H_0**

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



Type II Error

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

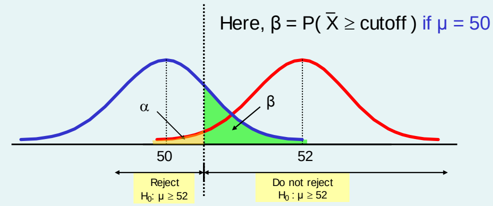


Basic Business Statistics, 11e © 2009 Prentice-Hall, Inc. Chap 9-61

Type II Error

(continued)

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



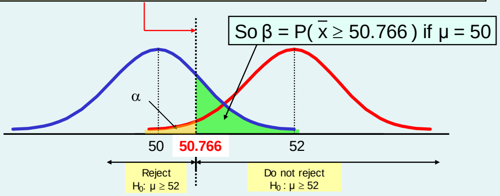
Basic Business Statistics, 11e © 2009 Prentice-Hall, Inc. Chap 9-62

Calculating β

- Suppose n = 64 , σ = 6 , and α = .05

$$\text{cutoff} = \bar{X}_\alpha = \mu - Z_\alpha \frac{\sigma}{\sqrt{n}} = 52 - 1.645\frac{6}{\sqrt{64}} = \boxed{50.766}$$

(for $H_0: \mu \geq 52$)



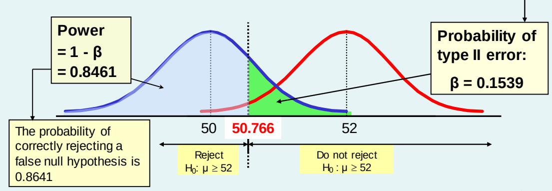
Basic Business Statistics, 11e © 2009 Prentice-Hall, Inc. Chap 9-63

Calculating β and Power of the test

(continued)

- Suppose n = 64 , σ = 6 , and α = 0.05

$$P(\bar{X} \geq 50.766 \mid \mu = 50) = P\left(Z \geq \frac{50.766 - 50}{6/\sqrt{64}}\right) = P(Z \geq 1.02) = 1.0 - 0.8461 = \boxed{0.1539}$$



Basic Business Statistics, 11e © 2009 Prentice-Hall, Inc. Chap 9-64

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

Basic Business Statistics, 11e © 2009 Prentice-Hall, Inc. Chap 9-65

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

Basic Business Statistics, 11e © 2009 Prentice-Hall, Inc. Chap 9-66

Chapter Summary

- Addressed hypothesis testing methodology
- Performed Z Test for the mean (σ known)
- Discussed critical value and p–value approaches to hypothesis testing
- Performed one-tail and two-tail tests

Chapter Summary

(continued)

- Performed t test for the mean (σ unknown)
- Performed Z test for the proportion
- Calculated the probability of a Type II error and the power of the test
- Discussed pitfalls and ethical issues