Chapter 10

Two-Sample Tests

Copyright © 2015, 2012, 2009 Pearson Education, Inc.

Two-Sample Tests

- In this chapter, we develop the required tools to use hypothesis testing for comparing the difference between:
  - The **means** of **two independent populations**
  - The **means** of **two related populations**
  - The **variances** of **two independent populations**

Copyright © 2015, 2012, 2009 Pearson Education, Inc.

Two-Sample Tests

Two-Sample Tests

Population Means, Independe-nt Samples	Population Means, Related Samples	Population Variances
Examples:		
Group 1 vs. Group 2	The same group before vs. after the treatment	Variance 1 vs. Variance 2

Copyright © 2015, 2012, 2009 Pearson Education, Inc.

Difference in the means with independent samples

Copyright © 2015, 2012, 2009 Pearson Education, Inc.

Difference Between Two Means

Population means, independent samples

*

Goal: Test hypothesis or form a confidence interval for the difference between two population means, $\mu_1 - \mu_2$

Copyright © 2015, 2012, 2009 Pearson Education, Inc.

Difference Between Two Means

Population means, independent samples *

Goal: Test hypothesis or form a confidence interval for the difference between two population means, $\mu_1 - \mu_2$

- σ_1 and σ_2 unknown, assumed equal
- σ_1 and σ_2 unknown, not assumed equal

Copyright © 2015, 2012, 2009 Pearson Education, Inc.

Difference Between Two Means

Population means, independent samples *

Goal: Test hypothesis or form a confidence interval for the difference between two population means, $\mu_1 - \mu_2$

- σ_1 and σ_2 unknown, assumed equal
- σ_1 and σ_2 unknown, not assumed equal

The point estimate for the difference is

$$\bar{X}_1 - \bar{X}_2$$

Copyright © 2015, 2012, 2009 Pearson Education, Inc.

Difference Between Two Means

Population means, independent samples *

Independent Condition

Sample selected from one population has no effect on the sample selected from the other population

- σ_1 and σ_2 unknown, assumed equal
- σ_1 and σ_2 unknown, not assumed equal

Copyright © 2015, 2012, 2009 Pearson Education, Inc.

Difference Between Two Means

Population means, independent samples *

Independent Condition

Sample selected from one population has no effect on the sample selected from the other population

σ_1 and σ_2 unknown, assumed equal

Use $\mathbf{S_p}$ (the pooled variance) to estimate unknown σ. Use a **Pooled-Variance t test.**

σ_1 and σ_2 unknown, not assumed equal

Use $\mathbf{S_1}$ and $\mathbf{S_2}$ to estimate unknown σ_1 and σ_2. Use a **Separate-variance t test**

Copyright © 2015, 2012, 2009 Pearson Education, Inc.

Hypothesis Tests for Two Population Means

Two Population Means, Independent Samples

Lower-tail test:

$H_0: \mu_1 \geq \mu_2$
$H_1: \mu_1 < \mu_2$

i.e.,

$H_0: \mu_1 - \mu_2 \geq 0$
$H_1: \mu_1 - \mu_2 < 0$

Upper-tail test:

$H_0: \mu_1 \leq \mu_2$
$H_1: \mu_1 > \mu_2$

i.e.,

$H_0: \mu_1 - \mu_2 \leq 0$
$H_1: \mu_1 - \mu_2 > 0$

Two-tail test:

$H_0: \mu_1 = \mu_2$
$H_1: \mu_1 \neq \mu_2$

i.e.,

$H_0: \mu_1 - \mu_2 = 0$
$H_1: \mu_1 - \mu_2 \neq 0$

Copyright © 2015, 2012, 2009 Pearson Education, Inc.

Hypothesis tests for $\mu_1 - \mu_2$

Two Population Means, Independent Samples

Lower-tail test:

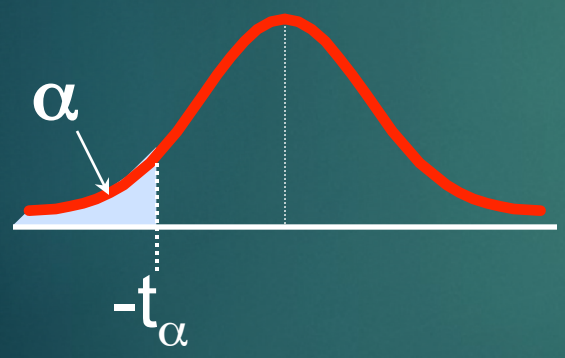
H_0: $\mu_1 - \mu_2 \geq 0$
H_1: $\mu_1 - \mu_2 < 0$

Upper-tail test:

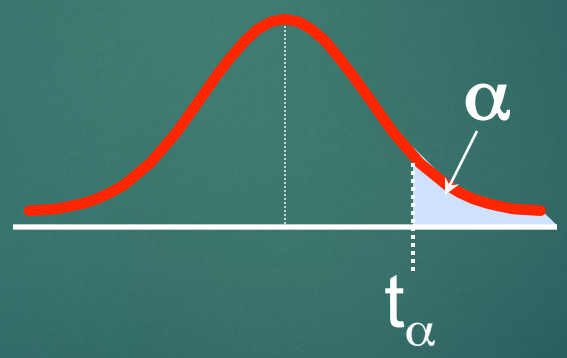
H_0: $\mu_1 - \mu_2 \leq 0$
H_1: $\mu_1 - \mu_2 > 0$

Two-tail test:

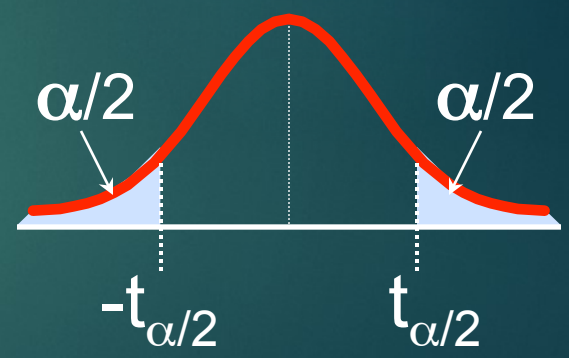
H_0: $\mu_1 - \mu_2 = 0$
H_1: $\mu_1 - \mu_2 \neq 0$



Reject H_0 if $t_{STAT} < -t_\alpha$



Reject H_0 if $t_{STAT} > t_\alpha$



Reject H_0 if $t_{STAT} < -t_{\alpha/2}$ or $t_{STAT} > t_{\alpha/2}$

Copyright © 2015, 2012, 2009 Pearson Education, Inc.

Hypothesis tests for $\mu_1 - \mu_2$ with σ_1 and σ_2 unknown and assumed equal

Population means, independent samples

- σ_1 and σ_2 unknown, assumed equal *
- σ_1 and σ_2 unknown, not assumed equal

Assumptions:

- Samples are randomly and independently drawn
- Populations are normally distributed or both sample sizes are at least 30
- Population variances are unknown but assumed equal

Copyright © 2015, 2012, 2009 Pearson Education, Inc.

Hypothesis tests for $\mu_1 - \mu_2$ with σ_1 and σ_2 unknown and assumed equal

Population means, independent samples

- σ_1 and σ_2 unknown, assumed equal *
- σ_1 and σ_2 unknown, not assumed equal

- The pooled variance is:

$$S_p^2 = \frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{(n_1 - 1) + (n_2 - 1)}$$

Copyright © 2015, 2012, 2009 Pearson Education, Inc.

Hypothesis tests for $\mu_1 - \mu_2$ with σ_1 and σ_2 unknown and assumed equal

Population means, independent samples

- σ_1 and σ_2 unknown, assumed equal *
- σ_1 and σ_2 unknown, not assumed equal

- The pooled variance is:

$$S_p^2 = \frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{(n_1 - 1) + (n_2 - 1)}$$

- The test statistic is:

$$t_{STAT} = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{S_p^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

- **Where t_{STAT} has d.f. = $(n_1 + n_2 - 2)$**

Copyright © 2015, 2012, 2009 Pearson Education, Inc.

Confidence interval for $\mu_1 - \mu_2$ with σ_1 and σ_2 unknown and assumed equal

Population means, independent samples

- σ_1 and σ_2 unknown, assumed equal *
- σ_1 and σ_2 unknown, not assumed equal

Point Estimate ± (CriticalValue) (Standard Error)

Copyright © 2015, 2012, 2009 Pearson Education, Inc.

Confidence interval for $\mu_1 - \mu_2$ with σ_1 and σ_2 unknown and assumed equal

Population means, independent samples

- σ_1 and σ_2 unknown, assumed equal *
- σ_1 and σ_2 unknown, not assumed equal

Point Estimate ± (CriticalValue) (Standard Error)

The confidence interval for $\mu_1 - \mu_2$ is:

$$\left(\bar{X}_1 - \bar{X}_2\right) \pm t_{\alpha/2}\sqrt{S_p^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}$$

Where $t_{\alpha/2}$ has d.f. = $n_1 + n_2 - 2$

Copyright © 2015, 2012, 2009 Pearson Education, Inc.

Pooled-Variance t Test Example

Given the data below, is there a difference in dividend yield (i.e. dividend, adjusted by price) between stocks listed on the NYSE & NASDAQ? You collect the following data:

	NYSE	NASDAQ
Number	**21**	**25**
Sample mean	**3.27**	**2.53**
Sample std dev	**1.30**	**1.16**

Copyright © 2015, 2012, 2009 Pearson Education, Inc.

Pooled-Variance t Test Example

Given the data below, is there a difference in dividend yield (i.e. dividend, adjusted by price) between stocks listed on the NYSE & NASDAQ? You collect the following data:

	NYSE	NASDAQ
Number	**21**	**25**
Sample mean	**3.27**	**2.53**
Sample std dev	**1.30**	**1.16**

Assuming both **populations** are **approximately normal** with **equal variances**, is there a difference in mean yield ($\alpha = 0.05$)?

Copyright © 2015, 2012, 2009 Pearson Education, Inc.

Pooled-Variance t Test Example:
Calculating the Test Statistic

H0: $\mu_1 - \mu_2 = 0$ i.e. ($\mu_1 = \mu_2$)
H1: $\mu_1 - \mu_2 \neq 0$ i.e. ($\mu_1 \neq \mu_2$)

The test statistic is:

$$t = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{S_p^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} = \frac{(3.27 - 2.53) - 0}{\sqrt{1.5021\left(\frac{1}{21} + \frac{1}{25}\right)}} = 2.040$$

$$S_p^2 = \frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{(n_1 - 1) + (n_2 - 1)} = \frac{(21-1)1.30^2 + (25-1)1.16^2}{(21-1) + (25-1)} = 1.5021$$

Copyright © 2015, 2012, 2009 Pearson Education, Inc.

Pooled-Variance t Test Example: Hypothesis Test Solution

H_0: $\mu_1 - \mu_2 = 0$ i.e. $(\mu_1 = \mu_2)$

H_1: $\mu_1 - \mu_2 \neq 0$ i.e. $(\mu_1 \neq \mu_2)$

$\alpha = 0.05$

df = 21 + 25 - 2 = 44

Critical Values: t = ± 2.0154

Copyright © 2015, 2012, 2009 Pearson Education, Inc.

Pooled-Variance t Test Example: Hypothesis Test Solution

H_0: $\mu_1 - \mu_2 = 0$ i.e. $(\mu_1 = \mu_2)$

H_1: $\mu_1 - \mu_2 \neq 0$ i.e. $(\mu_1 \neq \mu_2)$

$\alpha = 0.05$

df = 21 + 25 - 2 = 44

Critical Values: t = ± 2.0154

Test Statistic:

$$t = \frac{3.27 - 2.53}{\sqrt{1.5021\left(\frac{1}{21} + \frac{1}{25}\right)}} = \boxed{2.040}$$

Copyright © 2015, 2012, 2009 Pearson Education, Inc.

Pooled-Variance t Test Example: Hypothesis Test Solution

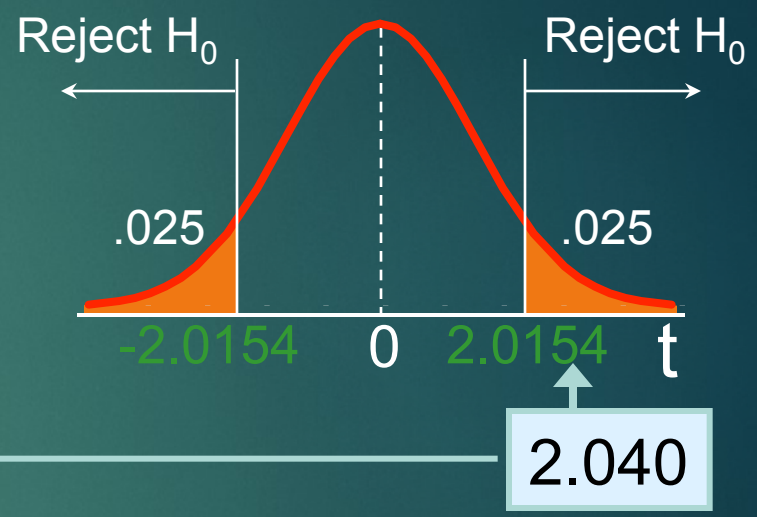
H_0: $\mu_1 - \mu_2 = 0$ i.e. $(\mu_1 = \mu_2)$

H_1: $\mu_1 - \mu_2 \neq 0$ i.e. $(\mu_1 \neq \mu_2)$

$\alpha = 0.05$

df = 21 + 25 - 2 = 44

Critical Values: t = ± 2.0154



Test Statistic:

$$t = \frac{3.27 - 2.53}{\sqrt{1.5021\left(\frac{1}{21} + \frac{1}{25}\right)}} = 2.040$$

Copyright © 2015, 2012, 2009 Pearson Education, Inc.

Pooled-Variance t Test Example: Hypothesis Test Solution

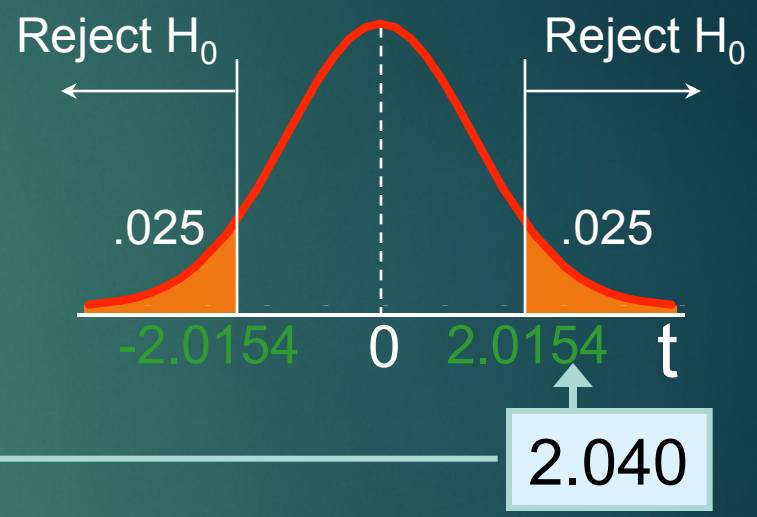
H_0: $\mu_1 - \mu_2 = 0$ i.e. $(\mu_1 = \mu_2)$

H_1: $\mu_1 - \mu_2 \neq 0$ i.e. $(\mu_1 \neq \mu_2)$

$\alpha = 0.05$

df = 21 + 25 - 2 = 44

Critical Values: t = ± 2.0154



Test Statistic:

$$t = \frac{3.27 - 2.53}{\sqrt{1.5021\left(\frac{1}{21} + \frac{1}{25}\right)}} = 2.040$$

Decision:

Reject H_0 at **a** = 0.05

Conclusion:

There is evidence of a difference in means.

Copyright © 2015, 2012, 2009 Pearson Education, Inc.

Pooled-Variance t Test Example: Confidence Interval for $\mu_1 - \mu_2$

Since we rejected H_0, can we be 95% confident that: $\mu_{NYSE} > \mu_{NASDAQ}$?

95% Confidence Interval for $\mu_{NYSE} - \mu_{NASDAQ}$

$$\left(\bar{X}_1 - \bar{X}_2\right) \pm t_{\alpha/2}\sqrt{S_p^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)} = 0.74 \pm 2.0154 \times 0.3628 = (0.009,\ 1.471)$$

Since 0 is less than the entire interval, we can be 95% confident that $\mu_{NYSE} > \mu_{NASDAQ}$

Copyright © 2015, 2012, 2009 Pearson Education, Inc.

Hypothesis tests for $\mu_1 - \mu_2$ with σ_1 and σ_2 unknown, not assumed equal

Population means, independent samples

- σ_1 and σ_2 unknown, assumed equal
- σ_1 and σ_2 unknown, not assumed equal *

Assumptions:

- Samples are randomly and independently drawn
- Populations are normally distributed or both sample sizes are at least 30
- Population variances are unknown and cannot be assumed to be equal

Copyright © 2015, 2012, 2009 Pearson Education, Inc.

Hypothesis tests for $\mu_1 - \mu_2$ with σ_1 and σ_2 unknown, not assumed equal

Population means, independent samples

- σ_1 and σ_2 unknown, assumed equal
- σ_1 and σ_2 unknown, not assumed equal *

The test statistic is:

$$t_{STAT} = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}}$$

Copyright © 2015, 2012, 2009 Pearson Education, Inc.

Hypothesis tests for $\mu_1 - \mu_2$ with σ_1 and σ_2 unknown, not assumed equal

Population means, independent samples

- σ_1 and σ_2 unknown, assumed equal
- σ_1 and σ_2 unknown, not assumed equal *

The test statistic is:

$$t_{STAT} = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}}$$

t_{STAT} has d.f. ν =

$$\nu = \frac{\left(\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}\right)^2}{\frac{\left(\frac{S_1^2}{n_1}\right)^2}{n_1 - 1} + \frac{\left(\frac{S_2^2}{n_2}\right)^2}{n_2 - 1}}$$

Copyright © 2015, 2012, 2009 Pearson Education, Inc.

Pooled-Variance t Test Example

Given the data below, is there a difference in dividend yield (i.e. dividend, adjusted by price) between stocks listed on the NYSE & NASDAQ? You collect the following data:

	NYSE	NASDAQ
Number	**21**	**25**
Sample mean	**3.27**	**2.53**
Sample std dev	**1.30**	**1.16**

Copyright © 2015, 2012, 2009 Pearson Education, Inc.

Pooled-Variance t Test Example

Given the data below, is there a difference in dividend yield (i.e. dividend, adjusted by price) between stocks listed on the NYSE & NASDAQ? You collect the following data:

	NYSE	NASDAQ
Number	**21**	**25**
Sample mean	**3.27**	**2.53**
Sample std dev	**1.30**	**1.16**

Assuming both **populations** are **approximately normal** with **unequal variances**, is there a difference in mean yield ($\alpha = 0.05$)?

Copyright © 2015, 2012, 2009 Pearson Education, Inc.

Separate-Variance t Test Example: Calculating the Test Statistic

H0: $\mu_1 - \mu_2 = 0$ i.e. $(\mu_1 = \mu_2)$
H1: $\mu_1 - \mu_2 \neq 0$ i.e. $(\mu_1 \neq \mu_2)$

The test statistic is:

$$t = \frac{\left(\bar{X}_1 - \bar{X}_2\right) - \left(\mu_1 - \mu_2\right)}{\sqrt{\left(\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}\right)}} = \frac{\left(3.27 - 2.53\right) - 0}{\sqrt{\left(\frac{1.30^2}{21} + \frac{1.16^2}{25}\right)}} = \boxed{2.019}$$

$$\nu = \frac{\left(\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}\right)^2}{\frac{\left(\frac{S_1^2}{n_1}\right)^2}{n_1 - 1} + \frac{\left(\frac{S_2^2}{n_2}\right)^2}{n_2 - 1}} = \frac{\left(\frac{1.30^2}{21} + \frac{1.16^2}{25}\right)^2}{\frac{\left(\frac{1.30^2}{21}\right)^2}{20} + \frac{\left(\frac{1.16^2}{25}\right)^2}{24}} = 40.57$$

→ Use degrees of freedom = 40

Copyright © 2015, 2012, 2009 Pearson Education, Inc.

Separate-Variance t Test Example: Hypothesis Test Solution

H_0: $\mu_1 - \mu_2 = 0$ i.e. $(\mu_1 = \mu_2)$

H_1: $\mu_1 - \mu_2 \neq 0$ i.e. $(\mu_1 \neq \mu_2)$

$\alpha = 0.05$

df = 40

Critical Values: t = ± 2.021

Copyright © 2015, 2012, 2009 Pearson Education, Inc.

Separate-Variance t Test Example: Hypothesis Test Solution

DCOVA

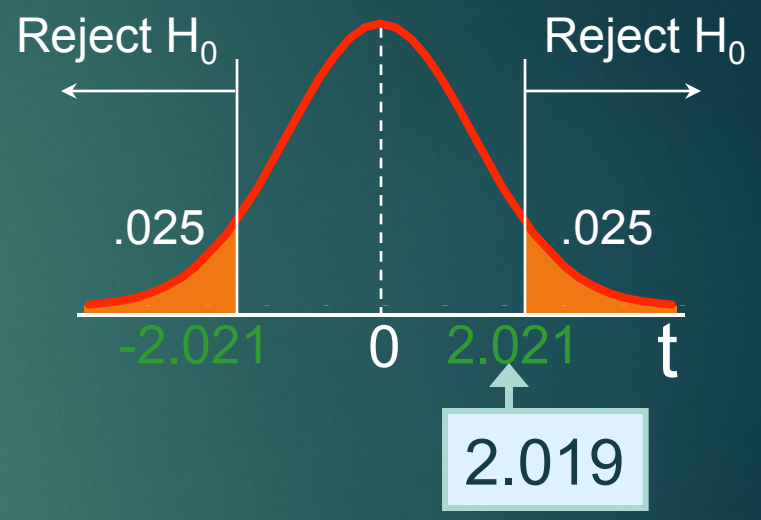
H_0: $\mu_1 - \mu_2 = 0$ i.e. $(\mu_1 = \mu_2)$

H_1: $\mu_1 - \mu_2 \neq 0$ i.e. $(\mu_1 \neq \mu_2)$

$\alpha = 0.05$

df = 40

Critical Values: t = ± 2.021



Copyright © 2015, 2012, 2009 Pearson Education, Inc.

Separate-Variance t Test Example: Hypothesis Test Solution

DCOVA

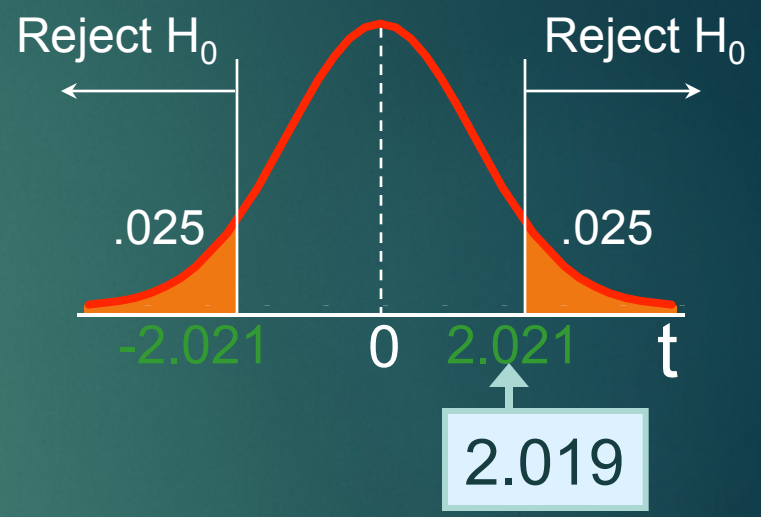
H_0: $\mu_1 - \mu_2 = 0$ i.e. $(\mu_1 = \mu_2)$

H_1: $\mu_1 - \mu_2 \neq 0$ i.e. $(\mu_1 \neq \mu_2)$

$\alpha = 0.05$

df = 40

Critical Values: t = ± 2.021



Decision:

Fail To Reject H_0

Conclusion:

There is insufficient evidence of a difference in means.

Copyright © 2015, 2012, 2009 Pearson Education, Inc.

Difference in the means with related samples

Copyright © 2015, 2012, 2009 Pearson Education, Inc.

The Difference in Time Spent

- A video game developer introduces a new feature to one of their games, and they would like to examine if this leads to any changes in the Time Spent by the gamers.
- To do so, they select a random sample from the gamers and observe follow them after introducing the new feature.

Copyright © 2015, 2012, 2009 Pearson Education, Inc.

Related Populations
The Paired Difference Test

- The Paired Difference Test examines the means of related populations
  - Repeated measures (before/after)
  - Use **the difference between** paired values:

$$D_i = X_{1i} - X_{2i}$$

Example: X_{1i} – Time spent by gamer i **before** the changes

X_{2i} – Time spent by gamer i **after** the changes

- Using the above difference, eliminates the variation among subjects

Copyright © 2015, 2012, 2009 Pearson Education, Inc.

Related Populations
The Paired Difference Test

- The Paired Difference Test examines the means of related populations
  - Repeated measures (before/after)
  - Use **the difference** between paired values:

$$D_i = X_{1i} - X_{2i}$$

- Assumptions:
  - Both populations are normally distributed
  - Or, if not normal, use large samples

Copyright © 2015, 2012, 2009 Pearson Education, Inc.

Related Populations
The Paired Difference Test

(continued)

DCOVA

The ith paired difference is D_i, where

$$D_i = X_{1i} - X_{2i}$$

The point estimate for the paired difference population mean μ_D is $\bar{D}$:

$$\bar{D} = \frac{\sum_{i=1}^{n} D_i}{n}$$

(n is the number of pairs in the paired sample)

Example: This is **the average of the changes in time spent** among the gamers in the sample

Copyright © 2015, 2012, 2009 Pearson Education, Inc.

Related Populations
The Paired Difference Test

(continued)

DCOVA

The ith paired difference is D_i, where

$$D_i = X_{1i} - X_{2i}$$

The sample standard deviation is S_D

(n is the number of pairs in the paired sample)

$$S_D = \sqrt{\frac{\sum_{i=1}^{n}(D_i - \bar{D})^2}{n-1}}$$

Example: This is **the average of the deviations from the average of the changes in time spent** among the gamers in the sample

Copyright © 2015, 2012, 2009 Pearson Education, Inc.

The Paired Difference Test: Finding the Test Stat: t_{STAT}

- The test statistic for μ_D is:

$$t_{STAT} = \frac{\overline{D} - \mu_D}{\frac{S_D}{\sqrt{n}}}$$

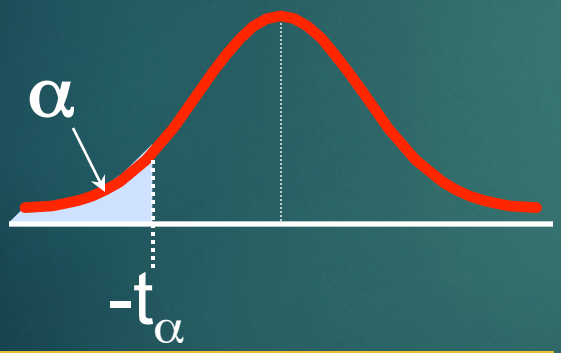
- Where t_{STAT} has $n - 1$ d.f.

Copyright © 2015, 2012, 2009 Pearson Education, Inc.

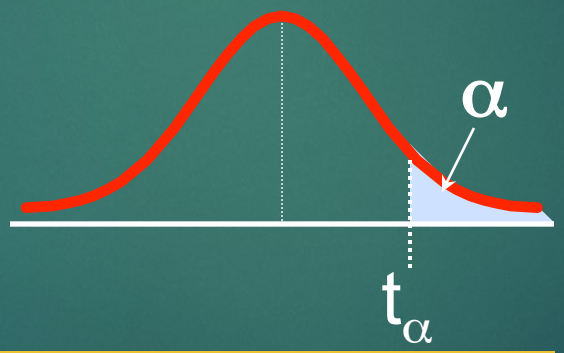
The Paired Difference Test: Possible Hypotheses

Paired Samples

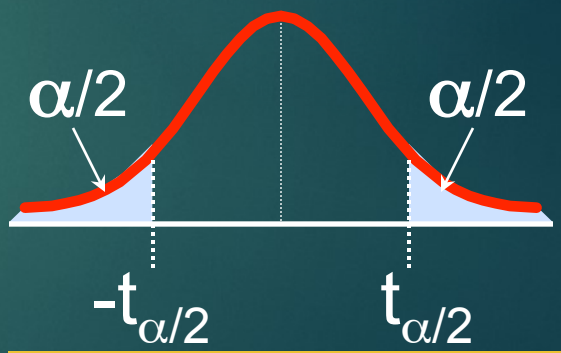
Lower-tail test:	Upper-tail test:	Two-tail test:
H_0: $\mu_D \geq 0$	H_0: $\mu_D \leq 0$	H_0: $\mu_D = 0$
H_1: $\mu_D < 0$	H_1: $\mu_D > 0$	H_1: $\mu_D \neq 0$



Reject H_0 if $t_{STAT} < -t_\alpha$



Reject H_0 if $t_{STAT} > t_\alpha$



Reject H_0 if $t_{STAT} < -t_{\alpha/2}$ or $t_{STAT} > t_{\alpha/2}$

Copyright © 2015, 2012, 2009 Pearson Education, Inc.

The Paired Difference Confidence Interval

The confidence interval for μ_D is

$$\overline{D} \pm t_{\alpha/2} \frac{S_D}{\sqrt{n}}$$

where $$S_D = \sqrt{\frac{\sum_{i=1}^{n} (D_i - \overline{D})^2}{n-1}}$$

Copyright © 2015, 2012, 2009 Pearson Education, Inc.

Paired Difference Test: Example

- Assume that the video game developers ask the gamers to pay for an upgraded component of the game that used to be free. The following table summarizes the time spent by five gamers.

The Gamers	Number of Hours Played: Before (1)	After (2)	(2) - (1) Difference, D_i
C.B.	6	4	- 2
T.F.	20	6	-14
M.H.	3	2	- 1
R.K.	0	0	0
M.O.	4	0	- 4
			-21

$$\bar{D} = \frac{\sum D_i}{n}$$

$$= -4.2$$

$$S_D = \sqrt{\frac{\sum (D_i - \bar{D})^2}{n-1}}$$

$$= 5.67$$

Copyright © 2015, 2012, 2009 Pearson Education, Inc.

Paired Difference Test: Example

- Has the new feature made a difference in the number of hours played (at the 0.01 level)?

H_0: $\mu_D = 0$
H_1: $\mu_D \neq 0$

Copyright © 2015, 2012, 2009 Pearson Education, Inc.

Paired Difference Test: Example

- Has the new feature made a difference in the number of hours played (at the 0.01 level)?

$H_0: \mu_D = 0$
$H_1: \mu_D \neq 0$

$\alpha = .01$ $\bar{D} = -4.2$

$t_{0.005} = \pm 4.604$

d.f. = n - 1 = 4

Copyright © 2015, 2012, 2009 Pearson Education, Inc.

Paired Difference Test: Example

- Has the new feature made a difference in the number of hours played (at the 0.01 level)?

H_0: $\mu_D = 0$
H_1: $\mu_D \neq 0$

$\alpha = .01$ $\quad \bar{D} = -4.2$

$t_{0.005} = \pm 4.604$

d.f. = n - 1 = 4

$$t_{STAT} = \frac{\bar{D} - \mu_D}{S_D/\sqrt{n}} = \frac{-4.2 - 0}{5.67/\sqrt{5}} = -1.66$$

Copyright © 2015, 2012, 2009 Pearson Education, Inc.

Paired Difference Test: Example

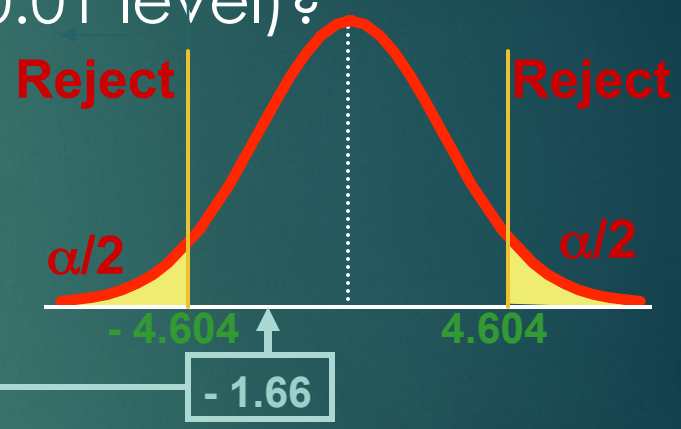
- Has the new feature made a difference in the number of hours played (at the 0.01 level)?

$H_0: \mu_D = 0$
$H_1: \mu_D \neq 0$

$\alpha = .01 \qquad \bar{D} = -4.2$

$t_{0.005} = \pm 4.604$

d.f. = n - 1 = 4



$$t_{STAT} = \frac{\bar{D} - \mu_D}{S_D/\sqrt{n}} = \frac{-4.2 - 0}{5.67/\sqrt{5}} = \boxed{-1.66}$$

Copyright © 2015, 2012, 2009 Pearson Education, Inc.

Paired Difference Test: Example

- Has the new feature made a difference in the number of hours played (at the 0.01 level)?

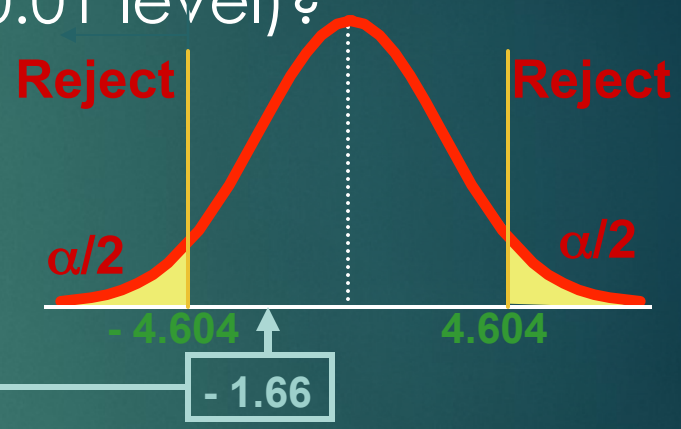
H_0: $\mu_D = 0$
H_1: $\mu_D \neq 0$

$\alpha = .01$ $\quad \bar{D} = -4.2$

$t_{0.005} = \pm 4.604$

d.f. = n - 1 = 4

$$t_{STAT} = \frac{\bar{D} - \mu_D}{S_D/\sqrt{n}} = \frac{-4.2 - 0}{5.67/\sqrt{5}} = \boxed{-1.66}$$



Decision: Do not reject H_0
(t_{stat} is not in the reject region)

Conclusion: There is not a significant change in the number of complaints.

Copyright © 2015, 2012, 2009 Pearson Education, Inc.

The Paired Difference Confidence Interval

$$\overline{D} \pm t_{\alpha/2} \frac{S_D}{\sqrt{n}}$$

$\overline{D}$ = -4.2, S_D = 5.67

$$99\%\ \text{CI for } \mu_D : -4.2 \pm 4.604 \frac{5.67}{\sqrt{5}}$$

$$= (-15.87, 7.47)$$

Copyright © 2015, 2012, 2009 Pearson Education, Inc.

The Ratio of Population Variances

Copyright © 2015, 2012, 2009 Pearson Education, Inc.

Testing for the Ratio Of Two Population Variances

Tests for Two Population Variances

*

Hypotheses

$H_0: \sigma_1^2 = \sigma_2^2$
$H_1: \sigma_1^2 \neq \sigma_2^2$

$H_0: \sigma_1^2 \leq \sigma_2^2$
$H_1: \sigma_1^2 > \sigma_2^2$

What is the Test Stat?

Copyright © 2015, 2012, 2009 Pearson Education, Inc.

Testing for the Ratio Of Two Population Variances

Tests for Two Population Variances *

F test statistic

Hypotheses	F_{STAT}
$H_0: \sigma_1^2 = \sigma_2^2$ $H_1: \sigma_1^2 \neq \sigma_2^2$	S_1^2 / S_2^2
$H_0: \sigma_1^2 \leq \sigma_2^2$ $H_1: \sigma_1^2 > \sigma_2^2$	

S_1^2 = Variance of sample 1 (the **larger** sample variance)

S_2^2 = Variance of sample 2 (the **smaller** sample variance)

Copyright © 2015, 2012, 2009 Pearson Education, Inc.

Testing for the Ratio Of Two Population Variances

Tests for Two Population Variances *

F test statistic

Hypotheses	F_{STAT}
$H_0: \sigma_1^2 = \sigma_2^2$ $H_1: \sigma_1^2 \neq \sigma_2^2$	S_1^2 / S_2^2
$H_0: \sigma_1^2 \leq \sigma_2^2$ $H_1: \sigma_1^2 > \sigma_2^2$	

Where:

S_1^2 = Variance of sample 1 (the **larger** sample variance)

n_1 = sample size of sample 1

S_2^2 = Variance of sample 2 (the **smaller** sample variance)

n_2 = sample size of sample 2

$n_1 - 1$ = **numerator degrees of freedom**

$n_2 - 1$ = **denominator degrees of freedom**

Copyright © 2015, 2012, 2009 Pearson Education, Inc.

The F Distribution

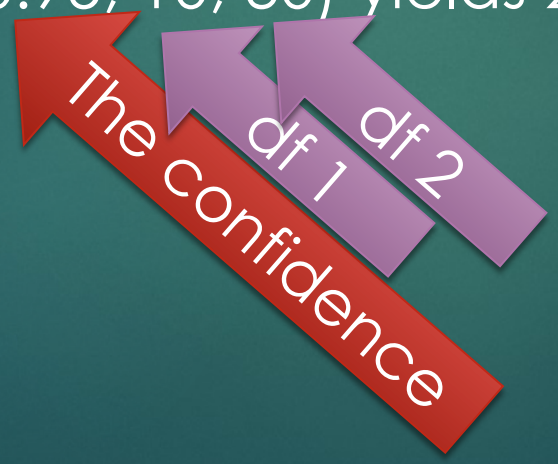
- The F critical value is found from the F table (E5)
- There are two degrees of freedom required: numerator and denominator
- The larger sample variance is always the numerator
- When $F_{STAT} = \frac{S_1^2}{S_2^2}$ $df_1 = n_1 - 1$; $df_2 = n_2 - 1$
- In the F table,
  - numerator degrees of freedom determine the column
  - denominator degrees of freedom determine the row

Copyright © 2015, 2012, 2009 Pearson Education, Inc.

The F Distribution

- The F critical value can also be found using **=F.INV** function in Excel
- The cumulative probability is required
- Also, two degrees of freedom are required: numerator and denominator

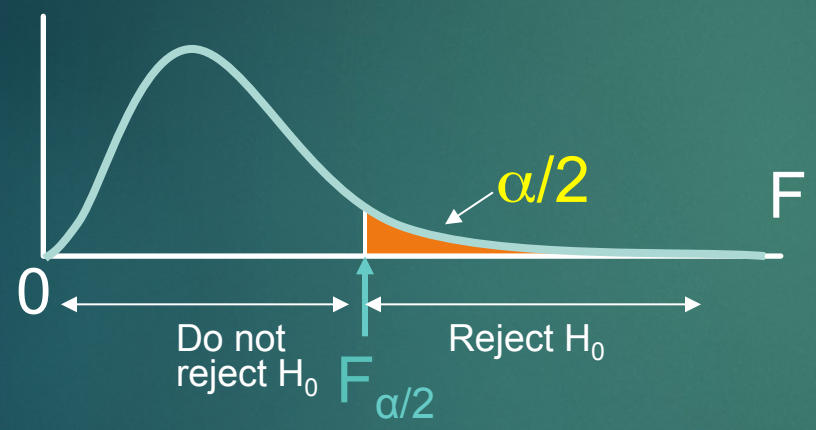
Example: =F.INV(0.95, 10, 30) yields 2.16



Copyright © 2015, 2012, 2009 Pearson Education, Inc.

Finding the Rejection Region

$H_0: \sigma_1^2 = \sigma_2^2$
$H_1: \sigma_1^2 \neq \sigma_2^2$

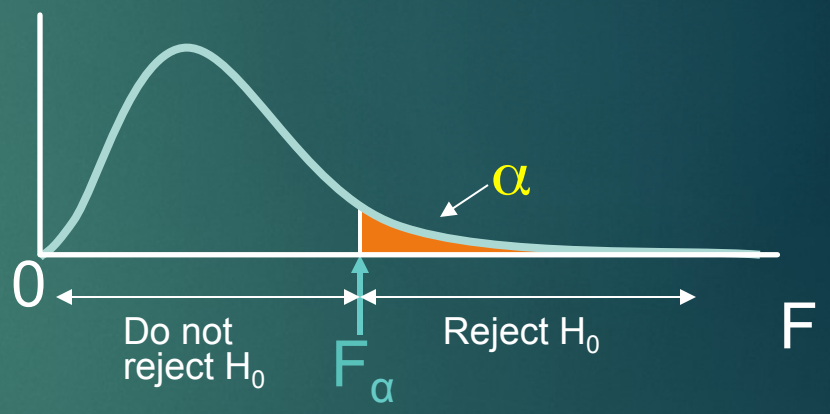


Reject H_0 if $F_{STAT} > F_{\alpha/2}$

Copyright © 2015, 2012, 2009 Pearson Education, Inc.

Finding the Rejection Region

$H_0: \sigma_1^2 \leq \sigma_2^2$
$H_1: \sigma_1^2 > \sigma_2^2$

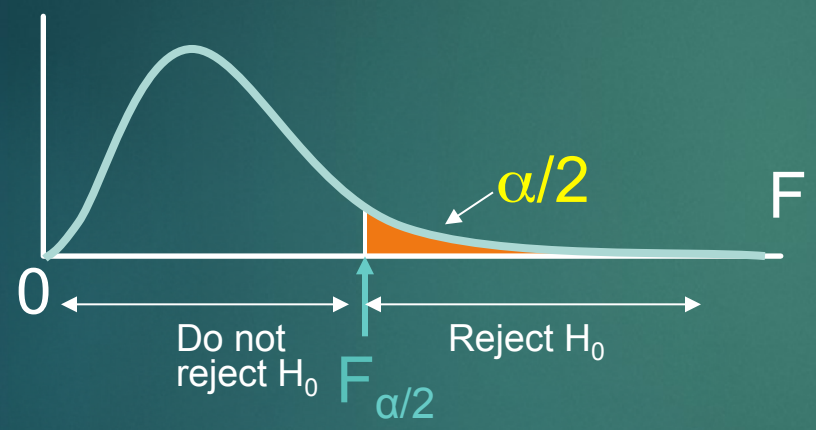


Reject H_0 if $F_{STAT} > F_\alpha$

Copyright © 2015, 2012, 2009 Pearson Education, Inc.

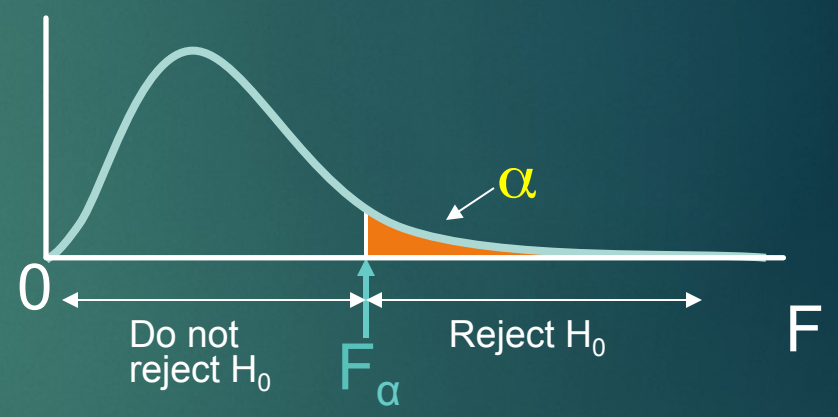
Finding the Rejection Region

$H_0: \sigma_1^2 = \sigma_2^2$
$H_1: \sigma_1^2 \neq \sigma_2^2$



Reject H_0 if $F_{STAT} > F_{\alpha/2}$

$H_0: \sigma_1^2 \leq \sigma_2^2$
$H_1: \sigma_1^2 > \sigma_2^2$



Reject H_0 if $F_{STAT} > F_{\alpha}$

Copyright © 2015, 2012, 2009 Pearson Education, Inc.

Example: Variability in Waiting Time

- The management of a local restaurant likes to test if the variation in the service time is different during lunch hours and dinner hours.
  - H_0: $\sigma^2_1 = \sigma^2_2$ (there is no difference between variances)
  - H_1: $\sigma^2_1 \neq \sigma^2_2$ (there is a difference between variances)

Copyright © 2015, 2012, 2009 Pearson Education, Inc.

Example: Variability in Waiting Time

- The management of a local restaurant likes to test if the variation in the service time is different during lunch hours and dinner hours.
  - H_0: $\sigma^2_1 = \sigma^2_2$ (there is no difference between variances)
  - H_1: $\sigma^2_1 \neq \sigma^2_2$ (there is a difference between variances)
- Samples:
  - Lunch: $n_1 = 40 \; and \; S_1^2 = 3.2$
  - Dinner: $n_2 = 40 \; and \; S_2^2 = 4.0$

Copyright © 2015, 2012, 2009 Pearson Education, Inc.

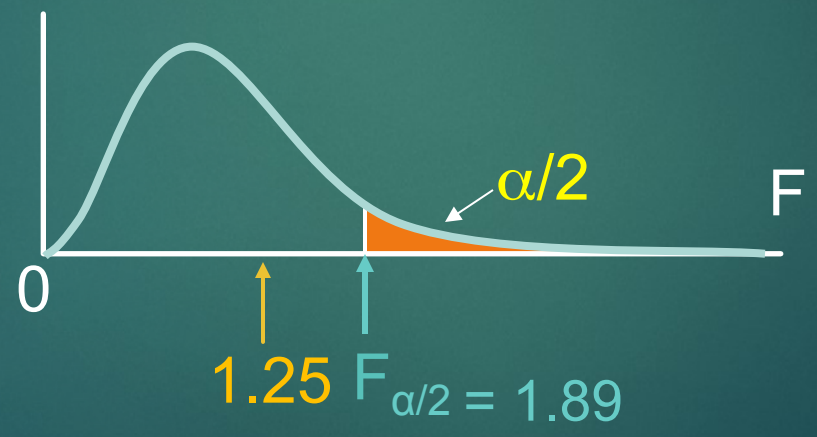
Example: Variability in Waiting Time

- F Stat=4/3.2=1.25
- The Critical Value for 0.05 Significance Level and 39 df for the numerator and 39 df for the denominator is given by F.INV(0.975,39,39), which yields 1.89.

Copyright © 2015, 2012, 2009 Pearson Education, Inc.

Example: Variability in Waiting Time

- F Stat=4/3.2=1.25
- The Critical Value = 1.89
- Since 1.25<1.89, we are not able to reject the null hypothesis with 95% confidence.



Not able to reject H_0 since $F_{STAT} < F_{\alpha/2}$

Copyright © 2015, 2012, 2009 Pearson Education, Inc.

F Test: Another Example

	NYSE	NASDAQ
Number	**21**	**25**
Mean	**3.27**	**2.53**
Std dev	**1.30**	**1.16**

Is there a difference in the variances between the NYSE & NASDAQ at the $\alpha = 0.05$ level?

Copyright © 2015, 2012, 2009 Pearson Education, Inc.

F Test: Another Example

- Form the hypothesis test:

  H_0: $\sigma^2_1 = \sigma^2_2$ (there is no difference between variances)

  H_1: $\sigma^2_1 \neq \sigma^2_2$ (there is a difference between variances)

Copyright © 2015, 2012, 2009 Pearson Education, Inc.

F Test: Another Example

- Form the hypothesis test:

  H_0: $\sigma^2_1 = \sigma^2_2$ (there is no difference between variances)

  H_1: $\sigma^2_1 \neq \sigma^2_2$ (there is a difference between variances)

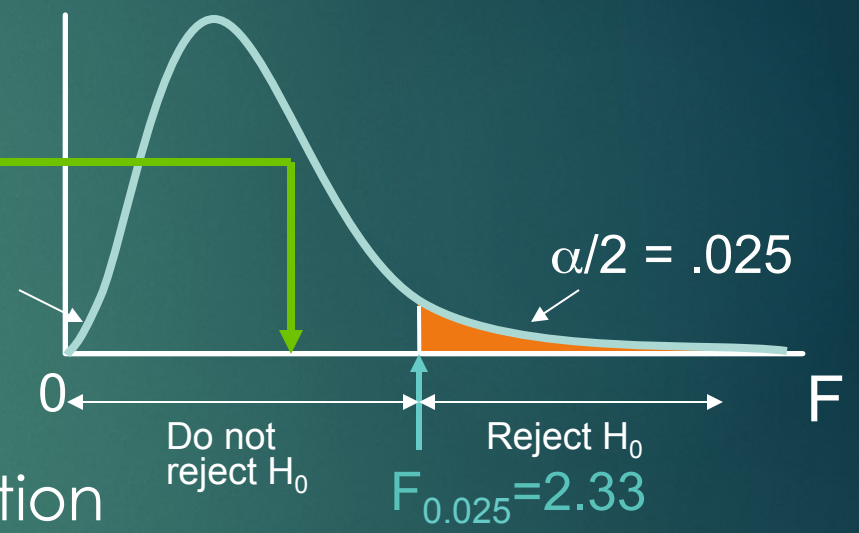
  - Find the F critical value for $\alpha = 0.05$:
    - Numerator d.f. = $n_1 - 1 = 21 - 1 = 20$
    - Denominator d.f. = $n2 - 1 = 25 - 1 = 24$
  - $F_{\alpha/2} = F_{.025,\ 20,\ 24} = 2.33$

Copyright © 2015, 2012, 2009 Pearson Education, Inc.

F Test: Another Example

- The test statistic is:

$$F_{STAT} = \frac{S_1^2}{S_2^2} = \frac{1.30^2}{1.16^2} = 1.256$$



- F_{STAT} = 1.256 is not in the rejection region, so we do not reject H_0
- Conclusion: There is not sufficient evidence of a difference in variances at α = .05

Copyright © 2015, 2012, 2009 Pearson Education, Inc.

F Test: Another Example

- Reminder:

  We compared the dividend yields (i.e. dividends, adjusted by prices) between the stocks listed on the NYSE & NASDAQ.

  - We rejected the null that the dividend yields are the same when we performed a Pooled Variance Test, assuming that the variances are equal.
  - We did not reject the null that the dividend yields are the same when we performed a Separate Variance Test, assuming that the variances are not equal.
- Given the new findings about variances, the findings that are based on Pooled Variance Test seem to be relevant.

Copyright © 2015, 2012, 2009 Pearson Education, Inc.