Chapter 10, Part A

Inference about Means and Proportions with Two Populations

- Inferences About the Difference Between Two Population Means: σ_1 and σ_2 Known
- Inferences About the Difference Between Two Population Means: σ_1 and σ_2 Unknown
- Inferences About the Difference Between Two Population Means: Matched Samples

Inferences about the Difference between Two Population Means: σ_1 and σ_2 Known

- Interval Estimation of $\mu_1 - \mu_2$
- Hypothesis Tests About $\mu_1 - \mu_2$

Estimating the Difference between Two Population Means (1 of 2)

- Let μ_1 equal the mean of population 1 and μ_2 equal the mean of population 2.
- The difference between the two population means is $\mu_1 - \mu_2$.
- To estimate $\mu_1 - \mu_2$, we will select a simple random sample of size n_1 from population 1 and a simple random sample of size n_2 from population 2.
- Let $\bar{x}_1$ equal the mean of sample 1 and $\bar{x}_2$ equal the mean of sample 2.
- The point estimator of the difference between the means of the populations 1 and 2 is $\bar{x}_1 - \bar{x}_2$.

Sampling Distribution of $\bar{x}_1 - \bar{x}_2$

- Standard Deviation (Standard Error)

$$\sigma_{\bar{x}_1-\bar{x}_2} = \sqrt{\frac{(\sigma_1)^2}{n_1} + \frac{(\sigma_2)^2}{n_2}}$$

where: σ_1 = standard deviation of population 1

σ_2 = standard deviation of population 2

n_1 = sample size from population 1

n_2 = sample size from population 2

CENGAGE

Interval Estimate of $\mu_1 - \mu_2$: σ_1 and σ_2 Known (1 of 4)

Interval Estimate

$$\left(\bar{x}_1 - \bar{x}_2\right) \pm z_{\alpha/2}\sqrt{\frac{(\sigma_1)^2}{n_1} + \frac{(\sigma_2)^2}{n_2}}$$

where:

$1 - \alpha$ is the confidence coefficient

Interval Estimate of $\mu_1 - \mu_2$: σ_1 and σ_2 Known (2 of 4)

Example: Homestyle Furniture

Homestyle sells furniture at two stores in Buffalo, New York: One is in the inner city and other in suburban shopping centre. There is difference in the types of furniture sold in each store, and manager believes this can be attributed to the difference in customer demographics.

The manager wants to investigate the difference in mean age of customers who shop at two stores.

Interval Estimate of $\mu_1 - \mu_2$: σ_1 and σ_2 Known (3 of 4)

Example: Homestyle Furniture

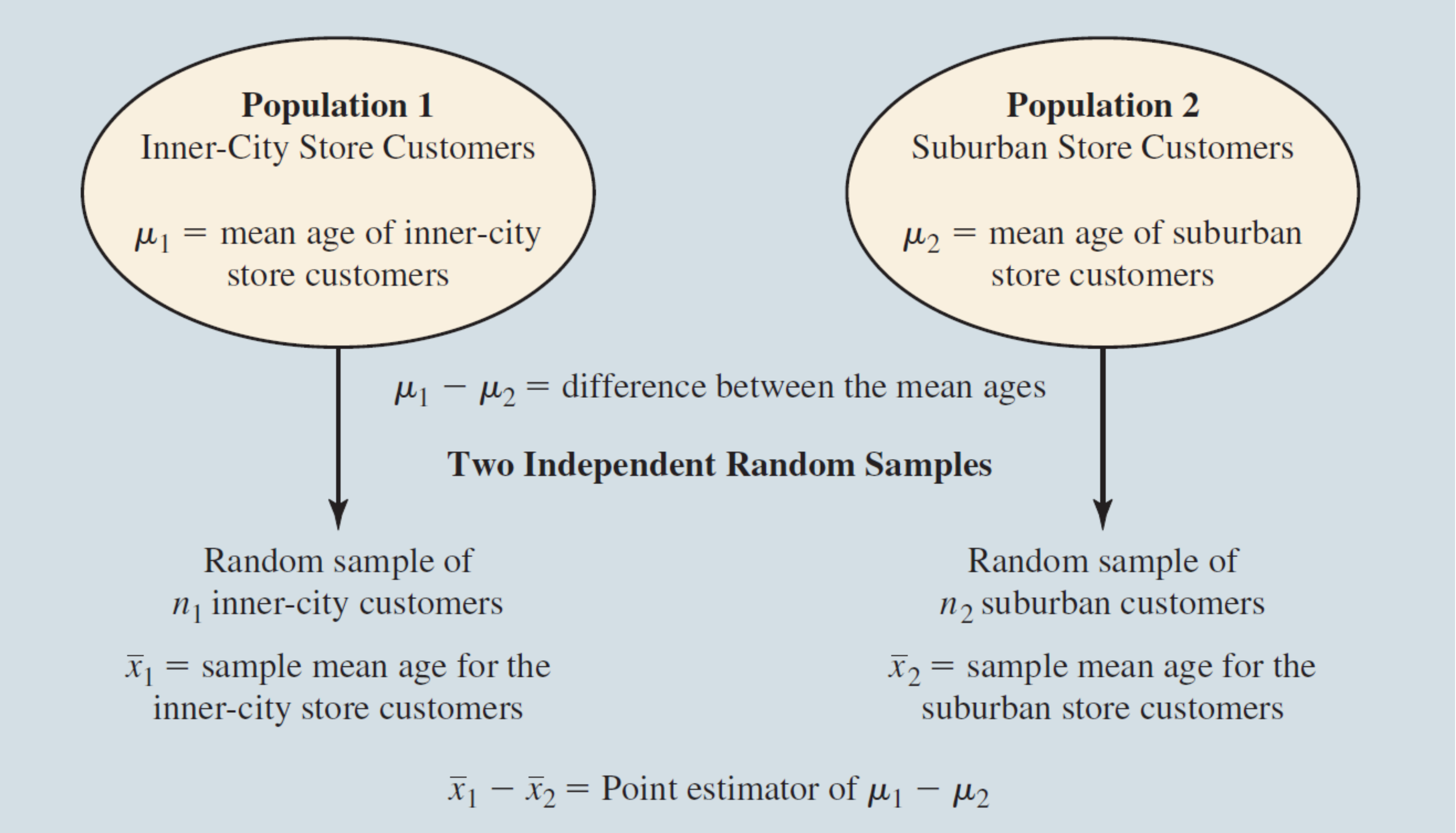
	Inner-City Store	Suburban Store
Sample Size	36	49
Sample Mean	40 years	35 years
Standard deviation	9 years	10 years
(Based on previous studies)		

Interval Estimate of $\mu_1 - \mu_2$: σ_1 and σ_2 Known (4 of 4)

Example: Homestyle Furniture

Let us develop a 95% confidence interval estimate of the difference between the mean age of the customers who shop at two stores.

Estimating the Difference between Two Population Means (2 of 2)



CENGAGE

Point Estimate of $\mu_1 - \mu_2$ (1 of 2)

Example: Homestyle Furniture

Point estimate of $\mu_1 - \mu_2 = \bar{x}_1 - \bar{x}_2 = 40 - 35 = 5$ years

where: $\mu_1 =$ mean age of inner-city store customers

$\mu_2 =$ mean age of suburban store customers

Interval Estimation of $\mu_1 - \mu_2$: σ_1 and σ_2 Known (1 of 2)

Example: Homestyle Furniture

$$\left(\bar{x}_1 - \bar{x}_2\right) \pm z_{\alpha/2}\sqrt{\frac{(\sigma_1)^2}{n_1} + \frac{(\sigma_2)^2}{n_2}} = 5 \pm 1.96\sqrt{\frac{(9)^2}{36} + \frac{(10)^2}{49}}$$

5 ± 4.06 or .94 years to 9.06 years

We are 95% confident that the difference between mean age of inner-city and suburban store customers is .94 years to 9.06 years.

Interval Estimation of $\mu_1 - \mu_2$: σ_1 and σ_2 Known (2 of 2)

Example: Homestyle Furniture

- Excel Formula and Value Worksheet

	A	B	C	D	E	F	G	H
1	Inner City	Suburban		Interval Estimate of Difference in Population Means:				
2	38	29		σ_1 and σ_2 Known Case				
3	46	35						
4	32	39			Inner City	Suburban		
5	23	10		Sample Size	=COUNT(A2:A37)	=COUNT(B2:B50)		
6	39	37		Sample Mean	=AVERAGE(A2:A37)	=AVERAGE(B2:B50)		
7	40	52						
8	35	40		Population Standard Deviation	9	10		
9	35	37		Standard Error	=SQRT(E8^2/E5+F8^2/F5			
10	36	43						
11	41	38		Confidence Coefficient	0.95			
12	32	28		Level of Significance	=1-E11			
13	38	37		z Value	=NORM.S.INV(1-E12/2)			
14	44	51		Margin of Error	=E13*E9			
15	50	23						
16	47	25		Point Estimate of Difference	=E6-F6			
17	59	37		Lower Limit	=E16-E14			
18	38	38		Upper Limit	=E16+E14			
36	44	19						
37	62	40						
49		22						
50		47						
51								

	A	B	C	D	E	F	G	H
1	Inner City	Suburban		Interval Estimate of Difference in Population Means:				
2	38	29		σ_1 and σ_2 Known Case				
3	46	35						
4	32	39			Inner City	Suburban		
5	23	10		Sample Size	36	49		
6	39	37		Sample Mean	40	35		
7	40	52						
8	35	40		Population Standard Deviation	9	10		
9	35	37		Standard Error	2.07			
10	36	43						
11	41	38		Confidence Coefficient	0.95			
12	32	28		Level of Significance	0.05			
13	38	37		z Value	1.960			
14	44	51		Margin of Error	4.06			
15	50	23						
16	47	25		Point Estimate of Difference	5			
17	59	37		Lower Limit	0.94			
18	38	38		Upper Limit	9.06			
36	44	19						
37	62	40						
49		22						
50		47						
51								

Hypothesis Tests about $\mu_1 - \mu_2$: σ_1 and σ_2 Known (1 of 8)

- Hypotheses

Left-tailed	Right-tailed	Two-tailed
$H_0\colon \mu_1 - \mu_2 \geq D_0$	$H_0\colon \mu_1 - \mu_2 \leq D_0$	$H_0\colon \mu_1 - \mu_2 = D_0$
$H_a\colon \mu_1 - \mu_2 < D_0$	$H_a\colon \mu_1 - \mu_2 > D_0$	$H_a\colon \mu_1 - \mu_2 \neq D_0$

- Test Statistic

$$z = \frac{(\bar{x}_1 - \bar{x}_2) - D_0}{\sqrt{\frac{(\sigma_1)^2}{n_1} + \frac{(\sigma_2)^2}{n_2}}}$$

Hypothesis Tests about $\mu_1 - \mu_2$: σ_1 and σ_2 Known (2 of 8)

Example: Training Centers

A standardized examination was given to the individuals who are trained at two different centres to evaluate the difference in education quality between them.

Let

$\mu_1 =$ The mean examination score for the population of individuals trained at center A

$\mu_2 =$ The mean examination score for the population of individuals trained at center B

Hypothesis Tests about $\mu_1 - \mu_2$: σ_1 and σ_2 Known (3 of 8)

Example: Training Centers

	A	B
Sample Size	30	40
Sample Mean	82	78
Standard deviation	10	10
(Based on previous studies)		

Hypothesis Tests about $\mu_1 - \mu_2$: σ_1 and σ_2 Known (4 of 8)

Example: Training Centers

Can we conclude, using $\alpha = .05$, that no difference exists between the training quality provided at the two centers?

Hypothesis Tests about $\mu_1 - \mu_2$: σ_1 and σ_2 Known (5 of 8)

Example: Training Centers

p -Value and Critical Value Approaches

1. Develop the hypotheses.

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

(two-tailed test)

where μ_1 = The mean examination score for the population of individuals trained at center A

μ_2 = The mean examination score for the population of individuals trained at center B

2. Specify the level of significance $\alpha = .05$

Hypothesis Tests about $\mu_1 - \mu_2$: σ_1 and σ_2 Known (6 of 8)

Example: Training Centers

p-Value and Critical Value Approaches

3. Compute the value of the test statistic.

$$z = \frac{(\bar{x}_1 - \bar{x}_2) - D_0}{\sqrt{\frac{(\sigma_1)^2}{n_1} + \frac{(\sigma_2)^2}{n_2}}}$$

$$z = \frac{(82 - 78) - 0}{\sqrt{\frac{(10)^2}{30} + \frac{(10)^2}{40}}} = \frac{4}{2.4152} = 1.66$$

Hypothesis Tests about $\mu_1 - \mu_2$: σ_1 and σ_2 Known (7 of 8)

Example: Training Centers

p-Value Approach

4. Compute the *p*–value.

   For $z = 1.66$, the area to the left is .9515.

   The area in the upper tail of the distribution is

$$1.0000 - .9515 = .0485$$

$$p\text{-value} = 2(.0485) = .0970$$

5. Determine whether to reject H_0.

$$\text{Because } p\text{-value} > \alpha = .05\text{, we cannot reject } H_0.$$

At the .05 level of significance, the sample evidence indicates there is no difference in quality between training centers.

Hypothesis Tests about $\mu_1 - \mu_2$: σ_1 and σ_2 Known (8 of 8)

Example: Training Centers

Critical Value Approach

4. Determine the critical value and rejection rule.

$$\text{For } \alpha = .05,\ z_{.025} = 1.96$$

$$\text{Reject } H_0 \text{ if } z \geq 1.96$$

5. Determine whether to reject H_0.

$$\text{Because } z = 1.66 < 1.96, \text{ we cannot reject } H_0.$$

At the .05 level of significance, the sample evidence indicates there is no difference in quality between training centers.

Excel's "*z*-Test: Two Sample for Means" Tool (1 of 3)

Step 1 Click the **Data** tab on the Ribbon.

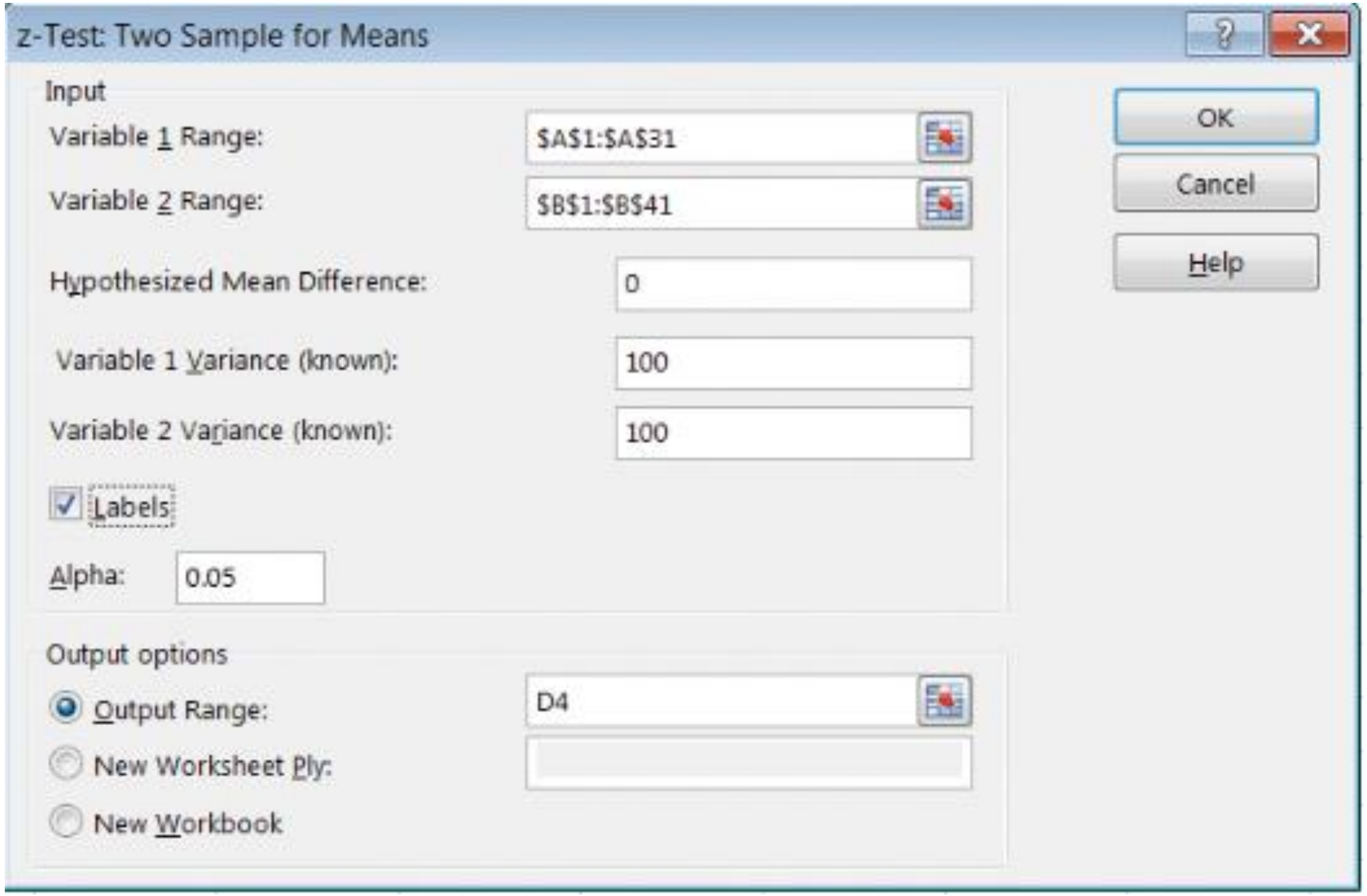
Step 2 In the **Analysis** group, click **Data Analysis.**

Step 3 Choose ***z*-Test: Two Sample for Means** from the list of Analysis Tools.

Step 4 When the *z*-Test: Two Sample for Means dialog box appears, as shown on the next slide.

Excel's "*z*-Test: Two Sample for Means" Tool (2 of 3)

- Enter *A1:A31* in the **Variable 1 Range** box.
- Enter *B1: B41* in the **Variable 2 Range** box.
- Enter *0* in **Hypothesized Mean Difference** box.
- Enter *100* in **Variable 1 and 2 Variance (known).**
- Select the check box for **Labels.**
- Enter *.05* in the **Alpha** box.
- Select **output range** and enter *D4.*
- Click **OK.**



Excel's "*z*-Test: Two Sample for Means" Tool (3 of 3)

- Excel results for the hypothesis test about equality of exam scores at two training centers.

	A	B	C	D	E	F	G
1	Center A	Center B					
2	97	64					
3	95	85					
4	89	72		z-Test: Two Sample for Means			
5	79	64					
6	78	74			*Center A*	*Center B*	
7	87	93		Mean	82	78	
8	83	70		Known Variance	100	100	
9	94	79		Observations	30	40	
10	76	79		Hypothesized Mean Difference	0		
11	79	75		z	1.6562		
12	83	66		P(Z<=z) one-tail	0.0488		
13	84	83		z Critical one-tail	1.6449		
14	76	74		P(Z<=z) two-tail	0.0977		
15	82	70		z Critical two-tail	1.9600		
16	85	82					
17	85	82					
29	88	65					
30	60	78					
31	73	66					
32		84					
40		80					
41		76					
42							

Inferences about the Difference between Two Population Means: σ_1 and σ_2 Unknown

- Interval Estimation of $\mu_1 - \mu_2$
- Hypothesis Tests about $\mu_1 - \mu_2$

Interval Estimation of $\mu_1 - \mu_2$: σ_1 and σ_2 Unknown (1 of 5)

When σ_1 and σ_2 are unknown, we will:

- Use the sample standard deviations s_1 and s_2 as estimates of σ_1 and σ_2, and
- Replace $z_{\alpha/2}$ with $t_{\alpha/2}$.

Interval Estimation of $\mu_1 - \mu_2$: σ_1 and σ_2 Unknown (2 of 5)

Interval Estimate

$$\bar{x}_1 - \bar{x}_2 \pm t_{\alpha/2}\sqrt{\frac{(s_1)^2}{n_1} + \frac{(s_2)^2}{n_2}}$$

where the degrees of freedom for $t_{\alpha/2}$ are:

$$df = \frac{\left(\frac{(s_1)^2}{n_1} + \frac{(s_2)^2}{n_2}\right)^2}{\frac{1}{n_1 - 1}\left(\frac{(s_1)^2}{n_1}\right)^2 + \frac{1}{n_2 - 1}\left(\frac{(s_2)^2}{n_2}\right)^2}$$

CENGAGE

Difference between Two Population Means: σ_1 and σ_2 Unknown (1 of 3)

Example: Clearwater National Bank

Clearwater National Bank wants to compare the account checking practices by the customers at two of its branch banks: Cherry Grove Branch and Beechmont Branch. A random sample of 28 and 22 checking accounts is selected from these branches respectively. The sample statistics are shown on the next slide.

Difference between Two Population Means: σ_1 and σ_2 Unknown (2 of 3)

Example: Clearwater National Bank

	Cherry Grove	Beechmont
Sample Size	28	22
Sample Mean	$1025	$910
Sample Standard Deviation	$150	$125

Difference between Two Population Means: σ_1 and σ_2 Unknown (3 of 3)

<u>Example</u>: Clearwater National Bank

Let us develop a 95% confidence interval estimate of the difference between the population mean checking account balances at the two branch banks.

Point Estimate of $\mu_1 - \mu_2$ (2 of 2)

Example: Clearwater National Bank

Point estimate of $\mu_1 - \mu_2 = \bar{x}_1 - \bar{x}_2 = 1{,}025 - 910 = 115$

where: $\mu_1 =$ mean checking account balance maintained by the population of Cherry Grove customers

$\mu_2 =$ mean checking account balance maintained by the population of Beechmont customers

Interval Estimation of $\mu_1 - \mu_2$: σ_1 and σ_2 Unknown (3 of 5)

Example: Clearwater National Bank

The degrees of freedom for $t_{\alpha/2}$ are:

$$df = \frac{\left[\frac{(150)^2}{28} + \frac{(125)^2}{22}\right]^2}{\frac{1}{28-1}\left[\frac{(150)^2}{28}\right]^2 + \frac{1}{22-1}\left[\frac{(125)^2}{22}\right]^2} = 47.8 = 47$$

with $\alpha/2 = .025$ and $df = 47$, $t_{\alpha/2} = 2.012$

Interval Estimation of $\mu_1 - \mu_2$: σ_1 and σ_2 Unknown (4 of 5)

Example: Clearwater National Bank

$$\bar{x}_1 - \bar{x}_2 \pm t_{\alpha/2}\sqrt{\frac{(s_1)^2}{n_1} + \frac{(s_2)^2}{n_2}}$$

$$1025 - 910 \pm 2.012\sqrt{\frac{(150)^2}{28} + \frac{(125)^2}{22}}$$

$$115 \pm 78 = \$37 \text{ to } \$193$$

We are 95% confident that the difference between the mean accounting checking balances maintained by the customers at Cherry Grove branch and the Beechmont branch is \$37 to \$193.

Interval Estimation of $\mu_1 - \mu_2$: σ_1 and σ_2 Unknown (5 of 5)

Example: Clearwater National Bank

- Excel Formula and Value Worksheet

	A	B	C	D	E	F	G
1	Cherry Grove	Beechmont		Interval Estimate of Difference in Population Means:			
2	1263	996.7		σ_1 and σ_2 Unknown Case			
3	897	897					
4	849	912			Cherry Grove	Beechmont	
5	891	894.9		Sample Size	=COUNT(A2:A29)	=COUNT(B2:B23)	
6	964	785		Sample Mean	=AVERAGE(A2:A29)	=AVERAGE(B2:B23)	
7	810	750.7		Sample Standard Deviation	=STDEV.S(A2:A29)	=STDEV.S(B2:B23)	
8	877	882.2					
9	899	1110		Estimate of Variance	=E7^2/E5+F7^2/F5		
10	847	907.2		Standard Error	=SQRT(E9)		
11	1070	1226.1					
12	1252	762.1		Confidence Coefficient	0.95		
13	920	835.5		Level of Significance	=1-E12		
14	1256	1048		Degrees of Freedom	=E9^2/((1/(E5-1))*(E7^2/E5)^2+(1/(F5-1)*(F7^2/F5)^2))		
15	1196	773.8		*t* Value	=T.INV.2T(E13,E14)		
16	1150	807		Margin of Error	=E15*E10		
17	1024	972					
18	1016	980		Point Estimate of Difference	=E6-F6		
19	1126	876.6		Lower Limit	=E18-E16		
20	1289	943		Upper Limit	=E18+E16		
21	1220	992.7					
22	912	704.3					
23	1026	962.9					
24	786						
25	989						
26	1133						
27	990						
28	999						
29	1049						
30							

	A	B	C	D	E	F	G
1	Cherry Grove	Beechmont		Interval Estimate of Difference in Population Means:			
2	1263	997		σ_1 and σ_2 Unknown Case			
3	897	897					
4	849	912			Cherry Grove	Beechmont	
5	891	895		Sample Size	28	22	
6	964	785		Sample Mean	1025	910	
7	810	751		Sample Standard Deviation	150	125	
8	877	882					
9	899	1110		Estimate of Variance	1513.8550		
10	847	907		Standard Error	38.9083		
11	1070	1226					
12	1252	762		Confidence Coefficient	0.95		
13	920	836		Level of Significance	0.05		
14	1256	1048		Degrees of Freedom	47.8		
15	1196	774		*t* Value	2.012		
16	1150	807		Margin of Error	78		
17	1024	972					
18	1016	980		Point Estimate of Difference	115		
19	1126	877		Lower Limit	37		
20	1289	943		Upper Limit	193		
21	1220	993					
22	912	704					
23	1026	963					
24	786						
25	989						
26	1133						
27	990						
28	999						
29	1049						
30							

Hypothesis Tests about $\mu_1 - \mu_2$: σ_1 and σ_2 Unknown (1 of 7)

- Hypotheses

Left-tailed	Right-tailed	Two-tailed
$H_0\colon \mu_1 - \mu_2 \geq D_0$	$H_0\colon \mu_1 - \mu_2 \leq D_0$	$H_0\colon \mu_1 - \mu_2 = D_0$
$H_a\colon \mu_1 - \mu_2 < D_0$	$H_a\colon \mu_1 - \mu_2 > D_0$	$H_a\colon \mu_1 - \mu_2 \neq D_0$

- Test Statistic

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - D_0}{\sqrt{\dfrac{(s_1)^2}{n_1} + \dfrac{(s_2)^2}{n_2}}}$$

Hypothesis Tests about $\mu_1 - \mu_2$: σ_1 and σ_2 Unknown (2 of 7)

Example: Computer software package

A new computer software package is developed to reduce the time required to design, develop and implement an information system. To evaluate the benefits a random sample of 24 system analysts is selected, 12 of them using current technology and other 12 using new software package.

Can we conclude, using a .05 level of significance, that the mean project completion time for system analysts using the new software package is lesser than the mean project completion time for system analysts using current technology?

Hypothesis Tests about $\mu_1 - \mu_2$: σ_1 and σ_2 Unknown (3 of 7)

<u>Example:</u> Computer software package

Summary Statistics		
Sample Size	12	12
Sample Mean	325	286
Sample SD	40	44

Current Technology	New Software
300	274
280	220
344	308
385	336
372	198
360	300
288	315
321	258
376	318
290	310
301	332
283	263

Hypothesis Tests about $\mu_1 - \mu_2$: σ_1 and σ_2 Unknown (4 of 7)

Example: Computer software package

p-Value approach

1. Develop the hypotheses. Right-tailed test:

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

$$H_a: \mu_1 - \mu_2 > 0$$

where:

$\mu_1 =$ the mean project completion time for system analysts using the current technology

$\mu_2 =$ the mean project completion time for system analysts using the new software package

Hypothesis Tests about $\mu_1 - \mu_2$: σ_1 and σ_2 Unknown (5 of 7)

Example: Computer software package

p-Value approach

2. Specify the level of significance $-\alpha = .05$

3. Compute the value of the test statistic.

$$t = \frac{(325 - 286) - 0}{\sqrt{\frac{(40)^2}{12} + \frac{(44)^2}{12}}} = 2.27$$

Hypothesis Tests about $\mu_1 - \mu_2$: σ_1 and σ_2 Unknown (6 of 7)

Example: Computer software package

p-Value approach

4. Compute the *p*-value.

   The degrees of freedom for t_α are:

$$df = \frac{\left[\frac{(40)^2}{12} + \frac{(44)^2}{12}\right]^2}{\frac{1}{12-1}\left[\frac{(40)^2}{12}\right]^2 + \frac{1}{12-1}\left[\frac{(44)^2}{12}\right]^2} = 21.8 = 21$$

CENGAGE

Hypothesis Tests about $\mu_1 - \mu_2$: σ_1 and σ_2 Unknown (7 of 7)

Example: Computer software package

Area in Upper Tail	.20	.10	.05	.025	.01	.005
t-Value (21 *df*)	0.859	1.323	1.721	2.080	2.518	2.831

$t = 2.27$

From the table, we see *p*-value is between .025 and .01.

5. Determine whether to reject H_0.

Because *p*-value $\leq \alpha = .05$, we reject H_0.

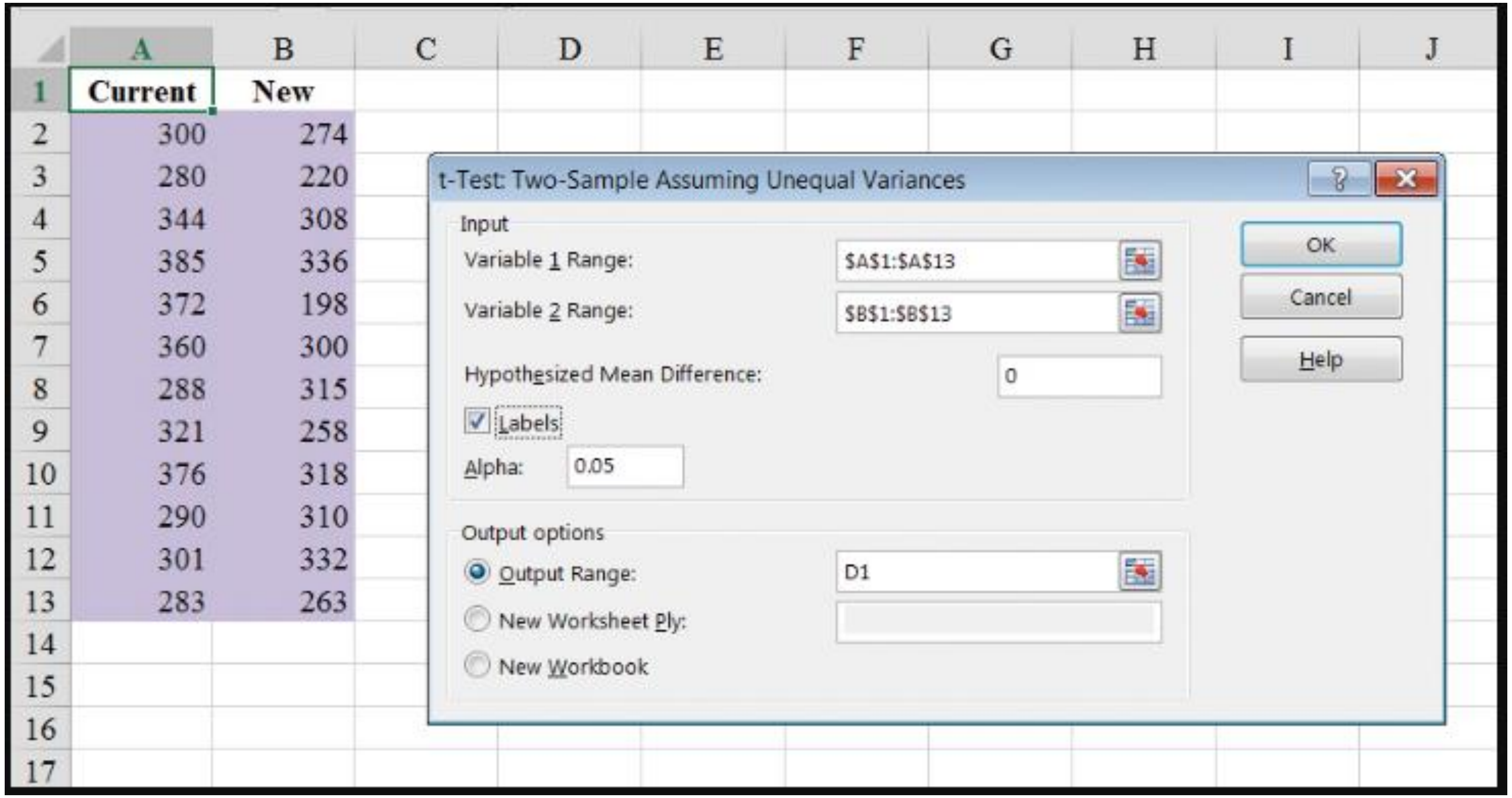
There is sufficient statistical evidence that $\mu_1 - \mu_2 > 0$ or $\mu_1 > \mu_2$ (i.e., new software package provides a smaller population mean completion time).

Excel's "*t*-Test: Two-Sample Assuming Unequal Variances" Tool (1 of 3)

- Step 1 Click the **Data** tab on the Ribbon.
- Step 2 In the **Analysis** group, click **Data Analysis.**
- Step 3 Choose ***t*-Test: Two-Sample Assuming Unequal Variances** from the list of Analysis Tools.
- Step 4 When the *t*-Test: Two-Sample Assuming Unequal Variances dialog box appears, as shown on the following slide:

Excel's "*t*-Test: Two-Sample Assuming Unequal Variances" Tool (2 of 3)

- Enter *A1:A13* in the **Variable 1 Range** box.
- Enter *B1: B13* in the **Variable 2 range** box.
- Enter *0* in **Hypothesized Mean Difference** box.
- Select **Labels.**
- Enter *.05* in the **Alpha** box.
- Select **Output Range** and enter *D1.*
- Click **OK.**



Excel's "*t*-Test: Two-Sample Assuming Unequal Variances" Tool (3 of 3)

- Excel Value Worksheet

	A	B	C	D	E	F	G
1	Current	New		t-Test: Two-Sample Assuming Unequal Variances			
2	300	274					
3	280	220			*Current*	*New*	
4	344	308		Mean	325	286	
5	385	336		Variance	1599.6364	1935.8182	
6	372	198		Observations	12	12	
7	360	300		Hypothesized Mean Difference	0		
8	288	315		df	22		
9	321	258		t Stat	2.2721		
10	376	318		P(T<=t) one-tail	0.0166		
11	290	310		t Critical one-tail	1.7171		
12	301	332		P(T<=t) two-tail	0.0332		
13	283	263		t Critical two-tail	2.0739		
14							

CENGAGE

Inferences about the Difference between Two Population Means: Matched Samples (1 of 6)

- With a matched-sample design, each sampled item provides a pair of data values.

- This design often leads to a smaller sampling error than the independent-sample design because variation between sampled items is eliminated as a source of sampling error.

Inferences about the Difference between Two Population Means: Matched Samples (2 of 6)

<u>Example</u>: Comparison of production methods

Two production methods are tested under similar conditions. A random sample of six workers is used.

Task Completion Times For a Matched Sample Design			
Worker	**Completion Time for Method 1 (Minutes)**	**Completion Time for Method 2 (Minutes)**	**Difference in Completion Times (d_i)**
1	6.0	5.4	.6
2	5.0	5.2	-.2
3	7.0	6.5	.5
4	6.2	5.9	.3
5	6.0	6.0	.0
6	6.4	5.8	.6

CENGAGE

Inferences about the Difference between Two Population Means: Matched Samples (3 of 6)

Example: Comparison of production methods

Each worker provides a pair of data values, one for each production method. The test is conducted to determine if the mean completion times differ between the two methods.

Inferences about the Difference between Two Population Means: Matched Samples (4 of 6)

p-value approach

1. Develop the hypotheses.

$$H_0 : \mu_d = 0$$
$$H_a : \mu_d \neq 0$$

Let μ_d = the mean of the difference in values for the population of workers

Inferences about the Difference between Two Population Means: Matched Samples (5 of 6)

p-value approach

2. Specify the level of significance. $\alpha = .05$

3. Compute the value of the test statistic.

$$\bar{d} = \frac{\sum d_i}{n} = \frac{(1.8)}{6} = .30$$

$$s_d = \sqrt{\frac{\sum (d_i - \bar{d})^2}{n-1}} = \sqrt{\frac{.56}{5}} = .335$$

$$t = \frac{\bar{d} - \mu_d}{s_d / \sqrt{n}} = \frac{.30 - 0}{.335/\sqrt{6}} = 2.20$$

Inferences about the Difference between Two Population Means: Matched Samples (6 of 6)

p-value approach

4. Compute the *p*-value.

   For $t = 2.20$ and $df = 5$, the *p*-value is between .10 and .05. (This is a two-tailed test, so we double the upper-tail areas of .05 and .025.)

Area in Upper Tail	.20	.10	.05	.025	.01	.005
t-Value (5 *df*)	0.920	1.476	2.015	2.571	3.365	4.032

$t = 2.20$

5. Determine whether to reject H_0.

   Because *p*-value $> \alpha = .05$, we cannot reject H_0.

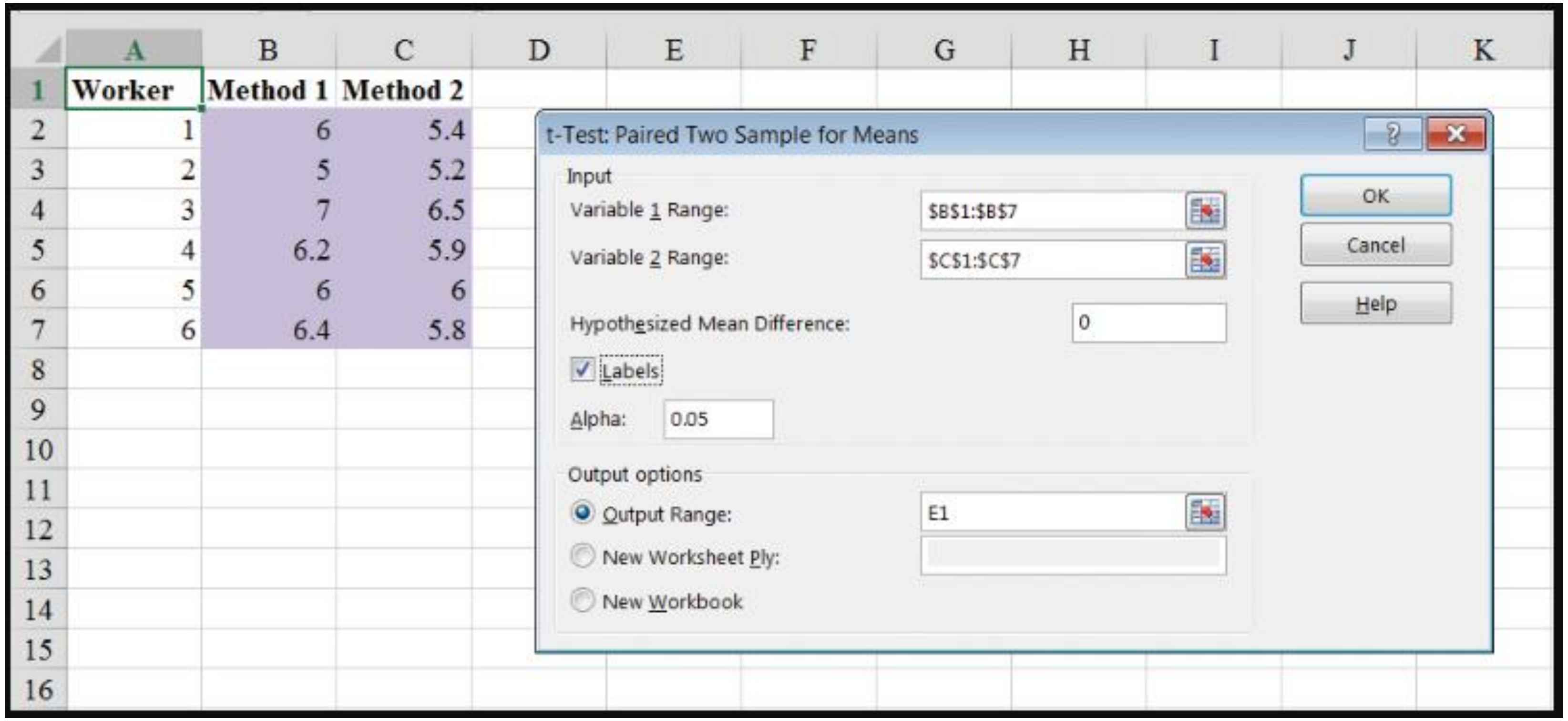
Excel's "*t*-Test: Paired Two Sample for Means" Tool (1 of 3)

- Step 1 Click the **Data** tab on the Ribbon.
- Step 2 In the **Analysis** group, click **Data Analysis.**
- Step 3 Choose ***t*-Test: Paired Two Sample for Means** from the list of Analysis Tools.
- Step 4 When the *t*-Test: Paired Two Sample for Means dialog box appears, as shown on the following slide:

CENGAGE

Excel's "*t*-Test: Paired Two Sample for Means" Tool (2 of 3)

- Enter *B1:B7* in the **Variable 1 Range** box.
- Enter *C1:C7* in the **Variable 2 Range** box.
- Enter *0* in **Hypothesized Mean Difference** box.
- Select **Labels.**
- Enter *.05* in the **Alpha** box.
- Select **Output Range** and enter *E1.*
- Click **OK.**



Excel's "*t*-Test: Paired Two Sample for Means" Tool (3 of 3)

- Excel Value Worksheet

	A	B	C	D	E	F	G	H
1	Worker	Method 1	Method 2		t-Test: Paired Two Sample for Means			
2	1	6	5.4					
3	2	5	5.2			*Method 1*	*Method 2*	
4	3	7	6.5		Mean	6.1	5.8	
5	4	6.2	5.9		Variance	0.428	0.212	
6	5	6	6		Observations	6	6	
7	6	6.4	5.8		Pearson Correlation	0.8764		
8					Hypothesized Mean Difference	0		
9					df	5		
10					t Stat	2.196		
11					P(T<=t) one-tail	0.0398		
12					t Critical one-tail	2.015		
13					P(T<=t) two-tail	0.0795		
14					t Critical two-tail	2.571		
15								

CENGAGE