Two-Sample Tests of Hypothesis

Chapter 11

McGraw-Hill/Irwin

GOALS

1. Conduct a test of a hypothesis about the difference between two independent population means.
2. Conduct a test of a hypothesis about the difference between two population proportions.
3. Conduct a test of a hypothesis about the mean difference between *paired* or *dependent observations*.
4. Understand the difference between *dependent and independent samples.*

Comparing two populations – Some Examples

1. Is there a difference in the mean value of residential real estate sold by male agents and female agents in south Florida?
2. Is there a difference in the mean number of defects produced on the day and the afternoon shifts at Kimble Products?
3. Is there a difference in the mean number of days absent between young workers (under 21 years of age) and older workers (more than 60 years of age) in the fast-food industry?
4. Is there is a difference in the proportion of Ohio State University graduates and University of Cincinnati graduates who pass the state Certified Public Accountant Examination on their first attempt?
5. Is there an increase in the production rate if music is piped into the production area?

Comparing Two Population Means

- No assumptions about the shape of the populations are required.
- The samples are from independent populations.
- The formula for computing the value of z is:

Use if sample sizes > 30
or if σ_1 and σ_2 are known

$$z = \frac{\overline{X}_1 - \overline{X}_2}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

Use if sample sizes > 30
and if σ_1 and σ_2 are unknown

$$z = \frac{\overline{X}_1 - \overline{X}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

Comparing Two Population Means - Example

The U-Scan facility was recently installed at the Byrne Road Food-Town location. The store manager would like to know if the mean checkout time using the standard checkout method is longer than using the U-Scan. She gathered the following sample information. The time is measured from when the customer enters the line until their bags are in the cart. Hence the time includes both waiting in line and checking out.

Customer Type	Sample Mean	Population Standard Deviation	Sample Size
Standard	5.50 minutes	0.40 minutes	50
U-Scan	5.30 minutes	0.30 minutes	100

EXAMPLE 1 *continued*

Step 1: State the null and alternate hypotheses.

(keyword: “longer than”)

H_0: $\mu_S \leq \mu_U$

H_1: $\mu_S > \mu_U$

Step 2: Select the level of significance.

The .01 significance level is stated in the problem.

Step 3: Determine the appropriate test statistic.

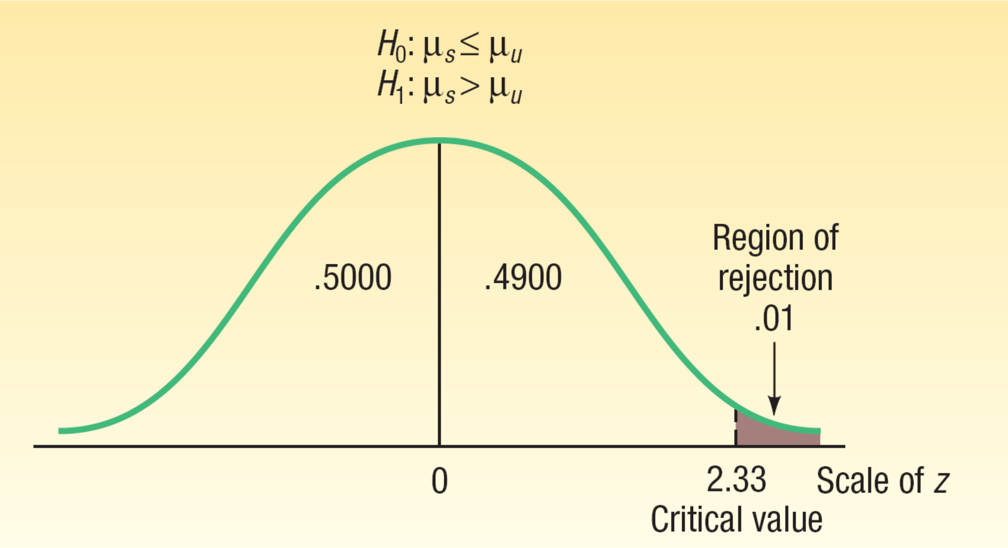
Because both population standard deviations are known, we can use *z-distribution* as the test statistic.

Example 1 *continued*

Step 4: Formulate a decision rule.

Reject H_0 if $Z > Z_\alpha$

$Z > 2.33$



Example 1 *continued*

Step 5: Compute the value of *z* and make a decision

$$z = \frac{\bar{X}_s - \bar{X}_u}{\sqrt{\frac{\sigma_s^2}{n_s} + \frac{\sigma_u^2}{n_u}}} = \frac{5.5 - 5.3}{\sqrt{\frac{0.40^2}{50} + \frac{0.30^2}{100}}} = \frac{0.2}{0.064} = 3.13$$

Customer Type	Sample Mean	Population Standard Deviation	Sample Size
Standard	5.50 minutes	0.40 minutes	50
U-Scan	5.30 minutes	0.30 minutes	100

The computed value of 3.13 is larger than the critical value of 2.33.

Our decision is to reject the null hypothesis. The difference of .20 minutes between the mean checkout time using the standard method is too large to have occurred by chance.

We conclude the U-Scan method is faster.

Two-Sample Tests about Proportions

EXAMPLES

- The vice president of human resources wishes to know whether there is a difference in the proportion of hourly employees who miss more than 5 days of work per year at the Atlanta and the Houston plants.
- General Motors is considering a new design for the Pontiac Grand Am. The design is shown to a group of potential buyers under 30 years of age and another group over 60 years of age. Pontiac wishes to know whether there is a difference in the proportion of the two groups who like the new design.
- A consultant to the airline industry is investigating the fear of flying among adults. Specifically, the company wishes to know whether there is a difference in the proportion of men versus women who are fearful of flying.

Two Sample Tests of Proportions

- We investigate whether two samples came from populations with an equal proportion of successes.

- The two samples are pooled using the following formula.

POOLED PROPORTION

$$p_c = \frac{X_1 + X_2}{n_1 + n_2} \qquad [11\text{–}3]$$

where:

X_1 is the number possessing the trait in the first sample.
X_2 is the number possessing the trait in the second sample.
n_1 is the number of observations in the first sample.
n_2 is the number of observations in the second sample.

Two Sample Tests of Proportions
continued

The value of the test statistic is computed from the following formula.

TWO-SAMPLE TEST OF PROPORTIONS

$$z = \frac{p_1 - p_2}{\sqrt{\frac{p_c(1 - p_c)}{n_1} + \frac{p_c(1 - p_c)}{n_2}}}$$

POOLED PROPORTION

$$p_c = \frac{X_1 + X_2}{n_1 + n_2}$$

Two Sample Tests of Proportions - Example

Manelli Perfume Company recently developed a new fragrance that it plans to market under the name Heavenly. A number of market studies indicate that Heavenly has very good market potential. The Sales Department at Manelli is particularly interested in whether there is a difference in the proportions of younger and older women who would purchase Heavenly if it were marketed. Samples are collected from each of these independent groups. Each sampled woman was asked to smell Heavenly and indicate whether she likes the fragrance well enough to purchase a bottle.

Two Sample Tests of Proportions - Example

Step 1: State the null and alternate hypotheses.

(keyword: “there is a difference”)

H_0: $\pi_1 = \pi_2$

H_1: $\pi_1 \neq \pi_2$

Step 2: Select the level of significance.

The .05 significance level is stated in the problem.

Step 3: Determine the appropriate test statistic.

We will use the z-distribution

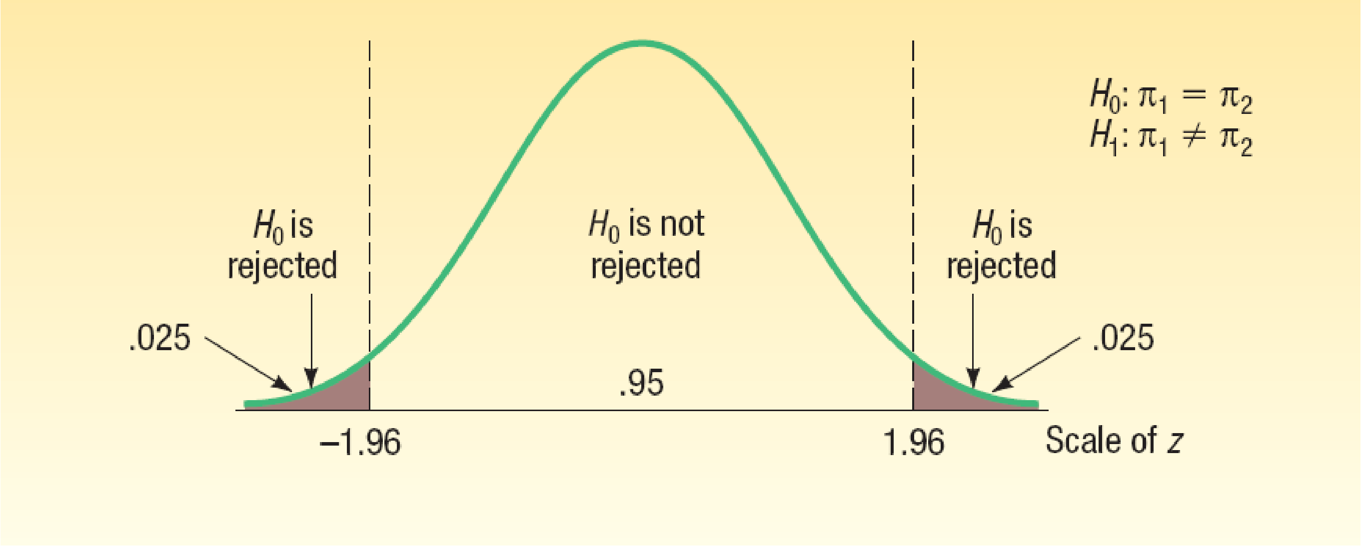
Two Sample Tests of Proportions - Example

Step 4: Formulate the decision rule.

Reject H_0 if $Z > Z_{\alpha/2}$ or $Z < -Z_{\alpha/2}$

$Z > Z_{.05/2}$ or $Z < -Z_{.05/2}$

$Z > 1.96$ or $Z < -1.96$



Two Sample Tests of Proportions - Example

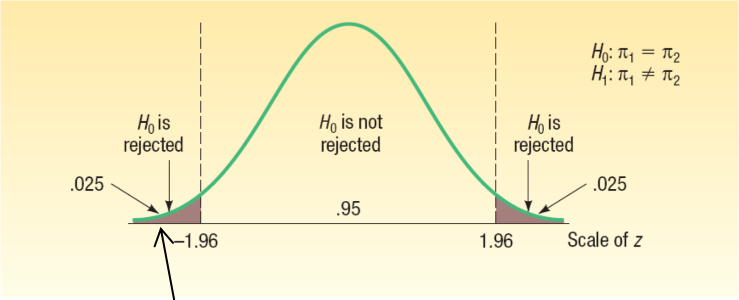
Step 5: Select a sample and make a decision

Let p_1 = young women p_2 = older women

$$p_1 = \frac{X_1}{n_1} = \frac{19}{100} = .19 \qquad p_2 = \frac{X_2}{n_2} = \frac{62}{200} = .31$$

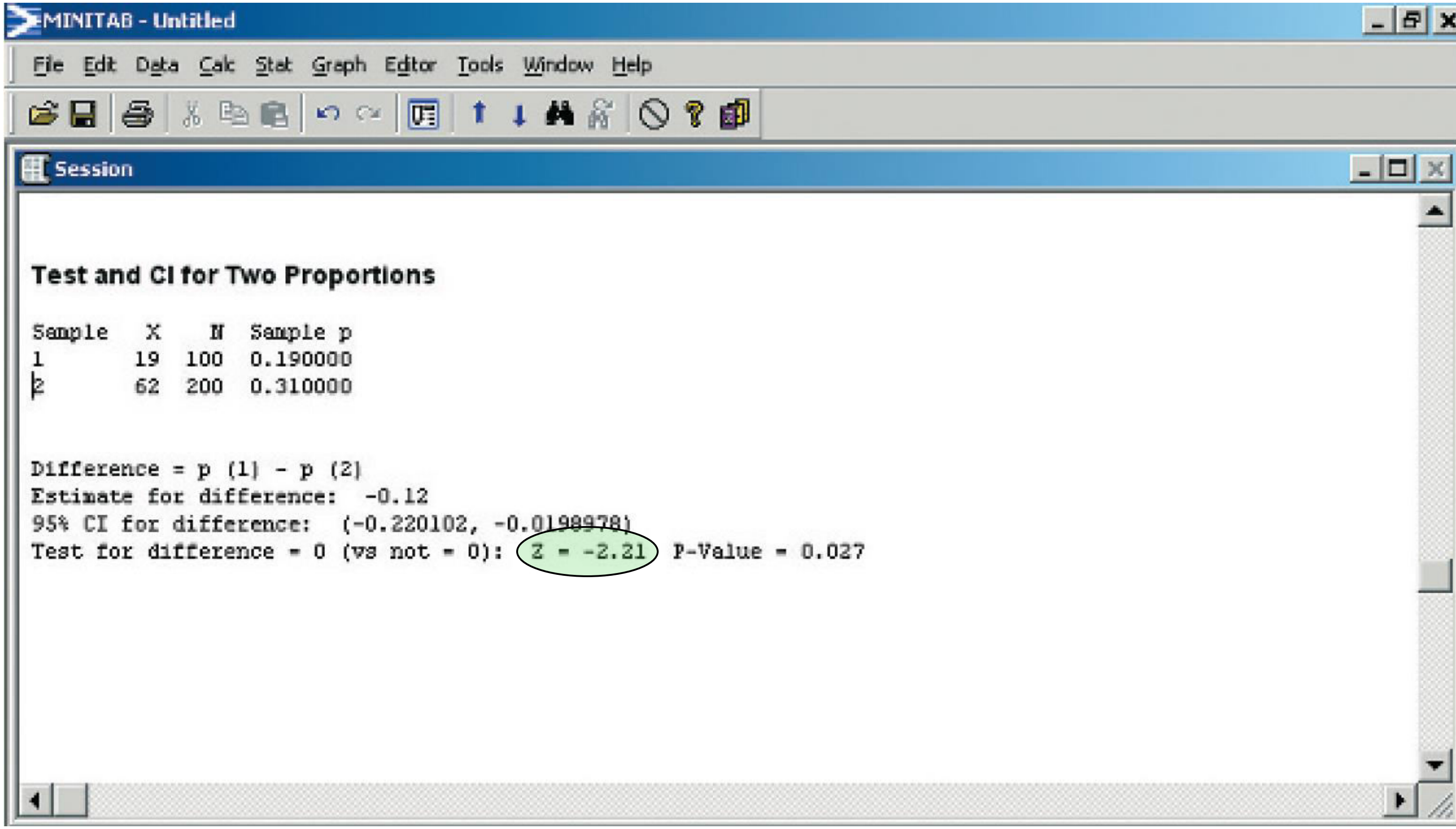
$$p_c = \frac{X_1 + X_2}{n_1 + n_2} = \frac{19 + 62}{100 + 200} = \frac{81}{300} = 0.27$$

$$z = \frac{p_1 - p_2}{\sqrt{\frac{p_c(1-p_c)}{n_1} + \frac{p_c(1-p_c)}{n_2}}} = \frac{.19 - .31}{\sqrt{\frac{.27(1-.27)}{100} + \frac{.27(1-.27)}{200}}} = -2.21$$



The computed value of **-2.21** is in the area of rejection. Therefore, the null hypothesis is rejected at the .05 significance level.
To put it another way, we reject the null hypothesis that the proportion of young women who would purchase Heavenly is equal to the proportion of older women who would purchase Heavenly.

Two Sample Tests of Proportions – Example (Minitab Solution)



Comparing Population Means with Unknown Population Standard Deviations (the Pooled *t*-test)

The *t* distribution is used as the test statistic if one or more of the samples have less than 30 observations. The required assumptions are:

1. Both populations must follow the normal distribution.
2. The populations must have equal standard deviations.
3. The samples are from independent populations.

Small sample test of means *continued*

Finding the value of the test statistic requires two steps.

1. Pool the sample standard deviations.

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

2. Use the pooled standard deviation in the formula.

$$t = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{s_p^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

Comparing Population Means with Unknown Population Standard Deviations (the Pooled *t*-test)

Owens Lawn Care, Inc., manufactures and assembles lawnmowers that are shipped to dealers throughout the United States and Canada. Two different procedures have been proposed for mounting the engine on the frame of the lawnmower. The question is: **Is there a difference in the mean time to mount the engines on the frames of the lawnmowers?**

The first procedure was developed by longtime Owens employee Herb Welles (designated as procedure 1), and the other procedure was developed by Owens Vice President of Engineering William Atkins (designated as procedure 2). To evaluate the two methods, it was decided to conduct a time and motion study.

A sample of five employees was timed using the Welles method and six using the Atkins method. The results, in minutes, are shown on the right.

Is there a difference in the mean mounting times? Use the .10 significance level.

Welles (minutes)	Atkins (minutes)
2	3
4	7
9	5
3	8
2	4
	3

Comparing Population Means with Unknown Population Standard Deviations (the Pooled *t*-test) - Example

Step 1: State the null and alternate hypotheses.

(Keyword: "Is there a *difference*")

H_0: $\mu_1 = \mu_2$

H_1: $\mu_1 \neq \mu_2$

Step 2: State the level of significance. The 0.10 significance level is stated in the problem.

Step 3: Find the appropriate test statistic.

Because the population standard deviations are not known but are assumed to be equal, we use the pooled *t*-test.

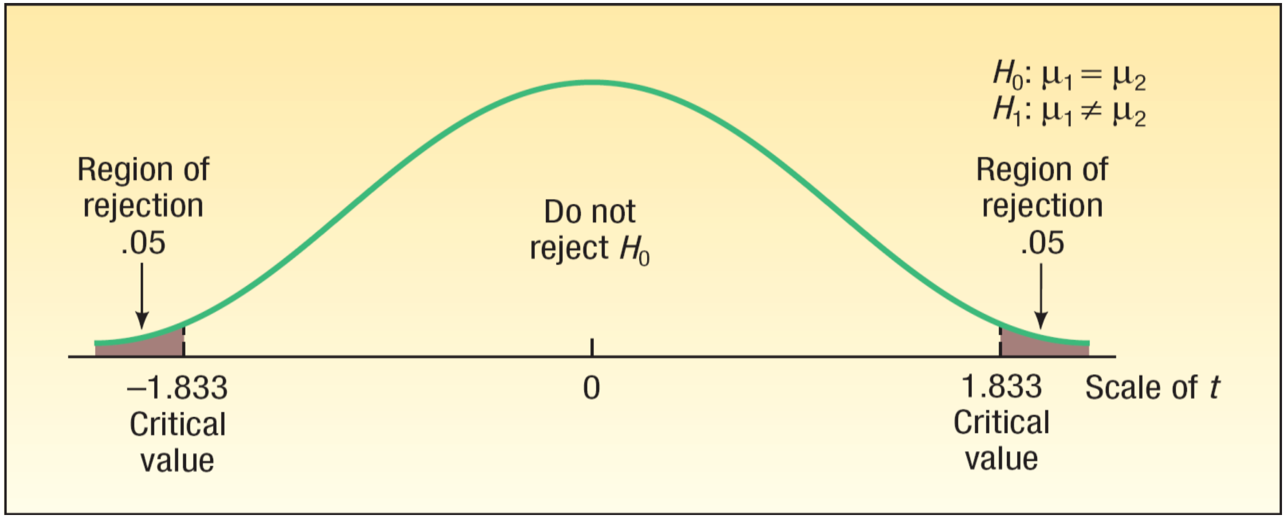
Comparing Population Means with Unknown Population Standard Deviations (the Pooled *t*-test) - Example

Step 4: State the decision rule.

Reject H_0 if $t > t_{\alpha/2, n_1+n_2-2}$ or $t < -t_{\alpha/2, n_1+n_2-2}$

$t > t_{.05,9}$ or $t < -t_{.05,9}$

$t > 1.833$ or $t < -1.833$



Comparing Population Means with Unknown Population Standard Deviations (the Pooled *t*-test) - Example

Step 5: Compute the value of *t* and make a decision

(a) Calculate the sample standard deviations

Welles Method		Atkins Method	
X_1	$(X_1 - \bar{X}_1)^2$	X_2	$(X_2 - \bar{X}_2)^2$
2	$(2-4)^2 = 4$	3	$(3-5)^2 = 4$
4	$(4-4)^2 = 0$	7	$(7-5)^2 = 4$
9	$(9-4)^2 = 25$	5	$(5-5)^2 = 0$
3	$(3-4)^2 = 1$	8	$(8-5)^2 = 9$
2	$(2-4)^2 = 4$	4	$(4-5)^2 = 1$
20	34	3	$(3-5)^2 = 4$
		30	22

$$\bar{X}_1 = \frac{\Sigma X_1}{n_1} = \frac{20}{5} = 4 \qquad \bar{X}_2 = \frac{\Sigma X_2}{n_2} = \frac{30}{6} = 5$$

$$s_1 = \sqrt{\frac{\Sigma(X_1 - \bar{X}_1)^2}{n_1 - 1}} = \sqrt{\frac{34}{5-1}} = 2.9155 \qquad s_2 = \sqrt{\frac{\Sigma(X_2 - \bar{X}_2)^2}{n_2 - 1}} = \sqrt{\frac{22}{6-1}} = 2.0976$$

(b) Calculate the ***pooled*** sample standard deviation

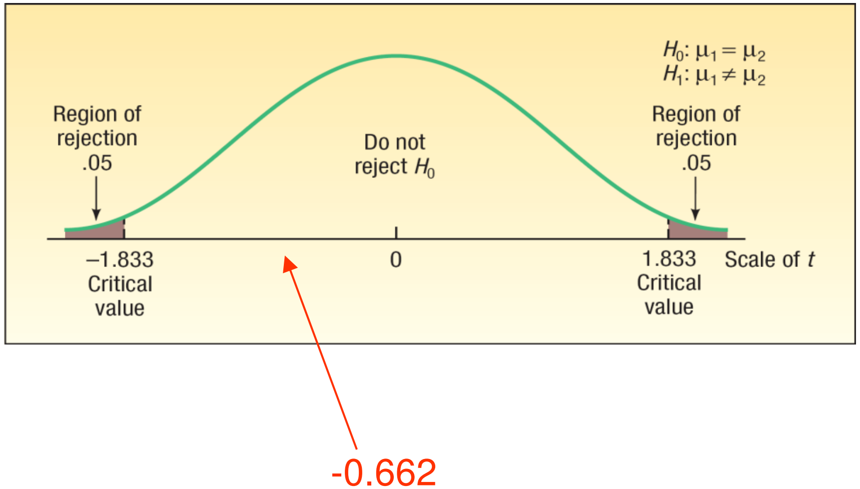
$$s_p^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2} = \frac{(5-1)(2.9155)^2 + (6-1)(2.0976)^2}{5 + 6 - 2} = 6.2222$$

Comparing Population Means with Unknown Population Standard Deviations (the Pooled *t*-test) - Example

Step 5: Compute the value of *t* and make a decision

(c) Determine the value of *t*

$$t = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{s_p^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} = \frac{4.00 - 5.00}{\sqrt{6.2222\left(\frac{1}{5} + \frac{1}{6}\right)}} = -0.662$$



The decision is not to reject the null hypothesis, because -0.662 falls in the region between -1.833 and 1.833.

We conclude that there is no difference in the mean times to mount the engine on the frame using the two methods.

Comparing Population Means with Unknown Population Standard Deviations (the Pooled *t*-test) - Example

welles and atkins

	A	B	C	D	E	F
1	Welles	Atkins		t-Test: Two-Sample Assuming Equal Variances		
2	2	3				
3	4	7			*Welles*	*Atkins*
4	9	5		Mean	4.000	5.000
5	3	8		Variance	8.500	4.400
6	2	4		Observations	5.000	6.000
7		3		Pooled Variance	6.222	
8				Hypothesized Mean Difference	0.000	
9				df	9.000	
10				t Stat	-0.662	
11				P(T<=t) one-tail	0.262	
12				t Critical one-tail	1.833	
13				P(T<=t) two-tail	0.525	
14				t Critical two-tail	2.262	
15						

Comparing Population Means with Unequal Population Standard Deviations

Compute the t-statistic shown on the right if it is not reasonable to assume the population standard deviations are equal.

The sample standard deviations s_1 and s_2 are used in place of the respective population standard deviations.

In addition, the degrees of freedom are adjusted downward by a rather complex approximation formula. The effect is to reduce the number of degrees of freedom in the test, which will require a larger value of the test statistic to reject the null hypothesis.

$$t = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

$$df = \frac{[(s_1^2/n_1) + (s_2^2/n_2)]^2}{\frac{(s_1^2/n_1)^2}{n_1 - 1} + \frac{(s_2^2/n_2)^2}{n_2 - 1}}$$

Comparing Population Means with Unequal Population Standard Deviations - Example

Personnel in a consumer testing laboratory are evaluating the absorbency of paper towels. They wish to compare a set of store brand towels to a similar group of name brand ones. For each brand they dip a ply of the paper into a tub of fluid, allow the paper to drain back into the vat for two minutes, and then evaluate the amount of liquid the paper has taken up from the vat. A random sample of 9 store brand paper towels absorbed the following amounts of liquid in milliliters.

8 8 3 1 9 7 5 5 12

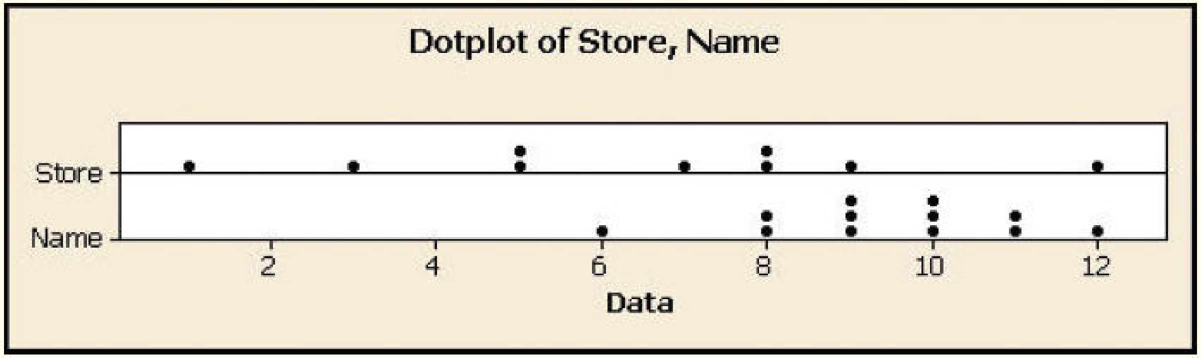
An independent random sample of 12 name brand towels absorbed the following amounts of liquid in milliliters:

12 11 10 6 8 9 9 10 11 9 8 10

Use the .10 significance level and test if there is a difference in the mean amount of liquid absorbed by the two types of paper towels.

Comparing Population Means with Unequal Population Standard Deviations - Example

The following dot plot provided by MINITAB shows the variances to be unequal.



The following output provided by MINITAB shows the descriptive statistics

Descriptive Statistics: Store, Name

Variable	N	Mean	StDev
Store	9	6.44	3.32
Name	12	9.417	1.621

Comparing Population Means with Unequal Population Standard Deviations - Example

Step 1: State the null and alternate hypotheses.

$H_0: \mu_1 = \mu_2$

$H_1: \mu_1 \neq \mu_2$

Step 2: State the level of significance.

The .10 significance level is stated in the problem.

Step 3: Find the appropriate test statistic.

We will use unequal variances *t*-test

Comparing Population Means with Unequal Population Standard Deviations - Example

Step 4: State the decision rule.

Reject H_0 if

$t > t_{\alpha/2\text{d.f.}}$ or $t < -t_{\alpha/2,\text{d.f.}}$

$t > t_{.05,10}$ or $t < -t_{.05,10}$

$t > 1.812$ or $t < -1.812$

$$df = \frac{[(s_1^2/n_1) + (s_2^2/n_2)]^2}{\frac{(s_1^2/n_1)^2}{n_1 - 1} + \frac{(s_2^2/n_2)^2}{n_2 - 1}} = \frac{[(3.32^2/9) + (1.621^2/12)]^2}{\frac{(3.32^2/9)^2}{9-1} + \frac{(1.621^2/12)^2}{12-1}} = \frac{1.4436^2}{.1875 + .0043} = 10.88$$

Descriptive Statistics: Store, Name

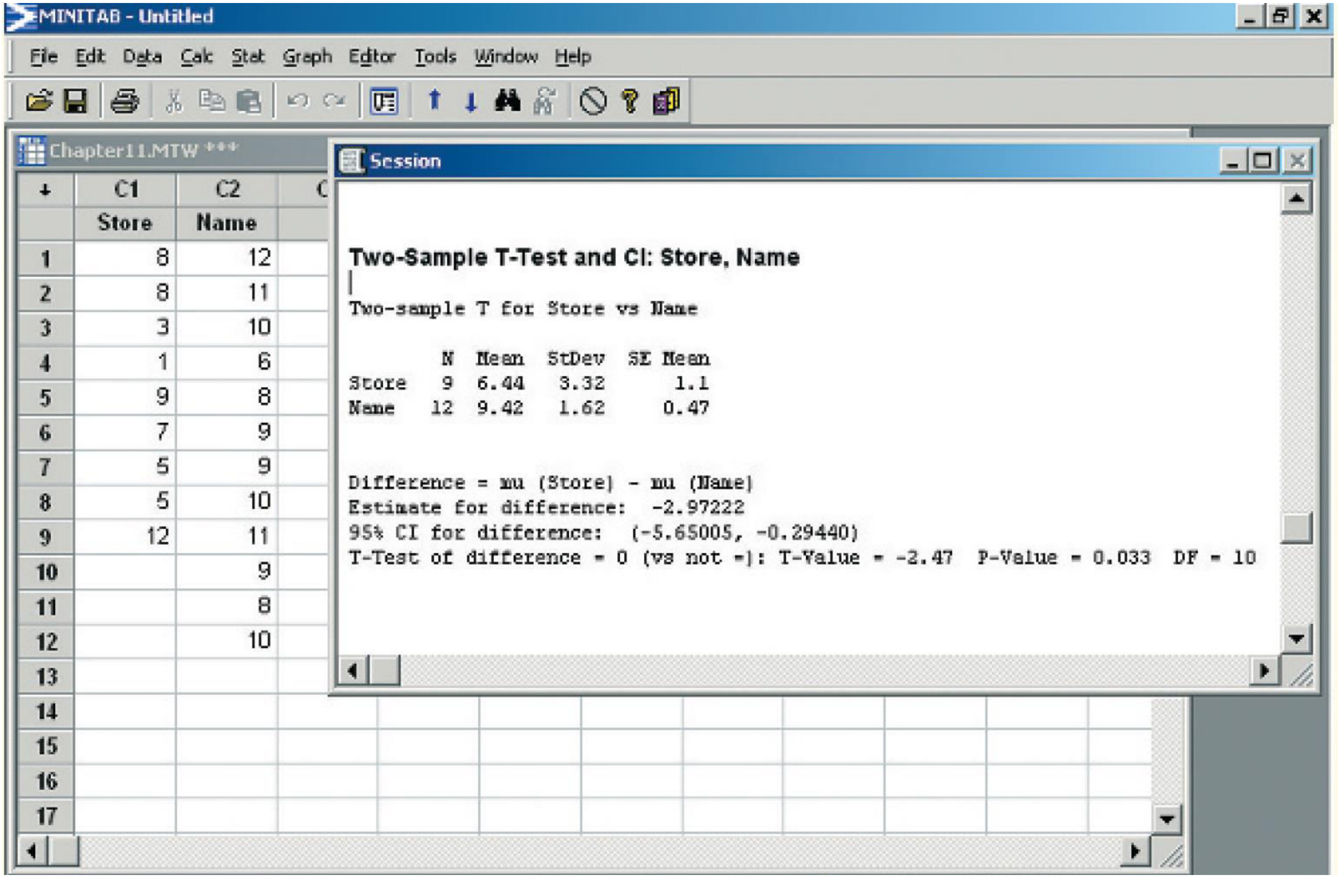
Variable	N	Mean	StDev
Store	9	6.44	3.32
Name	12	9.417	1.621

Step 5: Compute the value of t and make a decision

$$t = \frac{\bar{X}_1 + \bar{X}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} = \frac{6.44 - 9.417}{\sqrt{\frac{3.32^2}{9} + \frac{1.621^2}{12}}} = -2.478$$

The computed value of t is less than the lower critical value, so our decision is to reject the null hypothesis. We conclude that the mean absorption rate for the two towels is not the same.

Minitab


MINITAB - Untitled
File Edit Data Calc Stat Graph Editor Tools Window Help
Chapter11.MTW ***
C1 Store
C2 Name
Store: 8, 8, 3, 1, 9, 7, 5, 5, 12
Name: 12, 11, 10, 6, 8, 9, 9, 10, 11, 9, 8, 10
Session
Two-Sample T-Test and CI: Store, Name
Two-sample T for Store vs Name
N Mean StDev SE Mean
Store 9 6.44 3.32 1.1
Name 12 9.42 1.62 0.47
Difference = mu (Store) - mu (Name)
Estimate for difference: -2.97222
95% CI for difference: (-5.65005, -0.29440)
T-Test of difference = 0 (vs not =): T-Value = -2.47 P-Value = 0.033 DF = 10

Two-Sample Tests of Hypothesis: Dependent Samples

Dependent samples are samples that are paired or related in some fashion.

For example:

- If you wished to buy a car you would look at the *same* car at two (or more) *different* dealerships and compare the prices.
- If you wished to measure the effectiveness of a new diet you would weigh the dieters at the start and at the finish of the program.

Hypothesis Testing Involving Paired Observations

Use the following test when the samples are dependent:

$$t = \frac{\bar{d}}{s_d / \sqrt{n}}$$

Where

$\bar{d}$ is the mean of the differences

s_d is the standard deviation of the differences

n is the number of pairs (differences)

Hypothesis Testing Involving Paired Observations - Example

Nickel Savings and Loan wishes to compare the two companies it uses to appraise the value of residential homes. Nickel Savings selected a sample of 10 residential properties and scheduled both firms for an appraisal. The results, reported in $000, are shown on the table (right).

At the .05 significance level, can we conclude there is a difference in the mean appraised values of the homes?

Home	Schadek	Bowyer
1	235	228
2	210	205
3	231	219
4	242	240
5	205	198
6	230	223
7	231	227
8	210	215
9	225	222
10	249	245

Hypothesis Testing Involving Paired Observations - Example

Step 1: State the null and alternate hypotheses.

$H_0: \mu_d = 0$

$H_1: \mu_d \neq 0$

Step 2: State the level of significance.

The .05 significance level is stated in the problem.

Step 3: Find the appropriate test statistic.

We will use the *t*-test

Hypothesis Testing Involving Paired Observations - Example

Step 4: State the decision rule.

Reject H_0 if

$t > t_{\alpha/2,\, n-1}$ or $t < -t_{\alpha/2,n-1}$

$t > t_{.025,9}$ or $t < -t_{.025,\,9}$

$t > 2.262$ or $t < -2.262$

TABLE 11–2 A Portion of the *t* Distribution from Appendix B.2

	Confidence Intervals					
	80%	90%	95%	98%	99%	99.9%
	Level of Significance for One-Tailed Test					
df	0.100	0.050	0.025	0.010	0.005	0.0005
	Level of Significance for Two-Tailed Test					
	0.20	0.10	0.05	0.02	0.01	0.001
1	3.078	6.314	12.706	31.821	63.657	636.619
2	1.886	2.920	4.303	6.965	9.925	31.599
3	1.638	2.353	3.182	4.541	5.841	12.924
4	1.533	2.132	2.776	3.747	4.604	8.610
5	1.476	2.015	2.571	3.365	4.032	6.869
6	1.440	1.943	2.447	3.143	3.707	5.959
7	1.415	1.895	2.365	2.998	3.499	5.408
8	1.397	1.860	2.306	2.896	3.355	5.041
9	1.383	1.833	2.262	2.821	3.250	4.781
10	1.372	1.812	2.228	2.764	3.169	4.587

Hypothesis Testing Involving Paired Observations - Example

Step 5: Compute the value of *t* and make a decision

Home	Schadek	Bowyer	Difference, d	$(d - \bar{d})$	$(d - \bar{d})^2$
1	235	228	7	2.4	5.76
2	210	205	5	0.4	0.16
3	231	219	12	7.4	54.76
4	242	240	2	−2.6	6.76
5	205	198	7	2.4	5.76
6	230	223	7	2.4	5.76
7	231	227	4	−0.6	0.36
8	210	215	−5	−9.6	92.16
9	225	222	3	−1.6	2.56
10	249	245	4	−0.6	0.36
			46	0	174.40

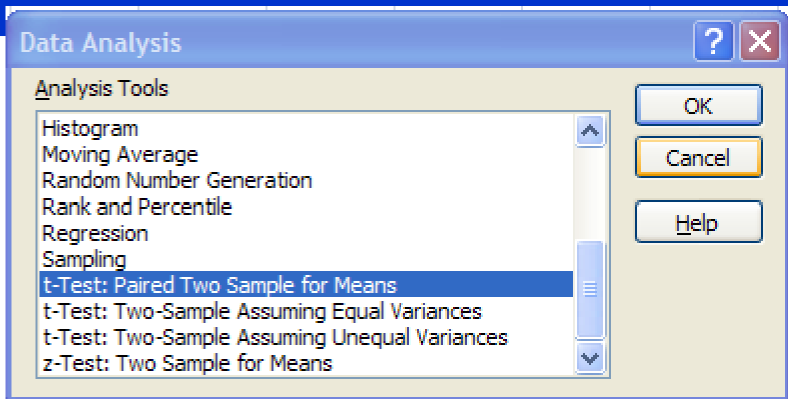
$$\bar{d} = \frac{\Sigma d}{n} = \frac{46}{10} = 4.60$$

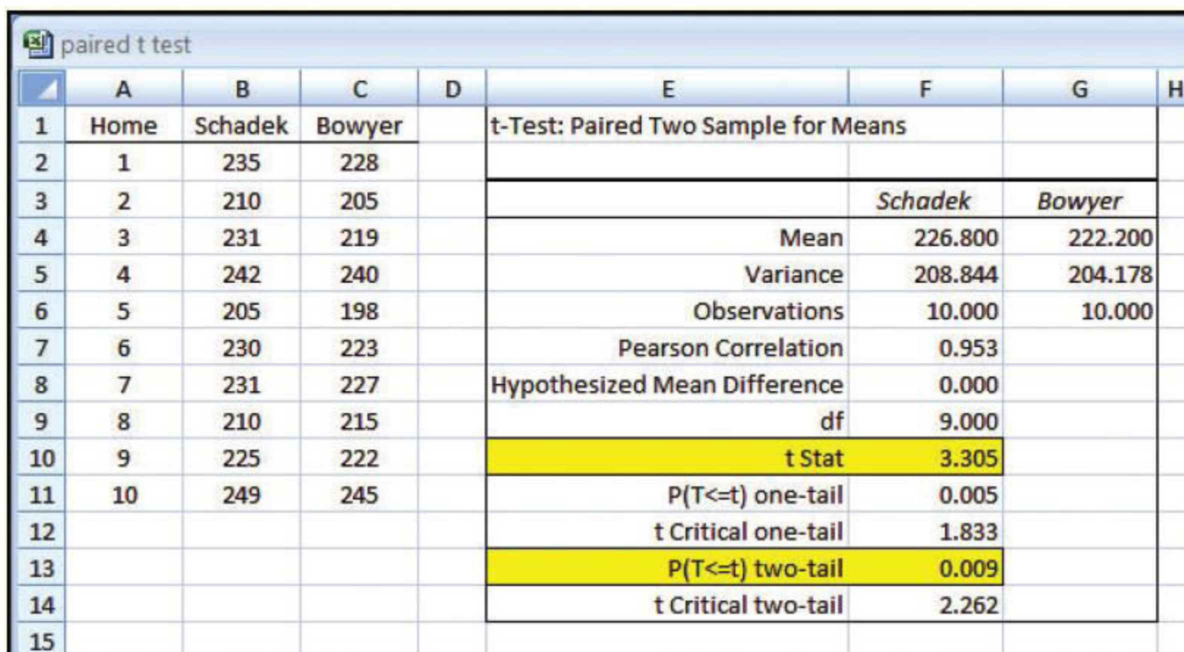
$$s_d = \sqrt{\frac{\Sigma(d - \bar{d})^2}{n - 1}} = \sqrt{\frac{174.4}{10 - 1}} = 4.402$$

$$t = \frac{\bar{d}}{s_d/\sqrt{n}} = \frac{4.6}{4.402/\sqrt{10}} = 3.305$$

The computed value of *t* is greater than the higher critical value, so our decision is to reject the null hypothesis. We conclude that there is a difference in the mean appraised values of the homes.

Hypothesis Testing Involving Paired Observations – Excel Example




paired t test

Home	Schadek	Bowyer
1	235	228
2	210	205
3	231	219
4	242	240
5	205	198
6	230	223
7	231	227
8	210	215
9	225	222
10	249	245

t-Test: Paired Two Sample for Means

	Schadek	*Bowyer*
Mean	226.800	222.200
Variance	208.844	204.178
Observations	10.000	10.000
Pearson Correlation	0.953	
Hypothesized Mean Difference	0.000	
df	9.000	
t Stat	3.305	
P(T<=t) one-tail	0.005	
t Critical one-tail	1.833	
P(T<=t) two-tail	0.009	
t Critical two-tail	2.262	

Dependent versus Independent Samples

How do we tell between dependent and independent samples?

1. Dependent sample is characterized by a measurement followed by an intervention of some kind and then another measurement. This could be called a "before" and "after" study.
2. Dependent sample is characterized by matching or pairing observation.

Why do we prefer dependent samples to independent samples?

By using dependent samples, we are able to reduce the variation in the sampling distribution.