Finding P-values for the F test statistic is somewhat more complicated since it requires looking through all the F tables (Table H in Appendix C) using the specific d.f.N. and d.f.D. values. For example, suppose that a certain test has $F = 3.58$, d.f.N. $= 5$, and d.f.D. $= 10$. To find the P-value interval for $F = 3.58$, you must first find the corresponding F values for d.f.N. $= 5$ and d.f.D. $= 10$ for α equal to 0.005, 0.01, 0.025, 0.05, and 0.10 in Table H. Then make a table as shown.

α	0.10	0.05	0.025	0.01	0.005
F	2.52	3.33	4.24	5.64	6.87

Now locate the two F values that the test value 3.58 falls between. In this case, 3.58 falls between 3.33 and 4.24, corresponding to 0.05 and 0.025. Hence, the P-value for a right-tailed test for $F = 3.58$ falls between 0.025 and 0.05 (that is, $0.025 < P\text{-value} < 0.05$). For a right-tailed test, then, you would reject the null hypothesis at $\alpha = 0.05$ but not at $\alpha = 0.01$. The P-value obtained from a calculator is 0.0408. Remember that for a two-tailed test the values found in Table H for α must be doubled. In this case, $0.05 < P\text{-value} < 0.10$ for $F = 3.58$.

Once you understand the concept, you can dispense with making a table as shown and find the P-value directly from Table H.

Example 9–16 Airport Passengers

The CEO of an airport hypothesizes that the variance in the number of passengers for American airports is greater than the variance in the number of passengers for foreign airports. At $\alpha = 0.10$, is there enough evidence to support the hypothesis? The data in millions of passengers per year are shown for selected airports. Use the P-value method. Assume the variable is normally distributed.

American airports		Foreign airports	
36.8	73.5	60.7	51.2
72.4	61.2	42.7	38.6
60.5	40.1		

Source: Airports Council International.

Solution

Step 1 State the hypotheses and identify the claim.

H_0: $\sigma_1^2 = \sigma_2^2$ and H_1: $\sigma_1^2 > \sigma_2^2$ (claim)

Step 2 Compute the test value. Using the formula in Chapter 3 or a calculator, find the variance for each group.

$s_1^2 = 246.38$ and $s_2^2 = 95.87$

Substitute in the formula and solve.

$$F = \frac{s_1^2}{s_2^2} = \frac{246.38}{95.87} = 2.57$$

Step 3 Find the P-value in Table H, using d.f.N. $= 5$ and d.f.D. $= 3$.

α	0.10	0.05	0.025	0.01	0.005
F	5.31	9.01	14.88	28.24	45.39

Since 2.57 is less than 5.31, the P-value is greater than 0.10. (The P-value obtained from a calculator is 0.234.)

Step 4 Make the decision. The decision is to not reject the null hypothesis since P-value > 0.10.

Step 5 Summarize the results. There is not enough evidence to support the claim that the variance in the number of passengers for American airports is greater than the variance in the number of passengers for foreign airports.

If the exact degrees of freedom are not specified in Table H, the closest smaller value should be used. For example, if $\alpha = 0.05$ (right-tailed test), d.f.N. = 18, and d.f.D. = 20, use the column d.f.N. = 15 and the row d.f.D. = 20 to get $F = 2.20$.

Note: It is not absolutely necessary to place the larger variance in the numerator when you are performing the F test. Critical values for left-tailed hypotheses tests can be found by interchanging the degrees of freedom and taking the reciprocal of the value found in Table H.

Also, you should use caution when performing the F test since the data can run contrary to the hypotheses on rare occasions. For example, if the hypotheses are H_0: $\sigma_1^2 \le \sigma_2^2$ (written H_0: $\sigma_1^2 = \sigma_2^2$) and H_1: $\sigma_1^2 > \sigma_2^2$, but if $s_1^2 < s_2^2$, then the F test should not be performed and you would not reject the null hypothesis.

Applying the Concepts 9–5

Variability and Automatic Transmissions

Assume the following data values are from the June 1996 issue of *Automotive Magazine*. An article compared various parameters of U.S.- and Japanese-made sports cars. This report centers on the price of an optional automatic transmission. Which country has the greater variability in the price of automatic transmissions? Input the data and answer the following questions.

Japanese cars		U.S. cars	
Nissan 300ZX	\$1940	Dodge Stealth	\$2363
Mazda RX7	1810	Saturn	1230
Mazda MX6	1871	Mercury Cougar	1332
Nissan NX	1822	Ford Probe	932
Mazda Miata	1920	Eagle Talon	1790
Honda Prelude	1730	Chevy Lumina	1833

1. What is the null hypothesis?
2. What test statistic is used to test for any significant differences in the variances?
3. Is there a significant difference in the variability in the prices between the Japanese cars and the U.S. cars?
4. What effect does a small sample size have on the standard deviations?
5. What degrees of freedom are used for the statistical test?
6. Could two sets of data have significantly different variances without having significantly different means?

See page 531 for the answers.

Exercises 9–5

1. When one is computing the F test value, what condition is placed on the variance that is in the numerator?
The variance in the numerator should be the larger of the two variances.

2. Why is the critical region always on the right side in the use of the F test? The larger variance is placed in the numerator of the formula; hence, $F \ge 1$.

3. What are the two different degrees of freedom associated with the F distribution?

4. What are the characteristics of the F distribution?

5. Using Table H, find the critical value for each.

a. Sample 1: $s_1^2 = 128$, $n_1 = 23$
Sample 2: $s_2^2 = 162$, $n_2 = 16$
Two-tailed, $\alpha = 0.01$

b. Sample 1: $s_1^2 = 37$, $n_1 = 14$
Sample 2: $s_2^2 = 89$, $n_2 = 25$
Right-tailed, $\alpha = 0.01$

c. Sample 1: $s_1^2 = 232$, $n_1 = 30$
Sample 2: $s_2^2 = 387$, $n_2 = 46$
Two-tailed, $\alpha = 0.05$

d. Sample 1: $s_1^2 = 164$, $n_1 = 21$
Sample 2: $s_2^2 = 53$, $n_2 = 17$
Two-tailed, $\alpha = 0.10$

e. Sample 1: $s_1^2 = 92.8$, $n_1 = 11$
Sample 2: $s_2^2 = 43.6$, $n_2 = 11$
Right-tailed, $\alpha = 0.05$

6. (ans) Using Table H, find the *P*-value interval for each *F* test value.

a. $F = 2.97$, d.f.N. = 9, d.f.D. = 14, right-tailed
b. $F = 3.32$, d.f.N. = 6, d.f.D. = 12, two-tailed
c. $F = 2.28$, d.f.N. = 12, d.f.D. = 20, right-tailed
d. $F = 3.51$, d.f.N. = 12, d.f.D. = 21, right-tailed
e. $F = 4.07$, d.f.N. = 6, d.f.D. = 10, two-tailed
f. $F = 1.65$, d.f.N. = 19, d.f.D. = 28, right-tailed
g. $F = 1.77$, d.f.N. = 28, d.f.D. = 28, right-tailed
h. $F = 7.29$, d.f.N. = 5, d.f.D. = 8, two-tailed

For Exercises 7 through 20, perform the following steps. Assume that all variables are normally distributed.

a. State the hypotheses and identify the claim.
b. Find the critical value.
c. Compute the test value.
d. Make the decision.
e. Summarize the results.

Use the traditional method of hypothesis testing unless otherwise specified.

7. Ages of Hospital Patients The average age of hospital inpatients has gradually increased to 52.5 years. Studies of two major health care systems found the following information. At the 0.05 level of significance is there sufficient evidence to conclude a difference between the two variances?

	System 1	System 2
Sample size	60	60
Sample mean	49.8	50.2
Sample standard deviation	5.4	7.6

Source: *New York Times Almanac.*

8. Museum Attendance A metropolitan children's museum open year-round wants to see if the variance in daily attendance differs between the summer and winter months. Random samples of 30 days each were selected and showed that in the winter months, the sample mean daily attendance was 300 with a standard deviation of 52, and the sample mean daily attendance for the summer months was 280 with a standard deviation of 65. At $\alpha = 0.05$ can we conclude a difference in variances?

9. Wolf Pack Pups Does the variance in average number of pups per pack differ between Montana and Idaho wolf packs? Random samples of packs were selected for each area, and the numbers of pups per pack were recorded. At the 0.05 level of significance, can a difference in variances be concluded?

Montana	4	3	5	6	1	2	8	2
wolf packs	3	1	7	6	5			
Idaho	2	4	5	4	2	4	6	3
wolf packs	1	4	2	1				

Source: www.fws.gov

10. Noise Levels in Hospitals In a hospital study, it was found that the standard deviation of the sound levels from 20 areas designated as "casualty doors" was 4.1 dBA and the standard deviation of 24 areas designated as operating theaters was 7.5 dBA. At $\alpha = 0.05$, can you substantiate the claim that there is a difference in the standard deviations?

Source: M. Bayo, A. Garcia, and A. Garcia, "Noise Levels in an Urban Hospital and Workers' Subjective Responses," *Archives of Environmental Health.*

11. Calories in Ice Cream The numbers of calories contained in $\frac{1}{2}$-cup servings of randomly selected flavors of ice cream from two national brands are listed here. At the 0.05 level of significance, is there sufficient evidence to conclude that the variance in the number of calories differs between the two brands?

Brand A		Brand B	
330	300	280	310
310	350	300	370
270	380	250	300
310	300	290	310

Source: *The Doctor's Pocket Calorie, Fat and Carbohydrate Counter.*

12. Winter Temperatures A random sample of daily high temperatures in January and February is listed below. At $\alpha = 0.05$ can it be concluded that there is a difference in variances in high temperature between the two months?

Jan.	31	31	38	24	24	42	22	43	35	42
Feb.	31	29	24	30	28	24	27	34	27	

13. Population and Area Cities were randomly selected from the list of the 50 largest cities in the United States (based on population). The areas of each in square miles are indicated below. Is there sufficient evidence to conclude that the variance in area is greater for eastern cities than for western cities at $\alpha = 0.05$? At $\alpha = 0.01$?

Eastern		Western	
Atlanta, GA	132	Albuquerque, NM	181
Columbus, OH	210	Denver, CO	155
Louisville, KY	385	Fresno, CA	104
New York, NY	303	Las Vegas, NV	113
Philadelphia, PA	135	Portland, OR	134
Washington, DC	61	Seattle, WA	84
Charlotte, NC	242		

Source: *New York Times Almanac.*

14. Carbohydrates in Candy The number of grams of carbohydrates contained in 1-ounce servings of randomly selected chocolate and nonchocolate candy is listed here. Is there sufficient evidence to conclude that there is a difference between the variation in carbohydrate content for chocolate and nonchocolate candy? Use $\alpha = 0.10$.

Chocolate	29	25	17	36	41	25	32	29
	38	34	24	27	29			
Nonchocolate	41	41	37	29	30	38	39	10
	29	55	29					

Source: *The Doctor's Pocket Calorie, Fat and Carbohydrate Counter.*

15. Tuition Costs for Medical School The yearly tuition costs in dollars for random samples of medical schools that specialize in research and in primary care are listed. At $\alpha = 0.05$, can it be concluded that a difference between the variances of the two groups exists?

Research			Primary care		
30,897	34,280	31,943	26,068	21,044	30,897
34,294	31,275	29,590	34,208	20,877	29,691
20,618	20,500	29,310	33,783	33,065	35,000
21,274			27,297		

Source: *U.S. News & World Report Best Graduate Schools.*

16. County Size in Indiana and Iowa A researcher wishes to see if the variance of the areas in square miles for counties in Indiana is less than the variance of the areas for counties in Iowa. A random sample of counties is selected, and the data are shown. At $\alpha = 0.01$, can it be concluded that the variance of the areas for counties in Indiana is less than the variance of the areas for counties in Iowa?

Indiana				Iowa			
406	393	396	485	640	580	431	416
431	430	369	408	443	569	779	381
305	215	489	293	717	568	714	731
373	148	306	509	571	577	503	501
560	384	320	407	568	434	615	402

Source: *The World Almanac and Book of Facts.*

17. Heights of Tall Buildings Test the claim that the variance of heights of tall buildings in Denver is equal to the variance in heights of tall buildings in Detroit at $\alpha = 0.10$. The data are given in feet.

Denver			Detroit		
714	698	544	620	472	430
504	438	408	562	448	420
404			534	436	

Source: *The World Almanac and Book of Facts.*

18. Elementary School Teachers' Salaries A researcher claims that the variation in the salaries of elementary school teachers is greater than the variation in the salaries of secondary school teachers. A sample of the salaries of 30 elementary school teachers has a variance of \$8324, and a sample of the salaries of 30 secondary school teachers has a variance of \$2862. At $\alpha = 0.05$, can the researcher conclude that the variation in the elementary school teachers' salaries is greater than the variation in the secondary school teachers' salaries? Use the *P*-value method.

19. Weights of Running Shoes The weights in ounces of a sample of running shoes for men and women are shown. Calculate the variances for each sample, and test the claim that the variances are equal at $\alpha = 0.05$. Use the *P*-value method.

Men			Women		
11.9	10.4	12.6	10.6	10.2	8.8
12.3	11.1	14.7	9.6	9.5	9.5
9.2	10.8	12.9	10.1	11.2	9.3
11.2	11.7	13.3	9.4	10.3	9.5
13.8	12.8	14.5	9.8	10.3	11.0

20. Daily Stock Prices Two portfolios were randomly assembled from the New York Stock Exchange, and the daily stock prices are shown below. At the 0.05 level of significance, can it be concluded that a difference in variance in price exists between the two portfolios?

Portfolio A	36.44	44.21	12.21	59.60	55.44	39.42	51.29	48.68	41.59	19.49
Portfolio B	32.69	47.25	49.35	36.17	63.04	17.74	4.23	34.98	37.02	31.48

Source: *Washington Observer-Reporter.*

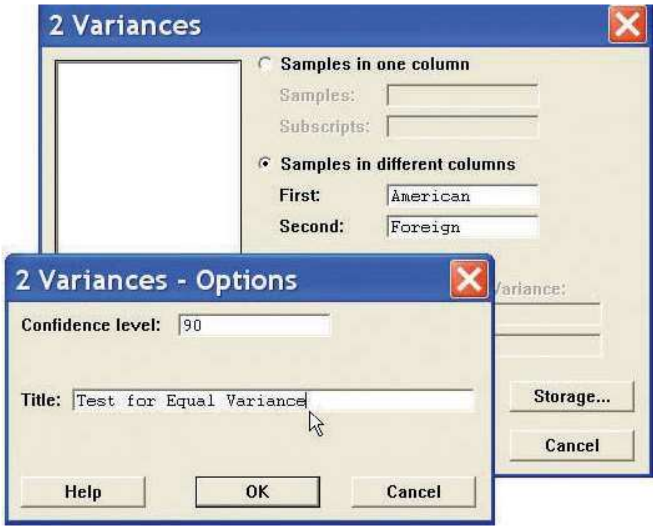
Technology *Step by Step*

MINITAB
Step by Step

Test for the Difference Between Two Variances

For Example 9–16, test the hypothesis that the variance in the number of passengers for American and foreign airports is different. Use the *P*-value approach.

American airports	Foreign airports
36.8	60.7
72.4	42.7
60.5	51.2
73.5	38.6
61.2	
40.1	



1. Enter the data into two columns of MINITAB.
2. Name the columns American and Foreign.
   a) Select **Stat>Basic Statistics>2-Variances.**
   b) Click the button for Samples in different columns.
   c) Click in the text box for First, then double-click C1 American.
   d) Double-click C2 Foreign, then click on [Options]. The dialog box is shown. Change the confidence level to **90** and type an appropriate title. In this dialog, we cannot specify a left- or right-tailed test.
3. Click [OK] twice. A graph window will open that includes a small window that says $F = 2.57$ and the *P*-value is 0.437. Divide this two-tailed *P*-value by 2 for a one-tailed test.

There is not enough evidence in the sample to conclude there is greater variance in the number of passengers in American airports compared to foreign airports.

TI-83 Plus or TI-84 Plus
Step by Step

Hypothesis Test for the Difference Between Two Variances (Data)

1. Enter the data values into L_1 and L_2.
2. Press **STAT** and move the cursor to TESTS.

3. Press **D (ALPHA X^{-1})** for 2-SampFTest. (The TI-84 uses E)
4. Move the cursor to Data and press **ENTER.**
5. Type in the appropriate values.
6. Move the cursor to the appropriate alternative hypothesis and press **ENTER.**
7. Move the cursor to Calculate and press **ENTER.**

Hypothesis Test for the Difference Between Two Variances (Statistics)

1. Press **STAT** and move the cursor to TESTS.
2. Press **D (ALPHA X^{-1})** for 2-SampFTest. (The TI-84 uses E)
3. Move the cursor to Stats and press **ENTER.**
4. Type in the appropriate values.
5. Move the cursor to the appropriate alternative hypothesis and press **ENTER.**
6. Move the cursor to Calculate and press **ENTER.**

Excel
Step by Step

F Test for the Difference Between Two Variances

Excel has a two-sample *F* test included in the Data Analysis Add-in. To perform an *F* test for the difference between the variances of two populations, given two independent samples, do this:

1. Enter the first sample data set into column A.
2. Enter the second sample data set into column B.
3. Select the Data tab from the toolbar. Then select Data Analysis.
4. In the Analysis Tools box, select F-test: Two-sample for Variances.
5. Type the ranges for the data in columns A and B.
6. Specify the confidence level Alpha.
7. Specify a location for the output, and click [OK].

Example XL9–4

At $\alpha = 0.05$, test the hypothesis that the two population variances are equal, using the sample data provided here.

Set A	63	73	80	60	86	83	70	72	82
Set B	86	93	64	82	81	75	88	63	63

The results appear in the table that Excel generates, shown here. For this example, the output shows that the null hypothesis cannot be rejected at an α level of 0.05.

F-Test Two-Sample for Variances		
	Variable 1	*Variable 2*
Mean	74.33333333	77.22222222
Variance	82.75	132.9444444
Observations	9	9
df	8	8
F	0.622440451	
P(F<=f) one-tail	0.258814151	
F Critical one-tail	0.290858004	

Summary

Many times researchers are interested in comparing two parameters such as two means, two proportions, or two variances. These measures are obtained from two samples, then compared using a z test, t test, or an F test.

- If two sample means are compared, when the samples are independent and the population standard deviations are known, a z test is used. If the sample sizes are less than 30, the populations should be normally distributed. (9–1)
- If two means are compared when the samples are independent and the sample standard deviations are used, then a t test is used. Both variances are assumed to be unequal. (9–2)
- When the two samples are dependent or related, such as using the same subjects and comparing the means of before and after tests, then the t test for dependent samples is used. (9–3)
- Two proportions can be compared by using the z test for proportions. In this case, each of n_1p_1, n_1q_1, n_2p_2, and n_2q_2 must all be 5 or more. (9–4)
- Two variances can be compared by using an F test. The critical values for the F test are obtained from the F distribution. (9–5)
- Confidence intervals for differences between two parameters can also be found.

Important Terms

Important Formulas

Formula for the z test for comparing two means from independent populations; σ_1 and σ_2 are known:

$$z = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}}$$

Formula for the confidence interval for difference of two means when σ_1 and σ_2 are known:

$$(\bar{X}_1 - \bar{X}_2) - z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} < \mu_1 - \mu_2 < (\bar{X}_1 - \bar{X}_2) + z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$$

Formula for the t test for comparing two means (independent samples, variances not equal), σ_1 and σ_2 unknown:

$$t = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\dfrac{s_1^2}{n_1} + \dfrac{s_2^2}{n_2}}}$$

and d.f. = the smaller of $n_1 - 1$ or $n_2 - 1$.

Formula for the confidence interval for the difference of two means (independent samples, variances unequal), σ_1 and σ_2 unknown:

$$(\bar{X}_1 - \bar{X}_2) - t_{\alpha/2}\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} < \mu_1 - \mu_2 < (\bar{X}_1 - \bar{X}_2) + t_{\alpha/2}\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$$

and d.f. = smaller of $n_1 - 1$ and $n_2 - 2$.

Formula for the t test for comparing two means from dependent samples:

$$t = \frac{\bar{D} - \mu_D}{s_D/\sqrt{n}}$$

where $\bar{D}$ is the mean of the differences

$$\bar{D} = \frac{\Sigma D}{n}$$

and s_D is the standard deviation of the differences

$$s_D = \sqrt{\frac{n\Sigma D^2 - (\Sigma D)^2}{n(n-1)}}$$

Formula for confidence interval for the mean of the difference for dependent samples:

$$\bar{D} - t_{\alpha/2}\frac{s_D}{\sqrt{n}} < \mu_D < \bar{D} + t_{\alpha/2}\frac{s_D}{\sqrt{n}}$$

and d.f. = $n - 1$.

Formula for the z test for comparing two proportions:

$$z = \frac{(\hat{p}_1 - \hat{p}_2) - (p_1 - p_2)}{\sqrt{\bar{p}\,\bar{q}\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

where

$$\bar{p} = \frac{X_1 + X_2}{n_1 + n_2} \qquad \hat{p}_1 = \frac{X_1}{n_1}$$

$$\bar{q} = 1 - \bar{p} \qquad \hat{p}_2 = \frac{X_2}{n_2}$$

Formula for confidence interval for the difference of two proportions:

$$(\hat{p}_1 - \hat{p}_2) - z_{\alpha/2}\sqrt{\frac{\hat{p}_1\hat{q}_1}{n_1} + \frac{\hat{p}_2\hat{q}_2}{n_2}} < p_1 - p_2$$

$$< (\hat{p}_1 - \hat{p}_2) + z_{\alpha/2}\sqrt{\frac{\hat{p}_1\hat{q}_1}{n_1} + \frac{\hat{p}_2\hat{q}_2}{n_2}}$$

Formula for the F test for comparing two variances:

$$F = \frac{s_1^2}{s_2^2} \qquad \text{d.f.N.} = n_1 - 1 \qquad \text{d.f.D.} = n_2 - 1$$

The larger variance is placed in the numerator.

Review Exercises

For each exercise, perform these steps. Assume that all variables are normally or approximately normally distributed.

a. State the hypotheses and identify the claim.
b. Find the critical value(s).
c. Compute the test value.
d. Make the decision.
e. Summarize the results.

Use the traditional method of hypothesis testing unless otherwise specified.

1. Driving for Pleasure Two groups of drivers are surveyed to see how many miles per week they drive for pleasure trips. The data are shown. At $\alpha = 0.01$, can it be concluded that single drivers do more driving for pleasure trips on average than married drivers? Assume $\sigma_1 = 16.7$ and $\sigma_2 = 16.1$. (9–1)

Single drivers					Married drivers				
106	110	115	121	132	97	104	138	102	115
119	97	118	122	135	133	120	119	136	96
110	117	116	138	142	139	108	117	145	114
115	114	103	98	99	140	136	113	113	150
108	117	152	147	117	101	114	116	113	135
154	86	115	116	104	115	109	147	106	88
107	133	138	142	140	113	119	99	108	105

2. Average Earnings of College Graduates The average yearly earnings of male college graduates (with at least a bachelor's degree) are \$58,500 for men aged 25 to 34. The average yearly earnings of female college graduates with the same qualifications are \$49,339. Based on the results below, can it be concluded that there is a difference in mean earnings between male and female college graduates? Use the 0.01 level of significance. (9–1)

	Male	Female
Sample mean	\$59,235	\$52,487
Population standard deviation	8,945	10,125
Sample size	40	35

Source: *New York Times Almanac.*

3. Communication Times According to the Bureau of Labor Statistics' American Time Use Survey (ATUS), married persons spend an average of 8 minutes per day on phone calls, mail, and e-mail, while single persons spend an average of 14 minutes per day on these same tasks. Based on the following information, is there sufficient evidence to conclude that single persons spend, on average, a greater time each day communicating? Use the 0.05 level of significance. (9–2)

	Single	Married
Sample size	26	20
Sample mean	16.7 minutes	12.5 minutes
Sample variance	8.41	10.24

Source: *Time* magazine.

4. Average Temperatures The average temperatures for a 25-day period for Birmingham, Alabama, and Chicago, Illinois, are shown. Based on the samples, at $\alpha = 0.10$, can it be concluded that it is warmer in Birmingham? (9–2)

Birmingham					Chicago				
78	82	68	67	68	70	74	73	60	77
75	73	75	64	68	71	72	71	74	76
62	73	77	78	79	71	80	65	70	83
74	72	73	78	68	67	76	75	62	65
73	79	82	71	66	66	65	77	66	64

5. Teachers' Salaries A sample of 15 teachers from Rhode Island has an average salary of \$35,270, with a standard deviation of \$3256. A sample of 30 teachers from New York has an average salary of \$29,512, with a standard deviation of \$1432. Is there a significant difference in teachers' salaries between the two states? Use $\alpha = 0.02$. Find the 98% confidence interval for the difference of the two means. (9–2)

6. Soft Drinks in School The data show the amounts (in thousands of dollars) of the contracts for soft drinks in local school districts. At $\alpha = 0.10$ can it be concluded that there is a difference in the averages? Use the *P*-value method. Give a reason why the result would be of concern to a cafeteria manager. (9–2)

Pepsi						Coca-Cola		
46	120	80	500	100	59	420	285	57

Source: Local school districts.

7. **High and Low Temperatures** March is a month of variable weather in the Northeast. The chart below records the actual high and low temperatures for a selection of days in March from the weather report for Pittsburgh, Pennsylvania. At the 0.01 level of significance, is there sufficient evidence to conclude that there is more than a 10° difference between average highs and lows? (9–3)

Maximum	44	46	46	36	34	36	57	62	73	53
Minimum	27	34	24	19	19	26	33	57	46	26

Source: www.wunderground.com

8. **Automobile Part Production** In an effort to increase production of an automobile part, the factory manager decides to play music in the manufacturing area. Eight workers are selected, and the number of items each produced for a specific day is recorded. After one week of music, the same workers are monitored again. The data are given in the table. At $\alpha = 0.05$, can the manager conclude that the music has increased production? (9–3)

Worker	1	2	3	4	5	6	7	8
Before	6	8	10	9	5	12	9	7
After	10	12	9	12	8	13	8	10

9. **Lay Teachers in Religious Schools** A study found a slightly lower percentage of lay teachers in religious secondary schools than in elementary schools. A random sample of 200 elementary school and 200 secondary school teachers from religious schools in a large diocese found the following. At the 0.05 level of significance is there sufficient evidence to conclude a difference in proportions? (9–4)

	Elementary	Secondary
Sample size	200	200
Lay teachers	49	62

Source: *New York Times Almanac.*

10. **Adopted Pets** According to the 2005–2006 National Pet Owners Survey, only 16% of pet dogs were adopted from an animal shelter and 15% of pet cats were adopted. To test this difference in proportions of adopted pets, a survey was taken in a local region. Is there sufficient evidence to conclude that there is a difference in proportions? Use $\alpha = 0.05$. (9–4)

	Dogs	Cats
Number	180	200
Adopted	36	30

Source: www.hsus.org

11. **Noise Levels in Hospitals** In the hospital study cited previously, the standard deviation of the noise levels of the 11 intensive care units was 4.1 dBA, and the standard deviation of the noise levels of 24 nonmedical care areas, such as kitchens and machine rooms, was 13.2 dBA. At $\alpha = 0.10$, is there a significant difference between the standard deviations of these two areas? (9–5)

Source: M. Bayo, A. Garcia, and A. Garcia, "Noise Levels in an Urban Hospital and Workers' Subjective Responses," *Archives of Environmental Health.*

12. **Heights of World Famous Cathedrals** The heights (in feet) for a random sample of world famous cathedrals are listed below. In addition, the heights for a sample of the tallest buildings in the world are listed. Is there sufficient evidence at $\alpha = 0.05$ to conclude that there is a difference in the variances in height between the two groups? (9–5)

Cathedrals	72	114	157	56	83	108	90	151	
Tallest buildings	452	442	415	391	355	344	310	302	209

Source: www.infoplease.com

13. **Paint Prices** Two large home improvement stores advertise that they sell their paint at the same average price per gallon. A random sample of 25 cans from store Y had a standard deviation of \$5.21, and store Z had a standard deviation of \$4.08 based on a sample of 20 cans. At $\alpha = 0.05$ can we conclude that the variances are different? How much less would store Z's standard deviation have to be in order to conclude a difference? (9–5)

Statistics Today

To Vaccinate or Not to Vaccinate? Small or Large?—Revisited

Using a *z* test to compare two proportions, the researchers found that the proportion of residents in smaller nursing homes who were vaccinated (80.8%) was statistically greater than that of residents in large nursing homes who were vaccinated (68.7%). Using statistical methods presented in later chapters, they also found that the larger size of the nursing home and the lower frequency of vaccination were significant predictions of influenza outbreaks in nursing homes.

Data Analysis

The Data Bank is found in Appendix D, or on the World Wide Web by following links from www.mhhe.com/math/stat/bluman/

1. From the Data Bank, select a variable and compare the mean of the variable for a random sample of at least 30 men with the mean of the variable for the random sample of at least 30 women. Use a z test.
2. Repeat the experiment in Exercise 1, using a different variable and two samples of size 15. Compare the means by using a t test.
3. Compare the proportion of men who are smokers with the proportion of women who are smokers. Use the data in the Data Bank. Choose random samples of size 30 or more. Use the z test for proportions.
4. Select two samples of 20 values from the data in Data Set IV in Appendix D. Test the hypothesis that the mean heights of the buildings are equal.
5. Using the same data obtained in Exercise 4, test the hypothesis that the variances are equal.

Chapter Quiz

Determine whether each statement is true or false. If the statement is false, explain why.

1. When you are testing the difference between two means, it is not important to distinguish whether the samples are independent of each other. False
2. If the same diet is given to two groups of randomly selected individuals, the samples are considered to be dependent. False
3. When computing the F test value, you always place the larger variance in the numerator of the fraction. True
4. Tests for variances are always two-tailed. False

Select the best answer.

5. To test the equality of two variances, you would use a(n) _______ test.
   a. z *c.* Chi-square
   b. t (*d.*) F
6. To test the equality of two proportions, you would use a(n) _______ test.
   (*a.*) z *c.* Chi-square
   b. t *d.* F
7. The mean value of the F is approximately equal to
   a. 0 (*c.*) 1
   b. 0.5 *d.* It cannot be determined.
8. What test can be used to test the difference between two sample means when the population variances are known?
   (*a.*) z *c.* Chi-square
   b. t *d.* F

Complete these statements with the best answer.

9. If you hypothesize that there is no difference between means, this is represented as H_0: _______. $\mu_1 = \mu_2$
10. When you are testing the difference between two means, the _______ test is used when the population variances are not known. t
11. When the t test is used for testing the equality of two means, the populations must be _______. Normal
12. The values of F cannot be _______. Negative
13. The formula for the F test for variances is _______. $\frac{s_1^2}{s_2^2}$

For each of these problems, perform the following steps.

a. State the hypotheses and identify the claim.
b. Find the critical value(s).
c. Compute the test value.
d. Make the decision.
e. Summarize the results.

Use the traditional method of hypothesis testing unless otherwise specified.

14. **Cholesterol Levels** A researcher wishes to see if there is a difference in the cholesterol levels of two groups of men. A random sample of 30 men between the ages of 25 and 40 is selected and tested. The average level is 223. A second sample of 25 men between the ages of 41 and 56 is selected and tested. The average of this group is 229. The population standard deviation for both groups is 6. At $\alpha = 0.01$, is there a difference in the cholesterol levels between the two groups? Find the 99% confidence interval for the difference of the two means.
15. **Apartment Rental Fees** The data shown are the rental fees (in dollars) for two random samples of apartments in a large city. At $\alpha = 0.10$, can it be concluded that the average rental fee for apartments in the east is greater than the average rental fee in the west? Assume $\sigma_1 = 119$ and $\sigma_2 = 103$.

East					West				
495	390	540	445	420	525	400	310	375	750
410	550	499	500	550	390	795	554	450	370
389	350	450	530	350	385	395	425	500	550
375	690	325	350	799	380	400	450	365	425
475	295	350	485	625	375	360	425	400	475
275	450	440	425	675	400	475	430	410	450
625	390	485	550	650	425	450	620	500	400
685	385	450	550	425	295	350	300	360	400

Source: *Pittsburgh Post-Gazette.*

16. Prices of Low-Calorie Foods The average price of a sample of 12 bottles of diet salad dressing taken from different stores is \$1.43. The standard deviation is \$0.09. The average price of a sample of 16 low-calorie frozen desserts is \$1.03. The standard deviation is \$0.10. At $\alpha = 0.01$, is there a significant difference in price? Find the 99% confidence interval of the difference in the means.

17. Jet Ski Accidents The data shown represent the number of accidents people had when using jet skis and other types of wet bikes. At $\alpha = 0.05$, can it be concluded that the average number of accidents per year has increased from one period to the next?

1987–1991			1992–1996		
376	650	844	1650	2236	3002
1162	1513		4028	4010	

Source: *USA TODAY.*

18. Salaries of Chemists A sample of 12 chemists from Washington state shows an average salary of \$39,420 with a standard deviation of \$1659, while a sample of 26 chemists from New Mexico has an average salary of \$30,215 with a standard deviation of \$4116. Is there a significant difference between the two states in chemists' salaries at $\alpha = 0.02$? Find the 98% confidence interval of the difference in the means.

19. Family Incomes The average income of 15 families who reside in a large metropolitan East Coast city is \$62,456. The standard deviation is \$9652. The average income of 11 families who reside in a rural area of the Midwest is \$60,213, with a standard deviation of \$2009. At $\alpha = 0.05$, can it be concluded that the families who live in the cities have a higher income than those who live in the rural areas? Use the *P*-value method.

20. Mathematical Skills In an effort to improve the mathematical skills of 10 students, a teacher provides a weekly 1-hour tutoring session for the students. A pretest is given before the sessions, and a posttest is given after. The results are shown here. At $\alpha = 0.01$, can it be concluded that the sessions help to improve the students' mathematical skills?

Student	1	2	3	4	5	6	7	8	9	10
Pretest	82	76	91	62	81	67	71	69	80	85
Posttest	88	80	98	80	80	73	74	78	85	93

21. Egg Production To increase egg production, a farmer decided to increase the amount of time the lights in his hen house were on. Ten hens were selected, and the number of eggs each produced was recorded. After one week of lengthened light time, the same hens were monitored again. The data are given here. At $\alpha = 0.05$, can it be concluded that the increased light time increased egg production?

Hen	1	2	3	4	5	6	7	8	9	10
Before	4	3	8	7	6	4	9	7	6	5
After	6	5	9	7	4	5	10	6	9	6

22. Factory Worker Literacy Rates In a sample of 80 workers from a factory in city A, it was found that 5% were unable to read, while in a sample of 50 workers in city B, 8% were unable to read. Can it be concluded that there is a difference in the proportions of nonreaders in the two cities? Use $\alpha = 0.10$. Find the 90% confidence interval for the difference of the two proportions.

23. Male Head of Household A recent survey of 200 households showed that 8 had a single male as the head of household. Forty years ago, a survey of 200 households showed that 6 had a single male as the head of household. At $\alpha = 0.05$, can it be concluded that the proportion has changed? Find the 95% confidence interval of the difference of the two proportions. Does the confidence interval contain 0? Why is this important to know?

Source: Based on data from the U.S. Census Bureau.

24. Money Spent on Road Repair A politician wishes to compare the variances of the amount of money spent for road repair in two different counties. The data are given here. At $\alpha = 0.05$, is there a significant difference in the variances of the amounts spent in the two counties? Use the *P*-value method.

County A	County B
$s_1 = \$11,596$	$s_2 = \$14,837$
$n_1 = 15$	$n_2 = 18$

25. Heights of Basketball Players A researcher wants to compare the variances of the heights (in inches) of four-year college basketball players with those of players in junior colleges. A sample of 30 players from each type of school is selected, and the variances of the heights for each type are 2.43 and 3.15, respectively. At $\alpha = 0.10$, is there a significant difference between the variances of the heights in the two types of schools?

Critical Thinking Challenges

1. The study cited in the article entitled "Only the Timid Die Young" stated that "Timid rats were 60% more likely to die at any given time than were their outgoing brothers." Based on the results, answer the following questions.

a. Why were rats used in the study?

b. What are the variables in the study?

c. Why were infants included in the article?

d. What is wrong with extrapolating the results to humans?

e. Suggest some ways humans might be used in a study of this type.

ONLY THE TIMID DIE YOUNG

DO OVERACTIVE STRESS HORMONES DAMAGE HEALTH?

ABOUT 15 OUT OF 100 CHILDREN ARE BORN SHY, BUT ONLY THREE WILL BE SHY AS ADULTS.

FEARFUL TYPES MAY MEET THEIR maker sooner, at least among rats. Researchers have for the first time connected a personality trait—fear of novelty—to an early death.

Sonia Cavigelli and Martha McClintock, psychologists at the University of Chicago, presented unfamiliar bowls, tunnels and bricks to a group of young male rats. Those hesitant to explore the mystery objects were classified as "neophobic."

The researchers found that the neophobic rats produced high levels of stress hormones, called glucocorticoids—typically involved in the fight-or-flight stress response—when faced with strange situations. Those rats continued to have high levels of the hormones at random times throughout their lives, indicating that timidity is a fixed and stable trait. The team then set out to examine the cumulative effects of this personality trait on the rats' health.

Timid rats were 60 percent more likely to die at any given time than were their outgoing brothers. The causes of death were similar for both groups. "One hypothesis as to why the neophobic rats died earlier is that the stress hormones negatively affected their immune system," Cavigelli says. Neophobes died, on average, three months before their rat brothers, a significant gap, considering that most rats lived only two years.

Shyness—the human equivalent of neophobia—can be detected in infants as young as 14 months. Shy people also produce more stress hormones than "average," or thrill-seeking humans. But introverts don't necessarily stay shy for life, as rats apparently do. Jerome Kagan, a professor of psychology at Harvard University, has found that while 15 out of every 100 children will be born with a shy temperament, only three will appear shy as adults. None, however, will be extroverts.

Extrapolating from the doomed fate of neophobic rats to their human counterparts is difficult. "But it means that something as simple as a personality trait could have physiological consequences," Cavigelli says.

—*Carlin Flora*

2. Based on the study presented in the article entitled "Sleeping Brain, Not at Rest," answer these questions.

a. What were the variables used in the study?

b. How were they measured?

c. Suggest a statistical test that might have been used to arrive at the conclusion.

d. Based on the results, what would you suggest for students preparing for an exam?