Difference between two population means with matched samples

MATCHED

We use the data on production times in Table 10.2 to illustrate the matched-samples procedure (file 'Matched.SAV' on the accompanying CD). The completion times for method 1 are entered into the first column of the PASW data file and the completion times for method 2 are entered into the second column. The PASW steps for matched samples are as follows:

Step 1 **Analyze > Compare Means > Paired-Samples T Test** [Main menu bar]

Step 2 Transfer Method 1 to **Variable1** in the **Paired variables** area [**Paired-Samples T Test** panel]
Transfer Method 2 to **Variable2** in the **Paired variables** area
Click **Options**

Step 3 Enter **95** in the **Confidence Interval** box [**Paired-Samples T Test:Options** panel]
Click **Continue**

Step 4 Click **OK** [**Paired-Samples T Test** panel]

Chapter 11

Inferences about Population Variances

Learning objectives

After studying this chapter and doing the exercises, you should be able to:

1. Construct confidence intervals for a population standard deviation or population variance, using the chi-squared distribution.
2. Conduct and interpret the results of hypothesis tests for a population standard deviation or population variance, using the chi-squared distribution.
3. Conduct and interpret the results of hypothesis tests to compare two population standard deviations or population variances, using the *F* distribution.

In the preceding four chapters we examined methods of statistical inference involving population means and population proportions. In this chapter we extend the discussion to situations involving inferences about population variances.

In many manufacturing applications, controlling the process variance is extremely important in maintaining quality. Consider the production process of filling containers with a liquid detergent product, for example. The filling mechanism for the process is adjusted so that the mean filling weight is 500 grams per container. In addition, the variance of the filling weights is critical. Even with the filling mechanism properly adjusted for the mean of 500 grams, we cannot expect every container to contain exactly 500 grams. By selecting a sample of containers, we can compute a sample variance for the number of grams placed in a container. This value will serve as an estimate of the variance for the population of containers being filled by the production process. If the sample variance is modest, the production process will be continued. However, if the sample variance is excessive, over-filling and under-filling may be occurring, even though the mean is correct at 500 grams. In this case, the filling mechanism will be re-adjusted in an attempt to reduce the filling variance for the containers.

In the first section we consider inferences about the variance of a single population. Subsequently, we will discuss procedures that can be used to make inferences comparing the variances of two populations.

11.1 Inferences about a population variance

Recall that sample variance is calculated as follows:

$$s^2 = \frac{\Sigma(x_i - \bar{x})^2}{n - 1} \qquad (11.1)$$

The sample variance (S^2) is a point estimator of the population variance σ^2. To make inferences about σ^2, the sampling distribution of the quantity $(n - 1)S^2/\sigma^2$ can be used, under appropriate circumstances.

Statistics in Practice

Takeovers and mergers in the UK brewing industry

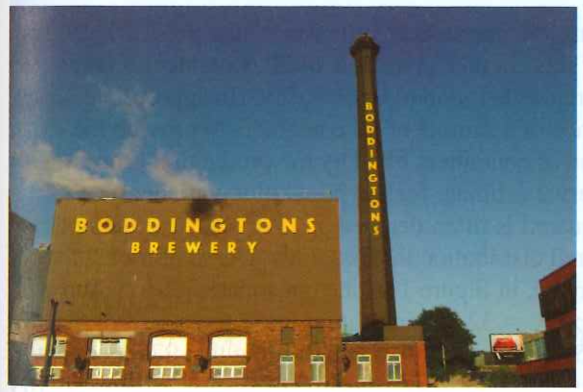
In the last 50 or so years, the structure of the UK brewing industry has changed radically. The *Statistical Handbook* of the British Beer & Pub Association records only 48 brewery companies in the UK in 2003, compared with 362 in 1950. Much of the re-structuring, involving mergers and takeovers, took place during the 1950s and 1960s. By 1969, the number of brewery companies had already fallen to 96. Alison Dean, of the City University Business School, undertook some research to characterize the firms that were the targets of takeovers in the 1945 to 1960 period. She studied samples of firms that were taken over during this period, firms that merged during the period, and firms that remained independent.

Dean's research questioned whether performance criteria were mainly responsible for making firms vulnerable to takeover, or whether non-performance criteria such as size and property holdings were more influential. Her broad conclusion was that performance criteria appeared to be relatively unimportant. Brewery companies typically own 'tied houses' – pubs leased to tenant managers on condition that the tenant buys beer from the brewery. Dean found that a distinguishing characteristic of taken-over firms was a low average asset value per tied house.

In respect of performance criteria, the research looked at profitability, earnings per share, dividend payments, liquidity and share price/net asset ratio. No statistically significant differences were found between the group of taken-over companies and the group of companies that remained independent. In two respects, though, there were significant differences between the group of merged companies and the other two groups. These two criteria were profitability and liquidity, and in both cases it was the standard deviation (rather than the mean of the performance criterion) that distinguished the group of merged companies. For example, in the case of liquidity, measured as the ratio of current assets to current liabilities, the mean values for the three groups of companies were quite similar: 1.6 for the taken-over companies, 1.5 for the merged companies and 1.7 for the independent companies. However, the standard deviation for this ratio was only 0.2 for the merged companies, compared with 0.6 for the independent companies and 0.8 for the taken-over companies. In other words, the group of merged companies seemed very homogeneous in respect of this performance characteristic.

The researcher investigated differences between standard deviations using a statistical test called the *F* test, which you will learn about in this chapter.

Boddington's famous strangeways brewery in Manchester, England.



Source: Alison Dean, The characteristics of takeover target firms: the case of the English Brewing Industry, 1945–1960. *Review of Industrial Organization*, 12, 579–591 (1997)

Sampling distribution of $(n - 1)S^2/\sigma^2$

When a simple random sample of size n is selected from a normal population, the sampling distribution of

$$\frac{(n - 1)S^2}{\sigma^2} \qquad (11.2)$$

has a chi-squared distribution with $n - 1$ degrees of freedom.

Figure 11.1 Examples of the sampling distribution of $(n-1)S^2/\sigma^2$ (chi-squared distribution)

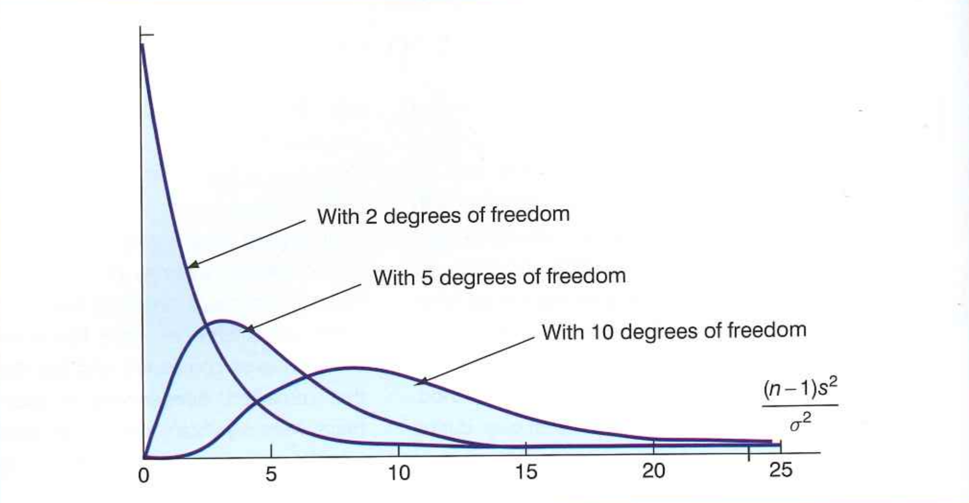


Figure 11.1 shows some possible forms of the sampling distribution of $(n-1)S^2/\sigma^2$. Because the sampling distribution of $(n-1)S^2/\sigma^2$ is a chi-squared distribution, under the conditions described above, we can use this distribution to construct interval estimates and do hypothesis tests about a population variance. Tables of areas or probabilities are readily available for the chi-squared distribution.

Interval estimation

Suppose we are interested in estimating the population variance for the production filling process mentioned at the beginning of this chapter. A sample of 20 containers is taken and the sample variance for the filling quantities is found to be $s^2 = 2.50$ (in appropriate units). However, we cannot expect the variance of a sample of 20 containers to provide the exact value of the variance for the population of containers filled by the production process. Our interest will be in constructing an interval estimate for the population variance.

The Greek letter chi is χ, so chi-squared is often denoted χ^2. We shall use the notation χ^2_α to denote the value for the chi-squared distribution that provides an area or probability of α to the *right* of the χ^2_α value. For example, in Figure 11.2 the chi-squared distribution with 19 degrees of freedom is shown, with $\chi^2_{0.025} = 32.852$ indicating that 2.5 per cent of the chi-squared values are to the right of 32.852, and $\chi^2_{0.975} = 8.907$ indicating that 97.5 per cent of the chi-squared values are to the right of 8.907. Refer to Table 3 of Appendix B and verify that these chi-squared values with 19 degrees of freedom are correct (19th row of the table).

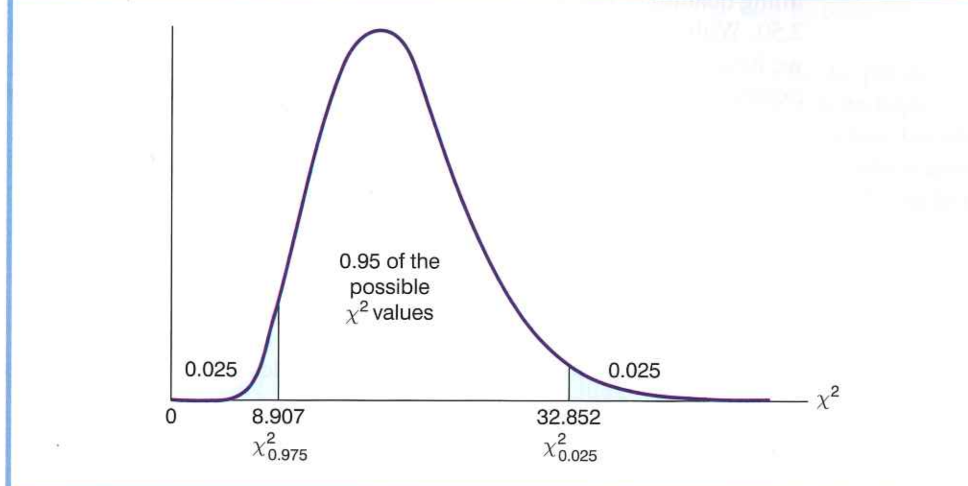
From Figure 11.2 we see that 0.95, or 95 per cent, of the chi-squared values are between $\chi^2_{0.975}$ and $\chi^2_{0.025}$. That is, there is a 0.95 probability of obtaining a χ^2 value such that

$$\chi^2_{0.975} \le \chi^2 \le \chi^2_{0.025}$$

We stated in expression (11.2) that the random variable $(n-1)S^2/\sigma^2$ follows a chi-squared distribution, therefore we can substitute $(n-1)s^2/\sigma^2$ for χ^2 and write

$$\chi^2_{0.975} \le \frac{(n-1)s^2}{\sigma^2} \le \chi^2_{0.025} \qquad (11.3)$$

Figure 11.2 A chi-squared distribution with 19 degrees of freedom



Expression (11.3) provides the basis for an interval estimate in that 0.95, or 95 per cent, of all possible values for $(n-1)\,S^2/\sigma^2$ will be in the interval $\chi^2_{0.975}$ to $\chi^2_{0.025}$. We now need to do some algebraic manipulations with expression (11.3) to develop an interval estimate for the population variance σ^2. Using the leftmost inequality in expression (11.3), we have

$$\chi^2_{0.975} \le \frac{(n-1)s^2}{\sigma^2}$$

So

$$\chi^2_{0.975}\sigma^2 \le (n-1)s^2$$

or

$$\sigma^2 \le \frac{(n-1)s^2}{\chi^2_{0.975}} \qquad (11.4)$$

Doing similar algebraic manipulations with the rightmost inequality in expression (11.3) gives

$$\frac{(n-1)s^2}{\chi^2_{0.025}} \le \sigma^2 \qquad (11.5)$$

Expressions (11.4) and (11.5) can be combined to provide

$$\frac{(n-1)s^2}{\chi^2_{0.025}} \le \sigma^2 \le \frac{(n-1)s^2}{\chi^2_{0.975}} \qquad (11.6)$$

Because expression (11.3) is true for 95 per cent of the $(n-1)s^2/\sigma^2$ values, expression (11.6) provides a 95 per cent confidence interval estimate for the population variance σ^2.

We return to the problem of providing an interval estimate for the population variance of filling quantities. Recall that the sample of 20 containers provided a sample variance of $s^2 = 2.50$. With a sample size of 20, we have 19 degrees of freedom. As shown in Figure 11.2, we have already determined that $\chi^2_{0.975} = 8.907$ and $\chi^2_{0.025} = 32.852$. Using these values in expression (11.6) provides the following interval estimate for the population variance.

$$\frac{(19)(2.50)}{32.825} \leq \sigma^2 \leq \frac{(19)(2.50)}{8.907}$$

or

$$1.45 \leq \sigma^2 \leq 5.33$$

Taking the square root of these values provides the following 95 per cent confidence interval for the population standard deviation.

$$1.20 \leq \sigma \leq 2.31$$

Note that because $\chi^2_{0.975} = 8.907$ and $\chi^2_{0.025} = 32.852$ were used, the interval estimate has a 0.95 confidence coefficient. Extending expression (11.6) to the general case of any confidence coefficient, we have the following interval estimate of a population variance.

Interval estimate of a population variance

$$\frac{(n-1)s^2}{\chi^2_{\alpha/2}} \leq \sigma^2 \leq \frac{(n-1)s^2}{\chi^2_{1-\alpha/2}} \qquad \textbf{(11.7)}$$

where the χ^2 values are based on a chi-squared distribution with $n - 1$ degrees of freedom and where $1 - \alpha$ is the confidence coefficient.

Hypothesis testing

Using σ^2_0 to denote the hypothesized value for the population variance, the three forms for a hypothesis test about a population variance are as follows:

$$H_0: \sigma^2 \geq \sigma^2_0 \qquad H_0: \sigma^2 \leq \sigma^2_0 \qquad H_0: \sigma^2 = \sigma^2_0$$
$$H_1: \sigma^2 < \sigma^2_0 \qquad H_1: \sigma^2 > \sigma^2_0 \qquad H_1: \sigma^2 \neq \sigma^2_0$$

These three forms are similar to the three forms we used to conduct one-tailed and two-tailed hypothesis tests about population means and proportions in Chapters 9 and 10.

Hypothesis tests about a population variance use the hypothesized value for the population variance and the sample variance s^2 to compute the value of a χ^2 test statistic. Assuming that the population has a normal distribution, the test statistic is:

Test statistic for hypothesis tests about a population variance

$$\chi^2 = \frac{(n-1)s^2}{\sigma^2_0} \qquad \textbf{(11.8)}$$

where χ^2 has a chi-squared distribution with $n - 1$ degrees of freedom.

After computing the value of the χ^2 test statistic, either the p-value approach or the critical value approach may be used to determine whether the null hypothesis can be rejected.

Here is an example. The Newcastle Metro Bus Company wants to promote an image of reliability by encouraging its drivers to maintain consistent schedules. As a standard policy the company would like arrival times at bus stops to have low variability. The company standard specifies an arrival time variance of four or less when arrival times are measured in minutes. The following hypothesis test is formulated to help the company determine whether the arrival time population variance is excessive.

$$H_0: \sigma^2 \leq 4$$
$$H_1: \sigma^2 > 4$$

In tentatively assuming H_0 is true, we are assuming that the population variance of arrival times is within the company guideline. We reject H_0 if the sample evidence indicates that the population variance exceeds the guideline. In this case, follow-up steps should be taken to reduce the population variance. We conduct the hypothesis test using a level of significance of $\alpha = 0.05$.

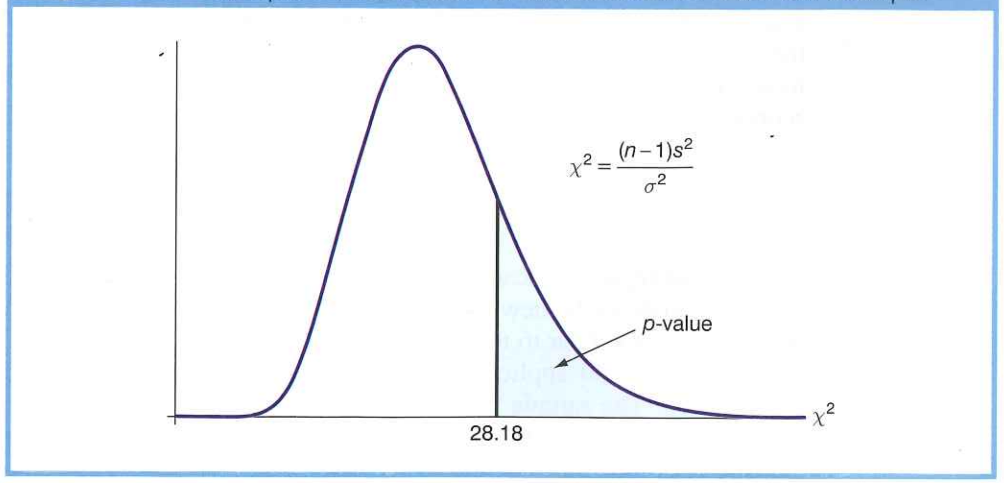
Suppose that a random sample of 24 bus arrivals taken at a city-centre bus stop provides a sample variance of $s^2 = 4.9$. Assuming that the population distribution of arrival times is approximately normal, the value of the test statistic is as follows.

$$\chi^2 = \frac{(n-1)s^2}{\sigma^2_0} = \frac{(24-1)4.9}{4} = 28.18$$

The chi-squared distribution with $n - 1 = 24 - 1 = 23$ degrees of freedom is shown in Figure 11.3. Because this is an upper tail test, the area under the curve to the right of the test statistic $\chi^2 = 28.18$ is the p-value for the test.

Like the t distribution table, the chi-squared distribution table does not contain sufficient detail to enable us to determine the p-value exactly. However, we can use the chi-squared distribution table to obtain a range for the p-value. For example, using Table 3 of Appendix B, we find the following information for a chi-squared distribution with 23 degrees of freedom.

Figure 11.3 Chi-squared distribution for the Newcastle Metro Bus example



Area in upper tail	0.10	0.05	0.025	0.01
χ^2 value (23 df)	32.007	35.172	38.076	41.638

$\chi^2 = 28.18$

Because $\chi^2 = 28.18$ is less than 32.007, the area in the upper tail (the p-value) is greater than 0.10. With the p-value $> \alpha = 0.05$, we cannot reject the null hypothesis. The sample does not support the conclusion that the population variance of the arrival times is excessive.

Because of the difficulty of determining the exact p-value directly from the chi-squared distribution table, a computer software package such as MINITAB, PASW or EXCEL is helpful. The sections at the end of the chapter describe the procedures used to show that with 23 degrees of freedom, $\chi^2 = 28.18$ provides a p-value $= 0.2091$.

As with other hypothesis testing procedures, the critical value approach can also be used to draw the conclusion. With $\alpha = 0.05$, $\chi_{0.05}$ provides the critical value for the upper tail hypothesis test. Using Table 3 of Appendix B and 23 degrees of freedom, $\chi_{0.05} = 35.172$. Consequently, the rejection rule for the bus arrival time example is as follows:

$$\text{Reject } H_0 \text{ if } \chi^2 \geq 35.172$$

Because the value of the test statistic is $\chi^2 = 28.18$, we cannot reject the null hypothesis.

In practice, upper tail tests as presented here are the most frequently encountered tests about a population variance. In situations involving arrival times, production times, filling weights, part dimensions and so on, low variances are desirable, whereas large variances are unacceptable. With a statement about the maximum allowable population variance, we can test the null hypothesis that the population variance is less than or equal to the maximum allowable value against the alternative hypothesis that the population variance is greater than the maximum allowable value. With this test structure, corrective action will be taken whenever rejection of the null hypothesis indicates the presence of an excessive population variance.

As we saw with population means and proportions, other forms of hypothesis test can be done. Let us demonstrate a two-tailed test about a population variance by considering a situation faced by a car driver licensing authority. Historically, the variance in test scores for individuals applying for driving licences has been $\sigma^2 = 100$. A new examination with a new style of test questions has been developed. Administrators of the licensing authority would like the variance in the test scores for the new examination to remain at the historical level. To evaluate the variance in the new examination test scores, the following two-tailed hypothesis test has been proposed.

$$H_0\colon \sigma^2 = 100$$
$$H_1\colon \sigma^2 \neq 100$$

Rejection of H_0 will indicate that a change in the variance has occurred and suggest that some questions in the new examination may need revision to make the variance of the new test scores similar to the variance of the old test scores.

A sample of 30 applicants for driving licences is given the new version of the examination. The sample provides a sample variance $s^2 = 162$. We shall use a level of significance $\alpha = 0.05$ to do the hypothesis test. The value of the chi-squared test statistic is as follows:

Table 11.1 Summary of hypothesis tests about a population variance

	Lower tail test	Upper tail test	Two-tailed test
Hypotheses	$H_0\colon \sigma^2 \geq \sigma_0^2$ $H_1\colon \sigma^2 < \sigma_0^2$	$H_0\colon \sigma^2 \leq \sigma_0^2$ $H_1\colon \sigma^2 > \sigma_0^2$	$H_0\colon \sigma^2 = \sigma_0^2$ $H_1\colon \sigma^2 \neq \sigma_0^2$
Test statistic	$\chi^2 = \frac{(n-1)s^2}{\sigma_0^2}$	$\chi^2 = \frac{(n-1)s^2}{\sigma_0^2}$	$\chi^2 = \frac{(n-1)s^2}{\sigma_0^2}$
Rejection rule: p-value approach	Reject H_0 if p-value $\leq \alpha$	Reject H_0 if p-value $\leq \alpha$	Reject H_0 if p-value $\leq \alpha$
Rejection rule: critical value approach	Reject H_0 if $\chi^2 \leq \chi^2_{1-\alpha}$	Reject H_0 if $\chi^2 \geq \chi^2_{\alpha}$	Reject H_0 if $\chi^2 \leq \chi^2_{1-\alpha/2}$ or if $\chi^2 \geq \chi^2_{\alpha/2}$

$$\chi^2 = \frac{(n-1)s^2}{\sigma_0^2} = \frac{(30-1)162}{100} = 46.98$$

Now, let us compute the p-value. Using Table 3 of Appendix B and $n - 1 = 30 - 1 = 29$ degrees of freedom, we find the following.

Area in upper tail	0.10	0.05	0.025	0.01
χ^2 value (29 df)	39.087	42.557	45.722	49.588

$\chi^2 = 46.98$

The value of the test statistic $\chi^2 = 46.98$ gives an area between 0.025 and 0.01 in the upper tail of the chi-squared distribution. Doubling these values shows that the two-tailed p-value is between 0.05 and 0.02. MINITAB, PASW or EXCEL can be used to show the exact p-value $= 0.0374$. With p-value $< \alpha = 0.05$, we reject H_0 and conclude that the new examination test scores have a population variance different from the historical variance of $\sigma^2 = 100$.

A summary of the hypothesis testing procedures for a population variance is shown in Table 11.1.

Exercises

Methods

1 Find the following chi-squared distribution values from Table 3 of Appendix B.

a. $\chi^2_{0.05}$ with df = 5 b. $\chi^2_{0.025}$ with df = 15 c. $\chi^2_{0.975}$ with df = 20
d. $\chi^2_{0.01}$ with df = 10 e. $\chi^2_{0.95}$ with df = 18

2 A sample of 20 items provides a sample standard deviation of 5.

a. Compute a 90 per cent confidence interval estimate of the population variance.
b. Compute a 95 per cent confidence interval estimate of the population variance.
c. Compute a 95 per cent confidence interval estimate of the population standard deviation.

3 A sample of 16 items provides a sample standard deviation of 9.5. Test the following hypotheses using $\alpha = 0.05$. What is your conclusion? Use both the *p*-value approach and the critical value approach.

$$H_0: \sigma^2 \le 50$$
$$H_1: \sigma^2 > 50$$

Applications

4 The variance in drug weights is critical in the pharmaceutical industry. For a specific drug, with weights measured in grams, a sample of 18 units provided a sample variance of $s^2 = 0.36$.

a. Construct a 90 per cent confidence interval estimate of the population variance for the weight of this drug.
b. Construct a 90 per cent confidence interval estimate of the population standard deviation.

5 The table below shows return-on-equity (ROE) figures for 2007, for sample of six companies listed on the Tel Aviv stock exchange (Source: Datastream, Thomson Financial).

Company	ROE (%)
Bezeq	25.83
Clal Industries	25.35
Harel Insurance	22.60
Koor Industries	28.72
Mizrahi Bank	17.10
Strauss Group	15.40

a. Compute the sample variance and sample standard deviation for these data.
b. What is the 95 per cent confidence interval for the population variance?
c. What is the 95 per cent confidence interval for the population standard deviation?

6 Because of staffing decisions, managers of the Worldview Hotel are interested in the variability in the number of rooms occupied per day during a particular season of the year. A sample of 20 days of operation shows a sample mean of 290 rooms occupied per day and a sample standard deviation of 30 rooms.

a. What is the point estimate of the population variance?
b. Provide a 90 per cent confidence interval estimate of the population variance.
c. Provide a 90 per cent confidence interval estimate of the population standard deviation.

7 The Fidelity Growth & Income mutual fund received a three-star, or neutral, rating from Morningstar. Shown here are the quarterly percentage returns for the five-year period from 2001 to 2005 (*Morningstar Funds 500*, 2006).

RETURN

	1st Quarter	2nd Quarter	3rd Quarter	4th Quarter
2001	−10.91	5.80	−9.64	6.45
2002	0.83	−10.48	−14.03	5.58
2003	−2.27	10.43	0.85	9.33
2004	1.34	1.11	−0.77	8.03
2005	−2.46	0.89	2.55	1.78

a. Compute the mean, variance, and standard deviation for the quarterly returns.
b. Financial analysts often use standard deviation of percentage returns as a measure of risk for stocks and mutual funds. Construct a 95 per cent confidence interval for the population standard deviation of quarterly returns for the Fidelity Growth & Income mutual fund.

8 In the file 'Travel' on the accompanying CD, there are estimated daily living costs (in euros) for a businessman travelling to 20 major cities. The estimates include a single room at a four-star hotel, beverages, breakfast, taxi fares, and incidental costs.

a. Compute the sample mean.
b. Compute the sample standard deviation.
c. Compute a 95 per cent confidence interval for the population standard deviation.

TRAVEL

City	Daily living cost	City	Daily living cost
Bangkok	242.87	Madrid	283.56
Bogota	260.93	Mexico City	212.00
Bombay	139.16	Milan	284.08
Cairo	194.19	Paris	436.72
Dublin	260.76	Rio de Janeiro	240.87
Frankfurt	355.36	Seoul	310.41
Hong Kong	346.32	Tel Aviv	223.73
Johannesburg	165.37	Toronto	181.25
Lima	250.08	Warsaw	238.20
London	326.76	Washington DC	250.61

9 To analyze the risk, or volatility, associated with investing in Chevron Corporation common stock, a sample of the monthly total percentage return for 12 months was taken. The returns for the 12 months of 2005 are shown here (*Compustat*, February 24, 2006). Total return is price appreciation plus any dividend paid.

CHEVRON

Month	Return (%)	Month	Return (%)
January	3.60	July	3.74
February	14.86	August	6.62
March	−6.07	September	5.42
April	−10.82	October	−11.83
May	4.29	November	1.21
June	3.98	December	−0.94

a. Compute the sample variance and sample standard deviation as a measure of volatility of monthly total return for Chevron.
b. Construct a 95 per cent confidence interval for the population variance.
c. Construct a 95 per cent confidence interval for the population standard deviation.

10 Part variability is critical in the manufacturing of ball bearings. Large variances in the size of the ball bearings cause bearing failure and rapid wear. Production standards call for a maximum variance of 0.0025 when the bearing sizes are measured in millimetres. A sample of 15 bearings shows a sample standard deviation of 0.066 mm.

a. Use $\alpha = 0.10$ to determine whether the sample indicates that the maximum acceptable variance is being exceeded.

b. Compute a 90 per cent confidence interval estimate for the variance of the ball bearings in the population.

11 The average standard deviation for the annual return of large cap stock mutual funds is 18.2 per cent (*The Top Mutual Funds*, AAII, 2004). The sample standard deviation based on a sample of size 36 for the Vanguard PRIMECAP mutual fund is 22.2 per cent. Construct a hypothesis test that can be used to determine whether the standard deviation for the Vanguard fund is greater than the average standard deviation for large cap mutual funds. With a 0.05 level of significance, what is your conclusion?

12 A sample standard deviation for the number of passengers taking a particular airline flight is 8. A 95 per cent confidence interval estimate of the population standard deviation is 5.86 passengers to 12.62 passengers.

a. Was a sample size of 10 or 15 used in the statistical analysis?

b. Suppose the sample standard deviation of $s = 8$ was based on a sample of 25 flights. What change would you expect in the confidence interval for the population standard deviation? Compute a 95 per cent confidence interval estimate of σ with a sample size of 25.

11.2 Inferences about two population variances

In some statistical applications we may want to compare the variances in product quality resulting from two different production processes, the variances in assembly times for two assembly methods, or the variances in temperatures for two heating devices. In making comparisons about the two population variances, we shall be using data collected from two independent random samples, one from population 1 and another from population 2. The two sample variances s_1^2 and s_2^2 will be the basis for making inferences about the two population variances σ_1^2 and σ_2^2. Whenever the variances of two normal populations are equal ($\sigma_1^2 = \sigma_2^2$), the sampling distribution of the ratio of the two sample variances is as follows.

Sampling distribution of S_1^2/S_2^2 when $\sigma_1^2 = \sigma_2^2$

When independent simple random samples of sizes n_1 and n_2 are selected from two normal populations with equal variances, the sampling distribution of

$$\frac{S_1^2}{S_2^2} \qquad (11.9)$$

has an F distribution with $n_1 - 1$ degrees of freedom for the numerator and $n_2 - 1$ degrees of freedom for the denominator; S_1^2 is the sample variance for the random sample of n_1 items from population 1, and S_2^2 is the sample variance for the random sample of n_2 items from population 2.

Figure 11.4 F distribution with 20 degrees of freedom for the numerator and 20 degrees of freedom for the denominator

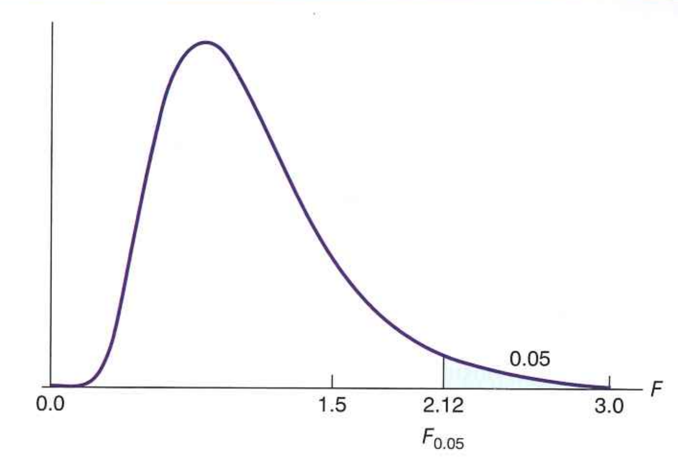


Figure 11.4 is a graph of the F distribution with 20 degrees of freedom for both the numerator and denominator. As can be seen from this graph, the F distribution is not symmetrical, and the F values can never be negative. The shape of any particular F distribution depends on its numerator and denominator degrees of freedom.

We shall use F_α to denote the value of F that provides an area or probability of α in the upper tail of the distribution. For example, as noted in Figure 11.4, $F_{0.05}$ identifies the upper tail area of 0.05 for an F distribution with 20 degrees of freedom for both the numerator and for the denominator. The specific value of $F_{0.05}$ can be found by referring to the F distribution table, Table 4 of Appendix B. Using 20 degrees of freedom for the numerator, 20 degrees of freedom for the denominator, and the row corresponding to an area of 0.05 in the upper tail, we find $F_{0.05} = 2.12$. Note that the table can be used to find F values for upper tail areas of 0.10, 0.05, 0.025 and 0.01.

We now show how the F distribution can be used to do a hypothesis test about the equality of two population variances. The hypotheses are stated as follows.

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

We make the tentative assumption that the population variances are equal. If H_0 is rejected, we will draw the conclusion that the population variances are not equal.

The hypothesis test requires two independent random samples, one from each population. The two sample variances are then computed. We refer to the population providing the *larger* sample variance as population 1. A sample size of n_1 and a sample variance of s_1^2 correspond to population 1, and a sample size of n_2 and a sample variance of s_2^2 correspond to population 2. Based on the assumption that both populations have a normal distribution, the ratio of sample variances provides the following F test statistic.

Test statistic for hypothesis tests about population variances with $\sigma_1^2 = \sigma_2^2$

$$F = \frac{s_1^2}{s_2^2} \tag{11.10}$$

Denoting the population with the larger sample variance as population 1, the test statistic has an F distribution with $n_1 - 1$ degrees of freedom for the numerator and $n_2 - 1$ degrees of freedom for the denominator.

Because the F test statistic is constructed with the larger sample variance in the numerator, the value of the test statistic will be in the upper tail of the F distribution. Therefore, the F distribution table (Table 4 of Appendix B) need only provide upper tail areas or probabilities.

We now consider an example. Midlands Schools is renewing its school bus service contract for the coming year and must select one of two bus companies, the Red Bus Company or the Route One Company. We shall assume that the two companies have similar performance for average punctuality (i.e. mean arrival time) and use the variance of the arrival times as a primary measure of the quality of the bus service. Low variance values indicate the more consistent and higher quality service. If the variances of arrival times associated with the two services are equal, Midlands Schools' managers will select the company offering the better financial terms. However, if the sample data on bus arrival times for the two companies indicate a significant difference between the variances, the administrators may want to give special consideration to the company with the better or lower variance service. The appropriate hypotheses follow.

SCHOOL BUS

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

If H_0 can be rejected, the conclusion of unequal service quality is appropriate. We shall use a level of significance of $\alpha = 0.10$ to do the hypothesis test. A sample of 26 arrival times for the Red Bus service provides a sample variance of 48 and a sample of 16 arrival times for the Route One service provides a sample variance of 20. Because the Red Bus sample provided the larger sample variance, we shall denote Red Bus as population 1. Using equation (11.10), the value of the test statistic is

$$F = \frac{s_1^2}{s_2^2} = \frac{48}{20} = 2.40$$

The corresponding F distribution has $n_1 - 1 = 26 - 1 = 25$ numerator degrees of freedom and $n_2 - 1 = 16 - 1 = 15$ denominator degrees of freedom. As with other hypothesis testing procedures, we can use the p-value approach or the critical value approach to reach a conclusion. Table 4 of Appendix B shows the following areas in the upper tail and corresponding F values for an F distribution with 25 numerator degrees of freedom and 15 denominator degrees of freedom.

Area in upper tail	0.10	0.05	0.025	0.01
F value ($df_1 = 25$, $df_2 = 15$)	1.89	2.28	2.69	3.28

$F = 2.40$

Because $F = 2.40$ is between 2.28 and 2.69, the area in the upper tail of the distribution is between 0.05 and 0.025. Since this is a two-tailed test, we double the upper tail area, which results in a p-value between 0.10 and 0.05. For this test, we selected $\alpha = 0.10$ as the level of significance, which gives us a p-value $< \alpha = 0.10$. Hence, the null hypothesis is rejected. This finding leads to the conclusion that the two bus services differ in terms of arrival time variances. The recommendation is that the Midlands Schools' managers give special consideration to the better or lower variance service offered by the Route One Company.

We can use MINITAB, PASW or EXCEL to show that the test statistic $F = 2.40$ provides a two-tailed p-value $= 0.0811$. With $0.0811 < \alpha = 0.10$, the null hypothesis of equal population variances is rejected.

To use the critical value approach to do the two-tailed hypothesis test at the $\alpha = 0.10$ level of significance, we select critical values with an area of $\alpha/2 = 0.10/2 = 0.05$ in each tail of the distribution. Because the value of the test statistic computed using equation (11.10) will always be in the upper tail, we only need to determine the upper tail critical value. From Table 4 of Appendix B, we see that $F_{0.05} = 2.28$. So, even though we use a two-tailed test, the rejection rule is stated as follows.

$$\text{Reject } H_0 \text{ if } F \geq 2.28$$

Because the test statistic $F = 2.40$ is greater than 2.28, we reject H_0 and conclude that the two bus services differ in terms of arrival time variances.

One-tailed tests involving two population variances are also possible. In this case, we use the F distribution to determine whether one population variance is significantly greater than the other. If we are using tables of the F distribution to compute the p-value or determine the critical value, a one-tailed hypothesis test about two population variances will always be formulated as an *upper tail* test:

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

This form of the hypothesis test always places the p-value and the critical value in the upper tail of the F distribution. As a result, only upper tail F values will be needed, simplifying both the computations and the table for the F distribution.

As an example of a one-tailed test, consider a public opinion survey. Samples of 31 men and 41 women were used to study attitudes about current political issues. The researcher conducting the study wants to test to see if women show a greater variation in attitude on political issues than men. In the form of the one-tailed hypothesis test given previously, women will be denoted as population 1 and men will be denoted as population 2. The hypothesis test will be stated as follows.

$$H_0: \sigma^2_{women} \leq \sigma^2_{men}$$
$$H_1: \sigma^2_{women} > \sigma^2_{men}$$

Rejection of H_0 will give the researcher the statistical support necessary to conclude that women show a greater variation in attitude on political issues.

With the sample variance for women in the numerator and the sample variance for men in the denominator, the F distribution will have $n_1 - 1 = 41 - 1 = 40$ numerator degrees of freedom and $n_2 - 1 = 31 - 1 = 30$ denominator degrees of freedom. We shall use a level of significance $\alpha = 0.05$ for the hypothesis test. The survey results provide a

Table 11.2 Summary of hypothesis tests about two population variances

	Upper tail test	Two-tailed test
Hypotheses	$H_0: \sigma_1^2 \le \sigma_2^2$ $H_1: \sigma_1^2 > \sigma_2^2$	$H_0: \sigma_1^2 = \sigma_2^2$ $H_1: \sigma_1^2 \ne \sigma_2^2$
	Note: Population 1 has the larger sample variance	
Test statistic	$F = \dfrac{s_1^2}{s_2^2}$	$F = \dfrac{s_1^2}{s_2^2}$
Rejection rule: p-value approach	Reject H_0 if p-value $\le \alpha$	Reject H_0 if p-value $\le \alpha$
Rejection rule: critical value approach	Reject H_0 if $F \ge F_\alpha$	Reject H_0 if $F \ge F_{\alpha/2}$

sample variance of $s_1^2 = 120$ for women and a sample variance of $s_2^2 = 80$ for men. The test statistic is as follows.

$$F = \frac{s_1^2}{s_2^2} = \frac{120}{80} = 1.50$$

Referring to Table 4 in Appendix B, we find that an F distribution with 40 numerator degrees of freedom and 30 denominator degrees of freedom has $F_{0.10} = 1.57$. Because the test statistic $F = 1.50$ is less than 1.57, the area in the upper tail must be greater than 0.10. Hence, we can conclude that the p-value is greater than 0.10. Using MINITAB, PASW or EXCEL provides a p-value $= 0.1256$. Because the p-value $> \alpha = 0.05$, H_0 cannot be rejected. Hence, the sample results do not support the conclusion that women show greater variation in attitude on political issues than men.

Table 11.2 provides a summary of hypothesis tests about two population variances. Research confirms that the F distribution is sensitive to the assumption of normal populations. The F distribution should not be used unless it is reasonable to assume that both populations are at least approximately normally distributed.

Exercises

Methods

13 Find the following F distribution values from Table 4 of Appendix B.

a. $F_{0.05}$ with degrees of freedom 5 and 10
b. $F_{0.025}$ with degrees of freedom 20 and 15
c. $F_{0.01}$ with degrees of freedom 8 and 12
d. $F_{0.10}$ with degrees of freedom 10 and 20

14 A sample of 16 items from population 1 has a sample variance $s^2 = 5.8$ and a sample of 21 items from population 2 has a sample variance $s^2 = 2.4$. Test the following hypotheses at the 0.05 level of significance.

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

a. What is your conclusion using the p-value approach?
b. Repeat the test using the critical value approach.

15 Consider the following hypothesis test.

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

a. What is your conclusion if $n_1 = 21$, $s_1^2 = 8.2$, $n_2 = 26$, $s_2^2 = 4.0$? Use $\alpha = 0.05$ and the p-value approach.
b. Repeat the test using the critical value approach.

Applications

16 Most individuals are aware of the fact that the average annual repair cost for a car depends on its age. A researcher is interested in finding out whether the variance of the annual repair costs also increases with the age of the car. A sample of 26 cars that were eight years old showed a sample standard deviation for annual repair costs of £170 and a sample of 25 cars that were four years old showed a sample standard deviation for annual repair costs of £100.

a. State the null and alternative hypotheses if the research hypothesis is that the variance in annual repair costs is larger for the older cars.
b. At a 0.01 level of significance, what is your conclusion? What is the p-value? Discuss the reasonableness of your findings.

17 On the basis of data provided by a salary survey, the variance in annual salaries for seniors in accounting firms is approximately 2.1 and the variance in annual salaries for managers in accounting firms is approximately 11.1. The salary data were provided in thousands of euros. Assuming that the salary data were based on samples of 25 seniors and 26 managers, test the hypothesis that the population variances in the salaries are equal. At a 0.05 level of significance, what is your conclusion?

18 Fidelity Magellan is a large cap growth mutual fund and Fidelity Small Cap Stock is a small cap growth mutual fund (*Morningstar Funds 500*, 2006). The standard deviation for both funds was computed based on a sample of size 26. For Fidelity Magellan, the sample standard deviation is 8.89 per cent; for Fidelity Small Cap Stock, the sample standard deviation is 13.03 per cent. Financial analysts often use the standard deviation as a measure of risk. Conduct a hypothesis test to determine whether the small cap growth fund is riskier than the large cap growth fund. Use $\alpha = 0.05$ as the level of significance.

19 Two new assembly methods are tested and the variances in assembly times are reported. Use $\alpha = 0.10$ and test for equality of the two population variances.

	Method A	Method B
Sample size	$n_1 = 31$	$n_2 = 25$
Sample variation	$s_1^2 = 25$	$s_2^2 = 12$

20 A research hypothesis is that the variance of stopping distances of cars on wet roads is greater than the variance of stopping distances of cars on dry roads. In the research study, 16 cars travelling at the same speeds are tested for stopping distances on wet roads and then tested for stopping distances on dry roads. On wet roads, the standard deviation of stopping distances is ten metres. On dry roads, the standard deviation is five metres.

a. At a 0.05 level of significance, do the sample data justify the conclusion that the variance in stopping distances on wet roads is greater than the variance in stopping distances on dry roads? What is the *p*-value?

b. What are the implications of your statistical conclusions in terms of driving safety recommendations?

21 The grade point averages of 352 students who completed a college course in financial accounting have a standard deviation of 0.940. The grade point averages of 73 students who dropped out of the same course have a standard deviation of 0.797. Do the data indicate a difference between the variances of grade point averages for students who completed a financial accounting course and students who dropped out? Use a 0.05 level of significance. *Note:* $F_{0.025}$ with 351 and 72 degrees of freedom is 1.466.

BAGS

22 The variance in a production process is an important measure of the quality of the process. A large variance often signals an opportunity for improvement in the process by finding ways to reduce the process variance. The file 'Bags' on the accompanying CD contains data for two machines that fill bags with powder. The file has 25 bag weights for Machine 1 and 22 bag weights for Machine 2. Conduct a statistical test to determine whether there is a significant difference between the variances in the bag weights for the two machines. Use a 0.05 level of significance. What is your conclusion? Which machine, if either, provides the greater opportunity for quality improvements?

For additional online summary questions and answers go to the companion website at www.cengage.co.uk/aswsbe2

Summary

In this chapter we presented statistical procedures that can be used to make inferences about population variances. In the process we introduced two new probability distributions: the chi-squared distribution and the F distribution. The chi-squared distribution can be used as the basis for interval estimation and hypothesis tests about the variance of a normal population.

We illustrated the use of the F distribution in hypothesis tests about the variances of two normal populations. With independent simple random samples of sizes n_1 and n_2 selected from two normal populations with equal variances, the sampling distribution of the ratio of the two sample variances has an F distribution with $n_1 - 1$ degrees of freedom for the numerator and $n_2 - 1$ degrees of freedom for the denominator.

Key formulae

Interval estimate of a population variance

$$\frac{(n-1)s^2}{\chi^2_{\alpha/2}} \le \sigma^2 \le \frac{(n-1)s^2}{\chi^2_{1-\alpha/2}} \qquad (11.7)$$

Test statistic for hypothesis tests about a population variance

$$\chi^2 = \frac{(n-1)s^2}{\sigma_0^2} \qquad (11.8)$$

Test statistic for hypothesis tests about population variances with $\sigma_1^2 = \sigma_2^2$

$$F = \frac{s_1^2}{s_2^2} \qquad (11.10)$$

Case problem Global economic problems in 2008

In 2008, particularly in the latter part of the year, there were global economic problems, including banking crises in a number of countries, clear indications of economic recession, and increased stock market volatility.

One method of measuring volatility in stock markets (a relatively unsophisticated one) is to calculate the standard deviation of percentage changes in stock market prices or share index levels (e.g. daily percentage changes or weekly percentage changes). Indeed, in many texts on finance, this is the first operational definition offered for the concept of 'volatility'.

The data in the file 'Share indices 2007-8' (on the accompanying CD) are samples of daily percentage changes in four well-known stock market indices during each of the years 2007 and 2008. The four indices are the FTSE 100 (London Stock Exchange, UK), the DAX 40 (Frankfurt Stock Exchange, Germany), the FTSE/JSE All Share (Johannesburg Stock Exchange, South Africa) and the ISE National 100 (Istanbul Stock Exchange, Turkey). The samples are of size 50 for each of 2007 and 2008.

The report you are asked to prepare below should be focused particularly on the question of whether the stock markets showed greater volatility in 2008 than in 2007.

SHARE INDICES 2007–8

FTSE 2007	DAX 2007	FTSE/JSE 2007	ISE 2007	FTSE 2008	DAX 2008	FTSE/JSE 2008	ISE 2008
-0.76	0.11	-0.16	-0.40	-1.37	-1.35	-2.62	-2.55
0.96	0.19	1.34	1.27	-0.85	0.08	-1.40	-1.04
0.49	-0.21	0.44	0.69	1.32	0.03	2.19	1.45
1.40	0.13	1.36	1.90	4.75	-0.24	5.28	5.81
-1.22	0.16	-0.89	-1.41	1.64	-0.26	1.44	2.03
1.47	-0.13	2.69	4.40	-0.60	0.01	-3.05	-0.97
-1.06	0.09	0.25	-1.46	-0.20	0.00	-1.17	0.08
0.39	0.06	-1.04	1.17	-1.52	-0.59	-2.13	-2.11
-0.58	0.13	-0.88	-0.25	-0.41	-0.33	-0.75	3.18
0.28	0.19	0.57	0.16	0.74	-0.38	-0.05	2.34

Analyst's report

1 Use appropriate descriptive statistics to summarize the daily percentage change data for each index in 2007 and 2008. What similarities or differences do you observe from the sample data?

2 Use the methods of Chapter 10 to comment on any difference between the population mean daily percentage change in each index for 2007 versus 2008. Discuss your findings.

3 Compute the standard deviation of the daily percentage changes for each share index, for 2007 and for 2008. For each share index, do a hypothesis test to examine the equality of population variances in 2007 and 2008. Discuss your findings.

4 What conclusions can you reach about any differences between 2007 and 2008?

Sales all year round to try and entice shoppers back to spending. © DBURKE/Alamy.

Software Section for Chapter 11

Population variances using MINITAB

MINITAB

Below we describe how to use MINITAB to do the F test to compare the variances of two populations. First we give some guidance on how MINITAB can be used to calculate p-values from either the χ^2 distribution or the F distribution, when the χ^2 statistic or the F statistic has been obtained using a calculator or with aid of a computer. At the end of Chapter 3, we showed how to use MINITAB to calculate sample standard deviations or sample variances.

Calculating *p*-values

At the end of Chapter 6, we showed how to use MINITAB to compute cumulative probabilities for the normal distribution. Similar steps can be used to obtain cumulative probabilities for the chi-squared or for the F distribution. These can then be used to calculate p-values for the tests described in the current chapter.

For example, in the Newcastle Metro Bus example in Section 11.1 (file 'BusTimes.MTW' on the accompanying CD), the chi-squared test statistic given by equation (11.8) is $\chi^2 = 28.18$. MINITAB can be used to compute an upper tail p-value (appropriate for H_0: $\sigma^2 \leq 4$ versus H_1: $\sigma^2 > 4$).

BUSTIMES

Step 1 **Calc > Probability Distributions > Chi-Square** [Main menu bar]

Step 2 Check **Cumulative Probability** [**Chi-Square** panel]
Enter **23** in the **Degrees of freedom** box
Check **Input constant** and enter **28.18** in the adjacent box
Click **OK**

This gives the cumulative probability 0.7909, which is the area under the curve to the left of $\chi^2 = 28.18$. The p-value is the upper tail area or probability, so p-value $= 1 - 0.7909 = 0.2091$. If the test is a lower tail test, the cumulative probability given by MINITAB is the p-value. For a two-tailed test, the p-value is double the lower or upper tail area, depending on whether the calculated χ^2 value is in the lower or upper tail area of the distribution.

A similar set of steps starting with **Calc > Probability Distributions > F** can be used to obtain p-values for the F distribution. In this case, the degrees of freedom for both the numerator and the denominator are entered at **Step 2**.

F-Test for two populations

We shall use the data for the Midlands Schools bus study in Section 11.2 (file 'SchoolBus.MTW' on the accompanying CD). The arrival times for Red Bus appear in column C1, and the arrival times for Route One appear in column C2. The following MINITAB procedure can be used to do the hypothesis test with hypotheses H_0: $\sigma_1^2 = \sigma_2^2$ and H_1: $\sigma_1^2 \neq \sigma_2^2$.

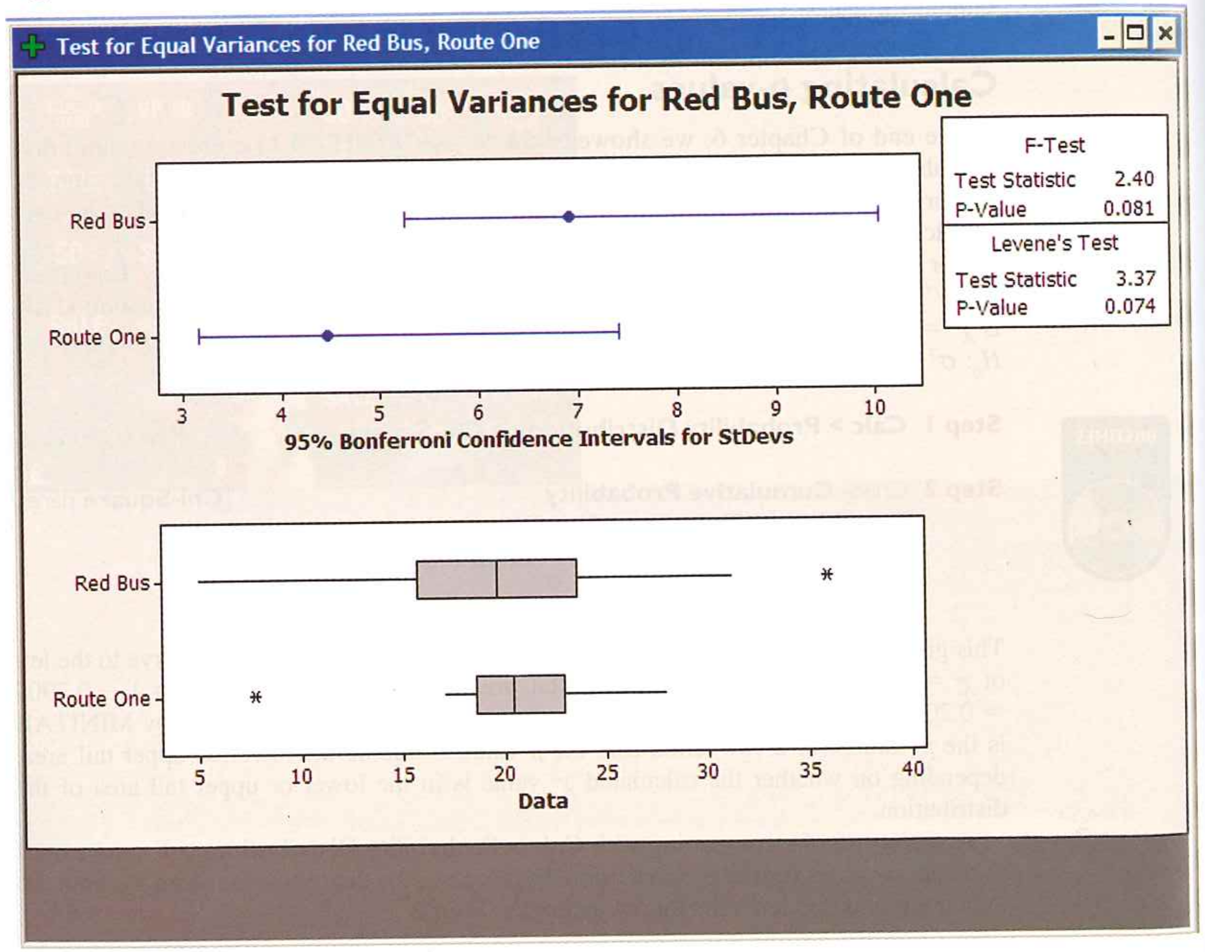
SCHOOLBUS

Step 1 **Stat > Basic Statistics > 2-Variances** [Main menu bar]

Step 2 Select **Samples in different columns** [**2-Variances** panel]
Enter **C1** in the **First** box
Enter **C2** in the **Second** box
Click **OK**

MINITAB produces both textual output in the Session window, and graphical output. The output refers to two tests, the first of which is the *F* test we have discussed in the present chapter. The graphical output is shown in Figure 11.5. This shows the test

Figure 11.5 MINITAB graphical output for the *F* test comparing two variances



statistic $F = 2.40$ and the *p*-value $= 0.081$. The *p*-value is for a two-tailed test. If this MINITAB routine is used for a one-tailed test, the two-tailed *p*-value should be halved to obtain the appropriate one-tailed *p*-value.

Population variances using EXCEL

EXCEL

Below we describe how to use EXCEL to do the *F* test to compare the variances of two populations. First we give some guidance on how EXCEL can be used to calculate *p*-values from either the χ^2 distribution or the *F* distribution, when the χ^2 statistic or the *F* statistic has been obtained using a calculator or with the aid of a computer. At the end of Chapter 3, we showed how to use EXCEL to calculate sample standard deviations or sample variances.

Calculating *p*-values

At the end of Chapter 6, we showed how to use EXCEL to compute probabilities for the normal distribution. EXCEL also has functions that give probabilities for the χ^2 and the *F* distributions. In the case of the χ^2 and *F* distributions, the EXCEL functions return probabilities in the right-hand tail area, so these are directly applicable as *p*-values for upper-tail tests.

BUSTIMES

For example, in the Newcastle Metro Bus example in Section 11.1 (file 'BusTimes.XLS' on the accompanying CD), the χ^2 test statistic given by equation (11.8) is $\chi^2 = 28.18$. The EXCEL function **CHIDIST**(χ^2, df) can be used to compute an upper tail *p*-value (appropriate for H_0: $\sigma^2 \leq 4$ versus H_1: $\sigma^2 > 4$). The first argument for the function is the value of the χ^2 test statistic, the second argument is the number of degrees of freedom. In this case, =CHIDIST(28.18, 23) will return the value 0.2091, which is the *p*-value for the test. If the test is a lower tail test, the probability given by EXCEL will need to be subtracted from 1 to obtain the *p*-value. For a two-tailed test, the *p*-value is double the lower or upper tail area, depending on whether the calculated χ^2 value is in the lower or upper tail area of the distribution.

The EXCEL function **FDIST**(F, df1, df2) returns the probability in the right-hand tail of the *F* distribution. The first argument for the function is the value of the *F* statistic, the second argument is the number of degrees of freedom for the numerator, and the third argument is the number of degrees of freedom for the denominator.

F-Test for two populations

SCHOOLBUS

We shall use the data for the Midlands Schools bus study in Section 11.2 (file 'SchoolBus.XLS' on the accompanying CD). The EXCEL worksheet has the label Red Bus in cell A1 and the label Route One in cell B1. The times for the Red Bus sample are in cells A2:A27 and the times for the Route One sample are in cells B2:B17. The steps to conduct the hypothesis test H_0: $\sigma_1^2 = \sigma_2^2$ versus H_1: $\sigma_1^2 \neq \sigma_2^2$ are as follows:

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

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

Figure 11.6 EXCEL output for the *F* test comparing two variances

D	E	F
F-Test Two-Sample for Variances		
	Red Bus	*Route One*
Mean	20.230769	20.24375
Variance	48.020615	19.999958
Observations	26	16
df	25	15
F	2.4010358	
P(F<=f) one-tail	0.0405271	
F Critical one-tail	2.2797293	

Step 3 Choose **F-Test Two-Sample for Variances**
Click **OK**

Step 4 Enter **A1:A27** in the **Variable 1 Range** box
[**F-Test Two Sample for Variances** panel]
Enter **B1:B17** in the **Variable 2 Range** box
Check **Labels**
Enter **0.05** in the **Alpha** box
(*Note:* This EXCEL procedure uses alpha as the area in the upper tail.)
Select **Output Range** and enter **D1** in the box
Click **OK**

The EXCEL results are shown in Figure 11.6. The output 'P(F <= f) one-tail' = 0.0405 is the one-tailed area associated with the test statistic $F = 2.40$. So the two-tailed p-value is $2(0.0405) = 0.081$. If the hypothesis test had been a one-tailed test, the one-tailed area in the cell labelled 'P(F <= f) one-tail' directly provides the information necessary to determine the p-value for the test.

Population variances using PASW

At the end of Chapter 6, we showed how to use PASW to compute cumulative probabilities for the normal distribution. Similar steps can be used to obtain cumulative probabilities for the chi-squared or for the F distribution. These can then be used to calculate p-values for the tests described in the current chapter. At the end of Chapter 3, we showed how to use PASW to calculate sample standard deviations or sample variances.

Calculating p-values for the χ^2 distribution

In the Newcastle Metro Bus example in Section 11.1 (file 'BusTimes. SAV' on the accompanying CD), the chi-squared test statistic given by equation (11.8) is $\chi^2 = 28.18$. PASW can be used to compute an upper tail p-value (appropriate for H_0: $\sigma^2 \leq 4$ versus H_1: $\sigma^2 > 4$).

BUSTIMES

Step 1 Enter **28.18** in the first row of the second column in the Data Editor [Data Editor]
Name this variable **Chisquared**

Step 2 **Transform > Compute** [Main menu bar]

Step 3 In the **Target Variable** box, enter the name **Cumprob**
[**Compute Variable** panel]
Select **CDF & Noncentral CDF** in the **Function Group** list
Double click on **Cdf:Chisq** in the **Functions and Special Variables** list, so that **CDF.CHISQ(?,?)** appears in the **Numeric Expression** box
Highlight the first argument (first question mark), then double-click on **Chisquared** in the variables list
Enter **23** (degrees of freedom) as the second argument (second question mark) – the **Numeric Expression** box should now have **CDF.CHISQ(Chisquared,23)**
Click **OK**

The value 0.7909 will be returned in the first row of column 3 in the Data Editor, now named **Cumprob**. This is a cumulative probability of 0.7909, i.e. the area under the curve to the left of $\chi^2 = 28.18$. The p-value is the upper tail area or probability, so p-value $= 1 - 0.7909 = 0.2091$. If the test is a lower tail test, the cumulative probability is the p-value. For a two-tailed test, the p-value is double the lower or upper tail area, depending on whether the calculated χ^2 value is in the lower or upper tail area of the distribution.

Calculating p-values for the F distribution

SCHOOLBUS

A similar set of steps to the above can be used to obtain p-values for the F distribution. We shall use the data for the Midlands Schools bus study in Section 11.2 (file 'SchoolBus.SAV' on the accompanying CD). The arrival times are in the first column of the PASW file, with codes for the two companies in the second column. Using PASW to calculate the variances yields 48.02 for the Red Bus arrival times and 20.00 for Route 1, giving an F ratio of 2.40.

Step 1 Enter **2.40** in the first row of the third column in the Data Editor [Data Editor]
Name this variable **Fratio**

Step 2 **Transform > Compute** [Main menu bar]

Step 3 In the **Target Variable** box, enter the name **Cumprob**
[**Compute Variable** panel]
Select **CDF & Noncentral CDF** in the **Function Group** list
Double click on **Cdf:F** in the **Functions and Special Variables** list, so that **CDF.CHISQ(?,?,?)** appears in the **Numeric Expression** box
Highlight the first argument (first question mark), then double-click on **Fratio** in the variables list

Enter **25** and **15** (degrees of freedom) as the second argument and third arguments (second and third question marks) – the **Numeric Expression** box should now have **CDF.CHISQ(Fratio,25,15)**
Click **OK**

The value 0.9594 will be returned in the first row of column 4 in the Data Editor, now named **Cumprob**. This is a cumulative probability, i.e. the area under the curve to the left of $F = 2.40$. The upper tail area is $1 - 0.9594 = 0.0406$. For a two-tailed test, as we did in Section 11.2, the p-value is twice this area, p-value $= 0.081$.

Chapter 12

Tests of Goodness of Fit and Independence