Step 2 Transfer the Rating variable to the **Test Variable(s)** box
[**One-Sample T Test** panel]
Enter **7** in the **Test Value** box
Click **OK**

The routine was illustrated in Chapter 8 using the credit card balance data for a sample of 85 households. The PASW results were displayed in Figure 8.10. Similar results are shown here in Figure 9.16 for the Munich Airport ratings, which are in the first column of the PASW data file ('AirRating.SAV' on the accompanying CD). The PASW routine constructs a two-tailed test. The *p*-value for a one-tailed test can be computed as half the two-tailed *p*-value shown in the output: 0.071/2 = 0.035.

Figure 9.16 PASW output for the Munich Airport rating hypothesis test

One-Sample Statistics

	N	Mean	Std. Deviation	Std. Error Mean
Rating	60	7.25	1.052	.136

One-Sample Test

	Test Value = 7					
					95% Confidence Interval of the Difference	
	t	df	Sig. (2-tailed)	Mean Difference	Lower	Upper
Rating	1.841	59	.071	.250	-.02	.52

Chapter 10

Statistical Inference About Means and Proportions with Two Populations

Learning objectives

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

1. Construct and interpret confidence intervals and hypothesis tests for the difference between two population means, given independent samples from the two populations:
   - 1.1 When the standard deviations of the two populations are known.
   - 1.2 When the standard deviations of the two populations are unknown.
2. Construct and interpret confidence intervals and hypothesis tests for the difference between two population means, given matched samples from the two populations.
3. Construct and interpret confidence intervals and hypothesis tests for the difference between two population proportions, given independent samples from the two populations.

In Chapters 8 and 9 we showed how to construct interval estimates and do hypothesis tests for situations involving a single population mean and a single population proportion. In this chapter we continue our discussion of statistical inference by showing how interval estimates and hypothesis tests can be developed for situations involving two populations, when the difference between the two population means or the two population proportions is of prime importance. For example, we may want to construct an interval estimate of the difference between the mean starting salary for a population of men and the mean starting salary for a population of women. Or we may want to conduct a hypothesis test to determine whether any difference is present between the proportion of defective parts in a population of parts produced by supplier A and the proportion of defective parts in a population of parts produced by supplier B. We begin our discussion of statistical inference about two populations by showing how to construct interval estimates and do hypothesis tests about the difference between the means of two populations when the standard deviations of the two populations are assumed known.

10.1 Inferences about the difference between two population means: σ_1 and σ_2 known

Let μ_1 denote the mean of population 1 and μ_2 denote the mean of population 2. We shall focus on inferences about the difference between the means: $\mu_1 - \mu_2$. To make an inference about this difference, we select a simple random sample of n_1 units from population 1 and a second simple random sample of n_2 units from population 2. The two samples, taken separately and independently, are referred to as independent simple random samples. In this section, we assume that information is available such that the two population standard deviations, σ_1 and σ_2, can be assumed known prior to collecting the samples. We refer to this situation as the σ_1 and σ_2 known case. In the following example we show how to compute a margin of error and develop an interval estimate of the difference between the two population means when σ_1 and σ_2 are known.

Statistics in Practice

Fisons Corporation

Fisons plc is a major company that manufactures pharmaceuticals, scientific equipment and horticultural products. The company's pharmaceutical division uses extensive statistical procedures to test and develop new drugs. The testing process usually consists of three stages: (1) pre-clinical testing, (2) testing for long-term usage and safety, and (3) clinical efficacy testing. At each successive stage, the chance that a drug will pass the rigorous tests decreases; however, the cost of further testing increases dramatically. Industry surveys indicate that on average the research and development for one new drug costs over €200 million and takes 12 years. Hence, it is important to eliminate unsuccessful new drugs in the early stages of the testing process, as well as identify promising ones for further testing. Statistics plays a major role in pharmaceutical research, where government regulations are stringent.

In pre-clinical testing, a two- or three-population statistical study is typically used to determine whether a new drug shows promise. The populations may consist of the new drug, a control, and a standard drug. The pre-clinical testing process begins when a new drug is sent to the pharmacology group for evaluation of efficacy – the capacity of the drug to produce the desired effects. As part of the process, a statistician is asked to design an experiment that can be used to test the new drug. The design must specify the sample size and the statistical methods of analysis. In a two-population study, one sample is used to obtain data on the efficacy of the new drug (population 1) and a second sample is used to obtain data on the efficacy of a standard drug (population 2). Depending on the intended use, the new and standard drugs are tested in such disciplines as neurology, cardiology and immunology. In most studies, the statistical method involves hypothesis testing for the difference between the means of the new drug population and the standard drug population. If a new drug lacks efficacy or produces undesirable effects in comparison with the standard drug, the new drug is rejected. Only new drugs that show promising comparisons with the standard drugs are forwarded to the long-term usage and safety testing programme.

Further data collection and multi-population studies are conducted in the long-term usage and safety testing programme and in the clinical testing programme. In the UK, the Medicines and Healthcare Products Regulatory Agency (MHRA) requires that statistical methods be defined prior to such testing to avoid data-related biases. In addition, to avoid human biases, some of the clinical trials are double or triple blind. That is, neither the participant nor the investigator knows what drug is administered to whom.

In this chapter you will learn how to construct interval estimates and do hypothesis tests about means and proportions with two populations. Techniques will be presented for analyzing independent random samples as well as matched samples.

Scientist doing pre-clinical testing on a new pharmaceutical drug. © Firefly Productions/First Collection.

Interval estimation of $\mu_1 - \mu_2$

Suppose a retailer such as Currys (selling TVs, DVD players, computers, photographic equipment and so on) operates two stores in Dublin: one is in the inner city and the other is in an out-of-town shopping centre. The regional manager noticed that products that sell well in one store do not always sell well in the other. The manager believes this may be attributable to differences in customer demographics at the two locations. Customers may differ in age, education, income and so on. Suppose the manager asks us to investigate the difference between the mean ages of the customers who shop at the two stores.

Let us define population 1 as all customers who shop at the inner-city store and population 2 as all customers who shop at the suburban store.

μ_1 = mean of population 1 (i.e. the mean age of all customers who shop at the inner-city store)

μ_2 = mean of population 2 (i.e. the mean age of all customers who shop at the out-of-town store)

The difference between the two population means is $\mu_1 - \mu_2$. To estimate $\mu_1 - \mu_2$, we shall select a simple random sample of n_1 customers from population 1 and a simple random sample of n_2 customers from population 2. We then compute the two sample means.

$\bar{x}_1$ = sample mean age for the simple random sample of n_1 inner-city customers

$\bar{x}_2$ = sample mean age for the simple random sample of n_2 out-of-town customers

The point estimator of the difference between the two populations is the difference between the sample means.

Point estimator of the difference between two population means

$$\bar{X}_1 - \bar{X}_2 \qquad (10.1)$$

Figure 10.1 provides an overview of the process used to estimate the difference between two population means based on two independent simple random samples.

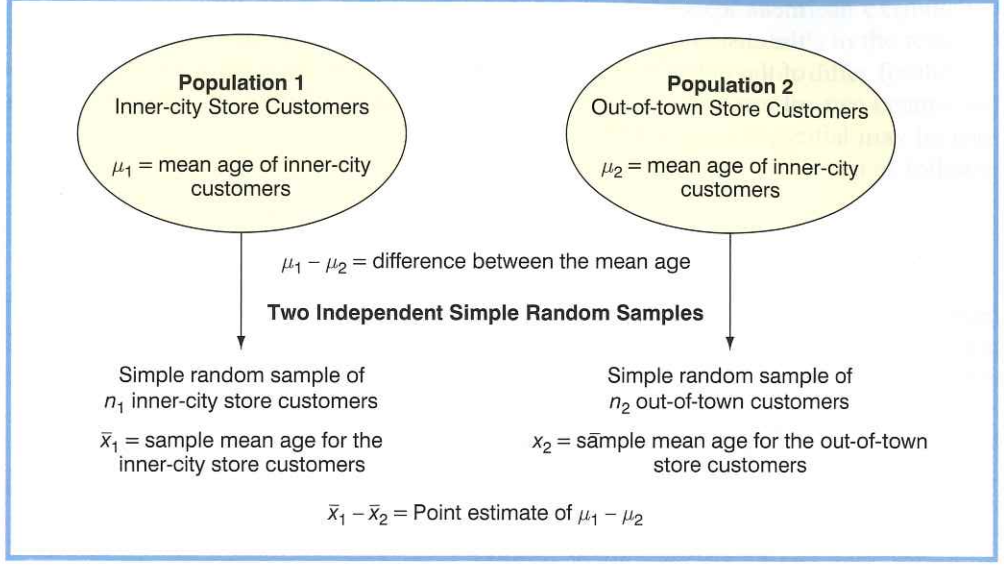
As with other point estimators, the point estimator $\bar{X}_1 - \bar{X}_2$ has a standard error that describes the variation in the sampling distribution of the estimator. With two independent simple random samples, the standard error of $\bar{X}_1 - \bar{X}_2$ is as follows:

Standard error of $\bar{X}_1 - \bar{X}_2$

$$\sigma_{\bar{X}_1 - \bar{X}_2} = \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} \qquad (10.2)$$

If both populations have a normal distribution, or if the sample sizes are large enough that the central limit theorem enables us to conclude that the sampling distributions of $\bar{X}_1$ and $\bar{X}_2$ can be approximated by a normal distribution, the sampling distribution of $\bar{X}_1 - \bar{X}_2$ will have a normal distribution with mean given by $\mu_1 - \mu_2$.

Figure 10.1 Estimating the difference between two population means



As we showed in Chapter 8, an interval estimate is given by a point estimate ± a margin of error. In the case of estimation of the difference between two population means, an interval estimate will take the form $(\bar{x}_1 - \bar{x}_2) \pm$ margin of error. When the sampling distribution of $\bar{X}_1 - \bar{X}_2$ is a normal distribution, we can write the margin of error as follows:

$$\text{Margin of error} = z_{\alpha/2}\sigma_{\bar{X}_1 - \bar{X}_2} = z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} \qquad (10.3)$$

Therefore the interval estimate of the difference between two population means is as follows:

Interval estimate of the difference between two population means: σ_1 and σ_2 known

$$(\bar{x}_1 - \bar{x}_2) \pm z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} \qquad (10.4)$$

where $1 - \alpha$ is the confidence coefficient.

Let us return to the example of the Dublin retailer. Based on data from previous customer demographic studies, the two population standard deviations are known with $\sigma_1 = 9$ years and $\sigma_2 = 10$ years. The data collected from the two independent simple random samples of the retailer's customers provided the following results.

	Inner city store	Out-of-town store
Sample size	$n_1 = 36$	$n_2 = 49$
Sample mean	$\bar{x}_1 = 40$ years	$\bar{x}_2 = 35$ years

Using expression (10.1), we find that the point estimate of the difference between the mean ages of the two populations is $\bar{x}_1 - \bar{x}_2 = 40 - 35 = 5$ years. We estimate that the customers at the inner-city store have a mean age five years greater than the mean age of the out-of-town customers. We can now use expression (10.4) to compute the margin of error and provide the interval estimate of $\mu_1 - \mu_2$. Using 95 per cent confidence and $z_{\alpha/2} = z_{0.025} = 1.96$, we have

$$(\bar{x}_1 - \bar{x}_2) \pm z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} = (40 - 35) \pm 1.96\sqrt{\frac{9^2}{36} + \frac{10^2}{49}} = 5 \pm 4.1$$

The margin of error is 4.1 years and the 95 per cent confidence interval estimate of the difference between the two population means is $5 - 4.1 = 0.9$ years to $5 + 4.1 = 9.1$ years.

Hypothesis tests about $\mu_1 - \mu_2$

Let us consider hypothesis tests about the difference between two population means. Using D_0 to denote the hypothesized difference between μ_1 and μ_2, the three forms for a hypothesis test are as follows:

$$\begin{array}{lll} H_0: \mu_1 - \mu_2 \geq D_0 & H_0: \mu_1 - \mu_2 \leq D_0 & H_0: \mu_1 - \mu_2 = D_0 \\ H_1: \mu_1 - \mu_2 < D_0 & H_1: \mu_1 - \mu_2 > D_0 & H_1: \mu_1 - \mu_2 \neq D_0 \end{array}$$

In many applications, $D_0 = 0$. Using the two-tailed test as an example, when $D_0 = 0$ the null hypothesis is H_0: $\mu_1 - \mu_2 = 0$, i.e. the null hypothesis is that μ_1 and μ_2 are equal. Rejection of H_0 leads to the conclusion that H_1: $\mu_1 - \mu_2 \neq 0$ is true, i.e. μ_1 and μ_2 are not equal.

The steps for doing hypothesis tests presented in Chapter 9 are applicable here. We must choose a level of significance, compute the value of the test statistic and find the p-value to determine whether the null hypothesis should be rejected. With two independent simple random samples, we showed that the point estimator $\bar{X}_1 - \bar{X}_2$ has a standard error $\sigma_{\bar{X}_1 - \bar{X}_2}$ given by expression (10.2), and the distribution of $\bar{X}_1 - \bar{X}_2$ can be described by a normal distribution. In this case, the test statistic for the difference between two population means when σ_1 and σ_2 are known is as follows.

Test statistic for hypothesis tests about $\mu_1 - \mu_2$: σ_1 and σ_2 known

$$z = \frac{(\bar{x}_1 - \bar{x}_2) - D_0}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \qquad \textbf{(10.5)}$$

Here is an example. As part of a study to evaluate differences in education quality between two training centres, a standardized examination is given to individuals who are trained at the centres. The difference between the mean examination scores is used to assess quality differences between the centres. The population means for the two centres are as follows.

μ_1 = the mean examination score for the population of individuals trained at centre A

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

EXAMSCORES

We begin with the tentative assumption that no difference exists between the average training quality provided at the two centres. Hence, in terms of the mean examination scores, the null hypothesis is that $\mu_1 - \mu_2 = 0$. If sample evidence leads to the rejection of this hypothesis, we shall conclude that the mean examination scores differ for the two populations. This conclusion indicates a quality differential between the two centres and suggests that a follow-up study investigating the reason for the differential may be warranted. The null and alternative hypotheses for this two-tailed test are written as follows.

$$\begin{array}{l} H_0: \mu_1 - \mu_2 = 0 \\ H_1: \mu_1 - \mu_2 \neq 0 \end{array}$$

The standardized examination given previously in a variety of settings always resulted in an examination score standard deviation near 10 points. We shall use this information to assume that the population standard deviations are known with $\sigma_1 = 10$ and $\sigma_2 = 10$. An $\alpha = 0.05$ level of significance is specified for the study.

Independent simple random samples of $n_1 = 30$ individuals from training centre A and $n_2 = 40$ individuals from training centre B are taken. The respective sample means are $\bar{x}_1 = 82$ and $\bar{x}_2 = 78$. Do these data suggest a difference between the population means at the two training centres? To help answer this question, we compute the test statistic using equation (10.5).

$$z = \frac{(\bar{x}_1 - \bar{x}_2) - D_0}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} = \frac{(82 - 78) - 0}{\sqrt{\frac{(10)^2}{30} + \frac{(10)^2}{40}}} = 1.66$$

Next let us compute the p-value for this two-tailed test. Because the test statistic z is in the upper tail, we first compute the area under the curve to the right of $z = 1.66$. Using the standard normal distribution table, the cumulative probability for $z = 1.66$ is 0.9515, so the area in the upper tail of the distribution is $1 - 0.9515 = 0.0485$. Because this test is a two-tailed test, we must double the tail area: p-value $= 2(0.0485) = 0.0970$. Following the usual rule to reject H_0 if p-value $\leq \alpha$, we see that the p-value of 0.0970 does not allow us to reject H_0 at the 0.05 level of significance. The sample results do not provide sufficient evidence to conclude that the training centres differ in quality.

In this chapter we shall use the p-value approach to hypothesis testing as described in Chapter 9. However, if you prefer, the test statistic and the critical value rejection rule may be used. With $\alpha = 0.05$ and $z_{\alpha/2} = z_{0.025} = 1.96$, the rejection rule using the critical value approach would be to reject H_0 if $z \leq -1.96$ or if $z \geq 1.96$. With $z = 1.66$, we reach the same 'do not reject H_0' conclusion.

In the preceding example, we demonstrated a two-tailed hypothesis test about the difference between two population means. Lower tail and upper tail tests can also be considered. These tests use the same test statistic as given in equation (10.5). The procedure for computing the p-value and the rejection rules for these one-tailed tests are the same as those presented in Chapter 9.

Practical advice

In most applications of the interval estimation and hypothesis testing procedures presented in this section, random samples with $n_1 \geq 30$ and $n_2 \geq 30$ are adequate. In cases where either or both sample sizes are less than 30, the distributions of the populations become important considerations. In general, with smaller sample sizes, it is more important for the analyst to be satisfied that it is reasonable to assume the distributions of the two populations are at least approximately normal.

Exercises

Methods

1 Consider the following results for two independent random samples taken from two populations.

Sample 1	Sample 2
$n_1 = 50$	$n_2 = 35$
$\bar{x}_1 = 13.6$	$\bar{x}_2 = 11.6$
$\sigma_1 = 2.2$	$\sigma_2 = 3.0$

a. What is the point estimate of the difference between the two population means?
b. Provide a 90 per cent confidence interval for the difference between the two population means.
c. Provide a 95 per cent confidence interval for the difference between the two population means.

2 Consider the following hypothesis test.

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

The following results are for two independent samples taken from the two populations.

Sample 1	Sample 2
$n_1 = 40$	$n_2 = 50$
$\bar{x}_1 = 25.2$	$\bar{x}_2 = 22.8$
$\sigma_1 = 5.2$	$\sigma_2 = 6.0$

a. What is the value of the test statistic?
b. What is the *p*-value?
c. With $\alpha = 0.05$, what is your hypothesis testing conclusion?

3 Consider the following hypothesis test.

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

The following results are for two independent samples taken from the two populations.

Sample 1	Sample 2
$n_1 = 80$	$n_2 = 70$
$\bar{x}_1 = 104$	$\bar{x}_2 = 106$
$\sigma_1 = 8.4$	$\sigma_2 = 7.6$

a. What is the value of the test statistic?
b. What is the *p*-value?
c. With $\alpha = 0.05$, what is your hypothesis testing conclusion?

Applications

4 A study of wage differentials between men and women reported that one of the reasons wages for men are higher than wages for women is that men tend to have more years of work experience than women. Assume that the following sample summaries show the years of experience for each group.

Men	Women
$n_1 = 100$	$n_2 = 85$
$\bar{x}_1 = 14.9$ years	$\bar{x}_2 = 10.3$ years
$\sigma_1 = 5.2$ years	$\sigma_2 = 3.8$ years

a. What is the point estimate of the difference between the two population means?
b. At 95 per cent confidence, what is the margin of error?
c. What is the 95 per cent confidence interval estimate of the difference between the two population means?

5 The Dublin retailer age study (used as an example above) provided the following data on the ages of customers from independent random samples taken at the two store locations.

Inner-city store	Out-of-town store
$n_1 = 36$	$n_2 = 49$
$\bar{x}_1 = 40$ years	$\bar{x}_2 = 35$ years
$\sigma_1 = 9$ years	$\sigma_2 = 10$ years

a. State the hypotheses that could be used to detect a difference between the population mean ages at the two stores.
b. What is the value of the test statistic?
c. What is the *p*-value
d. At $\alpha = 0.05$, what is your conclusion?

6 According to a report in *USA Today* on 13 February 2006, the average expenditure on Valentine's Day (14 February) was expected to be about $101. Do male and female consumers differ in the amounts they spend? The average expenditure in a sample survey of 40 male consumers was $135.67, and the average expenditure in a sample survey of 30 female consumers was $68.64. Based on past surveys, the standard deviation for male consumers is assumed to be $35, and the standard deviation for female consumers is assumed to be $20.

a. What is the point estimate of the difference between the population mean expenditure for males and the population mean expenditure for females?
b. At 99 per cent confidence, what is the margin of error?
c. Construct a 99 per cent confidence interval for the difference between the two population means.

10.2 Inferences about the difference between two population means: σ_1 and σ_2 unknown

In this section we extend the discussion of inferences about the difference between two population means to the case when the two population standard deviations, σ_1 and σ_2, are unknown. In this case, we shall use the sample standard deviations, s_1 and s_2, to estimate the unknown population standard deviations. When we use the sample standard deviations, the interval estimation and hypothesis testing procedures will be based on the *t* distribution rather than the standard normal distribution.

Interval estimation of $\mu_1 - \mu_2$

In the following example we show how to compute a margin of error and construct an interval estimate of the difference between two population means when σ_1 and σ_2 are unknown. The Union Bank is conducting a study designed to identify differences between cheque account practices by customers at two of its branches. A simple random sample of 28 cheque accounts is selected from the Northern Branch and an independent simple random sample of 22 cheque accounts is selected from the Eastern Branch. The current cheque account balance is recorded for each of the accounts. A summary of the account balances follows:

CHEQACCT

	Northern	Eastern
Sample size	$n_1 = 28$	$n_2 = 22$
Sample mean	$\bar{x}_1 = €1025$	$x_2 = €910$
Sample standard deviation	$s_1 = €150$	$s_2 = €125$

The Union Bank would like to estimate the difference between the mean cheque account balance maintained by the population of Northern customers and the population of Eastern customers. Let us develop the margin of error and an interval estimate of the difference between these two population means.

In Section 10.1, we provided the following interval estimate for the case when the population standard deviations, σ_1 and σ_2, are known.

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

With σ_1 and σ_2 unknown, we shall use the sample standard deviations s_1 and s_2 to estimate σ_1 and σ_2 and replace $z_{\alpha/2}$ with $t_{\alpha/2}$. As a result, the interval estimate of the difference between two population means is given by the following expression:

Interval estimate of the difference between two population means: σ_1 and σ_2 unknown

$$(\bar{x}_1 - \bar{x}_2) \pm t_{\alpha/2}\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} \qquad (10.6)$$

where $1 - \alpha$ is the confidence coefficient.

In this expression, the use of the *t* distribution is an approximation, but it provides excellent results and is relatively easy to use. The only difficulty that we encounter in using expression (10.6) is determining the appropriate degrees of freedom for $t_{\alpha/2}$. Statistical software packages compute the appropriate degrees of freedom automatically. The formula used is as follows:

Degrees of freedom for the t distribution using two independent random samples

$$df = \frac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)^2}{\left(\frac{1}{n_1 - 1}\right)\left(\frac{s_1^2}{n_1}\right)^2 + \left(\frac{1}{n_2 - 1}\right)\left(\frac{s_2^2}{n_2}\right)^2} \qquad (10.7)$$

Let us return to the Union Bank example and show how to use expression (10.6) to provide a 95 per cent confidence interval estimate of the difference between the population mean cheque account balances at the two branches. The sample data show $n_1 = 28$, $\bar{x}_1 = €1025$, and $s_1 = €150$ for the Northern Branch, and $n_2 = 22$, $\bar{x}_2 = €910$, and $s_2 = €125$ for the Eastern Branch. The calculation for degrees of freedom for $t_{\alpha/2}$ is as follows:

$$df = \frac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)^2}{\left(\frac{1}{n_1 - 1}\right)\left(\frac{s_1^2}{n_1}\right)^2 + \left(\frac{1}{n_2 - 1}\right)\left(\frac{s_2^2}{n_2}\right)^2} = \frac{\left(\frac{150^2}{28} + \frac{125^2}{22}\right)^2}{\left(\frac{1}{28 - 1}\right)\left(\frac{150^2}{28}\right)^2 + \left(\frac{1}{22 - 1}\right)\left(\frac{125^2}{22}\right)^2} = 47.8$$

We round the non-integer degrees of freedom *down* to 47 to provide a larger *t*-value and a more conservative interval estimate. Using the *t* distribution table with 47 degrees of freedom, we find $t_{0.025} = 2.012$. Using expression (10.6), we construct the 95 per cent confidence interval estimate of the difference between the two population means as follows.

$$(\bar{x}_1 - \bar{x}_2) \pm t_{\alpha/2}\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} = (1025 - 910) \pm 2.012\sqrt{\frac{150^2}{28} + \frac{125^2}{22}} = 115 \pm 78$$

The point estimate of the difference between the population mean cheque account balances at the two branches is €115. The margin of error is €78, and the 95 per cent confidence interval estimate of the difference between the two population means is 115 − 78 = €37 to 115 + 78 = €193.

The computation of the degrees of freedom (equation (10.7)) is cumbersome if you are doing the calculation by hand, but it is easily implemented with a computer software package. However, note that the expressions s_1^2/n_1 and s_2^2/n_2 appear in both expression (10.6) and equation (10.7). These values only need to be computed once in order to evaluate both (10.6) and (10.7).

Hypothesis tests about $\mu_1 - \mu_2$

Let us now consider hypothesis tests about the difference between the means of two populations when the population standard deviations σ_1 and σ_2 are unknown. Letting D_0 denote the hypothesized difference between μ_1 and μ_2, Section 10.1 showed that the test statistic used for the case where σ_1 and σ_2 are known is as follows. The test statistic, z, follows the standard normal distribution.

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

When σ_1 and σ_2 are unknown, we use s_1 as an estimator of σ_1 and s_2 as an estimator of σ_2. Substituting these sample standard deviations for σ_1 and σ_2 gives the following test statistic when σ_1 and σ_2 are unknown.

Test statistic for hypothesis tests about $\mu_1 - \mu_2$: σ_1 and σ_2 unknown

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - D_0}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} \quad \textbf{(10.8)}$$

The degrees of freedom for t are given by equation (10.7).

Consider an example, involving a new computer software package developed to help systems analysts reduce the time required to design, develop, and implement an information system. To evaluate the benefits of the new software package, a random sample of 24 systems analysts is selected. Each analyst is given specifications for a hypothetical information system. Then 12 of the analysts are instructed to produce the information system by using current technology. The other 12 analysts are trained in the use of the new software package and then instructed to use it to produce the information system.

This study involves two populations: a population of systems analysts using the current technology and a population of systems analysts using the new software package. In terms of the time required to complete the information system design project, the population means are as follow.

μ_1 = the mean project completion time for systems analysts using the current technology

μ_2 = the mean project completion time for systems analysts using the new software package

The researcher in charge of the new software evaluation project hopes to show that the new software package will provide a shorter mean project completion time, i.e. the researcher is looking for evidence to conclude that μ_2 is less than μ_1. In this case, the difference between the two population means, $\mu_1 - \mu_2$, will be greater than zero. The research hypothesis $\mu_1 - \mu_2 > 0$ is stated as the alternative hypothesis. The hypothesis test becomes

SOFTWARE TEST

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

We shall use $\alpha = 0.05$ as the level of significance. Suppose that the 24 analysts complete the study with the results shown in Table 10.1.

Table 10.1 Completion time data and summary statistics for the software testing study

	Current technology	New software
	300	274
	280	220
	344	308
	385	336
	372	198
	360	300
	288	315
	321	258
	376	318
	290	310
	301	332
	283	263
Summary statistics		
Sample size	$n_1 = 12$	$n_2 = 12$
Sample mean	$\bar{x}_1 = 325$ hours	$\bar{x}_2 = 286$ hours
Sample standard deviation	$s_1 = 40$	$s_2 = 44$

Using the test statistic in equation (10.8), we have

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - D_0}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} = \frac{(325 - 286) - 0}{\sqrt{\frac{40^2}{12} + \frac{44^2}{12}}} = 2.27$$

Computing the degrees of freedom using equation (10.7), we have

$$df = \frac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)^2}{\left(\frac{1}{n_1 - 1}\right)\left(\frac{s_1^2}{n_1}\right)^2 + \left(\frac{1}{n_2 - 1}\right)\left(\frac{s_2^2}{n_2}\right)^2} = \frac{\left(\frac{40^2}{12} + \frac{44^2}{12}\right)^2}{\frac{1}{12 - 1}\left(\frac{40^2}{12}\right)^2 + \left(\frac{1}{12 - 1}\right)\left(\frac{44^2}{12}\right)^2} = 21.8$$

Rounding down, we shall use a t distribution with 21 degrees of freedom. This row of the t distribution table is as follows:

Area in upper tail	0.20	0.10	0.05	0.025	0.01	0.005
t value (21 df)	0.859	1.323	1.721	2.080	2.518	2.831

$t = 2.27$

With an upper tail test, the p-value is the area in the upper tail to the right of $t = 2.27$. From the above results, we see that the p-value is between 0.025 and 0.01. Hence, the p-value is less than $\alpha = 0.05$ and H_0 is rejected. The sample results enable the researcher to conclude that $\mu_1 - \mu_2 > 0$, or $\mu_1 > \mu_2$. The research study supports the conclusion that the new software package provides a smaller population mean completion time.

Practical advice

The interval estimation and hypothesis testing procedures presented in this section are robust and can be used with relatively small sample sizes. In most applications, equal or nearly equal sample sizes such that the total sample size $n_1 + n_2$ is at least 20 can be expected to provide very good results even if the populations are not normal. Larger sample sizes are recommended if the distributions of the populations are highly skewed or contain outliers. Smaller sample sizes should only be used if the analyst is satisfied that the distributions of the populations are at least approximately normal.

Another approach sometimes used to make inferences about the difference between two population means when σ_1 and σ_2 are unknown is based on the assumption that the two population standard deviations are equal. You will find this approach as an option in MINITAB, PASW and EXCEL. Under the assumption of equal population variances, the two sample standard deviations are combined to provide the following 'pooled' sample variance s^2:

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

The t test statistic becomes

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - D_0}{s\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

and has $n_1 + n_2 - 2$ degrees of freedom. At this point, the computation of the p-value and the interpretation of the sample results are identical to the procedures discussed earlier in this section. A difficulty with this procedure is that the assumption that the two population standard deviations are equal is usually difficult to verify. Unequal population standard deviations are frequently encountered. Using the pooled procedure may not provide satisfactory results especially if the sample sizes n_1 and n_2 are quite different. The t procedure that we presented in this section does not require the assumption of equal population standard deviations and can be applied whether the population standard deviations are equal or not. It is a more general procedure and is recommended for most applications.

Exercises

Methods

7 Consider the following results for independent random samples taken from two populations.

Sample 1	Sample 2
$n_1 = 20$	$n_2 = 30$
$\bar{x}_1 = 22.5$	$\bar{x}_2 = 20.1$
$s_1 = 2.5$	$s_2 = 4.8$

a. What is the point estimate of the difference between the two population means?
b. What are the degrees of freedom for the t distribution?
c. At 95 per cent confidence, what is the margin of error?
d. What is the 95 per cent confidence interval for the difference between the two population means?

8 Consider the following hypothesis test.

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

The following results are from independent samples taken from two populations.

Sample 1	Sample 2
$n_1 = 35$	$n_2 = 40$
$\bar{x}_1 = 13.6$	$\bar{x}_2 = 10.1$
$s_1 = 5.2$	$s_2 = 8.5$

a. What is the value of the test statistic?
b. What are the degrees of freedom for the t distribution?
c. What is the p-value?
d. At $\alpha = 0.05$, what is your conclusion?

9 Consider the following data for two independent random samples taken from two normal populations.

Sample 1	10	7	13	7	9	8
Sample 2	8	7	8	4	6	9

a. Compute the two sample means.
b. Compute the two sample standard deviations.
c. What is the point estimate of the difference between the two population means?
d. What is the 90 per cent confidence interval estimate of the difference between the two population means?

Applications

AIRPORTS

10 The International Air Transport Association surveyed business travellers to determine ratings of various international airports. The maximum possible score was 10. Suppose 50 business travellers were asked to rate airport L and 50 other business travellers were asked to rate airport M. The rating scores follow.

Airport L

10 9 6 7 8 7 9 8 10 7 6 5 7 3 5 6 8 7 10 8 4 7 8 6 9
9 5 3 1 8 9 6 8 5 4 6 10 9 8 3 2 7 9 5 3 10 3 5 10 8

Airport M

6 4 6 8 7 7 6 3 3 8 10 4 8 7 8 7 5 9 5 8 4 3 8 5 5
4 4 4 8 4 5 6 2 5 9 9 8 4 8 9 9 5 9 7 8 3 10 8 9 6

Construct a 95 per cent confidence interval estimate of the difference between the mean ratings of the airports L and M.

11 Suppose independent random samples of 15 unionized women and 20 non-unionized women in a skilled manufacturing job provide the following hourly wage rates (€).

UNION

Union workers

22.40	18.90	16.70	14.05	16.20	20.00	16.10	16.30	19.10	16.50
18.50	19.80	17.00	14.30	17.20					

Non-union workers

17.60	14.40	16.60	15.00	17.65	15.00	17.55	13.30	11.20	15.90
19.20	11.85	16.65	15.20	15.30	17.00	15.10	14.30	13.90	14.50

a. What is the point estimate of the difference between mean hourly wages for the two populations?
b. Develop a 95 per cent confidence interval estimate of the difference between the two population means.
c. Does there appear to be any difference in the mean wage rate for these two groups? Explain.

12 The College Board provided comparisons of Scholastic Aptitude Test (SAT) scores based on the highest level of education attained by the test taker's parents. A research hypothesis was that students whose parents had attained a higher level of education would on average score higher on the SAT. During 2003, the overall mean SAT verbal score was 507 (*The World Almanac 2004*). SAT verbal scores for independent samples of students follow. The first sample shows the SAT verbal test scores for students whose parents are college graduates with a bachelor's degree. The second sample shows the SAT verbal test scores for students whose parents are high school graduates but do not have a college degree.

Student's Parents			
College Grads		High School Grads	
485	487	442	492
534	533	580	478
650	526	479	425
554	410	486	485
550	515	528	390
572	578	524	535
497	448		
592	469		

a. Formulate the hypotheses that can be used to determine whether the sample data support the hypothesis that students show a higher population mean verbal score on the SAT if their parents attained a higher level of education.
b. What is the point estimate of the difference between the means for the two populations?
c. Compute the *p*-value for the hypothesis test.
d. At $\alpha = 0.05$, what is your conclusion?

13 Periodically, Merrill Lynch customers are asked to evaluate Merrill Lynch financial consultants and services (2000 Merrill Lynch Client Satisfaction Survey). Higher ratings on the client satisfaction survey indicate better service, with 7 the maximum service rating. Independent samples of service ratings for two financial consultants are summarized here. Consultant A has ten years of experience while consultant B has one year of experience. Use $\alpha = 0.05$ and test to see whether the consultant with more experience has the higher population mean service rating.

Consultant A	Consultant B
$n_1 = 16$	$n_2 = 10$
$\bar{x}_1 = 6.82$	$\bar{x}_2 = 6.25$
$s_1 = 0.64$	$s_2 = 0.75$

a. State the null and alternative hypotheses.
b. Compute the value of the test statistic.
c. What is the *p*-value?
d. What is your conclusion?

14 Safegate Foods is redesigning the checkouts in its supermarkets throughout the country and is considering two designs. Tests on customer checkout times conducted at two stores where the two new systems have been installed result in the following summary of the data.

System A	System B
$n_1 = 120$	$n_2 = 100$
$\bar{x}_1 = 4.1$ minutes	$\bar{x}_2 = 3.4$ minutes
$s_1 = 2.2$ minutes	$s_2 = 1.5$ minutes

Test at the 0.05 level of significance to determine whether the population mean checkout times of the two systems differ. Which system is preferred?

15 Samples of final examination scores for two statistics classes with different instructors provided the following results.

Instructor A	Instructor B
$n_1 = 12$	$n_2 = 15$
$\bar{x}_1 = 72$	$\bar{x}_2 = 78$
$s_1 = 8$	$s_2 = 10$

With $\alpha = 0.05$, test whether these data are sufficient to conclude that the population mean grades for the two classes differ.

16 Educational testing companies provide tutoring, classroom learning, and practice tests in an effort to help students perform better on tests such as the Scholastic Aptitude Test (SAT). The test preparation companies claim that their courses will improve SAT score performances by an average of 120 points (*The Wall Street Journal*, 23 January 2003). A researcher is uncertain of this claim and believes that 120 points may be an overstatement in an effort to encourage students to take the test preparation course. In an evaluation study of one test preparation service, the researcher collects SAT score data for 35 students who took the test preparation course and 48 students who did not take the course.

	Course	No course
Sample mean	1058	983
Sample standard deviation	90	105

a. Formulate the hypotheses that can be used to test the researcher's belief that the improvement in SAT scores may be less than the stated average of 120 points.
b. Use $\alpha = 0.05$ and the data above. What is your conclusion?
c. What is the point estimate of the improvement in the average SAT scores provided by the test preparation course? Provide a 95 per cent confidence interval estimate of the improvement.
d. What advice would you have for the researcher after seeing the confidence interval?

10.3 Inferences about the difference between two population means: matched samples

Suppose employees at a manufacturing company can use two different methods to perform a production task. To maximize production output, the company wants to identify the method with the smaller population mean completion time. Let μ_1 denote the population mean completion time for production method 1 and μ_2 denote the population mean completion time for production method 2. With no preliminary indication of the preferred production method, we begin by tentatively assuming that the two production methods have the same population mean completion time. The null hypothesis is H_0: $\mu_1 - \mu_2 = 0$. If this hypothesis is rejected, we can conclude that the population mean completion times differ. In this case, the method providing the smaller mean completion time would be recommended. The null and alternative hypotheses are written as follows.

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

In choosing the sampling procedure that will be used to collect production time data and test the hypotheses, we consider two alternative designs. One is based on **independent samples** and the other is based on **matched samples**.

1 *Independent sample design*: A simple random sample of workers is selected and each worker in the sample uses method 1. A second independent simple random sample of workers is selected and each worker in this sample uses method 2. The test of the difference between population means is based on the procedures in Section 10.2.
2 *Matched sample design*: One simple random sample of workers is selected. Each worker first uses one method and then uses the other method. The order of the two methods is assigned randomly to the workers, with some workers performing method 1 first and others performing method 2 first. Each worker provides a pair of data values, one value for method 1 and another value for method 2.

In the matched sample design the two production methods are tested under similar conditions (i.e. with the same workers). Hence this design often leads to a smaller sampling error than the independent sample design. The primary reason is that in a matched sample design, variation between workers is eliminated because the same workers are used for both production methods.

Let us demonstrate the analysis of a matched sample design by assuming it is the method used to test the difference between population means for the two production

Table 10.2 Task completion times for a matched sample design

Worker	Completion time for Method 1 (minutes)	Completion time for Method 2 (minutes)	Difference in completion times (d_i)
1	6.0	5.4	0.6
2	5.0	5.2	−0.2
3	7.0	6.5	0.5
4	6.2	5.9	0.3
5	6.0	6.0	0.0
6	6.4	5.8	0.6

methods. A random sample of six workers is used. The data on completion times for the six workers are given in Table 10.2. Note that each worker provides a pair of data values, one for each production method. Also note that the last column contains the difference in completion times d_i for each worker in the sample.

The key to the analysis of the matched sample design is to realize that we consider only the column of differences. Therefore, we have six data values (0.6, −0.2, 0.5, 0.3, 0.0, 0.6) that will be used to analyze the difference between population means of the two production methods.

Let μ_d = the mean of the *difference* values for the population of workers. With this notation, the null and alternative hypotheses are rewritten as follows.

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

If H_0 is rejected, we can conclude that the population mean completion times differ. The d notation is a reminder that the matched sample provides *difference* data. The sample mean and sample standard deviation for the six difference values in Table10.2 follow. Other than the use of the d notation, the formulae for the sample mean and sample standard deviation are the same ones used previously in the text.

$$\bar{d} = \frac{\Sigma d_i}{n} = \frac{1.8}{6} = 0.30$$

$$s_d = \sqrt{\frac{\Sigma(d_i - \bar{d})^2}{n-1}} = \sqrt{\frac{0.56}{5}} = 0.335$$

With the small sample of $n = 6$ workers, we need to make the assumption that the population of differences has a normal distribution. This assumption is necessary so that we may use the t distribution for hypothesis testing and interval estimation procedures. Sample size guidelines for using the t distribution were presented in Chapters 8 and 9. Based on this assumption, the following test statistic has a t distribution with $n - 1$ degrees of freedom.

Test statistic for hypothesis test involving matched samples

$$t = \frac{\bar{d} - \mu_d}{s_d/\sqrt{n}} \qquad \textbf{(10.9)}$$

Let us use equation (10.9) to test the hypotheses H_0: $\mu_d = 0$ and H_1: $\mu_d \neq 0$, using $\alpha = 0.05$. Substituting the sample results $\bar{d} = 0.30$, $s_d = 0.335$, and $n = 6$ into equation (10.9), we compute the value of the test statistic.

$$t = \frac{\bar{d} - \mu_d}{s_d/\sqrt{n}} = \frac{0.30 - 0}{0.335/\sqrt{6}} = 2.20$$

Now let us compute the p-value for this two-tailed test. Because $t = 2.20 > 0$, the test statistic is in the upper tail of the t distribution. With $t = 2.20$, the area in the upper tail to the right of the test statistic can be found by using the t distribution table with degrees of freedom $= n - 1 = 6 - 1 = 5$. Information from the 5 degrees of freedom row of the t distribution table is as follows:

Area in upper tail	0.20	0.10	0.05	0.025	0.01	0.005
t value (5 df)	0.920	1.476	2.015	2.571	3.365	4.032

$t = 2.20$

We see that the area in the upper tail is between 0.05 and 0.025. Because this test is a two-tailed test, we double these values to conclude that the p-value is between 0.10 and 0.05. This p-value is greater than $\alpha = 0.05$, so the null hypothesis H_0: $\mu_d = 0$ is not rejected. MINITAB, EXCEL and PASW show the p-value as 0.080.

In addition we can obtain an interval estimate of the difference between the two population means by using the single population methodology of Chapter 8. At 95 per cent confidence, the calculation follows.

$$\bar{d} \pm t_{0.025} \frac{s_d}{\sqrt{n}} = 0.30 \pm 2.527 \left(\frac{0.335}{\sqrt{6}}\right) = 0.3 \pm 0.35$$

The margin of error is 0.35 and the 95 per cent confidence interval for the difference between the population means of the two production methods is − 0.05 minutes to 0.65 minutes.

In the example presented in this section, workers performed the production task with first one method and then the other method. This example illustrates a matched sample design in which each sampled element (worker) provides a pair of data values. It is also possible to use different but 'similar' elements to provide the pair of data values. For example, a worker at one location could be matched with a similar worker at another location (similarity based on age, education, gender, experience, etc.). The pairs of workers would provide the difference data that could be used in the matched sample analysis. A matched sample procedure for inferences about two population means generally provides better precision than the independent samples approach, therefore it is the recommended design. However, in some applications matching is not feasible, or perhaps the time and cost associated with matching are excessive. In such cases, the independent samples design should be used.

Exercises

Methods

17 Consider the following hypothesis test.

$$H_0: \mu_d \leq 0$$
$$H_1: \mu_d > 0$$

The following data are from matched samples taken from two populations.

Element	Population 1	Population 2
1	21	20
2	28	26
3	18	18
4	20	20
5	26	24

a. Compute the difference value for each element.
b. Compute $\bar{d}$.
c. Compute the standard deviation s_d.
d. Conduct a hypothesis test using $\alpha = 0.05$. What is your conclusion?

18 The following data are from matched samples taken from two populations.

Element	Population 1	Population 2
1	11	8
2	7	8
3	9	6
4	12	7
5	13	10
6	15	15
7	15	14

a. Compute the difference value for each element.
b. Compute $\bar{d}$.
c. Compute the standard deviation s_d.
d. What is the point estimate of the difference between the two population means?
e. Provide a 95 per cent confidence interval for the difference between the two population means.

Applications

19 In recent years, a growing array of entertainment options competes for consumer time. By 2004, cable television and radio surpassed broadcast television, recorded music, and the daily newspaper to become the two entertainment media with the greatest usage (*The Wall Street Journal*, 26 January 2004). Researchers used a sample of 15 individuals and collected data on the hours per week spent watching cable television and hours per week spent listening to the radio.

TVRADIO

Individual	Television	Radio	Individual	Television	Radio
1	22	25	9	21	21
2	8	10	10	23	23
3	25	29	11	14	15
4	22	19	12	14	18
5	12	13	13	14	17
6	26	28	14	16	15
7	22	23	15	24	23
8	19	21			

a. What is the sample mean number of hours per week spent watching cable television? What is the sample mean number of hours per week spent listening to radio? Which medium has the greater usage?

b. Use a 0.05 level of significance and test for a difference between the population mean usage for cable television and radio. What is the *p*-value?

20 A market research firm used a sample of individuals to rate the purchase potential of a particular product before and after the individuals saw a new television commercial about the product. The purchase potential ratings were based on a 0 to 10 scale, with higher values indicating a higher purchase potential. The null hypothesis stated that the mean rating 'after' would be less than or equal to the mean rating 'before'. Rejection of this hypothesis would show that the commercial improved the mean purchase potential rating. Use $\alpha = 0.05$ and the following data to test the hypothesis and comment on the value of the commercial.

	Purchase rating			Purchase rating	
Individual	After	Before	Individual	After	Before
1	6	5	5	3	5
2	6	4	6	9	8
3	7	7	7	7	5
4	4	3	8	6	6

21 StreetInsider.com reported 2002 earnings per share data for a sample of major companies (12 February 2003). Prior to 2002, financial analysts predicted the 2002 earnings per share for these same companies (*Barron's*, 10 September 2001). Use the following data to comment on differences between actual and estimated earnings per share.

Company	Actual	Predicted
AT & T	1.29	0.38
American Express	2.01	2.31
Citigroup	2.59	3.43
Coca Cola	1.60	1.78
DuPont	1.84	2.18
Exxon-Mobil	2.72	2.19
General Electric	1.51	1.71
Johnson & Johnson	2.28	2.18
McDonald's	0.77	1.55
Wal-Mart	1.81	1.74

EARNINGS

a. Use $\alpha = 0.05$ and test for any difference between the population mean actual and population mean estimated earnings per share. What is the *p*-value? What is your conclusion?

b. What is the point estimate of the difference between the two means? Did the analysts tend to underestimate or overestimate the earnings?

c. At 95 per cent confidence, what is the margin of error for the estimate in part (b)? What would you recommend based on this information?

22 A survey was made of Book-of-the-Month-Club members to ascertain whether members spend more time watching television than they do reading. Assume a sample of 15 respondents provided the following data on weekly hours of television watching and weekly hours of reading. Using a 0.05 level of significance, can you conclude that Book-of-the-Month-Club members spend more hours per week watching television than reading?

TVREAD

Respondent	Television	Reading	Respondent	Television	Reading
1	10	6	9	4	7
2	14	16	10	8	8
3	16	8	11	16	5
4	18	10	12	5	10
5	15	10	13	8	3
6	14	8	14	19	10
7	10	14	15	11	6
8	12	14			

10.4 Inferences about the difference between two population proportions

Let π_1 denote the proportion for population 1 and π_2 denote the proportion for population 2. We next consider inferences about the difference between the two population proportions: $\pi_1 - \pi_2$. To make an inference about this difference, we shall select two independent random samples consisting of n_1 units from population 1 and n_2 units from population 2.

Interval estimation of $\pi_1 - \pi_2$

In the following example, we show how to compute a margin of error and construct an interval estimate of the difference between two population proportions.

An accountancy firm specializing in the preparation of income tax returns is interested in comparing the quality of work at two of its regional offices. The firm will be able to estimate the proportion of erroneous returns by randomly selecting samples of tax returns prepared at each office and verifying the sample returns' accuracy. The difference between these proportions is of particular interest.

π_1 = proportion of erroneous returns for population 1 (office 1)
π_2 = proportion of erroneous returns for population 2 (office 2)
P_1 = sample proportion for a simple random sample from population 1
P_2 = sample proportion for a simple random sample from population 2

The difference between the two population proportions is given by $\pi_1 - \pi_2$. The point estimator of $\pi_1 - \pi_2$ is as follows.

Point estimator of the difference between two population proportions

$$P_1 - P_2 \quad (10.10)$$

The point estimator of the difference between two population proportions is the difference between the sample proportions of two independent simple random samples.

As with other point estimators, the point estimator $P_1 - P_2$ has a sampling distribution that reflects the possible values of $P_1 - P_2$ if we repeatedly took two independent random samples. The mean of this sampling distribution is $\pi_1 - \pi_2$ and the standard error of $P_1 - P_2$ is as follows:

Standard error of $P_1 - P_2$

$$\sigma_{P_1 - P_2} = \sqrt{\frac{\pi_1(1-\pi_1)}{n_1} + \frac{\pi_2(1-\pi_2)}{n_2}} \quad \textbf{(10.11)}$$

If the sample sizes are large enough that $n_1\pi_1$, $n_1(1 - \pi_1)$, $n_2\pi_2$ and $n_2(1 - \pi_2)$ are all greater than or equal to five, the sampling distribution of $P_1 - P_2$ can be approximated by a normal distribution.

As we showed previously, an interval estimate is given by a point estimate ± a margin of error. In the estimation of the difference between two population proportions, an interval estimate will take the form $p_1 - p_2 \pm$ margin of error. With the sampling distribution of $P_1 - P_2$ approximated by a normal distribution, we would like to use $z_{\alpha/2}\,\sigma_{P_1 - P_2}$ as the margin of error. However, $\sigma_{P_1 - P_2}$ given by equation (10.11) cannot be used directly because the two population proportions, π_1 and π_2, are unknown. Using the sample proportion p_1 to estimate π_1 and the sample proportion p_2 to estimate π_2, the margin of error is as follows.

$$\text{Margin of error} = z_{\alpha/2}\sqrt{\frac{p_1(1-p_1)}{n_1} + \frac{p_2(1-p_2)}{n_2}} \quad \textbf{(10.12)}$$

The general form of an interval estimate of the difference between two population proportions is as follows.

Interval estimate of the difference between two population proportions

$$(p_1 - p_2) \pm z_{\alpha/2}\sqrt{\frac{p_1(1-p_1)}{n_1} + \frac{p_2(1-p_2)}{n_2}} \quad \textbf{(10.13)}$$

where $1 - \alpha$ is the confidence coefficient.

Returning to the tax returns example, we find that independent simple random samples from the two offices provide the following information.

TAX PREP

Office 1	Office 2
$n_1 = 250$	$n_1 = 300$
Number of returns with errors = 35	Number of returns with errors = 27

The sample proportions for the two offices are:

$$p_1 = \frac{35}{250} = 0.14 \qquad p_2 = \frac{27}{300} = 0.09$$

The point estimate of the difference between the proportions of erroneous tax returns for the two populations is $p_1 - p_2 = 0.14 - 0.09 = 0.05$. We estimate that Office 1 has a 0.05, or 5 percentage points, greater error rate than Office 2.

Expression (10.13) can now be used to provide a margin of error and interval estimate of the difference between the two population proportions. Using a 90 per cent confidence interval with $z_{\alpha/2} = z_{0.05} = 1.645$, we have

$$(p_1 - p_2) \pm z_{\alpha/2}\sqrt{\frac{p_1(1-p_1)}{n_1} + \frac{p_2(1-p_2)}{n_2}}$$

$$= (0.14 - 0.09) \pm 1.645\sqrt{\frac{0.14(1-0.14)}{250} + \frac{0.09(1-0.09)}{300}} = 0.05 \pm 0.045$$

The margin of error is 0.045, and the 90 per cent confidence interval is 0.005 to 0.095.

Hypothesis tests about $\pi_1 - \pi_2$

Let us now consider hypothesis tests about the difference between the proportions of two populations. The three forms for a hypothesis test are as follows:

$$H_0: \pi_1 - \pi_2 \geq 0 \qquad H_0: \pi_1 - \pi_2 \leq 0 \qquad H_0: \pi_1 - \pi_2 = 0$$
$$H_1: \pi_1 - \pi_2 < 0 \qquad H_1: \pi_1 - \pi_2 > 0 \qquad H_1: \pi_1 - \pi_2 \neq 0$$

When we assume H_0 is true as an equality, we have $\pi_1 - \pi_2 = 0$, which is the same as saying that the population proportions are equal, $\pi_1 = \pi_2$. We shall base the test statistic on the sampling distribution of the point estimator $P_1 - P_2$.

In expression (10.11), we showed that the standard error of $P_1 - P_2$ is given by

$$\sigma_{P_1 - P_2} = \sqrt{\frac{\pi_1(1-\pi_1)}{n_1} + \frac{\pi_2(1-\pi_2)}{n_2}}$$

Under the assumption that H_0 is true as an equality, the population proportions are equal and $\pi_1 = \pi_2 = \pi$. In this case, $\sigma_{P_1 - P_2}$ becomes

Standard error of $P_1 - P_2$ when $\pi_1 = \pi_2 = \pi$

$$\sigma_{P_1 - P_2} = \sqrt{\frac{\pi(1-\pi)}{n_1} + \frac{\pi(1-\pi)}{n_2}} = \sqrt{\pi(1-\pi)\left(\frac{1}{n_1} + \frac{1}{n_2}\right)} \quad \textbf{(10.14)}$$

With π unknown, we pool, or combine, the point estimates from the two samples (p_1 and p_2) to obtain a single point estimate of π as follows:

Pooled estimate of π when $\pi_1 = \pi_2 = \pi$

$$p = \frac{n_1 p_1 + n_2 p_2}{n_1 + n_2} \quad \textbf{(10.15)}$$

This **pooled estimate of** π is a weighted average of p_1 and p_2.

Substituting p for π in equation (10.14), we obtain an estimate of $\sigma_{P_1 - P_2}$, which is used in the test statistic. The general form of the test statistic for hypothesis tests about the difference between two population proportions is the point estimator divided by the estimate of $\sigma_{P_1 - P_2}$.

Test statistic for hypothesis tests about $\pi_1 - \pi_2$

$$z = \frac{(p_1 - p_2)}{\sqrt{p(1-p)\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} \qquad \textbf{(10.16)}$$

This test statistic applies to large sample situations where $n_1\pi_1$, $n_1(1 - \pi_1)$, $n_2\pi_2$ and $n_2(1 - \pi_2)$ are all greater than or equal to five.

Let us return to the tax returns example and assume that the firm wants to use a hypothesis test to determine whether the error proportions differ between the two offices. A two-tailed test is required. The null and alternative hypotheses are as follows:

$$H_0: \pi_1 - \pi_2 = 0$$
$$H_1: \pi_1 - \pi_2 \neq 0$$

If H_0 is rejected, the firm can conclude that the error rates at the two offices differ. We shall use $\alpha = 0.10$ as the level of significance.

The sample data previously collected showed $p_1 = 0.14$ for the $n_1 = 250$ returns sampled at Office 1 and $p_2 = 0.09$ for the $n_2 = 300$ returns sampled at Office 2. The pooled estimate of π is

$$p = \frac{n_1p_1 + n_2p_2}{n_1 + n_2} = \frac{250(0.14) + 300(0.09)}{250 + 300} = 0.1127$$

Using this pooled estimate and the difference between the sample proportions, the value of the test statistic is as follows.

$$z = \frac{(p_1 - p_2)}{\sqrt{p(1-p)\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} = \frac{(0.14 - 0.09)}{\sqrt{0.1127(1 - 0.1127)\left(\frac{1}{250} + \frac{1}{300}\right)}} = 1.85$$

To compute the p-value for this two-tailed test, we first note that $z = 1.85$ is in the upper tail of the standard normal distribution. Using the standard normal distribution table, we find the area in the upper tail for $z = 1.85$ is $1 - 0.9678 = 0.0322$. Doubling this area for a two-tailed test, we find the p-value $= 2(0.0322) = 0.0644$. With the p-value less than $\alpha = 0.10$, H_0 is rejected at the 0.10 level of significance. The firm can conclude that the error rates differ between the two offices. This hypothesis test conclusion is consistent with the earlier interval estimation results that showed the interval estimate of the difference between the population error rates at the two offices to be 0.005 to 0.095, with Office 1 having the higher error rate.

Exercises

Methods

23 Consider the following results for independent samples taken from two populations.

Sample 1	Sample 2
$n_1 = 400$	$n_2 = 300$
$p_1 = 0.48$	$p_2 = 0.36$

a. What is the point estimate of the difference between the two population proportions?

b. Construct a 90 per cent confidence interval for the difference between the two population proportions.

c. Construct a 95 per cent confidence interval for the difference between the two population proportions.

24 Consider the hypothesis test

$$H_0: \pi_1 - \pi_2 \leq 0$$
$$H_1: \pi_1 - \pi_2 > 0$$

The following results are for independent samples taken from the two populations.

Sample 1	Sample 2
$n_1 = 200$	$n_2 = 300$
$p_1 = 0.22$	$p_2 = 0.1$

a. What is the p-value?

b. With $\alpha = 0.05$, what is your hypothesis testing conclusion?

Applications

25 In November and December 2008, research companies affiliated to the Worldwide Independent Network of Market Research carried out polls in 17 countries to assess people's views on the economic outlook. In the Canadian survey, conducted by Léger Marketing, 61 per cent of the sample of 1511 people thought the economic situation would worsen over the next three months. In the UK survey, conducted by ICM Research, 78 per cent of the sample of 1050 felt that economic conditions would worsen over that period. Provide a 95 per cent confidence interval estimate for the difference between the population proportions in the two countries. What is your interpretation of the interval estimate?

26 The Anwar Sadat Chair for Peace and Development carried out an opinion poll among adults in six African and Arab states in May 2004. The results show that 69 per cent of 400 respondents in Jordan felt that the war in Iraq had brought less democracy to the country, compared with 57 per cent of 700 respondents in Lebanon who had that view. Construct a 95 per cent confidence interval for the difference between the proportion of Jordanian adults who held this view and the proportion of Lebanese adults who held this view.

27 In a test of the quality of two television commercials, each commercial was shown in a separate test area six times over a one-week period. The following week a telephone survey was conducted to identify individuals who had seen the commercials. Those individuals were asked to state the primary message in the commercials. The following results were recorded.

	Commercial A	Commercial B
Number who saw commercial	150	200
Number who recalled message	63	60

a. Use $\alpha = 0.05$ and test the hypothesis that there is no difference in the recall proportions for the two commercials.

b. Compute a 95 per cent confidence interval for the difference between the recall proportions for the two populations.

28 UNITE/MORI published annual 'Student Experience Reports' from 2001 to 2005, based on face-to-face interviews carried out at a sample of UK universities. In 2001, it was reported that 74 per cent of 1103 respondents strongly agreed with the statement that 'going to university is a worthwhile experience'. The 2005 report says that 66 per cent of 1065 respondents strongly agreed with this statement. Test the hypothesis $\pi_1 - \pi_2 = 0$ with $\alpha = 0.05$. What is the *p*-value. What is your conclusion?

29 A large car insurance company selected samples of single and married male policyholders and recorded the number who made an insurance claim over the preceding three-year period.

Single policyholders	Married policyholders
$n_1 = 400$	$n_2 = 900$
Number making claims = 76	Number making claims = 90

a. Use $\alpha = 0.05$ and test to determine whether the claim rates differ between single and married male policyholders.

b. Provide a 95 per cent confidence interval for the difference between the proportions for the two populations.

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

Summary

In this chapter we discussed procedures for constructing interval estimates and doing hypothesis tests involving two populations. First, we showed how to make inferences about the difference between two population means when independent simple random samples are selected. We considered the case where the population standard deviations, σ_1 and σ_2, could be assumed known. The standard normal distribution z was used to develop the interval estimate and served as the test statistic for hypothesis tests. We then considered the case where the population standard deviations were unknown and estimated by the sample standard deviations s_1 and s_2. In this case, the t distribution was used to develop the interval estimate and served as the test statistic for hypothesis tests.

Inferences about the difference between two population means were then discussed for the matched sample design. In the matched sample design each element provides a pair of data values, one from each population. The difference between the paired data values is then used in the statistical analysis. The matched sample design is generally preferred to the independent sample design, when it is feasible, because the matched-samples procedure often improves the precision of the estimate.

Finally, interval estimation and hypothesis testing about the difference between two population proportions were discussed. Statistical procedures for analyzing the difference between two population proportions are similar to the procedures for analyzing the difference between two population means.

Key terms

Independent samples
Matched samples
Pooled estimator of π

Key formulae

Point estimator of the difference between two population means

$$\bar{X}_1 - \bar{X}_2 \quad (10.1)$$

Standard error of $\bar{X}_1 - \bar{X}_2$

$$\sigma_{\bar{X}_1 - \bar{X}_2} = \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} \quad (10.2)$$

Interval estimate of the difference between two population means: σ_1 and σ_2 known

$$(\bar{x}_1 - \bar{x}_2) \pm z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} \quad (10.4)$$

Test statistic for hypothesis tests about $\mu_1 - \mu_2$: σ_1 and σ_2 known

$$z = \frac{(\bar{x}_1 - \bar{x}_2) - D_0}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \quad (10.5)$$

Interval estimate of the difference between two population means: σ_1 and σ_2 unknown

$$(\bar{x}_1 - \bar{x}_2) \pm t_{\alpha/2}\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} \quad (10.6)$$

Degrees of freedom for the t distribution using two independent random samples

$$df = \frac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)^2}{\left(\frac{1}{n_1 - 1}\right)\left(\frac{s_1^2}{n_1}\right)^2 + \left(\frac{1}{n_2 - 1}\right)\left(\frac{s_2^2}{n_2}\right)^2} \quad (10.7)$$

Test statistic for hypothesis tests about $\mu_1 - \mu_2$: σ_1 and σ_2 unknown

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - D_0}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} \quad (10.8)$$

Test statistic for hypothesis test involving matched samples

$$t = \frac{\bar{d} - \mu_d}{s_d/\sqrt{n}} \quad (10.9)$$

Point estimator of the difference between two population proportions

$$P_1 - P_2 \quad (10.10)$$

Standard error of $P_1 - P_2$

$$\sigma_{P_1 - P_2} = \sqrt{\frac{\pi_1(1 - \pi_1)}{n_1} + \frac{\pi_2(1 - \pi_2)}{n_2}} \quad (10.11)$$

Interval estimate of the difference between two population proportions

$$(p_1 - p_2) \pm z_{\alpha/2}\sqrt{\frac{p_1(1 - p_1)}{n_1} + \frac{p_2(1 - p_2)}{n_2}} \quad (10.13)$$

Standard error of $P_1 - P_2$ when $\pi_1 = \pi_2 = \pi$

$$\sigma_{P_1 - P_2} = \sqrt{\frac{\pi(1 - \pi)}{n_1} + \frac{\pi(1 - \pi)}{n_2}} = \sqrt{\pi(1 - \pi)\left(\frac{1}{n_1} + \frac{1}{n_2}\right)} \quad (10.14)$$

Pooled estimate of π when $\pi_1 = \pi_2 = \pi$

$$p = \frac{n_1p_1 + n_2p_2}{n_1 + n_2} \quad (10.15)$$

Test statistic for hypothesis tests about $\pi_1 - \pi_2$

$$z = \frac{(p_1 - p_2)}{\sqrt{p(1 - p)\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} \quad (10.16)$$

Case problem Par Products

Par Products is a major manufacturer of golf equipment. Management believes that Par's market share could be increased with the introduction of a cut-resistant, longer-lasting golf ball. Therefore, the research group at Par has been investigating a new golf ball coating designed to resist cuts and provide a more durable ball. The tests with the coating have been promising.

GOLF

One of the researchers voiced concern about the effect of the new coating on driving distances. Par would like the new cut-resistant ball to offer driving distances comparable to those of the current-model golf ball. To compare the driving distances for the two balls, 40 balls of both the new and current models were subjected to distance tests. The testing was performed with a mechanical hitting machine so that any difference between the mean distances for the two models could be attributed to a difference in the two models. The results of the tests, with distances measured to the nearest metre, are available on the CD that accompanies the text, in the file 'Golf'.

Model		Model		Model		Model	
Current	New	Current	New	Current	New	Current	New
264	277	270	272	263	274	281	283
261	269	287	259	264	266	274	250
267	263	289	264	284	262	273	253
272	266	280	280	263	271	263	260
258	262	272	274	260	260	275	270
283	251	275	281	283	281	267	263
258	262	265	276	255	250	279	261
266	289	260	269	272	263	274	255
259	286	278	268	266	278	276	263
270	264	275	262	268	264	262	279

Managerial report

1 Formulate and present the rationale for a hypothesis test that Par could use to compare the driving distances of the current and new golf balls.

2 Analyze the data to provide the hypothesis test conclusion. What is the *p*-value for your test? What is your recommendation for Par Products?

3 Provide descriptive statistical summaries of the data for each model.

GOLF

4 What is the 95 per cent confidence interval for the population mean of each model, and what is the 95 per cent confidence interval for the difference between the means of the two populations?

5 Do you see a need for larger sample sizes and more testing with the golf balls? Discuss.

Software Section for Chapter 10

Inferences about two populations using MINITAB

We describe the use of MINITAB to construct interval estimates and do hypothesis tests about the difference between two population means and the difference between two population proportions. MINITAB provides both interval estimation and hypothesis test results within the same module, so the MINITAB procedure is the same for both types of inferences. In the examples that follow, we shall demonstrate interval estimation and hypothesis testing for the same two samples. We note that MINITAB does not provide a routine for inferences about the difference between two population means when the population standard deviations σ_1 and σ_2 are known.

Difference between two population means: σ_1 and σ_2 unknown

CHEQACCT

We shall use the data for the cheque account balances example presented in Section 10.2 (file 'CheqAcct.MTW' on the accompanying CD). The cheque account balances at the Northern Branch are in column C1, and the cheque account balances at the Eastern Branch are in column C2. In this example, we will use the MINITAB 2-Sample *t* procedure to provide a 95 per cent confidence interval estimate of the difference between population means for the cheque account balances at the two branch banks. The output of the procedure also provides the *p*-value for the hypothesis test: H_0: $\mu_1 - \mu_2 = 0$ versus H_1: $\mu_1 - \mu_2 \neq 0$.

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

Step 2 Check **Samples in different columns** [**2-Sample t (Test and Confidence Interval)** panel]
Enter **C1** in the **First** box
Enter **C2** in the **Second** box
Click **Options**

Step 3 Enter **95** in the **Confidence level** box [**2-Sample t - Options** panel]
Enter **0** in the **Test difference** box
Select **not equal** on the **Alternative** menu
Click **OK**

Step 4 Click **OK** [**2-Sample t (Test and Confidence Interval)** panel]

Figure 10.2 MINITAB output for the hypothesis test and confidence interval for the cheque account balances

Results for: CheqAcct.MTW

Two-Sample T-Test and CI: Northern, Eastern

```
Two-sample T for Northern vs Eastern

                              SE
            N  Mean  StDev  Mean
Northern   28  1025    150    28
Eastern    22   910    125    27


Difference = mu (Northern) - mu (Eastern)
Estimate for difference:  115.0
95% CI for difference:  (36.7, 193.2)
T-Test of difference = 0 (vs not =): T-Value = 2.95  P-Value = 0.005  DF = 47
```

The MINITAB output is shown above in Figure 10.2. The 95 per cent confidence interval estimate is (€37 to €193) as described in Section 10.2. The *p*-value = 0.005 shows the null hypothesis of equal population means can be rejected at the $\alpha = 0.01$ level of significance. Note that MINITAB used equation (10.7) to compute 47 degrees of freedom for this analysis. In other applications, the Options dialogue panel may be used to provide different confidence levels, different hypothesized values, and different forms of the hypotheses.

Difference between two population means with matched samples

MATCHED

We use the data on production times in Table 10.2 to illustrate the matched-sample procedure (file 'Matched.MTW' on the accompanying CD). The completion times for method 1 are entered into column C1 and the completion times for method 2 are entered into column C2. The MINITAB steps for a matched sample are as follows:

Step 1 **Stat > Basic Statistics > Paired t** [Main menu bar]

Step 2 Select **Samples in columns** [**Paired t (Test and Confidence Interval)** panel]
Enter **C1** in the **First sample** box
Enter **C2** in the **Second sample** box
Click **Options**

Step 3 Enter **95** in the **Confidence level** box [**Paired t - Options** panel]
Enter **0** in the **Test difference** box
Select **not equal** on the **Alternative** menu
Click **OK**

Step 4 Click **OK** [**Paired t (Test and Confidence Interval)** panel]

The **Paired t - Options** dialogue panel may be used to provide different confidence levels, different hypothesized values, and different forms of the hypotheses.

Difference between two population proportions

We shall use the data on tax return errors presented in Section 10.4 (file 'TaxPrep' on the accompanying CD). The sample results for 250 tax returns prepared at Office 1 are in column C1 and the sample results for 300 tax returns prepared at Office 2 are in column C2. Yes denotes an error was found in the tax return and No indicates no error was found. The procedure we describe provides both a 95 per cent confidence interval for $\pi_1 - \pi_2$ and hypothesis test results for H_0: $\pi_1 - \pi_2 = 0$ versus H_1: $\pi_1 - \pi_2 \neq 0$.

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

Step 2 Select **Samples in different columns**
[**2 Proportions (Test and Confidence Interval)** panel]
Enter **C1** in the **First** box
Enter **C2** in the **Second** box
Click **Options**

Step 3 Enter **95** in the **Confidence level** box [**2 Proportions - Options** panel]
Enter **0** in the **Test difference** box
Select **not equal** on the **Alternative** menu
Check **Use pooled estimate of p for test**
Click **OK**

Step 4 Click **OK** [**2 Proportions (Test and Confidence Interval)** panel]

The Options dialogue panel may be used to provide different confidence levels, different hypothesized values, and different forms of the hypotheses.

In the tax returns example, the data are qualitative. Yes and No are used to indicate whether an error is present. In modules involving proportions, MINITAB calculates proportions for the response coming second in alphabetic order. In the tax preparation example, MINITAB computes the proportion of Yes responses, which is the proportion we wanted. If MINITAB's alphabetical ordering does not compute the proportion for the response of interest, we can fix it. Select any cell in the data column, go to the MINITAB menu bar, and select **Editor > Column > Value Order**. This sequence will provide the option of entering a user-specified order. Simply make sure that the response of interest is listed second in the **define-an-order** box. MINITAB's **2 Proportion** routine will then provide the confidence interval and hypothesis testing results for the population proportion of interest.

Finally, we note that MINITAB's **2 Proportion** routine uses a computational procedure different from the procedure described in the text. Consequently, the MINITAB output can be expected to provide slightly different interval estimates and slightly different *p*-values. However, results from the two methods should be close and are expected to provide the same interpretation and conclusion.

Inferences about two populations using EXCEL

We describe the use of EXCEL to conduct hypothesis tests about the difference between two population means. We begin with hypothesis tests for the difference between the means of two populations when the population standard deviations σ_1 and σ_2 are known. (Routines are not available for interval estimation of the difference between two population means, nor for inferences about the difference between two population proportions.)

Difference between two population means: σ_1 and σ_2 known

We shall use the examination scores for the two training centres discussed in Section 10.1 (file 'ExamScores.XLS' on the accompanying CD). The label Centre A is in cell A1 and the label Centre B is in cell B1. The examination scores for Centre A are in cells A2:A31 and examination scores for Centre B are in cells B2:B41. The population standard deviations are assumed known with $\sigma_1 = 10$ and $\sigma_2 = 10$. The EXCEL routine will request the input of variances which are $\sigma_1^2 = 100$ and $\sigma_2^2 = 100$. The following steps can be used to conduct a hypothesis test about the difference between the two population means.

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

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

Step 3 Choose **z-Test: Two Sample for Means**
Click **OK**

Step 4 Enter **A1:A31** in the **Variable 1 Range** box
[**z-Test: Two Sample for Means** panel]
Enter **B1:B41** in the **Variable 2 Range** box
Enter **0** in the **Hypothesized Mean Difference** box
Enter **100** in the **Variable 1 Variance (known)** box
Enter **100** in the **Variable 2 Variance (known)** box
Select **Labels**
Enter **.05** in the **Alpha** box
Select **Output Range** and enter **C1** in the box
Click **OK**

The two-tailed *p*-value is denoted 'P(Z <= *z*) two-tail'. Its value of 0.0977 does not allow us to reject the null hypothesis at $\alpha = 0.05$.

Difference between two population means: σ_1 and σ_2 unknown

We use the data for the software testing study in Table 10.1 (file 'SoftwareTest.XLS' on the accompanying CD). The data are already entered into an EXCEL worksheet with the label Current in cell A1 and the label New in cell B1. The completion times for the current technology are in cells A2:A13, and the completion times for the new software are in cells B2:B13. The following steps can be used to conduct a hypothesis test about the difference between two population means with σ_1 and σ_2 unknown.

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

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

Step 3 Choose **t-Test: Two Sample Assuming Unequal Variances**
Click **OK**

Step 4 Enter **A1:A13** in the **Variable 1 Range** box
[**t-Test: Two Sample Assuming Unequal Variances** panel]
Enter **B1:B13** in the **Variable 2 Range** box
Enter **0** in the **Hypothesized Mean Difference** box
Select **Labels**
Enter **.05** in the **Alpha** box
Select **Output Range** and enter **C1** in the box
Click **OK**

The appropriate *p*-value is denoted 'P(T <= t) one-tail'. Its value of 0.017 allows us to reject the null hypothesis at $\alpha = 0.05$.

Difference between two population means with matched samples

MATCHED

We use the matched-sample completion times in Table 10.2 to illustrate (file 'Matched.XLS' on the accompanying CD). The data are entered into a worksheet with the label Method 1 in cell A1 and the label Method 2 in cell B1. The completion times for method 1 are in cells A2:A7 and the completion times for method 2 are in cells B2:B7. The EXCEL procedure uses the steps previously described for the *t*-Test except the user chooses the **t-Test: Paired Two Sample for Means** data analysis tool in step 3. The variable 1 range is A1:A7 and the variable 2 range is B1:B7.

The appropriate *p*-value is denoted 'P(T <= t) two-tail'. Its value of 0.08 does not allow us to reject the null hypothesis at $\alpha = 0.05$.

Inferences about two populations using PASW

PASW

We describe the use of PASW to construct interval estimates and do hypothesis tests about the difference between two population means, first for the independent-samples case when the population standard deviations σ_1 and σ_2 are unknown, and second for the matched-samples case. PASW provides both interval estimation and hypothesis test results within the same module, so the PASW procedure is the same for both types of inferences. In the examples that follow, we will demonstrate interval estimation and hypothesis testing for the same two samples. We note that PASW does not provide a routine for inferences about the difference between two population means when the population standard deviations σ_1 and σ_2 are known, nor a routine for inferences about the difference between two population proportions.

Difference between two population means: σ_1 and σ_2 unknown

CHEQACCT

We shall use the data for the cheque account balances example presented in Section 10.2 (file 'CheqAcct.SAV' on the accompanying CD). The cheque account balances at both the Northern and Eastern branches are in the first column of the PASW data file. The second column indicates whether the account balance is for a Northern Branch account or an Eastern Branch account, i.e. either Northern or Eastern appears in each row of the

second column. For an independent-samples analysis in PASW, the data must be set out in this way. In this example, we will use the PASW **Independent-Samples t** procedure to provide a 95 per cent confidence interval estimate of the difference between population means for the cheque account balances at the two branch banks. The output of the procedure also provides the *p*-value for the hypothesis test: H_0: $\mu_1 - \mu_2 = 0$ versus H_1: $\mu_1 - \mu_2 \neq 0$.

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

Step 2 Transfer the account balance to the **Test Variable(s)** box
[**Independent-Samples T Test** panel]
Transfer the branch variable to the **Grouping Variable** box
Click **Define Groups**

Step 3 Enter **1** (code for Northern branch) in the **Group 1** box [**Define Groups** panel]
Enter **2** (code for Eastern branch) in the **Group 2** box
Click **Continue**

Step 4 Click **Options** [**Independent-Samples T Test** panel]

Step 5 Enter **95** in the **Confidence Interval** box
[**Independent-Samples T Test:Options** panel]
Click **Continue**

Step 6 Click **OK** [**Independent-Samples T Test** panel]

Part of the PASW output is shown below in Figure 10.3. PASW computes two versions of the *t* procedure; one assuming equal variances for the two populations, the other not making this assumption. Only the latter, more general procedure has been described fully in this text. The results for this procedure are in the second row of the table. The 95 per cent confidence interval estimate is €37 to €193 as described in Section 10.2. The *p*-value = 0.005 shows the null hypothesis of equal population means can be rejected at the $\alpha = 0.01$ level of significance. Note that PASW used equation (10.7) to compute 47.8 degrees of freedom for this analysis.

Figure 10.3 PASW output for the hypothesis test and confidence interval for the cheque account balances

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means					95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
Account balance (€)	Equal variances assumed	1.571	.216	2.891	48	.006	115.014	39.777	35.037	194.991
	Equal variances not assumed			2.956	47.805	.005	115.014	38.908	36.775	193.252

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