The Open University

M140

Introducing statistics

Computer Book 3

This publication forms part of an Open University module. Details of this and other Open University modules can be obtained from the Student Registration and Enquiry Service, The Open University, PO Box 197, Milton Keynes MK7 6BJ, United Kingdom (tel. +44 (0)845 300 6090; email general-enquiries@open.ac.uk).

Alternatively, you may visit the Open University website at www.open.ac.uk where you can learn more about the wide range of modules and packs offered at all levels by The Open University.

To purchase a selection of Open University materials visit www.ouw.co.uk, or contact Open University Worldwide, Walton Hall, Milton Keynes MK7 6AA, United Kingdom for a brochure (tel. +44 (0)1908 858779; fax +44 (0)1908 858787; email ouw-customer-services@open.ac.uk).

The Open University, Walton Hall, Milton Keynes, MK7 6AA.

First published 2014.

Edited, designed and typeset by The Open University, using the Open University TEX System.

Printed in the United Kingdom by The Charlesworth Group, Wakefield.

SUP 03472 6

1.1

Contents

Introduction

This Computer Book describes all the computer activities you will be expected to complete during your study of M140 Units 10 to 12. Note that the numbering of the chapters, computer activities and figures follow straight on from the end of Computer Book 2. In this Computer Book, there are a few references to subsections in Computer Book 1 and Computer Book 2. You should be able to distinguish these as the numbering will be below 10. (For example, Subsection 7.3 is within Computer Book 2.)

10 Experiments

In this chapter, associated with Unit 10, you will learn how to do the following hypothesis tests using Minitab: the one-sample t-test, the two-sample t-test and the paired t-test. As you will discover when doing these tests, Minitab will also provide corresponding confidence intervals.

10.1 One-sample t-test

In Subsection 7.3 of this Computer Book, you learned how to do the one-sample z-test using Minitab. The procedure for the one-sample t-test works in a very similar way – the major difference is which dialogue box needs to be selected.

Computer activity 84 *Testing the yield from a variety of tomatoes*

In Subsection 4.1 of Unit 10, you learned how to use the one-sample t-test to test the following hypotheses

H_0: $\mu = 4$
H_1: $\mu \neq 4$,

where μ is the yield (in kg per plant) of one variety of outdoor bush tomatoes grown using a new fertiliser.

Yields from five tomatoes are given in the file **tomatoes.mtw**. Test the null hypothesis using a one-sample t-test by doing the following.

- Open the worksheet **tomatoes.mtw**.
- Click on **Stat**, choose **Basic Statistics** and choose **1-Sample t. . .** . This opens the **1-Sample t (Test and Confidence Interval)** dialogue box.
- It is possible to enter summary data (the sample size, mean and standard deviation) into the dialogue box in the same way that you did for the one-sample z-test. However, this means first working them

out from the data in the worksheet. Instead it is possible to work directly from the raw data. Do this now by making sure the **Samples in columns** option is selected and entering `yield` in the corresponding field.

- Select **Perform hypothesis test**. Then enter 4 in the **Hypothesized mean** field, as this is the value of the mean assumed by the null hypothesis. The completed dialogue box should be as in Figure 40.

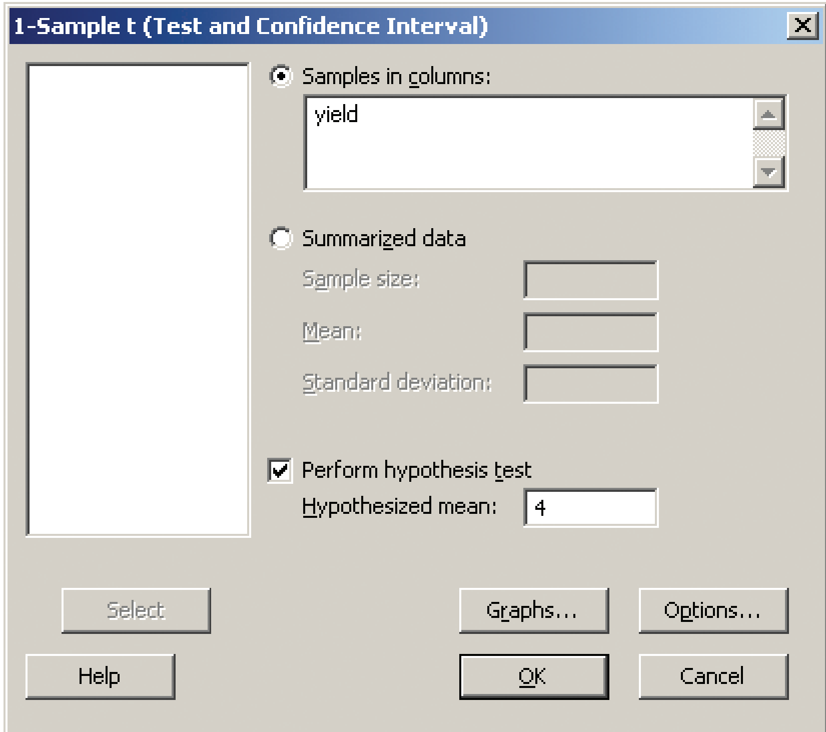


Figure 40 The **1-Sample t (Test and Confidence Interval)** dialogue box

- Click on **OK**.

The output from the test produced by Minitab is as follows.

```
Test of mu = 4 vs not = 4


Variable  N   Mean  StDev  SE Mean       95% CI            T      P
yield     5  3.280  0.497    0.222  (2.663, 3.897)  -3.24  0.032
```

Most of the output is the same as Minitab would produce for the one-sample z-test. The null and alternative hypotheses are restated, and the sample size, mean and standard deviation are given – reassuringly these are the same values as were given in the solution to Activity 16 in Unit 10. The estimated standard error (ESE) – `SE Mean` – is given as `0.222`, which matches that given in Example 11 of Unit 10.

Also in the output is the value `-3.24` for the test statistic, `T`, and the value `0.032` for the corresponding p-value, `P`. Using Table 1 in Subsection 6.2 (a copy of this table is also given in the Handbook), this p-value represents moderate evidence against the null hypothesis. That is, moderate evidence that the yield of this variety of tomato is not 4 kg per plant when the new fertiliser is used. (In fact, the yield appears to be less than 4 kg.)

Notice that as part of the output for the t-test, Minitab provides a 95% confidence interval for the mean yield (based unsurprisingly on the t-test, not the z-test). This interval, `(2.663, 3.897)`, is entirely below the hypothesised value of 4. So, based on the data, a plausible range for the mean yield is (2.66 kg, 3.90 kg) per plant.

In Computer activity 84, you saw that Minitab automatically calculates a 95% confidence interval for the population mean based on the t-test. In Minitab it is possible to obtain the confidence interval without the output for the t-test. This is done by making sure that the **Perform hypothesis test** option is *not* selected in the **1-Sample t (Test and Confidence Interval)** dialogue box. Note that when this is done, it does not matter if a number has been entered in the **Hypothesized mean** field.

The next computer activity gives you some practice in doing the one-sample t-test and obtaining the corresponding 95% confidence interval in Minitab.

Computer activity 85 *Measuring the speed of light*

The file **lightspeed.mtw** contains 23 measurements of the speed of light in air, made by Albert Michelson in 1882. (Data source: Stigler, S.M. (1977) 'Do robust estimators work with real data?', *The Annals of Statistics*, vol. 5, pp. 1055–1098.) Each of the measurements is in km/s over 299 000. It is reasonable for you to assume that the measurements come from a normal population distribution, because of the way they were taken.

(a) Using Minitab, obtain the 95% confidence interval for the speed of light in air.

(b) According to one source, the speed of light in air is 299 705 km/s. This corresponds to the following hypothesis:

$$H_0\colon\ \mu = 705$$
$$H_1\colon\ \mu \neq 705,$$

where μ is the speed of light in air in km/s over 299 000. Test this claim using a one-sample t-test. What value of the test statistic do you obtain and what is the p-value? Hence what conclusion do you draw?

(c) Does the confidence interval you obtained in part (a) match the conclusion you drew in part (b)? Why or why not?

10.2 Two-sample t-test

In Subsection 10.1, you learned how to do the one-sample t-test in Minitab. Using Minitab it is also possible to do the two-sample t-test, as you will discover in this subsection.

Computer activity 86 *Ball manoeuvres*

In Activities 10 and 12 of Unit 10 (Subsection 3.3) you tested whether the method of preparation made a difference to the length of time a child took to manoeuvre a ball round an obstacle course and into a hole. Those activities involved quite a bit of calculation by hand. Use Minitab to complete those same calculations by doing the following.

- Open the worksheet **obstaclecourse.mtw**. Notice in the worksheet that the data for the two groups are given in separate columns. This is not the only way data for the two-sample t-test can be structured in Minitab, but it is the way we will use for now.
- Obtain the **2-Sample t (Test and Confidence Interval)** dialogue box by clicking on **Stat**, choosing **Basic Statistics** and then **2-Sample t. . . .**
- In the dialogue box select the **Samples in different columns** option. Then in the **First** field enter `A` and in the **Second** field enter `B`. Also select the **Assume equal variances** option. The completed dialogue box is shown in Figure 41.

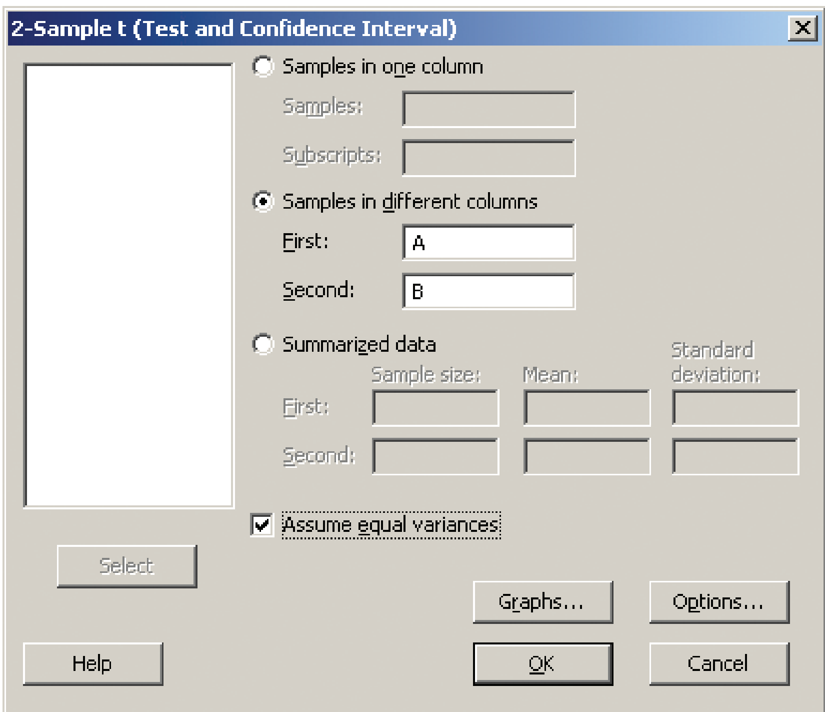


Figure 41 The **2-Sample t (Test and Confidence Interval)** dialogue box

- Click on **OK**.

The output produced by Minitab is as follows.

```
Two-sample T for A vs B

   N  Mean  StDev  SE Mean
A  5  5.00   2.55      1.1
B  5  7.00   3.08      1.4


Difference = mu (A) - mu (B)
Estimate for difference:  -2.00
95% CI for difference:  (-6.13, 2.13)
T-Test of difference = 0 (vs not =): T-Value = -1.12  P-Value = 0.296  DF = 8
Both use Pooled StDev = 2.8284
```

As with the one-sample t-test, Minitab starts by giving details about the test it has just done and gives some summary statistics. The output also automatically gives the 95% confidence interval for the difference between the two means. (In this case, the output makes it clear that the difference is taken to be $\mu_A - \mu_B$.) The last line gives the value of the pooled standard deviation – reassuringly the same as the value calculated in Activity 10 of Unit 10 (Subsection 3.3).

The value of the test statistic, t, and the associated p-value are given in the penultimate line of the output. The value $t = -1.12$ is the same as that calculated in Activity 12 of Unit 10, and a p-value of 0.296 matches the conclusion drawn in that activity – namely, there is little evidence that how the children were prepared for the obstacle course affected the length of time they took to negotiate the course.

Computer activity 87 *Effect of selection*

The file **fruitfly.mtw** contains data on the egg-laying capability of female fruit flies of the species *Drosophila melangoster*. (Data source: Sokal, R.R. and Rohlf, F.J. (1995) *Biometry: The Principles and Practice of Statistics in Biological Research*, 3rd edn, W.H. Freeman and Company.) Each data value is the number of eggs laid per female per day in the first 14 days of life. The females in the `selected` group have been selectively bred to change their resistance to DDT (an insecticide), the females in the `control` have not been selectively bred.

(a) Produce a single diagram that contains boxplots of the number of eggs laid per female per day for each of the two groups (**Graph** > **Boxplot**).

From the boxplots, is it reasonable to assume that in both groups the population distribution of the number of eggs laid per female per day is normal?

(b) Using the two-sample t-test, test the hypothesis that the selective breeding did not alter the egg-laying capability of this species of fruit fly.

(c) Is it reasonable to have used a pooled standard deviation in the test you did in part (b)?

(d) Give a 95% confidence interval for $\mu_s - \mu_c$, where μ_s is the egg-laying capability of selectively bred fruit flies and μ_c is the egg-laying capability of fruit flies that have not been selectively bred.

10.3 Matched-pairs t-test

Subsection 4.2 of Unit 10 focused on the matched-pairs t-test. There you learned that this t-test can be thought of as a particular form of the one-sample t-test – that is, a one-sample t-test where the data consist of differences (d), and the hypotheses are as follows:

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

In this subsection you will learn how to do this test using Minitab, without having to calculate the differences explicitly.

Computer activity 88 *Testing the effect on sleep by two forms of a drug*

In Subsection 4.2 of Unit 10, you considered whether two different forms, L and R, of a sleep-inducing drug, hyoscyamine hydrobromide, differ in their capacity to induce sleep. The data are given in **bromide.mtw**. Use Minitab to test the hypotheses

H_0: $\mu_{L-R} = 0$
H_1: $\mu_{L-R} \neq 0$,

where μ_{L-R} is the difference between sleep gain (in hours) when the L form of the drug was taken compared with when the R form of the drug was taken, by doing the following.

- Open the worksheet **bromide.mtw**. Notice how the data are structured in the worksheet. The data relating to each form of the drug are given in a separate column and each row gives the data for a particular patient.
- Obtain the **Paired t (Test and Confidence Interval)** dialogue box by clicking on **Stat**, choosing **Basic Statistics** and then **Paired t...** .
- In the **Paired t (Test and Confidence Interval)** dialogue box, make sure that **Samples in columns** is selected. Then in the **First sample** field enter `formL` and in the **Second sample** field enter `formR`. The completed dialogue box is shown in Figure 42.

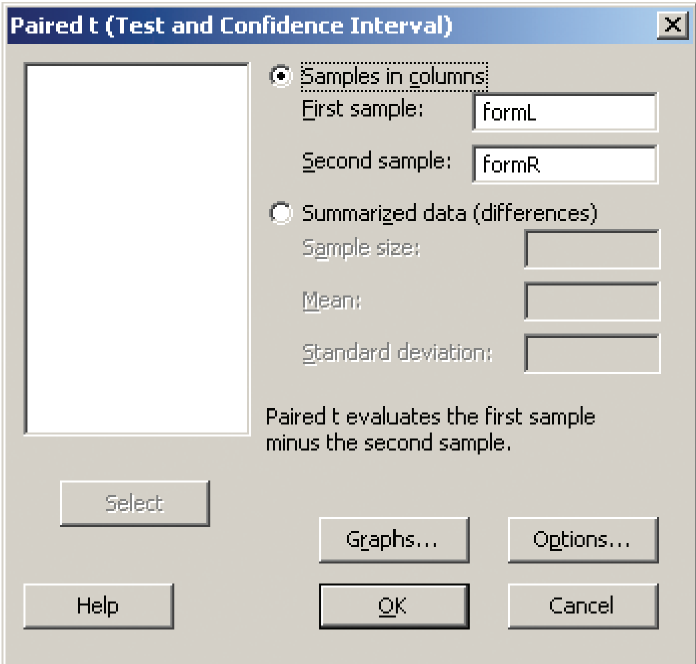


Figure 42 The **Paired t (Test and Confidence Interval)** dialogue box

- Click on **OK**.

The output produced by Minitab is as follows.

```
Paired T for formL - formR

             N   Mean  StDev  SE Mean
formL       10  2.330  2.002    0.633
formR       10  0.750  1.789    0.566
Difference  10  1.580  1.230    0.389


95% CI for mean difference: (0.700, 2.460)
T-Test of mean difference = 0 (vs not = 0): T-Value = 4.06  P-Value = 0.003
```

Notice that this output is structured in a similar way to the output for the one-sample t-test and the two-sample t-test. Some summary statistics are given, along with a 95% confidence interval and details about the test, particularly the value of the test statistic and the corresponding p-value.

So for these data and hypotheses, the test statistic is 4.06 and the p-value is 0.003. This means that there is strong evidence that the sleep gain is not the same for the two forms of hyoscyamine hydrobromide. In fact, it appears that the L form of the drug induces more sleep. An indication of just how much more sleep the L form of the drug induces is provided by the 95% confidence interval. The interval $(0.700, 2.460)$ means that it is plausible that the L form of the drug induces anywhere between 0.700 and 2.460 more hours of sleep than the R form of the drug.

Computer activity 89 *Comparing measuring devices*

In a study to compare measuring devices, readings from 10 samples were taken using two measuring devices, device A and device B. Of interest is whether it matters which machine is used to take the readings – that is, whether on average they both give the same value. This corresponds to the hypotheses

$$H_0\colon \mu_{A-B} = 0$$
$$H_1\colon \mu_{A-B} \neq 0,$$

where μ_{A-B} is the population mean difference in readings taken on device A compared with device B. The data are given in the file **devices.mtw**. (Data source: Hahn, G.J. and Nelson, W. (1970) 'A problem in the statistical comparison of measuring devices', *Technometrics*, vol. 12, pp. 95–102.)

(a) Briefly explain why the matched-pairs t-test is suitable for these data. (You may assume that differences in readings are normally distributed.)

(b) Using a matched-pairs t-test, test the hypothesis. Clearly state your conclusions.

(c) State a 95% confidence interval for the difference in readings given by device A compared with device B.

10.4 One-sided t-tests and z-tests

So far in this chapter, you have just been doing two-sided t-tests using Minitab. That is, the alternative hypothesis H_1 has been that a population mean, or difference between population means, is *not equal* to a stated value.

However, as you learned in Section 6 of Unit 10, sometimes a one-sided alternative hypothesis is more appropriate. That is, an alternative hypothesis of the form

$$H_1\colon \mu < A \quad \text{or} \quad H_1\colon \mu > A,$$

for a one-sample t-test or one-sample z-test, and of the form

$$H_1\colon \mu_A < \mu_B \quad \text{or} \quad H_1\colon \mu_A > \mu_B,$$

for a two-sample t-test.

It is possible to do one-sided tests in Minitab as well as two-sided ones. In fact, the means by which a one-sided test is obtained instead of the two-sided version is the same for the one-sample t-test, two-sample t-test, matched-pairs t-test and the one-sample z-test. To see how it is done for one of these tests we revisit an example introduced in Unit 10.

Computer activity 90 *Benefit of exercise*

In Activity 26 of Unit 10 (Section 6) some data were given on resting heart rate before and after a one-year exercise program. As explained in the solution to Activity 26, the following hypotheses are appropriate

$$H_0\colon \mu_d = 0$$
$$H_1\colon \mu_d < 0,$$

where μ_d denotes the population mean change in resting heart rate over the course of the exercise program. So a one-sided test is required. The data, which are given in the file **exercise.mtw**, consist of pairs of measurements for seven people. So this means that, in particular, a one-sided matched-pairs t-test is required.

To perform the test in Minitab, start in precisely the same way as you did for a two-sided matched-pairs t-test in Subsection 10.3:

- Open the worksheet **exercise.mtw**.
- Click on **Stat**, choose **Basic Statistics** and then **Paired t. . . .** The **Paired t (Test and Confidence Interval)** dialogue box appears.
- Make sure that **Samples in columns** is selected. In the **First sample** field enter `after` and in the **Second sample** field enter `before`. (In this case, `after` is chosen to be the first sample so that the differences taken correspond to `after` − `before`.)

Now there is one extra step to make sure that Minitab does a one-sided rather than two-sided test.

- Click on **Options. . .** in the **Paired t (Test and Confidence Interval)** dialogue box. The **Paired t - Options** dialogue box appears.
- The alternative hypothesis corresponds to $H_1\colon \mu_d < 0$, so in the **Paired t - Options** dialogue box select **less than** in the **Alternative** field. The completed dialogue box is shown in Figure 43.

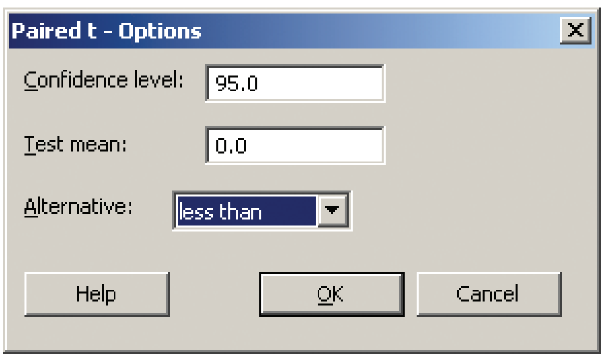


Figure 43 The **Paired t - Options** dialogue box

- Click on **OK** to return to the **Paired t (Test and Confidence Interval)** dialogue box and click on **OK** again.

The resulting Minitab output is:

```
Paired T for after - before

             N    Mean  StDev  SE Mean
after        7   69.57   2.64     1.00
before       7   72.00   2.83     1.07
Difference   7  -2.429  1.813    0.685


95% upper bound for mean difference: -1.097
T-Test of mean difference = 0 (vs < 0): T-Value = -3.54  P-Value = 0.006
```

As before with the matched-pairs t-test, the standard deviation (`StDev`), the sample size (`N`), sample mean (`Mean`) and ESE (`SE Mean`) are quoted for both variables and for the difference. However, notice this time that the statement of the alternative hypothesis is `< 0` rather than `not = 0`, confirming that Minitab has indeed done a one-sided test. A form of a confidence interval appropriate to the one-sided test (`95% Upper Bound`) is then given. In the solution to Activity 26 in Unit 10, the test statistic t was given as -3.545, which matches (apart from rounding) the value of t (`T-Value`) given by Minitab. Finally, the p-value given by Minitab (`P-Value`) is 0.006. So, from Table 1 (Subsection 6.2), this means there is strong evidence against the null hypothesis. That is, strong evidence that the population mean resting heart rate is lower after the exercise program compared with the resting heart rate before the exercise program.

The dialogue boxes you have used to perform the one-sample z-test, the one-sample t-test and the two-sample t-test, also have an **Options...** button. Clicking on this button generates a dialogue box similar to the **Paired t - Options** dialogue box, where you can change the entry in the **Alternative** field to switch between two-sided and one-sided versions of these other tests. (In Subsection 9.2, you have already used this dialogue box to specify which confidence interval from a z-test you wish Minitab to calculate.)

Computer activity 91 provides practice carrying out one of these other one-sided tests using Minitab.

Computer activity 91 *Do people in pain sleep for fewer hours a night?*

An American study of the sleeping problems associated with chronic widespread pain (fibromyalgia) considered many aspects, just one of which was the simple question *Do people with fibromyalgia sleep for fewer hours a night than people without?*. According to the study of a random sample of 744 American fibromyalgia patients, their mean sleeping time was 5.6 hours per night, with standard deviation 1.6 hours. Interest was in whether the population mean sleeping time of American fibromyalgia sufferers was less than the 'normative value' of 6.8 hours per night for the entire population of the USA.

(Source: Cappelleri, J.C. et al. (2009), 'Measurement properties of the Medical Outcomes Study Sleep Scale in patients with fibromyalgia', *Sleep Medicine*, vol. 10, pp. 766–770.)

(a) Write down the null and alternative hypotheses.

(b) Explain why it is appropriate to use a one-sample z-test to test the hypotheses you wrote down in part (a).

(c) Use a one-sided one-sample z-test to test the hypotheses you wrote down in part (a). Comment on your result.

Summary of Chapter 10

In this chapter, you have learned how to do various t-tests in Minitab: the two-sample t-test, the one-sample t-test and the matched-pairs t-test. The output from all these tests include corresponding 95% confidence intervals, thus providing a way to obtain these intervals using Minitab. You have also learned how to switch between doing two-sided and one-sided versions of these tests and also between two-sided and one-sided versions of the one-sample z-test.

11 Clinical trials

This chapter covers two practical elements associated with Unit 11. First, Subsection 11.1 demonstrates most of the randomisation methods for clinical trials that were discussed in Subsection 3.4 of Unit 11. Second, Subsection 11.2 uses the χ^2 test and a t-test to analyse some data from clinical trials. This revises what you learned in Units 8 and 10.

11.1 Randomisation in practice

In this subsection you will explore different schemes for randomising patients in a clinical trial. In Computer activities 93 to 95 you will use Minitab to assign patients to treatments systematically, randomly and then randomly but keeping the totals fixed.

By its very nature, randomisation leads to unpredictable results – at least in terms of which subjects gets assigned to each group. It is, however, possible to obtain exactly reproducible results by exploiting the means by which Minitab generates random numbers.

Random numbers produced by Minitab are in fact 'pseudo-random numbers' – that is, the numbers follow a predetermined sequence, but in such a way that they appear random. This means that by starting the sequence at the same place every time – 'setting the seed' – the same sequence of pseudo-random numbers will always be obtained. So first, in the next activity, you will learn how to set the seed for the random number generator in Minitab.

Computer activity 92 *Setting the seed*

(a) In a new worksheet in Minitab (**File** > **New**) use the same procedure as you did in Computer activity 40 (Subsection 4.2) to obtain a column of 30 numbers between 1 and 100, stored in column `C1` (**Calc** > **Random Data** > **Integer**).

(b) Now set the seed of the random number generator to take the value `10` by doing the following. (The value 10 just happened to be the value chosen by the module team. It has no significance other than it's a value chosen by somebody.)

- Click on **Calc** and then **Set base...** to open the **Set Base** dialogue box.
- Enter `10` in the **Set base of random data generator to** field. The completed dialogue box is shown in Figure 44.

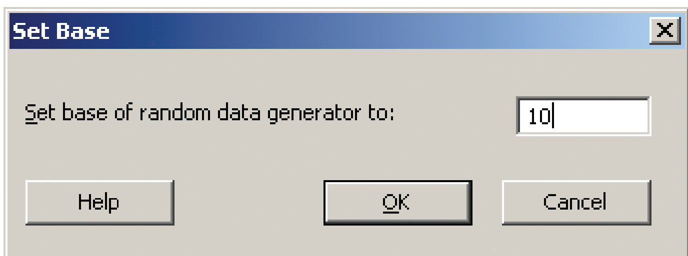


Figure 44 The **Set Base** dialogue box

- Click **OK**.

Generate another column of 30 random numbers between 1 and 100, this time storing them in column `C2`.

(c) Repeat part (b), this time storing the column of random numbers in `C3`.

(d) Generate a fourth column of 30 random numbers between 1 and 100, storing them in column `C4`. (This time do *not* reset the seed first.)

Suppose a clinical trial consisting of 100 participants spread over two centres is being set up, and these participants need to be randomised into the experimental and control groups. Open the worksheet **randomisation.mtw**, which contains two variables `patientID` and `centre`. The variable `patientID` is a patient identifier, and patients are simply numbered 1 to 100. The variable `centre` identifies the centre that the patient will be attending for the trial; there are two centres, numbered 1 and 2. Assignments to groups are to be numbered so that `0` indicates 'assign to the control group', and `1` indicates 'assign to the experimental group'.

In each of the following activities you will assign participants to treatment groups. Note that the solutions for each activity assume that the seed for the random number generator is set to take the value 42.

Computer activity 93 *Assigning patients systematically*

In this activity, you will assign trial participants to groups in a systematic way: odd patient IDs are assigned to the control group and even patient IDs are assigned to the experimental group.

- Click on **Calc**, then **Make Patterned Data** and then **Simple Set of Numbers. . . .** The **Simple Set of Numbers** dialogue box will now appear.
- Enter `treatment` in the **Store patterned data in** field.
- Enter `0` in the **From first value** field.
- Enter `1` in the **To last value** field.
- Amend **Number of times to list the sequence** to `50`.
- Click on **OK**.

Create a contingency table of the number of participants in each of the treatment groups in each centre by doing the following.

- Click on **Stat**, then **Tables** and then **Cross Tabulation and Chi-Square. . .** to obtain the **Cross Tabulation and Chi-Square** dialogue box.
- In the **For rows** field, enter `centre` and in the **For columns** field, enter `treatment`.

- Make sure that the **Counts** option is selected. The completed dialogue box is shown in Figure 45.

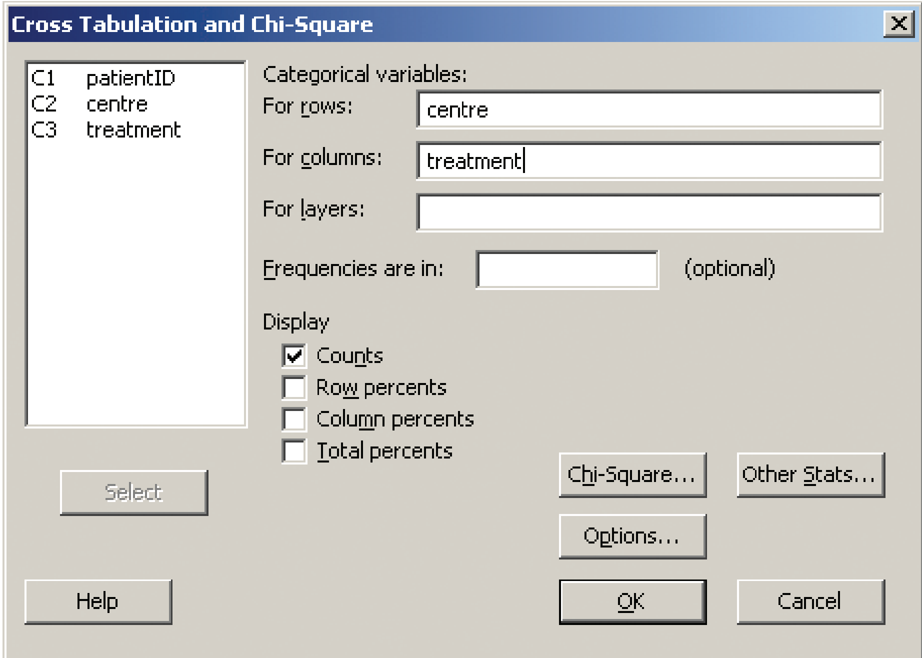


Figure 45 The **Cross Tabulation and Chi-Square** dialogue box

- Click on **OK**.

Notice that when you do this a contingency table appears in the Session window giving the number of participants assigned to each treatment group in each centre. The rows in this table correspond to centres and the columns correspond to treatments.

(a) Overall, how many participants were assigned to each treatment group?

(b) Is there balance in the number of participants assigned to each centre?

(c) If the seed for the random number generator is not set, will the assignment change?

Computer activity 94 *Assigning participants randomly*

In this activity, you will assign trial participants to treatment groups at random – with probability 0.5 of being assigned to either group, as if tossing a coin.

- Make sure that the worksheet **randomisation.mtw** is the active worksheet.
- If you want to get exactly the same answers as given in the solution, set the seed for the random number generator to the value 42 (**Calc** > **Set base**).

- Click on **Calc**, then **Random Data** and then **Binomial. . .** . The **Binomial Distribution** dialogue box will now appear.
- Enter `100` in the **Number of rows of data to generate** field.
- Enter `treatment` in the **Store in column(s)** field.
- For each participant, the outcome is like the outcome from a single coin toss. So enter `1` in the **Number of Trials** field.
- Enter `0.5` in the **Event probability** field. The completed dialogue box is shown in Figure 46.

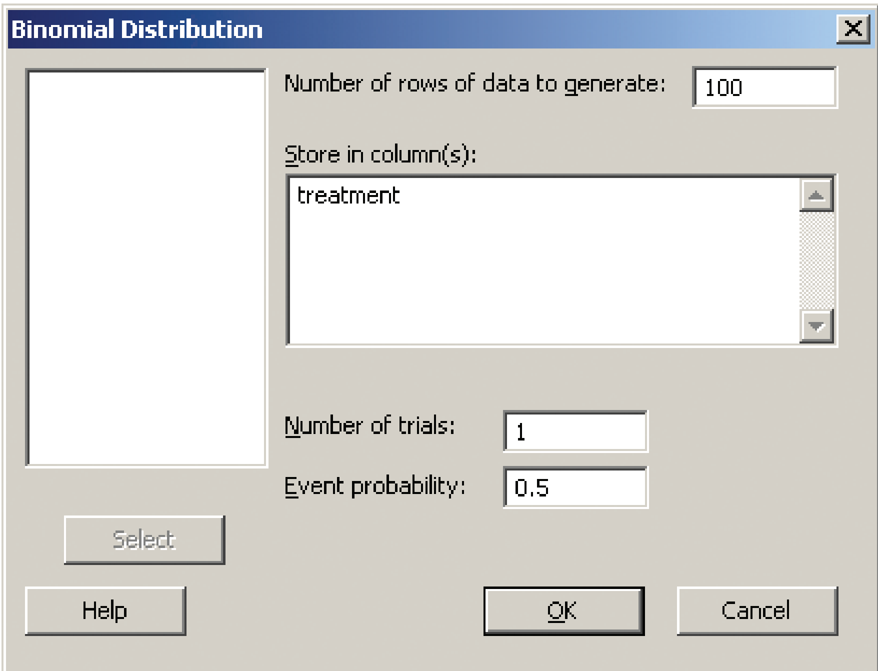


Figure 46 The **Binomial Distribution** dialogue box

- Click on **OK**.

Your randomisation should appear in the column labelled `treatment`. Create a contingency table of the numbers of participants assigned to each group at each centre (**Stat** > **Tables** > **Cross Tabulation and Chi-Square**).

What can you say about the balance of the groups?

Computer activity 95 *Assigning patients randomly but keeping the totals fixed*

You will now assign trial participants by fixing the total in each group (50 in each), but ordering randomly. This is done by first assigning treatments to participants systematically, and then shuffling these assignments.

- Make sure that the worksheet **randomisation.mtw** is the active worksheet.
- If you want to get exactly the same answers as given in the solution, set the seed for the random number generator to the value 42 (**Calc** > **Set base**).
- Create a column `treatment` in which participants are assigned to treatment groups systematically (**Calc** > **Make Patterned Data** > **Simple Set of Numbers**). (This was done in Computer activity 93.)
- Click on **Calc**, then **Random Data** and then **Sample From Columns. . .** . The **Sample From Columns** dialogue box will now appear.
- In Subsection 4.2, you used this same dialogue box to select a sample from a population. If a sample is specified to be the same size as the population, this results in the 'sample' just being a randomly shuffled copy of the 'population'. So in the **Sample From Columns** dialogue box, enter `100` in the **Number of rows to sample** field, `treatment` in the **From columns** field and `treatment` in the **Store samples in** field.
- Click on **OK**.

Your randomisation will appear in the column labelled `treatment`. Create a contingency table of the numbers of participants assigned to each group at each centre (**Stat** > **Tables** > **Cross Tabulation and Chi-Square**).

What can you say about the balance of the groups?

In Computer activity 95, you randomly assigned participants to treatment groups in such a way that the total number of participants in each treatment group was fixed. However, as you saw, this did not guarantee that the allocation was balanced. To achieve this, a stratified randomisation is required. There are different ways of achieving this using Minitab, however the most straightforward is just to do the same as you did in Computer activity 95 – but deal with each centre separately. We will not ask you to do stratified randomisation in M140.

This subsection has demonstrated most of the randomisation schemes described in Subsection 3.4 of Unit 11, which would be suitable for a group-comparative design. Matched-pairs and crossover designs require randomisation to be done in sets of pairs, but the procedure is similar. Remember that forcing balance may only be required when numbers within each strata are small. Packages are available that allow stratified randomisation to be carried out much more easily than in Minitab.

11.2 Analysis of data from clinical trials

In this subsection you will use Minitab to analyse data from a couple of clinical trials.

Computer activity 96 *A new treatment for pneumonia?*

A phase 3 clinical trial was carried out to test the effectiveness and safety of a new antibiotic drug, ceftaroline, against an existing antibiotic drug, ceftriaxone, in curing patients with pneumonia. Patients hospitalised with pneumonia were randomised to the experimental group – 613 received ceftaroline – or the control group – 615 received ceftriaxone. Participants who could be evaluated for a cure had to have received their treatment for between two and seven days and to have completed a test of cure visit (8 to 15 days after hospitalisation) with results that could be evaluated: 459 patients were evaluated for ceftaroline and 449 for ceftriaxone. The main effectiveness outcome was cure: bacteria that caused pneumonia in the patient was eliminated at the time of the 'test of cure' visit. Data on safety outcomes were also published, and the most common reported adverse event in this study was diarrhoea.

(Source: File, T.M. et al. (2010) 'Integrated analysis of FOCUS 1 and FOCUS 2: randomized, doubled-blinded, multicenter phase 3 trials of the efficacy and safety of ceftaroline fosamil versus ceftriaxone in patients with community-acquired pneumonia', *Clinical Infectious Diseases*, vol. 51, issue 12, pp. 1395–1405.)

These data are given in the Minitab worksheet **ceftaroline.mtw**. Open this worksheet. Notice the worksheet contains columns that relate to two different contingency tables. One contingency table, consisting of the columns `treatment`, `cure` and `nocure`, relates to the data on the effectiveness of ceftaroline. The other contingency table, consisting of the columns `treatment`, `diarrhoea` and `nodiarrhoea`, relates to the data on the safety of ceftaroline.

(a) Why is the χ^2 test an appropriate hypothesis test to be using with either of these contingency tables?

(b) Use Minitab to a carry out a χ^2 test to test the following null hypothesis:

H_0: No difference in pneumonia cure rate in patients taking ceftaroline and ceftriaxone.

(c) Use Minitab to a carry out a χ^2 test to test the following null hypothesis:

H_0: No difference in the chances of diarrhoea in patients taking ceftaroline and ceftriaxone.

In Unit 10, you learned about t-tests. In the following activity you will analyse some data from a clinical trial using a t-test.

Computer activity 97 *Does stretching improve movement of a hip joint?*

Ankylosing spondylitis is a form of arthritis that can affect the hip joint. A clinical trial was conducted at the Royal National Hospital for Rheumatic Diseases, in Bath, to see if some extra stretching exercises could be of benefit to such patients. The trial split subjects randomly into two groups: those given a standard treatment (the control group) and those given some stretching exercises as well as the standard treatment (the treatment group).

As part of this trial, the amount of lateral rotation (in degrees) in subjects' right hips was recorded before and after some treatment. Hence the change in lateral rotation could be worked out for every subject.

(Source: Chatfield, C. (1988) *Problem Solving: A statistician's guide*, Chapman & Hall.)

(a) What type of trial is this: crossover, matched-pairs or group-comparative? Justify your answer.

(b) What type of data are the 'changes in lateral rotation'?

(c) Given your answers to parts (a) and (b), which type of t-test is likely to be most appropriate for analysing the data: two-sample, one-sample or matched-pairs? (Assume here that any relevant population distribution is normal.)

(d) The null and alternative hypotheses are as follows:

$$H_0\colon\ \mu_c = \mu_t$$
$$H_1\colon\ \mu_c \neq \mu_t,$$

where μ_c is the population mean change in lateral rotation for patients given standard treatment and μ_t is the population mean change in lateral rotation for patients given stretching exercises and the standard treatment. Does this correspond to a two-sided or one-sided test?

(e) The data are given in the file **hips.mtw**. Use Minitab to test the hypotheses given in part (d). What are your conclusions about the effect of the stretching exercises.

(f) Give a 95% confidence interval for $\mu_t - \mu_c$. Why does this interval agree with the conclusion you came to in part (e)?

(g) When carrying out parts (e) and (f), you will have used a pooled standard deviation. Was this reasonable? Why or why not?

Summary of Chapter 11

In this chapter, you have learned how to implement different randomisation schemes that come up in clinical trials: systematic allocation, simple randomisation, random assignment and random assignment with fixed totals. You have also seen how the results from clinical trials can be analysed using the χ^2 test for contingency tables and by using a t-test.

12 Binomial distribution and two-sample tests

This chapter, which accompanies Unit 12, extends a couple of techniques in Minitab that you have already met. In Subsection 12.1, you use Minitab to explore the shape of the binomial distribution for different sample sizes and success probabilities. Then, in Subsection 12.2, you will learn how to perform two-sample (unpaired) t-tests that do not rely on the assumption that the samples come from populations whose variances are equal.

12.1 More on the binomial distribution

In Subsection 6.1, you learned how to use Minitab to calculate probabilities from a binomial distribution. However, Subsection 6.1 only dealt with binomial distributions where the probability of success (p) is 0.5. Here you will learn how to obtain probabilities from binomial distributions whatever the value of p. (Note that because p is a probability, only values between 0 and 1 make sense.)

One purpose of this subsection is to compare the results from the different computer activities. Hence it is best to do all the computer activities in this subsection in a single session.

Computer activity 98 *Binomial distribution with $p = 0.7$ and $n = 10$*

Suppose 'DontPanic' is a small driving school and that, of the people whom they teach, 70% pass the driving test at the first attempt. Further suppose that 10 of DontPanic's pupils are due to take the driving test for the first time next month. The number of these pupils who pass next month, x, follows a binomial distribution with $p = 0.7$ and $n = 10$. Obtain this probability distribution, in a column called `prob1`, by doing the following.

- Open a new worksheet in Minitab. (If you have not just started a new session, create a new worksheet via **File > New**.)
- In the worksheet, create a column called `x` with the entries 0, 1, ..., 10, to represent the values that x can take. (**Calc > Make Patterned Data > Simple Set of Numbers**. Then, in the **Simple**

Set of Numbers dialogue box, put `x` in the **Store patterned data in** field, `0` in the **From first value** field, `10` in the **To last value** field and make sure that the **In steps of** field contains `1`.)

- Click on **Calc**, then **Probability Distributions** and then **Binomial. . . .**
- In the **Binomial Distribution** dialogue box, select **Probability** as we wish to calculate probabilities.
- We are interested in a sample of 10 pupils, and the probability that a DontPanic pupil passes first time is 0.7. So enter `10` in the **Number of trials** field and enter `0.7` in the **Event probability** field.
- Fill in the rest of the dialogue box as you would for a binomial distribution when $p = 0.5$ and when the probabilities are to be stored in a column called `prob1` (see Computer activity 53, Subsection 6.1). That is, enter `x` in the **Input column** field and type `prob1` in the **Optional storage** field.

  The dialogue box should now look like Figure 47.

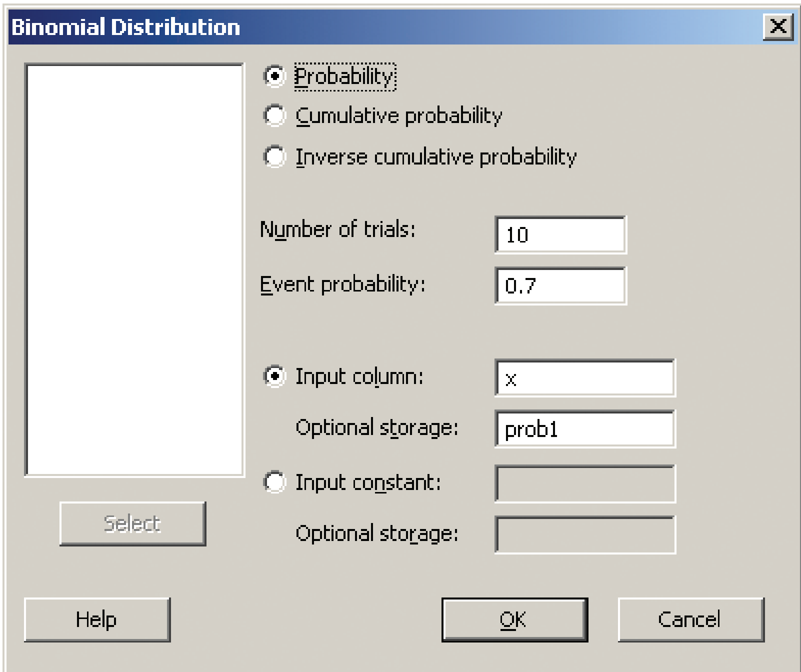


Figure 47 The **Binomial distribution** dialogue box

- Click on **OK**.

A new column of numbers appears in the worksheet. These numbers are the values of $P(x = 0)$, $P(x = 1), \ldots, P(x = 10)$. As an example, you should find that 0.266 828 is the probability that exactly 7 of the 10 pupils pass first time.

You will need the results in this worksheet for Computer activity 99. So, if possible, carry straight on to Computer activity 99.

Computer activity 99 *Picturing a binomial distribution*

(a) In Computer activity 54 (Subsection 6.1) you produced a bar chart of a binomial distribution when the probability of success is 0.5. By doing the same as you did in that computer activity, produce a bar chart of the binomial probability distribution you obtained in Computer activity 98 (**Graph** > **Bar Chart**). That is, produce a bar chart of the binomial distribution with $p = 0.7$ and $n = 10$.

(b) Describe the shape of the binomial distribution with $p = 0.7$ and $n = 10$: how many modes does it have and is it symmetric, left-skew or right-skew?

Computer activity 100 *Binomial distribution with $p = 0.9$ and $n = 10$*

Suppose 90% of the people who take driving lessons from DontPanic eventually pass the driving test in three or fewer attempts. Consider a random sample of 10 people taught by DontPanic who take their test for the first time in a particular month. Now let x denote the number of these people who eventually pass the driving test in three or fewer attempts.

(a) Obtain the probabilities for the number of people taught by DontPanic who pass the driving test in three or fewer attempts. That is, obtain the binomial distribution with $p = 0.9$ and $n = 10$. Store these probabilities in a column labelled `prob3` and then display these probabilities as a bar chart.

(b) Compare the shape of this binomial distribution ($p = 0.9$ and $n = 10$) with the binomial distribution you plotted in Computer activity 99 ($p = 0.7$ and $n = 10$).

In Computer activities 99 and 100, you have seen how the shape of the binomial distribution changes when p (but not n) changes. The distribution when $p = 0.9$ (and $n = 10$) is more (left-)skew than the distribution with $p = 0.7$ (and $n = 10$). This is because a binomial distribution is symmetric for $p = 0.5$ and becomes increasingly more skew as the probability of success moves further away from 0.5 (when the value of n does not change). When $p > 0.5$, the probability distribution is left-skew. When $p < 0.5$, the probability distribution is right-skew, and when $p = 0.5$ the probability distribution is symmetric.

What happens when the value of n changes? This is what you will consider in the next activity.

Computer activity 101 *Binomial distribution with $p = 0.7$ and $n = 60$*

Now consider a six-month period, rather than one month. Suppose 60 of DontPanic's pupils take the driving test for the first time in the next six months. The number who will pass, x, follows a binomial distribution with $p = 0.7$ and $n = 60$.

(a) Determine the probabilities of a binomial distribution with $p = 0.7$ and $n = 60$. Store these probabilities in a column labelled `prob6month` and display them as a bar chart.

(b) Comment on the shape of this distribution in comparison to: (i) a binomial distribution with $p = 0.7$ and $n = 10$, and (ii) a normal distribution.

In Computer activity 101 you will have found that the binomial distribution with $p = 0.7$ and $n = 60$ is more symmetric than the binomial distribution with $p = 0.7$ and $n = 10$, and that it has a similar shape to the normal distribution. In fact, the binomial distribution becomes steadily more symmetric as the sample size (n) increases, assuming the probability of success (p) remains the same. Further, the binomial distribution has approximately the same shape as a normal distribution when the sample size is large enough.

12.2 More on two-sample tests

In Section 5 of Unit 12, various tests for comparing two population means were reviewed and a new test was introduced. This latter test should be used when there are two samples (which are not matched pairs) from populations that have normal distributions and *we do not want to make the assumption that the variances of the two populations are equal*. We explore this test further in this section and compare its results with those obtained when the assumption is made that the two population variances are equal.

It is best to do Computer activities 102 to 104 in a single session. In Computer activity 102 you will draw boxplots for two samples of data and compare their variances. In Computer activity 103 you will then test for equality of the two population means using the new test (no assumption is made that the population variances are equal). Then, in Computer activity 104, you will again test whether the population means are equal, but under the assumption that the population variances are equal. You are then asked to comment on the result of that test in comparison with the result of the test in Computer activity 103.

In Computer activity 105 you perform the same two tests, but for data that is given in the form of summary statistics (as was done for the one-sample z-test in Subsection 7.3).

Computer activity 102 *Multiple boxplots of two samples*

Open the worksheet **Rtimes.mtw** and make sure it is the active window. This worksheet contains data of the time (in seconds) taken for aspirin tablets to release 50% of the pain-killing agent ('release times'). The tablets are from two manufacturers, A and B (with 10 from A and 20 from B). The times are given in the worksheet as `ManA` and `ManB`.

(a) Produce a single diagram that contains a boxplot for `ManA` and a boxplot for `ManB`. (The method for getting Minitab to draw such boxplots was described in Computer activity 36 in Subsection 3.3 (**Graph** > **Boxplot**).)

(b) On the basis of the boxplots in part (a):

- Is it reasonable to believe that each sample comes from a population that has a normal distribution?
- Could the two samples have come from populations with the same variance?

(c) Use Minitab to get the variances of `ManA` and `ManB`. (This was covered in Computer activity 32 in Subsection 3.2 (**Stat** > **Basic Statistics** > **Display Descriptive Statistics**).)

Compare the variances you obtain – is it reasonable to assume that they come from populations whose variances are equal?

Suppose that a regulator is interested in whether the times to release 50% of the active agent achieved by tablets from manufacturer A and from manufacturer B are, on average, the same. In other words, the regulator is interested in testing the null hypothesis

$$H_0\colon \mu_A = \mu_B,$$

against the alternative hypothesis

$$H_1\colon \mu_A \neq \mu_B,$$

where μ_A denotes the mean time taken by a tablet produced by manufacturer A to release 50% of the pain-killing agent and μ_B denotes the corresponding mean time taken by tablets produced by manufacturer B.

In Computer activity 102 you saw that the sample of tablets from manufacturer A is small (10 tablets) but that it is reasonable to assume that the distributions of release times for tablets from both manufacturers are normal. (Or at least it is not unreasonable to assume this!) These both suggest that the two-sample t-test is a reasonable test to use.

However, you also saw in Computer activity 102 that it is *not* reasonable to assume that the variances for the two populations are equal. So a version of the two-sample t-test that does not make this assumption is required. This is what you will use in Computer activity 103.

Computer activity 103 *Two-sample t-test and confidence interval when population variances are not equal*

A two-sample t-test is required that does not make the assumption that population variances are equal based on the sample data in **Rtimes.mtw** to test the hypothesis

$$H_0\colon\ \mu_A = \mu_B,$$

against the alternative hypothesis

$$H_1\colon\ \mu_A \neq \mu_B.$$

The procedure for doing this test in Minitab is very similar to the procedure for doing a two-sample t-test which does make the assumption of equal population variances (Subsection 10.2). So first do the following.

- With **Rtimes.mtw** as the active window, click on **Stat**, choose **Basic Statistics** and then **2-Sample t. . .** .
- In the resulting **2-Sample t (Test and Confidence Interval)** dialogue box, select **Samples in different columns** and then enter `ManA` in the **First** field and `ManB` in the **Second** field.

Now, for these data, the assumption that the population variances are equal is not reasonable. So for this two-sample t-test do the following.

- Make sure that **Assume equal variances** is *not* selected. The completed **2-Sample t (Test and Confidence Interval)** dialogue box is shown in Figure 48.

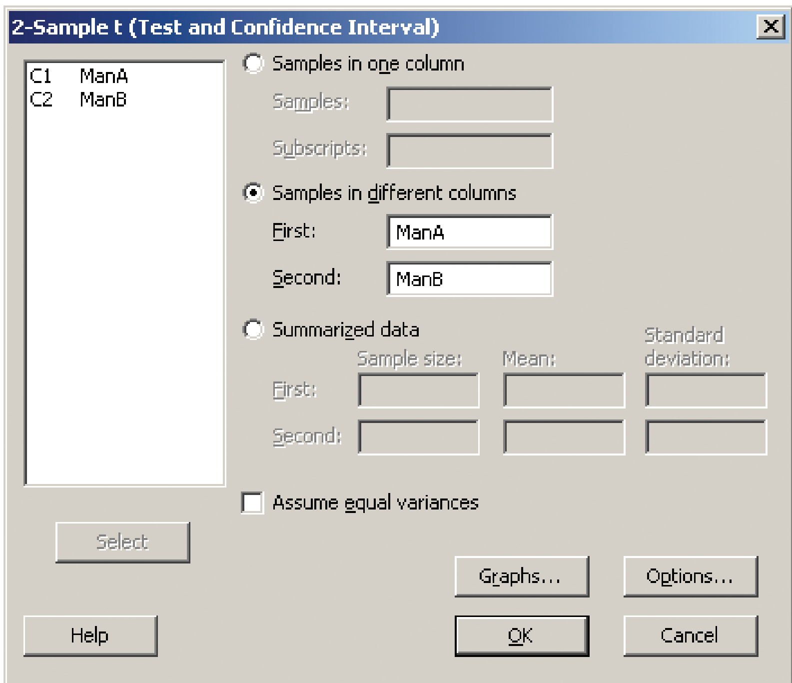


Figure 48 The **2-Sample t (Test and Confidence Interval)** dialogue box indicating that equal variances are not to be assumed

- Click on **OK**.

Using the output given by Minitab, answer the following.

(a) What is the value of the test statistic, and what is the value of the degrees of freedom that Minitab has calculated?

(b) What is the p-value from the test? What conclusion should be drawn from the test?

(c) What is the 95% confidence interval for $\mu_A - \mu_B$, the difference between the two population means? How does this relate to the result of the hypothesis test?

In Computer activity 103 the two sample t-test you performed did not make the assumption that the population variances were equal. But what difference does this assumption make? This is what you will explore in the next computer activity.

Computer activity 104 *Making the (unreasonable) assumption of equal population variances*

Test the same hypotheses as in Computer activity 103, but this time making the assumption that the samples come from two populations whose population variances are equal.

(a) What is the value of the test statistic, and what is the value of the degrees of freedom? How do these compare to the values found in Computer activity 103?

(b) What is the p-value from this test? If the assumptions made were correct, what conclusion would you draw from this test?

(c) Comment on the result of this test in relation to the result found in Computer activity 103.

Comparing the results obtained in Computer activities 103 and 104 illustrates the importance of checking whether population variances should be assumed to be equal. For these data and hypotheses, the conclusions drawn depend on whether or not the population variances are assumed to be equal.

So which set of conclusions are right? Well, in this case, the rule of thumb suggests that the population variances are not equal. So the conclusions drawn from the t-test which does not make this assumption (that is, little evidence that the population means are different) are more defensible.

However, to put this result in perspective, we should mention that often the two forms of the two-sample t-test yield the same conclusion. When this happens, it means that the assumption that the population variances are not equal is not all that important. In such cases, it does not really matter if they are or not!

The following computer activity gives a further comparison of the two forms of the two-sample t-test. Furthermore, in this activity you will perform the two-sample t-test by entering summary statistics to perform the test. Thus, it will be very similar to the way you performed the one-sample z-test in Minitab (Subsection 7.3).

Computer activity 105 *A further comparison of the two t-tests*

Suppose treatments A and B are two treatments for the common cold. Further suppose that summary statistics for the length of time (in days) some patients had symptoms when they were either taking treatment A or treatment B are given in Table 7.

Table 7 Summary statistics for the length of time with symptoms for patients with the common cold

	Sample size	Sample mean	Sample standard deviation
Treatment A	12	6.2	1.5
Treatment B	20	8.1	2.8

As in Computer activity 103 and 104, we want to test the null hypothesis

$$H_0\colon\ \mu_A = \mu_B,$$

against the alternative hypothesis

$$H_1\colon\ \mu_A \neq \mu_B,$$

where μ_A and μ_B are the means of the populations from which the samples were taken. This corresponds to testing whether, on average, the two treatments are equally effective.

Using the information in Table 7, it is not possible to decide if it is reasonable that the population distribution for the length of time with symptoms in each treatment group is normal. In such cases, other knowledge about the data has to be used. For example, in this case assume that clinicians believe that for either treatment the population distribution of the length of time with symptoms is normal.

(a) Perform a two-sample t-test that makes no assumption about the population variances being equal. Do this using the following steps.

- Obtain the **2-Sample t (Test and Confidence Interval)** dialogue box (**Stat** > **Basic Statistics** > **2-Sample t**).
- Select **Summarized data**.
- Fill in the boxes with the sample size, mean and standard deviation of each sample.

- Make sure that **Assume equal variances** is *not* selected. The completed dialogue box should be as in Figure 49.

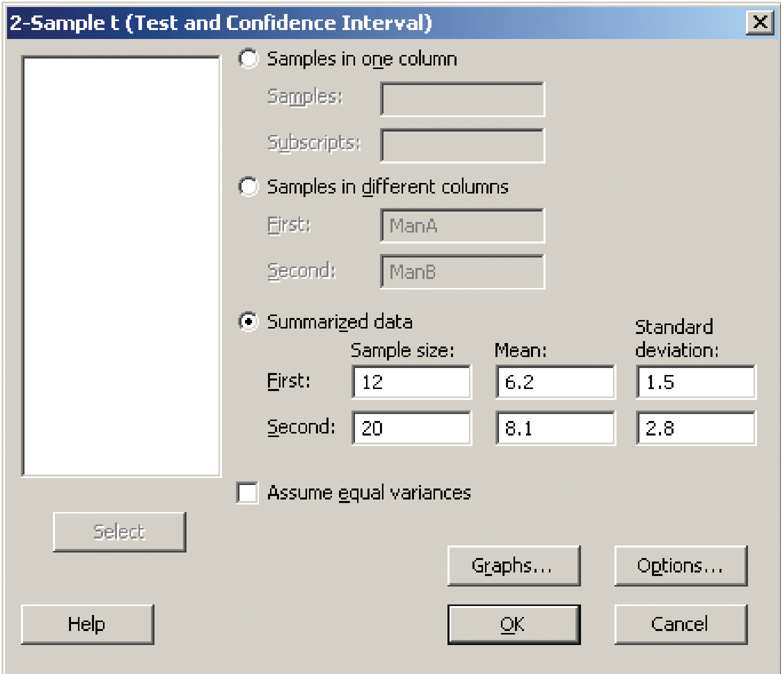


Figure 49 Dialogue box for a two-sample t-test using summary statistics

- Click on **OK**.

What is the p-value from the test? Assuming its underlying assumptions are satisfied, what conclusion should be drawn from the test?

(b) Perform a two-sample t-test, but this time make the assumption that the population variances are equal.

In other words, obtain the **2-Sample t (Test and Confidence Interval)** dialogue box again (**Stat** > **Basic Statistics** > **2-Sample t**). You should find that all of the fields are filled in for you. However, this time make sure that **Assume equal variances** *is* selected.

What is the p-value from this test? Assuming its underlying assumptions are satisfied, what conclusion should be drawn from the test?

(c) Compare the results of the two t-tests you have performed.

(d) Which of the two t-tests was the correct test to use for these data and hypotheses?

Summary of Chapter 12

In this chapter, you have extended your knowledge about two different tasks in Minitab: calculating binomial probabilities and performing a two-sample t-test.

You have learned how to calculate binomial probabilities whatever the value of the success probability, p, is assumed to be. You have seen that when p moves away from $p = 0.5$ and gets closer to $p = 1$ the probability distribution gets more skew. You have also seen that when the number of trials increases, the probability distribution becomes more symmetric.

For the two-sample t-test, you have extended the range of data to which it can be applied using Minitab. You have learned how to perform a version of this test that does not rely on the assumption of equal population variances. You have seen that whether or not this assumption is made can make a difference to the conclusions that are drawn from the test. You have also learned how to perform the test when the data are given in summarised form.

Minitab quick reference guide

The following list summarises the Minitab terms and commands used.

- Binomial probabilities (calculating): **Calc** > **Probability Distributions** > **Binomial**. (Computer activity 53 and Computer activity 55, Subsection 6.1; Computer activity 98, Subsection 12.1)
- Contingency table (creating): **Stat** > **Tables** > **Cross Tabulation and Chi-Square**. (Computer activity 93, Subsection 11.1)
- One-sided tests: In the dialogue box for the test, click on **Options...** and then select either **greater than** or **less than** in the **Alternative** field of the dialogue box. (Computer activity 90 and Computer activity 91, Subsection 10.4)
- Random allocation: **Calc** > **Random Data** > **Binomial**. (Computer activity 94, Subsection 11.1)
- Random allocation (keeping totals fixed): Use systematic allocation and then **Calc** > **Random Data** > **Sample from Columns**. (Computer activity 95, Subsection 11.1)
- Setting the seed (of the random number generator): **Calc** > **Set base**. (Computer activity 92, Subsection 11.1)
- Systematic allocation: **Calc** > **Make Patterned Data** > **Simple Set of Numbers**. (Computer activity 93, Subsection 11.1)
- t-test (matched-pairs): **Stat** > **Basic Statistics** > **Paired t**. (Computer activity 88, Subsection 10.3)
- t-test (one-sample): **Stat** > **Basic Statistics** > **1-Sample t**. (Computer activity 84, Subsection 10.1)
- t-test (two-sample): **Stat** > **Basic Statistics** > **2-Sample t**. If the assumption of equal population variances is to be made, make sure that **Assume equal variances** is selected in the dialogue box. (Computer activity 86, Subsection 10.2; Computer activity 103 and Computer activity 104, Subsection 12.2)

Solutions to computer activities

Solution to Computer activity 85

(a) The confidence interval is obtained by entering `speed` in the **Samples in columns** field of the **1-Sample t (Test and Confidence Interval)** dialogue box (**Stat** > **Basic Statistics** > **1-Sample t**). Note that the **Perform hypothesis test** option does not need to be selected.

The 95% confidence interval is given as `(709.9, 802.5)` and hence the 95% confidence interval for the speed of light in air (in km/s) is (299 709.9, 299 802.5).

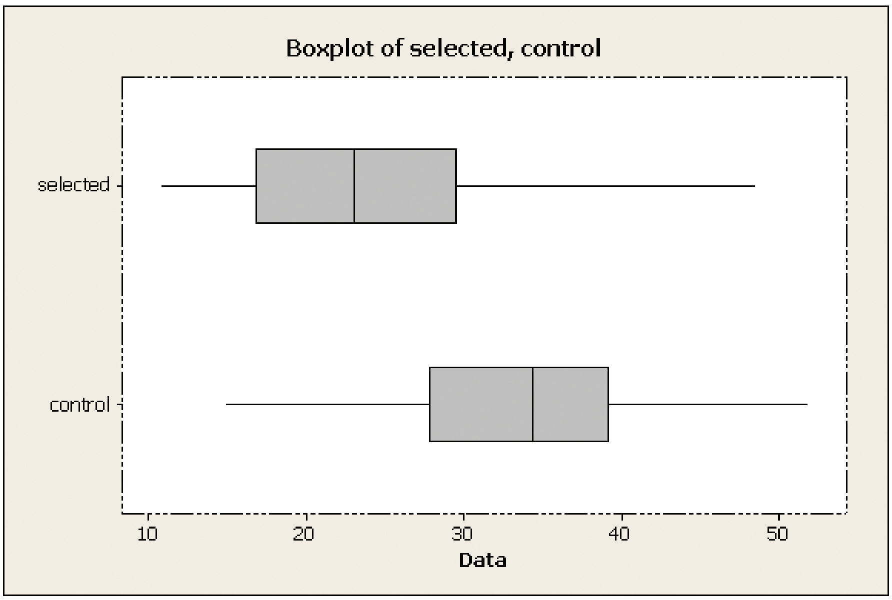
(b) This time in the **1-Sample t (Test and Confidence Interval)** dialogue box, the **Perform hypothesis test** option needs to be selected and `705` entered in the **Hypothesized mean** field.

The value of the test statistic (`T` in the output) is 2.29 and the p-value is 0.032. So, from Table 1 (Subsection 6.2), there is moderate evidence against the null hypothesis. This means there is moderate evidence that the claim $\mu = 705$ is not correct – at least for the conditions which applied when Michelson was taking his measurements. An alternative explanation for this result is that the method used by Michelson to measure the speed of light was slightly biased.

(c) The value 705 is not in the 95% confidence interval that was given by Minitab. So the p-value for the one-sample t-test must be less than 0.05 – as found in part (b).

Solution to Computer activity 87

(a) The diagram containing both boxplots is as follows.



These boxplots were obtained by entering both `selected` and `control` in the **Graph variables** field in the **Boxplot − Multiple Y's, Simple** dialogue box (**Graph** > **Boxplot** and selecting **Multiple Y's, Simple**). Also, by making sure that the **Transpose value and category scales** option is selected in the **Boxplot - Scale** dialogue box (obtained by clicking on the **Scale...** button).

Both boxplots appear to be reasonably symmetric and there are no potential outliers indicated. So it is reasonable to assume that the data in both groups are samples from normal population distributions.

(b) The output produced by Minitab is as follows.

```
Two-sample T for selected vs control

           N   Mean  StDev  SE Mean
selected  50  24.44   8.78      1.2
control   25  33.37   8.94      1.8


Difference = mu (selected) - mu (control)
Estimate for difference:  -8.93
95% CI for difference:  (-13.24, -4.62)
T-Test of difference = 0 (vs not =): T-Value = -4.13  P-Value = 0.000  DF = 73
Both use Pooled StDev = 8.8303
```

This was obtained by selecting the **Samples in different columns** and **Assume equal variances** options of the **2-Sample t (Test and Confidence Interval)** dialogue box (**Stat** > **Basic Statistics** > **2-Sample t**) and entering `selected` and `control` in the **First** and **Second** fields respectively.

So the test statistic is -4.13 and the corresponding p-value is given as 0.000. (Remember that a p-value of 0.000 means that $p < 0.0005$, not that p is exactly equal to zero.) This means that from Table 1 (Subsection 6.2), there is very strong evidence against the null hypothesis. In other words, there is very strong evidence that the egg-laying capabilities in the two groups of fruit fly are not the same. In fact, the egg-laying capability of selectively bred fruit flies appears to be lower than that of fruit flies that have not been selectively bred.

(c) The standard deviation of eggs laid in the selective breeding group is given as 8.78, and hence the variance is $8.78^2 = 77.0884$. Similarly, the variance of eggs laid in the group that had not been selectively bred group is 79.9236. The ratio of the bigger variance to the smaller is much less than 3. So, by the rule of thumb introduced in Subsection 3.3 of Unit 10, it is reasonable to have used the pooled standard deviation for the test completed in part (b).

(d) The 95% confidence interval for $\mu_s - \mu_c$ can be read off directly from the output obtained in part (b). It is $(-13.24, -4.62)$.

Solution to Computer activity 89

(a) The data consist of pairs of observations, one pair for each sample. The matched-pairs t-test is a good choice for such data because interest focuses on the difference within pairs, and not between pairs.

(b) The output from the matched-pairs t-test is as follows.

```
Paired T for A - B

             N    Mean  StDev  SE Mean
A           10   98.10  25.41     8.03
B           10  105.00  24.49     7.75
Difference  10   -6.90  10.87     3.44


95% CI for mean difference: (-14.67, 0.87)
T-Test of mean difference = 0 (vs not = 0): T-Value = -2.01  P-Value = 0.076
```

This output was obtained by selecting the **Samples in columns** option on the **Paired t (Test and Confidence Interval)** dialogue box (**Stat** > **Basic Statistics** > **Paired t**) and entering `A` and `B` in the **First sample** and **Second sample** fields respectively. (You may have to remove `formL` and `formR` from the **First sample** and **Second sample** fields first.)

The test statistic is therefore -2.01 and the corresponding p-value is 0.076. From Table 1 (Subsection 6.2) this means there is weak evidence that the two devices are not giving the same value on average.

(c) From the output obtained in (b), the 95% confidence interval for the difference in readings is $(-14.67, 0.87)$. Note that this interval contains 0, which is the value that corresponds to no difference between the devices. This ties in with the conclusion you drew in part (b) – there is only weak evidence of a difference between the devices.

Solution to Computer activity 91

(a) The null and alternative hypotheses are

$$H_0\colon\ \mu = 6.8$$
$$H_1\colon\ \mu < 6.8,$$

where μ is the population mean sleeping time (in hours) of American fibromyalgia sufferers.

(b) The hypotheses just involve one group of patients, American fibromyalgia sufferers. So an appropriate test is a one-sample test – either the one-sample z-test or the one-sample t-test.

As the sample size, 744, is much greater than 25, it is appropriate to apply the z-test.

(c) • In the **1-Sample Z (Test and Confidence Interval)** dialogue box (**Stat** > **Basic Statistics** > **1-sample Z**) select **Summarized data**.

- Type `744` in the **Sample size** field and `5.6` in the **Mean** field. Type `1.6` in the **Standard deviation** field. Make sure that **Perform hypothesis test** is selected. Type `6.8` in the **Hypothesized mean** field.
- Click on **Options...** in the **1-Sample Z (Test and Confidence Interval)** dialogue box.
- The alternative hypothesis corresponds to H_1: $\mu < 6.8$, so in the **1-Sample Z - Options** dialogue box, select **less than** in the **Alternative** field.
- Click on **OK** and on **OK** again.

The resulting Minitab output is:

```
Test of mu = 6.8 vs < 6.8
The assumed standard deviation = 1.6


                          95% Upper
  N    Mean  SE Mean        Bound       Z      P
744  5.6000   0.0587       5.6965  -20.46  0.000
```

The p-value associated with this one-sided z-test is, to three decimal places, 0.000. According to Table 1 (Subsection 6.2), this p-value, which is such that $0.001 \geq p$, gives very strong evidence against the null hypothesis. We conclude that there is very strong evidence that American fibromyalgia patients do indeed sleep for fewer hours per night, on average, than is the 'norm' for the entire USA population.

Solution to Computer activity 92

(a) When a member of the module team did this, the following numbers were generated.

45	67	38	80	6	50	78	72	75	15
61	62	65	50	77	4	28	27	93	59
98	47	47	74	27	66	49	45	71	85

This was done by entering the following in the **Integer Distribution** dialogue box (**Calc** > **Random Data** > **Integer**): `30` in the **Number of rows of data to generate** field, `C1` in the **Store in column(s)** field, `1` in the **Minimum value** field and `100` in the **Maximum value** field.

The numbers you get are likely to be different to these.

(b) When a member of the module team did this, the following numbers were generated.

6	44	59	58	100	77	54	55	94	72
42	92	94	92	30	57	21	42	11	70
9	12	92	40	67	50	75	78	34	37

This time, because the random number generator seed was first set to take the value 10, you should have generated exactly the same numbers.

(c) You should have ended up with exactly the same numbers in column `C3` as in column `C2`. Repeating part (b), including starting with the same seed, results in Minitab generating exactly the same 'random' numbers.

(d) You should have ended up with the following numbers in column `C4`.

```
87  44  94  73  22   1  70  57  50  75
46  58  86  73   1  29  12  65  99  31
95  54  16  78  35  64  21  86  96  86
```

These numbers are different to those generated in parts (b) and (c) because Minitab has started in a different place in its sequence of pseudo-random numbers. However, it turns out that the position chosen to start at is decided in an entirely deterministic way. So your set of 'random' numbers should be exactly the same as those obtained by the module team (and anyone else who does this activity!).

Solution to Computer activity 93

(a) Overall, 50 of the participants were assigned to the control group and 50 were assigned to the experimental group.

(b) Yes there is balance. There are 25 individuals in each of the control and experimental groups within each centre.

(c) There is no randomness in the assignment of participants to treatments. So not setting the seed for the random number generator will not make any difference.

Solution to Computer activity 94

There are 55 assigned to the control group and 45 to the experimental group. In centre 1, there are 29 assigned to the control group and 21 to the experimental group. In centre 2, there are 26 in the control group and 24 in the experimental group. This is not balanced, numbers in the experimental group are particularly low in centre 1.

Solution to Computer activity 95

There are 50 in each of the control and experimental groups, so this is balanced. In centre 1, there are 22 assigned to the experimental group, and in centre 2, there are 28 assigned to the experimental group. So this method of randomisation does not produce balance within centres.

Solution to Computer activity 96

(a) The χ^2 test is appropriate because the data are categorical.

(b) The output produced by Minitab for this χ^2 test is as follows.

```
Expected counts are printed below observed counts
Chi-Square contributions are printed below expected counts

         cure  nocure  Total
    1     387      72    459
       372.05   86.95
        0.600   2.570

    2     349     100    449
       363.95   85.05
        0.614   2.627

Total     736     172    908

Chi-Sq = 6.411, DF = 1, P-Value = 0.011
```

This was obtained by entering `cure` and `nocure` in the **Columns containing the table** field of the **Chi-Square Test (Table in Worksheet)** dialogue box (**Stat** > **Tables** > **Chi-Square Test (Two-Way Table in Worksheet)**).

The p-value for the test is 0.011, so by Table 1 (Subsection 6.2) there is moderate evidence that the treatment and being cured are not independent. Comparing the expected values with those observed, more patients in the ceftaroline group were cured than expected and fewer patients in the ceftriaxone were cured than expected. So there is a difference in pneumonia cure rate between the two treatment groups: ceftaroline appears to be more effective than ceftriaxone in curing pneumonia.

(c) This time the output produced by Minitab is as follows.

```
Expected counts are printed below observed counts
Chi-Square contributions are printed below expected counts

       diarrhoea  nodiarrhoea  Total
    1         26          587    613
           20.97       592.03
           1.209        0.043

    2         16          599    615
           21.03       593.97
           1.205        0.043

Total         42         1186   1228

Chi-Sq = 2.499, DF = 1, P-Value = 0.114
```

This is obtained in the same way as the output in part (a) was. The only difference is that this time `diarrhoea` and `nodiarrhoea` are entered in the **Columns containing the table** field.

This time there is little evidence against the null hypothesis ($p = 0.114$). So there is little evidence for a difference in the chances of diarrhoea between patients taking ceftaroline and ceftriaxone.

Solution to Computer activity 97

(a) This is a group-comparative trial. Each subject either received the standard treatment or the standard treatment plus stretching exercises, so it was not a crossover trial. Also, subjects in the trial were not specifically matched with other subjects in the trial, so it was not a matched-pairs trial.

(b) The 'changes in lateral rotation' are interval scale data as they correspond to measurements.

(c) Interval scale data from group-comparative trials can be analysed using two-sample t-tests.

(d) The hypotheses correspond to a two-sided test because the alternative hypothesis includes a 'not equals'. (For a one-sided test, the alternative hypothesis would either have included a 'greater than' or a 'less than'.)

(e) The output produced by Minitab is as follows.

```
Two-sample T for treatment vs control

            N  Mean  StDev  SE Mean
treatment  27  7.56   5.55      1.1
control    12  2.17   5.83      1.7


Difference = mu (treatment) - mu (control)
Estimate for difference:  5.39
95% CI for difference:  (1.43, 9.35)
T-Test of difference = 0 (vs not =): T-Value = 2.76  P-Value = 0.009  DF = 37
Both use Pooled StDev = 5.6337
```

This output was obtained by selecting the **Samples in different columns** option in the **2-Sample t (Test and Confidence Interval)** dialogue box (**Stat** > **Basic Statistics** > **2-Sample t**), entering `treatment` in the **First** field, entering `control` in the **Second** field and making sure that the **Assume equal variances** option is selected.

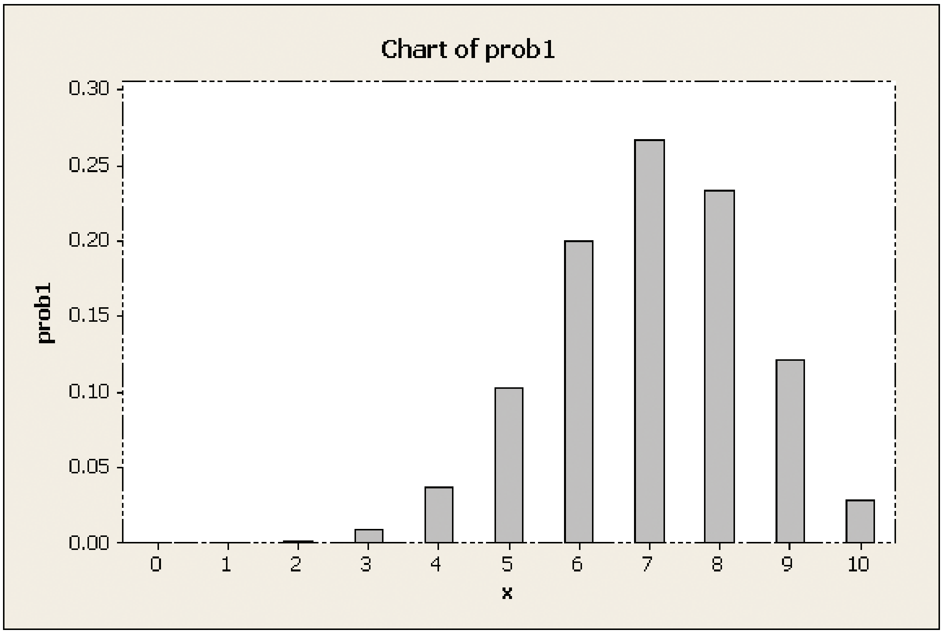
The p-value is between 0.01 and 0.001 so, using Table 1 (Subsection 6.2), there is strong evidence that the change in lateral rotation is not the same in the two groups. In fact, it appears that there is a greater change in lateral rotation when stretching exercises are given in addition to the standard treatment.

(f) From the output in (e), the 95% confidence interval for $\mu_t - \mu_c$ is $(1.43, 9.35)$. This agrees with the conclusion drawn in part (e) as the interval is entirely above zero.

(g) From the output produced by Minitab, the standard deviations for the change in lateral rotation in the treatment and control groups are as follows: $s_t = 5.55$ and $s_c = 5.83$. Thus the ratio of the larger to the smaller variance is $5.83^2/5.55^2 \simeq 1.10$. As this value is less than 3, it is reasonable to have used a pooled standard deviation in parts (e) and (f).

Solution to Computer activity 99

(a) The bar chart of the probability distribution obtained in Computer activity 98 is as follows.



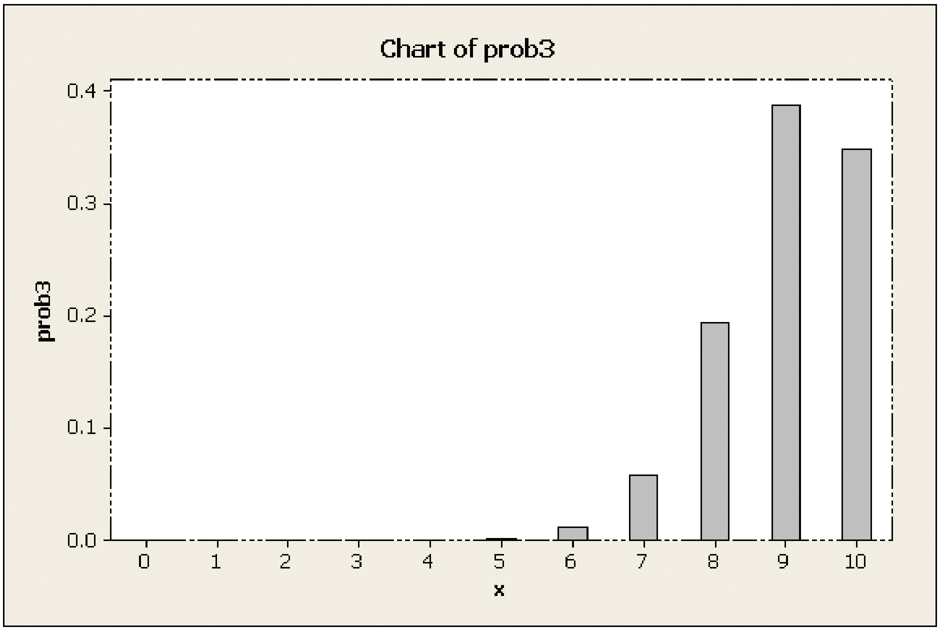
This bar chart was obtained by entering `prob1` in the **Graph variables** field and `x` in the **Categorical variable** field of the **Bar Chart – Values from a table, One column of values, Simple** dialogue box. (**Graph > Bar Chart**, select **Values from a table** for the **Bars represent** field and **Simple** as the form of bar chart.)

(b) The binomial distribution with $p = 0.7$ and $n = 10$ is unimodal, with the peak at about 7. In fact, all binomial distributions are unimodal and the peak is always near the value np. (In this case $np = 10 \times 0.7 = 7$.)

The binomial distribution with $p = 0.7$ and $n = 10$ is also left-skew.

Solution to Computer activity 100

(a) A bar chart of the probabilities is as follows



The probabilities were obtained by entering the following in the **Binomial Distribution** dialogue box (**Calc** > **Probability Distributions** > **Binomial**, with **Probability** selected): `0.9` in the **Event probability** field, `10` in the **Number of trials** field, `prob3` in the **Optional storage** field and `x` in the **Input column** field. As an example, 0.057 396 is the probability that exactly 7 of the pupils pass in three or fewer attempts.

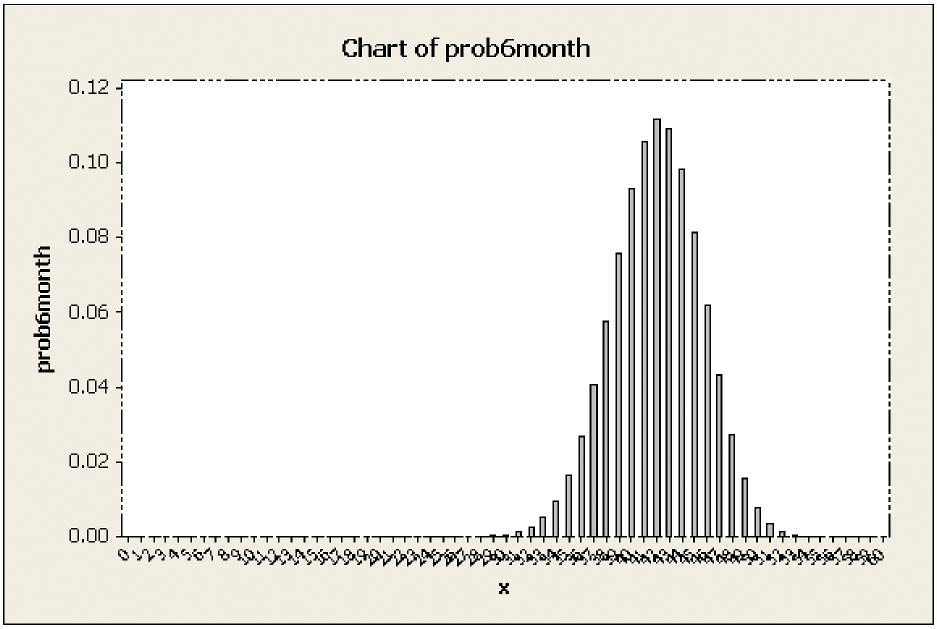
The bar chart was then obtained by entering `prob3` in the **Graph variables** field and `x` in the **Categorical variable** field of the **Bar Chart – Values from a table, One column of values, Simple** dialogue box. (**Graph** > **Bar Chart**, select **Values from a table** for the **Bars represent** field and **Simple** as the form of bar chart.)

(b) The binomial distribution with $p = 0.9$ and $n = 10$ is also unimodal, but this time the peak occurs around 9. This also corresponds to the value given by np.

The binomial distribution with $p = 0.9$ and $n = 10$ is more (left-)skew than the binomial distribution with $p = 0.7$ and $n = 10$.

Solution to Computer activity 101

(a) The bar chart that you should obtain is as follows.



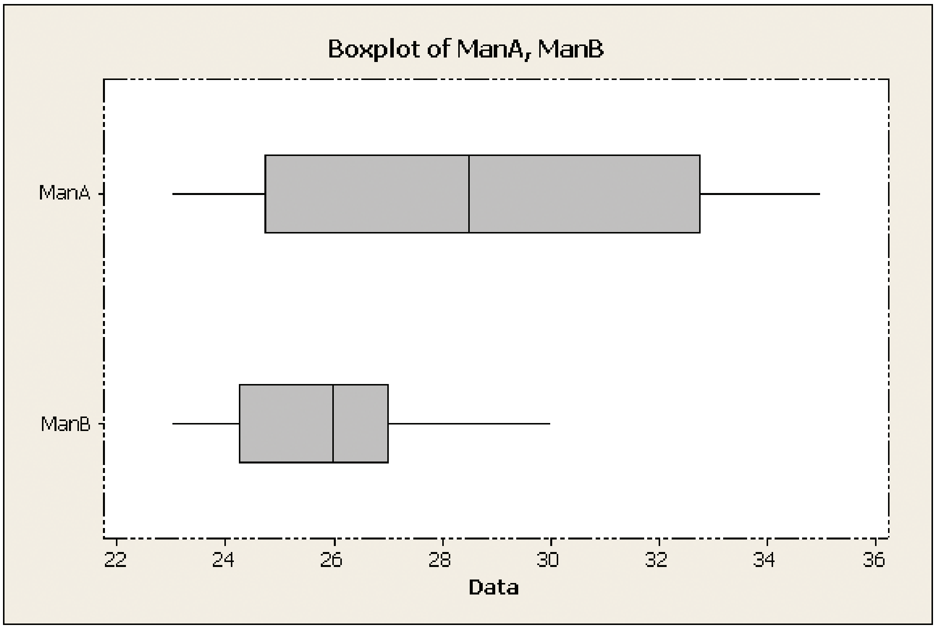
The probabilities of the binomial distribution with $p = 0.7$ and $n = 60$ were obtained following the same steps as in Computer activity 98, with the following exceptions: the values in the worksheet column labelled `x` run from 0 to 60, the value `60` is entered in the **Number of Trials** field and `prob6month` is entered in the **Optional storage** field. As an example, the 50th element of **prob6month**, corresponding to a value of $x = 49$, is 0.015 597.

The bar chart was then obtained following the same steps given in the solution to Computer activity 99, with the exception that `prob6month` was entered in the **Graph variables** field.

(b) The bar chart obtained in (a) shows that the distribution is almost symmetric. For (i), note that it is much more symmetric than the binomial distribution with $p = 0.7$ and $n = 10$ (see solution to Computer activity 99). For (ii), note that the distribution looks bell-shaped, like a normal distribution.

Solution to Computer activity 102

(a) The diagram that you obtain should look like the following.



This diagram was obtained by entering both `ManA` and `ManB` in the **Graph variables** field in the **Boxplot – Multiple Y's, Simple** dialogue box (**Graph > Boxplot** and selecting **Multiple Y's, Simple**). Also, by making sure that the **Transpose value and category scales** option is selected in the **Boxplot - Scale** dialogue box (obtained by clicking on the **Scale...** button).

(b)
- Neither boxplot has any extreme values and each sample has a distribution that seems reasonably symmetric. Thus, from the boxplots, the samples may come from populations that have normal distributions.
- The 'box' for the release times of tablets made by manufacturer A (`ManA`) is much bigger than the 'box' for the release times of tablets made by manufacturer B (`ManB`) and its whiskers cover a much bigger range (despite there being only 10 tablets in the sample from manufacturer A, while the sample from manufacturer B relates to 20 tablets). Hence the boxplots suggest it is unreasonable to believe that the samples come from populations whose variances are equal.

(c) The variance of `ManA` is 19.38 and the variance of `ManB` is 3.684.

These were obtained using the **Display Descriptive Statistics** dialogue box (**Stat > Basic Statistics > Display Descriptive Statistics**), with the variables `ManA` and `ManB` added to the **Variables** field and making sure **Variance** is selected in the **Display Descriptive Statistics - Statistics** dialogue box (obtained by clicking on the **Statistics...** button in the **Display Descriptive Statistics** dialogue box).

The ratio of the bigger variance to the smaller variance is

$$\frac{19.38}{3.684} \simeq 5.26.$$

Now, the rule of thumb first given in Subsection 3.3 of Unit 10 says that the population variances can be treated as equal if the larger variance is less than three times the size of the smaller variance. That is not the case with `ManA` and `ManB`.

So it is not reasonable to assume that the samples come from populations whose variances are equal. The release times of tablets from manufacturer A appear to be more spread out than the release times of tablets from manufacturer B.

Solution to Computer activity 103

(a) The output from the two-sample t-test for populations with unequal variances is as follows.

```
Two-sample T for ManA vs ManB

       N   Mean  StDev  SE Mean
ManA  10  28.60   4.40      1.4
ManB  20  26.00   1.92     0.43


Difference = mu (ManA) - mu (ManB)
Estimate for difference:  2.60
95% CI for difference:  (-0.65, 5.85)
T-Test of difference = 0 (vs not =): T-Value = 1.78  P-Value = 0.105  DF = 10
```

The test statistic, t, is given as 1.78. The value for the degrees of freedom is given as 10. Note that Minitab has used a complicated expression to come up with this value for the degrees of freedom, with truncation used to ensure that the final value is a whole number.

(b) Using the output given in part (a), the p-value is 0.105. According to Table 1 (Subsection 6.2), this corresponds to little evidence against the null hypothesis. In other words, there is little evidence that the tablets from the two manufacturers differ in the average time taken to release 50% of the active ingredient.

(c) Using the output given in part (a), the 95% confidence interval is $(-0.65\,\mathrm{s}, 5.85\,\mathrm{s})$. This interval contains the value 0, so the hypothesis test does not reject H_0 at the 5% significance level – consistent with the result found in part (b).

Solution to Computer activity 104

(a) The output from the two-sample t-test when the assumption of equal population variances is made is as follows.

```
Two-sample T for ManA vs ManB

       N   Mean  StDev  SE Mean
ManA  10  28.60   4.40      1.4
ManB  20  26.00   1.92     0.43


Difference = mu (ManA) - mu (ManB)
Estimate for difference:  2.60
95% CI for difference:  (0.26, 4.94)
T-Test of difference = 0 (vs not =): T-Value = 2.27  P-Value = 0.031  DF = 28
Both use Pooled StDev = 2.9544
```

This output was obtained in the same way as in Computer activity 103, with one crucial difference: in the **2-Sample t (Test and Confidence Interval)** dialogue box, **Assume equal variances** was selected.

The value of the test statistic is 2.27. This is higher than the value obtained when equal variances were not assumed. This difference is because the value of the ESE assuming a common variance must be bigger than the value of the ESE not assuming a common variance.

The value of the degrees of freedom is given as 28. This corresponds, as it should, to $n_1 + n_2 - 2$, where n_1 and n_2 are the two sample sizes.

So in this case, the degrees of freedom used when the assumption of a common variance is made is much larger than when the assumption is *not* made. (You saw in Computer activity 103 that the degrees of freedom was only 10 when the assumption of a common variance was not made.)

(b) Using the output given in part (a), the p-value from the test is 0.031. Thus this test rejects H_0 at the 5% significance level (and at the 3.1% significance level). If the assumptions underlying this test were correct, then from Table 1 (Subsection 6.2) there would be moderate evidence against H_0 – that is, moderate evidence that the tablets from the two manufacturers differ in the average time taken to release 50% of the active ingredient.

(c) Using the output given in part (a), the results of the two hypothesis tests are substantially different. With this dataset, the correct test (not assuming population variances are equal) finds little evidence of a difference, while moderate evidence of a difference is found when the dubious (in this case) assumption of equal population variances is made.

Solution to Computer activity 105

(a) The following is the output from the test that does not assume equal population variances.

```
Sample     N  Mean   StDev   SE Mean
1         12  6.20    1.50      0.43
2         20  8.10    2.80      0.63


Difference = mu (1) - mu (2)
Estimate for difference:  -1.900
95% CI for difference:  (-3.457, -0.343)
T-Test of difference = 0 (vs not =): T-Value = -2.50  P-Value = 0.018  DF = 29
```

The p-value from the test is 0.018. According to Table 1 (Subsection 6.2), this corresponds to moderate evidence against the null hypothesis. If the assumptions underlying the test are correct, there is moderate evidence that the population means differ. (The test assumes that observations come from populations that are approximately normally distributed.)

(b) The following is the output from the test that assumes the population variances are equal.

```
Sample     N  Mean   StDev   SE Mean
1         12  6.20    1.50      0.43
2         20  8.10    2.80      0.63


Difference = mu (1) - mu (2)
Estimate for difference:  -1.900
95% CI for difference:  (-3.694, -0.106)
T-Test of difference = 0 (vs not =): T-Value = -2.16  P-Value = 0.039  DF = 30
Both use Pooled StDev = 2.4063
```

The p-value from this test is 0.039. According to Table 1 (Subsection 6.2), this corresponds to moderate evidence against the null hypothesis. If its underlying assumptions are correct, there is again moderate evidence that the population means differ. (This test makes the assumptions made in part (a), and additionally assumes that the population variances are equal.)

(c) Although the values of the test statistic for these samples are different (-2.50 and -2.16), the values of the degrees of freedom are very similar (29 and 30).

The two tests give p-values that are fairly different (0.018 and 0.039). However, the resulting conclusions are the same.

Notice that the smaller p-value is given by the test that does not assume the population variances are equal. (It was the other way round in Computer activities 103 and 104.)

(d) The two variances are $1.5^2 = 2.25$ and $2.8^2 = 7.84$. As

$$\frac{7.84}{2.25} \simeq 3.48 > 3,$$

it should not be assumed that the two population variances are equal. Thus the test used in part (a) (that is, not assuming equal population variances) is the test that should be used.

However, the choice of test is not critical. Either version leads to the same conclusion – moderate evidence that the two treatments A and B are not equally effective. Treatment A appears to be more effective as it has the shorter average time with symptoms.