

MINITAB 14 Supplement for: Biostatistics for
the Health Sciences

R. Clifford Blair

and

Richard A. Taylor

December 30, 2007

© R. Clifford Blair and Richard A. Taylor 2005

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photocopying, recording, or otherwise, without the prior permission of R. Clifford Blair or Richard A. Taylor.

Preface

This manual is not an introduction to MINITAB[®] and contains only a portion of the material that would be contained therein.¹ Rather, this manual teaches the reader how to use MINITAB Release 14 for Windows to analyze data related to the problems addressed in *Biostatistics For The Health Sciences* by R. Clifford Blair and Richard A. Taylor.

To this end, MINITAB is used to analyze various of the data sets and problems contained in that text. The reader will be required, especially in the early part of the manual, to input selected data to the system. Additional data sets accompany this manual so that the reader will not be burdened with data entry once this skill is acquired.

It is assumed that the reader is familiar with the Microsoft Windows[®] operating system so that little instruction in this regard will be provided. It is further assumed that the reader is using MINITAB Release 14 rather than some earlier version.

Chapter 1 familiarizes the reader with some basic MINITAB concepts. Key among these is the entry and saving of data which is prerequisite to all that follows. Chapters 2 through 10 address the corresponding chapters in the text. That is, Chapter 2 in the manual addresses problems in Chapter 2 in the text and so on. Section headings in these latter chapters reflect the sections/pages in the text addressed in that section of the manual. Thus, while studying the text readers can quickly locate the comparable area in the manual by referring to the manual's table of contents. In addition, a manual appendix relates specific examples, data tables etc. in the text to specific pages in the manual. These provisions allow easy integration of MINITAB analyses as study of the text progresses.

Finally, we would caution that data analysis, whether by hand calculations or software is no substitute for a clear understanding of the underlying statistical concept. Such analysis can, however, be an invaluable aid to mastery of such concepts and is indispensable for research purposes.

¹The definitive introduction to the MINITAB[®] analysis system is found in *MINITAB[®] Handbook* by Barbara Ryan, Brian Joiner, and Jonathon Cryer published by Thomson.

Contents

1 Preliminaries	1
1.1 Introduction	1
1.2 Getting Started	1
1.3 Inputting Data	2
1.4 Performing Analyses	4
1.5 Editing and Printing Your Output	6
1.6 Using the Calculator	6
1.7 Graphs	8
Exercises	10
2 Descriptive Methods	11
2.1 Introduction	11
2.4 Distributions	11
2.5 Graphs	15
2.6 Numerical Methods	24
Exercises	30
3 Probability	31
3.1 Introduction	31
3.3 Contingency Tables	31
3.4 The Normal Curve	31
Exercises	33
4 Introduction to Inference and One Sample Methods	35
4.1 Introduction	35
4.2 Sampling Distributions	35
4.3 Hypothesis Testing	36
4.4 Confidence Intervals	42
Exercises	44
5 Paired Samples Methods	45
5.1 Introduction	45
5.2 Methods Related to Mean Difference	45
5.3 Methods Related to Proportions	46

5.4	Methods Related to Paired Samples Risk Ratios	48
5.5	Methods Related to Paired Samples Odds Ratios	49
	Exercises	52
6	Two Independent Samples Methods	53
6.1	Introduction	53
6.2	Methods Related to Differences Between Means	53
6.3	Methods Related to Proportions	55
6.4	Methods Related to Independent Samples Risk Ratios	56
6.5	Independent Samples Odds Ratios	56
	Exercises	57
7	Multi-Sample Methods	59
7.1	Introduction	59
7.2	The One-Way Analysis of Variance (ANOVA) F Test	59
7.3	The 2 By k Chi-Square Test	60
7.4	Multiple Comparison Procedures	61
	Exercises	63
8	The Assessment of Relationships	65
8.1	Introduction	65
8.2	The Pearson Product-Moment Correlation Coefficient	65
8.3	The Chi-Square Test for Independence	66
	Exercises	67
9	Linear Regression	69
9.1	Introduction	69
9.2	Simple Linear Regression	69
9.3	Multiple Linear Regression	70
	Exercises	73
10	Methods Based on the Permutation Principle	75
10.1	Introduction	75
10.3	Applications	75
	Exercises	78
A	Table Relating Text Items to Tasks	79

Chapter 1

Preliminaries

1.1 Introduction

The hand calculations you perform in conjunction with your studies of the statistical concepts in *Biostatistics For The Health Sciences* by R. Clifford Blair and Richard A. Taylor¹ are designed to enhance your understanding of the concepts under study. Modern statistical practice only rarely resorts to such methods for the analysis of data collected for research purposes. This is because (1) data sets collected in applied research contexts are often too large for hand calculations and (2) such calculations are prone to errors. Instead, researchers resort to tried and true statistical software that can be relied upon to provide quick, accurate results. The MINITAB system is one of the most popular of these.

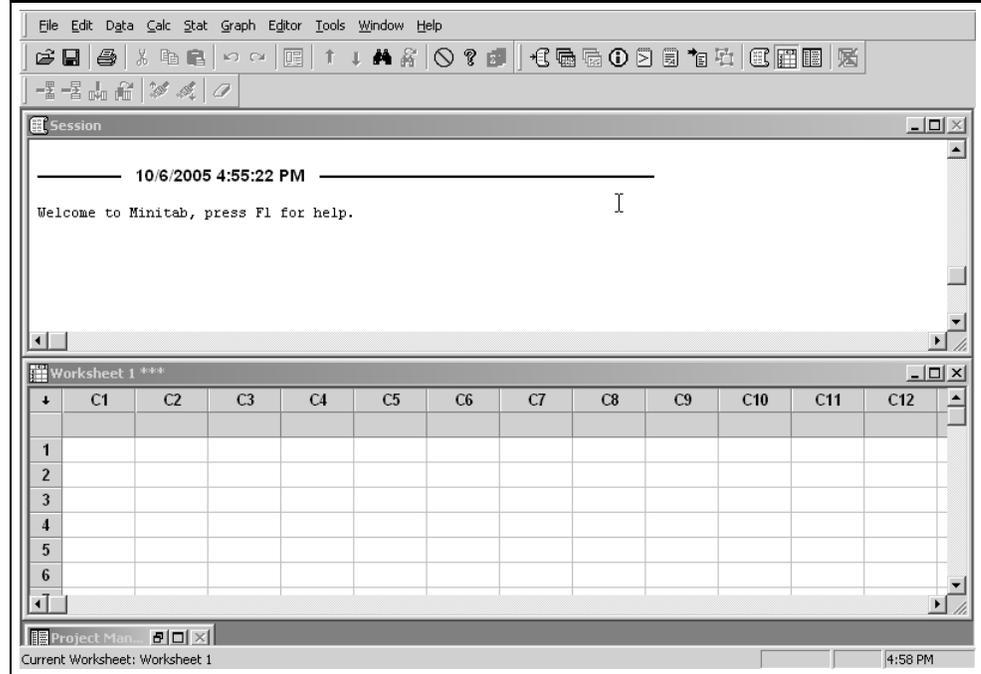
In this manual, you will learn to use MINITAB to analyze data associated with the problems outlined in the text. You will learn some basic MINITAB concepts in this chapter. Chapters 2 through 10 will address problems in the like numbered chapters in the text.

1.2 Getting Started

Before starting MINITAB, you should make yourself a directory in which to store your work. I've called my directory `bhs.dir`. You should then copy the files that accompany this manual into your directory.

To start MINITAB you can either double click the MINITAB shortcut icon on the desktop  or select the sequence **Start**⇒**Programs**⇒**MINITAB 14**. The notation **Start**⇒**Programs**⇒**MINITAB 14** means that you are to first select the Start menu, then Programs, and finally MINITAB 14. We will use this notation throughout this manual.

¹Henceforth referred to as “the text”.

Figure 1.1: Screen appearance when MINITAB is started.

Unless some of the default settings for MINITAB have been altered, the windows shown in Figure 1.1 will appear. These are but two of several windows utilized by MINITAB. You can move between these two windows by simply clicking on the desired window. Some of the commands available to you will depend on which window you are currently addressing.

The upper panel in Figure 1.1 is termed the Session Window and we will have more to say about it later. Before doing so, let's examine the lower panel which is called the Data Window.

1.3 Inputting Data

Logically, before you can analyze data, you must input it to the analysis system. This is done by either typing the data into a worksheet in the Data Window or loading a preexisting worksheet into the Data Window. Worksheets contain the data and may be viewed in the Data Window or stored in a directory. As you can see, the worksheet is similar to a spreadsheet and has many, though not all, of the characteristics of a spreadsheet. The rows of the worksheet are numbered 1, 2, 3, etc. while the columns are designated C1, C2, C3 etc. The intersection of a row and column is termed a cell. We will use the notation R_iC_j to designate the cell defined by the intersection of the i th row and j th column. Thus, R_2C_3

Figure 1.2: Screen appearance when data from Table 2.6 are entered.

	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12
1	x											
2		3										
3		3										
4		4										
5		4										
6		4										
7		4										
8		4										
9		5										
10		5										
11		5										
12												
13												
14												
15												
16												
17												
18												
19												

is the cell defined by the intersection of the second row and third column.

Let's use the data in Table 2.6 on page 34 of the text to demonstrate data entry. To begin, click cell R1C1. Notice that a dark frame appears around this cell. The frame indicates the cell in which data will be entered when you type. If you click on some other cell, the frame will appear around that cell. Now type the number 3,² which is the first value in Table 2.6, in cell R1C1. Because variables are entered columnwise, we will type the second entry, also a 3, in R2C1. To do this, you can click this cell and type the entry. You can also use the four keyboard direction keys to move from cell to cell. Alternatively, we can use an automatic feature that allows the frame to advance to the next cell for data entry each time we press **Enter**. To use this feature, notice the small arrow  at the left of the C1 designation. This arrow, called the **data-direction arrow**, will move the frame one cell downward when it points downward or one cell to the right when it points to the right. To change the direction in which the arrow points, simply click it. Now enter the remaining values from the first column of Table 2.6.

It is a good idea to name the variables in your worksheet. We will name our variable x because that was the only designation used in table 2.6 but we would usually want to use a more informative name such as "blood pressure" or "age." To enter the variable name, click on the empty cell just below the C1 designation. Type x in this cell. Your Data Window should now appear as shown in Figure 1.2.

It is wise to save ones work at frequent intervals so let's save our data at this point. To do this, we choose **File**⇒**Save Current Worksheet**.³ Let's name this worksheet **table 2.6 data**. Be sure you are saving into the directory you created earlier for this purpose (bhs_dir in my case). This worksheet is now available to us at any time we choose to load it into the Data Window.

To see that this is true, close the worksheet by selecting **File**⇒**Close Worksheet**. The **table 2.6 data** worksheet is replaced by a new blank worksheet. We can load **table 2.6 data** back into the Data Window by selecting **File**⇒

²We will use italics to indicate things you are to type or choose in windows.

³Do not use the **save** button  for this purpose.

Open Worksheet..., and choosing **table 2.6 data.MTW** from the file list. Note that MINITAB has appended the suffix MTW to the file. Now click **Open** and **table 2.6 data.MTW** reappears in the Data Window.

1.4 Performing Analyses

Now that we have entered our data, we can use MINITAB to perform analyses on these data. We will provide a simple example in this section but will give numerous examples in the chapters that follow.

MINITAB uses two methods for performing analyses and other tasks. The first is menu driven while the second makes use of Session commands. We will cover only the use of menus for analyses in this manual. To learn the Session command method, consult the MINITAB Handbook (see footnote 1 on page i).

Before beginning the analysis, be sure your data are displayed. If they are not displayed,⁴ select **File⇒Open Worksheet...**, navigate to the directory in which you saved **table 2.6 data** and open it.

Let's begin by calculating a few descriptive statistics on the variable we've labeled x . To do this, select **Stat⇒Basic Statistics⇒Display Descriptive Statistics...** A window opens. You can determine which statistics will be displayed by selecting **Statistics...** If you select this option, a window opens listing all the statistics that may be displayed with this command. Those that are checked will be displayed. Close this window by selecting **OK** or **Cancel**.

Though you only have one variable (x) in this data set, you will often have more. For this reason, you must tell MINITAB which variable(s) you wish to enter into the analysis. You can do this by double-clicking on x in the left hand window or typing x into the **Variables:** window. You can also enter the column number (C1 in this case). It is important to remember that names containing a space must be enclosed in single quotes. MINITAB will put these quotes around the variable name for you if you use the double-click method for entering the variable name into the **Variables:** window. Your screen appears as in Figure 1.3. Now select **OK** to perform the analysis.

Because we didn't change any of the statistics in the **Statistics...** window, results are displayed in the Session Window as shown in Figure 1.4. Displayed there are (a) the variable name (x), (b) the number of variable observations (11), (c) the number of missing variable observations (0), (d) the mean of x (4.000), (e) the standard error of the mean (0.234), (f) the standard deviation of x (0.775), (g) the minimum value of x (3.000), (h) the first quartile (3.000), (i) the median (4.000), (j) the third quartile (5.000) and, (k) the maximum value of x (5.000). You will learn about each of these in your text. If other statistics were checked in the **Statistics...** window you will have these statistics displayed.

⁴Your data will not be displayed if you've ended the MINITAB session or closed the worksheet.

Figure 1.3: Screen appearance for display of basic statistics.

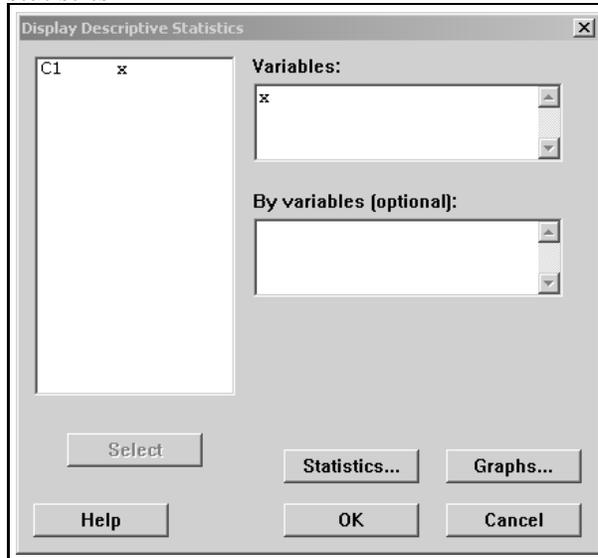


Figure 1.4: Display of basic statistics from x variable in text Table 2.6.

Results for: table 2.6 data.MTW										
Descriptive Statistics: x										
Variable	N	N*	Mean	SE Mean	StDev	Minimum	Q1	Median	Q3	Maximum
x	11	0	4.000	0.234	0.775	3.000	3.000	4.000	5.000	5.000

Figure 1.5: Edited output of basic statistics from x variable in text Table 2.6.

10/31/2005 10:09:30 AM										
Student Name: Cliff Blair										
Date: 10/30/05										
Assignment #: 3										
Results for: table 2.6 data.MTW										
Descriptive Statistics: x										
Variable	N	N*	Mean	SE Mean	StDev	Minimum	Q1	Median	Q3	Maximum
x	11	0	4.000	0.234	0.775	3.000	3.000	4.000	5.000	5.000

1.5 Editing and Printing Your Output

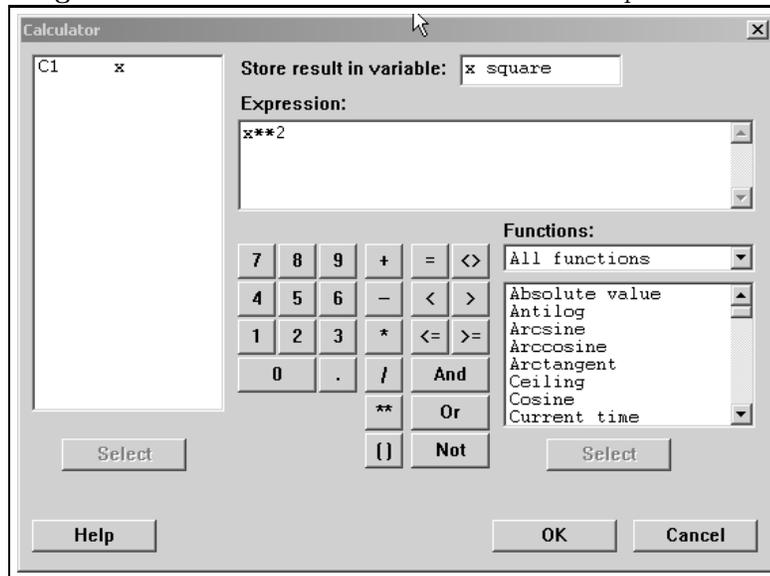
Once you have produced output in the Session Window, you will likely want to edit and/or print it. For example, you may wish to add your name, date and a class assignment number. You may then want to print out the result to turn in to your instructor.

To begin, be sure your editor is set to allow editing of output. To do this, click anywhere in the Session Window. (You can easily move from the Data Window to the Session Window and visa versa by simply clicking in the desired window.) Then select **E**ditor. If, in the list of options, **O**utput **E**ditable has a check mark by it, the output is already amenable for editing and you need do nothing more. If not, click by this option to place a check mark there. You can now carry out routine editing tasks such as typing, highlighting, cutting, copying, pasting etc. You can also cut and copy portions of the output to Microsoft Word[®] or other word processors if you so desire. If you wish to print portions of the output, simply highlight the part you want to print and use the usual printing commands. Figure 1.5 shows our edited output.

1.6 Using the Calculator

From time to time you may wish to use the variables in your data set to create new variables. This can be done through use of the calculator provided by MINITAB. To demonstrate its use, let's use it to create the second column in Table 2.6 which, as you can see, is the square of the first column.

To begin, be sure that **table 2.6 data.MTW** is the current worksheet. Then select **C**alc⇒**C**alculator.... The Calculator window opens. In the **Store result in variable:** window type *x square*. This will be the name of the new variable which will be placed in the next available column which is C2 in this case. In the **Expression:** window type x^{**2} . The symbol ****** means “raise to the power of”. Had we typed x^{**3} the new variable would be the cube of x. Thus, what we have typed in this case is equivalent to “x square= x^{**2} .” Figure 1.6 shows the calculator screen used for the calculation. Now click **OK**

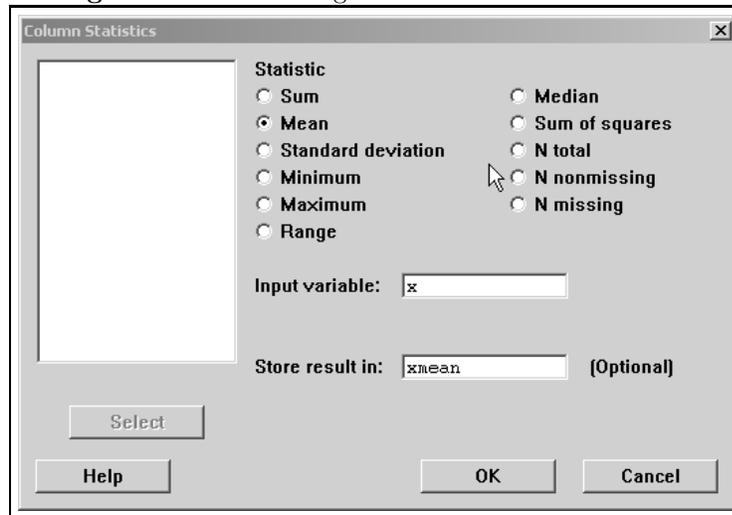
Figure 1.6: Calculator screen used to create the square of x .

to perform the calculation. Notice that the worksheet now contains the variable x square in C2. We can save our new data set by selecting **File**⇒**Save Current Worksheet**.

As a second example, let's generate the values in the column headed $x - \bar{x}$ in Table 2.6. These values are termed "deviation scores" and will be explained in text Chapter 2. To calculate these deviation scores, we must subtract the average (i.e. mean) of x from each value of x . We previously found that the average of x was 4.000 so we could write an expression to subtract this value from each of the x scores. But let's do this in a slightly more sophisticated way.

We will first calculate the mean of x by selecting **Calc**⇒**Column Statistics...** A window opens. Choose the button next to **mean** from the list of statistics. In the **Input variable:** window type x . This tells MINITAB that we are calculating the statistic on the x variable. In the **Store result in:** window type $xmean$. This tells MINITAB to name the mean value $xmean$. Your screen should appear as shown in Figure 1.7. Now click **OK**. Notice that the mean of x is displayed in the Session window. Let's save our work by selecting **File**⇒**Save Current Worksheet**. We can now use $xmean$ in the calculator to compute deviation scores.

Select **Calc**⇒**Calculator...** Notice that in addition to C1 and C2, the calculator window now shows K1 with the associated name of $xmean$. MINITAB refers to columns of data as C1, C2, etc. but refers to a single value (a constant) such as the mean of x as K1, K2, etc. Let's name the new variable *deviation* by typing this in the **Store result in variable:** window. Now type $x-xmean$ in the **Expression** window and click **OK**. C3 now contains the deviation scores

Figure 1.7: Calculating the mean of x .

as shown in Table 2.6. Save the worksheet.

1.7 Graphs

MINITAB can be used to create a wide variety of graph types ranging from the simple to the complex. You will employ some of these capabilities in Chapter 2. For the moment, we will provide only a simple example of one graph type by constructing a histogram of the x variable in Table 2.6.

To begin, be sure that **table 2.6 data.MTW** is the current worksheet. Now select **Graph**⇒**Histogram...**. Four windows appear. Click on the first one labeled **simple** to be sure it is darkened then select **OK**. A window labeled **Histogram-Simple** opens. Either double-click on x or type it into the **Graph variables:** window and select **OK**. A histogram of x appears as shown in Figure 1.8. Close the window without saving changes.

Figure 1.8: Histogram of x.

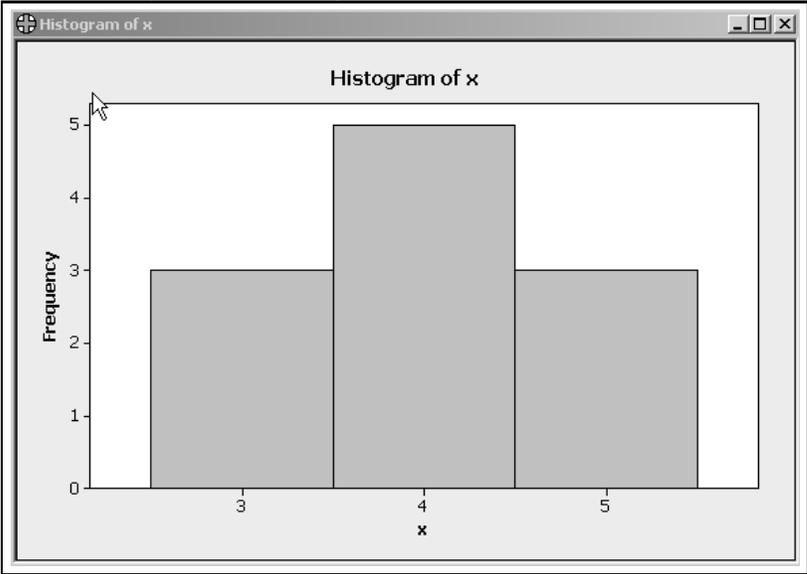


Figure 1.9: Reconstruction of Table 2.6.

↓	C1	C2	C3	C4	C5	C6
	x	x square	deviation	absolute deviation	squared deviation	
1	3	9	-1	1	1	
2	3	9	-1	1	1	
3	3	9	-1	1	1	
4	4	16	0	0	0	
5	4	16	0	0	0	
6	4	16	0	0	0	
7	4	16	0	0	0	
8	4	16	0	0	0	
9	5	25	1	1	1	
10	5	25	1	1	1	
11	5	25	1	1	1	

Exercises

- 1.1 Use the **Calculator...** to construct the remaining variables in Table 2.6. The result should appear as in Figure 1.9. (Hint: double-clicking on a function in the **Functions:** window places it in the **Expression:** window.)
- 1.2 Use **Display Descriptive Statistics...** to find the means of all the variables in Table 2.6. Don't generate any other statistics. Use **Display Descriptive Statistics...** only once. (Hint: You can place multiple variables in the **Variables** window by double-clicking on each.)
- 1.3 Label the output from Exercise 1.2 with your name and the current date. Print out the result.
- 1.4 Enter the data from text Table 7.1 into a new worksheet. Save the worksheet with the name **table 7.1 data**.
- 1.5 Find the means of the three diet groups in Table 7.1. Label them mean one, mean two and mean three. Save them in the **table 7.1 data** worksheet. (Don't forget that names containing spaces must be enclosed in single quotes.)

Chapter 2

Descriptive Methods

2.1 Introduction

In this chapter you will learn to use MINITAB to construct the graphs and calculate most of the statistics presented in text Chapter 2. You will also learn that in a few instances MINITAB and the text do not agree on the calculated values of certain descriptive statistics because different definitions of these statistics are used by the software and text. You will also be afforded an opportunity to practice data entry and data transformations.

2.4 Distributions (page 15)

Task 2.1

Use the data in Table 2.1 (**table 2.1 data.MTW**) to construct frequency, relative frequency, cumulative frequency, and cumulative relative frequency distributions as shown in Table 2.2.

Solution

Load the worksheet containing the Table 2.1 data by selecting **File⇒Open Worksheet...** and opening file **table 2.1 data.MTW**. Note first that the Pain Level variable is expressed in text rather than numeric form. Because the cumulative frequency and cumulative relative frequency distributions require that the Pain Level variable be ordered, we must inform MINITAB as to the low to high ordering of this variable before we can form these distributions. To do this, right-click on the Pain Level variable in the worksheet. A series of choices is displayed. Select **Column⇒Value Order**. A window opens. Select the option button for **User specified Order: New Order** is automatically highlighted in the **Choose an order:** window. The text values appear in a window labeled **Define an order (one value per line):**. Place these values in

Figure 2.1: Screen appearance for ordering pain level variable.

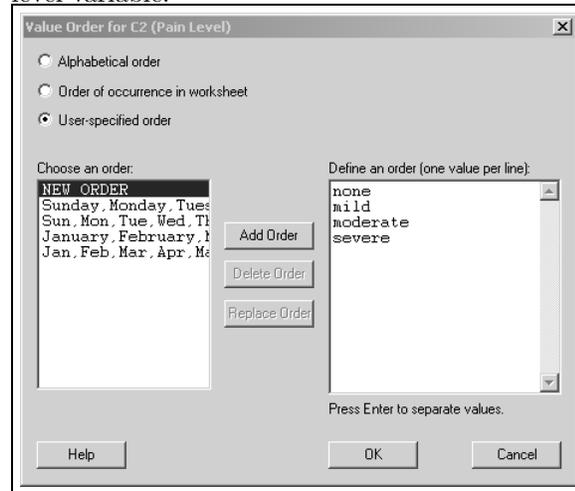


Figure 2.2: Frequency, cumulative frequency, relative frequency and cumulative relative frequency of perceived pain scores.

Pain Level	Count	CumCnt	Percent	CumPct
none	31	31	51.67	51.67
mild	17	48	28.33	80.00
moderate	8	56	13.33	93.33
severe	4	60	6.67	100.00
N=	60			

order from lowest to highest by editing the window entries. The screen should appear as shown in Figure 2.1. Now click **Add Order** and then **OK**.

Now MINITAB understands, at least in terms of order of magnitude, the meaning of none, mild, moderate and severe and will use this order when creating graphs and tables involving this variable. Save the worksheet.

To create the desired distributions select **Stat**⇒**Tables**⇒**Tally Individual Variables....** Now double-click on the Pain Level variable in the **Tally Individual Variables** window to place this variable in the **Variables:** window. Now check all four boxes in the **Display** list and click the **OK** button. The output is displayed in Figure 2.2.

Several points regarding this output should be noted. First, what your text refers to as Frequency, Relative Frequency, Cumulative Frequency and Cumulative Relative Frequency are labeled Count, Cumulative Count, Percent and Cumulative Percent by MINITAB. This is not unusual as many terms and styles of presentation are commonly used by both textbook authors and software developers. This may present some difficulties, but they tend to diminish as you

Figure 2.3: Simple frequency distribution of BP scores as shown in Table 2.3.

Tally for Discrete Variables: BP					
BP	Count				
86	1	109	1	126	7
90	1	110	3	127	3
92	2	111	3	128	3
93	1	114	2	129	3
95	3	115	6	130	5
96	2	116	3	131	3
97	2	117	1	132	4
98	2	118	2	133	8
99	1	119	2	134	5
100	4	120	3	135	8
102	1	121	1	136	3
103	1	122	3	137	11
105	2	123	2	138	3
106	1	124	4	139	6
107	2	125	4	140	4
			N=	144	

gain more experience with different software packages. Second, with the exception that MINITAB uses percents where the text uses proportions and the number of decimal places reported, the distributions provided by MINITAB are the same as those in text Table 2.2. Third, MINITAB orders the distributions in descending order while the text uses an ascending order. The style used by MINITAB is the more common. The text authors prefer the ascending ordering but either is correct and provides the same information.

Task 2.2

Use the data in worksheet **table 2.3 data.MTW** to construct the simple frequency distribution depicted in the text Table 2.3.

Solution

Because the blood pressure data in worksheet **table 2.3 data.MTW** are numeric as opposed to text, we need not go through the process of defining the variable order as was necessary with Task 2.1. Thus, we can construct the simple frequency distribution without any preliminaries.

After loading **table 2.3 data.MTW** we select **Stat**⇒**Tables**⇒**Tally Individual Variables...**, then double-click on **BP** to place it in the **Variables:** window. Place a check by **Counts** in the **Display** window and remove any checks by any other options in this list. Now select **OK** to produce the frequency distribution. The resultant frequency distribution is shown in Figure 2.3. Note that I have edited the output into three columns to save page space.

Task 2.3

Use the data in worksheet **table 2.3 data.MTW** to construct the grouped distributions depicted in the text Table 2.4.

Figure 2.4: Partial conversion table used to relate BP values to group intervals.

↓	C1	C2	C3-T	C4	C5	C6	C7	C8	C9	C10	C11	C12
	BP	x	y									
1	140	144	140-144									
2	128	143	140-144									
3	137	142	140-144									
4	143	141	140-144									
5	135	140	140-144									
6	129	139	135-139									
7	137	138	135-139									
8	143	137	135-139									
9	120	136	135-139									
10	119	135	135-139									
11	135	134	130-134									
12	118	133	130-134									
13	138	132	130-134									
14	109	131	130-134									
15	100	130	130-134									
16	129	129	125-129									
17	135	128	125-129									
18	100	127	125-129									
19	120	126	125-129									

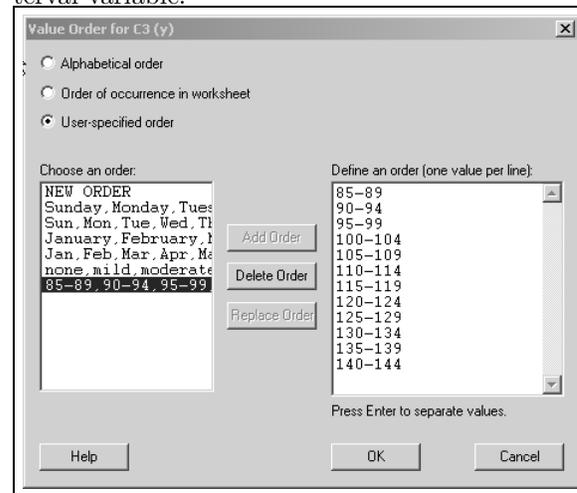
Solution

Unfortunately, MINITAB does not provide internal methods for constructing grouped distributions. So, in order to complete this task, we must first code the numeric values in **table 2.3 data.MTW** to form the discrete intervals shown in text Table 2.4. Because these categories are text based (even though they contain numbers) we will follow the same procedure used for Task 2.1 to order the text categories. We will then be able to construct the desired distributions as with previous tasks.

To begin the coding, we construct a conversion table that relates all BP values to specific category intervals. We will use C2 and C3 in our worksheet for this purpose. For convenience, let's label C2 as x and C3 as y . We now list all possible values of BP in x and the categories to which they relate in y . The (partial) result is shown in Figure 2.4.

We are now ready to use the conversion table to create the Interval variable. To accomplish this, select **Data⇒Code⇒Use Conversion Table....** The **Code-Use Conversion Table** window opens. Enter *BP* in the **Input column:** window. This tells MINITAB that it is the BP values that are to be converted into intervals. Now type *Interval* in the **Output column:** window. This names the new variable "Interval." Now type x in the **Column of Original Values:** window and y in the **Column of New Values:** window. This tells MINITAB the values of BP that correspond to the specified intervals. Now click **OK**. A new variable, named Interval, consisting of the specified intervals is created and placed in column C4 of the worksheet. Notice that each BP value is related to its corresponding interval.

Figure 2.5: Screen appearance for ordering Interval variable.



We've now reached the point where we began Task 2.1. That is, we must order the categories and construct the desired distributions. To do this, right-click on the Interval variable in the worksheet. A series of choices is displayed. Select **Column**⇒**Value Order**. A window opens. Select the option for **User specified Order**:. **New Order** is automatically highlighted in the **Choose an order**: window. The text values appear in a window labeled **Define an order (one value per line)**:. Place these values in order from lowest to highest by editing the window entries as shown in Figure 2.5. Now click **Add Order** and **OK**. MINITAB now understands that 85-89 is a lesser category than 90-94 which is less than 95-99 etc. Save the worksheet.

We are, now ready to construct the desired distributions. To create the distributions, select **Stat**⇒**Tables**⇒**Tally Individual Variables...**. Now double-click on the Interval variable in the **Tally Individual Variables** window to place this variable in the **Variables**: window. Check all four boxes in the **Display** list and click the **OK** button. The output is displayed in Figure 2.6. Save the worksheet, we will use it again later.

2.5 Graphs (page 19)

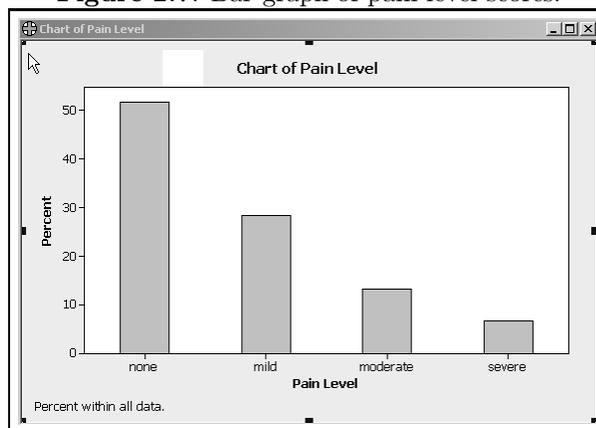
Task 2.4

Use the data in Table 2.1 (**table 2.1 data.MTW**) to construct the relative frequency bar graph depicted in text Figure 2.1.

Figure 2.6: Grouped frequency distributions for BP variable in 12 categories.

Interval	Count	CumCnt	Percent	CumPct
85-89	1	1	0.69	0.69
90-94	4	5	2.78	3.47
95-99	10	15	6.94	10.42
100-104	6	21	4.17	14.58
105-109	6	27	4.17	18.75
110-114	8	35	5.56	24.31
115-119	14	49	9.72	34.03
120-124	13	62	9.03	43.06
125-129	20	82	13.89	56.94
130-134	25	107	17.36	74.31
135-139	31	138	21.53	95.83
140-144	6	144	4.17	100.00
N=	144			

Figure 2.7: Bar graph of pain level scores.

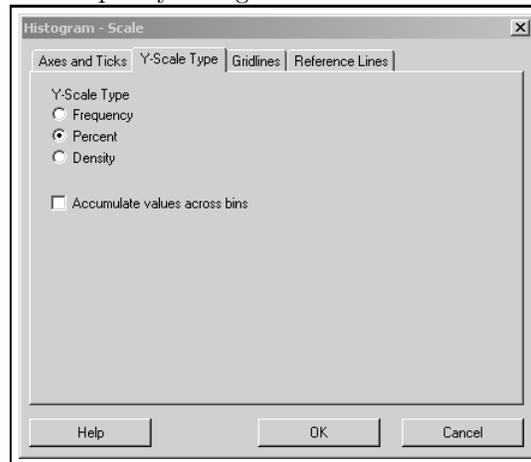


Solution

Open table 2.1 data.MTW as the active worksheet. Select **Graph**⇒**Bar Chart...** If the bar chart in the window labeled “Simple” is not selected, click on that window to select it and click **OK**. The **Bar Chart-Counts of unique values, Simple** window opens. Double-click on Pain Level to place it in the **Categorical variables:** window. Because the task specified that a *relative* frequency bar graph is to be constructed, we must inform MINITAB that the bars are to be shown in terms of proportions (or percents in this case). To do this, select **Bar Chart Options...** and place a check in the **Show Y as percent** box and click **OK**. You are now back at the **Bar Chart-Counts of unique values, Simple** window. Click **OK**. The graph appears as shown in Figure 2.7. Print this graph by selecting **File**⇒**Print Graph**.

MINITAB allows graphs to be edited if so desired. As one example, suppose we wish the title of the graph to be “Bar Chart of Pain Level” rather than

Figure 2.8: Informing MINITAB that a relative frequency histogram is to be constructed.



simply “Chart of Pain Level.” Just double-click on Chart of Pain Level at the top of the graph. The **Edit Title** window opens. Be sure the **Font** tab is selected. Now type in the desired title in the **Text:** window and click **OK**. The graph appears with the new title. You should experiment a bit to see what other alterations you can make to the graph.

Notice that the pain levels are ordered. How do you think MINITAB knew the order of this variable? (Hint: go back and look at Task 2.1.)

Task 2.5

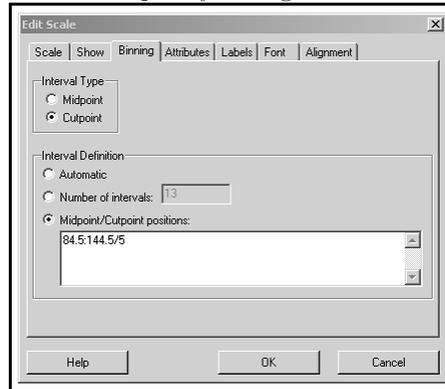
Use the data in Table 2.3 (**table 2.3 data.MTW**) to construct the relative frequency histogram depicted in text Figure 2.2.

Solution

Be sure **table 2.3 data.MTW** is the active worksheet. Then select **Graph**⇒**Histogram...** The **Histograms** window opens. Select the Simple histogram by clicking on the depiction in the **Simple** window. Click on **OK**. In the **Histogram-Simple** window, double-click on BP to move this variable into the **Graph variables:** window. If we wished to construct a (default) *frequency* histogram we would click **OK** at this point and continue, but because we wish to form a *relative frequency* histogram we must inform MINITAB that this is the case. To do so, select **Scale...** in the **Histogram-Simple** window. Then choose the **Y-Scale Type** tab and then click the **Percent** option. Your screen should appear as in Figure 2.8. Now click **OK** then click **OK** again. A relative frequency¹ histogram appears but it is not grouped as in text Figure 2.2.

¹Once again, MINITAB expresses the relative frequency as a percent rather than as a proportion as is done in the text. This will generally be the case so we will not continue to

Figure 2.9: Specifying intervals for a relative frequency histogram.



To form the appropriate grouping, double-click anywhere along the horizontal (i.e. x) axis of the histogram. The **Edit Scale** window opens. Choose the **Scale** tab and darken the **Position of ticks:** button. Remove any entries in the associated window and enter $84.5:144.5/5$. Click the **Binning** tab. Choose **Cutpoint** in the **Interval Type** list. This tells MINITAB to label the interval rather than the midpoint for each bar. Now choose **Midpoint/Cutpoint positions**. In the accompanying window you could type 84.5 89.5 94.5 etc. to label the graph but MINITAB provides a shortcut that will be easier in this case. Type $84.5:144.5/5$ in the window. This provides MINITAB with the range of values for the x axis (84.5 to 144.5) as well as the interval size (5). The screen appears as in Figure 2.9. Click **OK** and the histogram depicted in Figure 2.10 appears.

Task 2.6

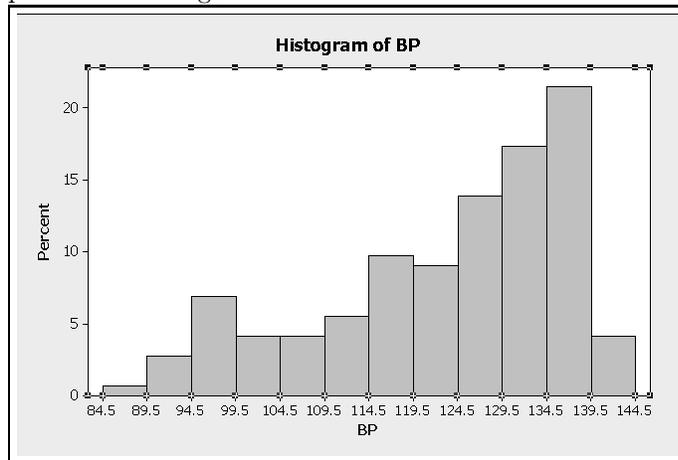
Use the data in Table 2.3 (**table 2.3 data.MTW**) to construct the relative frequency polygon depicted in text Figure 2.4.

Solution

This task can be approached in essentially the same manner as was used with Task 2.5 with the major difference being that we will ask MINITAB to provide connected dots rather than bars in the depiction. To this end, select **Graph**⇒**Histogram...** Choose the **Simple** histogram and click **OK**. In the **Histogram-Simple** window double-click on the BP variable to place it in the **Graph-variables:** window if it isn't there already. Now select **Scale...** and in the **Histogram-Scale** window choose the **Y-Scale Type** tab. Under this tab darken the **Percent** button so that MINITAB will plot percents rather than frequencies. Click **OK**. You are now back in the **Histogram-Simple** window.

remind you of this fact.

Figure 2.10: MINITAB rendering of the histogram depicted in text Figure 2.2.



From this window select **Data View...** Under the **Data Display** tab place a check by **Symbols** and uncheck any other checks. Now select the **Smoother** tab. Choose the **Lowess** button in the **Smoother** list and set the Degree of smoothing to 0 (zero). This causes MINITAB to connect the dots with line segments. Click **OK** then click **OK** again. A relative frequency polygon appears. But it does not reflect the twelve groupings.

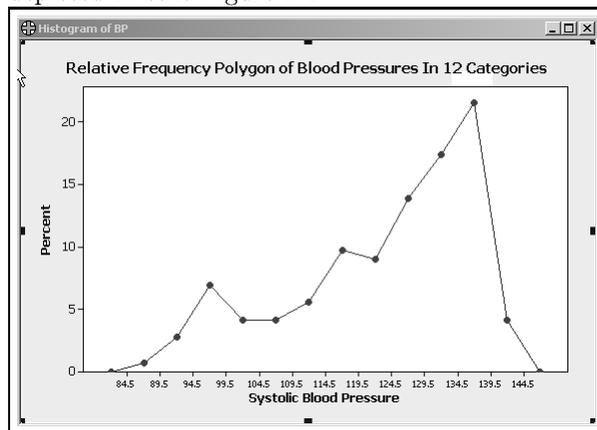
Double-click anywhere along the x axis. The **Edit Scale** window opens. Under the **Scale** tab, choose the **Position of ticks:** button. In the accompanying window type $84.5:144.5/5$. This tells MINITAB to place the first tick mark at 84.5, the last at 144.5 and every five units in between. Now choose the **Binning** tab. Be sure that **Midpoint** is selected under **Interval Type** so that the polygon dots will be placed at the midpoints of the intervals. Now choose the **Midpoint/Cutpoint positions:** button. We now enter the midpoints of our desired intervals as $82:147/5$ and click **OK**. The desired polygon appears.

Let's edit the graph just a bit. Double-click on the **Histogram of BP** title at the top of the graph. In the **Text:** box type *Relative Frequency Polygon of Blood Pressures in 12 Categories* and click **OK**. The graph appears with the new title. Now double-click on the BP designation at the bottom of the graph. Type *Systolic Blood Pressure* in the **Text:** box and click **OK**. The relative frequency polygon appears as in Figure 2.11.

Task 2.7

Use the data in Table 2.3 (**table 2.3 data.MTW**) to construct the cumulative relative frequency polygon depicted in text Figure 2.6.

Figure 2.11: MINITAB rendering of the polygon depicted in text Figure 2.4.



Solution

In formulating a solution for Task 2.3, we used a conversion table to assign subjects to specified intervals. We will use this technique again here to group subjects into the 12 intervals depicted in Figure 2.6. More specifically, we will assign the value 84.5 to all subjects with blood pressures of 85, 86, 87, 88 or 89, a value of 89.5 to all subjects with pressures of 90, 91, 92, 93 or 94 and so on with the five highest scale values being assigned 139.5. You will see the underlying rationale shortly.

The conversion table employed for the solution of Task 2.3 consisted of two variables which we named x and y . The first of these listed all possible blood pressure values while the second indicated the interval to which each of these should be assigned. We will use the x variable again here but we will replace the y variable with a new set of intervals.

Be sure that **table 2.3 data.MTW** is the current worksheet. If you saved it at the completion of Task 2.3, it should appear as in Figure 2.4 except that the variable named Interval has been added in C4-T. Let's clean up this worksheet by deleting the y and Interval variables. To do this, click and hold down the mouse button on C3-T and drag across to C4-T. This will select (darken) these two columns. Now select **Edit⇒Delete Cells**. The C3 and C4 columns are now empty.

To construct the new conversion table begin by labeling the C3 column y . Now enter the interval values in y as shown in Figure 2.12. This table indicates that the blood pressures listed in x will be converted to the values specified in y . Now let's perform the conversion.

Select **Data⇒Code⇒Use Conversion Table...** The **Code-Use Conversion Table** window opens. Double-click on BP to place it in the **Input column:** window. This tells MINITAB that it is the blood pressure values that

Figure 2.12: Specifying intervals for a cumulative relative frequency polygon.

↓	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10
	BP	x	y							
1	140	144	139.5							
2	128	143	139.5							
3	137	142	139.5							
4	143	141	139.5							
5	135	140	139.5							
6	129	139	134.5							
7	137	138	134.5							
8	143	137	134.5							
9	120	136	134.5							
10	119	135	134.5							
11	135	134	129.5							
12	118	133	129.5							
13	138	132	129.5							
14	109	131	129.5							
15	100	130	129.5							
16	129	129	124.5							
17	135	128	124.5							
18	100	127	124.5							

are being coded. Now type *Interval* in the **Output column:** window. This tells MINITAB to name the new variable Interval. Place an *x* in the **Column of Original Values:** window and a *y* in the **Column of New Values:** window. This tells MINITAB which *x* values correspond to which *y* values. Click **OK** and the Interval variable appears in C4.

Our strategy now involves calculating the cumulative relative frequency for the Interval variable. To this end, select **Stat**⇒**Tables**⇒**Tally Individual Variables...** In the **Tally Individual Variables** window double-click on the Interval variable to place it in the **Variables** window. In the **Display** list put a check by **Cumulative percents** and uncheck any others that have a check by them. Click **OK**. The resultant distribution appears in the Session Window and is shown in Figure 2.13.

We must now paste this result into the worksheet. To accomplish this, click and drag the cursor in the Session Window so as to select (darken) all of the output depicted in Figure 2.13. Now select **Edit**⇒**Copy** to copy the selected portion of the Session Window to the clipboard. Now click and hold down the mouse button on C5 in the worksheet and drag the cursor across to C6 so that both the C5 and C6 columns are selected (darkened). Select **Edit**⇒**Paste Cells**. Be sure that the **Use spaces as delimiters** button is selected and click **OK**. The cumulative relative frequency distribution now appears in columns C5 and C6 as shown in Figure 2.14.²

Let's now round off the CumPct values to the nearest integer. To do this, select **Calc**⇒**Calculator...** Double-click CumPct to place it in the **Store result in variable:** window. Place the cursor in the **Expression:** window and double-click on **round** in the **Functions:** window. (You will have to scroll down to find this function.) This enters "ROUND(number, num digits)" in the **Functions:**

²You can paste this distribution in C1 and C2 of a new worksheet if you prefer.

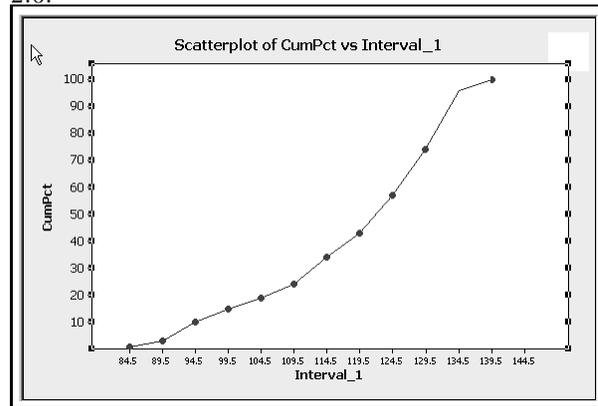
Figure 2.13: Cumulative relative frequency distribution for blood pressure measures in 12 intervals.

Interval	CumPct
84.5	0.69
89.5	3.47
94.5	10.42
99.5	14.58
104.5	18.75
109.5	24.31
114.5	34.03
119.5	43.06
124.5	56.94
129.5	74.31
134.5	95.83
139.5	100.00

Figure 2.14: Worksheet after pasting the cumulative relative frequency distribution into columns C5 and C6.

↓	C1	C2	C3	C4	C5	C6	C7
	BP	x	y	Interval	Interval_1	CumPct	
1	140	144	139.5	139.5	84.5	0.69	
2	128	143	139.5	124.5	89.5	3.47	
3	137	142	139.5	134.5	94.5	10.42	
4	143	141	139.5	139.5	99.5	14.58	
5	135	140	139.5	134.5	104.5	18.75	
6	129	139	134.5	124.5	109.5	24.31	
7	137	138	134.5	134.5	114.5	34.03	
8	143	137	134.5	139.5	119.5	43.06	
9	120	136	134.5	119.5	124.5	56.94	
10	119	135	134.5	114.5	129.5	74.31	
11	135	134	129.5	134.5	134.5	95.83	
12	118	133	129.5	114.5	139.5	100.00	
13	138	132	129.5	134.5			
14	109	131	129.5	104.5			

Figure 2.15: MINITAB rendering of the cumulative relative frequency polygon depicted in text Figure 2.6.



window. Edit this expression so it appears as `ROUND('CumPct',0)`. This tells MINITAB to round off the CumPct variable to the nearest whole number. Click **OK** and the CumPct values are rounded to the nearest whole number. We are now ready (at last) to create the graph.

Choose **Graph**⇒**Scatter Plot...** In the **Scatterplots** window choose the **With Connect Line** figure and click **OK**. Put CumPct in the **Y variables** window and Interval_1 in the **X variables** window and click **OK**. The cumulative distribution is presented on the screen.

In order to have the x axis scale conform with the scale in Figure 2.6, double-click on the x axis of the graph. In the **Edit Scale** window choose the **Position of ticks:** button. In the accompanying window type `84.5:144.5/5` which tells MINITAB to label the x axis from 84.5 to 144.5 in five unit increments. Click **OK**. The graph appears with the x axis properly scaled.

Finally, let's bring the y scale down so that zero will be closer to the x axis as in Figure 2.6. Double-click on the y axis. In the **Edit Scale** window be sure that the **Scale** tab is selected. Under **Scale Range** uncheck **Minimum:** and type `.1` in the accompanying window. This puts the bottom of the y scale near zero. Now click **OK** and the graph depicted in Figure 2.15 appears. Feel free to edit the graph titles to make them more meaningful.

Task 2.8

Construct a stem-and-leaf-display for the blood pressure data in text Table 2.3.

Solution

Be sure that **table 2.3 data.MTW** is the current worksheet. Select **Graph**⇒**Stem-and-leaf...** Double-click on BP to place it in the **Graph variables...** window and click **OK**. A stem-and-leaf-display appears in the Session Window as shown

Figure 2.16: MINITAB rendering of a stem-and-leaf display for the BP variable in text Table 2.3.

```

1  8  6
5  9  0223
15 9  5556677889
21 10 000023
27 10 556779
35 11 00011144
49 11 5555566678899
62 12 0001222334444
(20) 12 5555666666777888999
62 13 00000111222233333344444
37 13 555555566677777777888999999
6  14 000033

```

Figure 2.17: Edited MINITAB rendering of a stem-and-leaf display for the BP variable in text Table 2.3.

```

8  6
9  02235556677889
10 000023556779
11 000111445555566678899
12 0001222334444555566666777888999
13 0000011122223333333344444 555555566677777777888999999
14 000033

```

in Figure 2.16.

Note that MINITAB divides leaves in half so that digits 0 through 4 are placed on one line and 5 through 9 are placed on a second line. Also, the numbers on the far left are what MINITAB refers to as the **depth** of the stem. This column gives the cumulative frequency from the top of the figure down and from the bottom of the figure up. The number in parentheses gives the leaf count for the row containing the median.

Unlike other graphs produced by MINITAB, stem-and-leaf-displays are not provided in special graph windows but are printed in the Session Window. This means that they can be edited with ordinary cut and paste methods. Figure 2.17 shows the display after we've edited it to more closely conform with text Figure 2.7.

2.6 Numerical Methods (page 23)

Task 2.9

In text Example 2.4 on page 27 you are asked to find the median of the numbers 3, 2, 0, 2, 1, 1, 2, and 2. Two different methods are used that produce two different answers, namely, 2.00 and 1.75. It is argued that the better answer is provided by the method that produces the median value of 1.75. What answer is provided by MINITAB?

Solution

We begin by entering the numbers in C1 of a new worksheet. Name the variable x . Select **Stat**⇒**Basic Statistics**⇒**Display Descriptive Statistics...** Double-click on x to place it in the **Variables...** window. Now click on **Statistics...** Place a check by **Median** if there isn't one there already. Uncheck all other statistics and click **OK** then click **OK** again. The answer provided in the Session Window is 2.00.

Task 2.10

Find the mean, median, and standard deviation of the data in the table associated with Example 2.7 on page 30. Is the median reported by MINITAB the same as that reported in the text? Why do you think this is the case?

Solution

We will have to enter this data into a worksheet since none is provided. This process can be somewhat simplified when data are given in the form of a simple frequency distribution as in this case.

Open a new worksheet, if necessary, by selecting **File**⇒**New...** Be sure **MINITAB Worksheet** is highlighted and click **OK**. From the table on text page 30 we notice that there are three 2.4s in the table. To enter these select **Calc**⇒**Make Patterned Data**⇒**Arbitrary Set of Numbers...** In the **Store patterned data in:** window type $C1$ to store the 2.4s in C1. Type 2.4 in the **Arbitrary set of numbers:** window and 3 in the **List each value <blank> times** window. There should be a 1 in the **List the whole sequence <blank> times** window. Thus, we have informed MINITAB to place three 2.4 values in C1. Now click **OK**. Three values of 2.4 appear in C1.

The next value in the table is 2.3 and there are 40 of those so we will place 40 2.3s in C2. To enter these select **Calc**⇒**Make Patterned Data**⇒**Arbitrary Set of Numbers...** In the **Store patterned data in:** window type $C2$. Type 2.3 in the **Arbitrary set of numbers:** window and 40 in the **List each value <blank> times** window. There should still be a 1 in the **List the whole sequence <blank> times** window. Now click **OK**. C2 now contains 40 2.3s. We continue this process by placing 36 2.2s in C3, 21 2.1s in C4, 18 2.0s in C5, 8 1.9s in C6, 14 1.8s in C7 and 5 1.7s in C8. The first 18 rows of your worksheet should appear as in Figure 2.18.

In order to perform analyses, we must form these eight columns into a single column variable. To do this, select **Data**⇒**Stack**⇒**Columns...** Now double-click C1 through C8 in the **Stack Columns** window in order to enter these columns in the **Stack the following columns:** window. Now choose the **Column of current worksheet** button and enter $C1$ in the accompanying window. This set of instructions tells MINITAB to stack the values in C1 through C8 in C1. Now click **OK**. The values in C1 through C8 are now entered into C1. Let's clean up the worksheet by deleting C2 through C8 which we no longer need. To accomplish this, click on the C2 heading in column C2

Figure 2.18: Data from example 2.7.

↓	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10
1	2.4	2.3	2.2	2.1	2	1.9	1.8	1.7		
2	2.4	2.3	2.2	2.1	2	1.9	1.8	1.7		
3	2.4	2.3	2.2	2.1	2	1.9	1.8	1.7		
4		2.3	2.2	2.1	2	1.9	1.8	1.7		
5		2.3	2.2	2.1	2	1.9	1.8	1.7		
6		2.3	2.2	2.1	2	1.9	1.8			
7		2.3	2.2	2.1	2	1.9	1.8			
8		2.3	2.2	2.1	2	1.9	1.8			
9		2.3	2.2	2.1	2		1.8			
10		2.3	2.2	2.1	2		1.8			
11		2.3	2.2	2.1	2		1.8			
12		2.3	2.2	2.1	2		1.8			
13		2.3	2.2	2.1	2		1.8			
14		2.3	2.2	2.1	2		1.8			
15		2.3	2.2	2.1	2					
16		2.3	2.2	2.1	2					
17		2.3	2.2	2.1	2					
18		2.3	2.2	2.1	2					

and drag across to C8. This selects (darkens) columns C2 through C8. Now select **Edit**⇒**Delete Cells** and these columns disappear. The first 18 rows of your worksheet should now appear as in Figure 2.19.

To complete the task of finding the mean, median, and standard deviation of these data, select **Stat**⇒**Basic Statistics**⇒**Display Descriptive Statistics...** Double-click on C1 to place it in the **variables:** window. Click on the **Statistics...** bar to open the list of statistics to be calculated. Be sure that only **Mean**, **Standard deviation** and **Median** have check marks. Click **OK** then click **OK** again. The output is shown in Figure 2.20. As may be seen, the median reported here is 2.2 which contrasts with the value of 2.17 reported in the text as the answer for Example 2.7. Thus, it appears that MINITAB calculates the median as the middle score (or average of the two middle scores) rather than by the more sophisticated method given by Equation 2.5.

Task 2.11

Find the variance of the x variable in Table 2.6 (worksheet **Table 2.6 data.MTW**). Is the result provided by MINITAB the same as that given in the text for example 2.15 on page 36?

Solution

After loading worksheet **Table 2.6 data.MTW**, select **Stat**⇒**Basic Statistics**⇒**Display Descriptive Statistics...** Double-click on x in the **Display Descriptive Statistics** window in order to place it in the **Variables:** window. Now click on the **Statistics...** bar to open the **Descriptive Statistics-Statistics** window. Put a check by **Variance** and uncheck all others. Click **OK** then click **OK** again. The result, .60, which is the answer given in the text, is displayed in the Session Window.

Figure 2.19: Data from example 2.7 in a single column.

	C1	C2	C3	C4	C5	C6	C7	C8
1	2.4							
2	2.4							
3	2.4							
4	2.3							
5	2.3							
6	2.3							
7	2.3							
8	2.3							
9	2.3							
10	2.3							
11	2.3							
12	2.3							
13	2.3							
14	2.3							
15	2.3							
16	2.3							
17	2.3							
18	2.3							

Figure 2.20: Descriptive statistics for data from example 2.7.

Descriptive Statistics: C1			
Variable	Mean	StDev	Median
C1	2.1200	0.1813	2.2000

Task 2.12

Find the twenty-fifth, sixtieth, and seventy-fifth percentiles of the blood pressure data in worksheet **table 2.3 data.MTW**. Do the answers provided by MINITAB match those given in the solution to text Example 2.17 on page 38?

Solution

MINITAB does not provide a general method for finding percentiles.³ We can find the twenty-fifth, and seventy-fifth percentiles, however, because MINITAB gives the 1st and 3rd quartiles which, by definition, are the twenty-fifth and seventy-fifth percentiles.

To obtain these, select **Stat**⇒**Basic Statistics**⇒**Display Descriptive Statistics...** Double-click on BP in the **Display Descriptive Statistics** window in order to place it in the **Variables:** window. Now click on the **Statistics...** bar to open the **Descriptive Statistics-Statistics** window. Put a check by **First quartile** and **Third quartile**. Uncheck all other statistics. Click **OK** then click **OK** again. The result is displayed in the Session Window.

The first and third quartiles (twenty-fifth and seventy-fifth percentiles) are given as 115 and 135 respectively. The values given in the solution to Example 2.17 in the text are 114.67 and 134.63. The discrepancy is likely because, as with the median (fiftieth percentile), MINITAB employs a simpler method for finding percentiles than the one given in the text.

Task 2.13

Use MINITAB to perform the analyses outlined in Example 2.21 on page 43.

Solution

Open a new worksheet. Type the numbers *1, 3, 3, 9* in the first column. Name this variable *x*. Select **Calc**⇒**Standardize...** Enter *x* in the **Input columns(s):** window. Type *C2* in the **Store results in:** window. Be sure that the **Subtract mean and divide by std. dev.** button is selected and click **OK**. The z scores appear in C2. Notice that these z scores agree with those given in the text except that MINITAB gives the z score for nine as 1.44338 while the text gives 1.445. The discrepancy occurs because the text uses 3.46 as the standard deviation while MINITAB carries this calculation out to more decimal places.

Now use **Stat**⇒**Basic Statistics**⇒**Display Descriptive Statistics...** to find the mean and standard deviation of the z scores in C2. As expected, MINITAB gives these values as zero and one respectively.

Task 2.14

Use MINITAB to conduct the analyses alluded to in Example 2.23 on page 44.

³Percentiles can be found by use of macros but that is a topic beyond the scope of this manual.

Solution

Open the **table 2.8 data.MWS** worksheet. The data for distributions A, B and C are in columns with the like letter. Select **Stat**⇒**Basic Statistics**⇒**Display Descriptive Statistics....** Double-click A, B and C to place them in the **Variables:** window. Click **Statistics...** and place a check by **skewness** and uncheck all other statistics. Click **OK** and click **OK** again. Skew for the three distributions is displayed in the Session Window. MINITAB reports skew for the three distributions as 0.00, -0.77 and 0.77 while the solution for Example 2.23 gives skew of zero, $-.648$, and $.648$. Because there are different definitions of skew, it appears that MINITAB uses a different form than does the text. The definition provided by the text is probably more common.

Task 2.15

Use MINITAB to perform the analyses outlined in Example 2.24 on page 46.

Solution

Open the **table 2.9 data** worksheet. Follow the same steps outlined in the solution to Task 2.14 with **Kurtosis** being checked rather than **Skewness** in the **Descriptive Statistics-Statistics** window. The values obtained for distributions A and B respectively are 5.50 and -1.65 which differs markedly from the values of 5.04 and 1.26 given in the text. This again occurs because MINITAB used a different definition of kurtosis than that used in the text. The definition given in the text appears to be the more common.

Exercises

- 2.1 Use the data in worksheet **table 2.3 data.MTW** to construct the grouped distributions depicted in the text Table 2.5.
- 2.2 Construct the histogram depicted in text Figure 2.3.
- 2.3 Construct a relative frequency polygon similar to that depicted in Figure 2.4 but with eight instead of 12 categories.
- 2.4 Construct a cumulative relative frequency polygon similar to that depicted in Figure 2.6 but with eight instead of 12 categories.
- 2.5 Use MINITAB to find the mean and median of the following data.

<u>Score</u>	<u>frequency</u>
9	134
8	442
7	391
6	777

- 2.6 Calculate z scores for the BP variable in **table 2.3 data.MTW**. Find the mean, standard deviation and variance of the resultant z scores.
- 2.7 Use MINITAB to find the skew of the BP data in **table 2.3 data.MTW**.

Chapter 3

Probability

3.1 Introduction

In this chapter you will learn to use MINITAB to form contingency tables as well as to find areas under the normal curve.

3.3 Contingency Tables (page 52)

Task 3.1

Use worksheet **table 3.1 data.MWS** to construct text Table 3.1 located on page 54.

Solution

Notice that C1 of worksheet **table 3.1 data.MWS** characterizes each subject as being a smoker (S) or a non-smoker (SN) while C2 characterizes each subject as having some disease (D) or not having the disease (DN). In order to form these data into a contingency table, select **Stat**⇒**Tables**⇒**Cross Tabulation and Chi-Square...** Use double-clicks to enter *Smoking Status* in the **For rows:** window and *Disease Status* in the **For columns:** window. Check **Counts** in the **Display** list and click **OK**. The contingency table, with cell entries as depicted in text Figure 3.1, appears in the Session Window.

3.4 The Normal Curve (page 62)

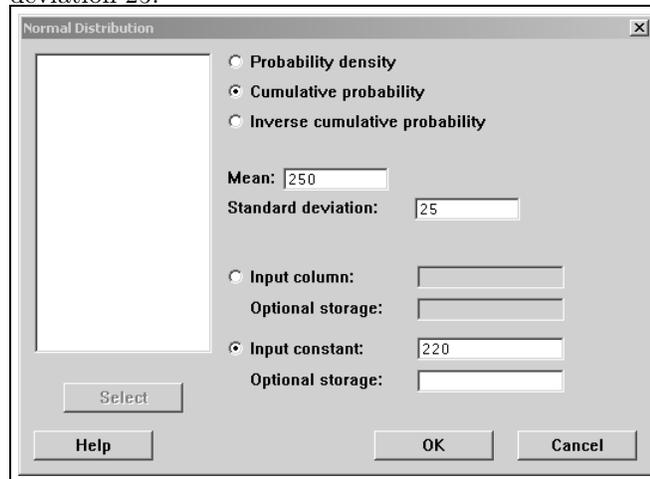
Task 3.2

Use MINITAB to find the area of the normal curve indicated in Example 3.9 on page 64.

Figure 3.1: MINTAB rendering of Table 3.1.

	D	DN	All
S	9	3	12
SN	2	6	8
All	11	9	20
Cell Contents:			Count

Figure 3.2: Screen appearance for finding the area below 220 of a normal curve with mean 250 and standard deviation 25.



Solution

MINITAB can find the area below any point on a normal distribution. To find the indicated area for this task, select **Calc**⇒**Probability Distributions**⇒**Normal...** The **Normal Distribution** window opens. Be sure the **Cumulative probability** button is selected. This tells MINITAB that the area below the specified value is to be calculated. Now enter the mean of 250 and standard deviation of 25 in the appropriate windows. Be sure the **Input constant:** button is selected and type 220 in the accompanying window. Your screen should appear as in Figure 3.2. Click **OK** and the area, 0.115070, appears in the Session Window. This agrees with the answer of .1151 given in the text.

Exercises

3.1 Construct a worksheet to represent the data in the accompanying contingency table. Then use the worksheet to construct the contingency table.

	B	\bar{B}	
A	1	3	4
\bar{A}	3	3	6
	4	6	

3.2 Use MINITAB in lieu of the normal curve table (text Appendix A) to find the areas of the normal curve indicated in Examples 3.10, 3.11 and 3.12.

Chapter 4

Introduction to Inference and One Sample Methods

4.1 Introduction

In this chapter you will use MINITAB to find binomial probabilities as well as probabilities associated with the normal curve. You will also learn to perform various hypothesis tests, including equivalence tests, and form certain confidence intervals. In addition, you will perform power calculations and compute required sample sizes to attain specified levels of power.

4.2 Sampling Distributions (page 75)

Task 4.1

Generate the binomial probabilities in Table 4.1 on page 83.

Solution

Enter the values 0 through 5 in C1 of a new worksheet. Then select **Calc**⇒**Probability Distributions**⇒**Binomial...** In the **Binomial Distribution** window select the **Probability** button. Type 5 in the **Number of trials:** window and .10 in the **Probability of success:** window. Select the **Input column:** button and type C1 in the accompanying window. This tells MINITAB the number of successes for which probabilities are to be calculated. If you wish to have the probabilities placed in C2, type C2 in the accompanying **Optional storage:** window or, if you prefer, leave this window blank and the probabilities will be printed in the Session Window. Now click **OK** and the probabilities, as shown in Figure 4.1, appear in the Session Window.¹

¹Or in C2 if you chose that option.

Figure 4.1: MINITAB rendering of binomial probabilities in Table 4.1.

Binomial with n = 5 and p = 0.1	
x	P(X = x)
0	0.59049
1	0.32805
2	0.07290
3	0.00810
4	0.00045
5	0.00001

Task 4.2

The Z value calculated in connection with Example 4.7 on page 86 was -5.33 . Use MINITAB to find the associated probability.

Solution

Select **Calc**⇒**Probability Distributions**⇒**Normal...** In the **Normal Distribution** window be sure the **Cumulative probability** button is selected. This tells MINITAB to give us the proportion of the curve that lies below the indicated Z value. Now enter 0.0 and 1.0 for the mean and standard deviation respectively. This is because we are dealing with a distribution of Z scores (i.e. the standard normal curve) which has mean zero and standard deviation one. Select the **Input constant:** button and enter -5.33 in the accompanying window and click **OK**. The result appears in the Session Window as 0.0000000. While the result cannot be exactly zero, it is zero to seven decimal places.

4.3 Hypothesis Testing (page 86)

Task 4.3

Use MINITAB to perform the one mean Z test outlined in Example 4.9 on page 92.

Solution

Select **Stat**⇒**Basic Statistics**⇒**1-Sample Z...** In the **1-Sample Z (Test and Confidence Interval)** window select the **Summarized data** button. This tells MINITAB that you will supply the sample size and mean rather than having MINITAB calculate it from a data sample. In the **Sample size:** window type 150 and in the **Mean:** window type 128.2 . In the **Standard deviation:** window type 40 then type 120 in the **Test mean:** window. We must now inform MINITAB as to the nature of the alternative hypothesis. To do this, select **Options...** and in the **Alternative:** window select *greater than*. Now click **OK** and click **OK** again. MINITAB displays obtained Z of 2.51 with

a one-tailed p -value of .006. Both numbers agree with those reported in the solution to Example 4.9.

Normally, we would not bother performing the critical Z versus obtained Z version of the test but in the spirit of learning more about MINITAB we will next find critical Z . To do this select **Calc**⇒**Probability Distributions**⇒**Normal...** Select the **Inverse cumulative probability** button in the **Normal Distribution** window. This tells MINITAB that you want a Z value defined by having a specified proportion of the normal curve lying below it. Because we are dealing with a one-tailed test with alternative of the form $H_A : \mu > 120$ and $\alpha = .01$, the desired Z value cuts off .01 in the upper tail of the curve which means that .99 of the curve lies below it. Choose the **Input constant:** button and enter .99 in the accompanying window. Be sure the **Mean:** and **Standard deviation:** are specified as 0.0 and 1.0 respectively and click **OK**. MINITAB gives critical Z (labeled X in the output) as 2.32635 which agrees with that given in the solution to Example 4.9.

Task 4.4

Use MINITAB to perform the one mean Z test outlined in Example 4.12 on page 97.

Solution

Select **Stat**⇒**Basic Statistics**⇒**1-Sample Z...** In the **1-Sample Z (Test and Confidence Interval)** window select the **Summarized data** button. This tells MINITAB that you will supply the sample size and mean rather than having MINITAB calculate it from a data sample. In the **Sample size:** window type *40* and in the **Mean:** window type *87.1*. In the **Standard deviation:** window type *21* then type *80* in the **Test mean:** window. We must now inform MINITAB as to the nature of the alternative hypothesis. To do this, select **Options...** and in the **Alternative:** window select *not equal* which specifies a two-tailed test. Now click **OK** and click **OK** again. MINITAB displays obtained Z of 2.14 with a two-tailed p -value of 0.032. Both numbers agree with those reported in the solution to Example 4.12.

To find critical Z , select **Calc**⇒**Probability Distributions**⇒**Normal...** Select the **Inverse cumulative probability** button in the **Normal Distribution** window. This tells MINITAB that you want a Z value defined by having a specified proportion of the normal curve lying below it. Because we are dealing with a two-tailed test with $\alpha = .10$, one of the desired Z values cuts off .05 in the upper tail of the curve which means that .95 of the curve lies below it. Choose the **Input constant:** button and enter .95 in the accompanying window. Be sure the **Mean:** and **Standard deviation:** are specified as 0.0 and 1.0 respectively and click **OK**. MINITAB gives critical Z as 1.64485. Because the test is two-tailed we know the critical values are ± 1.64485 which, with the exception of the number of decimal places employed, agrees with those given in the solution to Example 4.12.²

²Note that 1.64485 rounds to 1.64 when rounded to two decimal places. However, the value

Task 4.5

Use MINITAB to perform the one mean t test outlined in Example 4.14 on page 104.

Solution

Enter the data from Example 4.14 into C1 of a new worksheet. Label this variable x . Now select **Stat**⇒**Basic Statistics**⇒**1-Sample t...** In the **1-Sample t (Test and Confidence Interval)** window, be sure the **Samples in columns:** button is chosen. Type $C1$ in the accompanying window. Now type 8 in the **Test mean:** window and then click on **Options...** Specify **Alternative:** as *less than*. Now click **OK** and click **OK** again. MINITAB reports obtained t as -3.13 which agrees with the calculation in the text. The associated p -value is given as 0.010 .

Normally, we would not bother performing the critical t versus obtained t version of the test but in the spirit of learning more about MINITAB we will next find critical t . To do this select **Calc**⇒**Probability Distributions**⇒**t...** Select the **Inverse cumulative probability** button in the **t Distribution** window. This tells MINITAB that you want a t value defined by having a specified proportion of the t curve lying below it. Because we are dealing with a one-tailed test with alternative of the form $H_A : \mu < 8$ and $\alpha = .01$, the desired t value cuts off $.01$ in the lower tail of the curve. Type 6 in the **Degrees of freedom:** window. Choose the **Input constant:** button and enter $.01$ in the accompanying window and click **OK**. MINITAB gives critical t as -3.14267 which agrees with that given in the solution to Example 4.14.

Notice that this is a rare instance in which the decision *appears* to depend on whether you perform the p -value versus alpha version of the test or the obtained t versus critical t version of the test. Because the reported p -value is $.01$, you would reject on the basis of this criterion. However, MINITAB rounded this p -value to $.01$. The actual value is closer to 0.0102012 which would result in failure to reject the null hypothesis.

Task 4.6

Use MINITAB to perform the one sample test for a proportion outlined under the heading **one-tailed test** on page 108.

Solution

Select **Stat**⇒**Basic Statistics**⇒**1 Proportion...** Choose the **Summarized data** button and enter 10 for **Number of trials:** and 6 for **Number of events:**. Now click on **Options...** and enter $.38$ in the **Test proportion:** window. Select *greater than* in the **Alternative:** window and click **OK** twice. MINITAB reports an exact p -value of 0.135 which agrees with that reported in the text.

1.65 is used by convention because it is slightly more conservative.

Task 4.7

Use MINITAB to perform the one sample test for a proportion outlined in Example 4.19 on page 113.

Solution

Select **Stat**⇒**Basic Statistics**⇒**1 Proportion...**. Choose the **Summarized data** button and enter 8 for **Number of trials:** and 2 for **Number of events:**. Now click on **Options...** and enter .35 in the **Test proportion:** window. Select *not equal* in the **Alternative:** window and click **OK** twice. MINITAB reports an exact p -value of 0.721 which does not agree with the value of .85562 reported in the text. This is because MINITAB uses a much more complicated method for finding the p -value for two-tailed tests based on the binomial distribution.

Task 4.8

Use MINITAB to carry out the two-tailed equivalence test outlined in Example 4.25 on page 120.

Solution

Begin by entering the sample data into C1 of a new worksheet. We can now use MINITAB to conduct two one-tailed t tests as is done in the text. However, there is a simpler way to establish two-tailed equivalence which we will also demonstrate.³ We begin by conducting the tests as was done in the text.

To conduct Test One select **Stat**⇒**Basic Statistics**⇒**1 Sample t...**. Be sure the **Samples in columns:** button is selected and type C1 in the accompanying window. Type 2 in the **Test mean:** window. Now click on the **Options...** button and choose *less than* in the **Alternative:** window. Click **OK** and click **OK** again. MINITAB reports obtained t in the Session Window as -3.64 which agrees with the result given in the text. The reported p -value is 0.003 which results in rejection of the null hypothesis.

Test Two is carried out in the same manner with -2 replacing 2 in the **Test mean:** window and *greater than* replacing *less than* in the **Alternative:** window. Obtained t is reported as 2.18 which agrees with that given in the text. The p -value is given as 0.028 which results in rejection of the null hypothesis. Because both tests are significant, the null hypothesis of non-equivalence is rejected in favor of the alternative of equivalence.

A much simpler way of conducting a two-tailed equivalence test⁴ is by forming a $100(1 - 2\alpha)$ percent confidence interval and noting the relationship of L and U to EI_L and EI_U . If both L and U are between EI_L and EI_U , the

³The authors of your text presented two-tailed equivalence testing in the manner they did because they believed that, while more tedious, this method led to a better understanding of the equivalence concept.

⁴You may wish to delay learning this form of testing until you've studied Sections 4.4 and 4.5.

Figure 4.2: Confidence interval used to perform two-tailed equivalence test.

Variable	N	Mean	StDev	SE Mean	90% CI
x	10	-0.500000	2.173067	0.687184	(-1.759686, 0.759686)
					↑ ↑ L U

two-tailed equivalence null hypothesis is rejected otherwise, it is not rejected. You should be able to determine why this is true after you study Section 4.4 and especially Section 4.5. Let us demonstrate this technique on the current example.

To perform the shortcut two-tailed equivalence test, select **Stat**⇒**Basic Statistics**⇒**1 Sample t...** Be sure the **Samples in columns:** button is selected and type *CI* in the accompanying window. Click the **Options...** button and choose *not equal* in the **Alternative:** window. Now type 90 in the **Confidence level:** window. This figure is arrived at by noting that $100(1 - 2\alpha) = 100(1 - (2)(.05)) = 90$. Click **OK** and click **OK** again. The output is shown in Figure 4.2. We've marked the lower (L) and upper (U) limits of the confidence interval.

Because L of -1.759686 is greater than or equal to EI_L of -2 and U of 0.759686 is less than or equal to EI_U of 2 , the equivalence null hypothesis is rejected. Had either of these conditions not been met, the equivalence null hypothesis would not have been rejected.

Task 4.9

Use MINITAB to carry out the exact two-tailed equivalence test outlined in Example 4.26 on page 122.

Solution

We will first use MINITAB to test the equivalence null hypothesis by conducting Test One and Test Two as is done in the text. We will then use a shortcut method for conducting this test based on an exact confidence interval.⁵

To conduct Test One Select **Stat**⇒**Basic Statistics**⇒**1 Proportion...** Choose the **Summarized data** button and enter 8 for **Number of trials:** and 4 for **Number of events:**. Now click on **Options...** and enter .70 in the **Test proportion:** window. Select *less than* in the **Alternative:** window and click **OK** twice. MINITAB reports an exact p -value of 0.194 which agrees with the value reported in the text.

Conduct Test Two in the same manner as was done for Test One with the

⁵You will learn how to construct such intervals when you reach the discussion on page 149.

exceptions that $.30$ is substituted for $.70$ in the **Test proportion:** window and *greater than* is substituted for *less than* in the **Alternative:** window. The reported p -value is again 0.194 . Because both p -values are less than $\alpha = .20$, the equivalence null hypothesis is rejected.

For the shortcut method, Select **Stat**⇒**Basic Statistics**⇒**1 Proportion...** Choose the **Summarized data** button and enter 8 for **Number of trials:** and 4 for **Number of events:**. Now click on **Options...** and enter $100(1 - 2\alpha) = 100(1 - (2)(.20)) = 60$ in the **Confidence level:** window. Now enter *not equal* in the **Alternative:** window and click **OK** twice. The reported confidence interval is $(0.303226, 0.696774)$. Because this interval is completely contained in the equivalence interval of $.3$ to $.7$, the two-tailed equivalence null hypothesis is rejected.

Task 4.10

Compute power for the situation described in Example 4.32.

Solution

Select **Stat**⇒**Power and Sample Size**⇒**1-Sample Z...** In the **Power and Sample Size for 1-Sample Z** window, type 25 in the **Sample sizes:** window and -2 in the **Differences:** window. This is the difference $\mu_0 - \mu$. Be sure the **Power values:** window is blank. Type 20 in the **Standard deviation:** window then click the **Options...** button. Be sure the **Not equal** button is selected in the **Alternative hypothesis** list. Be sure the **Significance level:** is set to $.05$ and click **OK** twice. MINITAB gives the power calculation as 0.0790975 which is slightly different from the value of $.0721$ calculated in the text.

The discrepancy of $0.0790975 - .0721 = .0069975$ can be explained by reference to Panel B of Figure 4.25 on page 134 of the text. In this figure power is depicted as the darkly shaded portion of the alternative curve that lies in the lower tail critical region of the null curve. MINITAB adds to this figure the portion of the alternative curve that lies in the *right hand* critical region of the null curve the reasoning being that this too would bring about rejection of the false null hypothesis albeit for the wrong reason. That is, while $\mu < \mu_0$, rejection in the right hand tail would lead one to believe that $\mu > \mu_0$.⁶ We will leave calculation of the portion of the alternative curve that lies in the right hand critical region of the null curve as an exercise.

Task 4.11

Use MINITAB to find the power alluded to in Example 4.34 on page 135.

Solution

We will begin by demonstrating a method that seems a natural way to proceed

⁶Some authors refer to this as a Type III error while others simply treat it as part of the power calculation as does MINITAB.

but provides poor results and then show how better results can be achieved. To begin, select **Stat**⇒**Power and Sample Size**⇒**1 Proportion...** The **Power and Sample Size for 1 Proportion** window appears. Type *8* in the **Sample sizes:** window, *.5* in the **Alternative values of p:** window and *.35* in the **Hypothesized p:** window. Be sure the **Power values:** window is blank.

Now click **Options...** and choose the **Greater than** button. Type *.05* in the **Significance level:** window and click **OK** twice. Power is shown in the Session Window to be 0.235589 which is dramatically different from the value of .14454 calculated in the text. The problem arises because MINITAB uses an approximate method based on the normal curve for its calculation while the calculation in the text is exact. MINITAB will give us a better approximation if we replace the *.05* entry in the **Significance level:** window with *.02533* which is the exact level of significance calculated in the text. When we do this, MINITAB gives the power estimate as 0.154862 which is much closer to the value of .14454 calculated in the text. The lesson to be learned here is that if we are using an exact test for a proportion, we should calculate and use the exact level of significance rather than simply inputting the intended level.

Task 4.12

Use MINITAB to calculate the sample size described in Example 4.36 on page 136.

Solution

Select **Stat**⇒**Power and Sample Size**⇒**1-Sample Z...** The **Power and Sample Size for 1-Sample Z** window appears. Be sure the **Sample sizes:** window is blank. Type *2* in the **Differences:** window. This is the difference $\mu_0 - \mu$. Now type *.8* in the **Power values:** window and *4* in the **Standard deviation:** window. Click the **Options...** button and be sure the **Not equal** button is selected. Type *.01* in the **Significance level:** window and click **OK** twice. MINITAB prints the sample size of 47 in the Session Window which is the solution calculated in the text.

4.4 Confidence Intervals (page 137)

Task 4.13

Use MINITAB to construct the 95 percent confidence interval outlined in Example 4.37 on page 143.

Solution

Select **Stat**⇒**Basic Statistics**⇒**1-Sample Z...** Select the **Summarized data** button in the **1-Sample Z (Test and Confidence Interval)** window. Type *60* in the **Sample size:** window, *90.1* in the **Mean:** window and *16* in the **Standard deviation:** window. Now click the **Options...** button. Be

sure that the **Confidence level:** is set to 95.0 and the **Alternative:** is set to *not equal*. Click **OK** twice. The interval, $L = 86.0515$ and $U = 94.1485$ are provided in the Session Window. Both of these values agree with the results provided in the text.

Task 4.14

Use MINITAB to construct the confidence interval described in Example 4.41 on page 147.

Solution

Enter the data from Example 4.41 into C1 of a new worksheet. Select **Stat**⇒**Basic Statistics**⇒**1-Sample t...** Select the **Samples in columns:** button and type C1 in the accompanying window. Now click the **Options...** button and type 99.0 in the **Confidence level:** window. Be sure *not equal* is selected in the **Alternative:** window and click **OK** twice. MINITAB gives the confidence interval as (94.654, 120.146) which agrees with the result given in the text.

Task 4.15

Use MINITAB to form the exact and approximate confidence intervals constructed in Example 4.42 on page 150.

Solution

Select **Stat**⇒**Basic Statistics**⇒**1 Proportion...** Select the **Summarized data** button. Enter 10 in the **Number of trials:** window and $\frac{1}{4}$ in the **number of event:** window. Click the **Options...** button. Be sure that 95.0 is entered in the **Confidence level:** window and *not equal* is selected for the **Alternative:** window and click **OK** twice. MINITAB displays the confidence interval as (0.121552, 0.737622) which, given rounding, agrees with the exact computations carried out in the text.

To obtain the approximate result, repeat the above steps but check the **Use test and interval based on normal distribution** box under the **Options...** button. MINITAB gives the interval (0.096364, 0.703636) which, again, agrees with the approximate interval given in the text. It is noteworthy that MINITAB provides the following admonition in association with this result. “* NOTE * The normal approximation may be inaccurate for small samples.”

Exercises

- 4.1 Use MINITAB to reproduce the binomial probabilities in Table 4.5 on text page 109 on page 143.
- 4.2 In Example 4.8 on page 86, the solution to the problem involves finding the area of the normal curve that lies above a Z value of 1.36. The solution gives this area as .0895. Use MINITAB and, if necessary a calculator, to find this area.
- 4.3 Use MINITAB to perform the one mean Z test outlined in Example 4.11 on page 95 as well as to find critical Z for the test.
- 4.4 Use MINITAB to perform the one mean Z test outlined in Example 4.13 on page 98 as well as to find critical Z values for the test.
- 4.5 Use MINITAB to perform the one mean t test outlined in Example 4.16 on page 106 as well as to find critical t for the test.
- 4.6 Use MINITAB to perform the test outlined in Example 4.18 on page 111.
- 4.7 Use MINITAB to perform the test outlined in Example 4.20 on page 114.
- 4.8 Use MINITAB to perform the following test of hypothesis using both the method outlined in the text (i.e. Test One and Test Two) and the shortcut (i.e. confidence interval) method.

$$H_{0E} : \pi \leq .45 \text{ or } \pi \geq .55$$

$$H_{AE} : .45 < \pi < .55$$

where $\hat{p} = .47$ and $n = 500$. Use $\alpha = .05$.

- 4.9 In reference to Example 4.32, calculate the portion of the alternative curve that lies in the right hand critical region of the null curve thereby reconciling the difference of approximately .0069975 between the power calculation provided by MINITAB and that calculated in the text.
- 4.10 Use MINITAB to obtain a good estimate of the exact power calculated in Example 4.35.
- 4.11 Use MINITAB to form the confidence interval described in Example 4.38 on page 144.
- 4.12 Use MINITAB to form the confidence intervals described in Example 4.39 on page 144.
- 4.13 Use MINITAB to construct the confidence interval outlined in Exercise 4.22 on page 157.
- 4.14 Use MINITAB to construct the two-sided and one-sided confidence intervals alluded to in Example 4.43 on page 151.

Chapter 5

Paired Samples Methods

5.1 Introduction

In this chapter you will use MINITAB to perform tests of hypotheses and form confidence intervals for various paired samples methods. You will conduct paired samples t tests, McNemar's test, as well as tests on risk and odds ratios. You will also form confidence intervals for each of these statistics and perform equivalence tests.

5.2 Methods Related to Mean Difference (page 160)

Task 5.1

Use MINITAB to perform the paired samples t test described in Example 5.1 on page 162.

Solution

Load the worksheet containing the Table 5.1 data by selecting **File**⇒**Open Worksheet...** and opening file **table 5.1 data.MTW**. Now select **Stat**⇒**Basic Statistics**⇒**Paired t...**. In the **Paired t (Test and Confidence Interval)** window be sure the **Samples in columns** button is selected. In the **First sample:** window type *C2* then type *C1* in the **Second sample:** window. The sample designations are done in this fashion because MINITAB subtracts the second sample from the first sample when performing computations. Now click the **Options...** button and be sure that **Test mean:** is set to 0.0 and **Alternative:** is set to *not equal*. Click **OK** twice. MINITAB reports obtained t as "T-Value = 2.93" which is the value computed in the text. The p -value is given as 0.011 which is less than $\alpha = .05$ so that the null hypothesis is rejected.

Task 5.2

Use MINITAB to perform the equivalence test described in Example 5.4 on page 168.

Solution

Load the worksheet containing the Table 5.5 data by selecting **File**⇒**Open Worksheet...** and opening file **table 5.5 data.MTW**. The one-tailed equivalence test may be carried out by conducting a one-tailed paired samples t test as follows. Select **Stat**⇒**Basic Statistics**⇒**Paired t...** In the **Paired t (Test and Confidence Interval)** window be sure the **Samples in columns** button is selected. Type $C2$ in the **First sample:** window and $C1$ in the **Second sample:** window. Now click the **Options...** button. Type 6.0 in the **Test mean:** window and select *less than* in the **Alternative:** window and click **OK** twice. MINITAB gives obtained t as -2.61 which agrees with the value calculated in the text. The p -value is given as 0.009 which results in rejection of the null hypothesis.

Task 5.3

Use the paired data in Table 5.1 to construct a one-sided 95 percent confidence interval for the lower bound of μ_d .

Solution

Load the worksheet containing the Table 5.1 data by selecting **File**⇒**Open Worksheet...** and opening file **table 5.1 data.MTW**. Now select **Stat**⇒**Basic Statistics**⇒**Paired t...** In the **Paired t (Test and Confidence Interval)** window be sure the **Samples in columns** button is selected. In the **First sample:** window type $C2$ then type $C1$ in the **Second sample:** window. Now click the **Options...** button and be sure that **Confidence level:** is set to 95.0 and **Alternative:** is set to *greater than*.¹ Click **OK** twice. MINITAB reports the lower bound estimate as 2.55343 which is the value reported in the solution to Example 5.6.

5.3 Methods Related to Proportions (page 174)

Task 5.4

Use both approximate and exact methods to perform a two-tailed McNemar's test on the data in Table 5.7 on page 175 at the $\alpha = .05$ level.

¹Read the material under One-Tailed Hypothesis Tests And One-Sided Confidence Intervals on page 154 if you are not sure why the *greater than* option was chosen.

Figure 5.1: Cross tabulation of disease status for vaccine treated twins.

Rows: Vaccine Two		Columns:		
	D	D_bar	All	
D	2	2	4	
D_bar	6	6	12	
All	8	8	16	
Cell Contents:			Count	

Solution

Unfortunately, MINITAB does not provide a direct method for performing this test from the paired data. However, recalling that McNemar's test is simply a test for a proportion, we can use this facility of MINITAB to perform the test. We begin by finding the number of noninformative pairs, the number of pairs in which the vaccine one treated twin had disease but the vaccine two treated twin did not and the number of pairs in which the vaccine one treated twin had no disease while the vaccine two treated twin did develop the disease.

To begin, load the worksheet containing the Table 5.7 data by selecting **File**⇒**Open Worksheet...** and opening file **table 5.7 data.MTW**. Columns C1 and C2 contain the paired designations as to whether the twin contracted the disease (D) or did not contract the disease (D_bar) for the two vaccine groups. Now select **Stat**⇒**Tables**⇒**Cross Tabulation and Chi-Square...** Double-click on **Vaccine Two** to place this variable in the **For rows:** window (or alternatively, type C2 in the window) then double-click **Vaccine One** to place it in the **For columns:** window (or alternatively, type C1 in the window). Put a check in the **Counts** window and be sure the other items under **Display** are unchecked. Click **OK**. The output is shown in Figure 5.1.

Notice that for two pairs of twins both had the disease while for six pairs neither had the disease. Thus, there are eight noninformative pairs. Of the remaining eight pairs, there were six for whom the Vaccine One twin had disease while the Vaccine Two twin was disease free. We now have sufficient information to conduct the test.

We are now ready to perform McNemar's test. To this end, select **Stat**⇒**Basic Statistics**⇒**1 Proportion...** Choose the **Summarized data:** button then type 8 in the **Number of trials:** window and 6 in the **Number of events:** window. Now click the **Options...** button and be sure that .5 is specified in the **Test proportion:** window and *not equal* is specified in the **Alternative:** window. In order to perform the approximate test we place a check in the **Use test and interval based on normal distribution** window and click **OK** twice. MINITAB gives obtained Z as 1.41 which is the value computed in the

text on page 176.

The exact test is performed by following the above steps with the exception that the **Use test and interval based on normal distribution** window is not checked. In this case MINITAB gives the two-tailed p -value as .289 which is the value calculated in conjunction with Example 5.11 on page 179 of the text.

Task 5.5

Use the Outcome variable in Table 5.14 to perform the exact test outlined in Example 5.13 on page 183.

Solution

Enter the Outcome variable, minus the noninformative observations, from Table 5.14 into C1 of a new worksheet. Select **Stat**⇒**Basic Statistics**⇒**1 Proportion...** Select the **Samples in columns:** button and type C1 in the accompanying window. Click the **Options...** button and enter .4 in the **Test proportion:** window. Select *greater than* in the **Alternative:** window and click **OK** twice. MINITAB gives the p -value for the test as 0.514 which is the value calculated in the solution to Example 5.13.

Task 5.6

Construct the confidence intervals described in Example 5.15 on page 186.

Solution

Select **Stat**⇒**Basic Statistics**⇒**1 Proportion...** Select the **Summarized data button** and enter 10 in the **Number of trials:** window and 4 in the **Number of events:** window. Click on **Options...** and be sure that the **Confidence level:** window contains 95.0. Enter *not equal* in the **Alternative:** window and check the **Use test and interval based on normal distribution:** box to obtain the approximate interval. Now click **OK** twice. MINITAB displays the interval as (0.096364, 0.703636) which agrees with the calculations obtained in the solution to Example 5.15.

Repeat the above steps with the **Use test and interval based on normal distribution:** box unchecked. MINITAB gives the exact interval as (0.121552, 0.737622) which, except for rounding, again agrees with the calculations provided in the solution to 5.15.

5.4 Methods Related to Paired Samples Risk Ratios (page 190)

Task 5.7

Carry out the tasks outlined in Example 5.17 on page 191.

Solution

MINITAB does not provide a direct means for calculating a paired samples risk ratio. However, it can still be useful in such an endeavor in that it can be used to form the data into a two by two table thereby facilitating the calculation. This can be particularly useful for large data sets.

For the problem at hand, generate the table shown in Figure 5.1 of this manual by following the steps given in Task 5.4. You can then use the information given in this table to calculate the risk ratio via Equation 5.7 on page 191.

The test of significance conducted via Equation 5.4 was carried out in the solution to Task 5.4. To obtain the result provided by Equation 5.5, simply square the result given by Equation 5.4.

Task 5.8

Use MINITAB to perform the equivalence test as outlined in Example 5.20 on page 195.

Solution

Unfortunately, MINITAB does not provide a straightforward means for performing this test.

Task 5.9

Use MINITAB to form the confidence interval described in Example 5.22 on page 197.

Solution

Unfortunately, MINITAB does not provide a straightforward means for constructing this interval.

5.5 Methods Related to Paired Samples Odds Ratios (page 199)

Task 5.10

Use MINITAB to test the hypothesis $H_0 : OR = 1$ as described in Example 5.24 on page 201.

Solution

Using the fact that a test of $H_0 : OR = 1$ is equivalent to a test of $H_0 : \pi = .5$, we conduct the test as follows. Select **Stat**⇒**Basic Statistics**⇒**1 Proportion...** Choose the **Summarized data** button and enter 30 in the **Number of trials:** window and 25 in the **Number of events:** window. The number of trials is specified as 30 because after eliminating the noninformative pairs there are

$25 + 5 = 30$ remaining pairs that do provide information. Of these, in twenty-five pairs, the member with cancer had been exposed while the pair member without cancer was not exposed. We define this as the number of events.

Now click **Options...** and be sure that the **Test proportion:** is set to 0.5. Put a check in the **Use test and interval based on normal distribution** box and click **OK** twice. MINITAB performs the test as a Z test while a chi-square statistic is reported in the text. By squaring obtained Z of 3.65 we obtain, except for rounding, the chi-square statistic of 13.33 reported in the text.

Task 5.11

Perform the exact two-tailed equivalence test described in Example 5.27 on page 206.

Solution

Unfortunately, MINITAB does not provide a straightforward means of performing this test. However, by using the relationship between the paired samples odds ratio and proportions along with the so called shortcut method of performing equivalence tests introduced on page 39, the task can be performed without too much difficulty.

We begin by constructing a 90 (not 95) percent confidence interval for the estimation of π . To this end, select **Stat**⇒**Basic Statistics**⇒**1 Proportion...** Choose the **Summarized data** button and enter 15 in the **Number of trials:** window and 8 in the **Number of events:** window. Now click **Options...** and be sure that the **Confidence level:** is set to .90 and **Alternative:** is *not equal*. Because we are conducting an exact test, be sure the **Use test and interval based on normal distribution** box is not checked. Click **OK** twice. The confidence interval (0.299986, 0.756273) is displayed in the Session Window. Because this confidence interval is not contained in the equivalence interval $EI_L = .454$ and $EI_U = .545$, the equivalence null hypothesis is not rejected. This is the same conclusion reached in the Solution to Example 5.27.

Task 5.12

Construct the exact confidence interval described in Example 5.30 on page 211.

Solution

Unfortunately, MINITAB does not provide a straightforward means of constructing this interval. However, by using the relationship between the paired samples odds ratio and proportions, we can construct the CI for the associated proportion and then convert this interval for the estimation of π into an interval for estimating the paired samples *OR*.

To this end, select **Stat**⇒**Basic Statistics**⇒**1 Proportion...** Choose the **Summarized data** button and enter 24 in the **Number of trials:** window and 13 in the **Number of events:** window. Now click **Options...** and be sure that the **Confidence level:** is set to .95 and **Alternative:** is *not equal*. Because

we are conducting an exact test, be sure the **Use test and interval based on normal distribution** box is not checked. Click **OK** twice. The confidence interval (0.328208, 0.744470) is displayed in the Session Window. This is the same interval reported as part of the solution to Example 5.30. In order to convert this interval into an interval for the estimation of OR , you must apply Equation 5.17 (page 208) as is shown in the solution to Example 5.30.

Exercises

- 5.1 Use MINITAB to perform the test of hypothesis outlined in Example 5.2 on page 163.
- 5.2 Use MINITAB to perform the equivalence test outlined in Example 5.5 on page 169.
- 5.3 Use the paired data in Table 5.2 to form a 95 percent confidence interval for the estimation of $\mu_1 - \mu_2$ where μ_1 represents the mean for Treatment One and μ_2 represents the mean for Treatment Two.
- 5.4 Use MINITAB to perform the test outlined in Example 5.12 on page 180.
- 5.5 Use the Outcome variable in Table 5.16 to perform the tests alluded to in Example 5.14 on page 185.
- 5.6 Use MINITAB to construct the confidence intervals discussed in Example 5.16 on page 188.
- 5.7 Conduct the test outlined in Example 5.18.
- 5.8 Conduct the test of significance alluded to in Example 5.19 on page 192.
- 5.9 Use MINITAB to perform the exact two-tailed test of significance outlined in Example 5.25 on page 202.
- 5.10 Use MINITAB to construct the approximate confidence interval discussed in Example 5.28 on page 208.

Chapter 6

Two Independent Samples Methods

6.1 Introduction

In this chapter you will use MINITAB to perform tests of hypotheses and form confidence intervals for various two (independent) sample methods. You will conduct independent samples t tests and tests for differences between proportions. You will also form confidence intervals for each of these statistics and perform equivalence tests.

6.2 Methods Related to Differences Between Means (page 215)

Task 6.1

Conduct the independent samples t test described in Example 6.1 on page 220.

Solution

Open the **table 6.1 data.MTW** worksheet. Select **Stat**⇒**Basic Statistics**⇒**2-Sample t...** The **2-Sample t (Test and Confidence Interval)** window opens. Select the **Samples in different columns** button. Enter *C1* in the **First:** window and *C2* in the **Second:** window. Put a check in the **Assume equal variances** box. Click the **Options...** button. Be sure that the **Test difference:** window is set to *0.0* and the **Alternative:** window is set to *not equal*. Click **OK** twice. The result shown in Figure 6.1 appears in the Session Window.

Notice that obtained t is -0.81 which is the value calculated in the text. The associated p -value is 0.426 which results in failure to reject the null hypothesis

Figure 6.1: MINITAB output for an independent samples t test.

Two-sample T for Group One vs Group Two				
	N	Mean	StDev	SE Mean
Group One	15	131.8	12.5	3.2
Group Two	15	135.2	10.4	2.7

Difference = μ (Group One) - μ (Group Two)
 Estimate for difference: -3.40000
 95% CI for difference: (-12.02002, 5.22002)
 T-Test of difference = 0 (vs not =): T-Value = -0.81 P-Value = 0.426 DF = 28
 Both use Pooled StDev = 11.5245

at the .05 level.

Task 6.2

Conduct the equivalence test described in Example 6.3 on page 225.

Solution

Open the **table 6.1 data.MTW** worksheet. Select **Stat**⇒**Basic Statistics**⇒**2-Sample t...** The **2-Sample t (Test and Confidence Interval)** window opens. Select the **Samples in different columns** button. Enter $C1$ in the **First:** window and $C2$ in the **Second:** window. Put a check in the **Assume equal variances** box. Click the **Options...** button. Set the **Test difference:** window to -5.0 and the **Alternative:** window to *greater than*. Click **OK** twice. The result appears in the Session Window. Notice that obtained t of 0.38 is the same value calculated in the text. The associated p -value is 0.353 so that the equivalence null hypothesis is not rejected.

Task 6.3

Form the confidence interval described in Example 6.5 on page 228.

Solution

Open the **table 6.1 data.MTW** worksheet. Select **Stat**⇒**Basic Statistics**⇒**2-Sample t...** The **2-Sample t (Test and Confidence Interval)** window opens. Select the **Samples in different columns** button. Enter $C1$ in the **First:** window and $C2$ in the **Second:** window. Put a check in the **Assume equal variances** box. Click the **Options...** button. Be sure the **Confidence level:** is set to 95.0 and the **Alternative:** is set to *not equal*. Click **OK** twice. MINITAB prints the following in the Session Window.

“95% CI for difference: (-12.02002, 5.22002)”

This is the interval calculated in the text.

6.3 Methods Related to Proportions (page 230)

Task 6.4

Perform the two sample test for a difference between proportions addressed in Example 6.7 on page 233.

Solution

Select **Stat**⇒**Basic Statistics**⇒**2 Proportions...** In the **2 Proportions (Test and Confidence Interval)** window, choose the **Summarized data** button. In the **First:** window enter *314* for **Trials:** and *23* for **Events:**. In the **Second:** window enter *316* for **Trials:** and *39* for **Events:**. Now click the **Options...** button. Be sure that **Test difference:** is set to *0.0*, **Alternative:** is *not equal* and the **Use pooled estimate of p for test** box is not checked. Now click **OK** twice. MINITAB reports obtained Z as -2.12 , which is the result reported in the text, with an associated p -value of 0.034 . This value would be appropriate for a two-tailed test. For the one-tailed test addressed here we calculate $.034/2 = .017$ which causes the null hypothesis to be rejected.

Task 6.5

Perform the two-tailed equivalence test described in Example 6.9 on page 235.

Solution

For Test One, select **Stat**⇒**Basic Statistics**⇒**2 Proportions...** In the **2 Proportions (Test and Confidence Interval)** window, choose the **Summarized data** button. In the **First:** window enter *100* for **Trials:** and *7* for **Events:**. In the **Second:** window enter *100* for **Trials:** and *6* for **Events:**. Now click the **Options...** button. Set **Test difference:** to *0.04*, **Alternative:** to *less than* and be sure the **Use pooled estimate of p for test** box is not checked. Now click **OK** twice. MINITAB reports obtained Z as $-.86$, which is the result reported in the text, with an associated p -value of 0.195 .

For Test Two, follow the same steps as for Test One with the exception that **Test difference:** is set to *-.04* and **Alternative:** to *greater than*. MINITAB gives obtained Z as 1.43 , which is the value calculated in the text, with an associated p -value of 0.076 . Thus, the equivalence null hypothesis is not rejected.

Task 6.6

Construct the confidence interval described in Example 6.11 on page 236.

Solution

Select **Stat**⇒**Basic Statistics**⇒**2 Proportions...** In the **2 Proportions (Test and Confidence Interval)** window, choose the **Summarized data** button. In the **First:** window enter *299* for **Trials:** and *106* for **Events:**.

In the **Second:** window enter *313* for **Trials:** and *66* for **Events:**. Now click the **Options...** button. Be sure that **Confidence level:** is set to *95.0*, **Alternative:** is *not equal* and the **Use pooled estimate of p for test** box is not checked. Now click **OK** twice. MINITAB reports the confidence interval as (0.0730676, 0.214237). Notice that these values differ slightly from those calculated in the text. This is because the methods used by MINITAB and the text for constructing such intervals differ slightly. The method used in the text is slightly more conservative and, in the opinion of the authors, is preferable.

6.4 Methods Related to Independent Samples Risk Ratios (page 238)

Unfortunately, MINITAB does not implement the methods discussed in the text for dealing with independent samples risk ratios.

6.5 Methods Related to Independent Samples Odds Ratios (page 247)

Unfortunately, MINITAB does not implement the methods discussed in the text for dealing with independent samples odds ratios.

Exercises

- 6.1 The MINITAB output associated with Task 6.1 contains the statement “Both use Pooled StDev = 11.5245.” How does the value 11.5245 relate to the calculations carried out in the text in conjunction with Example 6.1?
- 6.2 Use MINITAB to conduct the test of hypothesis described in Example 6.2.
- 6.3 Perform the equivalence test described in Example 6.4 on page 226.
- 6.4 Find the confidence interval discussed in Example 6.6 on page 229.
- 6.5 Perform the test described in Example 6.8 on page 233.
- 6.6 Perform the equivalence test outlined in Example 6.10 on page 235.
- 6.7 Construct the one-sided confidence interval described in Example 6.12 on page 237.

Chapter 7

Multi-Sample Methods

7.1 Introduction

In this chapter you will use MINITAB to perform a One-Way ANOVA, a 2 by k chi-square test and Tukey's HSD test.

7.2 The One-Way Analysis of Variance (ANOVA) F Test (page 264)

Task 7.1

Use the data in Table 7.1 to compute MS_w , MS_b , obtained F and to produce an ANOVA table.

Solution

Load the **table 7.1b data.MTW** worksheet. Notice that C2 contains the weights of the 15 subjects while C1 contains an indicator (1, 2, or 3) as to which group the weight represents. Select **Stat**⇒**ANOVA**⇒**One-Way...** The **One-Way Analysis of Variance** window opens. Type *C2* in the **Response:** window and *C1* in the **Factor:** window. Click **OK**. The relevant portion of the MINITAB output is shown in Figure 7.1.

In the ANOVA table provided by MINITAB, between groups elements are labeled "Factor" while within group elements are labeled "Error." From the table we see that MS_w and MS_b are given as 625 and 935 respectively. Except for rounding, these are the values calculated in the text. Obtained F is given as 1.50 which is the value given in the text. The associated p -value is .263.

Figure 7.1: MINITAB rendering of ANOVA table for data in text Table 7.1.

Results for: table 7.1b data.MTW					
One-way ANOVA: Response versus Factor					
Source	DF	SS	MS	F	P
Factor	2	1870	935	1.50	0.263
Error	12	7500	625		
Total	14	9370			

Figure 7.2: Nutritional data arranged for chi-square analysis.

Worksheet 5 ***			
↓	C1	C2	C3
	Instruction	No Instruction	
1	23	39	
2	291	277	
3			

7.3 The 2 By k Chi-Square Test (page 276)

Task 7.2

Perform the chi-square test outlined in Example 7.6 on page 280.

Solution

Open a new worksheet. Label C1 “Instruction” and C2 “No Instruction.” Now enter the number of Low Birth Weight babies for the Instruction group in R1C1 and the number of No Low Birth Weight babies in R2C1. Enter the corresponding values for the No Instruction group in R1C2 and R2C2. Your worksheet should appear as in Figure 7.2.

To perform the analysis, select **Stat**⇒**Tables**⇒**Chi-Square Test (Table in Worksheet)**.... Enter *C1 C2* in the **Columns containing the table:** window and click **OK**. The output is shown in Figure 7.3. Notice that in addition to obtained chi-square with its associated p -value, the output also shows the observed and expected frequencies for each cell along with the contribution made by that cell to the overall chi-square value.¹

¹That is, the values summed in the calculation of obtained chi-square.

Figure 7.3: Chi-square analysis of nutritional data.

Chi-Square Test: Instruction, No Instruction			
Expected counts are printed below observed counts			
Chi-Square contributions are printed below expected counts			
	Instruction	No Instruction	Total
1	23	39	62
	30.90	31.10	
	2.020	2.008	
2	291	277	568
	283.10	284.90	
	0.221	0.219	
Total	314	316	630
Chi-Sq = 4.468, DF = 1, P-Value = 0.035			

7.4 Multiple Comparison Procedures (page 282)

Task 7.3

Perform the Tukey's HSD tests outlined in Example 7.9 on page 289.

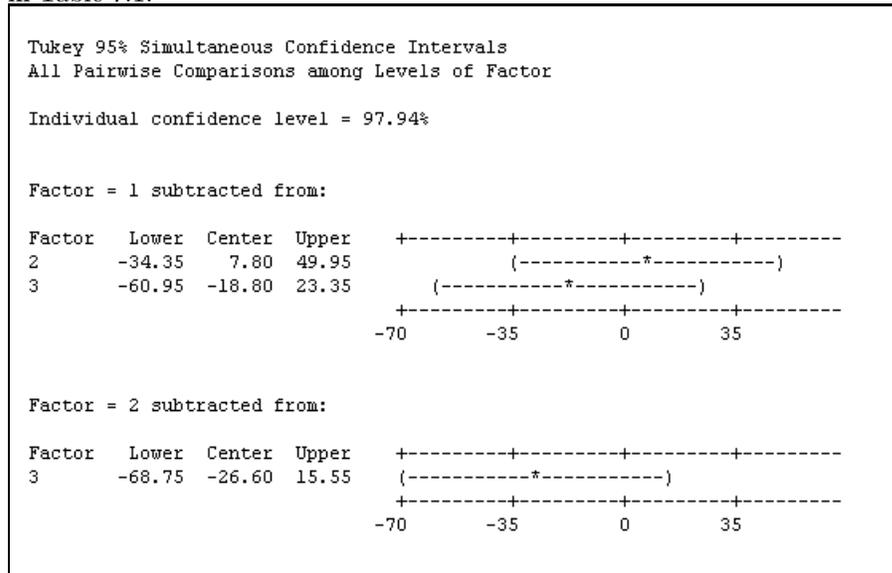
Solution

Load the **table 7.1b data.MTW** worksheet. Notice that C2 contains the weights of the 15 subjects while C1 contains an indicator (1, 2, or 3) as to which group the weight represents. Select **Stat**⇒**ANOVA**⇒**One-Way...** The **One-Way Analysis of Variance** window opens. Type *C2* in the **Response:** window and *C1* in the **Factor:** window. Click the **Comparisons...** button. Put a check in the **Tukey's, family error rate:** box and enter *.05* in the accompanying window. Click **OK** twice. In addition to the ANOVA table, the output shown in Figure 7.4 appears in the Session Window.

Notice that comparisons are made in terms of confidence intervals rather than hypothesis tests as is done in the text. Thus, the comparison $\mu_2 - \mu_1$ ² is represented by the interval $L = -34.35, U = 49.95$. Because zero is in the interval, the null hypothesis is not rejected. This is true of the remaining comparisons as well so that the conclusions provided by MINITAB are the same as those reached in the text.

²This comparison is represented as $\mu_1 - \mu_2$ in the text.

Figure 7.4: Results of Tukey's HSD tests conducted on the dieting study data in Table 7.1.



Exercises

- 7.1 Carry out the ANOVA analysis outlined in Example 7.3 on page 269.
- 7.2 Carry out the ANOVA analysis outlined in Example 7.4 on page 271.
- 7.3 Perform the chi-square analysis outlined in Example 7.5 on page 279.
- 7.4 Perform the multiple comparison tests outlined in Example 7.10 on page 290.

Chapter 8

The Assessment of Relationships

8.1 Introduction

In this chapter you will use MINITAB to calculate the Pearson product-moment correlation coefficient and to test the hypothesis $H_0 : \rho = 0$. Additionally, you will conduct the chi-square test for independence.

8.2 The Pearson Product-Moment Correlation Coefficient (page 295)

Task 8.1

Calculate the Pearson product-moment correlation coefficient for the data in Table 8.1 on page 297. Perform a two-tailed test of the hypothesis $H_0 : \rho = 0$ at the .05 level.

Solution

Load the **table 8.1 data.MTW** worksheet. Select **Stat**⇒**Basic Statistics**⇒**Correlation...** Enter x and y in the **Variables:** window and be sure the **Display p-values** box is checked. Click **OK**. MINITAB displays the correlation as 0.964 which is the value calculated in the text. The associated p -value of 0.000 results in reject of the null hypothesis $H_0 : \rho = 0$.

Figure 8.1: Chi-square analysis of data described in Example 8.7.

	C1	C2	C3	C4
	Vaccinated	Not Vaccinated	Unknown	
1	41	126	452	
2	202	210	440	
3	330	614	680	
4				

8.3 The Chi-Square Test For Independence (page 312)

Task 8.2

Perform the chi-square test described in Example 8.7 on page 314.

Solution

Enter the observed frequencies into a new worksheet as shown in manual Figure 8.1. Then select **Stat**⇒**Tables**⇒**Chi-Square Test (Table in Worksheet)**.... Enter *C1 C2 C3* in the **Columns containing the table:** window and click **OK**. The observed and expected frequencies for each cell of the table are printed along with the contribution made by each cell to the overall chi-square statistic. Obtained chi-square of 206.811 is reported with 4 degrees of freedom and an associated *p*-value 0.000.

Exercises

8.1 Calculate the correlation coefficient discussed in Example 8.2 on page 298.

Chapter 9

Linear Regression

9.1 Introduction

In this chapter, you will use MINITAB to construct simple and multiple regression models, generate \hat{R}^2 values and conduct tests of significance.

9.2 Simple Linear Regression (page 320)

Task 9.1

Construct the simple linear regression model alluded to in Example 9.1 on page 321.

Solution

Open the **table 8.2 data** worksheet. Select **Stat**⇒**Regression**⇒**Regression...** In the **Regression** window, type *C2* in the **Response:** window and *C1* in the **predictors:** window. Click **Options...** and be sure there is a check in the **Fit intercept** box. Click **OK** twice. MINITAB gives the regression equation as $\hat{y} = 0.874 + 0.706x$ which is the result given in the text.

Task 9.2

Find SS_{reg} , SS_{res} , and \hat{R}^2 for the data in Table 9.1 on page 323.

Solution

Open the **table 9.1 data** worksheet. Select **Stat**⇒**Regression**⇒**Regression...** In the **Regression** window, type *C1* in the **Response:** window and *C2* in the **predictors:** window. Click **Options...** and be sure there is a check in the **Fit intercept** box. Click **OK** twice. MINITAB gives the Regression sum of squares (SS_{reg}) as 107.44 and the Residual Error sum of squares (SS_{res}) as

8.29 which agrees with the values calculated in the text. \widehat{R}^2 is given as R-Sq = 92.8% which is the value calculated in the solution to Example 9.2 on page 324.

Task 9.3

Test the null hypothesis $R^2 = 0$ for a model used to predict y from x for the data in Table 9.1 on page 323.

Solution

Using the output from the solution to Task 9.2, we see that obtained F is reported as 168.49 with an associated p -value of 0.000. The value of 168.49 reported by MINITAB differs slightly from the value of 167.56 calculated in the text. This small difference appears to be the result of differences in rounding.

9.3 Multiple Linear Regression (page 329)

Task 9.4

Use the data in Table 9.2 on page 330 to construct a two predictor model with x_1 and x_2 being used to predict y .

Solution

Open the **table 9.2 data** worksheet. Select **Stat**⇒**Regression**⇒**Regression...** In the **Regression** window, type $C1$ in the **Response:** window and $C2$ $C3$ in the **predictors:** window. Click **Options...** and be sure there is a check in the **Fit intercept** box. Click **OK** twice. MINITAB gives the regression model as $\hat{y} = 2.69 + 0.404x_1 + 1.00x_2$ which agrees with the results calculated in the text.

Task 9.5

Test the null hypothesis $H_0 : \beta_1 = \beta_2 = 0$ for the model you constructed in Task 9.4.

Solution

From the output for the solution to Task 9.4, we note that obtained F for the regression model is 4.73 which is the value calculated in the text. The associated p -value is 0.031 which results in rejection of the null hypothesis for a test conducted at $\alpha = .05$. This is the same conclusion reached in the text.

Task 9.6

Use the data in Table 9.2 on page 330 to determine whether adding x_2 to a model that contains x_1 adds significantly to \widehat{R}^2 .

Figure 9.1: Sequential General Linear Model analysis of Table 9.2 data.

Analysis of Variance for y, using Sequential SS for Tests						
Source	DF	Seq SS	Adj SS	Seq MS	F	P
x1	1	150.83	14.19	150.83	5.90	0.032
x2	1	90.78	90.78	90.78	3.55	0.084
Error	12	306.79	306.79	25.57		
Total	14	548.40				

Solution

This problem can be approached in a number of different ways. For example, we could use MINITAB to find $\widehat{R}_{y,1}^2$ and $\widehat{R}_{y,12}^2$ then use a calculator to apply Equation 9.24. We will use a more direct, if somewhat less obvious, approach.

Select **Stat**⇒**ANOVA**⇒**General Linear Model...** The **General Linear Model** screen opens. Type y in the **Responses:** window and $x1$ $x2$ in the **Model:** window. To this point we have set up a model in which y is the dependent variable with $x1$ and $x2$ being predictors. Now click on **Covariates...** and enter $x1$ $x2$ in the **Covariates:** window. Among other things, this last action cause MINITAB to treat $x1$ and $x2$ as continuous variables rather than transforming them as might be done in other forms of analysis. Click **OK**. Now click **Options...** In the **Sum of Squares:** column, darken the **Sequential (Type I)** button. Click **OK** twice. The relevant portion of the output is shown in manual Figure 9.1. This output will require some explanation.

Of interest for the problem at hand are the headings labeled “Seq SS” (Sequential Sum of Squares), “F” (obtained F), and “P” (p -value). The Seq SS values represent the *increment* in SS_{reg} that occurs as each variable is added to the model.¹ Thus, with only x_1 in the model, SS_{reg} was 150.83. Adding x_2 to the model increased this value by 90.78. It will be instructive to convert these values into expressions of R^2 . This can be done by dividing each value by SS_y which is labeled in the printout as “Total.” For x_1 we have

$$\widehat{R}^2 = \frac{150.83}{548.40} = .2750$$

which is equal to $\widehat{R}_{y,1}^2$. Of more interest is the \widehat{R}^2 value obtained from x_2 which is

$$\widehat{R}^2 = \frac{90.78}{548.40} = .1655$$

which is equal to $\widehat{R}_{y,12}^2 - \widehat{R}_{y,1}^2$. The values of F reported in the output are the values of obtained F for a test of significance on each of the incremental sums of squares. It follows that obtained F of 3.55 is obtained F for a test conducted on

¹The order in which variables are added to the model is determined by the order in which they are listed in the **Model:** window.

$\widehat{R}_{y.12}^2 - \widehat{R}_{y.1}^2$. Indeed, this is the value calculated on page 337 of the text. The associated p -value of 0.084 means that we would fail to reject the null hypothesis $R_{y.12}^2 - R_{y.1}^2 = 0$ at the .05 level. This is the conclusion reached on page 337 of the text. Therefore, we are unable to demonstrate that adding x_2 to a model that contains x_1 adds to R^2 .

To further elucidate, suppose we were working with four predictors, x_1 , x_2 , x_3 and x_4 and had entered them in that order in the **Model:** box. The obtained F and associated p -values for each predictor would represent tests of the following null hypotheses.

<u>Source</u>	<u>Hypothesis</u>
x_1	$R_{y.1}^2 = 0$
x_2	$R_{y.12}^2 - R_{y.1}^2 = 0$
x_3	$R_{y.123}^2 - R_{y.12}^2 = 0$
x_4	$R_{y.1234}^2 - R_{y.123}^2 = 0$

Exercises

- 9.1 Find the regression equation for predicting y from x for the data in Table 9.1 on page 323.
- 9.2 Use MINITAB to construct the model described in Exercise 9.7 on page 341 of the text.
- 9.3 Test the null hypothesis $H_0 : R^2 = 0$ for the model you constructed in Exercise 9.2.
- 9.4 Use the data in Table 9.2 on page 330 to determine whether adding x_1 to a model that contains x_2 adds significantly to \widehat{R}^2 .

Chapter 10

Methods Based on the Permutation Principle

10.1 Introduction

In this chapter you will use MINITAB to perform Wilcoxon's signed-ranks test, Wilcoxon's rank-sum (Mann-Whitney) test, the Kruskal-Wallis test and Fisher's exact test.

10.3 Applications (page 348)

Task 10.1

Perform the Wilcoxon's signed-ranks test described in Example 10.15 on page 369.

Solution

Open the **table 10.7 data.MTW** worksheet. Notice that this file contains difference scores rather than the pre- and post-treatment values. Select **Stat**⇒**Nonparametrics**⇒**1-Sample Wilcoxon...** Type *C1* or *d* in the **Variables:** window. Darken the **Test median:** button and type *0.0* in the accompanying window if necessary. Be sure the **Alternative:** window is set to *not equal*. Click **OK**.

MINITAB reports the test *p*-value as 0.262 which results in failure to reject the null hypothesis at $\alpha = .05$. This is the same conclusion reached in the text. It should be noted, however, that *p*-values reported by MINITAB for Wilcoxon's signed-ranks are approximations rather than exact values.

Figure 10.1: Data arranged for Kruskal-Wallis test analysis.

	C1	C2	
	Factor	Response	
1	1	4	
2	1	9	
3	2	-1	
4	2	0	
5	3	7	
6	3	12	

Task 10.2

Perform the Wilcoxon's rank-sum test described in Example 10.20 on page 382.

Solution

Enter the data accompanying Example 10.20 into a new worksheet. Select **Stat**⇒**Nonparametrics**⇒**Mann-Whitney...**¹ The **Mann-Whitney** page opens. Type *C1* in the **First Sample:** window and *C2* in the **Second Sample:** window. Set **Alternative:** to *less than* and click **OK**.

MINITAB gives the one-tailed p -value as 0.3313. As noted in the solution to Example 10.20, the exact p -value is .35 so that the approximation provided by MINITAB is quite good even for this case in which sample size is small.

Task 10.3

Perform a Kruskal-Wallis test on the raw data in the table on page 393.

Solution

Enter the group number, i.e. 1, 2, or 3, of each of the 6 subjects in *C1* and their outcome measure in *C2*. The entries should appear as in Figure 10.1. Now select **Stat**⇒**Nonparametrics**⇒**Kruskal-Wallis...** The **Kruskal-Wallis** page opens. Enter *C2* in the **Response:** window and *C1* in the **Factor:** window. Click **OK**. MINITAB gives the p -value for the test as 0.156. As is shown in the solution to this problem (page 395), the exact p -value is .20 so that the MINITAB approximation is not too good. But this is to be expected with such small sample sizes.

¹Recall that the Mann-Whitney test is just a different form of the Wilcoxon rank-sum test.

Task 10.4

Perform a two-tailed Fisher's exact test on the data in Table 10.23 on page 405.

Solution

Open the **table 10.23 data.MTW** worksheet. Notice that C1 contains row numbers and C2 the column numbers of the cells that makeup Table 10.23. Taken together, these columns uniquely identify each cell of the table. Column C3 provides the frequency in each of the four cells.

Select **Stat**⇒**Tables**⇒**Cross Tabulation and Chi-Square...** The **Cross Tabulation and Chi-Square** page opens. Type *C1* in the **For rows:** window, *C2* in the **For columns:** window and *C3* in the **Frequencies are in:** window. Now click **Other Stats...** and place a check in the **Fisher's exact test for 2x2 tables** box. Click **OK** twice.

MINITAB gives the p -value for the test as 0.0889111 which is the value calculated for a two-tailed test in the solution to Example 10.30 on page 405. Notice that MINITAB did not provide the option of performing a one-tailed test.

Exercises

- 10.1 Conduct the test described in Example 10.16 on page 371.
- 10.2 Perform the test alluded to in Example 10.21 on page 383.
- 10.3 Conduct the test described in Example 10.22 on page 385.
- 10.4 Conduct the Kruskal-Wallis test described in Example 10.27 on page 395.
- 10.5 Conduct the test described in Example 10.31 on page 407.

Appendix A

Table Relating Text Items to Tasks

This table relates various portions of the text to specific tasks and the task page number.

<u>Text</u>	<u>Task</u>	<u>Page</u>
Table 2.2	2.1	11
Table 2.3*	2.2	13
Table 2.4	2.3	13
Figure 2.1	2.4	15
Table 2.3*	2.5	17
Figure 2.4	2.6	18
Figure 2.6	2.7	19
Example 2.4	2.9	24
Example 2.7	2.10	25
Table 2.6	2.11	26
Example 2.17	2.12	28
Example 2.21	2.13	28
Example 2.23	2.14	28
Example 2.24	2.15	29
Table 3.1	3.1	31
Example 3.9	3.2	31
Table 4.1	4.1	35
Example 4.7	4.2	36
Example 4.9	4.3	36
Example 4.12	4.4	37

<u>Text</u>	<u>Task</u>	<u>Page</u>
Example 4.14	4.5	38
page 108	4.6	38
Example 4.19	4.7	39
Example 4.25	4.8	39
Example 4.26	4.9	40
Example 4.32	4.10	41
Example 4.34	4.11	41
Example 4.36	4.12	42
Example 4.37	4.13	42
Example 4.41	4.14	43
Example 4.42	4.15	43
Example 5.1	5.1	45
Example 5.4	5.2	46
Table 5.1	5.3	46
Table 5.7	5.4	46
Example 5.13	5.5	48
Example 5.15	5.6	48
Example 5.17	5.7	48
Example 5.20	5.8	49
Example 5.22	5.9	49
Example 5.24	5.10	49
Example 5.27	5.11	50
Example 5.30	5.12	50
Example 6.1	6.1	53
Example 6.3	6.2	54
Example .6.5	6.3	54
Example 6.7	6.4	55
Example 6.9	6.5	55
Example 6.11	6.6	55
Table 7.1	7.1	59
Example 7.6	7.2	60
Example 7.9	7.3	61
Table 8.1	8.1	65
Example 8.7	8.2	66
Example 9.1	9.1	69
Table 9.1*	9.2	69
Table 9.1*	9.3	70
Table 9.2*	9.4	70
Table 9.2*	9.5	70

<u>Text</u>	<u>Task</u>	<u>Page</u>
Table 9.2*	9.6	70
Example 10.15	10.1	75
Example 10.20	10.2	76
page 393	10.3	76
Table 10.23	10.4	77

* Appears more than once in list.