

# QUICK INTRODUCTION TO MINITAB

The Stern School Statistics Group uses Minitab Release 14 for Windows<sup>®</sup> as its course software. This program was chosen specifically for use with courses B01.1305, C22.0103, C22.0001, and C22.0003. It should prove useful for other courses as well. Courses C22.0103, C22.0001, and C22.0003 will be using Student Minitab, which is based on release 14.

Minitab Release 14 for Windows<sup>®</sup> must run under a Windows<sup>®</sup> 98 (or higher) operating system. When running Minitab, do **Help** ⇒ **About Minitab**; this will show you the release number (currently 14.2). If you have release 14.0 or 14.1, you can get a free upgrade at [www.minitab.com](http://www.minitab.com).

## CONTENTS

Preliminary Remarks	page 2
Student Minitab	page 2
The Minitab Environment	page 2
Data Files Types	page 4
Location of Course Files	page 4
Special Considerations for Projects	page 4
Minitab's Windows	page 5
Wiping Minitab Clean	page 6
Reading Data into Minitab	page 7
Saving Minitab's Output	page 8
Preparing a Minitab Graph	page 9
Using Minitab to Do a Regression	page 11
Creating New Variables through Transformations	page 12
Using Minitab to Find Probabilities	page 13
Special Problems	page 15
Quick Index	page 16

This document was prepared by the Statistics Group of the IOMS Department. It is intended primarily as a quick introduction, including discussions of some common situations encountered in course work with B01.1305, C22.0103, C22.0001, and C22.0003. If you have comments or suggestions, please send them to [gsimon@stern.nyu.edu](mailto:gsimon@stern.nyu.edu).

## ***Preliminary Remarks***

In this document, we will use Windows<sup>®</sup> (with the name registration ® symbol) to refer to the Microsoft operating system. For most people this will either be Windows<sup>®</sup> XP or later. We will use windows (lowercase, with no registration symbol) to refer to the rectangular working areas that are set up on your computer screen.

Words and phrases which appear on Minitab screens will be written here in boldface. The double-line arrow  $\Rightarrow$  represents a mouse click. Thus, the notation **File**  $\Rightarrow$  **Open Worksheet**  $\Rightarrow$  indicates that you should click on **File** and then click on **Open Worksheet**.

Bold italics will be used in cross-referencing sections of this document.

The sunburst symbol  $\star$  will be used for subtle points. You might wish to skip these points at first reading.

## ***Student Minitab***

Student Minitab is bundled together with a number of undergraduate textbooks. Student Minitab is compatible with the regular release 14 series. The major limitations, compared to the full-featured program, are these:

- Student Minitab allows only 5 worksheets.

- Student Minitab allows a maximum of 10,000 cells per worksheet.

- Student Minitab does not have 3D graphs, the goodness-of-fit to the Poisson distribution, partial least squares, general linear models.

- Student Minitab has limited features for process control, design of experiments, reliability analysis, and multivariate analysis.

These limitations will not be a problem for coursework at Stern.

## ***The Minitab Environment***

Minitab is organized around *worksheets*, arrays of information in which rows represent data points (such as business firms) and columns represent variables (such as sales, earnings, employees).

- Excel easily creates cells with implicitly-computed values. Minitab can do this also, but extra effort is required.

- The results of Excel calculations (such as regression) are placed inside Excel's spreadsheet. Minitab places the results of calculations in a different window.

- Excel is casual about the distinction between an unfilled cell and missing data. Minitab treats all unfilled cells as missing data.

Excel is the program of choice for standard spreadsheet operations. Minitab is vastly superior to Excel for statistical tasks. Moreover, Excel is still plagued with errors resulting from inferior computational algorithms and from the inconsistent handling of empty cells.

In most cases, you will be working with one worksheet at a time. You can enter your own data to create a worksheet, or you can read a worksheet from memory. You can modify the worksheet and save the changes to memory. The modified worksheet can be saved with the original name or with a new name, and the file type can be changed as well. (See the section on *Data File Types*.) A saved worksheet consists only of the worksheet window and not any graphs or computations that were made from it.

A *project* consists of one or more worksheets. Project files are very powerful:

A project can keep multiple worksheets together.

Projects can hold user-created descriptions and comments. (See the section on *Project and Worksheet Descriptions*.)

Projects retain the contents of the session window. (See the section on *Minitab's Windows*.)

Projects can hold graphical output.

✿ Reports can be assembled on the ReportPad and saved with the project. Every information panel of Minitab will retain the commands from its most recent use.

✿ The worksheets of a project can be individually saved and opened later outside the context of the project. (This is recommended only for providing the data to others.)

Minitab acquires its worksheets in two basic ways:

When you open Minitab, you will start with a clean worksheet into which you can enter data. (If Minitab is holding previous work, you can wipe it clean with the sequence **File** ⇒ **New** ⇒ **Minitab Project**. This is noted also in the section *Wiping Minitab Clean*.)

You can read a worksheet from memory.

If a worksheet is open, you can still open a blank worksheet with the sequence **File** ⇒ **New** ⇒ **Minitab Worksheet** ⇒. This action does *not* erase the first worksheet.

If a worksheet is open, you can still read another worksheet from memory with the sequence **File** ⇒ **Open Worksheet** ⇒. This action does *not* erase the first worksheet.

If you are working with two or more logically-connected worksheets, you might want to save these as a project. Use **File** ⇒ **Save Project** ⇒ or perhaps **File** ⇒ **Save Project As** ⇒.

## ***Data File Types***

Most course files will be single spreadsheets. The following formats are used:

- \*.mtw            These are created by Minitab and read automatically into Minitab. These are encrypted.
- \*.mtp            There are “Minitab portable” files created by Minitab and read automatically into Minitab. These are ASCII text (not encrypted), so you can examine them in any text processor.
- \*.dat, \*.txt     These are ASCII text, nearly always created outside of Minitab. Minitab will require user-supplied guidance to read such files correctly. For example, Minitab needs to be told whether the file does or does not have variable names and needs to know how data fields are separated. Please see the section ***Reading Data into Minitab***.
- \*.xls            These are created by Excel. If the layout is simple, Minitab will read these automatically. Be aware that Excel sheets can have graphical content, and this may confuse Minitab.

✱ Minitab will also read files from dBase, Lotus 1-2-3, and Quattro Pro.

✱ You can save a worksheet in a format different from its original. Use **File ⇒ Save Worksheet As ⇒**.

Some files may be provided as projects. Simply use **File ⇒ Open Project ⇒**.

## ***Location of Course Data Files***

Course textbooks contain data disks, and you can copy the relevant information to another location, usually the C: drive of your computer. The data files for these Stern courses can also be obtained from this Web address:

<http://w4.stern.nyu.edu/ioms/docs/sg/data>

## ✱ ***Special Considerations for Projects***

The command **File ⇒ Project Description ⇒** will allow you to create a description for the project. You can also use this to edit a current description.

Be sure to find the project manager icon on the toolbar. It's the set of cascading windows, identified with the tag **Show Worksheets Folder**.

If you left-click on any worksheet name, that worksheet will become active. If you right-click on any worksheet name, you can edit its name with **Rename** and you can leave a description with **Set Description**.

There are cases in which you might want to copy a worksheet within a project. Start with **File** ⇒ **New** ⇒ **Minitab worksheet** ⇒ . At this point a blank worksheet will appear. Now go to the original worksheet and place the cursor in the name box for the first column, C1. Hold down the shift key and use the arrow keys to mark the entire worksheet as a block. Go to the blank new worksheet, place the cursor in the name box for the first column and press Ctrl+V. A copy will now appear in the new worksheet, and you can edit it as appropriate. It is strongly recommended that you use **Show Worksheets Folder** to name and describe your activities.

Another icon shown as cascading windows is identified with the tag **Show Graphs Folder**. This can be used in a similar fashion to name and describe a graph.

Worksheet descriptions cannot be used for worksheets saved in \*.mtp or \*.mtw format. These descriptions are only available for individual sheets within projects saved in \*.mpj format.

## ***Minitab's Windows***

Minitab works through three basic windows but has additional windows to deal with other features of your work. The three basic windows are these:

**Session window**...All textual (non-graphical) output will appear in the session window. As appropriate, headings will reflect the commands that you have actually created. Sometimes this window will also show the text of the Minitab commands (followed by >) equivalent to your mouse clicks. While the session window is active, do **Editor** ⇒ and then either select or deselect **Enable Commands**.

**Data window**...The data window holds the worksheet. This includes the data that you've entered or read into the program, and it may also include other variables (columns) or additional data points (rows) that were created or entered during your work.

- Project manager window...** The project manager window shows a schematic layout of all the work currently active (click on **Show Worksheets Folder**, the first icon with cascading windows)
- or
- a schematic layout of all the graphs (click on **Show Graphs Folder**, the second icon with cascading windows)
- or
- details about all the columns in the active worksheet (click on the **Show Info** icon, the letter **i** in a circle)

These basic Minitab windows can be made full screen, partial screen, or reduced to icons, but they cannot be closed. These windows are only closed when you quit Minitab.

Graph windows will be created as needed.

✿ The **History window** can be accessed by clicking on the **Show history** icon on the toolbar. This is the chevron symbol on a spiral notebook. This window provides a log of the activities, and it will include the command forms equivalent to all the mouse actions that you have done.

✿ The **ReportPad window** provides a place in which you can compose and edit details about all the work. You can reach this by clicking on the **Show ReportPad** icon, which is the red letter *A* on a spiral notebook. You can cut and paste from the session window to the reportpad, and you can also paste graphical output into this location.

You might like to explore some of the other features on the toolbar.

Click on the **Window** command or press Alt+W to get a list of available windows.

## ***Wiping Minitab Clean***

When you first start up Minitab, all its work areas will be empty. If you have finished working on a particular task and wish to start another, first be sure that you have saved the results in a form that you want. Then do **File** ⇒ **New** ⇒ **Minitab Project** ⇒. This act will empty out all the working areas.

## ***Reading Data into Minitab***

The command for this is **File** ⇒ **Open Worksheet** ⇒. You can then use any of the standard Windows<sup>®</sup> methods for bringing in this file. There are three things to watch for:

This is *not* **File** ⇒ **Open Project** ⇒, which only allows you to open projects.

Note whether a project is currently active. If so, any worksheet that you bring in will become part of this project. This may, at times, be exactly what you want. If you want the new file separated from other work, please see the section ***Wiping Minitab Clean***.

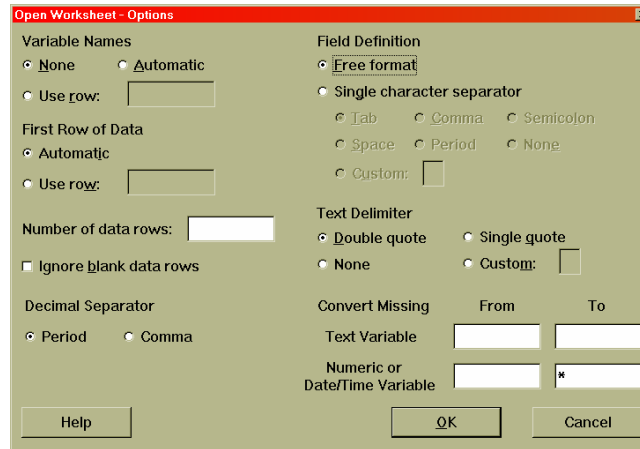
**Files of type** must acknowledge \*.mtw. The Minitab default is \*.mtw, \*.mpj and thus \*.mtw will be automatic. If \*.mtw does not show, click on the drop-down symbol ▼ to display all the file types.

The file will appear in the data window. All Minitab worksheet files, those with extension \*.mtw or \*.mtp, can be read in this fashion.

Suppose that you'd like to read a file in \*.dat format. Such files were not created with Minitab, and reading these files is not automatic. If you do not know the exact layout of the information in your \*.dat file, open that file with a text processor, such as WORD<sup>®</sup> or WordPad<sup>®</sup> or Notepad<sup>®</sup>. Watch out for

- whether or not the file has a row of variable names
- the strategy by which the file separates fields (open space for “free format,” tabs, commas, or other characters)
- whether or not alphabetic information contains internal blanks that might conflict with spaces used as separators
- the technique that is used to denote missing data
- \* whether numbers are entered in American format or in European format (in which the roles of decimal point and comma are interchanged)

Click **File** ⇒ **Open Worksheet** ⇒ . On the resulting panel, for **Files of type:** select **Data (\*.dat)**. This sequence will activate the **Options** feature, and clicking on **Options** will lead you to this panel:



The critical items here are **Variable names** and **Field definition**, but you need to be concerned about the other items as well.

Now click **OK** and then supply the file name. You can use **Preview** to see if the file will be read as anticipated.

**Important note:** The **Options** panel becomes available when you do one of these:

- select a file with extension \*.dat
- select a file with extension \*.txt
- select the **Data (\*.dat)** file type
- select the **Text (\*.txt)** file type

The **Options** panel is important! It provides exact details about the layout of the data file.

## ***Saving Minitab's Output***

You will of course want to save the results of your work. The usual methods are to cut and paste to another Windows<sup>®</sup> application or to save to file. Minitab 14 also allows you to save to the Minitab ReportPad.

Text output is moved through the copy or cut-and-paste methods common to Windows<sup>®</sup> applications. If you are working with a project file (\*.mpj) then you can also paste the selected text in the ReportPad. This is accessed through the icon which is the red letter *A* on a spiral notepad.

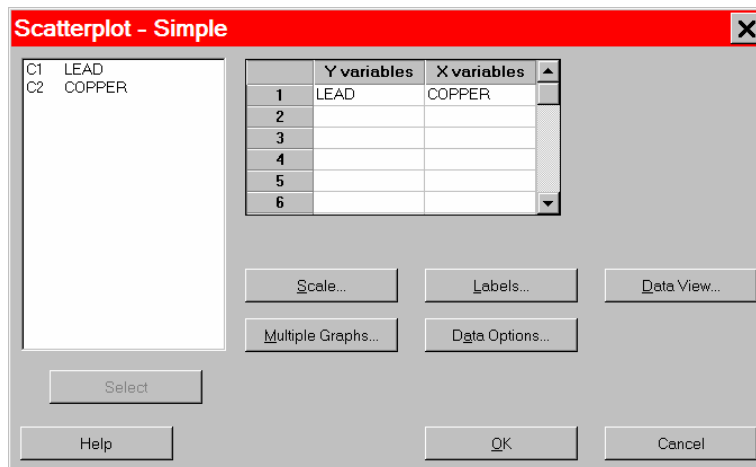
If a graph window is active, you can use **Edit** ⇒ **Copy graph** ⇒ or Ctrl+C to place the image on the Windows® clipboard. The image can then be pasted in another document or placed on the Minitab ReportPad.

✿ Pressing Alt+PrtScr copies a bitmap of the active window onto the Windows® clipboard. This can be useful if you wish to describe how you did something in Minitab. Indeed, the figures in this document were produced in exactly this way.

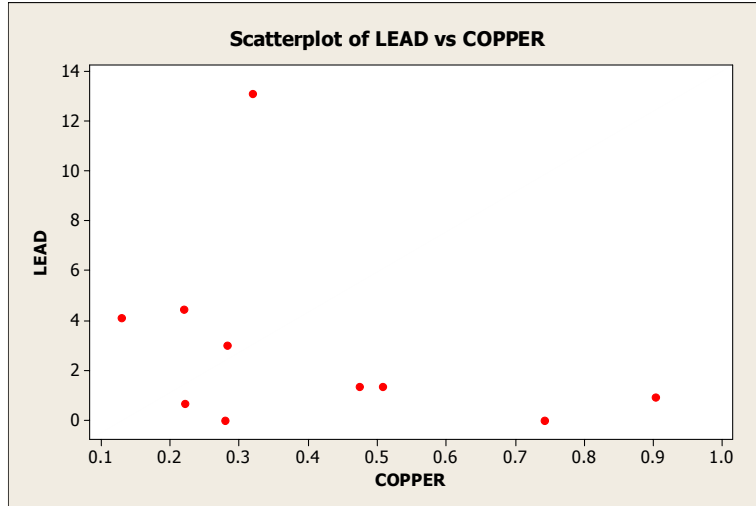
✿ If the active window is a dialog box, Alt+PrtScr copies just the dialog box to the Clipboard. If the active window is Minitab itself, Alt+PrtScr copies the Minitab frame window and all visible subwindows (such as the Data window and Graph windows) to the clipboard.

## *Preparing a Minitab Graph*

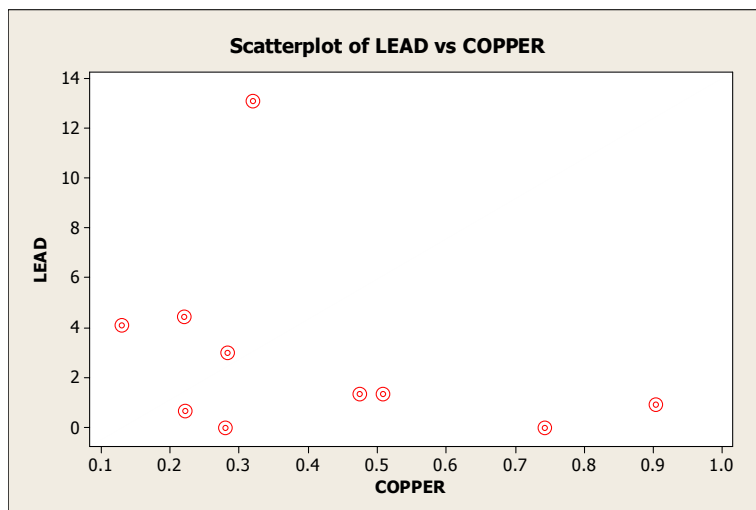
Let's show how to make a graph using the information in Minitab's data window. We'll use Exercise 7.27 from McClave-Benson-Sincich, 9<sup>th</sup> edition, as an illustration. The file name is LEADCOPP.mtw. The columns of the data give the lead concentration (in micrograms per liter) and copper concentration (in milligrams per liter) for ten residences of a subdivision near Tampa, Florida. We'd like our graph to have copper concentrations on the horizontal axis and lead concentrations on the vertical axis. Use **Graph** ⇒ **Scatterplot** ⇒ , then select **Simple**. and then put this into the resulting panel:



If you take the default settings, you'll get this:



There are many ways in which Minitab will allow you to change the appearance of this graph. Suppose that you wanted to replace the circles with another symbol. Double-click on any of the circles and the **Edit Symbols** panel will come up. You can then change the shape, size, and color of the graphing symbols. For instance, you could change the symbols to look like this:



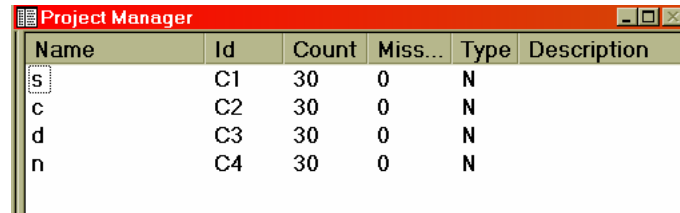
If you double-click any item on the graph (title, horizontal axis, variable name, ...) you will then be able to edit that item.

You can get additional editing power through **Tools** ⇒ **Toolbars** ⇒ **Graph Annotation Tools** ⇒. Editing features such as text insertion, lines, circles, ellipses, and so on will then be available on the Minitab toolbar when graph windows are active.

## Using Minitab to do a Regression

Let's illustrate the use of regression through the file XM1216.mtw from Hildebrand-Ott-Gray. This is a Minitab worksheet file, and it can be read automatically by Minitab.

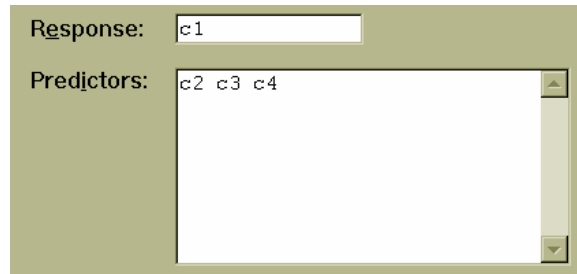
The Project Manager window (obtained by clicking on the **Show info** icon on the toolbar) will have this information:



Name	Id	Count	Miss...	Type	Description
s	C1	30	0	N	
c	C2	30	0	N	
d	C3	30	0	N	
n	C4	30	0	N	

This is summary information about the data set. It gives the variable names, the corresponding Minitab columns, number of values in each column (and how many of those are missing) and also the variable type. Here N stands for numeric. Column C1 is named s (for *sales*) and this will be the dependent variable in our regression.

To start the regression, do **Stat** ⇒ **Regression** ⇒ **Regression** ⇒. Set up the next panel to show this:



Response: c1

Predictors: c2 c3 c4

There are many tasks associated with regression, and only a few will be illustrated here. Click on **Graphs**, and then select **Residuals versus fits**. This diagnostic is a common choice in regression. Click **OK**. This will return you to the main regression panel, and you click **OK** again.

This action selects C1 as the dependent variable, with the other three variables used as independent variables. This has also asked to see a graph. The numerical output will appear in the session window, and the graph will be provided in a separate window. A little screen management will be needed to examine these.

If you wanted to see the individual residuals for your data set, you'd include an additional step from the panel reached by **Stat** ⇒ **Regression** ⇒ **Regression** ⇒. That step consists of clicking on **Storage** and then selecting **Residuals**. When the regression is completed, a new column RESI1 will appear in the data window. (Subsequent uses will create RESI2, RESI3, and so on.)

## *Creating New Variables through Transformations*

There will be many occasions in which you need to create new variables which are functions of other variables. For example, you might need to create  $\log(\text{SALES})$ , the base- $e$  logarithm of the variable SALES. The process is very simple. Click on **Calc**  $\Rightarrow$  **Calculator**  $\Rightarrow$ . In the box next to **Store result in variable:** type in the name Logsales. In the box for **Expression:** type in  $\text{LOGE}(\text{Sales})$  and click on **OK**.

As a result of this process, a new column Logsales will appear in the data window. You can of course overwrite an existing column, but then the original information in that column will be lost.

This use of the calculator option is simple, but there are some subtleties:

You need not remember function identifications such as LOGE for base- $e$  logarithms. The **Calculator** has a **Functions** panel that lists the available choices.

If you make an entry-by-entry calculation, as in the example above, then the result is a new column in the spreadsheet. If you calculate a single-valued function of a column (such as the median), then your transformation will create just one number, occupying one position in the data window spreadsheet.

✱ Logical operations take the values 0 and 1, so that you can use these to make transformations of the “if-then-else” form. Suppose, for example, that you’d like variable C3 to be defined as follows:

$$C3 = \begin{cases} 5 & \text{if } C2 \leq 5 \\ C2 & \text{if } C2 > 5 \end{cases}$$

The expression that makes this work is

$$5*(C2 \leq 5) + C2*(C2 > 5)$$

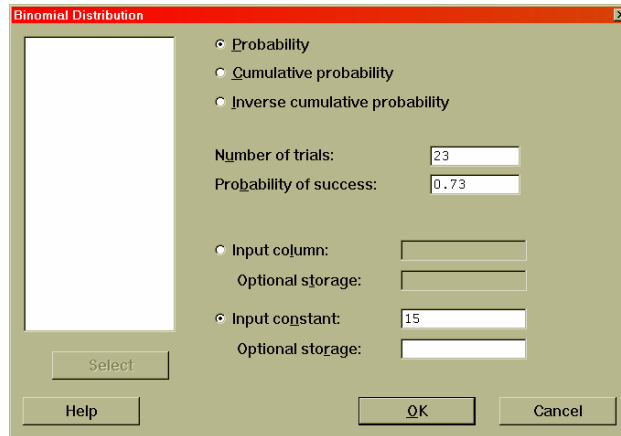
The yes-or-no logical statement  $(C2 \leq 5)$  will either take the value 0 or the value 1, as will the yes-or-no logical statement  $(C2 > 5)$ .

Missing values get passed along as missing. If you are creating a new column C6 from columns C2 and C3, and if a value is missing in row 28 for either C2 or C3, then the row 28 entry for C6 will also be missing, shown as \* in the data window.

- ✿ Invalid calculations end up as missing. Thus  $\text{LOGE}(0)$  will result in a missing value, as will  $\text{LOGE}(\text{negative number})$ . These will be shown as \* in the data window. Minitab 14 (but *not* earlier releases) gives you a warning when you create missing values in this style, and thus you get an opportunity to reconsider the appropriateness of the transformation.

## Using Minitab to Find Probabilities

You can use Minitab to find many probability calculations. For instance, if you want to find the probability that a binomial random variable with  $n = 23$  and  $p = 0.73$  takes the value 15, you call up **Calc**  $\Rightarrow$  **Probability Distributions**  $\Rightarrow$  **Binomial**  $\Rightarrow$ . Then set up the resulting panel as follows:



When you click **OK**, the session window will then show the following:

### Probability Density Function

Binomial with  $n = 23$  and  $p = 0.73$

x	P( X = x )
15	0.123375

This indicates that your desired probability is 0.123375.

If you had wanted to find the probability of a value *less than* or equal to 15, you'd do exactly the same as the above, except that you would click on the **Cumulative Probability** button.

That result shows up in the session window as

### Cumulative Distribution Function

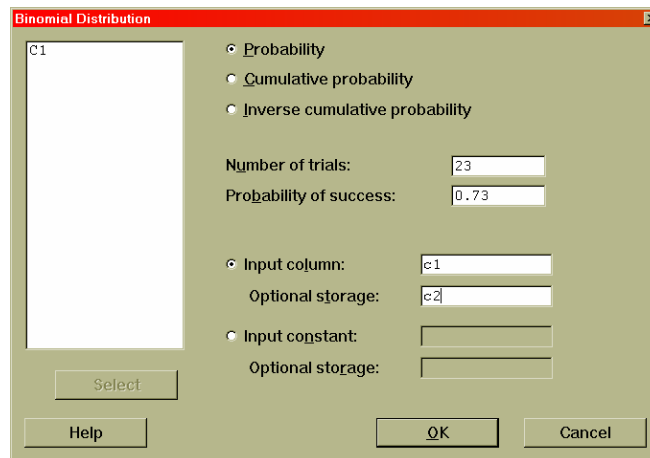
Binomial with  $n = 23$  and  $p = 0.73$

x	P( X <= x )
15	0.265051

This gives your probability as 0.265051.

It's also possible to give the complete probability distribution. Suppose that you wanted to see the whole set of probabilities for the binomial random variable with  $n = 23$  and  $p = 0.73$ . This is done by setting up a data column for the integers 0, 1, 2, ..., 23. Click on **Calc**  $\Rightarrow$  **Make Patterned Data**  $\Rightarrow$  **Simple Set of Numbers**  $\Rightarrow$ . Set up the resulting panel to ask for first value = 0, last value = 23, step = 1, with the patterned data stored in C1.

Then with **Calc**  $\Rightarrow$  **Probability Distributions**  $\Rightarrow$  **Binomial**  $\Rightarrow$ , set up the panel as indicated:



This illustration assumes that columns C1 and C2 are not being used for other purposes. At the completion of this command, the spreadsheet in the data window would have columns C1 and C2 showing the following information (slightly rearranged here):

0	0.000000	12	0.017214
1	0.000000	13	0.039381
2	0.000000	14	0.076053
3	0.000000	15	0.123375
4	0.000000	16	0.166785
5	0.000000	17	0.185680
6	0.000003	18	0.167341
7	0.000022	19	0.119063
8	0.000117	20	0.064382
9	0.000526	21	0.024867
10	0.001993	22	0.006112
11	0.006367	23	0.000718

Here we can see, for example, that the probability of exactly 15 successes is 0.123375.

The feature **Calc** ⇒ **Probability Distributions** ⇒ will allow you to deal with many different probability distributions.

## ✿ *Special Problems*

There are a number of problems commonly encountered by new users of Minitab. This list may provide some help.

I wanted to delete rows 3, 11, 19, and 21 from my spreadsheet, so I sequentially asked to delete these row numbers. However, Minitab did not remove the rows I wanted.

After each deletion, Minitab will renumber the remaining rows. If you remove row 3, the original row 11 will be moved up to row 10. Minitab will successfully remove the target rows if all are provided in a single operation. See page 12.

The data window was active. The cursor, unknown to me, was in row 14, column C3. As I started to type, I realized that the number originally in row 14, column C3 had disappeared. How do I get it back?

The Escape key will restore the original value.

I just read a text file, and then I discovered that the source used unusual phrases for missing data. The entire column was treated as **Text** and now I can't use it in numerical calculations. How do I recover this?

Click on **Data** ⇒ **Change Data Type** ⇒ **Text to Numeric...** The unusual phrases will be converted to \*, the missing data code. You will also receive a message in the session window indicating explicitly how Minitab did the conversion. The common missing data representations (\*, missing, N/A, unknown, ?) will be converted to \* automatically. Since the reading of missing data is vulnerable to confusion, the Minitab image of your file should be checked carefully.

One of the data columns shows four figures after the decimal point, but I'd really like to see some additional precision.

Click on the name both for this column. This will cause the whole column to be selected. Then **Editor** ⇒ **Format Column** ⇒ **Numeric** ⇒ will allow you to change the displayed precision.

## ***Quick Index***

This index locates facts that you might not easily find from the CONTENTS on page 1.

✱ defined, 2  
⇒ defined, 2  
> described, 5  
Alt+PrtScr, 9  
ASCII, 4  
Bitmap, 8  
Calculator, 12-13  
Change data type, 15  
Copy graph, 9  
Copy worksheet, 5  
Data window, 5, 7, 9, 11, 12, 13, 14, 15  
dBase, 4  
Deleting rows, 15  
Descriptions, 3, 4, 5  
Editor, 5  
Edit symbols, 10  
European format, 7  
Excel, 2, 3, 4  
Enable commands, 5  
Erase all work, 6  
Field definition, 8  
Files of Type, 7, 8  
Format column, 15  
Free format, 7  
Functions, 12  
Graph, 9  
Graph annotation, 10  
History window, 6  
Info window, 6, 11  
Logical operations, 12  
Lotus 1-2-3, 4  
Make Patterned Data, 14  
Minitab portable, 4  
Missing data, 2, 7, 11, 12, 13, 15  
Notepad<sup>®</sup>, 7  
Open Project, 7  
Open Worksheet, 2, 3, 8  
Options panel for file read, 8  
Preview (for file read), 8  
Probability Distribution, 14-15  
Projects, 3-6, 7, 8, 11  
Project descriptions, 3, 4  
Project manager, 4, 6, 11  
Quattro Pro, 4  
Rename, 5  
ReportPad, 3, 6, 8, 9  
Residuals, 11  
Residuals versus fits, 11  
Save Project, 3  
Save Project As, 3  
Scatterplot, 9  
Session window, 3, 5, 6, 11, 13, 14, 15  
Set Description, 5  
Show graphs folder, 8  
Show history, 6  
Show info, 6, 11  
Show worksheets folder, 4, 5, 6  
Student Minitab, 2  
Symbol (on graph), 10  
Variable names, 8  
Web address, 4  
Wiping Minitab clean, 6  
WORD<sup>®</sup>, 7  
WordPad<sup>®</sup>, 7  
Worksheet (defined), 2  
Worksheet descriptions, 3