MAS131/231: Introduction to Probability and Statistics
Practical 1: Introduction to Minitab 14

This document assumes that you have registered to use the UCS Windows XP PC clusters, and that you have worked through the induction booklet *Introduction to Oracle: The PC cluster in Merz Court*. If you haven't yet done this, make sure you have before the time of your practical session. This document does not assume experience of any particular software package, or any experience of programming.

Getting started

Before you can use Minitab, you must log into the Windows XP system. Sit down at a workstation and make sure the monitor is turned on. Press `Ctrl-Alt Delete` to bring up the logon box. Click with the mouse on the box marked `User name` and enter your user name. This is a bit like your Student ID — but not exactly. In particular, your user name will begin with a letter, whereas your Student ID will begin with a digit. Next click on the box marked `Password`, and enter your password. Next click on the *button* marked `OK`, in order to log into the network. You should end up by being presented with your Windows desktop, ready for use.

> *During the induction session you should have permanently "mapped" drive* `Q:`*, the course data directory for Maths and Stats students. If you didn't do this, you will need to map it now yourself, but don't attempt to do so if it is already mapped! To map the drive, first double-click on the* `My Documents` *icon on the desktop. Once Explorer has loaded, select* `Tools` → `Map Network Drive`*, then enter* `Q:` *for the drive, and* `\\campus\software\ mathematics and statistics\course data` *(type it carefully!) for the folder. Make sure that* `Reconnect at logon` *is ticked before clicking on* `Finish`*.*

Next load Minitab, by selecting `Start` → `Programs` → `Statistical Software` → `Minitab 14` → `Minitab 14`. This should load the Minitab application, which may take a few seconds.

About Minitab

Mintab is an easy-to-use statistical package which can carry out a wide variety of statistical tasks. Like the computer algebra package, Maple, you will use Minitab in many different courses during your time as an undergraduate, for many different purposes, so it is worth putting a little effort into familiarising yourself with the basics at an early stage. That way, you will be able to adapt to more sophisticated uses of the package, later in your course. There are three basic kinds of object that Minitab works with.

Data column: This is the most frequently used object type in Minitab. Columns are denoted `C1`, `C2`, `C3`, *etc*. They each store a collection of observations relating to a particular variable.

Constants: These are denoted `K1`, `K2`, `K3`, *etc*., and each store a single number of interest.

Data arrays: These are matrices of numbers, and are denoted by `M1`, `M2`, `M3`, *etc*.

All of the data columns, constants and data arrays relating to a particular problem are stored in a working environment called a *worksheet*. Worksheets can be saved to and loaded from disk, for later use. Several worksheets may be opened simultaneously, and these can be saved together as a *project* — see the section on saving and retrieving worksheets for further information.

There are two different ways to use Minitab. The simplest way is to select operations from Minitab's menus using the mouse. The other way is to type commands directly into Minitab via the `Session`

`Window`. In this module, we will mainly be concerned with the first method. However, when you select an operation using the menus, this is automatically translated into a Minitab command, which is displayed in the session window (if commands are enabled — see later) before being executed — examining the command displayed in the session window is a good way to learn Minitab commands. Minitab commands generally accomplish one of three things. They manipulate a worksheet, they perform an analysis, the results of which are displayed in the session window, or they produce a plot. We will see examples of each of these in this practical.

Data entry and loading a worksheet

When Minitab is first started, there are two main Mintab windows. The session window was described briefly in the previous section. The other main window is the `Data` or `Worksheet Window`. This allows you to view and edit the data columns of the current worksheet. Data can be entered directly into this window. For example, if you had 20 observations each on two variables, the observations on the first variable would go into the column marked `C1`, and observations on the second variable would go into the column marked `C2`. The first observation should go in the row marked `1`, *etc.* The space at the top of each column (directly below the column number) is provided for you to name the columns you use. You should get into the habit of always naming your columns, as it is easy to forget what each represents. After entering data, always check each entry carefully, as it is easy to make mistakes.

For this practical, we will use the questionnaire data you handed in at last week's lecture. This can be loaded by selecting `File → Open Worksheet...`, then selecting the `course data` drive (drive `Q:`), then directory `131`. From this directory, select the file `Prac1`, and then click on `OK` twice. If drive `Q:` is not available, you haven't mapped it yet — re-read **Getting started**. By looking at your data window, you should now see the questionnaire data in the first few columns of the worksheet. See if you can find the row of data corresponding to you.

Using the menus and commands

To tabulate discrete data, select `Stat → Tables → Tally Individual Variables...` from the Minitab menus, and then select the columns you wish to tabulate. For example, click on `C1`, then the button marked `Select`, then `C3`, then the button marked `Select`. Now click on `OK`. The tally results should be displayed in the session window.

Often, we will want to produce a new column of data, using existing columns. We will do this using the command interface. First, commands need to be enabled. To do this, first select the session window by clicking on it, and then select `Editor → Enable commands`. A `MTB >` prompt should appear in the session window. Enter the following session commands at the prompt. Be sure to enter your commands only at the very end of the session window, at a `MTB >` prompt, otherwise your command will be ignored — *you cannot go back and edit previous commands* and re-execute them, like you can in Maple.

```
MTB > let c21=c5*c5
MTB > let c22=c5**3
MTB > let c23=loge(c5)
MTB > let c24=sqrt(c5)
```

These commands produce some new columns representing transformations d^2, d^3, $\log(d)$ and $\sqrt{d}$ of the distance variable, d. Name your new columns by typing something appropriate at the top of each column (*eg.* `dist2`, `dist3`, `logdist`, `sqrtdist`).

Simple data summaries

To obtain some useful summary statistics, select `Stat` → `Basic Statistics` → `Display Descriptive Statistics`, then double-click on `C10, C11, C12` in order to select them before clicking on `OK`. This will give the number of observations, mean, median, trimmed mean, standard deviation, standard error of mean, minimum, maximum, lower and upper quartiles for columns 10, 11 and 12. Sometimes we would like to know how those statistics vary according to the level of some other categorical variable. For example, try calling up the descriptive statistics dialog box again, this time selecting just `C10`, clicking on the box marked `By variables`, and entering `C1` beside it. This produces summary statistics for male height and female height separately.

Simple plots

Stem-and-leaf plot

A stem-and-leaf plot for a variable can be produced by selecting `Graph` → `Stem & Leaf...`, then selecting the required variable (*eg.* `C10`), and then clicking on `OK`. This produces a stem-and-leaf plot in the session window. To produce separate plots for males and females re-do the above, but click in the box marked `By variable`, and then insert `c1` in the box to the right before clicking on `OK`. Selecting `Edit` → `Edit Last Dialog` (or pressing `Ctrl-E`) is useful when "re-doing" plots this way.

Scatter plot

A scatter plot can be produced by selecting `Graph` → `Scatterplot` → `OK`, then selecting the required y variable (say `C11`) and the required x variable (say `C10`). Before clicking on `OK`, you may select `Labels...`, and enter a title (say, *John Smith's scatter plot of foot size against height*). Clicking on `OK` twice should produce the plot.

If you want different symbols corresponding to different categorical groups, this can be achieved by selecting `With Groups`, then selecting the box marked `Categorial Variables...`, and entering (say) `C1`, before the final click on `OK`.

Plots such as these appear in their own window, and thus are not included in the session window. If you wish to save plots, these may be saved individually, or as part of a project (see below). To save a plot, select it and then `File` → `Save Graph As...`, and give an appropriate filename (after making sure an appropriate drive and directory has been selected — see later). To print a plot, select it (by clicking on it), and then `File` → `Print Graph` → `OK` (note that you are charged for printing).

Histogram

To produce a histogram, select `Graph` → `Histogram` → `OK` and the variable to be plotted (say `C10`). Titles, printing and saving are as above.

If we wanted to compare the distribution of height of males and females, we could form frequency polygons of the height distributions and overlay them. Select `Graph` → `Histogram` and then choose `With Outline and Groups` → `OK` . Then choose `C10` and select `C1` for the `Categorical Group Variables`. As ever, this basic graph can be customised in a variety of ways. In fact, Histograms have changed a lot in Minitab 14 from Minitab 13, and I haven't figured out how to do everything with them yet. You might want to experiment to see if you can get better frequency polygons, change the number of bins, select custom cut-points etc., but your practical group leaders may not be able to help you!

Box and whisker plots

Select Graph → Boxplot → OK, and the variable to be plotted (C10 say). A categorical grouping variable (C1 say) can be handled as for histograms. N.B. Minitab uses a slightly different definition of whiskers to that given in the lectures — the whiskers are limited in length, and extreme observations are individually plotted (though the behaviour can be changed — you might like to experiment with this).

Bar charts

Select Graph → Bar Chart → OK, and then select the variable to plot (C4 say).

Help facilities

Often while using Minitab you will not know, or may have forgotten, how to accomplish certain tasks. Fortunately Minitab has excellent "on-line" documentation that will allow you to find out how to achieve them.

In order to access the Help system, select Help → Help. This brings up the main Help window. Click on Basics → Introduction to Minitab and select something that takes your interest — for example, Data Windows. You can click on the button marked Back to go back. When you are finished using the Help system, close the window by clicking on the small cross in the top right-hand corner of it — *not* on the main Minitab window! If you have a specific topic or keyword in mind, click on the search tab in the Minitab Help window, and type (for example) graphs in the search box, and then select from the different possibilities. You might like to look up how to customise histograms, as this is something you are likely to want to know how to do.

Saving and retrieving worksheets and projects

When you have been working with Minitab, you will usually want to save the contents of your worksheet for future use. To save a worksheet, first click on it in order to make it active, and then select File → Save Current Worksheet As. Make sure that your current drive is H: (which appears as your user name), and give an appropriate name for the file before clicking on OK. On the Windows XP clusters, drive H: is synonymous with My Documents, so you may save your work in My Documents if you prefer — it makes no difference. Note that under Windows XP, file names may be longer than 8 letters, may be of mixed case, and may contains spaces and hyphens. You should try to avoid using other symbols, especially those which have special meaning (such as \) and are not permitted. Note that saving a worksheet only saves the worksheet contents. It does not save any plots you have produced, or the contents of the session window. To save your complete workspace, including the session window, all open worksheets, and any plots, select File → Save Project As... and select an appropriate directory and file name. This can be re-loaded at a later date by selecting File → Open Project..., or by clicking on the small yellow "open file" icon on the Minitab tool-bar. Projects are often more convenient than worksheets for a "project" you are currently working on. However, they are less useful for long term data storage, as the project files tend to be very large, and so you may eventually run out of disc storage space. Worksheets must be re-loaded by selecting File → Open Worksheet — they can't be opened by clicking on the open file icon on the toolbar.

Printing and copying files and plots

You can print the worksheet, the session window or any plots direct from your PC onto the laser printers in the same room. Remember that you are charged for printing, and that you will have to "top up" your printer funds at the UCS reception when your remaining funds get low. To print a window, click on it,

and then select File → Print ??? (where ??? will depend on what you are printing) and click on OK. Take care with printing — you don't want to print out a huge worksheet by mistake, when you actually just wanted to print out a one-page plot.

In fact, you will usually want to "Copy-and-Paste" a plot, or some session window output into a Microsoft Word document, and then print the Word document when it is complete. In Word, it is easy to format output and re-size plots for more efficient use of paper. For example, it is easy to copy several plots into a Word document, and arrange them so that they all print together on a single page.

Try this now by running Microsoft Word from the Start menu (Start → Programs → Office XP → Microsoft Word), without quitting Minitab. Then switch back to Minitab (click on the Minitab button on the task bar), select a Graph Window, and then choose Edit → Copy Graph. Switch back to Word (the Word button on the task bar) and Edit → Paste the graph into your Word document. By double-clicking on the graph you can resize it and edit it in other ways (such as adding a title).

You can also copy bits of the Session Window into Word. Scroll back through your Session Window output to find a Stem-and-Leaf plot. Click and hold the mouse button at the top of the plot, and then holding down the button, drag the pointer down to the bottom of the plot, highlighting it in the process. Select Edit → Copy in Minitab then Edit → Paste in Word as before. If you are pasting a graph into a Word document containing lots of text, you may wish to control the way text wraps around it. For example, selecting the plot in your Word document and choosing Format → Object... → Layout → Square → Left → OK is one possibility.

Exiting and logging out

When you are finished working with Minitab, it is important that you exit Minitab and then log out of the Windows XP network properly; this will ensure that all of your work is saved properly and that your files are not corrupted. To exit Mintab, select File → Exit. You will have an opportunity to save your project — you may wish to save this to your drive H: (or My Documents), which is your own personal file space, and will be available to you whichever machine you use, on any Windows XP cluster — not just the Oracle cluster in M418. It is recommended that you only use Windows XP clusters, and avoid other clusters (such as Windows NT or Windows 2000 clusters), as these behave differently, and are to be phased out.

Once you have exited Minitab, you can log out of the cluster by selecting Start → Log off ?????? (where ?????? is your user name), and then clicking on Yes. Don't do this until the end of your practical session!

And finally...

If you wish, you can find out more about Minitab, and MAS131 more generally, using the course web page. Load *Internet Explorer* (which should be on your desktop) and call up the School home page (the address is www.mas.ncl.ac.uk, but you should have added it to your "Favourites" during the induction session). You can get to the page for this module from the School home page by selecting Additional teaching information → Dr Wilkinson → MAS131. Add this page to your "Favourites", too. This page contains lots of information relevant to this course, so you should familiarise yourself with its contents, and check it regularly. If you are still unsure about using the web, or about how to read the module newsgroup, ncl.mas.131, now would be a good time to try it. Ask one of the practical session leaders if you need help.

Once you feel ready, you should attempt Exercise 1.