

QUEEN MARY, UNIVERSITY OF LONDON

MTH 4106

Introduction to Statistics

Practical 3

24 January 2012

If you didn't do so last week, please complete Practicals 1 and 2 before going on to do the following.

Today we will see how to use Minitab to plot and summarize one or two variables, quantitative or qualitative. This material will not be covered in lectures, so you are advised to save all the displays in a report pad, adding any comments that you find helpful, and then keep a printed copy of this report in your notes. You will be asked to do some of these in a later assignment.

Warning: open graphs take up a lot of working memory, because they retain a lot of information so that you can edit them. If you are satisfied with a graph, paste it into your report pad and delete the original graph. It takes up much less working memory in the report pad. If you have too many graphs open, Minitab may be unable to save your project.

1 (A new data set) Enter Minitab and then load the worksheet `Pulse.MTW` from the Minitab `Sample Data` folder.

This data set is supplied with Minitab. There is some information about the variables in the `Help`. Use

`Help` → `Help`

then double-click on `Data sets` then on `PULSE.MTW`. Read the description of the experiment and the variables. Make notes on which are quantitative and which are qualitative.

2 (One quantitative variable) We have previously used Minitab to draw box-plots and stem-and-leaf plots. Another possibility is a histogram. You can get the histogram for `Weight` by using

`Graph` → `Histogram...`,

then choosing `Simple`, clicking on `OK`, selecting the variable `Weight`, and clicking on `OK` again.

How does this histogram differ from those in lectures?

Another suitable plot is the dotplot. Use the `Graph` menu to draw a dotplot of the variable `Height`.

3 (One qualitative variable) The variable `Activity` is qualitative, with only three possible values. From the `Help`, make a note of what these values represent. (I think that this variable is ordinal, but Minitab uses the word ‘categorical’ for all qualitative variables.)

Such a variable is often displayed on a pie chart. Do this by using

`Graph` → `Pie Chart...`

and then entering `Activity` in the `Categorical variables` box. Click on `OK`.

Note that there is one value of `Activity` equal to 0, although the definition gave values of 1, 2 and 3 only!

Another suitable display for qualitative variables is a bar chart. Draw a bar chart for `Activity`. Do you find this more or less clear than the pie chart? How does a bar chart differ from a histogram?

4 (Too many graphs?) How many graphs are open in your Minitab session? Count those that are minimized and any that are hidden behind the **Session window** and **Data window** as well as those that are obvious. Now might be a good time to append those graphs to a report pad, save the report pad and close the graphs.

5 (Two quantitative variables) Now we examine the relationship between the variables **Height** and **Weight**. These are both quantitative variables, measured on the same set of people, so it is appropriate to show them on a scatterplot. Do this, with **Weight** as the Y variable and **Height** as the X variable. Adding a title to your graph by clicking on the box **Labels**.

Look at the scatterplot. To one decimal place, what do you think the sample correlation coefficient is?

Use

Stat → Basic Statistics ► Correlation

to find the sample correlation coefficient of **Height** and **Weight**. Is it close to what you guessed?

When two quantitative variables are shown on a scatterplot, we can show a plot for each variable separately along the relevant axis. These extra plots are known as *marginal plots*. Do a scatterplot of **Weight** versus **Height** with boxplots in the margins by using

Graph → Marginal Plot With Boxplots

and then filling in the dialogue box appropriately.

Boxplots are not the only way of displaying the marginal plots. Find two other possibilities, and make a note of them.

Use Minitab to draw both of these other types of marginal plot for the variables `Pulse1` and `Pulse2`, putting these on the `X` and `Y` axes respectively.

6 (One quantitative variable and one qualitative variable) Calculate a new quantitative variable `Diff` as the difference `Pulse2 - Pulse1`. We want to see if the qualitative variable `Ran` has any effect on the variable `Diff`. (Why?—read the description of the experiment again.)

We can draw two boxplots of `Diff` on the same scale, one for each value of `Ran`, by choosing `One Y, With Groups` in the dialogue box for boxplots and then entering `Ran` as the `categorical variable`. Can you explain why the two boxplots look so different?

7 (Two qualitative variables) The pie chart that you drew of `Activity` gives no information about whether levels of activity are different for men and women. As these are both qualitative variables, the best way to display the information is in a table. Use

`Stat` → `Tables` ► `Cross Tabulation and Chi-Square . . .`

Enter `Sex` for the `Rows` and `Activity` for the `Columns`. Do women have the same pattern of activity as men?

When you look at this table later, will you remember what the levels 1, 2, etc. represent? You may find it useful to edit the headings in the table. Can you do this from options in the dialogue box?

Don't forget to copy this table into your report pad and then save both the project and the report pad before you exit Minitab.