Generating Random Variates, Sampling Distributions, and Probability Plots
Wyatt Toolson
10/28/05

Generating Random Variates:
Generating random data points is easy in Minitab. There are a variety of different choices for different distributions, and you can even set specific options depending on which distribution you chose. You can record data in a column, and then perform all sorts of operations on it to pull out the statistics that you need.

To generate a column full of random data, first open up a new worksheet. Click on Calc -> Random Data.

This takes you to a list of all the different distributions that you can pick from. Some examples are Chi-Square, Normal, and Multivariate Normal. For example, let’s pick Normal. Selecting this distribution brings up a window where we can specify how many rows of data we want, the column we want to store the data in, the mean, and the standard deviation. Plug in some numbers and convince yourself that this does indeed resemble a Normal distribution. Now you can select the distributions and insert the values that your professors have asked for. After you have a list of data points, you can Display Descriptive Statistics as you did in the Introduction to Minitab handout and find out all sorts of useful information.

Sampling Distributions:
Another way to generate random data is to randomly select data points from a list. Minitab makes this very easy with just a few simple clicks. If you have a list of data points in a column, say that you randomly generated above, you can randomly select a number of points from the list, store them in a new column, and then obtain some statistical information from it.

To randomly select data points from a list, go to Calc -> Random Data -> Sample from Columns.

This brings you to a window with the option of how many points you want your random sample to be, the column to sample from, and the column to store the points in. Note that there is also a check box labeled Sample With Replacement. This option will put the data point back into the set of possibilities after it has been chosen. For example, say you have the set [1, 2, 3, 4, 5]. If you randomly pull out the number 2 with this option chosen, your set will still be [1, 2, 3, 4, 5]. However, if you pull 2 out without this option, your set will now be [1, 3, 4, 5]. Note that if you have selected this option, it is possible to generate a list of the same number as many times as you want. Make sure that you have chosen the right option for this, as it could greatly impact your results.
**Probability Plots:**
Minitab lets you make probability plots for a number of different distributions very easily. You can also pick a confidence interval or set the mean and standard deviation. In the following exercises from the book, you will make the probability plots asked for and then answer some short questions based off of what you see.

To make a Minitab probability plot, first open up the worksheet for the specified problem just like you did for the “Introduction to Minitab” handout. Make sure that commands are still enabled. Then, go to

Graphs-> Probability plot.

You then have the choice of either single or multiple. For all of the problems in this section, we will be using the simple plot. Make sure simple is selected and click ok. Now, in the box listing the possible columns, double click on the one you want to see a probability plot of. This should put the column name in the Graph Variables box.

Minitab’s default distribution is normal (Note: if you have already changed the distribution from normal, Minitab is now set to this new distribution. You will have to change it back to normal if the question asks for it.). However, if you need to change the distribution, click on the distribution button. This will probably pop up another one of those annoying Run-Time Error messages, but just click continue. From the Distribution pull down menu, select the distribution you want to try. Some of the options are Normal, Lognormal, Gamma, Exponential, and Weibull. You also see two boxes where you can enter a mean and standard deviation if you wish to. If you click on the Data Display tab, you can select whether you want to see a confidence interval or not, and what you want that confidence interval to be. Here we see the same graph, the first without confidence intervals, and the second with.

Displaying the confidence intervals makes it much easier to tell how good of an approximation a probability plot is. I suggest keeping them on unless the question specifically asks for them not to be there.

Once you’ve set the distribution you want to try, click ok. This takes you back to the Probability Plot – Single window. Click ok. This brings up a number of Run-Time Error
messages (I counted 6), but again, just click continue. This should bring up your beautiful probability plot, in all its glory.

Now that you know how to do these operations, perform the following tasks:

1. Generate a column of 1000 rows of U(0,10) values [store in column c1]
2. Plot a histogram of column c1. What do you notice about the shape of the histogram?
3. Find the sample mean and standard deviation of column c1.
4. Generate 10 more columns of 1000 rows of U(0,10) values [store in columns c3-c12]
5. Store in column c14 the row means for columns c3-c12 [using Calc -> Row Statistics]. (That is, column c14 will have 1000 rows. The entry in row i is the mean of the ten numbers stored in row i of columns c3-c12).
7. Find the sample mean and standard deviation of column c14. How do they compare to the sample mean and standard deviation of column c1 that you found in (3)?

You will need to hand in your responses to (2), (3), (6) and (7). In addition, you will need to have these in hand during class, as we will be discussing them.