CHAPTER **3**

Appendix

Simple Random Samples with Excel, JMP, Minitab, SPSS, CrunchIt!, R, and TI-83/-84 Calculators

Simple Random Samples

EXCEL

1. Assign a numerical label in the form of an integer to each population element. For example, as suggested with Example 3.8 in the text, we would assign the numbers 1–100 to the 100 population elements (the top 100 brands). If there is a data file with actual names to be sampled, you can open that and skip to step 3.
2. To put the numeric labels in a column (say A), place the cursor in the first cell of a blank column. Type a 1 in that cell. In the cell below, type "=a1+1." Pressing the enter key cell a2 should read as 2. Move the cursor back to cell A2 and highlight it. Now, grab the square at the lower right and drag the formula down to cell A100.
3. Click **Data ➔ Data Analysis ➔ Sampling ➔ OK**.
4. Enter the cell range of the numerical labels into the "Input Range" box.
5. Choose the "Random" option and input the number of random selections you wish Excel to make.

   Note: Repeated selections are possible, so it is advisable that you generate several more random numbers than the desired sample size.

6. Click **OK** and find the random numbers outputted to a new worksheet.

   Note: If you wish to sort the random numbers, you should first highlight the cells to be sorted and then click **Data ➔ Sort**. You can now pick the option to have the numbers sorted from smallest to largest. This makes locating any duplicates easy.

JMP

1. Assign a numerical label in the form of an integer to each population element. For example, as suggested with Example 3.8 in the text, we would assign the numbers 1–100 to the 100 population elements (the top 100 brands). If there is a data file with actual names to be sampled, you can open that and skip to step 3.
2. Click the red arrow next to the column headings. Select **New Column**. In the dialog box select the initialize method of **Sequential**. This should default to 1 as the lowest value. Enter the highest value (in this example, 100) and click **OK**.
3. Click **Tables ➔ Subset**.

4. Select "Random—sampling rate" or "Random—sample size" using the radio buttons. Enter the sampling proportion or size.
5. Under "Columns," choose to keep all columns or only the selected columns.
6. Click **OK**. The sample labels will be in a new worksheet.

MINITAB

1. If there is a data file, open that, or assign a numeric label in the form of an integer to each population element. For example, as suggested with Example 3.8 in the text, we would assign the numbers 1–100 to the 100 population elements (the top 100 brands). To enter the population labels easily, click **Calc ➜ Make Patterned Data ➜ Simple Set of Numbers**. Enter the name of a column to store the result (like C1). Enter 1 in the "From first value" box and N (the population size) in the "To last value" box. Click **OK**.
2. Click **Calc ➜ Random Data ➜ Sample from Columns**.
3. Enter the desired sample size in the "Number of rows to sample" box.
4. Enter the column name to be sampled in the "From Columns" box.
5. In "Store samples in" type in a new column or select an existing column to replace.
6. Click **OK**.

SPSS

1. Assign a numerical label in the form of an integer to each population element. For example, as suggested with Example 3.8 in the text, we would assign the numbers 1–100 to the 100 population elements (the top 100 brands). If there is a data file with actual names to be sampled, you can open that and skip to step 3.
2. Either enter all the number labels in a column, or use Excel to create this. There is no easy way to enter sequential values 1, 2, …, n in SPSS.
3. Click **Data ➜ Select Cases**.
4. Under "Select," pick "Random sample of cases."
5. Click "Sample" and specify percent or number of cases to sample, then click **Continue**.
6. Under "Output," select the desired option (usually "Copy selected cases to a new dataset"). If you are creating a new data set, enter a name.
7. Click **OK**.

CRUNCHIT!

CrunchIt! does not formally support random sampling, but a random sample (without replacement) can be obtained by creating a column of random values to allow the selection of only those individuals with labels matching a particular value.

1. Click **Insert ➜ Random Numbers ➜ Binomial**.
2. Enter 1 for n (this means values will be between 0 and 1).
3. For p, enter the proportion of data you wish to sample, for instance, 0.50.
4. For samples, enter the number of rows available or a value smaller if you wish to sample from only the first number of rows specified.
5. Click **Sample**.

A new variable will be created with values of 0 and 1—to use this variable to form a selection:

6. Click **Filtering, Exclude**.
7. In the "where" section, select the newly created variable.

8. In the final section, enter "0."

9. Click **OK**.

Rows not matching this criterion will be excluded from analysis.

Another option would be to generate uniform random numbers from 1 to N (the population size) and ignore the values after the decimal point.

TI calculators do not perform random sampling, but they can generate a random list of numbers to use in selecting a sample.

TI-83/-84

1. Press STAT ENTER for the Statistics list editor. Select an empty list or clear one.

2. Press ▲ to highlight the list name. In the input area at the bottom of the screen, you should see something like L1=.

3. Press MATH ◀ to PRB. Then select option 5:randInt(.

4. Enter the minimum desired value, the maximum, and the number of values to generate. Ask for a few more than you need because duplicates may occur. A finished command should look like the one shown below (to generate five numbers between 1 and 33, inclusive).

```
L1=randInt(1,33,5)
```

Use sample() to take a random sample of size n from a data set. An example of this command is shown and selects a sample of size n = 15 from data set named "mydata" into a subset named "mysamp:"

1. Open a data file with names of the individuals to be sampled or create a data frame with the values 1, 2, …, N. The command below will generate a variable called pop with values 1, 2, …, 100.

```
>pop <- c(1:100)
```

2. Create the sample with a command like this (samples 15 entries from data frame mydata:

```
>mysamp=mydata[sample(1:nrow(mydata),15,replace=F),]
```