Stat 463, Lab 1

September 8, 2006

1 In Class

In this tutorial, we will learn some of the basics of the statistical software R. We will learn how to read data in from a text file, how to handle variables, how to plot data, and how to save these plots.

1. Directories. Create a directory within your "home directory" entitled "timeseries". Create a directory within this "timeseries" directory entitled "lab1" using the "change dir..." option under the file menu.

2. Launch R. Change the working directory to "lab1".

3. Download files. Obtain the data file, "soi.dat" from Shumway/Stoffers web page. Put this file into "lab1". We want to import this data into R. One way to do this is to use the `scan()` command; simply type

   ```
   scan("soi.dat")
   ```

   R has read the numbers and output them to the screen. We would like to store the numbers so that we can work with them. To do this, we declare a variable `soi` and set it equal to the output of the `scan()` command, i.e. the list of numbers. Type

   ```
   soi=scan("soi.dat")
   ```

4. Command line. Now, let's get familiar with using the command line:

   (a) Type `ls()`. This command lists all the variables present in your workspace. Right now there should only be `soi`.

   (b) To check the contents of a variable, simply input the name of the variable, such as `soi`.

(c) Create a variable `temp` equal to 2 using the command `temp=2`

(d) Remove the variable `temp` by typing `rm(temp)`.

(e) Now, check to make sure that it's gone by typing `ls()`.

5. Easy plotting. Making attractive plots is one of R's strong suits.

   (a) Create a simple plot of the `roi` using `plot(roi)`.

   (b) We can expand on this by changing the x and y labels and adding a title.

   ```
   plot(soi,xlab="time", ylab="Southern Oscillation Index", main="Time Series Plot")
   ```

   (c) Now, download the file "recruit.dat", and load the data into the variable `recruit`.

6. Time series plots. Since these time series are related, we would like to plot `soi` and `recruit` on the same plot. Right now we have both variables as basic vectors. First, we want to convert them to time series objects, this will allow for more attractive and easier to interpret graphs.

   (a) Both series consist of monthly data starting in the year 1950. To convert `soi`, we use the command

   ```
   soi=ts(soi,start=1950,frequency=12)
   ```

   Note that by using this command we've overwritten what used to be in the variable `soi`. Use a similar command to convert `recruit`.

   (b) Now, plot `soi`. Notice the differences from how the graph appeared before.

   (c) We want to plot both series together. First, we need to split up the plot "device". We do this with the command

   ```
   par(mfrow=c(2,1))
   ```

   This will give two rows and one column.

   (d) Now, we plot both series by typing `plot(soi)` and then entering `plot(recruit)`.

   (e) We can save this plot in any format that we'd like (pdf, postscript, jpg, etc). Save this graphic as "joint.jpg".

7. Saving your work. You can save the current variables into a file using the following command

```
save.image(‘‘lab1.Rdata’’)
```

Clear the workspace. Check to make sure that no variables are left using `ls()`. Load the previous work by typing

```
load(‘‘lab1.Rdata’’)
```

Check to make sure you were successful using `ls()`. You can also use the "save workspace..." and "load workspace..." options under the file menu in the windows version of R. In addition, R will ask you if you want to save the workspace when you quit.

2 Homework

1. Read sections 1, 2, 3, 6, 7, and 8 of An Introduction of R (do this in front of your computer), and answer the following questions.

   (a) What's the difference in output between the commands `2*1:5` and `(2*1):5`? Why is there a difference?

   (b) If you wanted to enter the odd numbers from 1 to 19 in the variable `x`, what command would you use?

   (c) If you create a variable using the following command `y=c(-1,2,-3,4,-5)`, what command would put the positive values of `y` into the variable `z`?

   (d) What R command would give you the 95th percentile for a chi-squared distribution with 10 degrees of freedom?

   (e) Generate a vector of 1000 standard normal random variables using the command `x=rnorm(100)`, use R to give a five number summary of your simulated data; what is the mean and variance of your `x` variable? Make and print a histogram for this data.

2. From Shumway and Stoffer's web page, download "gas.dat".

   (a) Import the data from "gas.dat", and make sure to convert it to time series type. (You can get a little more info about the data in exercise 2.9 on page 82.)

   (b) Make a plot of the data. Make sure to label the plot and the x and y axis appropriately. Print out the plot to turn in.