Introduction to course and to R

Leo Schalkwyk

June 2, 2011

What is R and why does it have such a silly name?

- `http://en.wikipedia.org/wiki/R` - what the encyclopedia says
- descended from S, developed at Bell Labs (home of C)
- S was developed by John Chambers, described in several books

R is an implementation of S

- so is S-Plus (a commercial product)
- the two are very similar
- we have a site license for S-plus and you could use it identically
- I'll mention the few differences between R and S-plus as we come to them

Why (not) use R

- +very powerful computing environment for data analysis
- +under active development by enthusiastic volunteers
- +lots of specialised software packages available
- +you can distribute your own code
- +superb graphics capabilities
- -hard to get started
- -documentation style requires getting used to
- -aimed at statisticians

Why (not) use S+

- +very powerful computing environment for data analysis
- +under active development by a company
- +lots of specialised software packages available
- +superb graphics capabilities
- +some help in getting started from menus
- -you can distribute your own code but others need S+ to use it
- -R packages will probably run but need to be ported

What are the characteristics of R

- command line interface similar to unix shell
  - somewhat C - like syntax
  - command-line editing
  - online help in man-page tradition
- programming environment designed for interactive use
- functional programming language similar to scheme
- object oriented features
- high and low level graphics

Installation of R

- http://www.r-project.org
- http://www.stats.bris.ac.uk/R/

Functions

- functions are the only way of making something happen
- name(args)
- easy to define your own:
  - myfu <- function (a,b) {a+b}
- args may not be required but () are always required
  - sum(1,2)
  - q() # quit R
  - q # print q function
- the basic arithmetic operators are convenience exceptions
  - '+'(1,2)
  - 1+2

Playing with R: basic arithmetic

- if you type in an expression R will evaluate and print-
- 5*2
- 5^2
- 100%%3
- 100%%3
- multiply and divide come before add and subtract

vectorise me!

- variable names start with a letter
- assign values to variable names with ¡-
- basic data structure is the vector
- a <- 1:100
- b <- c(1,0) # c is a function that combines its arguments
- all of the basic functions work on vectors
- a + 0.5
- a + a

recycling

- operations on vectors of unequal length recycle the shorter one
- a <- 1:100
- b <- c(1,0)
- a + b # add 1 to odd numbered positions only

indexing: index vectors

- a[50] # 50
- a[-50] # everything but 50
- a[100:1] # backwards
- a[c(1:100,100:1)] # count down then up
- a[1:10*10] # every tenth number

indexing: logical vectors

- b <- a%%2==1
- a[b] # odd only

Exercises:

- calculate the sum of odd numbers up to 1000
- what's the largest multiple of 13 under 1000
- calculate pi using the Gregory-Leibniz series