Solving the Unsolvable Through Scientific Computing: Explorations in the Best Uses of Popular Mathematics Software

Talitha Washington, Howard University
Edray Goins, Purdue University
Luis Melara, Shippensburg University
Alejandra Alvarado, Eastern Illinois University

16 September 2016
Tapia Conference

Abstract

Many times one is posed with the need to implement a mathematical software program to solve or investigate a phenomenon. The workshop organizers are faculty in mathematics who often use these programs to carry out research applied to biology, and engineering as well as the development of numerical schemes. This two-hour session will provide best practices on how to best choose and implement popular mathematics software such as Mathematica, MATLAB (or GNU Octave), R along with RStudio, and Sage. The format will encourage both participating hands-on (using your own laptop with pre-installed software) as well as observing and asking questions.

Outline of Talk

Computational Tools

R
www.r-project.org
The Engine

RStudio
www.rstudio.com
Th Graphical Front End

History

- In the mid 1980's, statistical software S was developed at AT&T using the language Scheme
- About 1992, Ross Ihaka and Robert Gentleman of the University of Auckland, New Zealand developed R, an implementation of S.
- R is named after them and as a successor to S
- By 1995, R became a GNU project (free software and collaboration)
- Source code: C, Fortran, and R

Why R?

- R, is a free open statistical software package that is increasingly finding use in statistical analysis and is rapidly becoming the industry standard.
- Easy to download on Windows, Macs, Unix and easy to install (each less than 5 minutes on MacBook)
- R is a language and environment for statistical computing and graphics with many built in functions.
- Extensive Help, FAQ and material.
- Two R conferences: useR! - International R User Conference and DSC - Directions in Statistical Computing
- The R Journal; A peer-reviewed, open-access publication of the R Foundation for Statistical Computing

The R environment

R is an integrated suite of software facilities for data manipulation, calculation and graphical display.

- calculations on arrays, in particular, matrices
- tools for data analysis
- graphical facilities
- programming language
- effective data handling and storage facility

RStudio Interface

Console; calculations, name variables, functions

```
> #this is a comment
> 2+2
[1] 4
> log(exp(1))
[1] 1
> x=1:10
> x
[1] 1 2 3 4 5 6 7 8 9 10
> y=c("blue","red","green","yellow")
> y
[1] "blue"   "red"    "green"  "yellow''
> test<-function(u,v){return(u-v)}
> test(1,2)
[1] -1
```

Packages

R has many classical and modern statistical techniques implemented.

Data sets and R-functions are stored in libraries of packages. Only when a package is loaded are its contents available.

This is done both for efficiency and to aid package developers, who are protected from name clashes with other code.

To see the packages in your library:

```
> library()
```

You can install a package:

```
> install.packages("hypergeo")
```

Data Sets

R has about 100 built-in data sets (in the package 'datasets')

```
> data()
```

The following is built in the package

```
> AirPassengers
```

Some data comes with a package

```
> data(package="rpart")
```

If you have your own data set, you can call it

```
> read.delim("Desktop/Folder/filename.txt")
```

New R Script

Click on 'File', then 'New File', then 'R Script'

An editor window will appear above console

Execute file: click on 'Source'

Save your entire workspace, variables, functions.

Environment data files are saved with the extension '.Rdata'

All the commands used in an R session saved in an '.Rhistory' file

Monty Hall Problem: solve by running simulation

Suppose you're on a game show, and you're given the choice of three doors: Behind one door is a car; behind the others, goats. You pick a door, say No. 1, and the host, who knows what's behind the doors, opens another door, say No. 3, which has a goat. He then says to you, "Do you want to pick door No. 2?"
Is it to your advantage to switch your choice?
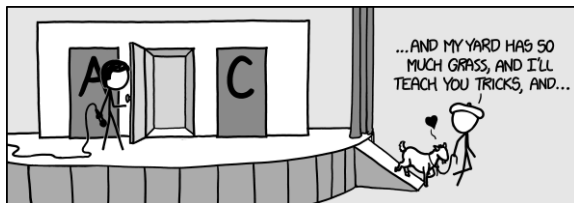
xkcd.com

References

- W. N. Venables, D. M. Smith, and the R Core Team, *An Introduction to R: A Programming Environment for Data Analysis and Graphics*, Version 3.3.1 (2016-06-21), http://cran.wustl.edu/doc/manuals/R-intro.pdf
- R Core Team, *R: A Language and Environment for Statistical Computing*, (2016).
- Hystad, Grethe, *Tutorial for the integration of the software, R, with introductory statistics*, (2014).
- Free online R editor and console, http://www.r-fiddle.org
- SAS (Statistical Analysis System) software, not free

Thank you!

Questions?