King Fahd University of Petroleum and Minerals

Department of Mathematics & Statistics

STAT 591

Special Topics in Statistics (Title: Statistical Computing)

Term 151

Objectives

The aim of the course is for students to develop a deep understanding and set of skills in modern statistical and computational methods. The course will focus on the use of R, a statistical computing language and environment for handling complex data relevant issues and graphing. The course also involves the generation of random numbers to simulate real situations and to solve problems where mathematical analysis is hard or intractable.

Rationale

The field of statistics has undergone a major change, as new and novel techniques of statistical modeling and data analysis are continually required for handling data from many different fields of life. Advanced techniques such as random number generation to simulate a real process, bootstrapping, variance reduction techniques, numerical integration, optimization and the use of Markov chains can be very helpful in handling real data. R is most widely used for teaching statistics classes at universities all over the world because students can freely use the statistical computing tools. With a wide range of built in functions, further functionality has been added with the help of more than 6000 contributed packages from many disciplines of life such as business, econometrics, finance, actuarial statistics etc.

Time Table:

Monday and Wednesday: 6:30 – 7:45 pm   Build/6, Room 101

Office Hours:

Sunday and Tuesday: 10:00 – 11:30 am   Build/5, Room 318

Or by appointment

Assessment and Grading:

- Assignments and Class Participation (35%) – Regular assignments will be given to the students
- Major Exam (30 %)
- Final Exam   (35 %)

Course Topics

In this course, we will cover the use of R statistical language for handling data relevant issues and graphing. The course involves in teaching advanced computational techniques to the students in an easy to understandable format.

A tentative list of the topics is given below:

- Introduction to R statistical language
- R basics
- R programming: functions, data structures loops, efficient computational techniques
- Numerical computation
- Optimization
- Factors and frames
- Input-output in R
- Monte Carlo Simulations
- Graphics in R
- Graphics in R at an advanced level
- Programing efficiency
- Advanced Topics
  - Bootstrapping
  - Variance reduction techniques
  - Markov Chain Monte Carlo
- Creating new R packages
  - Some useful packages
- Miscellaneous Topics
  - Regression in R
  - Time Series in R
  - Quality Control in R
  - Multivariate statistics in R
- Writing manuscripts and thesis using LaTex

Text Books and Other References:

There is no recommended text book for the course but one can get a guide from the following reference books. Class lectures/notes/manuals will be provided.

Books

1. Rizzo, M. L. (2007). Statistical Computing with R, Chapman & Hall/CRC, The R Series, USA.

2. Givens, G. H. and Hoeting, J. A. (2013). Computational Statistics, 2nd Edition, Wiley, New Jersey, USA.

3. Maindonald, J. and Braun, J. (2010). *Data Analysis and Graphics using R*, Cambridge University Press, Cambridge, UK.

4. Hothorn, T. and Everitt, B. S. (2014). *A handbook of Statistical Analyses using R*, 3rd Edition, Chapman and Hall/CRC Press, Florida, USA.

5. Stowell, S. (2014). *Using R for Statistics.* Apress.

6. Verzani, J. (2011). *Getting Started with RStudio*, O'Reilly, USA.

7. Ihaka, R. and Gentleman, R. (1996). *R: A language for data analysis and graphics*. *Journal of* Computational and Graphical Statistics, 5(3):299—314

8. Gandrud, C. (2013). *Reproducible Research with R and RStudio*. Chapman & Hall/CRC The R series. Chapman & Hall/CRC Press, Boca Raton, FL