University of Manitoba
Department of Statistics

STAT 4600: Computational Statistics

Fall Term 2015

Course Title: Computational Statistics (CRN: 15629)

Instructor: Alexandre Leblanc
Office: 324 Machray Hall
Phone: 474-6273
E-mail: alex_leblanc@umanitoba.ca

Course Schedule: Mon/Wed/Fri 11:30 am to 12:30 pm (Slot 6), in 316 Machray Hall.

Office Hours: Tuesday and Friday, from 1:30 to 3:00 pm, or by appointment.

Prerequisites:

The only formal prerequisite for this course is STAT 3800 (C), or permission of the instructor. Previous programming experience is not required, but is definitely a big asset. Familiarity with computer systems is expected.

General Goals and Outline of covered topics for this Course:

This course will be focusing on

- computational/numerical aspects of more traditional Statistics,
- some modern computer-intensive methods used in Statistics,
- learning the fundamentals of the R statistical software.

Obviously, you will need to make extensive use of the computer throughout the whole course for your homework/assignments but also, sometimes, for class.

Time permitting, the course will proceed as follows:

1. Introduction to R
2. Principles of Monte Carlo Simulation
3. The Bootstrap
4. Permutation Tests
5. (Pseudo) Random Number Generation
6. Basic Numerical Integration Methods
7. Some Methods of Statistical Inference
8. Numerical Optimization and Root Finding
9. Advanced Techniques for Random Number Generation

Statistical Package:

As mentioned above, you will need to make extensive use of the computer throughout the whole course for your homework/assignments.

Specifically, we will use the statistical software R which can be downloaded for free (for Windows or Mac, including PDF documentation) from the *Comprehensive R Archive Network* at: `http://cran.r-project.org/`

Textbook(s):

The course will be based roughly on

- Rizzo, M.L. (2008), *Statistical Computing with R*, Chapman & Hall/CRC.

which will be placed on reserve at the Science Library. However, you are not expected to buy the textbook.

The course material can be found in many other useful references that are available for free download (in PDF) through the online catalogue of the University of Manitoba Libraries. Some interesting references are:

Computational Statistics:

- Gentle, J.E. (2009), *Computational Statistics*, Springer.
- Lange, K. (2010), *Numerical Analysis for Statisticians*, Springer.

Statistics using R:

- Marin, J.-M. and Robert, C.P. (2014), *Bayesian Essentials with R*, Springer.
- Robert, C.P. and Casella, G. (2011), *Introducing Monte Carlo Methods with R*, Springer.
- Albert, J. (2007), *Bayesian Computation with R*, Springer.

Programming in R:

- Lafaye de Micheaux, P., Drouilhet, R. and Liquet, B. (2013), *The R Software: Fundamentals of Programming and Statistical Analysis*, Springer.
- Chambers, J.M. (2008), *Software for Data Analysis: Programming with R*, Springer.

Other useful online resources:

Other useful material can be found online at:

- `http://www.rstudio.com` (for downloading R Studio)
- `http://rmarkdown.rstudio.com` (to learn about R Markdown)
- `http://r-pkgs.had.co.nz/style.html`
  (a simple R style guide that is generally good to follow)
- `http://cran.r-project.org` (R software, documentation and contributed packages)

Breakdown of the Marks: Assignments (6 to 10)   40%
                                          Midterm Test   20%
                                          Final Exam   40%

About the Midterm Test and Final Exam:

The test and exam should each have a take-home part that will require the use of the computer.

The Midterm Test will be scheduled in mid-February, the exact date to be determined later. However, it will be scheduled outside of regular class time (probably in the afternoon) and should be 3 hours in duration. The Final Exam will be held on a date to be selected later by the Registrar's office and should be 4 hours in duration. Both test and exam will have a programming part using the R software and R Markdown.

About Assignments:

Assignments will be handed out at a rate of about one assignment per one or two weeks of class, for of total of up to 10 assignments. You will also be required to do some reading/homework (not to be turned in for marking).

Unless otherwise noted, all work (assignments and projects) must be turned in
1. electronically, by e-mailing the instructor,
2. in hard copy, in class.

All work must be submitted in the R Markdown format, which allows for incorporating R code within regular text, and should be written in a style that smoothly integrates prose, code, tables and graphics. All your code should be *human-readable*.

Make sure your submission properly compiles with "Knit HTML" before turning it in. Every file you submit should have your name and be properly labeled (Assignment 1, etc.)

Work submitted without following the above guidelines will not be marked and given a grade of 0.

Collaboration, Copying, and Plagiarism

Students are encouraged to discuss course material, including assignments. However, each student must hand in his or her own copy of each assignment with personalized solutions, including comments, discussions, explanations and interpretations, and R code. Copying from anywhere, including other students, books and the web,
1. does not help you to learn the course material,
2. constitutes a case of academic dishonesty and could have serious consequences.

If you are unclear on what is acceptable, please ask the instructor.

Final Note:

All course notes, assignments, tests, etc. are the intellectual property of your instructor. Reproduction or distribution of these materials is strictly forbidden without the consent of the Department of Statistics.