MATH 6359 - Statistical Computing

Fall 2017

Instructor:	Andrey Skripnikov	**Time:**	TuTh 1:00PM – 2:30PM
Email:	andreysk@math.uh.edu	**Classroom:**	SEC 201

Office Hours: Room PGH 214, TuTh 12:00PM - 12:50PM, or by appointment.

Main References: Two books that will serve as cornerstones for the course are

- Peter Dalgaard, *Introductory Statistics in R, 2nd ed.*, Springer, 2008.
- Maria L. Rizzo, *Statistical Computing with R*, Chapman & Hall/CRC, 2008

Dalgaard presents a very nice introduction to *R statistical programming* and guides you through various examples for some of the ubiquitous (and not so ubiquitous) statistical modeling tasks, all while using *R* as the main tool. Meanwhile, Rizzo emphasizes using *R* for the statistical computing techniques in particular, such as bootstrapping, Monte Carlo simulations, probability density estimation, etc.

Objectives: Course will consist of two parts: 1) getting acquainted with *R statistical programming* (introducing syntax, writing functions, making graphs, conducting basic statistical analysis), 2) outlining the most renown statistical computing methods and implementing them via *R*. The first part serves an introductory purpose and may prove useful for any student having to deal with data in his or her research (be it medicine, marketing, agricultural/biological engineering, etc), as it touches upon such popular topics in statistical analysis as t-tests, group comparisons, linear and logistic regression, time series analysis, and others. Second part takes on statistical computing techniques that may aid in data science research. In particular, methods are introduced for numerical integration and hypothesis testing (e.g. Monte Carlo approach), random variable generation, probability density estimation (e.g kernel methods), constructing confidence intervals (bootstrap, jackknife). All-in-all, at the end of the course, a successful student should be able to:

- use *R statistical programming* as a tool for conducting research and data analysis
- have a solid conceptual grasp on numerous statistical modeling approaches and respective data applications
- comprehend and implement the main statistical computing techniques for their data science research goals
- writing their own *R* functions and potentially developing new computational methods in the future

Prerequisites (more like encouragements): Being familiar with at least one programming language. An undergraduate-level understanding of probability, statistics, calculus and linear algebra is assumed.

Software: Make sure to download *R* and *RStudio* (which can't be installed without *R*) before the course starts. Use the link `https://www.rstudio.com/products/rstudio/download/` to download it from the mirror appropriate for your platform. Let me know via email in case you encounter difficulties.

Tentative Course Outline:

- R syntax (expressions, objects, functions, vectors, etc)
- R coding basics (R environment, flow control)
- Importing/exporting data, working with data frames in R
- Probability and distribution functions in R
- Graphics in R (histograms, boxplots, tables and others)
- Descriptive statistics and hypothesis testing in R (e.g. one-/two-sample and paired t-tests)
- Linear regression and ANOVA in R
- Linear models and regression diagnostics in R
- Logistic regression in R
- Time series analysis in R
- Methods for generating random variables
- Monte Carlo numerical integration
- Monte Carlo methods in statistical inference
- Bootstrap and Jackknife estimation methods
- Probability density estimation

Grading Policy: Attendance and participation (10%), Homework (30%), Project (30%), Final (30%).

Course Policy:

- Use of laptops during the class, especially the first part of the course, is encouraged solely for the purpose of practicing your *R* programming skills. Refrain from using your laptop for anything else.
- There will be bi-weekly homework. Late homework will warrant a $d \times 25\%$ penalty (subtracting $d \times 25\%$ of full possible points from your actual score), where d - number of days past due.
- Homework solutions must be typeset in a form of a report (using LaTeX or *Word*).
- Most outstanding work may be considered for demonstration to the rest of the class as an exemplary standard, upon student's permission (with implied bonus mark for the student).
- All homework solutions must be submitted in class (along with the quiz).
- You may discuss homework problems with other students, but you must write up your homework independently in your own words.
- Your lowest homework-quiz score will be dropped when calculating your final homework-quiz grade.
- The exams may or may not be take-home. If not, by default, all exams (midterms and final) are closed book, and you are not allowed to use any electronic devices such as mobiles and tablets.

Academic Honesty: Please click the link below for the full explanation of the Academic Honesty policy and procedure:

http://www.uh.edu/provost/policies/honesty/_documents-honesty/academic-honesty-policy.pdf

Definitions:

"Academic dishonesty" means employing a method or technique or engaging in conduct in an academic endeavor that contravenes the standards of ethical integrity expected at the University of Houston or by a course instructor to fulfill any and all academic requirements. Academic dishonesty includes but is not limited to, the following:

Plagiarism

1. Representing as ones own work the work of another without acknowledging the source (plagiarism). Plagiarism includes copying verbatim text from the literature, whether printed or electronic, in all assignments including field.

Cheating and Unauthorized Group Work

2. Openly cheating in an examination, as copying from anothers paper;

3. Being able to view during an examination, quiz or any in-class assignment an electronic device that allows communication with another person, access to unauthorized material, access to the internet, or the ability to capture an image, unless expressly permitted by the instructor;

4. Using and/or possessing crib notes, as unauthorized use of notes or the like to aid in answering questions during an examination;

5. Giving or receiving unauthorized aid during an examination, such as trading examinations, whispering answers, and passing notes, and using electronic devices to transmit or receive information;

6. Securing another to take a test in the students place. Both the student taking the test for another and the student registered in the course are at fault;

Fabrication, Falsification, and Misrepresentation

7. Changing answers or grades on a test that has been returned to a student in an attempt to claim instructor error;

8. Using anothers laboratory results as ones own, whether with or without the permission of the owner;

9. Falsifying results in laboratory experiments;

10. Misrepresenting academic records or achievements as they pertain to course prerequisites or corequisites for the purpose of enrolling or remaining in a course for which one is not eligible;

11. Representing oneself as a person who has earned a degree without having earned that particular degree

Stealing and Abuse of Academic Materials

12. Stealing, as theft of tests or grade books, from faculty offices or elsewhere, or knowingly using stolen tests or materials in satisfaction of exams, papers, or other assignments; this includes the removal of items posted for use by the students;

13. Mutilating or stealing library materimaterials; misshelving materials with the intent to reduce accessibility to other students;

Complicity in Academic Dishonesty

14. Failing to report to the instructor or departmental hearing officer an incident which the student believes to be a violation of the academic honesty policy;

Academic Misconduct

15. Any other conduct which a reasonable person in the same or similar circumstances would recognize as dishonest or improper in an academic setting.

Process:

Students shall have the responsibility of reporting incidents of alleged academic dishonesty to the instructor of record involved or to the appropriate authority if the alleged act is not associated with a specific class within 5 class days of the incident. Faculty or instructor of record shall have the responsibility of reporting incidents of alleged academic dishonesty through their college hearing officer within 5 class days of the incident. The faculty should include the recommended sanction in the report. The college hearing officer will notify the student of the report and recommended sanction. The student can accept the sanction and waive a hearing or request a college hearing. A hearing shall be set within 10 days and would be consist of two faculty and three students chosen by the hearing officer.

UH CAPS Statement: Counseling and Psychological Services (CAPS) can help students who are having difficulties managing stress, adjusting to the demands of a professional program, or feeling sad and hopeless. You can reach CAPS (www.uh.edu/caps) by calling 713-743-5454 during and after business hours for routine appointments or if you or someone you know is in crisis. No appointment is necessary for the Let's Talk program, a drop-in consultation service at convenient locations and hours around campus. http://www.uh.edu/caps/outreach/lets_talk.html