STT802/EPI 853B: Statistical Computing

Instructors:
Hyokyoung Grace Hong, Room #C435 Wells Hall (hhong@stt.msu.edu) &
Gustavo de los Campos, # 1311, IQ-Building, 775 Woodlot Dr. (gustavoc@msu.edu)

Time/Place: M/W 3:00-4:20PM, E109 Fee Hall

Prerequisites: Mathematical statistics (STT 442 or equivalent) and Linear Algebra (MTH 309 or equivalent).

Grading: Class Participation (10%), HW (30%), Midterm (30%), Final (30%)

<table>
<thead>
<tr>
<th>Score (5)</th>
<th>&lt;50</th>
<th>50-59</th>
<th>60-69</th>
<th>70-74</th>
<th>75-79</th>
<th>80-84</th>
<th>85-89</th>
<th>&gt;90</th>
</tr>
</thead>
<tbody>
<tr>
<td>Grade</td>
<td>0</td>
<td>1</td>
<td>1.5</td>
<td>2</td>
<td>2.5</td>
<td>3</td>
<td>3.5</td>
<td>4</td>
</tr>
</tbody>
</table>

Course Description:
In this course we will cover computational methods commonly used in statistics, including algorithms used for fitting linear and non-linear regressions via ordinary least squares and maximum likelihood, simulation of random variables, Monte Carlo and resampling methods.

Topics:

Module 1 (Introduction)
  a) Introduction to R
  b) Data preparation and descriptive analyses

Module 2 (Statistical models)
  a) Linear regression models
  b) Generalized linear models
  c) Survival models
  d) Quantile regression models

Module 3 (High-dimensional data analysis)
  a) Regularized regression methods
  b) Variable screening methods
Module 4 (Maximum likelihood)
   a) Maximization using general purpose optimization algorithms
   b) The EM-Algorithm

Module 5 (Monte Carlo Methods)
   c) Sampling Random Variables
   d) Power Analysis

Module 6 (Resampling methods)
   a) Bootstrap (application: SEs for odds ratios in logistic regression)
   b) Permutation test (p-values for odds ratios in logistic regression)

**Software:** The course will be mostly based on R.

**Approach:** Although the focus of the course is on computational methods, for each topic we will first describe the problem from a statistical perspective. Derivations will be presented in class and students are expected to take their own notes. Scripts for computations will be developed in class and a summary will be posted online. Students are expected to bring their own laptops. If you do not have access to a laptop, please check with the instructor to get access to one.

**Evaluation:** The evaluation will be based on HW (approximately 1 per module) and two in-class exams.

**Textbook:** There is no required textbook. We do not have a required textbook. The following are very good textbooks that will guide you learning about statistical analyses in R.

- An Introduction to Statistical Learning with Applications in R, by J. Gareth, D. Witten, T. Hastie and R. Tibshirani.

**Academic Honesty:** The Departments of Epidemiology & Biostatistics and Statistics & Probability adheres to the policies of academic honesty as specified in the General Student Regulations 1.0, Protection of Scholarships and Grades, and in the All-University of Integrity of scholarship and Grades which are included in Spartan Life: Student Handbook and Resource Guide. Students who plagiarize will receive a grade 0.0 on the homework, exam or quiz.

**ADA:** To arrange for accommodation a student should contact the Resource Center for People with Disabilities at http://www.rcpd.msu.edu/ or (517)353-9642
Disclaimer: Changes on the syllabus/important dates will be announced in class and on the course web site. It is students’ responsibility to keep up with any changed policies and assignments.