Fall Semester 2013
Biostatistics (BST) 695 - Statistical Computing with R

Instructor:
Nianjun Liu, PhD
Office: 420A Ryals Public Health Building
Phone: 975-9190
Email: nliu@uab.edu

Meetings:
Wednesday 9:05am-10:50am
308 Ryals Public Health Building

Office Hours:
420A Ryals Public Health Building
(By appointment if extra office hours are needed)

Credits:
2 credit hours

Introduction:
In the past decade, computing becomes more and more important in modern statistics. R, an open source language and environment for statistical computing and graphics, gains increasing popularity in recent years. Knowledge of statistical computing and R is becoming increasing important in statistics. This course will introduce R and how to perform basic statistical computing using R to the M.S. or Ph.D. students in Biostatistics. The course requires mathematical and statistical skills including calculus, linear algebra, probability theory, and basic inference, in addition to some basic computer knowledge.

Prerequisites:
BST 621, BST 622, BST 626, BST 631 and 632 (Introductory Probability and Inference) or equivalent.
Appropriate background in calculus and matrix algebra.

Course Description:
This course is mainly focused on R and how to use R to conduct basic statistical computing. The course contains three themes: R programming, introduction to high performance computing, and basics of statistical computing.

Objectives and Competencies:

<table>
<thead>
<tr>
<th>Learning Objectives</th>
<th>Competencies</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Introduction to R programming, especially how to use R to access, manipulate, and manage datasets.</td>
<td></td>
</tr>
<tr>
<td>2. Utilize common computer programs to aid in analysis, description, and presentation of statistical data and results</td>
<td>BST 8, BST 28</td>
</tr>
<tr>
<td>3. Define and distinguish between numerical algorithms useful in biostatistics</td>
<td>BST 41</td>
</tr>
<tr>
<td>4. Understand and implement numerical optimization and integration methods</td>
<td></td>
</tr>
</tbody>
</table>
Evaluation:
Students will be evaluated during the course through homework and a take-home final exam. Homework problems will be handed in and graded. Students are encouraged to work together on homework assignments to understand the concepts in the problems; however, it is expected that each student will turn in an assignment that reflects their own independent work. Exam will be completed independently, with no collaboration allowed. Assignments are expected to be turned in on time. Five points will be deducted each day for late homework, unless there are extenuating circumstances. There may be rare exceptions made to this policy on an individual basis, provided that this is worked out in advance with the instructor.

Grading Scale:
Assigned work will be weighted as follows:
- Homework: 50%
- Final Exam: 50%

Final grades will be assigned as follows:
- ≥90%: A
- [80%, 90%): B
- [70%, 80%): C
- <70%: F

Suggested Texts:

Other References:

Online Resources:
1. R website
Special Instructions:
Students are not required to attend every meeting, as long as they finish the homework assignments and exam as required.

Lecture Schedule (tentative):

Lecture 1
Course Overview
Introduction to the R/S
(B&M Chapters 1 & 2, ZIM Chapter 1, Teetor Chapters 1-3)

Lecture 2
Introduction to the R Language
(Teetor Chapters 4-7, B&M Chapters 2-4, ZIM Chapters 2, 3, 6, 9, Paradis Chapter 3, Rizzo Appendix B, Adler Chapters 9 & 12)

Lecture 3
Introduction to the R Language, continued
(Teetor Chapters 4-7, B&M Chapters 2-4, ZIM Chapters 2, 3, 6, 9, Paradis Chapter 3, Rizzo Appendix B, Adler Chapters 9 & 12)

Lecture 4
Introduction to the R Language, continued
(Teetor Chapters 4-7, B&M Chapters 2-4, ZIM Chapters 2, 3, 6, 9, Paradis Chapter 3, Rizzo Appendix B, Adler Chapters 9 & 12)

Lecture 5
Graphics with R
(Teetor Chapter 10, ZIM Chapters 5, 7, 8, Paradis Chapter 4, B&M Chapter 3, Adler Chapters 14-15, Murrell)

Lecture 6
Statistical Analysis with R
(Teetor Chapters 9 & 11, Paradis Chapter 5, Adler Chapters 17-21)

Lecture 7
R Functions and Packages
(Gentleman Chapters 6, 7, 9, Paradis Chapter 6, B&M Chapter 4, ZIM Chapter 6, JMR Chapters 5 & 8)

Lecture 8
Basics of UNIX/LINUX

Lecture 9
Cluster Environment and Grid Computing

Lecture 10
Computational Linear Algebra
(B&M Chapter 6)

Lecture 11
Computational Linear Algebra (continued) and Numerical
Optimization
(B&M Chapter 6 & 7, JMR Chapters 10 & 12)

Lecture 12
Numerical Optimization, continued
(B&M Chapter 7, JMR Chapter 12)

Lecture 13
Numerical Optimization (continued) and Numerical Integration
(B&M Chapter 7, JMR Chapter 12)

Lecture 14
Numerical Integration
(B&M Chapter 7, JMR Chapter 11)
Take-home Final Exam Distributed

Dec.
Final Exam Due

Accessibility:
Any student with a disability that may need accommodations in order to successfully complete all requirements of this course should visit the Office of Disability Support Services, located in Room 516 of the Hill University Center, extension 4-4205 or at dss@uab.edu. This office is responsible for registering students and ensuring the University’s compliance with Section 504 of the Rehabilitation Act and the American with Disabilities Act. Once registered, this office will then inform faculty members of all courses in which the student is enrolled of the students status and the specific nature of any accommodations required. Any student requiring such accommodation should discuss this with the course master and assure that the appropriate correspondence is sent from the Office of Disability Support Services.

Honor Code:
As a student in the School of Public Health, you are subject to the SOPH Student Honor Code which can be found in its entirety at http://www.soph.uab.edu/default.aspx?id=844. You are responsible to understand the contents of the Honor Code and to abide by it. Academic dishonesty: Cheating includes but is not limited to the unauthorized use of notes, books or other sources of information; copying the work of another or allowing someone to copy the work of another student during a formal academic exercise (e.g. take home examination, homework assignment or written essay). Plagiarism is the undocumented use of other authors’ words, texts, images, and ideas that don’t come from your own head. Making up sources, altering numbers, statistics, or just a few words of a document is considered plagiarism. Poor documentation or paraphrasing of a source is also considered plagiarism. Unauthorized collaboration is working with others without the specific permission of the instructor on assignments that will be submitted for a grade. This rule applies to in-class or take-home tests, papers, labs, or homework assignments. Students may not collaborate without faculty authorization.

Any violations of the Honor Code will be punished to the full extent allowable under the SOPH Honor Code.