BIOST 2094 - Statistical Computing in R
Spring 2011

Instructor: Nicholas Christian, MS
E-mail: njc23@pitt.edu
Office Hours: After class and by appointment

Instructor: Abdus S. Wahed, PhD
E-mail: wahed@pitt.edu

Course Homepage: http://www.pitt.edu/∼njc23/
Time and Place: Tuesday 10:00-11:55 AM, WWPH 1201

Course Summary

BIOST 2094 is an introduction to statistical computing using R. This course will have two components. In the first part of the course you will learn how to write efficient and transparent programs in R. In the second part of the course, you will learn about packages and functions that are used in statistical analysis as well as techniques for managing data and using graphs to visually describe data. Prerequisite: BIOST 2041, 2042.

Learning Objectives

1. Be able to write efficient transparent programs in R
2. Produce clear and effective graphical descriptions of data
3. Import, export, and manipulate datasets
4. Analyze data using descriptive and inferential statistics
5. Analyze data through model fitting

Textbooks

There is *no required textbook* for this course. However, the following books are recommended as references:

- John Chambers, *Software for Data Analysis: Programming with R*, Springer, 2008.
- Michael J. Crawley, *The R Book*, Wiley, 2007.
- Brian Everitt and Torsten Hothorn, *A Handbook of Statistical Analyses Using R*, Chapman & Hall/CRC, 2006.
- Owen Jones, Robert Maillardet, Andrew Robinson, *Introduction to Scientific Programming and Simulation Using R*, Chapman & Hall/CRC, 2009.

Student Evaluation and Grading Scale

You will be graded on,

- Homework assignments (40%)
- One independent mid-term project (30%)
- One final group project (30%)

The cut-offs for computing letter grades will be: A, 100%-90%; B, 89%-80%; C, 79%-70%; D, 69%-60%; and F, <60%. Plus-minus grades will be assigned by dividing the respective intervals into thirds.

Homework

There will be approximately 4 homework assignments. You are encouraged to work in groups, but remember to submit your own unique work.

For each assignment, turn-in a hard copy of your `R` source file along with a brief write-up of the solutions (do not submit raw output). Also submit via e-mail (njc23@pitt.edu) a copy of your `R` source file. No late homework will be accepted unless it is due to some excused absence.

Mid-Term and Final Projects

For the mid-term project, you will work independently to perform a simulation study in `R`. For the final project, you will work in small groups (likely 4 to 5 students per group) to analyze a dataset using `R`. Each project consists of a technical report along with `R` code. The final project will also include a class presentation. Details of the mid-term and final projects will be handed out later in the semester.

Lecture Schedule

Date		Topic
January 11, 2011		Lecture 1: Getting Started with `R`
January 18, 2011		Lecture 2: Data Structures and Data Types
January 25, 2011		Lecture 3: Basic Programming
February 1, 2011	Homework 1 Due	Lecture 4: Advanced Programming
February 8, 2011		Lecture 5: Basic Graphics
February 15, 2011	Homework 2 Due	Lecture 6: Advanced Graphics
February 22, 2011		Lecture 7: Numerical Methods
March 1, 2011	Mid-Term Project Due	Lecture 8: Dataframes
March 8, 2011	**Spring Break**	No Class
March 15, 2011		Lecture 9: Basic Statistical Methods
March 22, 2011	Homework 3 Due	Lecture 10: Linear and Generalized Linear Models
March 29, 2011		Lecture 11: Survival Models
April 5, 2011	Project Proposal Due	Lecture 12: Longitudinal Models
April 12, 2011	Homework 4 Due	`R` Server - Chi Song (Chuck)
April 19, 2011	Final Project Due	Group Presentations
April 26, 2011	Final Project Due	Group Presentations

Students with Disabilities

If you have any disability for which you may require accommodation, you are encouraged to notify both your instructor and the Office of Disability Resources and Services, 216 William Pitt Union (412-648-7890) during the first two weeks of the term.

Academic Integrity Statement

All students are expected to adhere to the school's standards of academic honesty. Any work submitted by a student for evaluation must represent his/her own intellectual contribution and efforts. The GSPH policy on academic integrity, approved by EPCC on 10/14/08, which is based on the University policy, is available online at http://www.publichealth.pitt.edu/interior.php?pageID=126. The policy includes obligations for faculty and students, procedures for adjudicating violations, and other critical information. Please take the time to read this policy.

Students committing acts of academic dishonesty, including plagiarism, unauthorized collaboration on assignments, cheating on exams, misrepresentation of data, and facilitating dishonesty by others, will receive sanctions appropriate to the violation(s) committed. Sanctions include, but are not limited to, reduction of a grade for an assignment or a course, failure of a course, and dismissal from GSPH.

All student violations of academic integrity must be documented by the appropriate faculty member; this documentation will be kept in a confidential student file maintained by the GSPH Office of Student Affairs. If a sanction for a violation is agreed upon by the student and instructor, the record of this agreement will be expunged from the student file upon the student's graduation. If the case is referred to the GSPH Academic Integrity Hearing Board, a record will remain in the student's permanent file.