36-711: Statistical Computing  
Fall 2003  
TTh 9:00–10:20 Porter Hall A22  

Course Policies and Syllabus  

Essential Information  

Instructor: Brian Junker  
132E Baker Hall  
(412) 268-8874  
brian@stat.cmu.edu  

Teaching Assistant: Alex Rojas  
229C Baker Hall  
(412) 268-7826  
arojas@stat.cmu.edu  

Office Hours: TBA  
OR BY APPOINTMENT.  

Course Description  

This course will cover a wide range of topics that are likely to be of use to a graduate student or researcher who needs to use and develop statistical methods. The course has three major functions: It is an immigration course, in which you learn some basic tools for everyday statistical computing and document preparation in our Department. We will focus on Splus (or R) and LaTeX, which you will use throughout your career here at Carnegie Mellon. It is a sampling of numerical methods in statistics. We will concentrate on optimization, ideas from numerical linear algebra, numerical integration and Monte Carlo Methods in Splus (or R) and C. It is a survey of special topics that you will find useful as you pursue MS and PhD degrees in our Department. These include interfacing Splus and C, exploring methods of computational statistics, and developing strategies for tackling computational problems in statistics.  

Texts  


Prerequisites  

If you have had a solid undergrad course in mathematical probability and statistics, and you have some familiarity with calculus, linear algebra, and computer programming\(^1\), you are ready to go in this course. In addition I will assume most of you are concurrently registered for 36-701 (Perspectives on Statistics and Statistics Instruction) 36-705 (Intermediate Statistics) and 36-707 (Regression Analysis).  

\(^1\)You don’t need lots of experience programming, but you should know roughly how to “read” a program in Fortran, C, Pascal or something similar; this will help when you have to teach yourself details of using Splus, Fortran, C, etc., that we will not have time to go into in lecture.
Topics

Here are the general themes or topics I will discuss in the course. Each topic motivates us to learn a new technique in statistical computing or computational statistics. We won’t cover any topic in much depth, but in some cases I’ve listed some subtopics that I may try to say a little about.

- Crash Course in UNIX, Splus and \LaTeX.
- Numerical Solution of Nonlinear Equations.
- Numerical Linear Algebra.
- Numerical Maximization.
- Numerical Integration.
- Simulation Methods.
- Other Topics.

In addition we will meet many familiar topics along the way, such as computer arithmetic, truncation and rounding, error, condition number of a calculation, etc.

Other Texts and Resources

As you can see the required and recommended texts only cover a small part of the course. A variety of other resources will be available to you during the course:

▷ Course Web Site

The course web site is http://www.stat.cmu.edu/~brian/711. This is the place to go for the most current course information (when is the next assignment due, what is it, etc. In addition links to a variety of resources on the World Wide Web are available there.

▷ Other Books

At the end of this syllabus is a list of books that are also useful in statistical computing. They are mostly available at the Engineering and Science (E&S) Library in Wean Hall (except where indicated).

In addition, members of the Statistics Department are free to browse the DeGroot Library in 229 Baker Hall. Books and journals are checked out and returned on the honor system, usually just long enough to xerox the relevant pages.

Finally, you may find that the right introductory book for you (for example if you are learning C for the first time) is available at the University of Pittsburgh bookstore near Fifth and Bigelow, Border’s in the malls, Barnes & Noble in Squirrel Hill, amazon.com, etc.

▷ In the Department

Everyone in the department does statistical computing and has ideas, opinions and resources that you may be interested in. Whether in the hallways, offices or computer labs—or in their online webpages (e.g. http://www.stat.cmu.edu/~hseltman)—the people in the department can help.
Homework

- A full homework will be handed out approximately every two weeks, for a total of 5–7 assignments over the course of the semester.
- Smaller sets of exercises will be done in class, or assigned for discussion in the next class.
- Toward the end of the semester there may be a larger homework assignment that will count as an end-of-semester exam or project.

Answers to specific questions and summary of program results and performance should be well organized and clearly written and presented. Program code, if required, should be in an appendix, and well documented.

Discussing homework assignments with your fellow students is encouraged. However, you will learn statistical computing well only if you think about how you would solve the assignment before discussing it, and if you share as few line of code as possible.

Also, you should be aware that the best way to learn statistical computing is to “play”. This means, e.g. trying various modifications of the program and/or data, and seeing if what you expect really happens. In addition learn to make good use of on-line help, reference books, and the world wide web.

Participation

You are strongly encouraged to participate in class discussion. This helps you become more comfortable with the material, and, at the same time, gives other members of the class the benefit of your ideas and perspective. In particular, you should ask questions whenever you have them. Your questions show me what I have made clear and what needs to be clarified—and, consequently, they help me to teach more effectively.

Grading

If a final project is assigned, the grading scheme will be, approximately:

- Homework 50%
- Class Participation and Finger Exercises 25%
- Final Project 25%

If no final project is assigned, the grading scheme will be, approximately:

- Homework 67%
- Class Participation and Finger Exercises 33%

Getting in Touch with Us

The easiest and most reliable way to get in touch with us is by electronic mail. Feel free to send mail at any time to brian@stat.cmu.edu or arojas@stat.cmu.edu.
Plagiarism

It is sometimes unclear what behavior is appropriate when discussing assignments with classmates in an interactive learning environment. Feel free to come talk to me if you have any questions or comments.

Copying answers (even incorrectly) from another student, allowing another student to copy from you, or any other form of cheating are typically grounds for course failure. We are obliged in these situations to report the incident to the appropriate University authorities.

Physically disabled and learning disabled students

The Office of Equal Opportunity Services provides support services for both physically disabled and learning disabled students. For individualized academic adjustment based on a documented disability, contact Equal Opportunity Services at eos@andrew.cmu.edu or (412) 268-268-2012.

Other Books
