Biost 578 B
Introductory Applied Data Analysis

Syllabus

Instructor : Scott S. Emerson, M.D., Ph.D., Professor of Biostatistics
Office : HSB F665
Phone : 543-1044 (Biostatistics)
Email : semerson@u.washington.edu
Office hours : by appointment

Time and Place : Lectures : MWF 9:30 - 10:20 HSB T635

Prerequisites : Biost 515 or 518; Stat 513 or permission of instructor
(This course is not intended for students who have taken Biost/Stat 570.)

Text : None

Class Web Pages: http://faculty.washington.edu/~semerson/b578b

The web page will be used to post data sets, homework assignments, etc. I urge you to check this site regularly. Questions that are submitted to me (via email or otherwise) that I think might be of general interest will have their answers posted on the web page, as well.

Assignments :

Weekly data analyses for discussion, along with a brief outline of the tables and figures which would be presented in a complete report of the data analysis. At the end of the quarter, a complete report of a data analysis will be submitted, and the student will defend their analysis in an oral exam similar to the second year applied exam.

Grading(CR/NC) :

Written homeworks	50%
Project	50%

Assignments will include weekly written Statistical Analysis Plans, class discussion of analyses, and a single formal data analysis report.

Course Objectives

Students will gain experience in the general approach to a data analysis and in the application of the statistical methods learned in their prior coursework. Emphasis will be placed on developing a statistical analysis plan appropriate for the scientific question to be addressed. The course will primarily use the data arising from a recently completed randomized clinical trial in a chronic liver disease. Scientific issues to be addressed include:1) Design of the sequential clinical trial; 2) Description of the baseline characteristics of participants; 3) Comparison of adverse event rates; 4) Analyses of the primary endpoints of time to event data; 5) Longitudinal analyses of secondary laboratory markers; and 6) Assessment of quality control issues in measurements.

Statistical issues of particular interest include: 1) Sequential sampling (this study was terminated early); 2) Competing risks; 3) Informative missingness in the longitudinal trials; and 4) Interrater reliability.

At the end of this course, a student should have made significant progress toward being able to:

1. Demonstrate an organized approach to the analysis of data gathered to address a scientific question.
2. Develop a statistical analysis plan to address the question, including specification of
   a. suitable descriptive analyses of the data,
   b. an appropriate model to analyze the data to address the scientific question, including
      - identification of the dependent variable, the functional of its distribution which is to be modeled, and the link function,
      - identification of the independent variables and any transformations (including dummy variables) that might be necessary,
      - modeling of interactions as necessary, and
      - appropriate consideration of variables to include in the model.
3. Perform the specified analyses.
4. Make statistical inference about the generalizability of the analysis results to a larger population.
5. Present the results of your analysis to a statistically naive reader, including a full interpretation of all parameter estimates taking into account the functional being modeled, the link function, and any transformations of the predictors.

I welcome student suggestions regarding ways in which these goals can be best achieved. If you have questions regarding the content or structure of the class, please feel free to talk (or write) to me at any time during the quarter.