Applied Statistical Modeling for Medicine and Public Health

University of Kentucky
BST 764, Fall 2011
Credit: 3.0

Lecture: 12:30 p.m - 1:45 p.m., Tuesdays & Thursdays
Room 207, College of Public Health (CPH 207)

Instructor: Patrick Breheny, Ph.D.
Office: Room 203C, College of Public Health Building
Phone: 218-2077
e-mail: patrick.breheny@uky.edu
Office hours: Whenever I'm in my office, or by appointment

Course description: The goal of this course is to introduce useful statistical models not otherwise encountered in the core courses of the Ph.D. program in Epidemiology and Biostatistics. In my opinion, the six most important topics are:

1. Penalized regression
2. Discriminant analysis
3. Nonparametric regression (kernel methods and splines)
4. Classification and regression trees
5. Principal components analysis (and related ideas)
6. Robust regression

In addition to discussing these new types of models, we will also introduce two important ways of evaluating and performing inference on models: cross-validation and the bootstrap. We will spend 2-3 weeks on each topic. At the end of the course, it is my goal that you will be able to:

1. Understand the ideas and concepts that the above methods are based on
2. Apply these methods to real data
3. Distinguish the strengths and weaknesses of these methods compared to each other and compared to more traditional methods
4. To better understand what research involving statistical methodology consists of, how it proceeds, and how it is communicated

Text:

- Hastie T., Tibshirani R., and Friedman, J. (2009) *The Elements of Statistical Learning*. Second edition. Springer.
- (suggested) Harrell F. (2001) *Regression Modeling Strategies*. Springer.

Prerequisite: BST 675; BST 760.

Course website: The course notes, assignments, data sets, and other relevant materials will be made available on the course web site:
`http://web.as.uky.edu/statistics/users/pbreheny/764-F11/index.html`

Grading: Your grade will be based on six projects, one over each of the six course topics. Each project will consist of three sections: (1) Mathematical concepts and derivations (2) Simulation study (3) Analysis of real data.

The grading scale will be as follows:

85-100	A
70-84	B
55-69	C
Below 54	E

Proofreading: This is the first time I have taught this course, and my notes are likely to have typographical errors. If you see them, please tell me about them! I will announce corrections on the course webpage, or if the mistake is critical, send out an e-mail.

Electronic communication: I will occasionally send e-mails to the class (to the account listed for you in the campus directory), so please check that account regularly.

Plagiarism: You may discuss projects with fellow classmates, but all writing should be your own. The University of Kentucky takes plagiarism seriously, and has in place a number of rather severe academic sanctions, a summary of which can be found at `http://www.uky.edu/Ombud/acadoffenses/index.htm`

Complaints: Students with suggestions or complaints should see me first, and if we cannot come to an agreement, I will direct you to the head of the department.

Disabilities: If anyone has a disability requiring special accommodations, please let me know as soon as possible, so that these arrangements can be made.

I look forward to getting to know you, and I hope that we have a great semester together!