Department of Biostatistics
Brown University

Linear and Generalized Linear Models

PHP 2601 **FALL 2013**

Instructor: Xi Luo
Email: xluo@stat.brown.edu
Office: 121 S Main St, Rm 746
Telephone: 3-6321
Office hours: TTh 2:30-4:00 or by appointment

Classes meet: 121 S Main St, Rm 636

Teaching Assistant: TBD
121 S Main St
Regular office hours: TBD

Course Overview

Linear and generalized linear models are essential for the practice of statistics in many different fields. They are the backbones for models and methods with a wide range of applications, from health insurance premium to disease diagnosis.

This graduate-level course introduces the rigorous theory and practice of linear and generalized linear models. Models for continuous, binary, and count data will be introduced. The emphasis is on building skills for performing independent data analysis, from understanding the mathematics and assumptions, to graphical and numerical tools for model diagnosis, and to advanced techniques for model-fitting and decision-making. This class uses Canvas.

The prerequisites for this class are linear algebra, probability, statistical inference, and applied regression, all at the college level or above.

Textbooks

Lecture notes and assignments.

Seber and Lee, *Linear Regression Analysis*, 2nd Edition (2003, Wiley).

McCullagh and Nelder, *Generalized Linear Models*, 2nd Edition (1989, Chapman & Hall/CRC).

Agresti, *Categorical Data Analysis*, 2nd Edition (2002, Wiley).

Computing

Most of the analysis in class will be carried out in *R*. You are welcome to use other software packages (*SAS*, *STATA*, *S-plus* and etc) for homework assignments and exams, however there will be limited resources to help you on these software packages.

Grading
Your overall grades will be determined as follows:
background survey (5%), homeworks (20%), midterm project (25%), and final project and presentation (50%).

You are encouraged to discuss with the instructor as soon as possible (by Nov 1) about your final project. It usually involves analyzing an interesting dataset, producing a written report, and delivering an oral presentation. Biostat PhDs are also expected to rigorously analyze a method not discussed in lectures.

Module	Day	Topic	Text
Review	R, 9/5	Introduction	SL 1-2
	T, 9/10	Matrix, MVN	
LM Basics	R, 9/12	Estimation & assumptions	SL 3, 9
	T, 9/17		
	R, 9/19		
	T, 9/24		
	R, 9/26	Hypothesis testing	SL 4
	T, 10/1	Confidence intervals	SL 5, 6
LM Advanced	R, 10/3	Departures & remedies	SL 10
	T, 10/8		
	R, 10/10	Model selection & prediction	SL 12
	T, 10/15		
ANOVA	R, 10/17	ANOVA	SL 8
	T, 10/22		
	F, 11/1	Midterm project due	
GLM Basics	R, 10/24	Model	MN 1-3
	T, 10/29		
	R, 10/31	Binary data	MN 4
	T, 11/5		
	R, 11/7	Contingency table	A 2, 3, 8
	T, 11/12		
GLM Advanced	R, 11/14	Polytomous data	MN 5
	T, 11/19		
	R, 11/21	Log-linear models	MN 6
	T, 11/26		
	R, 11/28	*Thanksgiving: no class*	
GLM assumptions	T, 12/3	Model checking and diagnosis	MN 12
	R, 12/5		
Final	12/8--12/12	Presentations	