

ILRST 6190 Hierarchical Regression Models

Instructor

Dr. William Anderson
1176 Comstock Hall
wla6@cornell.edu

Lectures

M,W 8:40AM-9:55AM in Ives 112

Course Website

BTRY 6520 at
<http://www.blackboard.cornell.edu>

Course Overview

This course will foster a solid working knowledge in modern methods of regression analysis - hierarchical regression models. Simply put, the name stems from two sources. First, it follows from the structure of the data, e.g. people clustered within states. Second, from the hierarchical structure of these models, e.g. parameters of the within state regressions and from the hyperparameters at the uppermodel level. This class is designed for Ph.D. and advanced masters students. Students in ILR, (bio)statistics, economics and finance will especially find the fruits of this course beneficial in research. We will introduce statistical tools and methods to study Nature the way she is: complex and nonlinear.

The class is applied in nature, with an occasional look at the theory behind hierarchical models. The intuition behind these models is typically difficult to grasp, while the mathematics is straight forward. We will concentrate on the intuition. Examples are provided throughout the course and students will implement the techniques covered using real and simulated data sets. Covered topics include:

- Review of classical regression
- Review of logistic regression
- Review of Poisson regression
- Hierarchical Linear Models:
 - Linear structures, both nested and non-nested
- Generalized Hierarchical Linear Models:
 - logistic regression, both nested and non-nested
 - Poisson regression, both nested and non-nested

Time and sufficient student interest permitting:

- Penalized splines
- Hierarchical Bayesian regression models

Note that the “Review” is not just about model building and model checking. It is about the intuition behind these classical methods and why they fail to handel longitudinal data, correlated data, nonlinear relationships, etc. This will set the stage for why hierarchical models work with these types of data structures and why they should be always used.

Dedicated students should come away with an in depth understanding of the statistical and computing concepts related to modern regression methods - hierarchical regression models, as well as a thorough comprehension of how and when to implement the various techniques in practice.

The software used will be R; it is free (<http://cran.r-project.org/>) and very powerful. In my opinion, R is the future of statistical computing and well worth the time and effort learning the program. You are welcome to use the software of your choice, e.g. STATA, SAS, etc.

Prerequisites

- A graduate level course covering multivariate regression, and the associated prior courses.
- Prior statistical programming is helpful, e.g. C++, MATLAB, R, STATA, SAS

Textbooks

This is a good introductory book, one that you will want to keep on your bookshelf for reference:

Gelman, Andrew and Jennifer Hill (2006). *Data Analysis Using Regression and Multi-level/Hierarchical Models*, Cambridge University Press, New York.

Other books are given on the reference material list.

Evaluation

- 10% Homework
- 40% Mid-semester project
- 50% Final project

Homework

About 4 homework assignments will be given regularly throughout the semester.

Office Hours

Determined in class the first week.