Washington University, St. Louis

Jeff Gill
Voice: 314-935-9012
Fax: 314-935-5856

Political Science 430/Mathematics 425C/Public Health Sciences M19-530: Multilevel Modeling for Quantitative Research Fall 2015 Seminar

Thursday, 4:00-6:00 PM, Location TBD.

- Course Description: This **3-credit course** covers statistical model development with explicitly defined hierarchies. Such multilevel specifications allow researchers to account for different structures in the data and provide for the modeling of variation between defined groups. The course begins with simple nested linear models and proceeds on to non-nested models, multilevel models with dichotomous outcomes, and multilevel generalized linear models. In each case, a Bayesian perspective on inference and computation is featured. The focus on the course will be practical steps for specifying, fitting, and checking multilevel models with much time spent on the details of computation in the R and bugs environments.

- Competencies: At the conclusion of this course participants will: be able to specify and estimate multilevel (hierarchical) models with linear and nonlinear outcomes, treat missing data in a principled and correct manner using multiple imputation, gain facility in the R and bugs statistical languages, know how to compute the appropriate sample size and power calculations for multilevel models, gain exposure to Bayesian approaches including MCMC computation, and be able to assess model reliability and fit in complex models.

- Prerequisite Details: This course assumes a knowledge of basic statistics as taught in a first year undergraduate or graduate sequence. Topices should include: probability, cross-tabulation, basic statistical summaries, and linear regression in either scalar or matrix form. Knowledge of R, basic matrix algebra and calculus is helpful.

- Course Grade: The final grade will be based on two components: weekly attendance and participation (20%) and exercises (80%). Graduate students will have one additional component of their exercise grade that constitutes 30 points out of the 80 points total: submission of an analysis of real research using a multilevel model applied to data in their field along with 5-10 pages of discussion to include a description of the data, model diagnostics, and the subsequent findings. Consider this assignment to be the start of a research manuscript to be eventually submitted to a an academic journal. Graduate students will still submit all exercises assigned below in addition to this work.

- Office Hours: Friday 8-10, and by appointment.

- Incompletes: Due to the scheduled nature of the course, no incompletes will be given.

- Teaching Assistant: TBD. Office Hours: TBD from X:XX to X:3X, Seigle XXX.

- Required Reading: Gelman and Hill, "Data Analysis Using Regression and Multilevel/Hierarchical Models (Cambridge University Press 2007). Some papers will be available at jstor.org or distributed by the instructor. Readings should be completed before class.

- Topics (subject to minor change):
  - **August 27:** Introduction To the Course and Motivation.
    - Reading: Gelman & Hill, Chapters 1, 2, and 9, R Tutorial online, Intro code from the lecture. Chapter 9 code from the lecture.
    - Exercises: Gelman & Hill 2.2, 2.3, 9.4, 9.13.
  - **September 3 (no class) and September 10:** Linear and Generalized Linear Models Review.
    - Reading: Gelman & Hill, Chapters 3 and 4, Chapter 3-4 code from the lecture,

Binomial PMF likelihood grid search.
    - Exercises: Gelman & Hill 3.4, 4.4, 5.4, 6.1.
- **September 17:** Multilevel Structures and Multilevel Linear Models: the Basics.
    - Reading: Gelman & Hill, Chapters 11 and 12, Introductory Chapter (Gill and Womack, Forthcoming The SAGE Handbook of Multilevel Modeling). Chapter 11-12 code from the lecture.
    - Exercises: Gelman & Hill 11.4, 12.2, 12.5.
- **September 24:** Multilevel Linear Models: Varying Slopes, Non-Nested Models and Other Complexities.
    - Reading: Gelman & Hill, Chapter 13, Chapter 13 code from the lecture.
    - Exercises: Gelman & Hill 13.2, 13.4, 13.5.
- **October 1:** Multilevel Logistic Regression, Multilevel Generalized Linear Models.
    - Reading: Gelman & Hill, Chapter 14 (skip Section 14.3), Chapter 15, Chapter 14 code from the lecture.
    - Exercises: Gelman & Hill 14.5, 14.6, 15.1, 15.2.
- **October 8:** Multilevel Modeling in Bugs and R: the Basics, MCMC Theory.
    - Reading: Gelman & Hill, Chapter 16, Bayesian Estimation Case Study (Gill and Witko 2012), R to JAGS code for the model (get data from the download site: http://jgill.wustl.edu/downld), Chapter 16 code from the lecture.
    - Exercises: Gelman & Hill 16.1, 16.2, 16.3, 16.8.
- **October 15:** Fitting Multilevel Linear and Generalized Linear Models in Bugs and R, MCMC Coding.
    - Reading: Gelman & Hill, Chapter 17, Chapter 17 code from the lecture.
    - Exercises: Gelman & Hill Rerun 16.3 using the instructions in 17.2 and 17.3, 17.5.
- **October 22:** Likelihood and Bayesian Inference, Computation, MCMC Diagnostics and Customization.
    - Reading: Gelman & Hill, Chapter 18
    - Exercises: Gelman & Hill 18.1, 18.2, 18.4.
- **October 29:** Treatment of Missing Data.
    - Reading: Gelman & Hill, Chapter 25, Paper by van Buuren and Groothuis-Oudshoorn, Chapter 25 code from the lecture.
    - Exercises: missing data problems.
- **November 5:** Understanding and Summarizing the Fitted Models, Multilevel Analysis of Variance.
    - Reading: Gelman & Hill, Chapter 21 and 22, Chapter 21 code from the lecture, Chapter 22 code from the lecture.
    - Exercises: 21.1, 21.3, 21.4, 22.1.
- **November 12:** Model Checking and Comparison.
    - Reading: Gelman & Hill, Chapter 24, Chapter 24 code from the lecture.
    - Exercises: 24.1, 24.4.
- **November 19:** Sample Size and Power Calculations.
    - Reading: Gelman & Hill, Chapter 20, Chapter 20 code from the lecture.
    - Exercises: 20.1, 20.2, 20.3.
- **November 26:** Thanksgiving Holiday.
- **December 3:** Causal Inference Using Regression on the Treatment Variable.
    - Reading: Gelman & Hill, Chapter 9, Chapter 9 code from the lecture.
    - Exercises: TBD.