Government 2003: Bayesian Hierarchical Models

Thursday 10-12, CGIS North K450

Jeff Gill

- Course Description: This course covers Bayesian 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.

- Learning Outcomes: At the conclusion of this course participants will: be able to specify and estimate Bayesian 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 Bayesian 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 graduate sequence. Topices should include: probability, cross-tabulation, basic statistical summaries, and linear regression in matrix form. and *knowledge of R*. Exposure to 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%). Exercises are due one week after assignment on the syllabus.

- Office Hours: Wednesday 10-12, in the CGIS South building, Room S407.

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

- Teaching Fellow: Jonathan Homola. Office Hours: TBD.

- Required Text: Gelman and Hill, "Data Analysis Using Regression and Multilevel/Hierarchical Models (Cambridge University Press 2007). Other readings will be papers will made available at jstor.org or distributed by the instructor on this syllabus/webpage. Readings should be completed before class listed on the syllabus.

- Topics (subject to minor change):
    - **January 25:** Introducing Bayesian Inference.
        - *Background Reading.* Gelman & Hill: Chapters 1-2.
        - Refreshing R Skills (only if necessary): Starting R, Writing Functions, Making Plots,Sampling.
        - *Weekly Reading.* Issues in Inference: Gill, Jeff. The Insignificance of Null Hypothesis Significance Testing. Political Research Quarterly, vol. 52, no. 3, 1999, pp.647-674, JSTOR. Click here.,
        - Leamer, Edward E. Let's Take the Con Out of Econometrics. The American Economic Review, vol. 73, no. 1, 1983, pp. 31-43, JSTOR. Click here.,
        - King, Gary. 1986. How Not to Lie With Statistics: Avoiding Common Mistakes in Quantitative Political Science. American Journal of Political Science, 30: pp. 666-687. Click here, slides from the lecture.
        - *Exercises.* Gelman & Hill 2.2, 2.3, 2.4.

    - **February 1:** Detailed Linear Model Theory Review.
        - *Reading.* Gelman & Hill: Chapters 3-6.
        - Guy Lebanon's Summary (Georgia Tech),
        - Chapter 3-4 code from the lecture,
        - Binomial PMF likelihood grid search.
        - *Exercises.* Gelman & Hill: 3.4, 4.4, 5.4, 6.1.

    - **February 8:** Multilevel Structures and Multilevel Linear Models. *Note: we are back in K031 this meeting.*
        - *Reading.* Gelman & Hill: Chapters 11 and 12,
        - Gill, Jeff and Andrew J. Womack. The Multilevel Model Framework. In The SAGE Handbook of Multilevel Modeling. Scott, Marc A, Jeffrey S Simonoff and Brian D Marx (eds). London: SAGE Publications Ltd, 2013. Pp. 3-20. SAGE Research Methods. Click here.
        - Chapter 11-12 code from the lecture.
        - *Exercises.* Gelman & Hill: 11.4, 12.2, 12.5.

    - **February 15:** Multilevel Linear Models: Varying Slopes, Non-Nested Models and Other Complexities. Running Bayesian Regression Models.
        - *Reading.* Gelman & Hill: Chapter 13. Chapter 13 code from the lecture.
        - *Exercises.* Gelman & Hill: 13.2, 13.4, 13.5.

- **February 22:** Multilevel Modeling in Bugs and R: the Basics, MCMC Theory.
  - *Reading.* Gelman & Hill: Chapter 16, Bayesian Estimation Case Study (Gill and Witko, JPART 2012),
  - R to JAGS code for the model (get data from Dataverse),
  - Chapter 16 code from the lecture.
  - *Exercises.* Gelman & Hill: 16.1, 16.2, 16.3, 16.8.

- **March 1:** Multilevel Logistic Regression Models, Multilevel Bayesian GLMs.
  - *Reading,* Gelman & Hill: Chapter 14 (skip Section 14.3), Chapter 15. Chapter 14 codefrom the lecture.
  - *Exercises.* Gelman & Hill: 14.5, 14.6, 15.1, 15.2.

- **March 8:** 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.

- **March 15:** No Class: Spring break.

- **March 22:** Likelihood and Bayesian Inference, Computation, MCMC Diagnostics and Customization.
  - *Reading.* Gelman & Hill: Chapter 18.
  - *Exercises.* Gelman & Hill: 18.1, 18.2, 18.4.

- **March 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.

- **April 5:** No Class: MPSA meeting in Chicago.

- **April 12:** 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.

- **April 19:** Model Checking and Comparison.

        - *Reading.* Gelman & Hill: Chapter 24, Chapter 24 code from the lecture.
        - *Exercises.* 24.1, 24.4.

    - **April 26:** Sample Size and Power Calculations (voluntary extra session).
        - *Reading.*: Gelman & Hill: Chapter 20, Chapter 20 code from the lecture.
        - *Exercises.* 20.1, 20.2, 20.3.