Introduction to Applied Bayesian Statistics for Social Scientists
ICPSR Summer Program in Quantitative Methods
June 23-June 28, 2008

Instructor: Scott M. Lynch, Princeton University, slynch@princeton.edu
Time: 9:30-12:30 (lecture) and 1:30-4:30 (lab)
Location: University of North Carolina at Chapel Hill

Course Description

Bayesian statistics has received a growing amount of attention over the last decade in social science research, largely because of (1) the growth in the use of hierarchical modeling in social science coupled with the ease with which Bayesian modeling can handle such models, and (2) the development of Markov chain Monte Carlo (MCMC) simulation methods that have simplified parameter estimation. In this workshop, we will first review mathematical statistics and the classical approach to model construction and parameter estimation using maximum likelihood methods. Next, we will develop Bayes' Theorem and show its extension to probability distributions and modeling. Following the exposition of this theoretical material, we will spend considerable time presenting MCMC methods for parameter estimation, especially the Gibbs sampler and the random walk Metropolis algorithm. We will see how these methods can be used for a variety of models commonly used in social science research, including the OLS regression model, generalized linear models, hierarchical models, and multivariate models. The workshop will be highly-applied with an emphasis on using R and WinBugs for conducting Bayesian analysis. A key focus will be on comparing the Bayesian approach with the classical approach and showing the advantages of using the Bayesian approach in model evaluation and comparison, and inference. Familiarity with statistics at the level of multiple linear regression (as taught in most social science departments) is assumed. Familiarity with matrix algebra and the basic concepts underlying differential and integral calculus will be helpful, although we will briefly review these topics.

Format

The course will meet mornings and afternoons for one week. The morning sessions will involve lecture; the afternoon sessions will involve computer application. I will provide exercises in the afternoon involving both analytic and computer-based solutions.

Reading

Lynch, Scott M. (2007). *Introduction to Applied Bayesian Statistics and Estimation for Social Scientists*. New York: Springer.

Schedule

<u>Monday 6/23</u>

Calculus and Matrix Algebra review; probability theory and review of classical statistics; Introduction to R programming (Appendix A and B, Chapters 1 and 2)[*]

<u>Tuesday 6/24</u>

Bayes' Theorem for point probabilities and distributions; prior distributions; summarizing posterior distributions analytically; logic of sampling and two simple sampling methods (Chapter 3, Chapter 4 pp. 77-88)

<u>Tuesday 6/25</u>

Markov chain Monte Carlo simulation methods—Gibbs sampling and Metropolis Hastings Algorithms, monitoring performance, and making inference (Chapter 4 and 5, Chapter 6 pp. 131-153)

<u>Wednesday 6/26</u>

Posterior predictive simulation, model evaluation and comparison; linear regression modeling (Chapter 6 pp. 154-164, Chapter 7)

<u>Thursday 6/27</u>

Generalized linear modeling (Chapter 8)

<u>Friday 6/28</u>

Hierarchical modeling, random effects modeling, using WinBugs (Chapter 9)

[*] If we cover this material quickly enough, we may shift the schedule up to allow us to cover multivariate modeling (Chapter 10) on Friday.