Advanced Bayesian Models for the Social Sciences

Instructors:
Week 1&2: Skyler J. Cranmer
Department of Political Science
University of North Carolina, Chapel Hill
skyler@unc.edu

Week 3&4: Jong Hee Park
Department of Political Science
University of Chicago
jhp@uchicago.edu

TA: Tobias Heinrich (Rice University)
tobias.heinrich@gmail.com

Date and Time: TBD

Office Hours: Cranmer: TBD
              Park: TBD
              Heinrich: TBD

Description and Schedule:
This course covers the theoretical and applied foundations of Bayesian statistical analysis at a level that goes beyond the introductory course at ICPSR. Therefore knowledge of basic Bayesian statistics (such as that obtained from the Introduction to Applied Bayesian Modeling for the Social Sciences workshop) is assumed. First, we will discuss model checking, model assessment, and model comparison, with an emphasis on computational approaches. Second, the course will cover Bayesian stochastic simulation (Markov chain Monte Carlo) in depth with an orientation towards deriving important properties of the Gibbs sampler and the Metropolis Hastings algorithms. Extensions and hybrids will be discussed. The third and fourth modules will focus on applications of Bayesian statistics in social science data analysis. The topics include Bayesian change-point models, dynamic linear models, Bayesian item response theory (IRT) models, and Bayesian hierarchical models for panel data. Throughout the workshop, estimation with modern programming software (R and BUGS) will be emphasized.

Week I: Bayesian Model Checking, Assessment and Comparison.
Skyler Cranmer (University of North Carolina, Chapel Hill)

The first week has three components: assessing model quality, comparing models in a Bayesian context, and standard statistical computing tools that are useful for Bayesian analysis. The emphasis is on in-depth technical understanding of the mathematical statistics that justify and govern the use of these tools.

Monday: Quick Review of Bayesian Inference
1.   This is not intended to be (nor will it be) a substitute for an introductory Bayes course. Rather it will be a refresher to make sure we're all on the same page.
2.   Essential Reading: Gill (2007) Chapters 1-4 or equivalent

Tuesday: The Bayesian Prior
1.   Bayesian Shrinkage
2.   (Many) Types of Priors
3.   Essential Reading: Gill (2007) Chapter 5

Wednesday: Assessing Model Quality
1.   Global Sensitivity Analysis
2.   Local Sensitivity Analysis
3.   Global Robustness
4.   Local Robustness
5.   Comparing Data to the Posterior Predictive Distribution
6.   Essential Reading: Gill (2007) Chapter 6

Thursday: Model Comparison
1.   Posterior Probability Comparison
2.   Cross-Validation
3.   Bayes Factors
4.   AIC, BIC, DIC
5.   Software Issues
6.   Essential Reading: Gill (2007) Chapter 7

Friday: Introduction to Monte Carlo Integration
1.   Rejection Sampling
2.   Classical Numerical Integration
3.   Importance Sampling
4.   Mode finding and the EM Algorithm
5.   Essential Reading: Gill (2007) Chapter 8

Optional Additional Reading (for the week):
a.   Carlin, B. P. and Chib, S. (1995). ``Bayesian Model Choice via Markov Chain Monte Carlo Methods.'' Journal of the Royal Statistical Society, Series B 57, 473-484.

    b.   Dempster, A. P., Laird, N. M., and Rubin, D. B. (1977). ``Maximum Likelihood from Incomplete Data via the EM Algorithm.'' Journal of the Royal Statistical Society, Series B 39, 1-38.

    c.   Kennedy, W. J. and Gentle, J. E. (1980). Statistical Computing. New York: Marcel Dekker.

    d.   Metropolis, N. and Ulam, S. (1949). ``The Monte Carlo Method.'' Journal of the American Statistical Association 44, 335-3.

    e.   Mooney, C. Z. (1997). Monte Carlo Simulation. Thousand Oaks, CA: Sage.

    f.   Rubin, D. B. (1987). ``A Noniterative Sampling/Importance Resampling Alternative to the Data Augmentation Algorithm for Creating a Few Imputations When Fractions of Missing Information Are Modest: the SIR Algorithm.'' Discussion of Tanner & Wong (1987). Journal of the American Statistical Society 82, 543-546.

Week II: Markov Chain Monte Carlo.
Skyler Cranmer (University of North Carolina, Chapel Hill)

This week we continue our focus on computational techniques. We will expand on the idea of Monte Carlo integration introduced last week and then discuss Markov chains, Markov Chain Monte Carlo, MCMC algorithms (esp. Metropolis-Hastings and Gibbs Sampling), and conclude by discussing convergence diagnostics.

Monday: Markov Chains
    1. What are Markov Chains?
    2. Some Simple Examples
    3. Marginal Distributions
    4. Properties of Markov Chains
    5. The Ergodic Theorem
    6. Essential Reading: Gill (2007) Chapter 9

Tuesday: Gibbs Sampling
    1. The Gibbs Sampler
    2. Software Topic: Bayesian Analysis with MCMCpack
    3. Essential Reading: Gill (2007) Chapter 9

Wednesday: Metropolis-Hastings
    1. The Metropolis-Hastings Algorithm
    2. The Hit-and-Run Algorithm
    3. Software Topic: Bayesian Analysis with WinBUGS
    4. Essential Reading: Gill (2007) Chapter 9

Thursday: Convergence Diagnostics
    1. Trace Plots
    2. Running mean plots
    3. Density/HPD plots
    4. The Geweke Diagnostic

5. The Gelman and Rubin Diagnostic
6. The Raftery and Lewis Diagnostic
7. The Heidelberger and Welch Diagnostic
8. Essential Reading: Gill (2007) Chapter 12

Friday: Convergence Continued
1. Finish anything we didn't cover on Thursday
2. Software topic: Using the CODA and BOA packages in R

Optional Additional Reading (for the week):

a.   Casella, G. and George, E. I. (1992). ``Explaining the Gibbs Sampler.'' The American Statistician 46, 167-174.

b.   Gelfand, A. E. and Smith, A. F. M. (1990). ``Sampling-Based Approaches to Calculating Marginal Densities.'' Journal of the American Statistical Association 85: 389-409.

c.   Geman, S. and Geman, D. (1984). ``Stochastic Relaxation, Gibbs Distributions and the Bayesian Restoration of Images.'' IEEE Transactions on Pattern Analysis and Machine Intelligence 6, 721- 741.

d.   Geyer, C. J. (1992). ``Practical Markov Chain Monte Carlo.'' Statistical Science 7, 473-511.

e.   Hastings, W. K. (1970). ``Monte Carlo Sampling Methods Using Markov Chains and Their Applications.'' Biometrika 57, 97-109.

f.   Jackman, S. (2000). ``Estimation and Inference via Bayesian Simulation: An Introduction to Markov Chain Monte Carlo.'' American Journal of Political Science 44, 375-404.

g.   Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H., and Teller E. ``Equation of State Calculations by Fast Computing Machine.'' Journal of Chemical Physics 21, 1087-1091.

h.   Peskun, P. H. (1973). ``Optimum Monte Carlo Sampling Using Markov Chains.'' Biometrika 60, 607-612.

i.   Tierney, L. (1994). ``Markov Chains for Exploring Posterior Distributions.'' Annals of Statistics 22, 1701-1728.

j.   Cowles, M. K., Roberts, G. O., and Rosenthal, J. S. (1999). ``Possible Biases Induced by MCMC Convergence Diagnostics.'' Journal of Statistical Computation and Simulation 64, 87-104.

k.   Gelfand, A. E. and Sahu, S. K. (1994). ``On Markov Chain Monte Carlo Acceleration.'' Journal of Computational and Graphical Statistics 3, 261-276.

l.   Gelman, A., Rubin, D. B. (1992). ``Inference from Iterative Simulation Using Multiple Sequences.'' Statistical Science 7, 457- 511.

m.   Geyer, C. J. (1992). ``Practical Markov Chain Monte Carlo.'' Statistical Science 7, 473-511.

n.   Zellner, A. and Min, C-K. (1995). ``Gibbs Sampler ConvergenceCriteria.'' Journal of the American Statistical Association 90, 921-927.

Week III: Bayesian Analysis of Historical Data
Jong Hee Park (University of Chicago)

Monday: Time in Social Science Methodology

Nathaniel Beck and Jonathan N. Katz, 2011, "Modeling Dynamics in Time-Series Cross-Section Political Economy Data," *Annual Review of Political Science*

Gregory J. Wawro_ Ira Katznelson, 2012, "Political Science and History: Enhancing the Methodological Repertoire" Working Paper.

Bruce Western and Meredith Kleykamp, 2004, "A Bayesian Change Point Model for Historical Time Series Analysis", *Political Analysis*, 12(4).

Tuesday: Markov Model and Change-point Analysis

Markov model, Markov transition model, Changepoint problem

Simon Jackman, "In and Out of War and Peace: Transitional Models of International Conflict" Working Paper.

Quandt, R.E. 1958. "The estimation of the parameters of a linear regression system obeying two separate regimes". *Journal of American Statistical Association*: 873 – 880

Chow, G.C. (1960). "Tests of Equality between Sets of Coefficients in Two Linear Regressions," *Econometrica*, 52, 211-22.

Spirling, Arthur. 2007. "Bayesian Approaches for Limited Dependent Variable Change Point Problems." *Political Analysis* 15(4): 387-405.

Wednesday: Hidden Markov Model and its Extensions

Fruhwirth-Schnatter, Sylvia. 2006. *Finite Mixture and Markov Switching Models*. New York: Springer-Verlag. Ch.1-3.

Chib, Siddhartha. 1996. "Calculating Posterior Distributions and Modal Estimates in Markov Mixture Models." *Journal of Econometrics* 75(1): 79-98.

Chib, Siddhartha. 1998. "Estimation and Comparison of Multiple Change-Point Models." *Journal of Econometrics* 86(2): 221-241.

Binary probit, ordered probit, and Poisson extensions

Jong Hee Park. 2011. "Changepoint Models for Binary and Ordinal Probit Models: An Application to Bank Rate Policy in the Interwar Period", *Political Analysis*, Vol.19, No.2, 188-204.

Jong Hee Park. 2010. "Structural Change in the U.S. Presidents' Use of Force Abroad." *American Journal of Political Science*. Vol.54, No.3.

Thursday: Dynamic Linear Models and its Extensions

Mike West & Jeff Harrison, 1997, *Bayesian Forecasting and Dynamic Models*, Springer (2nd Ed.) Ch.2.-5.

Fruhwirth-Schnatter, Sylvia, and Helga Wagner. 2006. "Auxiliary Mixture Sampling for Parameter-driven Models of Time Series of Small Counts with Applications to State Space Modelling." *Biometrika* 93(4):827-841.

Friday: Bayesian Intervention Analysis

(Background Reading) Box, G. E. P., and G. C. Tiao. 1975. "Intervention Analysis with Applications to Economic and Environmental Problems." *Journal of the American Statistical Association* 70 (03):70–79.

Jong, Piet de, and Jeremy Penzer. 1998. "Diagnosing Shocks in Time Series." *Journal of the American Statistical Association* 93 (06): 796–806.

Jong Hee Park, "Bayesian Intervention Analysis" Working Paper

Week IV: Bayesian Hierarchical Modeling
Jong Hee Park (University of Chicago)

Monday: Bayesian Panel Model

Rubin, Donald, 1980. "Using Empirical Bayes Techniques in the Law School Validity Studies" *Journal of the American Statistical Association*, 75(372): 801- 816

Siddartha Chib, "Panel Data Modeling and Inference: A Bayesian Primer" In L. Matyas, P. Sevestre (eds.), *The Econometrics of Panel Data*

Tuesday: Bayesian Panel Analysis using BUGS

Andrew Gelman, "Multilevel (Hierarchical) modeling: what it can and can't do." *Technometrics* 48: 432-435.

Andrew Gelman and Iain Pardoe, "Bayesian measures of explained variance and pooling in multilevel (hierarchical) models." *Technometrics* 48, 241-251.

A method of post-stratification

David Park, Andrew Gelman, and Joseph Baufumi, 2004, "Bayesian Multilevel

Estimation with Poststratification: State-Level Estimates from National Polls," *Political Analysis*, 12(4):375-385.

Jeffrey R. Lax and Justin H. Phillips, 2009, "How Should We Estimate Public Opinion in The States?" *American Journal of Political Science*, 53:107-121.

Wednesday: Bayesian Estimation of Item Response Theory Model

IRT Models as (a special type of) Binary Response Panel Models

(Background Reading) Keith Poole and Howard Rosenthal. 1985. "A Spatial Model of Legislative Roll Call Analysis," *American Journal of Political Science* , 357-384

Simon Jackman, 2001. "Multidimensional Analysis of Roll Call Data via Bayesian Simulation: Identification, Estimation, Inference, and Model Checking." *Political Analysis*. 9: 229-240.

Joshua Clinton, Simon Jackman, and Doug Rivers, 2004, "The Statistical Analysis of Roll Call Voting: A Unified Approach," *American Political Science Review* 98: 355-370.

Joseph Bafumi, Andrew Gelman, David Park, and Noah Kaplan. 2005, "Practical Issues in Implementing and Understanding Bayesian Ideal Point Estimation." *Political Analysis* 13: 171-187.

Thursday: How to Deal with Temporality in Panel Data

Andrew Martin and Kevin Quinn. 2002. "Dynamic Ideal Point Estimation via Markov Chain Monte Carlo for the US Supreme Court, 1954-1999" *Political Analysis* 10: 134-153.

Jong Hee Park. 2012. "A Unified Method for Dynamic and Cross-Sectional Heterogeneity: Introducing Hidden Markov Panel Models" *American Journal of Political Science*.

Jong Hee Park, 2011. "Analysis of Preference Changes using Bayesian Changepoint Item Response Theory Model," To appear in Steve Brooks, Andrew Gelman, Galin Jones and Xiao-Li Meng, eds., *Handbook of Markov Chain Monte Carlo*, Chapman & Hall/CRC Press.

Xun Pang, 2010. "Modeling Heterogeneity and Serial Correlation in Binary Time-Series Cross-sectional Data: A Bayesian Multilevel Model with AR(p) Errors" *Political Analysis*.