Advanced Bayesian Modeling

Daniel Stegmueller
Department of Political Science, Duke University
Office Hours: alternating 10am and 10pm, M-F
daniel.stegmueller@duke.edu

Course Information

Meeting Times: 1:00-3:00pm Daily
Meeting Location: Zoom
Teaching Assistant: TBA
TA Office Hours & TBA
Class Website: Canvas

Course Outline:

This course covers the theoretical and applied foundations of Bayesian statistical analysis at a level that goes beyond the introductory course at ICPSR. Knowledge of basic Bayesian statistics (such as that obtained from the Introduction to Applied Bayesian Modeling for the Social Sciences workshop) is assumed. We begin with a deeper discussion of the logic of Bayesian inference, the role of subjective priors, the likelihood and the posterior distribution. We then discuss predictive distributions, model checking, and model comparison. The course will cover Bayesian stochastic simulation (Markov chain Monte Carlo) in depth with an orientation towards deriving important properties of the Gibbs sampler, the Metropolis Hastings algorithms and more recent approaches. Extensions and hybrids will be discussed. Then we will focus on applications of Bayesian statistics in social science data analysis. The topics include bayesian hierarchical models for cross-sections and panel data, factor analysis models, item response theory models, semiparametric models, and instrumental variable and treatment effect models. Throughout the workshop, estimation with a variety of tools will be emphasized.

Readings:

Required for purchase (or that you at least have access to it any time):

- Gill, J. (2014). *Bayesian Methods: A Social and Behavioral Sciences Approach*, 3rd edition. Chapman & Hall/CRC.

All other readings (if needed) will be available on the course website.

Topics

I Introduction Monday

- Course details and modalities
- How did it all start?
- Basic ideas

II Fundamental principle of Bayesian inference Tuesday

- Probability theory review
- Subjective probability
- Dutch book theorem and the principle of *coherence*
- Bayes theorem
- Prior-posterior analysis

III The three key players: likelihood, priors, and posteriors. Wednesday

- The role of the likelihood function
- Joint and marginal posterior distribution
- Identification
- Bayes risk and Bayesian estimators
- Posterior predictive distributions
- Prior distribution choices
- Sensitivity analysis

IV Model comparison and criticism Thursday

- The marginal likelihood
- Bayesian model comparison
- Bayesian model averaging

V Markov Chain Monte Carlo Friday-Monday

- Simulation estimation
- Markov Chain Monte Carlo theory
- Construction of MCMC samplers
- Gibbs sampling
- Metropolis-Hastings sampling
- Hamiltonian Monte Carlo
- Diagnostics & words of caution

VI The Bayesian Linear Model — Tuesday

- The linear model
- Robust regression via t-errors
- Bayesian tobit model
- Diagnostics, model comparison, and prediction
- Conjugate and nonconjugate priors

VII Binary and count outcomes — Wednesday

- Bayesian probit and logit models
- Latent data augmentation
- Diagnostics via latent Bayesian residuals
- Bayesian estimation of Poisson and negative binomial models
- Dealing with complete separation in binary data models
- Prediction and effective graphical presentation

VIII Discrete choice models — Thursday-Friday

- Ordered choice models
- Priors and sampling strategies for latent variable cutpoints
- Multinomial and Conditional Logit Models: Principles and Bayesian estimation
- Identification problems in discrete choice models
- Multinomial Probit Models: Principles and various Bayesian estimation strategies
- Sampling unidentified models
- Understanding prior choices
- Diagnostics and model comparison

IX Multivariate outcomes — Monday

- The Bayesian Seemingly Unrelated Regression Model
- Multivariate Probit Model
- Prior choices and estimation
- Identification issues

X Hierarchical Models — Tuesday-Wednesday

- The Bayesian Hierarchical Linear Model
- Hierarchical Logit/Probit Models
- Understanding and choosing variance component priors
- Advantages/disadvantages of Bayesian vs. Frequentist hierarchical models
- Heterogeneity in units

- Heterogeneity in effects
- Functional form assumptions in hierarchical models
- 'Mr.P': Multilevel regression and post-stratification

XI Models for panel data Thursday

- Hierarchical models
- Modeling heterogeneity over units
- Correlated random effects
- Distributional assumptions and flexibe random effects
- Serially correlated residuals
- Autoregressive models

XII Latent factor models Friday

- Factor analysis for multivariate normal data
- Factor analysis for mixed data
- Structural models
- Distributional assumptions

XIII IRT / ideal point models Monday

- Ideal-point / item-response theory models
- Identification issues in IRT models
- Prior choices
- IRT models with covariates

XIV Semiparametric regression Tuesday

- Flexible prior distributions
- Finite mixtures of normals
- Dirichlet process priors
- Computational issues and shortcuts
- Identification issues

XV Treatment effect models Wednesday-Thursday

- Bayesian modeling of selection and endogeneity issues
- The classical instrumental variables estimator
- Bayesian instrumental variable model: priors and estimation
- Advantages of Bayesian IV in the presence of weak instruments
- Robust Bayesian IV via flexible error distributions