CEP 991B Bayesian Data Analysis
Spring 2003

Instructor: Kim Maier

Email: kmaier@msu.edu
Phone: 355-8538
Class time: Tu 12:40 – 3:30
Classroom: 109 Olds
Office Hours: By Appointment
Office: 451 Erickson

Website: *blackboard.msu.edu*

Course Content

This course provides an introduction to the method of model estimation using Bayesian data analysis techniques. In contrast to the traditional frequentist (likelihood) approach, Bayesian inference assumes that the data are fixed and the model parameters are random entities and have distributions. Students will learn the basics of the Bayesian approach, MCMC approaches to estimating data parameters, and how to use various Bayesian-related statistical software packages (BUGS and R). The course will begin with an introduction to the Bayesian approach and will concentrate on applying this approach to specific statistical models. An emphasis will be placed on evaluation of model fit and interpretation and synthesis of the estimation results and model fit. Knowledge of basic algebra, matrix algebra, and some calculus will be helpful, but not necessary.

Grading

Grades are criterion-referenced. That is, grades will be assigned based on the percent of the total possible points that you receive on the assignments [4.0 > 90%, 3.5 > 80%, 3.0 > 70%]. The 100 total points are distributed as follows:

Homework	10 points
Model Presentation	20 points
Class Project	50 points
Presentation of Project	20 points
Total	100 points

Homework: Several short homework assignments will be given throughout the semester. The format of the homework will vary from summarizing extra readings to short computational examples.

Model Presentation: You will prepare a presentation outlining the Bayesian approach for a specific statistical model. This presentation will be given during a class period. You will need to present the material to your classmates and prepare for them relevant material such as overheads, handouts, etc., that will describe the general approach for that particular statistical model, share results from an actual analysis using real data, and share and discuss diagnostics of the model fit for the MCMC algorithms you used to perform the analysis; please provide your classmates with all relevant computer code in a handout.

Class Project: The class project (to be completed individually) will consist of a paper that would be appropriate for publication.

Presentation of Project: You will present the findings of your report to the class. The presentation will

function as the final exam for this class.

Except in extreme circumstances, assignments must be submitted on the day that they are due.

Readings

The book will be available at the MSU campus bookstore, SBS, and at the College Store in Hagadorn Plaza.

Required: Gill, Jeff (2002). *Bayesian Methods: A Social and Behavioral Sciences Approach*. Boca Raton, FL: Chapman & Hall/CRC.

Recommended: Carlin, Bradley P. and Louis, Thomas A. (2000). *Bayes and Empirical Bayes Methods for Data Analysis* (2nd ed.). Boca Raton, FL: Chapman & Hall/CRC.

Best, N. G., Spiegelhalter, D. J., Thomas, A., & Brayne, C. E. G. (1996). Bayesian analysis of realistically complex models. *Journal of the Royal Statistical Society, Series A (Statistics in Society), 159*(2), 323-342.

Blyth, S. (1995). The dead of the Gulag: An experiment in statistical investigation. *Applied Statistics, 44*(3), 307-321.

Brooks, S. P. (1998). Markov chain monte carlo method and its application. *Statistician, 47*(1), 69-100.

Casella, G., & George, E. I. (1992). Explaining the Gibbs Sampler. *The American Statistician, 46*(3), 167-174.

Chib, S., & Greenberg, E. (1995). Understanding the Metropolis-Hastings algorithm. *The American Statistician, 49*(4), 327-335.

Gelfand, A. E., Hills, S. E., Racine-Poon, A., & Smith, A. F. M. (1990). Illustration of Bayesian inference in Normal data models using Gibbs sampling. *Journal of the American Statistical Association, 85*(412), 972-985.

Jackman, S. (2000). Estimation and inference via Bayesian simulation: An introduction to markov chain monte carlo. *American Journal of Political Science, 44*(2), 375-404.

Liu, C. (1996). Bayesian robust multivariate linear regression with incomplete data. *Journal of the American Statistical Association, 91*(435), 1219-1227.

Lynn, R., & Vanhanen, T. (2001). National IQ and economic development: A study of eighty-one nations. *The Mankind Quarterly, 41*(4), 415-435.

Murphy. (1997). How to read the statistical methods literature: A guide for students. *The American Statistician, 51*(2), 155-157.

Seltzer, M. H., Wong, W. H., & Bryk, A. S. (1996). Bayesian analysis in applications of hierarchical models: Issues and methods. *Journal of Educational and Behavioral Statistics, 21*(2), 131-167.

Stigler, S. M. (1982). Thomas Bayes's Bayesian inference. *Journal of the Royal Statistical Society, Series A (General), 145*(2), 250-258.

Tanner, M. A., & Wong, W. H. (1987). The calculation of posterior distributions by data augmentation. *Journal of the American Statistical Association, 82*(398), 528-540.

Western, B., & Jackman, S. (1994). Bayesian inference for comparative research. *The American Political Science Review, 88*(2), 412-423.

Western, B. (1999). Bayesian analysis for sociologists: An introduction. *Sociological Methods and Research, 28*(1), 7-34.

Software

We will use various Bayesian-related software in this class. All are freeware/shareware that you will be able to download via the Internet.

R: A freeware/shareware version of the statistical program, *S-Plus*. Various sites have free manuals and guides to using *R*. Download from http://cran.r-project.org/

CODA: Diagnostic tools to assess the MCMC algorithms. Source code implemented in *R (S-Plus)*. Download from http://www.mrc-bsu.cam.ac.uk/bugs/classic/coda04/readme.shtml

BUGS: Bayesian MCMC algorithm program. Download from http://www.mrc-bsu.cam.ac.uk/bugs/

SCHEDULE FOR CLASSES & READINGS

Dates	Topic	Readings	Overheads*
1/07	• Introduction • Review of Probability Theory	G1 C1	C1.pdf
1/14	• Review of matrix algebra • Background of Bayes • Likelihood Inference	G2 Matrix1.pdf* Matrix2.pdf* Stigler.pdf*	C2.pdf Matrix1.pdf Matrix2.pdf
1/21	• Generalized Linear Model • An Introduction to R	G2 Murphy, 1997**	C3.pdf
1/28	• The Bayesian Setup	G3 Blyth, 1995**	C3.pdf
2/04	• The Normal & Student t Models	G4 Lynn.pdf* Morris 1977.pdf*	C4.pdf
2/11	• Priors	G5 Morris 1974.pdf* Best, 1996**	C5.pdf
2/18	• Model Selection	G6 Liu, 1996**	C6.pdf
2/25	• Hypothesis Testing	G7	C7.pdf
3/11	• MCMC Basics	G8 Tanner, 1987** (Nana) Gelfand, 1990**	C8.pdf
3/18	• MCMC Basics and Not-So-Basics	G9 Casella, 1992** Chib, 1995** Western Jackman.pdf* Jackman.pdf*	C9.pdf
3/25	• Diagnosing Convergence	Brooks Roberts.pdf	C9.pdf
4/01	• Hierarchical Linear Models	Seltzer, 1995	C9.pdf
4/08	• Limited Dependent Variables	G10	C10.pdf

4/15	• Latent Variables	Patz & Junker, 1999a (Kevin) Patz & Junker, 1999b	C11.pdf
4/22	• CANCELLED		
4/29	• Missing Data, Mixture Models	Maier, 2002	C12.pdf

*Available on Blackboard, under Course Documents (pdf format). **Available on Blackboard, under External Links