Syllabus for Statistics 157: Bayesian Statistics
University of California, Berkeley, Spring 2009

T/Th 11-12:30, 332 Evans Hall

Instructor: Dr. Cari Kaufman
Office: 315 Evans Hall
Office Hours: TBA
e-mail: cgk@stat.berkeley.edu
Phone: 510-643-0915

Course Overview

This is a seminar course on statistical inference from a Bayesian viewpoint, with an emphasis on computation. The Bayesian approach to statistics historically predates the "classical" or frequentist statistical methods you may have seen in other classes, but it did not gain widespread popularity until the introduction of new algorithms for sampling-based numerical integration, which have made it possible to fit more complicated Bayesian models. My main goals for this course are

- to help you gain a solid understanding of the basic Bayesian approach to inference, based on expressing all uncertainty in terms of conditional probability distributions,
- to introduce you to common practice in fitting and interpreting Bayesian models in applied problems, including hierarchical model specification and computational techniques,
- to give you practice in reading statistics articles from the literature and presenting statistical ideas to others.

Some Important Notes on the Structure of This Course

This course is structured differently than a traditional mathematical statistics course with weekly problem sets. Since it is a seminar, you will need to be an active participant in the course to do well. Practically speaking, each class period will require a specified advance reading, and most will require one or two problems to be completed prior to class. The class period itself will be a combination of a traditional lecture and a more interactive discussion of the reading and the assigned problems, with groups of students taking turns presenting some of the ideas. This course format does not work well without the involvement of all participants. For this reason, attendance is mandatory, and completion of the assigned problem(s) for each day is a key part of your grade.

Prerequisites

I assume you have had prior exposure to some linear algebra, calculus (univariate ok, preferably multivariate) and to probability at the level of Statistics 134. Exposure to statistical inference – for example, in Statistics 135 – is helpful but not required. We will be using the `R` language and environment for statistical computing, particularly in later parts of the course. If you have not used `R` before, you will need to work on your own to become familiar with the basics. If you have taken Statistics 133 (Concepts in Computing with Data) or are taking it concurrently with this course, you should be well equipped.

Textbooks and Journal Articles

Many readings will be taken from the required textbook:

J. Gill (2008). Bayesian Methods: A Social and Behavioral Sciences Approach, Second Edition. Chapman & Hall.

You may also find the following textbooks useful for reference. The first is on reserve in the Math/Stat library.

A. Gelman, J.B. Carlin, H.S. Stern and D.B. Rubin (2004). Bayesian Data Analysis, Second Edition. Chapman & Hall.

W. J. Braun and D. J. Murdoch (2007). A First Course in Statistical Programming with R. Cambridge University Press.

Copies of journal articles will be handed out during class, and electronic copies will be posted on bSpace. If you do not want paper copies, please let me know.

Grading

Your final grade will be a weighted average of grades in the following areas:

- 40% Completion of daily assignments
- 20% Midterm 1 - in class
- 20% Midterm 2 - take home data analysis project
- 20% Final Exam - cumulative

Tentative Schedule

Date	Topics	Readings
Jan. 20	Course Introduction, Some History	
Jan. 22	Probability Review, Bayes' Rule	Malakoff (1999), Gill 1.3-1.4
Jan. 27	Exponential Families, Likelihoods	Gill 2.1-2.2
Jan. 29	Prior and Posterior Distributions	Gill 2.3
Feb. 3	Conjugate Priors	Gill 5.1-5.3
Feb. 5	Models for Normal Data	Gill 3.1-3.4
Feb. 10	Multivariate Normal, Shrinkage	Gill 3.5, 5.8
Feb. 12	Bayesian Linear Models	Gill 4.1-4.2
Feb. 17	Informative and Noninformative Priors	Gill 5.4-5.5
Feb. 19	Subjective or Objective Bayes?	Berger (2006) or Goldstein (2006)
Feb. 24	MIDTERM 1	—
Feb. 26	Monte Carlo Integration	Gill 8.1-8.2
Mar. 3	Rejection and Importance Sampling	Smith and Gelfand (1992)
Mar. 5	Markov Chains	Gill 9.1-9.2
Mar. 10	The Gibbs Sampler	Casella and George (1992)
Mar. 12	Hierarchical Models	Gill 10.1-10.3
Mar. 17	Exchangeability; Linear Models Revisited	Gill 10.4-10.5
Mar. 19	More Complicated MCMC Algorithms	Gilks (1996)
Mar. 24	NO CLASS	—
Mar. 26	NO CLASS	—
Mar. 31	Practical MCMC, Part 1	Gill 12.1-12.2.2 (inclusive)
Apr. 2	Practical MCMC 2, MIDTERM 2 ASSIGNED	Gill 12.2.3.1, 12.2.3.2, 12.3
Apr. 7	Empirical Bayes	Gill 10.6
Apr. 9	Catch-up, MIDTERM 2 DUE	—
Apr. 14	Sensitivity Analysis	Gill 6.2
Apr. 16	Hypothesis Testing	Gill 7.1-7.2
Apr. 21	The Bayes Factor	Goodman (1999)
Apr. 23	Model Choice vs. Model Averaging	Gill 7.4, 7.5, 6.5
Apr. 28	Stochastic Variable Selection	George and McCulloch (1993)
Apr. 30	The Kalman Filter	Meinhold and Singpurwalla (1983)
May 5	Sequential Monte Carlo	Doucet et al. (2001)
May 7	Why Be Bayesian?	Gelman (2008)

Additional Course Policies

Email: Please post questions regarding course content to the bSpace forum, rather than emailing me directly, so that other students may benefit from the exchange as well.

Academic integrity: Any material submitted by you and that bears your name is presumed to be your own original work. I encouraged to work together on the readings and practice problems, but the actual writeup must be your own. Any evidence of cheating and plagiarism will be subject to disciplinary action.

References

Berger, J. (2006). The case for objective Bayesian analysis. *Bayesian Analysis*, 1:385–402.

Casella, G. and George, E. (1992). Explaining the Gibbs Sampler. *The American Statistician*, 46:167–174.

Doucet, A., deFreitas, N., and Gordon, N. (2001). An Introduction to Sequential Monte Carlo Methods. In Doucet, A., deFreitas, N., and Gordon, N., editors, *Sequential Monte Carlo Methods in Practice*. Springer.

Gelman, A. (2008). Objections to Bayesian Statistics. *Bayesian Analysis*, 3:445–450.

George, E. and McCulloch, R. (1993). Variable Selection via Gibbs Sampling. *Journal of the American Statistical Association*, 88:881–881.

Gilks, W. (1996). Full Conditional Distributions. In Gilks, W., Richardson, S., and Spiegelhalter, D., editors, *Markov Chain Monte Carlo in Practice*. Chapman & Hall.

Goldstein, M. (2006). Subjective Bayesian Analysis: Principles and Practice. *Bayesian Analysis*, 1:403–420.

Goodman, S. (1999). Toward Evidence-Based Medical Statistics. 2: The Bayes Factor. *Annals of Internal Medicine*, 130:1005–1013.

Malakoff, D. (1999). Statistics: Bayes Offers a 'New' Way to Make Sense of Numbers. *Science*, 286:1460.

Meinhold, R. and Singpurwalla, N. (1983). Understanding the Kalman Filter. *The American Statistician*, 37:123–127.

Smith, A. and Gelfand, A. (1992). Bayesian Statistics without Tears: A Sampling-Resampling Perspective. *The American Statistician*, 46:84–88.