36-724 Applied Bayesian Methods Spring 2002
Syllabus, Course Policies & Schedule

INSTRUCTOR:	**Michele DiPietro** Cyert Hall 108 `dipietro@stat.cmu.edu` 268-1287
TA:	**Philip Lee** Old Student Center 218 `plee@stat.cmu.edu` (412) 268-1889

OFFICE HOURS:	By appointment (Email is the best way to contact me)
TA OFFICE HOURS:	Tuesday 3:00 - 5:00 PM
WEB PAGE:	On BlackBoard `http://www.stat.cmu.edu/blackboard`
LECTURES:	Monday, Wednesday 11:30AM-12:50 PM CFA 212

COURSE DESCRIPTION

If you know Bayes' theorem you know (almost) everything there is to know about Bayesian statistics. Why then a whole course in Bayesian methods? For a long time, Bayesian statistics was nothing more than an intellectual exercise, not good for data analysis and shunned by "real" statisticians. It wasn't until the 1950s that the Bayesian approach started to gain some respect, thanks to Jeffreys, Savage, DeFinetti, and others. Even after it became more popular, the Bayesian approach was severely limited because of its computational unfeasibility. Whereas frequentist estimation relies on maximization of the likelihood, Bayesian estimation depends on the posterior distribution, which can be derived only through integration, a more challenging operation. With the advent of numerical integration and Markov Chain Monte Carlo (MCMC) techniques, the posterior distribution can be explored via simulations and very complex problems can be tackled. Therefore, this is an exciting time to study Bayesian statistics, both for the momentous advances and the research possibilities that are still open. The focus of the course will be on parametric methods, and even though the Bayesian approach can be embedded in a decision-theoretic framework, the emphasis of the course is decidedly applied (but still making use of mathematical statistics). Topics in this 7-week course will include:

- Philosophical considerations regarding the Bayesian and frequentist approaches
- Methods for deriving prior distributions, including noninformative priors
- Prior-posterior analysis, including Empirical Bayes methods, asymptotics, and
- Simulation methods, including Monte Carlo integration and MCMC
- Bayesian inference, including point estimation and credible intervals
- Diagnostics and model checking, including sensitivity analysis, Bayes factors and other model selection techniques
- Predictions

Objectives

Upon successful completion of the course, students should be able to:

- Build models for data
- Apply Bayesian inferential techniques to extract evidence from the data
- Write computer code to carry out required simulations
- Evaluate the adequacy of their model
- Communicate their results in writing to stakeholders

Prerequisites

I assume you are familiar with mathematical statistics at the level of 36-705 (Intermediate Statistics) and linear regression at the level of 36-707 (Linear Regression). This course also involves a fair amount of computing. While you are free to use any package or language you are comfortable with (Matlab, C, C++, BUGS or others), I will use S-plus in my handouts or in the code I will provide for you. For this reason, I will also assume familiarity with S-plus at the level of 36-711 (Statistical Computing).

Textbooks

There are no required textbooks. There are two *recommended* textbooks, the first for the theoretical part and the second for simulation methods:

Carlin, B. P., and Louis, T. A. (2000). *Bayes and Empirical Bayes Methods for Data Analysis.* Chapman and Hall.

Robert, C. P., and Casella, G. (1999). *Monte Carlo Statistical Methods.* Springer.

In addition, I will occasionally hand out in class some material from

Kass, R. E., and Wasserman, L. *A Short Course in Bayesian Statistics* (in preparation).

Course Policies

Attendance: Attendance is expected at all classes. Because there is no required textbook, the lectures will weave together different sources; students will be responsible for everything covered in class, therefore participation is in your best interest.

Homework: Six homework assignments will be handed out on Wednesdays; they will be due in class one week later (the following Wednesday).

Re-grade requests must be submitted in writing to the instructor, with a brief paragraph explaining the nature of the problem.

Collaboration: Collaboration among students is expected and even encouraged as an additional learning opportunity. However, the write-up of each problem has to be your own. My suggestion is to work on the assignment first and then ask around if you are stuck or to compare methods.

CHEATING AND PLAGIARISM: Cheating and/or plagiarism will not be tolerated. Please see the CMU Student Handbook, p. 7-8, for definitions of cheating and plagiarism, and the severe consequences of such behaviors or read the University Policy on Academic Integrity online at
`gollum.mac.cc.cmu.edu/univ_policy/documents/Cheating.html`

WRITING: Statisticians never work alone: they always analize data for other professionals, or are involved in article and grant writing (not to mention thesis writing!). Clear communication is crucial to interdisciplinarity and fund raising. For these reasons most homework assignments will include the writing of a report. You are expected to tie together all the evidence from your analysis in a coherent fashion, appropriate style and good grammar. Your grade will be affected by the clarity of your writing.

FINAL EXAM: There is no final exam.

SOFTWARE AND WEB RESOURCES

- I will assume everyone in the course has an account on the Statistics Department workstations. If you are not a member or guest of the department you should be able to get by with an Andrew account, but you will be somewhat on your own.
- BUGS (Bayesian inference Using Gibbs Sampling) is an S-Plus-like system for posterior simulation of hierarchical models. Information and documentation on BUGS, including download and install information is available at `http://www.mrc-bsu.cam.ac.uk/bugs/welcome.shtml`.
- Most of the data sets used in the Carlin and Louis book are available online at `http://www.biostat.umn.edu/~brad/data.html`.
- All the handouts, assignments and solutions for this course will be store on the class Web site, `http://www.cmu.edu/blackboard`.

Tentative Schedule

Date	Title	Topic
1/14	Through the looking glass	Bayes' Theorem; pros and cons of Bayesian and frequentist approach; examples
1/16	American π	Priors: elicitation, conjugate, and non-informative/reference
1/21	What do Martin Luther King and Thomas Bayes have in common?	
1/23	With a little help from my friends	Empirical Bayes; posterior analysis and inference
1/28	You can leave your hat on	Asymptotic approximation of posteriors: Bayesian central limit theorem and delta method
1/30	You are the weakest link. Goodbye.	Rejection sampling and other simulation methods
2/4	The Full Monty [Carlo]	Importance sampling and sampling importance resampling
2/6	[Markov] Chain of fools	MCMC methods: Gibbs sampler and Metropolis-Hastings algorithm
2/11	Oh Behave!	Improving and diagnosing convergence
2/13	Hey buddy, my Y's are up here!	Data augmentation and efficient simulation
2/18	Supermodel (You better work!)	Hierarchical models
2/20	May the Schwarz be with you!	Bayes factors and other model selection techniques
2/25	Does this model make my tails look fat?	Posterior predictive distributions and model checking
2/27	Steel magnolias	Robustness and sensitivity analysis

Graphic Organization of the Course