BAYESIAN METHODS
3-th term 2002

Instructor: Dr. Francesca Dominici

Office: Bloomberg School of Public Health, Room E-3148

Office phone: 410-6145107

e-mail: fdominic@jhsph.edu

Class Meetings: Tuesday-Thursday 10:30-12:00

Prerequisites: Introduction to Statistical Theory I,II.

TEXTBOOKS

Required: *Bayesian Data Analysis*, by A.Gelman,J.B. Carlin, H.S. Stern and D.Rubin, Chapman and Hall.

Helpful: *Markov Chain Monte Carlo in Practice*, by W.R Gilks, S.Richardson and D.J. Spiegelhalter, Chapman and Hall. *Bayes and Empirical Bayes Methods for Data Analysis*, by B.P. Carlin and T.A. Louis.

COURSE OUTLINE

This course is intended to illustrate the current approaches to Bayesian modeling and computation in statistics.

We begin with a description of simple familiar models such as those based on normal and binomial distributions, to illustrate concepts such us conjugate and non informative prior distributions. We then study more advanced tools in Bayesian modeling. A variety of models are considered, including linear regression, hierarchical models (random effects models), generalized linear models, mixed models. We discuss some aspects of modern Bayesian computational methods, including Markov Chain Monte Carlo methods such us the Gibbs Sampler and Metropolis Hastings algorithm, their implementation and monitoring.

Examples of real statistical analyses are investigated throughout the course.

Week #	Topics	Chap.
1	Setting up a probability model, Bayes'rule, posterior means and variances, binomial model	1 2.1-2.5
2	Standard univariate models including the normal model, conjugate and noninformative prior distributions	2.6-2.8
3	Multiparameters models, normal with unknown mean and variance, the multivariate normal distribution, multinomial models. Computation and simulation from arbitrary posterior distribution in two parameters	3
4	Hierarchical models, estimating populations parameters from the data, rat tumor rates application	5.1-5.3
5	Study design, missing data	7.1-7.5
6	Posterior simulation and integration Markov Chain Simulation	10.1-10.5 11.1-11.5
7	Linear models: simple and hierarchical formulation, SAT coaching example, selection of explanatory variables, hierarchical logistic regression model, Generalized linear models	13 14
8	Final week: student presentations	

INTENDED AUDIENCE

The course will be of interest to:

Masters and Ph.D students in biostatistics or other quantitative fields who recognize the need for advanced modeling tools in their thesis research.

Practicing statisticians and biostatisticians seeking to learn methods for analyzing complex data sets from industry, agriculture, biomedical science, public health, and other field from a Bayesian perspective.

METHOD OF STUDENT EVALUATION

The grade will be determined by a collection of homework problems and a final project.
The problems will include theoretical questions from the text, and several data analysis problems from various sources. Assignments will be made as appropriate throughout the course, and will generally be due two weeks after they are assigned.
The final project involves giving a brief classroom presentation of some subtopic of interest to the student.