Syllabus: STA 216 Fall 2008

Generalized Linear Models

Instructor: David Dunson
218 Old Chemistry,
dunson@stat.duke.edu

Teaching Assistant: P. Richard Hahn
hahn@stat.duke.edu

FIRST HALF (Prior to midterm 10/8):

1. **GLM basics**: components, exponential family, model fitting, frequent inference: analysis of deviance, stepwise selection, goodness of fit

2. **Bayesian inference in GLMs (basics)**: priors, posterior, comparison with frequentist approach, posterior computation, MCMC strategies (*Gibbs, Metropolis-Hastings*)

3. **Binary & categorical response data**:
   (a) **Basics**: link functions, form of posterior, approximations, Gibbs sampling via adaptive rejection
   (b) **Latent variable models**: Threshold formulations, probit models, discrete choice models, logistic regression & generalizations, data augmentation algorithms (*Albert & Chib + other forms*)

4. **Model uncertainty**: Bayesian formulation, variable selection, stochastic search, Bayesian hypothesis testing, applications to binary response models, shrinkage

5. **Survival analysis**: discrete hazard models, time-varying predictors, time-varying coefficients, smoothing, order-constraints

SECOND HALF

1. **Count data**: Poisson & over-dispersed Poisson log-linear models, prior distributions, applications

2. **Hierarchical GLMs**: random effects, generalized linear mixed models, longitudinal data analysis, hierarchical data structures, accounting for dependence in multivariate observations

3. **Mixtures of GLMs**: defining more flexible classes of models through mixtures of GLMs, Gaussian mixture models, linear regression with mixtures for residual distribution, mixtures of experts, computation via MCMC, label switching

4. **Latent class models**: Finite mixture models for random effects, flexible modeling of longitudinal data using latent class trajectory models
5. **Latent factor models**: factor analysis for multivariate Gaussian data, latent trait models for multivariate modeling of mixed discrete & continuous data, latent factor regression, structural equation models

6. **Missing data**: problem formulation, selection, pattern mixture & shared parameter approaches, examples, definitions of MCAR, MAR, ignorable missingness, accounting for missing predictors in GLMs

**Student Responsibilities:**

- **Assignments**: Outside reading and problems sets will typically be assigned after each class (10%)
- **Mid-term examination**: An in-class closed-book mid term examination will be given (30%)
- **Project**: Students will be expected to write-up and present results from a data analysis project (30%)
- **Final examination**: The final examination will be out of class (30%)

**Comments on Computing:**

- Course will have an applied emphasis
- Students will be expected to implement frequentist & Bayesian analyses of real and simulated data examples
- It is not required that students use a particular computing package
- Emphasis will be given to R/S-PLUS & Matlab, with code sometimes provided