University of California, Santa Cruz
Department of Statistics
Baskin School of Engineering

STAT 209: Generalized Linear Models (Fall 2019)

General course information

Instructor: Athanasios Kottas
Office: 523 Engineering 2 Building
Phone: 459-5536; E-mail: thanos@soe.ucsc.edu

Course web page: `https://stat209-fall19-01.courses.soe.ucsc.edu`

Lectures: Tuesday, Thursday 1:30pm–3:05pm (Baskin Engineering 169)
Office hours: Monday 3-4pm (or by appointment)

Course description and background: This is a graduate-level course on the theory, methods and applications of Generalized Linear Models (GLMs). Emphasis will be placed on statistical modeling, building from standard normal linear models, extending to GLMs, and briefly covering more specialized topics. With regard to inference, prediction, and model assessment, we will study both likelihood and Bayesian methods. In particular, within the Bayesian modeling framework, we will discuss practically important hierarchical extensions of the standard GLM setting.

Note that this is a course on methods for GLMs, rather than a course on using software for data analysis with GLMs. Students will be expected to be familiar with statistical software R, which will be used to illustrate the methods with data examples as part of the homework assignments (and exam). For data analysis problems involving Bayesian GLMs, you will be expected to write your own programs to fit Bayesian models using Markov chain Monte Carlo posterior simulation methods (R will suffice for this).

Knowledge of distribution theory and likelihood inference at the level of STAT 205B (or STAT 205) will be assumed. Also required is background on Bayesian modeling and computing at the level of STAT 206B, and preferably, STAT 207. Course STAT 208 provides relevant background, but is not strictly required.

Grading: The course grade will be based on homework assignments and (likely) a take-home final exam. Homework sets (and the exam) will include a combination of methodology and data analysis problems.

Tentative syllabus: We will cover topics from the following.

1. Introduction to GLMs

* Statistical modeling in the context of GLMs
* Exponential dispersion family of distributions (definitions, properties, and examples)
* Components of a GLM, examples of GLMs

2. Likelihood inference for GLMs

∗ Likelihood estimation (iterative weighted least squares) and inference (asymptotic interval estimates)

∗ Model diagnostics (residuals for GLMs, model comparison criteria)

3. Regression models for categorical responses

∗ Models for binary responses (dose-response modeling, probit and logit models)

∗ Poisson regression and log-linear models

∗ Basic ideas for modeling of contingency tables

∗ Multinomial response models for nominal or ordinal responses

4. Bayesian GLMs

∗ General setting, examples, priors for GLMs

∗ Markov chain Monte Carlo posterior simulation methods for GLMs

∗ Bayesian residual analysis and model choice

∗ Hierarchical GLMs, overdispersed GLMs, generalized linear mixed models

As time permits, we will briefly discuss ideas for: analysis of longitudinal and clustered data; generalized additive models; and spatial generalized linear models.

Reading/References: The lectures will be based on material taken from books and articles from the related literature. There is no required textbook. The course webpage will include relevant references as needed. Some handouts and notes will be provided.

Books that will be used for the lectures include:

• Agresti, A. (2002). *Categorical Data Analysis (Second Edition)*. Wiley.

• Agresti, A. (2007). *An Introduction to Categorical Data Analysis (Second Edition)*. Wiley.

• Johnson, V.E. and Albert, J.H. (1999). *Ordinal Data Modeling*. Springer.

• McCullagh, P. and Nelder, J.A. (1989). *Generalized Linear Models (Second Edition)*. London: Chapman & Hall.

Further references include:

• D. Dey, S.K. Ghosh, B.K. Mallick (editors) (2000). *Generalized Linear Models: A Bayesian Perspective*. (Biostatistics (New York, N.Y.), 5.) Marcel Dekker.

• Dobson, A.J. (2002). *An Introduction to Generalized Linear Models (Second Edition)*. Chapman & Hall.

• Fahrmeir, L. and Tutz, G. (1994). *Multivariate Statistical Modelling Based on Generalized Linear Models*. New York: Springer-Verlag.

• Gill, J. (2001). *Generalized Linear Models: A Unified Approach*. Series: Quantitative Applications in the Social Sciences, Sage University Papers, Thousand Oaks.

• Hastie, T.J. and Tibshirani, R.J. (1990). *Generalized Additive Models*. London: Chapman and Hall.

• Hoffmann, J.P. (2003). *Generalized Linear Models, An Applied Approach*. Pearson Allyn & Bacon.

• Jorgensen, B. (1997). *The Theory of Dispersion Models*. Chapman and Hall.

• Lindsey, J.K. (1997). *Applying Generalized Linear Models*. New York: Springer.

• Myers, R.H., Montgomery, D.C. and Vining, G.G. (2001). *Generalized Linear Models: With Applications in Engineering and the Sciences*. Wiley.