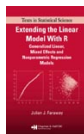


**Political Science 582: Quantitative Analysis in Political Science II, Fall 2011, Seigle Hall L016**

1. Course Description: This course extends what you did in previous methods courses by focusing more on nonlinear model forms. These are typically called "generalized linear models," although for historical reasons people in political science call them "maximum likelihood models." The principle we will care about is how to adapt the standard linear model that you know so that a broader class of outcome variables can be accommodated. These include: counts, dichotomous outcomes, bounded variables, and more. There is a strong theoretical basis for the models that we will use. Also, the bulk of the learning in the course will take place outside of the classroom by reading, practicing using statistical software, replicating the work of others, and doing problem sets. Keep in mind that the skills attained in this course are those that the discipline of political science expects of any self-declared data-oriented researcher.

The second aspect of the course is focused on the statistical package **R** which is completely free for downloading for Mac, Unix, Linux and that other platform at [CRAN](http://CRAN), the Comprehensive R Archive Network. **R** is an implementation of the **S** language, which is the default computational tool for research statisticians. Quite simply **R** is the most powerful, extensively featured, and capable statistical computing tool that has ever existed on this planet. And as mentioned, its free. We will not use Stata; don't ask.

2. Prerequisite Details: The only official prerequisite for this course is QPA II. However, each student should be familiar with: basic probability theory, statistical inference, hypothesis testing, and least squares estimation. The course will also assume a working knowledge of calculus and linear algebra at the level of *Essential Mathematics for Political and Social Research*. Jeff Gill, 2006, Cambridge University Press. Since students come to the course with varying levels of experience with statistical packages like **R**, some may spend quite a bit of time learning basic programming skills. If you suspect that you are in this group, it will pay to spend some time with a basic text such as *An R and S-Plus Companion to Applied Regression*. John Fox, 2002, Sage.
3. Course Grade: The final grade will be based on three components: problem sets (50%), a replication assignment (25%), and an exam (25%) on MLE theory and basic models. The problem sets will be a combination of analytical and computational assignments and given in each meeting. For the replication assignment, find a published work in your field of interest, obtain the data, and exactly replicate the author's model results. It is usually easier to find an article that uses the readily available datasets in the discipline (COW, ANES, GSS, etc.), but some authors are forthcoming about distributing their data if asked. The relevant model should be one of the nonlinear forms studied in this course. Gary King has some useful tips and links to his PS paper on the subject [here](#).
4. Office Hours: By appointment.
5. Incompletes: None given
6. Teaching Assistant: TBD<>. Office Hours: TBD
7. Homework: assigned each day and due the following week at classtime. No late homework accepted. All homework must be LaTeX'd.
8. Required Text:



Title: Extending the Linear Model with R: Generalized Linear, Mixed Effects and Nonparametric Regression Models.  
 Author: Faraway.  
 Publisher: Chapman & Hall/CRC.  
 Edition: First.  
 ISBN: 158488424.  
 Also: [Practical Regression and Anova using R](#).

9. Optional Texts (these are for background; see me before making any purchases):
  1. Title: A Guide to Econometrics.  
 Author: Kennedy.  
 Publisher: MIT Press, 2003.  
 Edition: Fifth or Sixth.  
 ISBN: 0-262-61183-X.
  2. Title: Generalized Linear Models: A Unified Approach.  
 Author: Gill.  
 Publisher: Sage, 2001.  
 Edition: First.  
 ISBN: 0761920552.
  3. Title: Modern Applied Statistics with S.  
 Author: Venables and Ripley  
 Publisher: Springer-Verlag, 2003.  
 Edition: Fourth.  
 ISBN: 0387954570.
  4. Title: An Introduction to R: Notes on R: A Programming Environment for Data Analysis and Graphics.  
 Author: R Development Core Team  
 Available (free) online [here](#)  
 Version 1.1, June 15, 2000
  5. Title: Linear Models with R.  
 Author: Faraway.  
 Chapman & Hall/CRC

Edition: First.  
ISBN: 1-58488-425-8.

12. List of Topics/Dates:

1. **September 9.** A Review of R Basics and the Linear Model.

*Reading:*

- Faraway, Chapter 1.
- [R Language Fundamentals](#).
- [How Not to Lie With Statistics](#), by Gary King and Ellie Powell.

*Homework:*

- [Problem Set #1](#).

2. **September 16.** Uncertainty, Inference, and Hypothesis Testing.

*Reading:*

- Leamer, [Let's Take the Con Out of Econometrics](#)
- McCloskey, [The Loss Function Has Been Mislaid: The Rhetoric of Significance Tests](#)
- Gill, [The Insignificance of Null Hypothesis Significance Testing](#)
- TPM (The Political Methodologist) Volume 11, No. 2, articles: (1) Jackman, (2) Anderson, et al., (3) Gill (pages 20-26). Available at: The Society for Political Methodology [webpage](#)

*Homework:*

- [Problem Set #2](#).

3. **September 23.** The Likelihood Model of Inference.

*Reading:*

- [Binomial PMF likelihood grid search](#).
- [Model syntax summary](#).

*Homework:*

- [Problem Set #3](#).

4. **September 30.** Models for Dichotomous Outcomes.

*Reading:*

- Faraway, Chapter 2.
- [Example Code](#).

*Homework:*

1. Faraway, Chapter 2, Exercises 1-7. For Exercise 2.2, download the wbca.txt data from <http://www.maths.bath.ac.uk/~jjf23/ELM/>. Also for Exercise 2.2, do **not** use the step function in part (b), use your own intuition).
2. Find a dataset with a dichotomous outcome that you are interested in. Run an appropriate glm model in R and submit the output with a paragraph defending the model fit.

5. **October 7.** Models for Count Outcomes.

*Reading:*

- Faraway, Chapter 3.
- [Poisson Example](#).
- [Negative Binomial Example](#).

*Homework:*

1. Faraway Chapter 3, Exercises 1-7.

6. **October 14.** Models for Contingency Tables.

*Reading:*

- Faraway, Chapter 4.
- [R code for Chapter 4](#),
- [contrasts](#).

*Homework:*

1. Faraway, Chapter 4, Exercises 1-7.

7. **October 21.** Models For Ordered and Unordered Categorical Data.

*Reading:*

- Faraway, Chapter 5.
- [R code for Chapter 5](#),

*Homework:*

1. Faraway Chapter 5, Exercises 1-6.
2. Consider a proportional odds model using the logit link function with only one explanatory variable in addition to the constant. Express the odds ratio (i.e. not-logged) for a one-unit change in the explanatory variable. What does this simplify to?

8. **October 28.** Exam On Fundamentals.

9. **November 4.** How to Handle Missing Data in Models. The EM Algorithm and Multiple Imputation.

*Reading:*

- [Multiple Imputation in R](#).
- [Detroit example in R](#).

*Homework:*

1. [Problem Set #4](#).

10. **November 11.** The GLM Theory and the Exponential Family Form.

*Reading:*

- Faraway, Chapter 6.

*Homework:*

1. Faraway Chapter 6, Exercises 1-5.

11. **November 18.** Other GLMs, Quasi-Likelihood Estimation.

*Reading:*

- Faraway, Chapter 7.

*Homework:*

1. Faraway Chapter 7, Exercises 1-7.

12. **November 25.** Holiday.

13. **December 2.** Random Effects.

*Reading:*

- Faraway, Chapter 8.

*Homework:*

1. Faraway Chapter 8, Exercises 1-9.
2. Turn in replications.

14. **December 9.** *Submission and Discussion of Replications.*