



Political Science 582: Quantitative Analysis in Political Science II, Fall 2013, Seigle Hall L016

1. Course Description: This course extends what you did in previous methods courses by focusing on nonlinear model forms for the outcome variable. 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](#), 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 I. 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. Knowledge of **R** is assumed.
3. Course Grade: The final grade will be based on three components: problem sets (40%), a replication assignment (30%), and an exam (30%) on MLE theory and basic models. The exam covers material from the first 7 weeks of the course plus the assigned readings (Faraway and articles). Consequently, we will discuss the readings in as much detail as the class desires. The problem sets will be a combination of analytical and computational assignments and given in each meeting. See Alicia Uribe's [tips](#) on success with the problem sets. 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](#), and a recent success story (publication) is described by two of his students . [All submitted work must be from LaTeX source](#).
4. Office Hours: Friday 9-10.
5. Incompletes: None given
6. Teaching Assistant: Taeyong Park. Office hours: 4:30-6:00 PM on Tuesdays in Seigle 278.
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:



Faraway Book

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.

10. List of Topics/Dates:

1. **August 30**. No Class (APSA Meeting).

Reading:

- Faraway, Chapter 1.
- [R Tutorial](#).
- [How Not to Lie With Statistics](#), by Gary King and Ellie Powell,
- 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
 - [R code from the lecture](#).
- Homework:
- [Problem Set #1](#).
2. **September 6.** 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 webpage](#)
 - Tressoldi et al. [High Impact = High Statistical Standards? Not Necessarily So](#)
 - Wetzels, et al., [Statistical Evidence in Experimental Psychology An Empirical Comparison Using 855 t Tests](#)
 - [R code from the lecture](#).
- Homework:
- [Problem Set #2](#).
3. **September 13.** The Likelihood Model of Inference.
- Reading:
- Faraway, Appendix A.
 - [Binomial PMF likelihood grid search](#).
 - [Model syntax summary](#).
 - [R code from the lecture](#).
- Homework:
- [Problem Set #3](#).
4. **September 20.** Models for Dichotomous Outcomes.
- Reading:
- Faraway, Chapter 2.
 - Altman, [The cost of dichotomising continuous variables](#).
 - [R Code for Chapter 2](#).
- 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. **September 27.** Models for Count Outcomes.
- Reading:
- Faraway, Chapter 3.
 - [Poisson Example](#).
 - [Negative Binomial Example](#).
 - [R code for Chapter 3](#),
- Homework:
1. Faraway Chapter 3, Exercises 1-7.
6. **October 4.** Models for Contingency Tables.
- Reading:
- Faraway, Chapter 4.
 - [R code for Chapter 4](#),
 - [contrasts](#).
- Homework:
1. Faraway, Chapter 4, Exercises 1-7.
7. **October 11.** 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 18.** No Class: Fall Break.
9. **October 25.** How to Handle Missing Data in Models. The EM Algorithm and Multiple Imputation.
- Reading:
- [Paper](#) by van Buuren and Groothuis-Oudshoorn,
 - [Multiple Imputation in R](#),
 - [R code](#) from the lecture.
- Homework:
1. [Problem Set #4](#).
10. **November 1. Exam On Fundamentals**
- Homework:
1. Turn in an electronic copy of your replication data and one regression model using these data (necessary to sit for the midterm).
11. **November 8.** The GLM Theory and the Exponential Family Form.
- Reading:
- Faraway, Chapter 6.
 - [R code from the lecture](#).
- Homework:
1. Faraway Chapter 6, Exercises 1-5.

12. **November 15.** Other GLMs, Quasi-Likelihood Estimation.

Reading:

- Faraway, Chapter 7.
- [R code from the lecture](#).

Homework:

1. Faraway Chapter 7, Exercises 1-7.

13. **November 22.** Nonparametric Regression, Additive Models.

Reading:

- Faraway, Chapters 11-12,
- [R code from the lecture](#).

Homework:

1. Faraway Chapter 11, Exercises 1-5; Faraway Chapter 12, Exercises 1-5.

14. **November 29.** Thanksgiving Holiday.

15. **December 6.** [Submission and Presentation of Replications](#).