COURSE DESCRIPTION: This course builds upon the analytical and applied foundations of PSC 404 and 405, taking the latter’s emphasis on the classical linear model as its point of departure. Because the classical linear regression model is inappropriate for data that arises in many interesting areas of political science, students need additional statistical tools in order to conduct rigorous empirical research. In this course, students will learn methods to analyze models and data for event counts, durations, censoring, truncation, selection, multinomial ordered/unordered categories, and strategic choices – in other words, all the other data out there. From time to time, we will also venture into semi-parametric methods, nonparametric methods, and machine learning, especially when those topics complement the MLE techniques we are studying.

A major goal of the course will be to teach students how to develop new models and techniques for analyzing issues they encounter in their own research. “Canned” statistical routines are often not appropriate for most of the micro-level models we develop as political science researchers. Students will therefore be required to program their own statistical routines (primarily in R).

PREREQUISITES: PSC 404 and 405, or the equivalent.

COURSE REQUIREMENTS: Course grades will be based on a series of homeworks (45%), a course paper (50%), and participation (5%). The exercises will consist primarily of programming and data analysis. Students are encouraged to work in groups of any size, so long as that size is no greater than two.

READINGS: Students are responsible for keeping up with the reading each week. I post my lecture notes and will provide links or copies of articles from time to time. In addition, students should read the appropriate chapters in the following, many of which are available in the star lab:

- The star lab introduction to R.
COURSE OUTLINE:

1. **R Programming and Monte Carlo Simulation**

2. **Maximum Likelihood Estimation**

3. **Count Models and Issues in Nonlinear Models**

   **Count Data**

   **Standard Errors and Confidence Intervals**

   **Interaction Terms in Nonlinear Models**

**Homework Reading**


4. **Duration Models**

   **Parametric Models**

**Cox Proportional Hazard Models**


**Grouped Binary Duration Data**


**Supplemental Reading on Splines**


• Ridgeway, Greg. “Splines.”


5. **Censoring and Truncation**


**Recommended**


6. **Selection Models**


Recommended:

Homework Reading:

7. Categorical Data and Random Utility Models

8. Strategic Models, part I

Bounded Rationality and the Quantal Response Equilibrium (QRE)

Private Information, Regression, and Misspecification

Recommended:

9. Model Discrimination

10. Strategic Models, part II

Bargaining Models

Signaling Models


• Penalized Estimators: Ridge Regression, LASSO, Adaptive LASSO
• Kenkel & Signorino working papers.
• Neural nets
• Recursive Partitioning and Regression Trees
• Random Forests

12. Parallel Computing in R

Important Dates:
Topic and Data OK’d Nov 6 (One paragraph)
Rough Draft Due Nov 25 (Tuesday before Thanksgiving Break)
Comments Returned Dec 2 (Tuesday after Thanksgiving Break)
In-Class Presentations Dec 9, 11 (Randomly assigned)
Final Paper Due 10:30am, Thursday, Dec 18