MLE Syllabus

PLSC 606J, Fall 2024

Instructor

- Dave Clark
- LNG 57; LN2430
- dclark@binghamton.edu
- clavedark
- office hours: M 1:30-3:30pm

Course details

- Fall 2024
- Wednesday
- 9:40-12:40
- LNG 332 Bing SSEL

Seminar Description

This 4 credit hour seminar is a survey of maximum likelihood methods and their applications to empirical political questions. It presumes students have a detailed and intuitive knowledge of least squares, probability theory, basic skills in scalar and matrix algebra, and a basic understanding of calculus. The course will deal mainly in understanding the principles of maximum likelihood estimation, under what conditions we move away from least squares, and what particular models are appropriate given observed data. The seminar will focus on application and interpretation of ML models and linking theory to statistical models. The course emphasizes coding and data viz in R and Stata.

The class meets one time per week for three hours. My office hours are designed to be homework help hours where I'll work in the grad lab with any of you who are working on the exercises. The most productive pathway for this class is for you to get in the habit of working together, and those office hours are a good time for this.

Course Purpose

This seminar fulfills the advanced quantitative methods requirement in the Ph.D. curriculum. The method of maximum likelihood underlies a majority of quantitative models in Political Science; this class teaches students to be astute consumers of such models, and how to implement and interpret ML models. These are crucial skills for dissertations in Political Science, and for producing publishable quantitative research.

Learning Objectives

Students will encounter an array of maximum likelihood models in this course. By the end of the course, students will have mastered the theory of maximum likelihood sufficient to write and program likelihood functions in R; they will be able to choose, estimate, and interpret appropriate models, model specifications, and model evaluation tools given their data; and they will be able to produce sophisticated quantities of interest (e.g. predicted probabilities, expected values, confidence intervals) via a variety of techniques including simulation and end point transformation. Students will also be able to present model findings verbally and graphically.

R Resources

There are lots of good, free R resources online. Here are a few:

- Modern Statistics with R
- R for Data Science
- R Markdown: The Definitive Guide
- R Graphics Cookbook
- Advanced R
- R Markdown Cookbook
- Data Science:A First Introduction
- The Big Book of R

Class Meetings, Office Hours, Assignments

The course will meet this fall entirely in-person in the Social Science Experiment Lab on Wednesdays 9:40am-12:40pm.

Office hours are Mondays 1:30pm-3:30pm. I'll likely hold these in the grad work room to help with your assignments. For an appointment, email me and we'll sort out a time.

All assignments should be turned in on Brightspace - please submit ::

- PDFs generated from *LaTeX* or R Markdown (Quarto).
- annotated R scripts.
- where necessary, data.

Assignments should be instantly replicable - running the code file should produce all models, tables, plots, etc.

Reading

The reading material for the course is important because it often demonstrates application of various MLE models; seeing how folks apply these and how they motivate their applications is really informative, something you cannot miss. We often won't directly discuss the readings, but don't let that imply they're not important. If I get the sense we're not keeping up with reading, expect the syllabus to change to incorporate quizzes or other accountability measures.

Reading for the course will consist of several books and articles (listed by week below). The books listed below also have Amazon links - you'll find most of these cheaper used online.

Required

- Box-Steffensmeier, Janet and Jones, Brad. 2004. Event History Modeling. Cambridge. ISBN 0521546737
- J. Scott Long. 1997. Regression Models for Categorical and Limited Dependent Variables. Sage Publications Inc. ISBN 0803973748
- Ward, Michael D. and John S. Ahlquist. 2018 Maximum Likelihood for Social Science. Cambridge. ISBN 978-1316636824.

Recommended

Useful, but not required (though some required reading in the first one):

- Gary King. 1998. Unifying Political Methodology. University of Michigan Press. ISBN 0472085549
- J. Scott Long. 2014. Regression Models for Categorical Dependent Variables Using Stata. 3rd Ed. Stata Press. ISBN 1597181110 (this book is good for practical/applied examples even if R is your primary language)

Gary King's book is regarded as seminal in developing ML applications in political science. Scott Long's is a similarlyaccessible treatment of a host of ML models and applications (and the Stata book is a great applied companion). Together, these two books are probably the most important on the syllabus as they are both accessible, but comprehensive and technical enough to be useful. Ward and Ahlquist's book is a new overview of applied ML in a political science setting. Box-Steffensmeier and Jones is a thorough and accessible treatment of hazard models in a variety of empirical settings.

Additional Resources

Other useful books include:

- Cameron, A. Colin and Trivedi, Pravin K. 1998. Regression Analysis of Count Data. Cambridge. ISBN 0521635675
- Maddala, Gregory. 1983. Limited Dependent and Qualitative Variables in Econometrics. Cambridge. ISBN 0521338255
- Paul D Allison - Event History Analysis : Regression for Longitudinal Event Data. Sage Publications Inc. ISBN 0803920555
- Tim Futing Liao - Interpreting Probability Models : Logit, Probit, and Other Generalized Linear Models. Sage Publications Inc. ISBN 0803949995
- John H Aldrich and Forrest D Nelson - Linear Probability, Logit, and Probit Models. Sage Publications Inc. ISBN 0803921330
- Fred C Pampel - Logistic Regression : A Primer. Sage Publications Inc. ISBN 0761920102
- Vani Kant Borooah - Logit and Probit : Ordered and Multinomial Models. Sage Publications Inc. ISBN 0761922423
- Scott R Eliason - Maximum Likelihood Estimation : Logic and Practice. Sage Publications Inc. ISBN 0803941072
- Richard Breen - Regression Models : Censored, Sample Selected, or Truncated Data. Sage Publications Inc. ISBN 0803957106
- Krishnan Namboodiri - Matrix Algebra : An Introduction. ISBN 0803920520

Course Requirements and Grades

The seminar requires the following:

- Problem sets - 60% total
- Mechanism papers - 40%

Please note that all written assignments must be submitted as PDFs either compiled in *LaTeX* or in R markdown (Quarto).

You'll complete a series of problem sets, mostly applied. How many will depend on how things move along during the term. Regarding the problem sets - the work you turn in for the problem sets should clearly be your own, but I urge you to work together - doing so is a great way to learn and to overcome problems.

The mechanism papers are a series of three short papers you'll write during the semester aimed at learning to identify and describe causal mechanisms, then at producing a causal mechanism. More on these early in the term.

A word about completeness - attempt everything. To receive a passing grade in the course, you must finish all elements of the course, so all problem sets, all exams, papers, etc. To complete an element, you must at least attempt all parts of the element - so if a problem set has 10 problems, you must attempt all 10 or the assignment is incomplete, you've not completed every element of the course, and you cannot pass. I realize there may be problems you have trouble with and even get wrong, but you must try - the bottom line is don't turn in incomplete work.

Grades will be assigned on the following scale:

Grade	Range	Grade	Range
A	94-100%	C+	77-79%
A-	90–93%	C	73-76%
B+	87–89%	C-	70-72%
B	83-86%	D	60-69%
B-	80-82%	F	<60%

Course Policies

Attendance

Attendance is expected, and is essential if you're to succeed in this class.

Academic Integrity

Ideas are the currency in academic exchange, so acknowledging where ideas come from is important. Acknowledging the sources of ideas also helps us identify an idea's lineage which can be important for understanding how that line of thought has developed, and toward promoting future growth. As graduate students, you should have a good understanding of academic honesty and best practices. Here are details of Binghamton's honesty policy.

Course Schedule

Week 1, Aug 21 – Binary y Variables I - probit/logit, QI

- Ward & Alhlquist, 2018. *Maximum Likelihood for Social Science.* Chapter 1, 2, 4

- J. Scott Long. 1997. *Regression Models for Categorical and Limited Dependent Variables.* Chapter 3.

Week 2, Aug 28 – Likelihood Theory and ML Estimation

- Gary King. 1998. *Unifying Political Methodology.* Chapter 1-4
- J. Scott Long. 1997. *Regression Models for Categorical and Limited Dependent Variables}. Chapters 1-2.

Week 3, Sept 4 – Binary y Variables II - symmetry, fit, diagnostics, prediction

- Ward & Alhlquist, 2018. *Maximum Likelihood for Social Science.* Chapter 3, 5, 6, 7

- Nagler (1994)
- King and Zeng (2001)
- Franklin and Kosaki (1989)
- C. Zorn (2005)

Week 4, Sept 11 – Binary y Variables III (discrete hazards)

- Ward & Alhlquist, 2018. *Maximum Likelihood for Social Science.* Chapter 11
- Beck, Katz, and Tucker (1998)
- Carter and Signorino (2010)

Week 5, Sept 18 – Binary y Variables IV - variance, order

- J. Scott Long. 1997. *Regression Models for Categorical and Limited Dependent Variables.* Chapter 5.
- Franklin (1991)
- Alvarez and Brehm (1995)
- Clark and Nordstrom (2005)

Week 6, Sept 25 – Assumptions and Specification - interactions, functional form, measurement of y

- J. Scott Long. 1997. *Regression Models for Categorical and Limited Dependent Variables.* Chapter 5.
- Clark, Nordstrom, and Reed (2008)
- Clarke and Stone (2008)
- Berry, Golder, and Milton (2012)
- Brambor, Clark, and Golder (2006)

Week 7, Oct 2 – No class, Yom Kippur

Week 8, Oct 9 – Choice Models I (Unordered y Variables) - MNL, MNP, CL (IIA)

- Ward & Alhlquist, 2018. *Maximum Likelihood for Social Science.* Chapter 9

- J. Scott Long. 1997. *Regression Models for Categorical and Limited Dependent Variables.* Chapter 6.
- Alvarez and Nagler (1998)
- Lacy and Burden (1999)
- C. J. W. Zorn (1996)

Week 9, Oct 16– Choice Models II (Unordered Dependent Variables continued, and systems of eqs, ordered)

- Ward & Alhlquist, 2018. *Maximum Likelihood for Social Science.* Chapter 8
- J. Scott Long. 1997. *Regression Models for Categorical and Limited Dependent Variables.* Chapter 5.
- Franklin and Kosaki (1989)

Week 10, Oct 23 – Event Count Models I - poisson, dispersion

- Ward & Alhlquist, 2018. *Maximum Likelihood for Social Science.* Chapter 10
- J. Scott Long. 1997. *Regression Models for Categorical and Limited Dependent Variables.* Chapter 8.1, 8.2.
- Gowa (1998)
- Fordham (1998)

Week 11, Oct 30 – Event Count Models II - negative binomial, zero-altered

- J. Scott Long. 1997. *Regression Models for Categorical and Limited Dependent Variables.* Chapter 8.3-8.7.
- C. J. W. Zorn (1998)
- Clark (2003)

Week 12, Nov 6 – no class, Peace Science

Week 13, Nov 13 – Continuous Time Hazard Models I - parametric, semi-parametric models

- Ward & Alhlquist, 2018. *Maximum Likelihood for Social Science.* Chapter 11
- Janet Box-Steffensmeier and Brad Jones. 2004. *Event History Modeling.* Chs. 1-4
- J. M. Box-Steffensmeier, Arnold, and Zorn (1997)

Week 14, Nov 20 – Continuous Time Hazard Models II - parametric models, special topics

- Janet Box-Steffensmeier and Brad Jones. 2004. *Event History Modeling.* Chs. 5-11
- C. J. W. Zorn (2000)
- Bennett and Stam (1996)
- J. Box-Steffensmeier, Reiter, and Zorn (2003)

Week 15, Nov 27 – no class, Thanksgiving

Week 16, Dec 4 – Censored/Truncated Variables, Samples - selection models - J. Scott Long. 1997. *Regression Models for Categorical and Limited Dependent Variables}. Chapter 7

- Reed (2000)
- Signorino (1999)
- Timpone (1998)

References

Alvarez, R. Michael, and John Brehm. 1995. "American Ambivalence Towards Abortion Policy: Development of a Heteroscedastic Probit Model of Competing Values." *American Journal of Political Science* 39 (4): 1055–82.

Alvarez, R. Michael, and Jonathan Nagler. 1998. "When Politics and Models Collide: Estimating Models of Multiparty Elections." *American Journal of Political Science* 42 (1): 55–96.

Beck, Nathaniel, Jonathan M. Katz, and Richard Tucker. 1998. "Beyond Ordinary Logit: Taking Time Seriously in Binary Time-Series Cross-Section Models." *American Journal of Political Science* 42 (4): 1260–88.

Bennett, D. Scott, and Allan C. Stam. 1996. "The Duration of Interstate Wars, 1816-1985." *American Political Science Review* 90: 239–57.

Berry, William D, Matt Golder, and Daniel Milton. 2012. "Improving Tests of Theories Positing Interaction." *Journal of Politics* 74 (3): 653–71.

Box-Steffensmeier, Janet M., Laura W. Arnold, and Christopher J. W. Zorn. 1997. "The Strategic Timing of Position Taking in Congress: A Study of the North American Free Trade Agreement." *American Political Science Review* 91 (2): 324–38.

Box-Steffensmeier, Janet, Dan Reiter, and Christopher J. W. Zorn. 2003. "Nonproportional Hazards and Event History Analysis in International Relations." *Journal of Conflict Resolution* 47 (1): 33–53.

Brambor, T., W. R. Clark, and M. Golder. 2006. "Understanding Interaction Models: Improving Empirical Analyses." *Political Analysis* 14 (1): 63–82.

Carter, David B, and Curtis S Signorino. 2010. "Back to the Future: Modeling Time Dependence in Binary Data." *Political Analysis* 18 (3): 271–92.

Clark, David H. 2003. "Can Strategic Interaction Divert Diversionary Behavior? A Model of US Conflict Propensity." *Journal of Politics* 65 (4): 1013–39.

Clark, David H., and Timothy Nordstrom. 2005. "Democratic Variants and Democratic Variance: How Domestic Constraints Shape Interstate Conflict." *Journal of Politics* 67 (1): 250–70.

Clark, David H., Timothy Nordstrom, and William Reed. 2008. "Substitution Is in the Variance: Resources and Foreign Policy Choice." *American Journal of Political Science* 52 (4): 763–73.

Clarke, K., and R. Stone. 2008. "Democracy and the Logic of Political Survival." *American Political Science Review* 102 (3).

Fordham, Benjamin. 1998. "Partisanship, Macroeconomic Policy, and u.s. Uses of Force, 1949-1994." *Journal of Conflict Resolution* 42 (4): 418–39.

Franklin, Charles H. 1991. "Eschewing Obfuscation? Campaigns and the Perception of u.s. Senate Incumbents." *American Political Science Review* 85 (4): 1193–1214.

Franklin, Charles H., and Liane C. Kosaki. 1989. "Republican Schoolmaster: The u.s. Supreme Court, Public Opinion, and Abortion." *American Political Science Review* 83 (3): 751–71.

Gowa, Joanne. 1998. "Politics at the Water's Edge: Parties, Voters and the Use of Force Abroad." *International Organization* 52 (2): 307–25.

King, G., and L. Zeng. 2001. "Explaining Rare Events in International Relations." *International Organization* 55 (03): 693–715.

Lacy, Dean, and Barry C. Burden. 1999. "The Vote-Stealing and Turnout Effects of Ross Perot in the 1992 u.s. Presidential Election." *American Journal of Political Science* 43 (1): 233–55.

Nagler, Jonathan. 1994. "Scobit: An Alternative Estimator to Logit and Probit." *American Journal of Political Science* 38 (1): 230–55.

Reed, William. 2000. "A Unified Statistical Model of Conflict Onset and Escalation." *American Journal of Political Science* 44 (1): 84–93.

Signorino, Curtis S. 1999. "Strategic Interaction and the Statistical Analysis of International Conflict." *American Political Science Review* 93 (2): 279–98.

Timpone, Richard. 1998. "Structure, Behavior, and Voter Turnout in the United States." *American Political Science Review* 92: 145–58.

Zorn, Christopher. 2005. "A Solution to Separation in Binary Response Models." *Political Analysis* 13 (2): 157–70.

Zorn, Christopher J. W. 1996. "Choice Models in Theory and Practice: Independence of Irrelevant Alternatives and the Decision to Litigate."

———. 1998. "An Analytic and Empirical Examination of Zero-Inflated and Hurdle Poisson Specifications." *Sociological Methods and Research* 26 (3): 368–400.

———. 2000. "Modeling Duration Dependence." *Political Analysis* 8 (4): 367–80.