POL 2519
Quantitative Methods and Data Analysis
Fall 2025

PREREQUISITE: POL2504 and POL 2507, or equivalent.

Instructor: Dr. Mark Nieman, mark.nieman@utoronto.ca

Time and Location: M 2–4pm, see Acorn/Quercus

Student Hours: M 11:45am–1:45pm, Sidney Smith 3018

Overview and Objectives

This class introduces a variety of statistical techniques for limited and categorical dependent variables relevant to political science research. The course builds on materials from POL 2504 and POL 2507, designed for PhD students advancing their methodological skills and applying statistical methods to their own research. The course covers a wide variety of estimators, including those for binary, ordered, count, multivariate, and temporal outcome variables. We also discuss issues relating to censoring and non-random sample selection. Additional and related topics will be covered as necessary.

The primary course objective is for students to understand how, when, and why to use various estimators in applied research. A large portion of class time (approximately 40% of classes) is devoted to hands-on use and interpretation of relevant estimators using computational methods. All sample code for assignments are written in the statistical program R. Some familiarity with the concepts of probability theory, linear algebra, and linear regression is assumed, but these topics are reviewed as necessary.

Required Texts

Long, J. Scott. 1997. *Regression Models for Categorical and Limited Dependent Variables*. Sage Publications.

Recommended:

King, Gary. 1989. *Unifying Political Methodology: The Likelihood Theory of Statistical Inference*. Ann Arbor, MI: University of Michigan Press.

Software

- R (latest version) https://www.r-project.org/
- RStudio (latest version) https://rstudio.com/products/rstudio/download/
- R and RStudio frequently put out new versions; we will keep up by periodically checking and updating our software (to avoid various errors and inconsistencies).

Grading Scale

Final grades are assigned in accordance with the university grading scale.

Requirements

Grades are based on homework assignments and classroom participation.

- **Homework Assignments**: 80%

  A total of 5 graded homework assignments (equally weighted) are assigned throughout the semester. *No late assignments will be accepted, unless agreed on prior to the deadline.* The only exceptions/arrangements will be made for cases of extreme adversities. Students should complete assignments using resources described, discussed, and applied from class; e.g., use only those R packages discussed in class.

  Students can choose to complete these assignments on their own or work together in groups of two. Assignments completed in groups must clearly list the names of each student in the group (each student can submit a group assignment individually, or a designated group member may submit the assignment for the group). All group members will receive the same grade. Whether a student completes an assignment individually or as a part of the group, they are responsible for understanding every part of the answer they provided. *In the case where there is a disconnect between a student's performance on homeworks and their in-class participation, the instructor will schedule an oral assessment, during which the student will be asked to explain some of the answers given in homeworks.*

  Students have an option to re-submit *one assignment within one week after the assignment was graded.* Students can only re-submit an assignment that received a grade lower than an "A-". All assignments must be re-submitted via email to the instructor with a short description of how the assignment has been improved. This option is not available for missed assignments or those that received a grade of 0.

- **Classroom Participation**: 20%

  Classroom contribution is dependent on a student's attendance, positive and thoughtful contribution to discussion, and regular short pass/no pass in-class problem sets.

Course Policies

Student Responsibilities in the Learning Process: Students are expected to complete all required readings on a topic prior to completing that topic's assessment and complete all assessments on time. This means accessing the materials with sufficient time to complete assessments prior to deadlines. In the event that a student has questions concerning the material, they should formulate specific questions to ask via office hours or email with sufficient time for a response prior to assessment deadlines (i.e., questions should be sent at least 24 hours prior to a deadline, excluding weekends).

Classroom Conduct: Students are expected to participate in class in a thoughtful and respectful manner while in the pursuit of knowledge accumulation. Generally, this means engaging with one another's ideas and treating others as one would like to be treated, as well as *not* treating others how one would *not* like to be treated. Please see university policies on freedom of speech and discrimination and harassment.

Accommodations: Please discuss any special needs with the instructor start of the semester, for example to request reasonable accommodations if an academic requirement conflicts with religious practices and/or observances. Those seeking accommodations based on disabilities should complete the appropriate documentation with Student Life Programs and Services.

Academic Misconduct: All acts of dishonesty in any work constitute academic misconduct; please see the University's guidelines—including ways to avoid inadvertent misconduct—and rules of procedures regarding misconduct. The Student Disciplinary Regulations will be followed in the event of academic misconduct.

A special note on plagiarism. Plagiarism is the act of representing, directly or indirectly, another person's work as one's own. It may involve presenting someone's work, wholly or partially; quoting without acknowledging the source material; copying and submitting another person's work (including code); and similar infractions.

Copyright: Course materials, including recorded lectures and slides, are the instructor's intellectual property covered by the Copyright Act, RSC 1985, c C-42. Course materials posted on Quercus are for registered students only and may not be posted to other websites or media without the express permission of the instructor. Unauthorized reproduction, copying, or use of online recordings constitute copyright infringement.

Course Outline

Week 1: Binary Outcome Variables I

Welcome and introduction; Linear probability model; Logistic regression.

Long, Ch 1–3.

Berry, William D., Jacqueline HR DeMeritt, and Justin Esarey. 2010. Testing for Interaction in Binary Logit and Probit Models: Is a Product Term Essential? *American Journal of Political Science* 54(1): 248–266.

Week 2: Binary Outcome Variables II

Latent variable approach; Probit regression; Model fit, comparison, and diagnostics.

Long, Ch 4.

Zorn, Christopher. 2005. A Solution to Separation in Binary Response Models. *Political Analysis* 13(2): 157–170.

Week 3: Ordered Outcome Variables

Ordered logit and probit; Parallel regression assumption.

Long, Ch 5.

Week 4: Multinomial Outcome Variables

Multinomial logit; Independent of Irrelevant Alternatives assumption; Multinomial probit; Conditional logit.

Long, Ch 6.

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.

Week 5: Count Outcome Variables

Poisson regression; Negative binomial model; Model comparison and fit.

Long, Ch 8, Sections 1–2 (Poisson), Sections 3–7 (negative binomial).

King, Gary, 1988. Statistical Models for Political Science Event Counts: Bias in Conventional Procedures and Evidence for the Exponential Poisson Regression Model. *American Journal of Political Science* 32: 838–863.

Week 6: Temporal Dependence

Duration analysis; Temporal splines and polynomials with binary outcomes.

Box-Steffensmeier, Janet M., and Bradford S. Jones. 1997. Time is of the Essence: Event History Models in Political Science. *American Journal of Political Science* 41(4): 1414–1461.

Carter, David B. and Curtis S. Signorino. 2010. Back to the Future: Modeling Time Dependence in Binary Data. *Political Analysis* 18(3): 271–292.

Williams, Laron K. 2018. Temporal Dependence and the Sensitivity of Quantities of Interest: A Solution for a Common Problem. *International Studies Quarterly* 62(4): 892–902.

Week 7: Non-Constant Variance with Latent Variables

Heterskedastic probit; Heteroskedastic ordered and multinomial models.

Williams, Richard. 2009. Using Heterogeneous Choice Models to Compare Logit and Probit Coefficients Across Groups. *Sociological Methods & Research* 37(4): 531–559.

Week 8: Correlated Outcomes I

Bivariate probit; Multivariate probit.

Greene, William. 1993. Bivariate and Multivariate Probit. *Econometric Analysis*. Upper Saddle River, NJ: Prentice-Hall, 2nd Ed. Chapter 21, Section 6.

Week 9: Correlated Outcomes II

Recursive bivariate probit; Partial observability models; Censored probit.

Kimball, Anessa L. 2006. Alliance Formation and Conflict Initiation: The Missing Link. *Journal of Peace Research* 43(4): 371–389.

Xiang, Jun. 2010. Relevance as a Latent Variable in Dyadic Analysis of Conflict. *Journal of Politics* 72(2): 484–498.

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

Week 10: Selection on Unobservables

Types of missing data; Data censoring and truncation; Sample selection models.

Long, Ch 7.

Hug, Simon. 2003. Selection Bias in Comparative Research: The Case of Incomplete Data Sets. *Political Analysis* 11(3): 255–274.

Nieman, Mark David, Olga Chyzh, and Douglas M. Gibler. 2023. Modeling Structural Selection in Disaggregated Event Data. *Unpublished manuscript.*

Week 11: Strategic Models

Empirical implications of theoretical models; Strategic selection processes.

Signorino, Curtis S. 1999. Strategic Interaction and the Statistical Analysis of International Conflict. *American Political Science Review* 93(2): 279–297.

Signorino, Curtis S. 2003. Structure and Uncertainty in Discrete Choice Models. *Political Analysis* 11(4): 316–344.

Kenkel, Brandon and Curtis S. Signorino. 2014. Estimating Extensive Form Games. *R. Journal of Statistical Software* 56(8): 1—27.

Nieman, Mark David. 2018. Strategic Binary Choice Models with Partial Observability. *Statistical Sinica* 28(4): 2089-–2105