

# PSCI 7155: Maximum Likelihood

University of Colorado Boulder

Fall 2018

**Time:** Tuesday, 4:35-7:05  
**Location:** KTCH 1B31  
**Instructor:** Dr. Andrew Q. Philips  
**Office:** KTCH 144  
**Email:** andrew.philips@colorado.edu  
**Office hours:** Monday 1:00-2:45 or by appointment

**COURSE DESCRIPTION:** Models utilizing maximum likelihood (ML) estimation are ubiquitous in political science. This course is designed to introduce you to a variety of ML estimators that you are likely to encounter in your research. We will start with a simple overview of probability models and show why we might need such an estimation technique. After discussing likelihood theory, we then move to more “novel” types of data, such as linear models with censoring and truncation, binary choice models, count models, and duration models. Much of the class will be spent on learning how to interpret these models and implement them in a statistical program. The last two weeks are special topics that we will decide on during the first weeks of class, and may cover topics such as bounded and compositional data, interpreting linear and non-linear interactions, and imputing missing data. These last topics will be chosen based on time constraints and class interests.

By the end of this course you should be able to:

- Understand what is going on “under the hood” when estimating maximum likelihood models.
- Identifying which parametric model best fits your research question and the underlying data-generating process.
- Be able to interpret a variety of ML models and present your results using predictions, substantive interpretations, marginal effects, simulations, etc.
- Apply what you have learned to your own research.

**PREREQUISITES:** This an graduate level course; students should have a background in at least introductory (and preferably) advanced regression statistics (i.e., Data I and Data II). Students from other graduate programs must check with me before signing up for this course.

**SOFTWARE:** We will use R in this course. Although familiarity with R is not necessary, it is a plus. Those unfamiliar with this program may want to purchase or borrow the suggested textbooks that cover working with R, although there are copious amounts of information available for free online. Please download both R (<https://cran.r-project.org/>) and RStudio (<https://www.rstudio.com/>) before the first class session.

**GRADES:** Course grades will be based on the following. Participation is worth 15% of the final grade. An in-class presentation of the student’s original research paper is worth 10%, as is the colleague critique. About halfway through the semester, a mid-term exam will be given that is worth 20% of the final grade. Last, the original research paper will comprise 45% of the final grade. Note that there are no opportunities for extra credit, nor is there a final exam.

Participation	15%
In-Class Presentation	10%
Colleague Critique	10%
Midterm Exam	20%
Original Research Paper	45%

The following scale will be used to turn numerical grades into letter ones. Note that I will round up a letter should your grade fall on the number (but on or above 0.5) between two letters (e.g., 89.5 up to 90 rounds up to an A-).

#### Grade Scale

A	95-100
A-	90-94
B+	87-89
B	84-86
B-	80-83
C+	77-79
C	74-76
C-	70-73
D+	67-69
D	64-66
D-	60-63
F	0-59

**PARTICIPATION:** Participation is an integral component of graduate courses. Students are expected to come to every class *having already read the assigned readings for that day*, and should be prepared to discuss them. Graduate-level courses only are successful when all participate actively in the discussion.

**MIDTERM EXAM:** About halfway through the semester there will be a mid-term exam. There are two portions to this. First is an in-class, open book exam, where you are free to consult your notes, readings, etc. The second part is a take-home applied exam which you should use R or other programs to produce your answer. Although you may consult help files, notes, and the internet for the second portion of the midterm, *you cannot consult with other graduate students*. This will be considered a form of cheating.

**ORIGINAL RESEARCH PAPER:** By the end of the class, students should have a solo-authored research article-length working paper that either: a.) a purely methodological paper (less common), or b.) an applied research paper that utilizes at least one of the maximum likelihood models discussed in the course. Since one semester is a short time in which to write such a paper, it is crucial to get some semblance of a research topic early in the semester. I encourage students to meet with me early on to discuss potential topics. Half-way through the semester, there will be a mandatory “check-in” to ensure all students have a feasible research topic. I am open to this paper being submitted for credit in another course, or a substantial re-write from a previous semester, but this needs to be cleared by me (and by the other professor). Note that the most important part of the research paper for the purposes of this class is the theory, hypotheses, research design, and presentation and interpretation of the results. Students are strongly encouraged to write the paper using  $\LaTeX$ .

**COLLEAGUE CRITIQUE:** After all research papers are due, I will assign each student to read and critique *two* of their fellow students’ papers. Students should treat this just as they would an invitation

to peer review a manuscript. Theory, research design, data, methods, etc,... should be critiqued.

**IN-CLASS PRESENTATIONS:** Instead of a final exam, students will give in-class presentations on their research papers. This will be held after the colleague critiques are due, so students will have some time to make changes as necessary. These will be presented in a conference-style format, with the time length determined by the number of students in the course.

**ATTENDANCE AND LATE POLICY:** Attendance is a key component of succeeding in graduate school. I provide slides for each class, but we will have a much more comprehensive discussion than what appears on the slide. Attendance is mandatory.

Assignments are due on the day listed in the syllabus. Late assignments will not be accepted.

**REQUIRED TEXTS:** The following texts are required for the course. Any additional readings will be made available to you on the first day of class or as needed.

- Long, J. Scott. 1997. *Regression models for categorical and limited dependent variables*. Thousand Oaks, CA. Sage.
- Eliason, Scott R. 1993. *Maximum likelihood estimation: Logic and practice*. Thousand Oaks, CA. Sage.
- King, Gary. 1998. *Unifying political methodology: The likelihood theory of statistical inference*. University of Michigan Press.

Note that it is expected to read the week's required readings before coming to class.

**RECOMMENDED TEXTS:** The following texts are not required, but may be helpful to some. In the schedule below there are additional texts in the "suggested readings".

- Box-Steffensmeier, Janet M., and Bradford S. Jones. 2004. *Event history modeling: A guide for social scientists*. Cambridge University Press.
- Cameron, Adrian Colin and Pravin K. Trivedi. 2013. *Regression analysis of count data*. Second Edition. Cambridge University Press.
- Cameron, Adrian Colin, and Pravin K. Trivedi. 2009. *Microeconometrics using Stata Volume 5*. College Station, TX: Stata Press.
- Gelman, Andrew, and Jennifer Hill. *Data analysis using regression and multilevel/hierarchical models*. Cambridge University Press, 2006.
- Hosmer, D., Stanley Lemeshow, and Susanne May. 2008. *Applied survival analysis: Regression modeling of time to event data*. Second Edition. New York: Wiley.
- Long, J. Scott and Jeremy Freese. 2014. *Regression models for categorical dependent variables using Stata*. Third Edition. Stata Press.
- Monogan, James E. 2015. *Political analysis using R*. Springer.
- Philips, Andrew Q. 2016. "L<sup>A</sup>T<sub>E</sub>X: A brief introduction". Available [here](#).
- Forbes, Catherine, Merran Evans, Nicholas Hastings, and Brian Peacock. 2010. *Statistical distributions*. Fourth Edition. Wiley.
- Greene, William. 2011. *Econometric analysis*. 7th Edition. Upper Saddle River, NJ: Prentice-Hall.

## TENTATIVE SCHEDULE:

### Week 1: Introduction to Probability Models

Required Readings:

- King Chapters 1 and 2
- Eliason Chapter 1 (pp. 1-18)
- Long Chapter 1

Suggested Readings:

- Burns, Patrick. 2011. *The R Inferno*.
- Philips, Andrew Q. 2018. "R: A brief introduction"

### Week 2: Estimation: Looking Under the Hood

Properties of ML models,

Required Readings:

- King Chapter 3
- Eliason Chapters 1 (pp. 18-21), and 3

### Week 3: Generalized Linear Model

Required Readings:

- King Chapter 4
- Eliason Chapter 2

### Week 4: Censoring and Truncation, Selection Models and More

Required Readings:

- Eliason Chapter 5
- Reed, William, and David H. Clark. 2000. "War initiators and war winners: The consequences of linking theories of democratic war success." *Journal of Conflict Resolution* 44(3): 378-395.

Suggested Readings:

- White, Halbert. 1980. "A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity." *Econometrica: Journal of the Econometric Society*: 817-838.
- Tobin, James. 1958. "Estimation of relationships for limited dependent variables." *Econometrica: Journal of the Econometric Society*: 24-36.
- Sigelman, Lee and Langche Zeng. 2000. "Analyzing censored and sample-selected data with Tobit and Heckit models." *Political Analysis* 8(2): 167-182.

### Week 5: Binary Choice Models

Logit and probit Required Readings:

- Long Chapter 3
- King Chapter 5

## Week 6: Binary Choice Models Continued

Heteroskedastic probit, binary choice interpretation, and more

Required Readings:

- 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): 1055-1082.
- Ai, Chunrong, and Edward C. Norton. 2003. "Interaction terms in logit and probit models". *Economics Letters* 80:123-129.
- Mood, Carina. 2010. "Logistic regression: Why we cannot do what we think we can do, and what we can do about it." *European Sociological Review* 26:67-82.

Suggested Readings:

- Greenhill, Brian, Michael D. Ward, and Audrey Sacks. 2011. "The separation plot: A new visual method for evaluating the fit of binary models." *American Journal of Political Science* 55(4): 991-1002.
- Nagler, Jonathan. 1994. "Scobit: An alternative estimator to logit and probit." *American Journal of Political Science* 38(1):230-255.
- Yatchew, Adonis and Zvi Griliches. 1985. "Specification error in probit models." *The Review of Economics and Statistics* 67(1):134-139.

## Week 7: Ordered and Multinomial Choice Models

Ordered and multinomial logit and probit

Required Readings:

- Long Chapters 5 and 6
- Eliason Chapter 4
- Whitten, Guy D. and Harvey D. Palmer. 1996. "Heightening comparativists' concern for model choice: Voting behavior in Great Britain and the Netherlands". *American Journal of Political Science*: 231-260.
- Alvarez, R.M. and J. Nagler. 1995. "Economics, issues and the Perot candidacy: Voter choice in the 1992 presidential election". *American Journal of Political Science*: 714-744.
- Dow, Jay K. and Endersby, James W. 2004. "Multinomial probit and multinomial logit: A comparison of choice models for voting research." *Electoral Studies* 23: 107-122.

Suggested Readings:

- Imai, Kosuke and David A. van Dyk. 2005. "A Bayesian analysis of the multinomial probit model using marginal data augmentation." *Journal of Econometrics* 124(2): 311-334.
- 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.
- Duch, Raymond M. and Harvey D. Palmer. 2004. "It's not whether you win or lose, but how you play the game: Self-interest, social justice, and mass attitudes toward market transition." *American Political Science Review* 98(3):437-452

## Week 8: Midterm Exam

Part I (in class)

Part II (take-home) **Due by 11:59pm on Friday, October 19**

## Week 9: Event Count Models

Poisson, negative binomial, dealing with overdispersion and zero-inflation

Required Readings:

- Long Chapter 8
- King, Gary. 1989. "Event count models for international relations: Generalizations and applications." *International Studies Quarterly* 33(2): 123-147.
- 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* 32:838-63.

Suggested Readings:

- Cameron, Adrian Colin and Pravin K. Trivedi. 2013. *Regression analysis of count data*. Second Edition. Cambridge University Press.

## Week 10: Event Count Models Continued

Required Readings:

- Wilson, Matthew C., and James A. Piazza. 2013. "Autocracies and terrorism: Conditioning effects of authoritarian regime type on terrorist attacks." *American Journal of Political Science* 57(4): 941-955.
- Ruggeri, Andrea, Theodora-Ismene Gizelis, and Han Dorussen. 2013. "Managing mistrust: An analysis of cooperation with UN peacekeeping in Africa." *Journal of Conflict Resolution* 57(3):387-409.

## Week 11: Duration Models

Required Readings:

- 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*: 1414-1461.
- Berliner, Daniel. 2014. "The political origins of transparency." *The Journal of Politics* 76(2): 479-491.
- Bueno de Mesquita, Bruce, and Randolph M. Siverson. 1995. "War and the survival of political leaders: A comparative study of regime types and political accountability." *American Political Science Review* 89(2):841-55.

Suggested Readings:

- Berry, Frances Stokes, and William Berry. 1990. "State lottery adoptions as policy innovations: An event history analysis." *American Political Science Review* 84(2): 395-415.
- Box-Steffensmeier, Janet M., and Bradford S. Jones. 2004. *Event history modeling: A guide for social scientists*. Cambridge University Press.
- Hosmer, D., Stanley Lemeshow, and Susanne May. 2008. *Applied survival analysis: Regression modeling of time to event data*. Second Edition. New York: Wiley.

## Week 12: Duration Models Continued

Required Readings:

- Box-Steffensmeier, Janet M., and Christopher JW Zorn. 2001. "Duration models and proportional hazards in political science." *American Journal of Political Science*: 972-988.
- Park, Sunhee, and David J. Hendry. 2015. "Reassessing Schoenfeld residual tests of proportional hazards in political science event history analyses." *American Journal of Political Science* 59(4): 1072-1087.
- Kiefer, Nicholas M. 1988. "Economic Duration Data and Hazard Functions." *Journal of Economic Literature* 26:646-79.
- Kropko, Jonathan and Jeffrey J. Harden. 2017. "Beyond the hazard ratio: Generating expected durations from the Cox proportional hazards model." *British Journal of Political Science* 1-18.

#### Suggested Readings:

- Alt, James E., Gary King and Curtis S. Signorino. 2001. "Aggregation among binary, count, and duration models: Estimating the same quantities from different levels of data." *Political Analysis* 9(1):21-44.
- Box-Steffensmeier, Janet M., Suzanna De Boef and Kyle A. Joyce. 2007. "Event dependence and heterogeneity in duration models: The conditional frailty model." *Political Analysis* 15:237-256.

### **Week 13: NO CLASS (Thanksgiving Break)**

### **Weeks 14 and 15: Special Topics**

Weeks 14 and 15 are special topics. We may cover some/all of the following, depending on time and the interests of the class:

- Linear and Non-Linear Interactions and Interpretation
  - Rainey, Carlisle. 2016. "Compression and Conditional Effects: A Product Term Is Essential When Using Logistic Regression to Test for Interaction." *Political Science Research and Methods*: 1-19.
  - Berry, William D., Jacqueline H. R. 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: 248-66.
  - Berry, William D., Jacqueline H.R. DeMeritt, and Justin Esarey. 2016. "Bias and overconfidence in parametric models of interactive processes." *American Journal of Political Science* 60(2): 521-539.
  - Norton, E.C., H. Want, and C. Ai. 2004. "Computing Interaction Effects and Standard Errors in Logit and Probit Models." *Stata Journal* 4(2):154-67.
- Bounded and Compositional Data
  - Tomz, Michael, Joshua A. Tucker, and Jason Wittenberg. 2002. "An easy and accurate regression model for multiparty electoral data." *Political Analysis* 10(1): 66-83.
  - Philips, Andrew Q., Amanda Rutherford, and Guy D. Whitten. 2016. "Dynamic pie: A strategy for modeling trade-offs in compositional variables over time". *American Journal of Political Science* 60(1):268-283.
  - Philips, Andrew Q., Amanda Rutherford, and Guy D. Whitten. 2016. "dynsimpie: A command to examine dynamic compositional dependent variables." *Stata Journal* 16(3): 662-677.
  - Funk, Kendall D., and Andrew Q. Philips. 2018. "Representative budgeting: Women mayors and the composition of spending in local governments." *Political Research Quarterly*: 1-15.
  - Katz, J.N. and King, G. 1999. "A statistical model for multiparty electoral data." *American Political Science Review* 93(01):15-32.
- Hierarchical Models and Random Effects
  - Steenbergen, Marco R., and Bradford S. Jones. 2002. "Modeling multilevel data structures." *American Journal of Political Science* 46(1): 218-237.
  - Gelman, Andrew, and Jennifer Hill. *Data analysis using regression and multilevel/hierarchical models*. Cambridge University Press, 2006.
  - King, Gary and Margaret E. Roberts. 2015. "How robust standard errors expose methodological problems they do not fix, and what to do about it." *Political Analysis* 23: 159-179.
  - Arceneaux, K. and Nickerson, D.W., 2009. "Modeling certainty with clustered data: A comparison of methods." *Political Analysis* 17(2):177-190.
  - Bowers, Jake and Katherine W. Drake. 2005. "EDA for HLM: Visualization when probabilistic inference fails." *Political Analysis* 13:301-326.
- Missing Data Imputation and the EM Algorithm
  - Honaker, James, Gary King, and Matthew Blackwell. 2011. "Amelia II: A program for missing data." *Journal of Statistical Software* 45(7):1-47.
  - Rubin, Donald B. 1976. "Inference and missing data" *Biometrika* 63(3):581-592.
- Uncertainty, Significance, Hypothesis Testing, and Presenting and Simulating Models
  - Efron, Bradley. 2005. "Bayesians, frequentists, and scientists." *Journal of the American Statistical Association* 100(469):1-5.

- Gill, Jeff. 1999. "The insignificance of null hypothesis significance testing" *Political Research Quarterly* 52(3):647-674.
- Gelman, A. and Stern, H., 2006. "The difference between 'significant' and 'not significant' is not itself statistically significant". *The American Statistician* 60(4):328-331.
- King, Gary, Michael Tomz and Jason Wittenberg. 2000. "Making the most of statistical analyses: Improving interpretation and presentation" *American Journal of Political Science* 44(2):347-361.
- Imai, Kosuke, Gary King, and Olivia Lau. 2008. "Zelig: Everyone's statistical software."
- Hanmer, Michael J., and Kerem Ozan Kalkan. 2013. "Behind the curve: Clarifying the best approach to calculating predicted probabilities and marginal effects from limited dependent variable models." *American Journal of Political Science* 57(1):263-277.
- King, Gary. 1991. "Calculating standard errors of predicted values based on non-linear functional forms." *The Political Methodologist* 4(2):2-4.

## **Week 14: Special Topics**

## **Week 15: Special Topics**

Research papers due Thursday, December 6, by Noon.

## **Week 16: Student Presentations**

Student critiques due Monday, December 10 by 11:59pm.

## **SYLLABUS CHANGES**

I reserve the right to make changes to the syllabus during the course of the semester as needed and will make the most updated copy available to you and announce said changes during class.

**Last updated:** August 22, 2018

## **UNIVERSITY-MANDATED STATEMENTS**

### **Accommodation for disabilities**

If you qualify for accommodations because of a disability, please submit your accommodation letter from Disability Services to your faculty member in a timely manner so that your needs can be addressed. Disability Services determines accommodations based on documented disabilities in the academic environment. Information on requesting accommodations is located on the Disability Services website. Contact Disability Services at 303-492-8671 or dsinfo@colorado.edu for further assistance. If you have a temporary medical condition or injury, see Temporary Medical Conditions under the Students tab on the Disability Services website.

### **Classroom behavior**

Students and faculty each have responsibility for maintaining an appropriate learning environment. Those who fail to adhere to such behavioral standards may be subject to discipline. Professional courtesy and sensitivity are especially important with respect to individuals and topics dealing with race, color, national origin, sex, pregnancy, age, disability, creed, religion, sexual orientation, gender identity, gender expression, veteran status, political affiliation or political philosophy. Class rosters are provided to the instructor with the student's legal name. I will gladly honor your request to address you by an alternate name or gender pronoun. Please advise me of this preference early in the semester so that I may make appropriate changes to my records. For more information, see the policies on classroom behavior and the Student Code of Conduct.



## **Honor code**

All students enrolled in a University of Colorado Boulder course are responsible for knowing and adhering to the Honor Code. Violations of the policy may include: plagiarism, cheating, fabrication, lying, bribery, threat, unauthorized access to academic materials, clicker fraud, submitting the same or similar work in more than one course without permission from all course instructors involved, and aiding academic dishonesty. All incidents of academic misconduct will be reported to the Honor Code ([honor@colorado.edu](mailto:honor@colorado.edu); 303-492-5550). Students who are found responsible for violating the academic integrity policy will be subject to nonacademic sanctions from the Honor Code as well as academic sanctions from the faculty member. Additional information regarding the Honor Code academic integrity policy can be found at the Honor Code Office website.

## **Sexual misconduct, discrimination, harassment and/or related retaliation**

The University of Colorado Boulder (CU Boulder) is committed to fostering a positive and welcoming learning, working, and living environment. CU Boulder will not tolerate acts of sexual misconduct (including sexual assault, exploitation, harassment, dating or domestic violence, and stalking), discrimination, and harassment by members of our community. Individuals who believe they have been subject to misconduct or retaliatory actions for reporting a concern should contact the Office of Institutional Equity and Compliance (OIEC) at 303-492-2127 or [cureport@colorado.edu](mailto:cureport@colorado.edu). Information about the OIEC, university policies, anonymous reporting, and the campus resources can be found on the OIEC website. Please know that faculty and instructors have a responsibility to inform OIEC when made aware of incidents of sexual misconduct, discrimination, harassment and/or related retaliation, to ensure that individuals impacted receive information about options for reporting and support resources.

## **Religious holidays**

Campus policy regarding religious observances requires that faculty make every effort to deal reasonably and fairly with all students who, because of religious obligations, have conflicts with scheduled exams, assignments or required attendance. In this class, I will try to accommodate your requests, but you must contact me early in the semester. See the campus policy regarding religious observances for full details.