GOV 385L: Advanced Statistical Analysis, Fall 2024, TTH 5:00-6:30pm (37485), MEZ 1.210
SDS 385.14: Maximum-Likelihood Statistics, Fall 2024, TTH 5:00-6:30pm (57560), MEZ 1.210

Instructor: Tse-min Lin
Office Hours: TTH 3:30-5:00pm (BAT 4.144)
E-Mail: tml@austin.utexas.edu

Course Overview:

Assuming that students have already learned the linear regression model, this course introduces more advanced statistical models, including discrete choice models, event count models, and, optionally, models with limited dependent variables and event history models. All these models are widely used, and our emphasis will be their applications in political science. Since the estimation of these models relies mainly on the maximum likelihood method of estimation, we will first introduce MLE together with its mathematical prerequisites, mainly calculus and probability distributions. We will use R and/or Stata for statistical analysis and Mathematica for mathematical analysis.

Maximum Likelihood Estimation

- is an estimation method like OLS or Bayesian estimation;
- is particularly useful when the dependent variable follows a special distribution (such as Bernoulli, binomial, Poisson, negative binomial, exponential, Weibull, or multinomial) or when the model to be estimated is nonlinear;
- can accommodate limited and truncated distributions;
- is very flexible in *(re)parameterization*, i.e., the parameter of a distribution can be considered as a (linear or nonlinear) function of some covariates with a set of other parameters or it can be considered as a random variable with other parameters;
- can be applied to models arisen out of substantive considerations;
- produces estimates with nice statistical properties.

Grading Policy:

In addition to regular homework assignments, you are required to write a research paper based on a statistical procedure introduced in this class. The topic of the paper is your own choice, but you should work closely with the instructor in developing ideas, formulating models, acquiring data, and carrying out the analysis. **By Week 10, you should submit a short version of the proposal, in the form of "bullet points," to Canvas as "Discussion". (See Appendix 1: A Guide to Proposal Writing.) By Week 12, you are required to turn in a complete proposal.** ***The final paper should include a methodological appendix detailing the methods used in the research.*** The proposal and the final paper will account for a significant proportion of your final grade. Plus/minus grades will be assigned for the final grade.

Homework Assignments (5 sets): 20%
Paper Proposal (Week 10 & 12): 20%
Final Paper: 60%

Required Texts:

- G. King.1998. *Unifying Political Methodology: The Likelihood Theory of Statistical Inference*. Michigan.

- J. Xu. 2022. *Modern Applied Regressions: Bayesian and Frequentist Analysis of Categorical and Limited Response Variables with R and Stan*. Chapman and Hall. (Digital version available at UT Library's website.)
- A packet of journal articles and book chapters.

Optional but Strongly Recommended Texts:

- S. R. Eliason. 1993. *Maximum Likelihood Estimation: Logic and Practice.* Sage.
- T. F. Liao. 1994. *Interpreting Probability Models*. Sage.

Useful Texts for Mathematical and Statistical Preparation:

- R. J. Larsen and M. L. Marx. 1981 (1986, 2001). *An Introduction to Mathematical Statistics*. 1st (2nd, 3rd) ed. Prentice Hall.

Students with Disabilities:

Students with disabilities may request appropriate academic accommodations from the Division of Diversity and Community Engagement, Services for Students with Disabilities, 512-471-6259. For more information, visit http://diversity.utexas.edu/disability/.

Policy on Academic Integrity:

Students who violate University rules on academic dishonesty are subject to disciplinary penalties, including the possibility of failure in the course and / or dismissal from the University. Since such dishonesty harms the individual, all students, and the integrity of the University, policies on academic dishonesty will be strictly enforced. For further information, please visit the Student Conduct and Academic Integrity website at: http://deanofstudents.utexas.edu/conduct/.

Sharing of Course Materials is Prohibited:

No materials used in this class, including, but not limited to, lecture hand-outs, videos, assessments (quizzes, exams, papers, projects, homework assignments), in-class materials, review sheets, and additional problem sets, may be shared online or with anyone outside of the class unless you have my explicit, written permission. Unauthorized sharing of materials promotes cheating. It is a violation of the University's Student Honor Code and an act of academic dishonesty. I am well aware of the sites used for sharing materials, and any materials found online that are associated with you, or any suspected unauthorized sharing of materials, will be reported to Student Conduct and Academic Integrity in the Office of the Dean of Students. These reports can result in sanctions, including failure in the course.

FERPA and Class Recordings:

Class recordings are reserved only for students in this class for educational purposes and are protected under FERPA. The recordings should not be shared outside the class in any form. Violation of this restriction by a student could lead to Student Misconduct proceedings.

Accommodations for Religious Holidays:

By UT Austin policy, you must notify me of your pending absence at least fourteen days prior to the date of observance of a religious holy day. If you must miss a class, an examination, a work assignment, or a project in order to observe a religious holy day, you will be given an opportunity to complete the missed work within a reasonable time after the absence.

Course Outline and Reading Assignments:
(#: materials available online, at the library, or upon request - for duplicating)

Week 1: Introduction

The Evolution of Political Methodology (Slides)
Understanding P-Values (Slides)
A. Gelman and E. Lokan. 2014. "The Statistical Crisis in Science." *American Scientist*, 102(6).
R. L. Wasserstein and N. A. Lazar. 2016. "The ASA's Statement on p-Values: Context, Process, and Purpose." *The American Statistician*, 70(2):129-133.
Benjamin et al. 2017. "Redefine Statistical Significance." *Nature Human Behaviour*, 01 September 2017.
Lakens et al. 2018. "Justify Your Alpha: A Response to 'Redefine Statistical Significance'". Accepted by *Nature Human Behaviour*, January 2018. (https://psyarxiv.com/9s3y6).
Amrhein et al. 2019. "Scientists Rise up against Statistical Significance." *Nature*, 20 March 2019. (https://www.nature.com/articles/d41586-019-00857-9).
U. Simonsohn, L. D. Nelson, & J. P. Simmons. 2014. "P-curve: A key to the File-Drawer". *Journal of Experimental Psychology: General*, 143(2), 534-547.
S. Dominus, "When the Revolution Came for Amy Cuddy." *The New York Times Magazine*, October 18, 2017.

Week 2: Mathematical Tools: ChatGPT, R, & Mathematica

Use ChatGPT to Write R Code (Slides)
Xu, 1
Wolfram Documentation (https://reference.wolfram.com/language/)

Week 3-5: Probability Distributions

King, 1; 3

R. J. Larsen and M. L. Marx. 1981. "Special Distributions." Chapter 4 of *An Introduction to Mathematical Statistics and Its Applications*, Prentice Hall.

Week 6-8: Statistical Inference and the Maximum Likelihood Estimation

King, 2; 4
Eliason,1-6
R. J. Larsen and M. L. Marx. 2001. "Estimation." Chapter 5 of *An Introduction to Mathematical Statistics*, 3rd ed. Prentice Hall.

Week 9-10: Binary Choice Models

Xu, 2
King, 5.1-5.3
Liao, 1-3

Week 11-12: Advanced Discrete Choice Models

Xu, 3
King, 5.4
Liao, 4-7

Multinomial Probit/Logit

R. M. Alvarez and J. Nagler, 1995. "Economics, Issues and the Perot Candidacy: Voter Choice in the 1992 Presidential Election." *AJPS*, 39:714-744.
K. M. Quinn, A. D. Martin, and A. B. Whitford. 1999. "Voter Choice in Multi-Party Democracies: A Test of Competing Theories and Models." *AJPS*, 43: 1231–1247.
J. K. Dow and J. W. Endersby. 2004. "Multinomial Probit and Multinomial Logit: A Comparison of Choice Models for Voting Research." *Electoral Studies*, 23:107-122.

Heteroscedastic Probit

R. M. Alvarez and J. Brehm. 1995. "American Ambivalence Towards Abortion Policy: Development of a Heteroskedastic Probit Model of Competing Values." *AJPS*, 39:1055-82.
R. M. Alvarez and J. Brehm. 1997. "Are American Ambivalent Towards Racial Policies?" *AJPS*, 41:345-374.
C. H. Achen. 2002. "Toward a New Political Methodology: Microfoundations and ART." *Annual Review of Political Science*, 5:423–450. (*Read pp. 444-445.)
T. Lin. 2011. "Information and Ideological Structure in Spatial Voting." *Taiwan Journal of Democracy*, 7:1-24.

Ordered Probit/Logit and Heteroscedastic Ordered Probit/Logit

R. Williams, 2019. "Ordered Logit Models: Basic and Intermediate Topics." https://www3.nd.edu/~rwilliam/stats3/Ologit01.pdf
R. D. McKelvey and W. Zavoina. 1975. "A Statistical Model for the Analysis of Ordinal Level Dependent Variables." *Journal of Mathematical Sociology*, 4:103-120.
W. E. Becker and P. E. Kennedy. 1992. "A Graphical Exposition of the Ordered Probit." *Econometric Theory*, 8:127-131.
R. M. Alvarez and J. Brehm. 1998. "Speaking in Two Voices: American Equivocation about the Internal Revenue Service." *AJPS*, 42:418-452.
R. A. Hart, Jr. 2000. "Democracy and the Successful Use of Economic Sanctions." *Political Research Quarterly*, 53:267-284.

Conditional Logit

D. McFadden. 1973. "Conditional Logit Analysis of Qualitative Choice Behavior." In P. Zarembka, ed., *Frontiers in Econometrics*. New York: Academic.
R. M. Alvarez and J. Nagler, 1998. "When Politics and Models Collide: Estimating Models of Multiparty Elections." *AJPS*, 42:55-96.

Nested Logit

Born, R. 1990. "Surge and Decline, Negative Voting and the Midterm Loss Phenomenon: A Simultaneous Choice Analysis." *AJPS*, 34:615-645.

Other Choice Models

M. S. Sanders. 1999. "Unified Models of Turnout and Vote Choice for Two-Candidate and Three-Candidate Elections." *Political Analysis*, 7:89-116.
C. S. Signorino. 1999. "Strategic Interaction and the Statistical Analysis of International Conflict." *APSR*, 93:279-297.
C. S. Signorino. 2003. "Structure and Uncertainty in Discrete Choice Models." *Political Analysis*, 11:316-344.
J. L. Carson. 2003. "Strategic Interaction and Candidate Competition in U.S. House Elections: Empirical Applications of Probit and Strategic Probit Models." *Political Analysis*, 11:368-380.
M. A. Bas, C. S. Signorino, and R. W. Walker. 2008. "Statistical Backwards Induction: A Simple Method for estimating Recursive Strategic Models." *Political Analysis*, 16:21-40.

Week 13: Event Count Models

Xu, 4
King, 5.5-5.10
Liao, 8

On International Relations

L. F. Richardson. 1944. "The Distribution of Wars in Time." *Journal of Royal Statistical Society*, 107:242-250.
R. M. Siverson. and G. T. Duncan. 1976. "Stochastic Models of International Alliance Initiation, 1885-1965." In D. A. Zinnes and J. V. Gillespie, eds., *Mathematical Models in International Relations*.
W. W. Davis and G. T. Duncan. 1978. "The Dynamics of Warfare: 1816-1965." *AJPS*, 22:772-792.
H. W. Houweling and J. B. Kune. 1984. "Do Outbreaks of War Follow a Poisson-Process?" *Journal of Conflict Resolution*, 28:51-61.
G. King. 1989. "Event Count Models for International Relations: Generalizations and Applications." *International Studies Quarterly*, 33:123-147.
K. Benoit. 1996. "Democracies Really Are More Pacific (in General)." *Journal of Conflict Resolution*." 40:636-657.

General & Miscellaneous

G. King. 1987. "Presidential Appointments to the Supreme Court: Adding Systematic Explanation to Probabilistic Description." *APQ*, 15:373-386.
G. King 1988. "Statistical Models for Political Science Event Counts: Bias in Conventional Procedures and Evidence for the Exponential Poisson Regression Model." *AJPS*, 32:838-863.
G. King 1989. "Variance Specification in Event Count Models: From Restrictive Assumptions to a Generalized Estimator." *AJPS*, 33:762-784.
T. Y. Wang, W. J. Dixon, E. N. Muller, and M. A. Seligson. 1993. "Inequality and Political Violence Revisited." *APSR*, 87: 979-993.

Week 14: Fall Break

Week 15: GLM with Multilevel & Panel Data

Xu, 6.1
W. Greene. 2004. "The Behavior of the Maximum Likelihood Estimator of Limited Dependent Variable Models in the Presence of Fixed Effects." *Econometrics Journal*, 7: 98-119.
C. Cameron and P. K. Trivedi. 2005. *Microeconometrics: Methods and Applications*. Cambridge University Press.
B. Liseo. 2015. "Likelihoods that Eliminates Nuisance Parameters." *International Encyclopedia of the Social and Behavioral Sciences*, 2nd Edition. 14: 120-124.

Appendix 1. A Guide to Proposal Writing

The short version of your proposal should be a list of the following information, in "bullet points":

- **research question**
- **dependent variable, its scale of measurement and distribution**
- **independent variables, their scales of measurement**
- **unit of analysis and data structure, including N**
- **model (ideally a formal representation of your model)**
- **method**

The complete version of your proposal should be organized as follows:

(1) The "big picture": What is your research question? What do you seek to explain? Specifically, what is your dependent variable?

(2) The significance of the project: Why is it important/interesting to conduct the investigation? Include a literature review here to show that your project contributes to the accumulation of knowledge.

(3) Concepts, theory, and propositions: Discuss the key concepts in your research question. Use those concepts to build a theory: what causes the variation of your dependent variable and why? Deduce propositions from your theory. If possible, present a formal theory or model.

(4) Research design:

- unit of analysis
- data source (how and where you will collect data)
- measurement (operational definitions, reliability, validity, etc.)
- model (empirical model to be estimated with data)
- (estimation) method
- testable hypotheses

(5) Preliminary results (if any)

(6) Research plan and schedule