GOV 385L: Advanced Statistical Analysis, Fall 2010 (38810)

Prof. Tse-min Lin Office: BAT 4.144 Office Hours: TTH 3:30-5:00
Phone: 232-7248 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 models with categorical dependent variables, models with limited dependent variables, event count models, 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 probability distributions. We will use STATA for statistical analysis and MAPLE for symbolic algebra.

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;
- can be easily implemented in statistical software such as STATA;
- produces estimates with nice statistical properties.

Examples:

- $Y \sim \text{Bernoulli}(\pi)$ with π=normal_cdf($x'\beta$), which leads to the probit model
- $Y \sim \text{Bernoulli}(\pi)$ with π=logistic_cdf($x'\beta$), which leads to the logit model
- $Y \sim N(\mu,\sigma^2)$ with $\mu=x'\beta$, which leads to linear regression
- $Y \sim \text{TruncatedN}(\mu,\sigma^2)$ with $\mu=x'\beta$, which leads to the tobit model
- $Y \sim N(\mu,\sigma^2)$ with $\mu=x'\beta$ and $\sigma^2=x'\beta$ (or $\sigma^2=\exp(x'\beta)$), which leads to linear regression with heteroscedasticity
- $Y \sim \text{Poisson}(\lambda)$ with $\lambda=\exp(x'\beta)$, which leads to Poisson regression
- $Y \sim \text{Poisson}(\lambda)$ with $\lambda \sim \text{gamma}(\varphi,\sigma^2)$ and $\varphi= \exp(x'\beta)$, which leads to negative binomial regression
- $Y = x'\beta + \alpha z^{\gamma}+\varepsilon$

A substantive example is the choice between voting for one of two (or more) candidate and abstention:

$$Vote_i = \begin{cases} \text{Candidate A if } U_i(A) - U_i(B) > T_i \\ \text{Candidate B if } U_i(B) - U_i(A) > T_i \\ \text{Abstention if } -T_i \le U_i(A) - U_i(B) \le T_i \end{cases} \quad \text{where } T_i = x_i'\beta$$

This threshold model of voting can be estimated with MLE. See M. S. Sanders, 1999, "Unified Models of Turnout and Vote Choice for Two-Candidate and Three-Candidate Elections," *Political Analysis*, 7:89-116.

Grading Policy:

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 discuss your ideas with the instructor early in the semester to obtain his approval. Depending on substantive merits, topics based "simplistic" methods may not be acceptable. In Week 10, you are required to turn in and present a paper proposal. A suggested outline of the proposal is attached at the end of this syllabus. You should work closely with the instructor in developing ideas, formulating models, acquiring data, and carrying out the analyses. *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, although the completion of 3-5 sets of homework will also be required. Plus/minus grades will be assigned for the final grade.

Paper Proposal & Its Presentation (Week 10): 20%
Final Paper (Due December 3): 60%
Homework Assignments (3-5 sets): 20%

Required Texts:

- S. R. Eliason. 1993. *Maximum Likelihood Estimation: Logic and Practice.* Sage.
- G. King.1998. *Political Methodology: The Likelihood Theory of Statistical Inference*. Michigan.
- T. F. Liao. 1994. *Interpreting Probability Models*. Sage.
- A packet of journal articles and book chapters (available online or on Blackboard).

Strongly Recommended Texts:

- J. M. Box-Steffensmeier and B. S. Jones. 2004. *Event History Analysis*. Cambridge.
- W. H. Greene. 2008. *Econometric Analysis*. 6th ed. Pearson & Prentice Hall.
- J. S. Long and J. Freese. 2006. *Regression Models for Categorical Dependent Variables Using Stata*. 2nd ed. Stata Press.

Useful Texts for Mathematical and Statistical Preparation:

- J. Fox. 2009. *A Mathematical Primer for Social Statistics*. Sage.
- J. Gill. 2006. *Essential Mathematics for Political and Social Research*. Cambridge.
- R. J. Larsen and M. L. Marx. 1981 (1986, 2001). *An Introduction to Mathematical Statistics*. 1st (2nd, 3rd) ed. Prentice Hall.
- M. H. DeGroot.. 1986/1987. *Probability and Statistics*, 2nd ed. Addison Wesley.
- M. H. DeGroot and M. J. Schervish. 2002. *Probability and Statistics*, 3rd ed. Addison Wesley.

Students with Disabilities:

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

University Honor Code:

http://registrar.utexas.edu/catalogs/gi09-10/ch01/index.html

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.

Emergency Evacuation Policy:

Occupants of buildings on The University of Texas at Austin campus are required to evacuate buildings when a fire alarm is activated. Alarm activation or announcement requires exiting and assembling outside.

Familiarize yourself with all exit doors of each classroom and building you may occupy. Remember that the nearest exit door may not be the one you used when entering the building.

Students requiring assistance in evacuation shall inform their instructor in writing during the first week of class.

In the event of an evacuation, follow the instruction of faculty or class instructors.

Do not re-enter a building unless given instructions by the following: Austin Fire Department, The University of Texas at Austin Police Department, or Fire Prevention Services office.

Behavior Concerns Advice Line (BCAL): 232-5050
Emergency Information Web Site: http://www.utexas.edu/emergency

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

Week 1: Introduction

Week 2: Mathematical Tools: Calculus with Maple

Fox, Chapters 2

Week 3-4: Probability Distributions

King, Chapters 1; 3
Fox, Sections 3.1-3.3
Greene, Appendix B

R. J. Larsen and M. L. Marx. 1981. "Special Distributions." Chapter 4 of *An Introduction to Mathematical Statistics and Its Applications*, Prentice Hall.
M. H. DeGroot.. 1986/1987. "Special Distributions." Chapter 5 of *Probability and Statistics*, 2nd ed. Addison Wesley.

Week 5-6: Statistical Inference and the Maximum Likelihood Estimation

King, Chapters 2; 4
Fox, Sections 3.4-3.7; Chapter 4
Greene, Appendices C & D
Eliason, Chapters 1-6

R. J. Larsen and M. L. Marx. 2001. "Estimation." Chapter 5 of *An Introduction to Mathematical Statistics*, 3rd ed. Prentice Hall.

Week 7: Logit and Probit Models

King, Sections 5.1-5.3
Greene, Sections 23.1-23.4
Liao, Chapters 1-3

Week 8: Multinomial, Ordered, and Heteroscedastic Probit/Logit Models

King, Section 5.4
Greene, Sections 23.10.1; 23.11.1; 23.11.3; 23.11.5
Liao, Chapters 4-6

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. 2005. "Information and Ideological Structure in Spatial Voting," Typescript.

Ordered Probit and Heteroscedastic Ordered Probit

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.
B. J. Lee. 2007. "Exploring the Determinants of Transitional Justice: Ordered Regression Models." Typscript.

Week 9: Conditional Logit, Nested Logit, and Other Choice Models

Greene, Sections 23.11.2; 23.11.4; 23.11.7
Liao, Chapter 7

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.

Week 10: Proposal Presentations

Week 11: Event Count Models

King, Sections 5.5-5.10
Greene, Sections 25.1; 25.2
Liao, Chapter 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.
G. King 1989. "A Seemingly Unrelated Poisson Regression Model." *Sociological Methods and Research*, 17:235-255.
T. Y. Wang, W. J. Dixon, E. N. Muller, and M. A. Seligson. 1993. "Inequality and Political Violence Revisited." *APSR*, 87: 979-993.

Week 12: Models with Limited Dependent Variable

King, Chapter 9
Greene, Sections 24.1-24.3

Continuous Models

K. B. Grier, M. C. Munger, and B. E. Roberts. 1994. "The Determinants of Industry Political Activities, 1978-1986." *APSR*, 88:911-926.
P. Paolino. 2001. "Maximum Likelihood Estimation of Models with Beta-Distributed Dependent Variables." *Political Analysis*, 9:325-346.

Zero-Inflated Discrete Models

D. B. Hall. 2000. "Zero-Inflated Poisson and Binomial Regression with Random Effects: A Case Study." *Biometrics*, 56:1030-1039.
M. N. Harris and X. Zhao. 2007. "A Zero-Inflated Ordered Probit Model with an Application to Modelling Tobacco Consumption." *Journal of Econometrics*, 141(2):1073-1099.

Week 13-14: Event History Models

Greene, Section 25.6
Box-Steffensmeier & Jones [Optional]

On Government Survival

C. Cioffi-Revilla. 1984. "The Political Reliability of Italian Governments." *APSR*, 78:318-337.
E. C. Browne, J. P. Frendreis, and D. W. Gleiber. 1986. "The Process of Cabinet Dissolution." *AJPS*, 30:628-650.
K. Strom, E. C. Browne, J. P. Frendries, and D. W. Gleiber. 1988. "Contending Models of Cabinet Stability." *APSR*, 82:923-941.
G. King, J. E. Alt, N. E. Burns, and M. Laver. 1990. "A Unified Model of Cabinet Dissolution in Parliamentary Democracies." *AJPS*, 34:846-871.
P. Warwick. 1992. "Rising Hazards: An Underlying Dynamics of Parliamentary

Government." *AJPS*, 36:857-876.
P. Warwick and S. T. Easton. 1992. "The Cabinet Stability Controversy: New Perspectives on a Classic Problem." *AJPS*, 36:122-146.
P. Warwick. 1992. "Economic Trends and Government Survival in West European Parliamentary Democracies." *APSR*, 86:875-887.
T. Lin and M. Guillen. 1999. "The Rising Hazards of Party Incumbency: A Discrete Renewal Analysis." *Political Analysis*, 7:31-58.

On Elite Circulation

T. W. Casstevens. 1989. "The Circulation of Elites: A Review and Critique of a Class of Models." *AJPS*, 33:294-317.
H. Bienen and N. van de Walle. 1992. "A Proportional Hazard Model of Leadership Duration." *JOP*, 54:686-717.
J. N. Katz and B. R. Sala. 1996. "Careerism, Committee Assignments, and the Electoral Connection." *APSR*, 90:21-33.
J. M. Box-Steffensmeier and B. S. Jones. 1997. "Time is of the Essence: Event History Models in Political Science." *AJPS*, 41:1414-1461.
C. J. Finocchiaro and T. Lin. 2004. "The Hazards of Incumbency: An Event History Analysis of Congressional Tenure," Typescript.

On International Relations

D. S. Bennett. 1997. "Testing Alternative Models of Alliance Duration, 1816-1984." *AJPS*, 41:846-878.
N. Beck, J. N. Katz, and R. Tucker. 1998. "Taking Time Seriously: Time- Series Cross-Section Analysis with a Binary Dependent Variable." *AJPS*, 42:1260-1288.
H. Dorussen and J. Mo. 2001. "Ending Economic Sanctions." *Journal of Conflict Resolution*, 45: 395-426.

General & Miscellaneous

F. S. Berry and W. D. Berry. 1990. "State Lottery Adoptions as Policy Innovations: An Event History Analysis." *APSR*, 84:395-415.
J. M. Box-Steffensmeier and B. S. Jones. 1997. "Time is of the Essence: Event History Models in Political Science." *AJPS*, 41:1414-1461.
C. J. W. Zorn. 2000. "Modeling Duration Dependence." *Political Analysis*. 8:367-380.
J. M. Box-Steffensmeier and C. J. W. Zorn. 2001. "Duration Models and Proportional Hazards in Political Science." *AJPS*, 45:972-988.
J. M. Box-Steffensmeier and C. J. W. Zorn. 2002. "Duration Models for Repeated Events." *JOP*, 64:1069-1094.

Week 15: Review

Suggested Proposal Outline

(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