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

See Appendix 1 for examples of MLE.

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 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. By Week 12, you are required to turn in a paper proposal. (See Appendix 2: A Guide for Proposal Writing.) 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.

**Plus/minus grades will be assigned for the final grade.**

Paper Proposal (Week 12): 20%
Final Paper (Due December 14 by midnight): 60%
Homework Assignments (5 sets): 20%
Required Texts:

- J. S. Long and J. Freese. 2006. *Regression Models for Categorical Dependent Variables Using Stata*. 2nd ed. Stata Press. (This book serves as a manual of STATA, but it is also a very useful reference for the statistical models introduced in this course.)
- A packet of journal articles and book chapters (available online or on Blackboard).

Optional but Strongly Recommended Texts:

- A STATA Companion to Political Analysis. 2nd ed. CQ Press.

Useful Texts for Mathematical and Statistical Preparation:


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/](http://www.utexas.edu/diversity/ddce/ssd/).

University Honor Code:

[http://registrar.utexas.edu/catalogs/gi09-10/ch01/index.html](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 2-3:** Mathematical Tools: Calculus

# Fox, 2

**Week 4-6:** Probability Distributions

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


**Week 7-8:** Statistical Inference and the Maximum Likelihood Estimation

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


**Week 9-10:** Binary Choice Models

King, 5.1-5.3  
Greene, 17.2-17.3  
Liao, 1-3
**Week 11-12: Advanced Discrete Choice Models**

King, 5.4
Greene, 18.2.1-18.2.10; 18.3.1-18.3.2
Liao, Chapters 4-7

**Multinomial Probit/Logit**


**Heteroscedastic Probit**


**Ordered Probit and Heteroscedastic Ordered Probit**


**Conditional Logit**


**Nested Logit**


**Other Choice Models**


**Week 13-14: Event Count Models**

King, 5.5-5.10  
Greene, 18.4.1-18.4.6  
Liao, 8

**On International Relations**


**General & Miscellaneous**


**Week 14-15: Optional Topics**

**Optional Topic 1: Models with Limited Dependent Variables**

King, 9  
Greene, 19.2-19.3

**Continuous Models**

**Zero-Inflated Discrete Models**


**Optional Topic 2: Event History Models**

Greene, 19.4
Box-Steffensmeier & Jones [Optional]

**On Government Survival**


**On Elite Circulation**


**On International Relations**

AJPS, 41:846-878.

General & Miscellaneous

Appendix 1. Examples of Maximum Likelihood Estimation

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 \( \pi = \text{normal}_\text{cdf}(x'\beta) \), which leads to the probit model
- \( Y \sim \text{Bernoulli}(\pi) \) with \( \pi = \text{logistic}_\text{cdf}(x'\beta) \), which leads to the logit model
- \( Y \sim \text{N}(\mu, \sigma^2) \) 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
- \( Y \sim \text{N}(\mu, \sigma^2) \) with \( \mu = x'\beta \) and \( \sigma^2 = z'\gamma \) (or \( \sigma^2 = \exp(z'\gamma) \)), 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 \( \lambda \sim \text{gamma}(\phi, \sigma^2) \) and \( \phi = \exp(x'\beta) \), which leads to negative binomial regression
- \( Y = x'\beta + \alpha z' + \varepsilon \)

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

\[
\text{Vote}_i = \begin{cases} 
\text{Candidate A if } U_i(A) - U_i(B) > T_i \\
\text{Candidate B if } U_i(B) - U_i(A) > T_i \\
\text{Abstention if } -T_i \leq U_i(A) - U_i(B) \leq T_i 
\end{cases}
\]

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.
Appendix 2. A Guide for Proposal Writing

(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