POS 6933 – Fall 2019

Topics in Political Methodology: Maximum Likelihood Estimation (MLE)
Department of Political Science, University of Florida

Monday: Periods 5-7; Room: MAT 006

Instructor: Prof. Badredine Arfi
Office: 221 Anderson Hall
Office Hours: M 10:00am – 11:00pm
W 1:30pm – 3:30pm
Or: by appointment

Phone: (352) 273 2357
Email: barfi@ufl.edu

COURSE DESCRIPTION AND OBJECTIVES

The linear model is a useful but limited tool for statistically studying most political and social phenomena. The dependent variables of interest to social scientists do not usually fit the assumptions of the Classical Linear Regression Model (CLRM). Methodologists have developed sophisticated methods to address many important failures of the CLRM. To analyze, for example, data with qualitative and ordinal dependent variables, count models as well as situations where there is a selection bias we need to go beyond the limitations of the CLRM. The maximum likelihood estimation – MLE – methodology is a general approach that enables researchers to estimate models with these and other difficulties. The maximum likelihood is a methodology that refers to a general estimation strategy, which means that it refers to a different way for thinking about data and parameters. In brief: we have a maximum likelihood methodology and there are many statistical models that are built using this methodology to address specific statistical problems.

The students are expected to understand the theory of maximum likelihood estimation and how it has been used to develop a number of important statistical models that political scientists use to study politics. The students are also expected to use these powerful statistical models to research problems that interest them.

The course consists of two parts. The first one deals with a number of models suited for analyzing problems with categorical and limited dependent variables. The second part consists of an introduction to the methodology of event history modeling.

REQUIRED TEXTS

1. J. Scott Long and Jeremy Freese. 2006. Regression Models for Categorical Dependent Variables Using Stata. Stata Press. Available on reserve at Library West.

2. Mario Cleves, William Gould, Roberto G. Gutierrez, and Yulia V. Marchenko. 2010. An Introduction to Survival Analysis Using Stata, Third Edition. Stata Press.

HIGHLY RECOMMENDED TEXTS

1. J. Scott Long. 1997. Regression Models for Categorical and Limited Dependent Variables. Sage. Available on reserve at Library West.

2. Janet M. Box-Steffensmeier and Bradford S. Jones. 2004. Event History Modeling: A Guide for Social Scientists. Cambridge University Press. Available on reserve at Library West.

ADDITIONAL MATERIAL ON CONVAS

1. Additional readings and materials are posted on e-learning site for the course.

REQUIREMENTS AND ASSESSMENT

The requirement for this course is simple (as always): work diligently and persistently. This includes attending classes, doing the readings carefully before the seminar meets, and working regularly on the computer applications, the extra readings, and the research paper. Each student should expect to be spending many hours learning how to excel in using the Stata software commonly used to estimate the models discussed in class.

There will be a number of homework assignments that the students must complete and turn in. The homework assignments are due on the specified dates; no late submission is acceptable. There will also be one final take-home exam.

A major component of the course evaluation will be a term research paper. Each student will produce a manuscript of high quality using an appropriate modelling strategy.

DISTRIBUTION OF GRADES

1. **10%: Weekly Stata self-training**
   - Read Long & Freese Stata book and submit complete Stata log (.smcl) and do files showing all the work for chapters 1-3 together, then chapters 4, 5, 6, 7, 8 individually. **Make sure you do read the preface of the book first!**
   - Read Cleves et al. book and submit a complete Stata log (.smcl) file showing all the work for chapters 1–7 together, then chapters 8, 9, 10, 11, 12, 13, 14, 15 individually.
   - Each week one of the students will briefly present to, and discuss with, the class his/her replications of the Stata work as compiled in the log/do files.

2. **10%: Extra-Reading Paper Presentation**
   - Each student will be required to present in a summary form one paper posted in the additional readings on canvas.
   - The presentation is expected to focus on the methodology of the paper and engage in a critique/assessment of the methods used to analyze the empirical data.

3. **30%: Take-Home Final Examination**
   - The final exam is a take-home and open-book, open-computer, open-anything-but-another-human-being (physical or virtual).

4. **40%: A Replication Paper**
   - Each student is required to choose in consultation with the instructor (see down below) a paper (published in the last 5 years) using an MLE methodology.
   - The student is required to replicate the results of the paper and go beyond (see down blow for details).
   - The goal is to produce a high-quality, potentially publishable manuscript, using a model (or models) discussed in the course.

5. **10%: Paper Presentation.**
   - Each student will present his/her paper on the last day of classes of the semester. The presentation will consist of a ppt presentation for about 10 minutes followed by 5 minutes of Q & A.

COMPUTER REQUIREMENTS

All models covered in this class will be estimated using the Stata software package. It is a must that you have a UF account and password so that you can use UFApps online. Stata is available on UFApps.

IMPORTANT NOTES

- The instructor reserves the right to change or amend any part or aspect of this document should a need for doing so emerge at any point in time during the semester.

- Students requesting classroom accommodation for disabilities must register with the Dean of Students Office and provide documentation from this office.

- All students are required to abide by UF standards of academic honesty laid out in the Student Honor Code, posted at http://www.dso.ufl.edu/sccr/process/student-conduct-honor-code/

- Students are expected to provide professional and respectful feedback on the quality of instruction in this course by completing course evaluations online via GatorEvals. Guidance on how to give feedback in a professional and respectful manner is available at https://gatorevals.aa.ufl.edu/students/. Students will be notified when the evaluation period opens, and can complete evaluations through the email they receive from GatorEvals, in their Canvas course menu under GatorEvals, or via https://ufl.bluera.com/ufl/. Summaries of course evaluation results are available to students at https://gatorevals.aa.ufl.edu/public-results/.

SPECIFICS ON THE REPLICATION PAPER

In order for the instructor to provide guidance in the preparation of the paper, you will be required to turn in various brief intermediate papers throughout the semester.

Each student must:

1. Find a published paper that interests you and that applies a statistical method comparable to the material covered in this course. **Date: September 16**

2. Obtain the data from ICPSR or elsewhere or the author if at all possible. **Date: October 7**

3. Replicate the published results as nearly as possible. **Date: November 4**

4. You **must extend** the analysis in some way. You could, for example:
   - Suggest a more appropriate functional form for the estimation and re-estimate.
   - Argue that one or a set of important variables were omitted and conduct the analysis anew.
   - Argue that the results are likely to be sensitive to sample selection or variable measurement etc. and then conduct appropriate analyses to address that possibility.
   - Extend the data or use a different data set to test the theory.
   - Any other good idea that you might have.

   The final paper should be 15-20 pages long, including the bibliography.
   Date: December 2.

5. **Note on Final Submission of the Paper**:
   Students are required to submit to canvas (in addition to the paper) an annotated smcl log life displaying their complete work, an annotated do file listing all the commands that one would use to replicate the results of the paper, and the final dta file used for the paper (including both replication and extension works)

IMPORTANT DATES

Classes Begin (for this course)	**Wed, August 21**
Holidays *No classes*	**September 2: Labor Day** **October 4 - 5: Homecoming** **November 11: Veterans Day** **November 27 - 30: Thanksgiving**
Classes End	**Wed, December 4**
Reading Days - no classes	**December 5-6**
Final Exams	**December 7-13**

Session	Book	TOPIC
1	SL	• Chap 3: Binary Outcomes: Linear Probability, Probit and Logit Models
2	SL	• Chap 4: Hypothesis Testing and Goodness of Fit
3	SL	• Chap 5: Ordinal Outcomes: Ordered Logit and Ordered Probit Analysis
4	SL	• Chap 6: Nominal Outcomes: Multinomial Logit and Related Models
5	SL	• Chap 7: Limited Outcomes: Tobit Model • Heckman Model and Other Sample Selection Models
6	SL	• Chap 8: Count Outcomes: Regression Models for Counts
7	BSJ	• Chap 2: The Logic of Event History Analysis • Chap 3: Parametric Models for Single-Spell Duration Data
8	BSJ	• Chap 4: The Cox Proportional Hazards Model
9	BSJ	• Chap 5: Models for Discrete Data
10	BSJ	• Chap 6: Issues in Model Selection
11	BSJ	• Chap 7: Inclusion of Time-Varying Covariates
12	BSJ	• Chap 8: Diagnostic Methods for the Event History Model
13	BSJ	• Chap 9: Some Modeling Strategies for Unobserved Heterogeneity
14	BSJ	• Chap 10: Models for Multiple Events
15		• Presentations

SL: J. Scott Long, **BSJ**: Box-Steffensmeier and Jones.