PS 43100 Maximum Likelihood

Brehm
Pick 318A
702–8075
Office Hours: Mon, 10-noon; Weds, 2-4pm
jjbrehm@uchicago.edu

This course is principally a course in maximum likelihood estimation of problems in political science (with possible diversions into basic measurement methods). One advantage of maximum likelihood approaches is that it allows us to more realistically model political phenomena. With MLE approaches, we can examine phenomena such as turnout in elections or votes for candidates as if they were a series of choices. We can approach the progression of economic conditions or the status of the president over time, where we let time become a part of the process, rather than simply something that we "control." We can examine the number of attacks on terrorist camps or the onset and duration of conflict.

The idea of a "likelihood" refers to the joint probability of a set of known events according to some probability distribution. The maximum of the likelihood distribution is one way of estimating the parameters of that distribution (such as a mean, a variance, a count, or a duration), and with something called a "link function," even regression–like parameters that would account for higher or lower levels of the parameter.

Although the concept of "likelihood" may be mysterious to you at the moment, likelihood is a useful bridge concept between several competing schools of probability. One notion of probability, the frequentist idea, describes probabilities in terms of the long-run chances of some set of events. The frequentist idea of probability forms the foundation for a lot of the statistics you should already know, comprising ideas like "hypothesis testing," "confidence intervals" and "statistical significance." You would have used tools like t-tests, ANOVAs, ordinary least squares regressions, and some extensions. It is a powerful and quite dominant approach in statistical analysis of social phenomena.

A second notion of probability, subjective probability or Bayesian inference, is based on the idea that probabilities describe subjective assessments of how often an event will occur, even if the event itself describes a never-to-be-replicated scenario (such as the 2018 November mid-term elections, or the 2016 vote on Brexit). (While there may well be elements of all presidential elections that are common, or all exits from economic treaties, the specific election happens only once). A core idea here is to derive estimates of the probability of a theory given the data, and is usually covered in great detail in subsequent courses. We will use some of these core ideas in this class, too.

A third notion of probability derives from information theory, and asks how well do my predictions about these uncertain events generalize to other uncertain events. The general idea is to make claims about the quality of a model based upon its predictive validity.

In all three of these approaches, the idea of a likelihood is a unifying principle. You need likelihood ideas in order to make claims about the generalizability of the sample that you have for the theories you wish to test in a frequentist frame. You need a strong foundation in likelihood modeling in order to be able to produce strong subjective posterior claims based upon assertions of prior probabilities (i.e., your posterior claims are based on the product of the priors and the likelihood). You need likelihood modeling in order to be able to make claims about the "fitness" of an extension of the sample that you have to an out-of-sample generalization. Likelihood modeling is an essential next step in learning about the methods of quantitative social science research.

My agenda for this class also intends to help advanced students like yourself learn how to handle the dirty work of data analysis. We will learn how to work with a variety of different analytical tools, both for generating estimates of the relationships and for interpreting the results. We will learn how to extract and use data from the public archives. We will work over several different methods for presenting analytical results.

The assignments for this course are two-fold. You will have (nearly) weekly assignments which apply methods that we have covered during the week. You may work on your assignments as a team (up to two people), and please submit the work jointly. You will also have a final assignment where you pick the problem which interests you, analyze it using an appropriate method that we develop in this course (although you are permitted to venture into new territory if that is best for your problem), and write up the work. (You must complete the final assignment on your own). I anticipate that all of your write-ups, including the final one, will be around five pages in length. I will weight each assignment, including the final, equally for determining your grade.

It is especially important that you be active participants in this class. I am sure that some of the slides may be unclear. If you are confused there is an excellent chance that my slides are wrong and and equally excellent chance that you are not alone in your confusion.

I will make all the slides for the course available as PDFs which you can then print and annotate. The slides will all be available under "Lectures" on Canvas. All of the datasets, non-text readings, and replication files will be available in the "Files" section on Canvas as well.

There is a brand–new text for the course by John Ahlquist and Michael Ward, but the latest word is that the book is unavailable until mid-November at the earliest. Hence, I still assign King's (very good) *Unifying Political Methodology.* There are strengths to both approaches: King's book presents very clear derivations of the likelihood and excellent examples, and the Ahlquist and Ward book uses more up-to-date material, emphasizes R as an inference engine (with code), if it skimps on the derivations of the likelihoods. **If you are studying for the Political Science quantitative prelim, you should pick up both books.**

There is one required book for the course, and several recommended texts:
- Ahlquist and Ward, *Maximum Likelihood for Social Science* (required)
- King, *Unifying Political Methodology* (required)
- Hayduk, *Structural Equation Modeling with LISREL* (recommended)
- Maddala, *Limited Dependent Variables in Econometrics* (recommended)

There are a couple of websites that you may find helpful (and a lot less expensive than the best print versions):
- WolframAlpha: `wolframalpha.com`
- Wikipedia: search for "Probability Distributions"

In most ways, this syllabus is a lie: we will need to cover some basic material up through choice problems (because this class of problems is ubiquitous in the social sciences), but the subsequent topics should reflect your interests and we will adjust the topics accordingly. I will let you know of the material we anticipate covering in the subsequent class. That said, the usual pattern in the past is that we have covered choice models, count models, duration models, and then more flexibility for a few weeks depending on our speed and your (collective) interests.

2 Oct   <u>Introduction</u>
- King, Ch. 1 – 2.
- Ahlquist and Ward, Ch. 1.
- Gegenrizer, G. *Mindless Statistics* (recommended)
- Verzani, J. *Simple R* (if you need a refresher)

4 Oct   <u>Densities; General Method</u>
- King, Ch. 3.

9 Oct   <u>Properties</u>
- King, Ch. 4.
- Ahlquist and Ward, Ch. 2

16 Oct   <u>Dichotomous Outcomes</u>
- King, Ch. 5 (first half)
- Ahlquist and Ward, Ch. 3.
- Maddala, Ch 2 (first half)

18 Oct   <u>Interpretation of Probits and Logits</u>
- Ahlquist and Ward, Ch. 5 – 6
- King, Tomz, and Wittenberg. "Making the Most of Statistical Analyses," *AJPS* 44:341–355 (JSTOR)

23 Oct   <u>Alternative Links</u>
- Brehm and Gates, "Policing Police Brutality"

25 Oct   <u>Ordered Outcomes</u>
- Ahlquist and Ward, Ch. 8.
- Maddala, Ch. 2–3
- Maddala, Ch. 2 (second half)

30 Oct – 3 Nov   <u>Unordered Outcomes</u>
- Ahlquist and Ward, Ch. 9.
- Alvarez and Nagler, "Economics, Issues and the Perot Candidacy," *AJPS* 39:714–744 (JSTOR).
- Maddala, Ch. 3.

                     <u>Other topics, we select as appropriate to class interests</u>
- Counts (Ahlquist and Ward, Ch. 10; King, Ch. 5 (second half))

- Tabular data (King, Ch. 6)
- Latent variables (Psychometric models) (King, Ch. 8, 11)
- LISREL (Hayduk)
- Intro to computational methods (lasso, ridge)
- Duration (event history) (Ahlquist and Ward, Ch. 11)
- Time–series in ML (Ahlquist and Ward, Ch. 11)

Probably not until MLE II, whenever it is that I can offer it
- Compositional data
- Censored data (Ahlquist and Ward, Ch 12; King, Ch 10)
- Heteroskedasticity in ML
- Item Response Theory (Really best taken with Bob Gibbons)