Institut for Political Science—University of Zürich

MAXIMUM LIKELIHOOD ESTIMATION

Lucas Leemann

January 2010

Contact Information:
Lucas Leemann, Institut für Politikwissenschaft, Uni Zürich. Email: `lleemann@gmail.com`. Office hours: Daily, TBA.

Course Information:
Lecture times: Every day 10.15-12.00. Lecture location: TBA. Recitation/Lab times: Every day 15.30-17.30. Recitation location: TBA. Language of instruction: English or German (depends on students), slides will be in English.

Overview

The work horse of quantitative social sciences is the regression model. In introductory classes students usually learn about the linear regression and learn about *ordinary least squares (OLS)*. OLS can be used to estimate the linear regression model and its cousin (*weighted least squares*) can be used for more sophisticated linear regression models. Nevertheless, in many application we need a more powerful estimation technique. This technique is *Maximum Likelihood Estimation*. With MLE we can estimate many models and especially the standard models such as Logit/Probit, ordered Logit/Probit, multinomial Logit/Probit, Tobit, selection models, and duration models. Therefore MLE is a highly important topic for all students interested in quantitative political science.

A thorough understanding of MLE is valuable for three reasons. First, it is far easier to learn new models if the underlying statistical theory is clear. Second, sometimes one would like to estimate a model that is not *pre-canned*. The only option thus is to write up the code yourself and this requires an understanding of MLE theory. Finally, Bayesian statistics is very popular and seems

to be gaining even more popularity among political scientists. But Bayesian statistics builds on MLE and most Bayesian classes require basic knowledge of MLE.

This class offers an introduction to *maximum likelihood estimation* and certain applications. We will also see the link between the economic utility theory and basic stochastic decision models. By the end of the two weeks, participants will be able to program their own ML estimator and have a deeper understanding of basic models such as Logit, Probit, ordered Logit/Probit, multinomial Logit/Probit, quantal response equilibria models and more.

Basic Idea

Eight days is not a lot of time to cover the material. Therefore this class does not try to mimic a standard graduate student MLE class. Instead of covering a maximum amount of models I will rather spend more time on a thorough coverage of the statistical background. I believe that participants can understand and learn to apply more sophisticated models on their own if they have a clear understanding of the basics. I will therefore emphasize rather theory than application.

Books - Manuscripts

The main book will be Scott Long's *Regression Models for Categorical and Limited Dependent Variables*. Participants may already buy this book and familiarize themselves with the topic. Several other books are very helpful for this topic and will prove helpful during class. I will send out an e-mail mid December announcing which books or manuscripts I can offer as pdf documents and which books should be purchased.

- **Buy now:**

  Long, J. Scott. 1997. *Regression Models for Categorical and Limited Dependent Variables*. Thousand Oaks, CA: Sage.

  Eliason, Scott R. 1993. *Maximum Likelihood Estimation: Logic and Practice*. Newbury Park, CA: Sage.

- **Probably as pdf or hard copy:**

  Steenbergen, Marco R. 2007. *Maximum Likelihood for Political Analysis: Theory and Applications*. Chapel Hill: Department of Political Science, University of North Carolina. (pdf)

  Davidson, Russell and James G. MacKinnon. 2004. *Econometric Theory and Methods*. New York: Oxford University Press. **(p. 1-35)** (pdf)

Steenbergen, Marco. 2007. *Probability Distributions.* Chapel Hill: Department of Political Science, University of North Carolina. (pdf)

- **You might want to take a look at:**

  King, Gary. 1989. *Unifying Political Methodology. The Likelihood Theory of Statistical Inference.* New York: Cambridge University Press.

  Faraway, Julian. 2006. *Extending Linear Models with R: Generalized Linear, Mixed Effects and Nonparametric Regression Models.* Boca Raton: Chapman & Hall/CRC.

- **Refresher mathematics:**

  Chiang, Alpha C. and Kevin Wainwright. 2005. *Fundamental Methods of Mathematical Economics.* New York: McGraw Hill.

- **Statistics background:**

  DeGroot, Morris H. and Mark J. Schervish. 2002. *Probability and Statistics.* Reading, MA: Addison-Wesley, 2nd edition.

- **Software:**

  Stata: Long, J. Scott, and Jeremy Freese. 2006. *Regression Models for Categorical Dependent Variables Using Stata.* College Station: Stata Press, 2nd ed.

  R: Fox, John. 2002. *An R and S-PLUS Companion to Applied Regression.* New York: Sage Publication.

  R: Venables, William N. and Brian D. Ripley. 2002. *Modern Applied Statistics with S-PLUS.* Fourth ed. New York: Springer-Verlag.

  R: Kuhnert, Petra and Bill Venables. 2005. *An Introduction to R: Software for Statistical Modelling & Computing.* CSIRO Mathematical and Information Sciences Cleveland, Australia. (pdf)

Grading:

- *1st Take-Home Exam:* The first take-home exam will cover all the materials we discuss in the first week. It will be available on Thursday and has to be submitted by Monday morning.

- *2nd Take-Home Exam:* The second exam will be handed out on the last day after class and is due in 48 hours. Alternatively this exam might be converted into a in-class exam and take place in the second week.
- *Application with Real Data:* A short four two five page paper showing the students ability to use the learned skills. Any social science data and one of the discussed models can be used. Special emphasis is on interpretation and illustration of the results. Students are encouraged to use either Stata or R. The deadline for this final assignment will be 14 days after class.

Schedule:

This course takes place during the first two weeks of January. To maximize the benefit we will meet twice a day, for a class lecture in the morning from 10.15 to 12.00 and for the lab from 15.30-17.30. Altogether we will have 32 hours of class and lab, I expect all participants to spend a considerable amount of time to prepare and review the material. It is demanding and there is a substantial amount of reading to be covered.

January 4th, 2009

Class:
Basics of mathematical statistics. What is a distribution? We will focus on the normal distribution, the logistic distribution, and the Bernoulli distribution. Short repetition of the linear regression model. A first maximum likelihood estimation.
Lab:
The first part is an introduction to R for those students who do not know R yet. In the second part we will see how one can use simulations to test estimation performance. This procedure is known as *Monte Carlo simulation.*

Readings:

- Davidson, Russell and James G. MacKinnon. 2004. *Econometric Theory and Methods.* New York: Oxford University Press. **(skim p. 1-35)**
- Steenbergen, Marco. 2007. *Probability Distributions.* Chapel Hill: Department of Political Science, University of North Carolina. **(skim p. 1-35)**

January 5th, 2009

Class:
Theory of maximum likelihood estimation. Estimation of a parameter, estimation of the variance of the parameter estimate. Simple hypothesis tests. Derivation of simple parameter estimates for Bernoulli and normal distribution.
Lab:
We will see how one can write a function in R. We will use a numerical optimizer

and find maxima and minima of different functions. We will carry out a maximum likelihood estimation of a single parameter for a Bernoulli distribution and the normal distribution.

Readings:

- Eliason, Scott R. 1993. *Maximum Likelihood Estimation: Logic and Practice*. Newbury Park, CA: Sage. **p. 1-25**
- Long, J. Scott. 1997. *Regression Models for Categorical and Limited Dependent Variables*. Thousand Oaks, CA: Sage. **(Kapitel 3.1-3.6)**
- Steenbergen, Marco R. 2007. *Maximum Likelihood for Political Analysis: Theory and Applications*. Chapel Hill: Department of Political Science, University of North Carolina. **(tba)**
- * Train, Kenneth E. 2003. *Discrete Choice Methods with Simulation*. New York: Cambridge UNiversity Press. **Kapitel 3.1-3.3**
- *Long, J. Scott, and Jeremy Freese. 2006. *Regression Models for Categorical Dependent Variables Using Stata.* College Station: Stata Press.

January 6th, 2009

Class:
If the outcome variable is binary (e.g. yes or no vote) we are usually reluctant to estimate a linear probability model. We will briefly revisit the LPM and see how one would apply it. We will then turn to Logit and Probit. Both models will be derived and thereby we will cover stochastic utility theory. Participants will be introduced to the latent variable framework.

Lab: We write/program are own Logit and Probit functions in R. For the lab we will use data on the Titanic passengers and estimate different models – we will see that the survival probability depended on gender, boarding class and age.

Readings:

- Long, J. Scott. 1997. *Regression Models for Categorical and Limited Dependent Variables*. Thousand Oaks, CA: Sage. **(Kapitel 3.1-3.6)**
- Steenbergen, Marco R. 2007. *Maximum Likelihood for Political Analysis: Theory and Applications*. Chapel Hill: Department of Political Science, University of North Carolina. **(tba)**

January 7th and 8th, 2009

No class
The annual Swiss Political Science Association meeting takes place on Thursday and Friday. We will have no class. Use this time (and the weekend) for the take-home exam.

January 11th, 2009

Class:
Second part of maximum likelihood theory. We will discuss alternative hypothesis tests, such as the likelihood ratio test and the Wald test. Several measures of model fit will be introduced and we will cover different ways of interpreting estimation results (change in predicted probabilities and marginal effects).
Lab:
I will discuss the solutions of the take-home exam. Hypothesis tests in R. We will focus on the Wald test as it is the more complicated of the two tests. Most exercises will be based on the binary models from Wednesday.

Readings:

- Long, J. Scott. 1997. *Regression Models for Categorical and Limited Dependent Variables.* Thousand Oaks, CA: Sage. **(Kapitel 3.7, 4.1 & 4.3)**
- Eliason, Scott R. 1993. *Maximum Likelihood Estimation: Logic and Practice.* Newbury Park, CA: Sage. **p. 28-34**
- Steenbergen, Marco R. 2007. *Maximum Likelihood for Political Analysis: Theory and Applications.* Chapel Hill: Department of Political Science, University of North Carolina. **tba**

January 12th, 2009

Class:
Many possible outcome variables in the social sciences are ordered. Typically these variables take on three, five or seven categories and we know the ordering but not the distance between two answer categories. A typical example for such a variable is the answer to how much one favors policy A over policy B where the answer possibilities are 'less', 'equal', or 'more'. The correct model for such an outcome variable may be the ordinal Logit or ordinal Probit model.
Lab:
We will write/program our own ordered Logit and ordered Probit model. We will see what the crucial assumption (PRA) is and what its implications are.

Readings:

- Long, J. Scott. 1997. *Regression Models for Categorical and Limited Dependent Variables.* Thousand Oaks, CA: Sage. **(Kapitel 5.1-5.7)**
- Steenbergen, Marco R. 2007. *Maximum Likelihood for Political Analysis: Theory and Applications.* Chapel Hill: Department of Political Science, University of North Carolina. **(tba)**
- * McKelvey RD, Zavoina W (1975). "A Statistical Model for the Analysis of Ordinal Level Dependent Variables," *Journal of Mathematical Sociology* 4: 103-120.

January 13th, 2009

Class:
An outcome variable can have more than two answer categories but no ordering among the answers. In such a case we cannot estimate an ordered model. A multinomial Logit model is one possibility. We will cover the MNL model and we will see a formal test of PRA.

Lab:
We will program/write code to estimate our own MNL models. We will also return to Tuesday's topic and see a formal test of the parallel regression assumption.

Readings:

- Long, J. Scott. 1997. *Regression Models for Categorical and Limited Dependent Variables*. Thousand Oaks, CA: Sage. **(Kapitel 6.1-6.7)**
- Steenbergen, Marco R. 2007. *Maximum Likelihood for Political Analysis: Theory and Applications*. Chapel Hill: Department of Political Science, University of North Carolina. **(tba)**

Special Topics

The last two days are reserved for two special applications. Both these applications are highly illustrative of how flexible maximum likelihood is. This section might be less of practical value but will definitely contribute to the deeper understanding of this estimation framework and its potential.

January, 14th 2009

Class:
All models we have encountered so far are built upon the *iid* assumption. We have assumed that the random part is identically and independently distributed. If we are to take anything seriously that the school of new institutionalism is teaching us, then it has to be the *fact* that actors do not decide in a vacuum. Rather political actors take the strategic situation into consideration when making decisions. But if actors are behaving strategically, we cannot assume that observations are independent. The *Quantal Response Equilibrium* model was first presented for experimental data by McKelvey and Palfrey. Signorino subsequently extended it for observational data.

Lab:
We will use this model to estimate models on peace and war. There are two possibilities to estimate such a model and we will cover at least the easier (but less elegant) procedure. If possible we will apply the more sophisticated solution and derive a customized statistical model.

Readings:

- Signorino, Curtis S. 1999. "Strategic Interaction and the Statistical Analysis of International Conflict." *American Political Science Review*, 93(2): 279-297.
- Signorino, Curtis S. and Kuzey Yilmaz. 2003. "Strategic Misspecification in Regression Models." *American Journal of Political Science*, 47(3): 551-566.
- Bas, Muhammet Ali, Curtis S. Signorino, Robert W. Walker. 2008. "Statistical Backwards Induction: A Simple Method for Estimating Recursive Strategic Models." *Political Analysis*, 16: 21-40.
- * McKelvey, Richard D. and Thomas R. Palfrey. 1995. "Quantal Response Equilibria in Normal Form Games." *Games and Economic Behavior*, 10: 6-38.
- * McKelvey, Richard D. and Thomas R. Palfrey. 1996. "A Statistical Theory of Equilibrium in Games." *The Japanese Economic Review*, 47(2): 186-209.
- * McKelvey, Richard D. and Thomas R. Palfrey. 1998. "Quantal Response Equilibria for Extensive Form Games." *Experimental Economics*, 1: 9-41.

January 15th, 2009

Class:
Maximum likelihood estimation is extremely flexible and will most likely suit any statistical needs you have. Today's class will focus on an example of this flexibility. We will start with an explicit theory (spatial model) and derive an estimator.

Lab:
We will implement the morning's class and write/program code to estimate parameters of the spatial model.

Readings:

- Train, Kenneth E. 2003. *Discrete Choice Methods with Simulation.* New York: Cambridge University Press. **Kapitel 3.1-3.3**
- *Leemann, Lucas. 2009. "Catch Me If You Can! Recovering Sophisticated Voting" *Paper presented at the MPSA Conference*, Palmer House, Chicago.