PS350C: Mathematical and Statistical Models of Choice

Description

This course primarily considers mathematical and statistical models of individual choice behavior. Emphasis will be placed both on the derivations and motivations for statistical models of choice. Empirical applications will be reviewed to motivate and illustrate different models. Readings will be from political science, mathematical psychology, econometrics, statistics, and sociology. Topics include multinomial choice models and extensions, including strategic choice, dynamic choice, and the modeling of aggregated choices. Maximum likelihood will be the primary method of estimation considered, with a brief overview of simulation methods in a classical framework.

Contact Information

Jonathan Wand, Assistant Professor
    Office: Encina Hall West, Room 308
    Office hours: W 3-4 or appointment
    Email: `wand(at)stanford.edu`

Stephen Jessee, PhD Candidate
    Office: Encina Hall, 4th Floor
    Office hours: TBA
    Email: `sjessee(at)stanford.edu`

Times and location

Lectures (Wand): MW 900-1050, location: 200-202 (Rm 202 Lane Hall) ** revised **
Section (Jessee): Th 1000-1050, location: Encina Hall 464

Prerequisites

Prerequisite: PS350A (mathematical statistics) and PS350B (linear models), or comparable courses in econometrics or statistics.

Computational homework assignments will assume familiarity with a software package for statistical computation and plotting. All computation will be presented and discussed in terms of **R**.

Course Requirements

- Weekly assignments, partly analytical and partly computational in **R**

- Two part final exam: in-class analytical component and take-home computational component

Texts

In addition to articles assigned for particular weeks, there are two required texts. While articles are an important component of the course, it is sometimes handy to have a common text which can be referenced. Train's book covers core material in a clear fashion, although estimation by simulation is not our focus. The second book, by Long, has been a student favorite, covers material at a more basic level, and will used as a supplemental resource. Shorthand names are listed in **bold**.

Also listed below are useful recommended resources for the course and further reading. These will likely be of interest as your interest in the materials and related topics grows in the future. Mathematical sophistication ranges from generally accessible (*) to difficult (***). Some of these books are out of print (oop), but available at the library and in used form on-line (e.g., www.abebooks.com).

Required

Train Train, Kenneth. 2003. *Discrete Choice Methods with Simulation*. Cambridge University Press

Long Long, J. Scott. 1997. *Regression models for categorical and limited dependent variables*. Thousand Oaks: Sage

Recommended supplements for using **R**

MB Maindonald, John and John Braun. 2003. *Data Analysis and Graphics Using R*. Cambridge

Fox Fox, John. 2002. *An R and Splus Companion to Applied Regression*. Thousand Oaks: Sage

Recommended reading on choice

Luce, R. Duncan. 1959. *Individual choice behavior: a theoretical analysis*. NY: Wiley (* ; oop)

Ben-Akiva, Moshe and Stenven Lerman. 1985. *Discrete Choice Analysis: Theory and Applications to Travel Demand*. Cambridge, MA: MIT Press (*)

Maddala, G. S. 1983. *Limited-Dependent and Qualitative Variables*. Cambridge University Press (*)

Recommended reading for general econometrics with choice models

Davidson, Russell and James MacKinnon. 1993. *Estimation and Inference in Econometrics*. NY: Oxford (*)

Amemyia, Takshi. 1985. *Advanced Econometrics*. Cambridge, MA: Harvard University Press (**)

Recommended reading on extreme value theory

Embrechts, Paul, Claudia Kluppelberg, and Thomas Mikosch. 1997. *Modeling Extremal Events*. NY: Springer (**)

Resnick, Sidney I. 1987. *Extreme Values, Regular Variation, and Point Processes*. New York: Springer-Verlag (*** ; oop)

Course outline by week

1. INTRODUCTION AND OVERVIEW (March 31)

   I will present examples of research from different fields to illustrate types of questions and topics which are related to this course. Technical details of the works discussed will be considered during the remainder of the course.

   We will also discuss the overall enterprise that we undertake as part of this course. McFadden's article is a useful starting point, as a review and critique of important ideas.

   Reading:

   - McFadden, Daniel. 1999. Rationality for Economists? *Journal Of Risk And Uncertainty,*, 19(1/3):73–105.
     URL http://www.kluweronline.com/issn/0895-5646/

2. MATHEMATICAL PSYCHOLOGY AND MODELS OF CHOICE (April 5 and 7)

   We will review the foundational work from mathematical psychology on the properties of choice behavior and qualitative judgments. We will begin with Luce's axiomatic approach and Thurstone's measurement approach. Though in appearance quite different, these models share a number of close connections, particularly for pairs of choices. Yellot derives the circumstances under which they are equivalent and the circumstances under which they are unique.

   Emphasis will be put on deriving the relationship between distributional assumptions and particular choice models. We will also discuss the motivations for particular distributional assumptions.

   Readings April 5:

   - Luce, R. Duncan. 1959. *Individual choice behavior: a theoretical analysis*. NY: Wiley (Chapter 1)
     URL http://www.stanford.edu/class/polisci350c/classonly/luce1959.pdf
   - Thurstone, L. L. 1927. A Law of Comparative Judgement. *Pychological Review*, 34:272–86.
     URL http://www.psycinfo.com/library/display.cfm?uid=1994-28135-001&format=pdf

   Readings April 7:

   - Yellot, John I. Jr. 1977. The Relationship between Luce's Axiom, Thurstone's Theory of Comparative Judgment, and the Double Exponential Distribution. *Journal of Mathematical Psychology*, 15:109–144
     URL http://www.stanford.edu/class/polisci350c/classonly/yellot1977.pdf
   - (opt) McFadden, Danel. 1981. Structural Discrete Probability Models Derived from Theories of Choice. In Charles F. Manski and Daniel L. McFadden, editors, *Structural Analysis of Discrete Data and Econometric Applications*, Cambidge, MA: MIT Press, chapter 5. pages 198–272.
     URL http://emlab.berkeley.edu/discrete/ch5.pdf
   - (opt) McFadden, Daniel. 1974. Conditional logit analysis of qualitative choice behavior. In P Zarembka, editor, *Frontiers of Econometrics*, New York: Acadmic Press. pages 105–42.
     URL http://emlab.berkeley.edu/reprints/mcfadden/zarembka.pdf

Assignment, April 7 A.1 distributed

R: MB, Chap 2; Fox, Chap 3

3. MATHEMATICAL PROPERTIES OF BINARY CHOICE (April 12)

   We will parameterize a model for Bernoulli (i.e., binary) choices as a function of coefficients and individual specific variables. Relationships to OLS regression will be drawn, but estimation issues are deferred until later in the course.

   Emphasis will be on interpreting coefficients, and understanding how they vary across distributional assumptions used in the Bernoulli model.

   Readings:

   - Train, Chap 3.1–3.6
   - Long, Chap 3.1–3.4, 3.7–3.8
   - Londregan, John, Henry Bienen, and Nicolas van de Walle. 1995. Ethnicity and Leadership Succession in Africa. *International Studies Quarterly*, 39:1–25
   - (for problem set replication) Fearon, James D. and David D. Laitin. 2003. Ethnicity, Insurgency, and Civil War. *American Political Science Review*, 97(1):75–90

4. LIKELIHOOD AND ESTIMATION OF BINARY CHOICE (April 14)

   We will consider methods for estimating the coefficients in the Bernoulli choice model. Emphasis will be placed on derivations and understanding features of the likelihood functions for the models we have seen. Introductory discussion of optimization methods will also be covered. We will return to the topic of maximum likelihood estimation and consider it more rigorously in later weeks.

   Readings:

   - Train, Chap 3.7, Chap 8
   - Long, 3.5–3.6
   - (opt) Pratt, John W. 1981. Concavity of the Log Likelihood. *Journal of the American Statistical Association*, 76(373):103–6.
     URL http://links.jstor.org/sici?sici=0162-1459%28198103%2976%3A373%3C103%3ACOTLL%3E2.0.CO%3B2-X

   Assignments, April 14 A.1 due, A.2 distributed

5. GENERALIZATION OF CHOICE MODELS (April 19 and 21)

   We will derive generalizations for multiple choice categories, including nested and ordered categories.

   Readings:

   - Train Chap 3 and 4
   - Long Chap 5 (and 6)
   - (Indifference): Sanders, Mitchell S. 1999. Unified models of turnout and vote choice for two-candidate and three candidate elections. In Walter R. Mebane Jr., editor, *Political Analysis*, Ann Arbor: University of Michigan Press, volume 7. pages 89–116
     URL http://www.stanford.edu/class/polisci350c/classonly/sanders.pdf

- (GEV): Mebane, Walter R. Jr. 2000. Coordination, Moderation and Institutional Balancing in American Presidential and House Elections. *American Political Science Review*, 94(1):37–57

- (Ordered probit): Krehbiel, Keith and Douglas Rivers. 1988. The Analysis of Committee Power: An Application to Senate Voting on the Minimum Wage. *American Journal of Political Science*, 32:1151–74

Assignments, April 21 A.2 due, A.3 distributed

6. ESTIMATION AND INFERENCE (April 26 and 28)

   Now that we understand some models, the issue is how to estimate the probabilities conditional on some observed variables, compare competing models, and draw inference. To this end we will spend the week on the theory of maximum likelihood estimation. We will also cover estimates of the covariance matrix and inference. Classical test statistics will be reviewed.

   Readings:

   - Train, Chap 8
   - Long, Chap 4
   - (opt) Davidson and McKinnon Chapter 8, 13
   - (opt) Vuong, Q. 1989. Likelihood Ratio Tests for Model Selection and Non-Nested Hypotheses. *Econometrica*, 57:307–33

   Assignments, April 28 A.3 due, A.4 distributed

7. STRATEGIC CHOICE (May 3 and 5)

   We consider models where choices of individuals are affected by the choices of others in an extensive form game.

   Readings:

   - McKelvey, Richard and Thomas Palfrey. 1998. Quantal Response Equilibria for Extensive Form Games. *Experimental Economics*, 1:9–41
   - Signorino, Curtis S. 1999. Strategic Interaction and the statistical analysis of international conflict. *American Political Science Review*, 93(2):279–297
   - Lewis, Jeffrey and Kenneth Schultz. 2003. Revealing preferences: empirical estimation of a crisis bargaining game with incomplete information. *Political Analysis*, 11:345–67

   Assignments, May 5 A.4 due, A.5 distributed

8. DURATION MODELS (May 10)

   This week we use the choice framework to model a discrete duration process, with a brief look at continuous time duration models too.

   Readings:

   - Thompson, Jr., W. A. 1977. On the Treatment of Grouped Observations in Life Studies. *Biometrics*, 33(3):463–70
   - Box-Steffensmeier, Janet M. and Christopher J. W. Zorn. 2001. Duration Models and Proportional Hazards in Political Science. *American Journal of Political Science*, 45(4):972–88

- Beck, Nathaniel, Jonathan N. Katz, and Richard Tucker. 1998. Taking Time Seriously in Binary Time-Series–Cross-Section Analysis. *American Journal of Political Science*, 42(4):1260–1288
- Oneal, Russett, and Berbaum. 2002. Causes of Peace: Democracy, Interdependence, and International Organizations, 1885-1992. *International Studies Quarterly*, ?(?):1
- (opt) Fortna, Virginia Page. 2003. Scraps of Paper? Agreements and the Durability of Peace. *International Organization*, 57(2)

9. DYNAMIC CHOICE (May 12)

We consider models of individuals being faced with repeated choices among a fixed set of alternatives (e.g., presidential vote choice).

- Wand, Jonathan. 2004. Utility Maximization and Intertemporal Choice. Working paper.
- (opt) Hillygus, Sunshine and Simon Jackman. 2003. Voter Decision Making in Election 2000. *American Journal of Political Science*, 47(3):583–96
- (opt) Dagsvik, John K. 2002. Discrete Choice in Continuous Time: Implications of an intertemporal version of the IIA property. *Econometrica*, 70(2):817–31
- (opt) Resnick, Sidney and Rishin Roy. 1990. Multivariate Extremal Processes, Leader Processes and Dynamic Choice Models. *Advances in Applied Probability*, 22:309–331

Assignments, May 12 A.5 due, A.6 distributed

10. HETEROGENEITY AND RANDOM COEFFICIENTS (May 17)

We consider alternative specifications and some flexible models. How much can we loosen some of the restrictions on the traditional choice models?

Readings:

- Train Chap 8
- Rivers, Douglas. 1988. Heterogeneity in Models of Electoral Choice. *American Journal of Political Science*, 32(3):737–757

11. SPECIFICATION TESTS FOR MODELS OF CHOICE (May 19)

How do you know that your model is appropriate for a particular application? Which assumptions have testable implications? This is not an area which receives much attention in theory or in practice. This is a difficult and usually disappointing area for tests, particularly for evaluating IIA. Connections to mixed models will also be considered.

Readings:

- McFadden and Hausman. 1984. "Specification Tests for the Multinomial Logit Model" Econometrica, vol. 52, No. 5, 1219-1240.
- Fry, Tim R. L.; Harris, Mark N. 1998. "Testing for Independence of Irrelevant Alternatives: Some Empirical Results" *Sociological Methods & Research.* 26(3):401–423

Assignments, May 19 A.6 due, A.7 distributed

12. MODELS OF AGGREGATED CHOICES AND COUNTS (May 24)

Much data is not available at the individual level for privacy or practical reasons. We consider methods for analyzing counts of choices.

Readings:

- Long, Chap 8
- Wand, Jonathan, Kenneth Shotts, Jasjeet S. Sekhon, Walter R. Mebane, Jr., Michael Herron, and Henry E. Brady. 2001. The Butterfly Did It: The Aberrant Vote for Buchanan in Palm Beach County, Florida. *American Political Science Review*, 95:793–810 URL http://wand.stanford.edu/research/apsr2001.pdf
- (opt) Mebane, Walter R., Jr. and Jasjeet S. Sekhon. 2004. Robust Estimation and Outlier Detection for Overdispersed Multinomial Models of Count Data. *American Journal of Political Science*. Forthcoming

13. STATISTICAL AND NUMERICAL ISSUES... AND CAUTIONARY TALES (May 26 and June 2)

This final topic is aimed at tempering your hubris and unbounded enthusiasm for statistical modeling with some cautionary tales. These will cover pitfalls and limitations in order to warn of specific problems and to hopefully encourage you to undertake any new project in a careful and humble manner.

Readings:

- Achen, Christopher H. 2002. Toward a New Political Methodology: Microfoundations and ART. *Annual Review of Political Science*, 5:423–50
- McCullough, B. D. and H. D. Vinod. 2003. Verifying the solution from a nonlinear solver: A case study. *American Economic Review*, 93(3):873
- King, Gary and Langche Zeng. 2001. Logistic Regression in Rare Events Data. *Political Analysis*:137–63

Assignments, May 26 A.7 due, A.8 distributed
Assignments, June 2 A.8 due

14. FINAL EXAM

Class will meet during university assigned exam period (TBA) for in-class component of exam. Take-home component will be handed out at end of in-class component and is due in the political science office 48 hours later. Both the in-class and take-home component are to be completed individually, without any collaboration or assistance from others.