

Categorical Data Analysis – SOC507

Spring 2007

TuTh 10:40-11:55 AM

3 credit hours

Coor 5501

Instructor: Scott T. Yabiku

scott.yabiku@asu.edu

www.public.asu.edu/~syabiku

480-965-3943

Coor 5578

Office Hours: Friday 10-11 AM

Description

This course teaches you practical skills for interpreting and performing research with categorical data. Statistical techniques used to analyze categorical data include logistic regression, log-linear models, ordered logit and probit models, and models for unordered outcomes. I emphasize practical, hands-on skills with computing exercises in all of the above. The prerequisite for this course is a graduate course in statistics—something that has given you a working knowledge of data analysis, including hypothesis testing and linear regression.

Computing

I will be showing you how to use these techniques almost exclusively using SAS. You will need access to the unix stats machines (stats.asu.edu). Please do the following, if you have not done so previously:

1. Sign up for unix stats access. Go to <https://selfsub.asu.edu/selfsub> and subscribe to the service called "Statistics Unix Cluster." Once you do this, you will have access about 30 min later.
2. Download and install SSH Secure Shell software if you want to be able to work from your own PC. It's at <http://www.asu.edu/it/security/software/html/ssh/index.htm> and is necessary to connect to stats.asu.edu.
3. Increase your AFS quota to 50 MB. See <http://www.asu.edu/it/tempe/cac/accounts/quota.htm>

It is possible to work with SAS on a Windows PC, but I cannot offer help with how to use this interface or transport data from unix to PC data formats—you're on your own. Note that prior SAS programming experience is not necessary, but it does help and can reduce levels of frustration and anxiety. You may have to invest extra time if you have no prior experience.

Course Calendar

Topic	Reading	Assignments
Introduction and Basic Skills		
1/16 Introduction	PX Chapter 1, Allison Chapter 1	
1/18 Review of Linear Regression Model	PX Chapter 2, Allison Chapter 2 (5-11)	
1/23 Using SAS on Unix	<u>Meet in computer lab</u>	
Logit and Probit Models		
1/25 Introduction	PX Chapter 3, Allison Chapter 2 (11-80)	Give Asn 1
1/30 Continued		
2/1 with SAS	<u>Meet in computer lab</u>	
2/6 Applications	*Hummer (1993), *Akin et al (1981), *Bennett and Xie (2003), *Yabiku (2000)	
Loglinear Models for Contingency Tables		
2/8 Introduction	PX Chapter 4, Allison Chapter 10	Asn 1 + brief proposal due,
2/13 Continued		Give Asn 2
2/15 Continued		

2/20	Continued	Qian(1997)	
2/22	with SAS	<u>Meet in computer lab</u>	
2/27	Applications	*Ishii-Kuntz (1991), *Hayes (1991), *Bearman and Deane (1992), *Morris (1991), Raymo and Xie (2000)	

Models for Ordinal Dependent Variables

3/1	Introduction	PX Chapter 6, Allison Chapter 6	Asn 2 due , Give Asn 3
3/6	with SAS	<u>Meet in computer lab</u>	
3/8	Applications	*Ghimire et al. (2006), *Manning and Smock (1999), *Uzzi and Barsness (1998), *Cooksey and Fondell (1996)	
			Asn 3 due 3/18 (electronically)
3/20	Catch up/Review		
3/22	Midterm (closed book, in class)		

Models for Unordered Dependent Variables

3/27	Introduction	PX Chapter 7, Allison Chapter 5	Give Asn 4
3/29	No class & no office hours Fri 3/30		
4/3	with SAS	<u>Meet in computer lab</u>	Full proposal due
4/5	Applications	*Zimmer et al. (1998), *Malhotra (1991), *Mare (1980), *South and Crowder (1997), Zhang and Hoffman (1993)	

Brief Introduction to Further Topics

4/10	Models for Rates	PX Chapter 5, Heaton and Call (1995)	Asn 4 due , Give Asn 5
4/12	Discrete Choice Analysis	Allison Chapter 7, Davis et al. (2002)	
4/17	Poisson Regression	Allison Chapter 9, Pearce and Haynie (2004)	
4/19	Catch up + presentations		
4/24	4 Presentations		Asn 5 due
4/26	4 Presentations		
5/1	4 Presentations		Final Paper Due

Texts and Readings

1. Powers, Daniel A. and Yu Xie. 1999. *Statistical Methods for Categorical Data Analysis*. San Diego: Academic Press. [Referred to as PX in the syllabus]
2. Allison, Paul D. 1999. *Logistic Regression Using the SAS System*. Cary, NC: SAS Institute. [Referred to as Allison in the syllabus]

In addition, class notes are posted on my.asu.edu class webpage prior to class. Print and bring to class.

I will assign the articles listed below. Most are available through www.jstor.org; if not on JSTOR then see class website (articles with #). Readings preceded by * will be presented by a class member.

Logit and Probit Models

*Racial Differentials in Infant Mortality in the U.S.: An Examination of Social and Health Determinants
Robert A. Hummer, *Social Forces*, Vol. 72, No. 2. (Dec., 1993), pp. 529-554.

*The Determinants of Breast-Feeding in Sri Lanka
John Akin; Richard Bilborrow; David Guilkey; Barry M. Popkin; Daniel Benoit; Pierre Cantrelle; Michele Garenne; Pierre Levi, *Demography*, Vol. 18, No. 3 (Aug., 1981), pp. 287-307

*Revisiting Racial Differences in College Attendance: The Role of Historically Black Colleges and Universities
Pamela R. Bennett; Yu Xie, *American Sociological Review*, Vol. 68, No. 4 (Aug., 2003), pp. 567-580

* #Family History and Pensions: The Relationships Between Marriage, Divorce, Children, and Private Pension Coverage
Scott T. Yabiku, 2000, *Journal of Aging Studies* 14(3): 293-312.

Loglinear Models

* The Structure of Opportunity: Middle-Class Mobility in England, 1548-1689
Peter S. Bearman; Glenn Deane, *American Journal of Sociology*, Vol. 98, No. 1 (Jul., 1992), pp. 30-66

* Religious Identification and Marriage Patterns in Australia
Bernadette C. Hayes, *Journal for the Scientific Study of Religion*, Vol. 30, No. 4 (Dec., 1991), pp. 469-478

*Association Models in Family Research
Masako Ishii-Kuntz, *Journal of Marriage and the Family*, Vol. 53, No. 2. (May, 1991), pp. 337-348.

* #A Log-Linear Modeling Framework for Selective Mixing
Martina Morris, 1991, *Mathematical Biosciences*, 107: 349-377.

Temporal and Regional Variation in the Strength of Educational Homogamy
James M. Raymo; Yu Xie, *American Sociological Review*, Vol. 65, No. 5 (Oct., 2000), pp. 773-781

Breaking the Racial Barriers: Variations in Interracial Marriage Between 1980 and 1990
Zhenchao Qian, *Demography*, Vol. 34, No. 2. (May, 1997), pp. 263-276.

Ordinal Outcomes

* #Social Change, Premarital Non-family Experience and Spouse Choice in an Arranged Marriage Society
Dirgha J. Ghimire, William G. Axinn, Scott T. Yabiku, and Arland Thornton, *American Journal of Sociology*, 2006, 111: 1181-1218

*New Families and Nonresident Father-Child Visitation

Wendy D. Manning; Pamela J. Smock, *Social Forces*, Vol. 78, No. 1. (Sep., 1999), pp. 87-116.

*Contingent Employment in British Establishments: Organizational Determinants of the Use of Fixed-Term Hires and Part-Time Workers

Brian Uzzi; Zoe I. Barsness, *Social Forces*, Vol. 76, No. 3. (Mar., 1998), pp. 967-1005.

*Spending Time with His Kids: Effects of Family Structure on Fathers' and Children's Lives

Elizabeth C. Cooksey; Michelle M. Fondell, *Journal of Marriage and the Family*, Vol. 58, No. 3. (Aug., 1996), pp. 693-707.

Unordered Outcomes

*Educational Attainment and Transitions in Functional Status among Older Taiwanese

Zachary Zimmer; Xian Liu; Albert Hermalin; Yi-Li Chuang, *Demography*, Vol. 35, No. 3. (Aug., 1998), pp. 361-375.

*Gender and Changing Generational Relations: Spouse Choice in Indonesia

Anju Malhotra, *Demography*, Vol. 28, No. 4. (Nov., 1991), pp. 549-570.

*Social Background and School Continuation Decisions

Robert D. Mare, *Journal of the American Statistical Association*, Vol. 75, No. 370. (Jun., 1980), pp. 295-305.

*Residential Mobility Between Cities and Suburbs: Race, Suburbanization, and Back-to-the-City Moves

Scott J. South; Kyle D. Crowder, *Demography*, Vol. 34, No. 4. (Nov., 1997), pp. 525-538.

#Discrete-Choice Logit Models: Testing the IIA Property

Junsen Zhang; Saul D. Hoffman, *Sociological Methods and Research* (1993), Vol., 22, No. 2, pp. 193-213.

Other Topics

Modeling Family Dynamics with Event History Techniques

Tim B. Heaton; Vaughn R. A. Call, *Journal of Marriage and the Family*, Vol. 57, No. 4. (Nov., 1995), pp. 1078-1090.

Domestic and International Migration from Rural Mexico: Disaggregating the Effects of Network Structure and Composition

Benjamin Davis; Guy Stecklov; Paul Winters, *Population Studies*, Vol. 56, No. 3 (Nov., 2002), pp. 291-309

Intergenerational Religious Dynamics and Adolescent Delinquency

Lisa Pearce; Dana Haynie, *Social Forces* (2004), Vol. 82 No. 4, pp. 1553-1572.

Description of Course Requirements

The course requirements include 5 homework assignments, a midterm, leading class discussion on one or more articles, and a final paper project. In the final paper project, you will analyze data using methods for categorical data analysis; the topic and data are your choice. You will also briefly present your findings to the class.

To help you make progress on your final paper project, I require that you hand in research proposals (these proposals are ungraded; they are simply for your benefit). The first proposal is brief (a paragraph is fine), and you should describe generally what your topic is, what the dependent and independent concepts are, why you expect an association between them, and what kind of data you might use. The second proposal is more detailed. By this time, you should know what dataset you are going to use, what your variables probably will be, and what statistical method you need. This full proposal should be about 2-3 pages.

Grading

The grade is determined by the following components:

Homework (30%)

Midterm (30%)

Final Paper Project (30%)

Participation and Presentations (10%)

Once a final percentage is calculated, a letter grade is assigned: A=91-100%, B=81-90%, C=71-80%, D=61-70%, E=60% and less. The evaluation of work follows the standard University definitions: A=Excellent, B=Good, C=Average, D=Passing, E=Failure. I do not use plus-minus grading.

Note: These letter grade thresholds may be adjusted downward (i.e., to your advantage). They will not be adjusted higher.

Prerequisites

Experience with regression models and data analysis is helpful. An example of this is a course such as SOC505 (Applied Regression Analysis). Ask me if you have any questions about whether or not the course will be appropriate for you.

Other Relevant Course policies

Missed exams and late assignments:

I will not accept missed exams or late assignments without penalty, and they cannot be made up. Don't start working on homework the day before and ask for an extension—you won't get it.

Missed classes due to university-sanctioned activities:

Academic Affairs Policy Manual, 304-02: "Students who participate in university-sanctioned activities that require classes to be missed, should be given opportunities to make up examinations and other graded in-class work. However, absence from class or examinations due to university-sanctioned activities does not relieve students from responsibility for any part of the course work required during the period of the absence. ... Students should inform their instructors early in the semester of required class absences."

Class participation and attendance:

I will not keep a formal record of attendance. Keep in mind, however, that attending class, participating in discussions, and asking questions during class will help you learn.

Extra credit:

There is no extra credit.

Academic dishonesty:

It is a serious violation to cheat on exams, hand in work that is not your own, or plagiarize others' materials. Working on homework together, of course, is acceptable and a good way to learn. But copying others' work without contributing is not a good way to learn. Possible penalties include but are not limited to failure of the assignment, failure of the course, and failure of the course due to academic dishonesty. Additional penalties at the College or University level are possible.

Disability Accommodation:

Students with disabilities must identify themselves before the end of the second week of class.

Changes: This syllabus may be subject to minor changes; any changes will be both announced in class and posted on the class web page.