Course Title: APSTA-GE 2997: Generalized Linear and Multilevel Growth Curve Models

Course Instructor: Ying Lu

Number of Credits: 2

Meeting Pattern: once a week, 3 hours; 7 consecutive weeks.

Course Description:

This course is a second year course in advanced statistical techniques that covers useful quantitative tools in health, education, social science and policy research. Assuming a strong foundation in regression and the general linear model, this course focuses on data analysis that utilizes models for categorical, discrete or limited outcomes, as well as introducing growth curve modeling of these same outcome types. Examples are drawn from broad areas of applied research settings. In this course students will also learn the principles of likelihood-based inference, which will assist them in some of the more advanced statistics courses.

Course Prerequisites:

- REQUIRED: APSTA-GE 2003 Intermediate Statistical Methods or an equivalent course that covers linear regression analysis at intermediate level.
- RECOMMENDED: APSTA-GE 2042 Multilevel Models: Nested Data – this course pairs naturally with APSTA-GE 2997.

Learning Objectives:

By the end of the course, students will be able to:

1. Learn about the different characteristics of categorical and discrete outcomes and their inherent challenges when part of a data analysis.
2. Be introduced to and master the use of a wide range of statistical models that deal with categorical, discrete and limited outcomes, such as logistic/probit regression, Poisson regression models through examples.
3. Be introduced to and master the use of statistical models for longitudinal (growth curve) data in the context of limited dependent outcomes.
4. Build practical skills to apply appropriate statistical models and analyze data using statistical software.
5. Develop writing skills for quantitative research through class projects.
6. Gain deeper understanding of the statistical theory, specifically likelihood-based inference that underlies statistical practices.

Course Format:

2 hours of lecture followed immediately by 1 hour lab

Course Outline:

	Topic	Reading
Week 1	Review of Logistic model as introduction to maximum likelihood inference	C1(H)

Week 2	Introduction to Generalized Linear Models and likelihood theory; Probit model	C2(H) C2-4(HM) C3 (H) C9-11(HM)
Week 3	Ordinal and multinomial model	C4-5(H) C15-16(HM)
Week 4	Logistic regression with mixed effects	TBD
Week 5	Poisson Regression and overdispersion	C6(H) C12-14(HM)
Week 6	Negative binomial; Poisson models with random effects as a form of overdispersion	TBD
Week 7	Negative binomial with random effects	TBD

Required Readings and/or Text:
[H] John P. Hoffmann. 2004 *Generalized Linear Models: An Applied Approach*. Pearson. James [HM] Hardin and Joseph Hilbe. 2007. *Generalized Linear Models and Extensions using Stata.* Second edition, Stata Press. (Stata reference book)

Course Requirements:
The grade for this course will be determined as follows: 2 problem sets (20% each for a total of 40%), one data analysis projects (60% each).

The problem sets are designed to give students practice with the analytical tools and statistical software introduced in class. Roughly the problem sets will be assigned during weeks 3 and 5.

The data analysis projects are designed to give students experience analyzing real data and writing quantitative articles. The project reports need to be typed (in a word processor) and should be professional in appearance. The project will be released at the 5th week and it is due one week after the class ends. Students are expected to work on the projects independently. They are equivalent to take-home exams and there will be no in-class exams.

Grading Scale (cutpoints):
93% and up: A
90%: A-
87%: B+
83%: B
80%: B-
77%: C+
73%: C
70%: C-
67%: D+
63%: D
60%: D-
<60%: F

A=Excellent
The quality of project and homework is very high: thorough and careful data management, accurate implementations of statistical models and interpretations of the results, clear and consistent quantitative writing in final project.
B= Good
The quality of project and homework is satisfactory: decent data management, mostly accurate implementations of statistical models and interpretations of the results, clear quantitative writing in final project.
C=Average
The quality of project and homework is barely satisfactory: evidence of data management, reasonably accurate implementations of statistical models and interpretations of the results, complete final project.
D=Unsatisfactory
The quality of project and homework is unsatisfactory: lack of evidence of data management, inaccurate implementations of statistical models and interpretations of the results, complete final project.
F= Failed

Academic Integrity:

All students are responsible for understanding and complying with the NYU Steinhardt Statement on Academic Integrity. A copy is available at: http://steinhardt.nyu.edu/policies/academic_integrity.

Students with Disabilities:

Students with physical or learning disabilities are required to register with the Moses Center for Students with Disabilities, 726 Broadway, 2nd Floor, (212-998-4980 and online at http://www.nyu.edu/csd) and are required to present a letter from the Center to the instructor at the start of the semester in order to be considered for appropriate accommodation.