Course Title: APSTA-GE 2997: Advanced Data Analysis in Health and Policy Research: Generalized Linear Models

Number of Credits: 2

Meeting Pattern: once a week, 3 hours; 7 consecutive weeks beginning the 8th class week of the semester in 2014.

Course Description:

This course is a second year course in advanced statistical techniques that covers useful quantitative tools in health 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 that are commonly seen in health and policy studies. Examples are drawn from broad areas of health and policy research including determinants of self-reported health status or factors influencing number of clinic visits. 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: RESCH-GE 2003 Intermediate Statistical Methods or an equivalent course that covers linear regression analysis at intermediate level.

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 in public health and policy studies.
3. Build practical skills to apply appropriate statistical models and analyze data using statistical software.
4. Develop writing skills for quantitative research through class projects.
5. 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:

<table>
<thead>
<tr>
<th>Week</th>
<th>Topic</th>
<th>Reading</th>
</tr>
</thead>
<tbody>
<tr>
<td>Week 1</td>
<td>Linear regression review</td>
<td>C1(H)</td>
</tr>
<tr>
<td>Week 2</td>
<td>Introduction to Generalized</td>
<td>C2(H) C2-4(HM)</td>
</tr>
<tr>
<td>Week 3</td>
<td>Logistic model</td>
<td>C3 (H) C9-11(HM)</td>
</tr>
<tr>
<td>-------------</td>
<td>----------------</td>
<td>------------------</td>
</tr>
<tr>
<td>Week 4</td>
<td>Probit model</td>
<td>Same as above</td>
</tr>
<tr>
<td>Week 5</td>
<td>Ordinal and multinomial model</td>
<td>C4-5(H) C15-16(HM)</td>
</tr>
<tr>
<td>Week 6</td>
<td>Contingency tables and the loglinear model</td>
<td>TBD</td>
</tr>
<tr>
<td>Week 7</td>
<td>Poisson Regression and overdispersion</td>
<td>C6(H) C12-14(HM)</td>
</tr>
</tbody>
</table>

**Required Readings and/or Text:**


**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.