Categorical Data Analysis

PC 419

Wednesday 9am – 11:45am

Instructor  Stefany Coxe, Ph.D.

Office:  DM 275
Office hours:  by appointment
Email:  stefany.coxe@fiu.edu
Website:  http://faculty.fiu.edu/~scoxe

NOTE: Anything on this syllabus is subject to change at the Instructor’s discretion.

Course Description
This course covers topics related to statistical analysis of categorical outcome variables, focusing on methods used in the social sciences. Topics include chi-square and other non-parametric methods for categorical outcomes, the generalized linear model (GLiM, including logistic regression, Poisson regression, and survival analysis), and repeated measures extensions of GLiM (such as GEE and generalized linear mixed models). You will analyze, interpret, and write up results using these methods.

Prerequisite
Graduate coursework in analysis of variance and linear regression. We will cover a variety of topics in this course, but all of them build on a basic ANOVA and regression (general linear model) framework. A course covering multivariate statistics (such as PSY 5246C) is highly recommended but it is not required.

Teaching Assistant
Our teaching assistant, Kelly Cromer, will be available as an additional resource. You can contact her via email: kcromer@fiu.edu

Software
We will use SPSS and SAS in this course. I expect you to be able to use one of these software packages (e.g., open datasets, transform variables, conduct simple analyses, etc.). I will teach specific procedures for this course.

Blackboard
Lecture notes, computer code, assignments, and other materials will be posted on Blackboard. You should have them available during class.

Textbook
Generalized Linear Models for Categorical and Continuous Limited Dependent Variables
Michael Smithson and Edgar C. Merkle
ISBN: 978-1466551732
Other readings: Additional helpful readings are listed at the end of the syllabus.
Assignments

**Homework:** 5 assignments, 12% each

Five homework assignments covering the five major topic areas of the course: (1) non-parametric tests, (2) two- and three-way contingency tables; (3) generalized linear models for binary, ordered, and unordered categories; (4) generalized linear models for rates and counts; and (5) models for clustered and repeated-measures categorical outcomes.

The assignments involve running several analyses, making some decisions based on the analyses, interpreting output, and presenting the results in tables/figures and text.

**Assignment grading.** Homework assignments will be graded as if they are journal submissions; that is, you will receive written feedback and grades of ACCEPT (100), MINOR REVISIONS (95), MAJOR REVISIONS (85), REJECT (75). You may re-submit up to two assignments for which you receive a REJECT grade. For the re-submission, you should provide the original submission, your revised submission, and a “letter to the editor” explaining where you erred and how you revised the errors. Please contact me if you want more information about the revision procedure.

**Final Project:** 5% for proposal, 10% for presentation, 25% for final paper

I will post several datasets with categorical variables to Blackboard, along with brief descriptions of each. You will propose a project using one of these datasets or your own data. This will culminate in a short paper. I want you to focus on developing research questions and mapping them onto appropriate analyses. More details to follow throughout the semester.

**Proposal.** 2 to 3 page proposal due March 22. Similar to a thesis or dissertation proposal, but shorter.

**Presentation.** Presentations on April 12 and 19. (If you are unable to present on one of these dates, please contact me ASAP to make sure you get scheduled appropriately.) I expect that your analyses should be complete (or nearly so) at this point; preparing the presentation should help you organize your thoughts to write the paper. The main purpose of this presentation is to give you practice presenting your analysis findings in a group setting. Approximately 10 minutes.

**Paper.** The final paper is due April 26.

**Letter Grades**

<table>
<thead>
<tr>
<th>Grade</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>&gt;= 93.00</td>
<td>A</td>
</tr>
<tr>
<td>90.00 - 92.99</td>
<td>A-</td>
</tr>
<tr>
<td>87.00 - 89.99</td>
<td>B+</td>
</tr>
<tr>
<td>83.00 - 86.99</td>
<td>B</td>
</tr>
<tr>
<td>80.00 - 82.99</td>
<td>B-</td>
</tr>
<tr>
<td>77.00 - 79.99</td>
<td>C+</td>
</tr>
<tr>
<td>73.00 - 76.99</td>
<td>C</td>
</tr>
</tbody>
</table>
Course and University Policies

Attendance and Late Policy

• I shouldn’t have to tell you to attend every class. This is graduate school.

• Assignments are late if they are turned in after the due date and time. A 5 point late penalty will be deducted for each 24 hour period late — maximum score of 95/100 if 1 day late, maximum score of 90/100 if 2 days late, etc.

• Legitimate, verifiable cases of illness and emergencies, religious holy days, and conference travel can be accommodated. You need to contact me as soon as possible to make arrangements.

Drop Dates
Tuesday, January 17: Last day to drop courses or withdraw from the University without incurring financial liability for tuition and fees
Monday, March 20: Deadline to drop a course with a DR grade

Special Needs
Any student with a disability or other special need that may require special accommodations for this course should make this known to the instructor during the first week of class.

Disability Resource Center
Graham Center (GC) 190
(305) 348-3532
drcupgl@fiu.edu
drc.fiu.edu

Academic Misconduct
Florida International University is a community dedicated to generating and imparting knowledge through excellent teaching and research, the rigorous and respectful exchange of ideas, and community service. All students should respect the right of others to have an equitable opportunity to learn and to honestly demonstrate the quality of their learning. Therefore, all students are expected to adhere to a standard of academic conduct, which demonstrates respect for themselves, their fellow students, and the educational mission of the University. All students are deemed by the University to understand that if they are found responsible for academic misconduct, they will be subject to the Academic Misconduct procedures and sanctions, as outlined in the Student Handbook.

Academic Dishonesty
Please refer to your student handbook for a description of what constitutes academic dishonesty.

NOTE: Anything on this syllabus is subject to change at the Instructor’s discretion.
## Tentative Course Outline

<table>
<thead>
<tr>
<th>Date</th>
<th>Topics</th>
<th>Readings</th>
<th>Assignments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jan 11</td>
<td>Introduction</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Jan 18</td>
<td>Non parametric</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Jan 25</td>
<td>2 way cont</td>
<td></td>
<td>HW1</td>
</tr>
<tr>
<td>Feb 01</td>
<td>3 way cont</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Feb 08</td>
<td>GLIM Chapter 1</td>
<td></td>
<td>HW2</td>
</tr>
<tr>
<td>Feb 15</td>
<td>GLIM Chapter 2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Feb 22</td>
<td>GLIM Chapters 3 and 4</td>
<td></td>
<td>HW3</td>
</tr>
<tr>
<td>Mar 01</td>
<td>GLIM Chapter 5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mar 08</td>
<td>Intro to GEE and GLMM</td>
<td>Chapter 8</td>
<td>HW4</td>
</tr>
<tr>
<td>Mar 15</td>
<td>SPRING BREAK</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mar 22</td>
<td>GEE Chapter 8</td>
<td></td>
<td>Proposal</td>
</tr>
<tr>
<td>Mar 29</td>
<td>GLMM Chapter 8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Apr 05</td>
<td>Censoring?</td>
<td>Chapter 7</td>
<td>HW5</td>
</tr>
<tr>
<td>Apr 12</td>
<td>Presentations</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Apr 19</td>
<td>Presentations</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Apr 26</td>
<td>FINALS WEEK</td>
<td></td>
<td>Final paper</td>
</tr>
</tbody>
</table>
Extended Reading list

These are additional resources if you want to learn more about a specific topic. I used many of these resources when developing the course.

**Linear regression**


**Non parametric statistics**


**Contingency tables**


**Proportions, binary outcomes, and ordered categories**


Counts and frequencies


Clusters and repeated measures (GEE and GLMM)


