Applied Regression: Generalized Linear Models for Educational Research

Syllabus

Mondays & Wednesdays 5/21-6/20

Instructor:

Dr. M. Shane Tutwiler
710 Chafee
shane_tutwiler@uri.edu

Office hours:

By appointment

Required text:


Recommended texts:


Course Overview:

Building on the topics covered in EDP 613, the focus of this course is on expanding the multiple regression framework to include non-continuous outcome variables. Questions in educational research often involve trends in data that are non-linear. For example, one might ask the following (with suggested analytic method):

“What was the impact of a technology-enhanced reading curriculum on student learning across twenty schools in a given district?” (multilevel model)

“Did students who used a science-based video game lose interest in the technology over time?” (multilevel model for change)
“Controlling for key demographic factors, does prior knowledge predict the probability of students passing required exit examinations?” (logistic regression)

“What demographic factors predict teacher exit from the workforce over time?” (discrete time survival analysis)

“Do students with divorced parents experience more expulsions in a given year, controlling for free and reduced-price lunch status and gender?” (Poisson / negative binomial regression)

In this class, students will develop an understanding of how to identify, fit, and interpret the appropriate regression-based approach (generalized linear models) to answer a variety of important research questions. To do so, we will work through a variety of real-world data examples. The course will be divided into three general topics, as follows:

I. Continuous Outcomes
   a. The ubiquity of regression models & OLS review
   b. OLS residual assumptions (and what to do when they fail)
   c. Non-independent residuals – Multilevel Model
   d. Time as a predictor – Multilevel Model for Change

II. Categorical Outcomes
   a. Why linear models don’t (usually) fit binary outcomes
   b. Fitting & interpreting Logistic Regression models
   c. Testing nested Logistic Regression models
   d. Time to event occurrence – Discrete Time Survival Analysis

III. Counts of Events
   a. To transform or not to transform?
   b. Fitting & interpreting the Poisson Regression model
   c. Testing nested Poisson Regression models
   d. Over-dispersed count data (mean > standard deviation) – Negative Binomial Regression

Lectures:

Each lecture will begin with a discussion of an assigned reading, usually based on the analytic method discussed in the previous class. We will then use “real world” data examples to demonstrate the utility of the proposed methods during each lecture, and model appropriate interpretations and language throughout.

Workshop / Review Sessions

Application of statistical package in analyzing quantitative data will be emphasized in workshop sessions with hands-on experience of jamovi, a free, powerful software tool similar to SPSS. Workshops will be conducted in jamovi, but students who wish to use SPSS to complete their homework may do so (though technical support may not be as readily available).
**Course requirements:**

1. **Exams (50%)**
   - Two exams: midterm and final.
   - Open-book, open-note exams

2. **Attendance & Participation (20%)**
   - Attend all class meetings
   - Read assigned chapters from required textbook(s)
   - Read assigned research articles and complete reading guides
   - Answer reading guide questions when randomly called upon (or via online forum)

3. **Data Analytic Memos (DAMs) (30%)**
   - Analyze data to answer research questions.
   - Compose salient responses understandable and usable by naïve but intelligent readers

**Grades:**

Course grades will be based on the following final point ranges:

<table>
<thead>
<tr>
<th>Grade</th>
<th>Points</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>95 – 100</td>
</tr>
<tr>
<td>A-</td>
<td>90 – 94</td>
</tr>
<tr>
<td>B+</td>
<td>87 – 89</td>
</tr>
<tr>
<td>B</td>
<td>83 – 86</td>
</tr>
<tr>
<td>B-</td>
<td>80 – 82</td>
</tr>
<tr>
<td>C+</td>
<td>77 – 79</td>
</tr>
<tr>
<td>C</td>
<td>73 – 76</td>
</tr>
<tr>
<td>C-</td>
<td>70 – 72</td>
</tr>
<tr>
<td>D+</td>
<td>67 – 69</td>
</tr>
<tr>
<td>D</td>
<td>63 – 66</td>
</tr>
</tbody>
</table>

Attendance at ALL classes is both important and expected. If you have a valid reason for missing a class, it is your responsibility to contact the instructor and to get notes and information from other students.

If you have a documented disability that requires accommodations, please make an appointment with me prior to the third class meeting. We will discuss how to meet your individual needs to ensure your full participation and fair assessment procedures.
## Class Schedule

**Note:** JC. = Jamovi Chapter, Art. = Article, ALDA = Singer & Willett Chapter, GH=Gelman & Hill Chapter

<table>
<thead>
<tr>
<th>Week</th>
<th>Date</th>
<th>Reading</th>
<th>Topic</th>
<th>Note</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>5/21</td>
<td>JC. 3&amp;12 Art.1 (Quantitative Research) Optional: GH 3,4,6.1</td>
<td>I Continuous Outcomes (a-b) The ubiquity of regression models &amp; OLS review OLS residual assumptions (and what to do when they fail) Workshop</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>5/23</td>
<td>Art.2 (Linear Regression) Optional: ALDA 3&amp;4 GH 1,11,12</td>
<td>I Continuous Outcomes (c-d) Non-independent residuals – Multilevel Model Time as a predictor – Multilevel Model for Change Workshop</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>5/28</td>
<td>Art. 3 (Longitudinal Regression) Optional: G&amp;H 5,6,3</td>
<td>II Categorical Outcomes (a-b) Why linear models don’t (usually) fit binary outcomes Fitting &amp; interpreting logistic regression models Workshop</td>
<td>DAM1 due</td>
</tr>
<tr>
<td>2</td>
<td>5/30</td>
<td>Art. 4 (Logistic Regression) Optional: ALDA 9-11</td>
<td>II Categorical Outcomes (c-d) Testing nested logistic regression models Discrete time survival analysis (time as a predictor in logistic regression) Workshop</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>6/4</td>
<td></td>
<td>Review / Workshop</td>
<td>DAM2 due</td>
</tr>
<tr>
<td>3</td>
<td>6/6</td>
<td></td>
<td>Midterm (I a-d, II a-d)</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>6/11</td>
<td>Art. 5 (DTSA) Optional: GH 6.2</td>
<td>III Counts of Events (a-b) To transform or not to transform? Fitting &amp; interpreting the Poisson regression model Workshop</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>6/13</td>
<td>Art. 6 (Poisson Regression)</td>
<td>III Counts of Events (c-d) Testing nested Poisson regression models Options for over-dispersed data (negative binomial regression) Workshop</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>6/18</td>
<td></td>
<td>Review / Workshop</td>
<td>DAM3 due</td>
</tr>
<tr>
<td>5</td>
<td>6/20</td>
<td></td>
<td>Final (III a-d)</td>
<td></td>
</tr>
</tbody>
</table>

The above schedule is only approximate and may change to allow for expanding or reducing the coverage of certain topics.