Instructor: Dr. Zhehui Luo, B627 W. Fee Hall, 353-8623 × 161, email: zluo@msu.edu

Time and Location: Fr 12:40-3:30, A131 E. Fee Hall

Office Hours: Fr 4:00-5:00 or by appointment

Objectives: The first 60% of the course will focus on modeling count data using regression techniques, including Poisson regression, negative binomial regression, zero-inflation models, hurdle models, endogenous regressors, and longitudinal count data models. For PhD students in epidemiology who have not taken the comprehensive exams, these are materials for the exam.

The next 40% of the course will be like seminars. We will cover special topics suggested by former students of the course, including missing data, bootstrap, and quantile regression. A brief (10-15 pages double space) term paper is expected from each student by 5:00 p.m., Thursday, 12/15. The paper may take one of two formats:

- An empirical investigation that utilizes one or more models/techniques you learned in class on a real dataset, or

- A critical review of recent developments in a relevant topic of the biostatistics literature.

To avoid wasting your time and effort, please submit a 1-page paper outline by noon, Friday, 10/14. If you choose the second format, here are some potential topics:

1. Missing data in effect modifiers or confounders in generalized linear models,
2. Count data regression with endogenous categorical regressors, or
3. Quantile regressions with missing data in covariates.

Resources:


- Hilbe JM (2011). Negative Binomial Regression. 2nd Ed. Cambridge University Press. For data, errata, extensions, etc...


Materials: Lecture notes, slides, data and programs will be distributed or D2L.

Grading: Homework assignments (30%), a mid-term exam (40%) and a term paper (30%). The distribution of grades will be as follows.

\[
\begin{align*}
4.0 &= [90 - 100] & 3.5 &= [80 - 90] & 3.0 &= [70 - 80] \\
2.5 &= [60 - 70] & 2.0 &= [50 - 60] & 1.5 &= [40 - 50]
\end{align*}
\]
Tentative Course Content

<table>
<thead>
<tr>
<th>WEEK</th>
<th>DATE</th>
<th>TOPIC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Week 1</td>
<td>9/2</td>
<td>Introducing count data models</td>
</tr>
<tr>
<td>Week 2</td>
<td>9/9</td>
<td>Estimation and inference</td>
</tr>
<tr>
<td>Week 3</td>
<td>9/16</td>
<td>Assessing model fit</td>
</tr>
<tr>
<td>Week 4</td>
<td>9/23</td>
<td>Overdispersion and negative binomial models</td>
</tr>
<tr>
<td>Week 5</td>
<td>9/30</td>
<td>The problem with zeros</td>
</tr>
<tr>
<td>Week 6</td>
<td>10/7</td>
<td>Finite mixture models</td>
</tr>
<tr>
<td>Week 7</td>
<td>10/14</td>
<td>Endogeneity and some solutions</td>
</tr>
<tr>
<td>Week 8</td>
<td>10/21</td>
<td>Longitudinal count data</td>
</tr>
<tr>
<td>Week 9</td>
<td>10/28</td>
<td>Midterm exam</td>
</tr>
<tr>
<td>Weeks 10, 11</td>
<td>11/4, 11</td>
<td>Missing data</td>
</tr>
<tr>
<td>Weeks 12, 13</td>
<td>11/16, 25</td>
<td>Bootstrap</td>
</tr>
<tr>
<td>Weeks 14, 15</td>
<td>12/2, 9</td>
<td>Quantile regression</td>
</tr>
</tbody>
</table>

\textsuperscript{a} Open add period ends: 9/7 by 8:00 pm.
\textsuperscript{b} Need to reschedule the 9/9 class due to conflict with comprehensive exams.
\textsuperscript{c} Last day to drop with refund: 9/26 by 8:00 pm.
\textsuperscript{d} Last day to drop with no grade reported: 10/19 by 8:00 pm.
\textsuperscript{e} Need to reschedule the 11/25 class due to Thanksgiving holiday.

Academic integrity

All Michigan State University policies regarding academic integrity apply. For details, see here.

Collaboration on homework assignments is allowed, but each student must turn in written answers that reflect his or her own understanding of the materials. Academically dishonest behaviors will not be tolerated.