**Statistics 6650**  
**Discrete Data Analysis**  
**Spring 2020**

**Course Information:**  
Class Number 20232  
Lectures TR 12:40 – 2:30 p.m.  
136 Baker Systems Engineering

**Instructor:** Dr. Elizabeth A. Stasny, 319 Cockins Hall, 614-292-0784, stasny.1@osu.edu  
**Office Hours (Stasny):** Wednesday 2:15 – 3:15 p.m., and by prior appointment

**Course Assistant:** Mr. Azriel Krongauz, krongauz.1@buckeyemail.osu.edu  
**Office Hours (Krongauz):** Thursday 5:20-6:20 p.m., 134 Cockins Hall, and by prior appointment


You may also read the textbook online (one at a time) via the e-book at the OSU library. A link to the worldcat summary page is: http://osu.worldcat.org/title/introduction-to-categorical-data-analysis/oclc/368083243?ht=edition, and a direct link to the e-book that should work from computers on campus is: http://site.ebrary.com/lib/ohiostate/docDetail.action?docID=10278250.

This text is essentially a subset of the much more substantive *Categorical Data Analysis*. Make sure you are looking at the correct textbook for assignments.

Supplemental reading assignments may be posted on the course website.

**Other References:**
- Bishop, Fienberg, and Holland (1975), *Discrete Multivariate Analysis*  
- Christensen (1990), *Log-Linear Models*  
- Fienberg (1980), *The Analysis of Cross-Classified Categorical Data*  
- Hosmer and Lemeshow (1989), *Applied Logistic Regression*  
- Santner and Duffy (1989), *The Statistical Analysis of Discrete Data*

**Prerequisites:** excellent background in material from the Statistics courses 5302, 6450, 6950, PubHBio 6203 (or permission from the instructor), knowledge of regression, ANOVA, basic matrix algebra, and basic calculus.

**Computing:** I plan to use the Statistical package R for the computing needs in the course. R is freely available for most operating systems (http: //www.r-project.org/). If you are not already familiar with R, there are online tutorials available. I recommend getting started with swirl. This site http://swirllstats.com/students.html has instructions for downloading R, RShiny, and opening the swirl tutorial system. Once inside, you can learn the basics of working with R via the tutorial "R Programming: The basics of programming in R."

**Course Requirements:** You are responsible for all material covered in class, in assigned readings, and on homework assignments. You are expected to attend all classes.  
**Note:** for a graduate class, you should expect to spend three hours working outside of class for every hour of class time.
**Homework:** Assignments will be due at the start of the class period approximately weekly during the term. You must ensure that your homework easy to grade. Any computer output must be edited and annotated; graphs and plots must be clearly labeled and discussed in the text of the homework. Problems that are out of order or with parts not clearly identified may not receive full credit. It can be helpful to discuss course material with your colleagues outside of class, but your submissions must be your own and should demonstrate your personal understanding of the problems. **No late assignments will be accepted.** Solutions will be available on Carmen.

**Exams:**
- **Midterm:** Scheduled for Thursday, January 30, during the regular class time. A single sheet (8.5" x 11") of handwritten notes and a stand-alone calculator may be used for the exam.
- **Final Exam:** Scheduled for Tuesday, February 25, during the regular class time. Two sheets (8.5" x 11") of handwritten notes and a stand-alone calculator may be used for the exam.
- **Both Exams:** No cell phones, tablets, or other electronic devices will be permitted; nothing that obscures your eyes may be worn.

**Grades:** The final numerical grade will be determined as follows:

<table>
<thead>
<tr>
<th>Component</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Homework (lowest grade dropped)</td>
<td>25%</td>
</tr>
<tr>
<td>Midterm Exam</td>
<td>35%</td>
</tr>
<tr>
<td>Final Exam</td>
<td>35%</td>
</tr>
<tr>
<td>Class Participation*</td>
<td>5%</td>
</tr>
</tbody>
</table>

*You may earn two points towards the five possible points for class participation by giving me a good discrete data set from your research, the press, or a source other than a statistics text book!*
**Tentative Schedule of Topics**

(Actual schedule will evolve as the term progresses)

<table>
<thead>
<tr>
<th>Dates</th>
<th>Topic</th>
<th>Chapters in Text</th>
</tr>
</thead>
<tbody>
<tr>
<td>January 7</td>
<td>Introduction</td>
<td>Chapter 1</td>
</tr>
<tr>
<td>January 9 – January 21</td>
<td>Classical Methods</td>
<td>Chapter 2, selected sections Chapter 3</td>
</tr>
<tr>
<td>January 23 – February 13</td>
<td>Log-Linear Models</td>
<td>Chapter 7, selected sections Chapter 8</td>
</tr>
<tr>
<td>February 18 – February 20</td>
<td>Logistic Regression</td>
<td>Chapter 4</td>
</tr>
</tbody>
</table>

I also encourage you to read Chapter 11 for the historical perspective – it's fun reading!