
Home > SYLLABUS

Course Web Pages: Start at [http://www.stat.uiowa.edu/~jblang/s220](http://www.stat.uiowa.edu/~jblang/s220)

Lecture: MWF 11:30 - 12:20 14 SH
Instructor: Joseph B. Lang, 207 SH, 335-3129, joseph-lang@uiowa.edu
Office Hours: M 12:30-1:30, W 9:30-10:30, R 1:30-2:30, or by appointment
Pre-Requisites: 22S:165 and 22S:194 (or equivalents)
Department, College: Statistics and Actuarial Science, Liberal Arts and Sciences
DEO: Luke Tierney, 241 SH, 335-0712, luke-tierney@uiowa.edu
Main Office: 241 Schaeffer Hall

<table>
<thead>
<tr>
<th>Required Text</th>
<th>Supplementary Texts</th>
<th>Course Description and Objectives</th>
<th>Course Organization</th>
<th>Course Outline/Pace</th>
</tr>
</thead>
<tbody>
<tr>
<td>Required Text</td>
<td>Grading</td>
<td>Misc</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Required Text:


Supplementary Texts:


Course Description:

This course surveys theory and methods for the analysis of categorical response and count data. The course begins with an overview of likelihood-based inference for categorical data analysis. Methods for describing and analyzing contingency tables are surveyed. These include loglinear modeling of association structures, the Cochran-Mantel-Haenszel approach to detecting conditional association, linear predictor modeling using weighted least squares, and tests of marginal homogeneity. Dichotomous response models such as the logistic regression model will be described and applied in several settings including cohort and case-control studies. Poisson regression models will be used to analyze rate data from event history studies. Ordinal and polychotomous response models such as the cumulative and multinomial logit models will also be introduced. Time permitting, these regression models will be adapted and extended to accommodate longitudinal data. [See the course outline below.]

The statistical package SAS (e.g. PROC CATMOD, FREQ, GENMOD, and LOGISTIC) will be used extensively in this course. The freeware package R (or its commercial relative Splus) will be used to perform less standard analyses or small-scale simulation studies.

Note: SAS, Splus, and R are available on the HP machines in the UNIX Computing Lab (346 SH). SAS is also available in the Myers Computing Lab (41 SH). Students can purchase a personal copy of SAS, version 8, for around $10! The software R can be downloaded from http://cran.us.r-project.org to your personal computer.

Course Objectives:

The student who successfully completes this course should have a reasonable grasp of the theoretical foundations of categorical data analysis. As examples, the student will be able to derive and work with sampling distributions of sufficient counts. He or she will understand and be able to apply basic asymptotic techniques (e.g. multivariate central limit theorem and
delta method). He or she will be familiar with a variety of estimation methods and the basic properties of the corresponding estimators. He or she will be familiar with a variety of methods for analyzing categorical or count data and understand in what settings they are applicable. The successful student will have a working knowledge of the SAS procedures PROC FREQ, GENMOD, LOGISTIC, and CATMOD.

Course Organization:

**Lecture.** We will cover many of the topics introduced in Agresti (2002) and include several topics from Stokes, Davis, and Koch (2000). Although the majority of the meetings will be in the lecture/discussion format, several meetings will be set aside for student presentations and discussion.

**Homework.** Homework problems (applied, computational, and theoretical) will be assigned on a regular basis. Some will be handed in and graded. Solutions to select problems will be handed back to you or posted on the web. The textbook (Agresti 2002) includes a supporting website: [http://web.stat.ufl.edu/~aa/cda/cda.html](http://web.stat.ufl.edu/~aa/cda/cda.html), which includes solutions to odd numbered problems.

**Projects.** There will be two take-home projects (due 11:30am, Fri, March 9 and 11:30am, Mon, May 7). The projects will require you to address some combination of applied and theoretical problems. You must work alone on these two projects.

Course Outline/Pace [approx number of lectures in brackets]:

I. Intro to Categorical Data Analysis (AA Chapter 1) [10]

II. Intro to Contingency Tables (AA Ch. 2, 3, 10) [8]

III. Basic Asymptotic Tools [3]

IV. Regression Models for Categorical and Count Responses (AA Chapters 4-7)

   A. Dichotomous Response Models [9]
   B. Poisson Regression [1]
   C. Ordinal Response Models [3]

IV. Loglinear Models for Contingency Tables (AA Ch. 8-9) [6]

V. Regression Models for Correlated Categorical Responses (AA Ch. 11,12) [3]

   A. Multivariate Dichotomous Regression
   B. Multivariate Ordinal Regression

Course Guidelines and Policies:

**Reading Ahead.** It is vitally important that you read ahead. If the material in a lecture is completely new to you, you will find it very difficult to get much out of lecture.

**Participation and Attendance.** Students are expected to attend, and participate in, class. You will be asked many questions, and you will be strongly encouraged to ask lots of questions.
**Working Together.** Unless instructed otherwise (e.g. for the two class projects), you may work together on the homework problems. However, you must write up your own solutions in your own words. If you are personally asked to please write up your own solutions and subsequently turn in material that is obviously in the same words as a fellow student, the work will be considered to be plagiarized. Plagiarism will be dealt with according to the policies of the University.

**Late Homework.** Unless otherwise instructed, homework is due at 11:30 AM. Late homework has a half-life of 24 hours; that is you get 50% credit if it is handed in late, but within 24 hours of the due time; you get 25% credit for the next 24 hours; etc. Homework not handed in directly to me must be handed in to a department secretary (located in 241 SH). The homework must include a hand-in time and date, and must be signed by the department secretary. (It follows that you cannot hand in homework after the main office is closed.)

**Grading Questions.** Questions about grading must be asked within one week of the graded work's return.

**Grading and Components for Evaluation**

Your final score $S$ will be computed as $S = 0.5H + 0.2P_1 + 0.2P_2 + 0.1P$, where $H =$ percent correct on homework, $P_1$ and $P_2$ are the scores on the two take-home projects, and $P =$ participation score.

Letter grades (including +'s and -'s) will be awarded according to a 90-80-70-60 schedule (e.g. if $S \geq 90$ then a grade of A- or better will be awarded). Class participation will be considered when a student "falls on the borderline" between two grades. These are guaranteed cutoffs, so it is possible (but unlikely) that everyone receives an 'A.' I do, however, reserve the right to lower (but not raise) the cutoffs. Note that with this grading scheme you are not "graded on a curve," and so you are not competing with fellow students. Therefore, you are not penalized for working together to better understand concepts.

**Miscellaneous**

**Textbook (Agresti 2002) support:**

Start at http://web.stat.ufl.edu/~aa/cda/cda.html. (Includes datasets and solutions to odd numbered problems.)

**Help outside of class:**

I have regular office hours. Sometimes it is effective to ask specific questions via email.

Course web pages; start at http://www.stat.uiowa.edu/~jblang/s220.

**Help with SAS:**

- Creating (temporary and permanent) SAS data sets
- Some helpful links...
  - SAS in a nutshell, SAS basics for Windows, SAS INSIGHT (PDF file), SAS INSIGHT intro, SAS basics (inc. INSIGHT)
  - or go to search engine Google, enter the key words "SAS basics"

- Use SAS Options statement

- Using SAS ODS (Output Delivery System)

Help with R software:

- An Introduction to R, by Elizabeth Slate and Elizabeth Hill.
- R for CDA. Go to http://www.stat.ufl.edu/~presnell/Teaching/sta4504-2000sp/R/

Information for students with disabilities: I would like to hear from anyone who has a disability which may require some modification of seating, testing, or other class requirements so that appropriate arrangements can be made. Please see me about this in my office as soon as possible. (First visit the University of Iowa Student Disability Services webpage for information about the procedures for requesting accommodations.)

Class Policy: Class policies are governed by the College of Liberal Arts and Sciences (CLAS), the administrative home of this course. Policy details can be found in the CLAS student handbook.

Complaint Procedure: I am reminded to tell you that there are procedures available to you if you have complaints about the oral communication competence of the instructor in this course, the teaching competence of his TA's, or other aspects of the obligations of faculty and TA's that are covered by University policy. Ask me or one of the administrative personnel in the Department of Statistics and Actuarial Science for more information. You can also find relevant information in the CLAS student handbook, especially Section IX, Student Rights and Responsibilities.

Academic Misconduct: The College of LAS considers academic fraud, dishonesty, and cheating serious academic misconduct. I adhere to the College's policy, a summary of which follows: A student who plagiarizes [in this course, taking answers from the student solution manual is not plagiarism] or cheats on any assignment or exam faces penalties that may include an 'F' on the work or an 'F' in the course. If I suspect academic fraud there are very specific procedures I am obligated to follow. These are outlined in Section IX, Student Rights and Responsibilities of the CLAS Student Academic Handbook. Please see the additional university policies that also describe (a) complaint procedures and (b) student academic (mis)conduct.

I hope you all have an enjoyable and successful semester. Good luck in all of your courses.

This page was last updated: 01/17/2007 10:13:48 AM -0600 (Joseph B. Lang)