PUBHLTH 490ST (3 credits)
Telling Stories with Data: Statistics, Modeling, and Data Visualization
Spring 2016 :: T/Th 2:30-3:45 :: Arnold 120

INSTRUCTOR
Professor Nicholas G Reich Ph.D.
425 Arnold House
(413) 545-4534
nick [at] umass.edu
Reich Lab website
on twitter: @reichlab

Office Hours: TBD
Course website Piazza site

MATERIALS

Required Textbook

Kaplan, Daniel T. 2011. *Statistical Modeling: A Fresh Approach*.

Recommended Textbooks (all freely available online)

Faraway JJ. 2002. *Practical Regression and Anova using R*.

James G, Witten D, Hastie T, and Tibshirani R. 2014. *An Introduction to Statistical Learning*.

Diez D, Barr C, and Çetinkaya-Rundel M. 2012. *OpenIntro Statistics, 2nd Ed.*.

Software

R :: r-project.org (or just Google "r")

RStudio :: rstudio.org

Prerequisites

One of any of the following introductory stats courses taught at UMass: BIOSTAT 391B, STAT 111, STAT 240, STAT 501, ResEcon 212, PSYCH 240. If you have not taken an intro stats course at UMass but still want to enroll in this course, you are encouraged to petition the instructor for permission, especially if any of the following apply: (a) you have taken AP Stats in high school, (b) you have taken a college-level intro stats course just not one of the ones listed above, or (c) you are confident in your quantitative skills and your ability to succeed in a fast-paced, advanced introductory course.

Course Description

The aim of this course is to provide students with the skills necessary to tell interesting and useful stories in real-world encounters with data. Specifically, they will develop the statistical and programming expertise necessary to analyze datasets with complex relationships between variables. Students will gain hands-on experience summarizing, visualizing, modeling, and analyzing data. Students will learn how to build statistical models that can be used to describe and evaluate multidimensional relationships that exist in the real world. Specific methods covered will include linear, logistic, and Poisson regression. This course will introduce students to the R statistical computing language and by the end of the course will require substantial independent programming. To the extent possible, the course will draw on real datasets from biological and biomedical applications. This course is designed for students who are looking for a second course in applied statistics/biostatistics (e.g. beyond BIOSTATS 391B or STAT 240), or an accelerated introduction to statistics and modern statistical computing.

Learning Goals *(By the end of the course students will be able to...)*

- understand and critique statistical model equations as representations of a given real-world setting,
- independently formulate, fit, and interpret statistical models to weigh evidence for/against hypotheses about associations between variables,
- diagnose the appropriateness or "goodness-of-fit" of a given model,
- independently write code in R, the language of modern statistical computing,
- create powerful data visualizations that reveal features of data or fitted models,
- write concise, professional, and reproducible statistical analysis reports using `knitr` and RMarkdown.

Types of Assignments and Activities, with Grade Contributions

Homework (35%)
Quizzes (15%)
Midterm exam (20%)
Final Project (20%)
Participation/citizenship (10%)

Course Schedule

Week 1	Introduction, motivation, and overview
Week 2	Understanding and visualizing data
Week 3	The Language of Models
Week 4	Model formulas and coefficients
Week 5	Fitting models to data
Week 6	Confidence and uncertainty in models
Week 7	Midterm review and exam
Week 8	Logistic Regression
Week 9	Poisson Regression
Week 10	Smooth splines
Week 11	Case study: NHANES dataset
Week 12	Case study: NCHS dataset
Week 13	Case study: Infant Health
Week 14	Final projects

Grading Scale

Grade	Percentage
A	93-100
A-	90-92
B+	87-89
B	83-86
B-	80-82
C+	77-79
C	73-76
C-	70-72
D+	67-69
D	63-66
D-	60-62
F	0-59

Council on Education for Public Health (CEPH) Course Competencies

- Distinguish among the different measurement scales and the implications for selection of statistical methods to be used based on these distinctions.
- Describe conceptual frameworks (statistical literacy) in biostatistics
- Apply biostatistical methods to the design of studies in public health.
- Use computers to appropriately store, manage, manipulate and process data for a research study using modern software.
- Apply descriptive techniques commonly used to summarize public health data.
- Describe the basic concepts of probability, random variation and selected, commonly used, probability distributions.
- Select and perform the appropriate descriptive and inferential statistical methods in selected basic study design settings.
- Describe appropriate methodological alternatives to commonly used statistical methods when assumptions are violated.
- Integrate analysis strategies in biostatistics with principles and issues in epidemiology. literature
- Develop written and oral presentations based on statistical analyses for both public health professionals and educated lay audiences.
- Apply statistical methods to solve problems in the health sciences and carry out theoretical research in statistical methodology.

Academic Honesty Policy Statement

Since the integrity of the academic enterprise of any institution of higher education requires honesty in scholarship and research, academic honesty is required of all students at the University of Massachusetts Amherst. Academic dishonesty is prohibited in all programs of the University. Academic dishonesty includes but is not limited to: cheating, fabrication, plagiarism, and facilitating dishonesty. Appropriate sanctions may be imposed on any student who has committed an act of academic dishonesty. Instructors should take reasonable steps to address academic misconduct. Any person who has reason to believe that a student has committed academic dishonesty should bring such information to the attention of the appropriate course instructor as soon as possible. Instances of academic dishonesty not related to a specific course should be brought to the attention of the appropriate department Head or Chair. The procedures outlined below are intended to provide an efficient and orderly process by which action may be taken if it appears that academic dishonesty has occurred and by which students may appeal such actions. Since students are expected to be familiar with this policy and the commonly accepted standards of academic integrity, ignorance of such standards is not normally sufficient evidence of lack of intent. For more information about what constitutes academic dishonesty, please see the Dean of Students' website.

Disability Statement

The University of Massachusetts Amherst is committed to making reasonable, effective and appropriate accommodations to meet the needs of students with disabilities and help create a barrier-free campus. If you are in need of accommodation for a documented disability, register with Disability Services to have an accommodation letter sent to your faculty. It is your responsibility to initiate these services and to communicate with faculty ahead of time to manage accommodations in a timely manner. For more information, consult the Disability Services website.