University of Wisconsin-Madison
Regression Methods for Population Health
Pop Hlth/BMI 552
3 credits

Canvas Course URL
https://canvas.wisc.edu/courses/86717

Meeting Time and Location:
4:00 - 5:15pm
2305 (or 2309) Signe Skott Cooper Hall (Nursing)

Instructional Mode:
All Face-to-Face

How Credit Hours are Met:
This class meets for twenty-eight 75-minute lectures (on average, two per week) and fourteen 75-minute lab sessions (on average, one per week) over the spring semester. It also carries the expectation that students will work on course learning activities (reading, writing, problem sets, studying, etc) for about 3 hours out of classroom for every lecture.

Course Instructor

Ronald Gangnon
Professor, Biostatistics and Population Health
Room 603 WARF Office Building
ronald@biostat.wisc.edu
Office hours: By appointment

Teaching Assistant

Lan Luo
Room 785 WARF Office Building
lluo24@wisc.edu
Office hours: 2:30-3:30 Mondays and Thursday, or By appointment

Official Course Description

Introduction to the primary statistical tools used in epidemiology and health services research; multiple linear regression, logistic regression and survival analysis.

Prerequisites

Pop Hlth/BMI 451 and Pop Hlth/BMI 551

Learning Objectives

The course is an accessible introduction to the primary statistical tools used in epidemiology and health services research: multiple linear regression for continuous outcomes, logistic regression for binary outcomes and Cox proportional hazards regression for time-to-event outcomes, with emphasis on their proper use and interpretation. Analyses of a variety of real data sets will receive considerable attention. Non-essential mathematical complexities will be avoided. The development of the ability to interpret results and to evaluate critically the methods used is of paramount importance.

Upon successful completion of the course, students will be able to

1. State the assumptions underlying linear, logistic, survival and Poisson regression models, recognize and address violations of those assumptions, and estimate and interpret regression models to answer epidemiologic and public health research questions.
2. Critique uses of linear, logistic, survival and Poisson regression models in the epidemiologic and public health literature.
3. Translate epidemiologic concepts into statistical modeling assumptions, and explain statistical modeling assumptions in epidemiologic terms.
4. Recognize applications that require methods beyond their expertise, and identify resources to learn about more advanced techniques.

"Statistics is, or should be, about scientific investigation and how to do it better, but many statisticians believe it is a branch of mathematics…. Now I agree that the physicist, the chemist, the engineer, and the statistician can never know too much mathematics, but their objectives should be better physics, better chemistry, better engineering, and in the case of statistics, better scientific investigation. Whether in any given study this implies more or less mathematics is incidental." –George Box

Course Requirements and Evaluation

1. Staying current with the assigned readings and participating in class.
2. Participating in the lab sessions and completing the lab exercises.
3. One data analysis assignment.
4. One cumulative final exam.

Letter grades will be assigned based on natural break points in the distribution of cumulative scores.

Textbooks

Regression with Linear Predictors. 2010. Andersen PK and Skovgaard LT. Springer: New York. ISBN: 978-1-4614-2627-1. EBook.

Applied Linear Regression, 4th Edition. 2014. Weisberg S. Wiley: Hoboken, NJ. ISBN: 978-1-118-38608-8. EBook.

Multivariable Analysis: A Practical Guide for Clinicians and Public Health Researchers, 3rd Edition. 2011. Katz MH. Cambridge University Press: Cambridge. ISBN: 978-0-521-14107-9. Ebook.

Software

You are welcome to use any statistical software for the data analysis assignments and lab exercises. The instructors will provide support for SAS, and course instruction will use SAS. If you choose to use other software, instructors will provide as much assistance as they can, but support cannot be guaranteed.

There are several options for getting access to SAS software at no charge. SAS University Edition is the recommended option. It is available for all computer operating systems (Windows, Linux, Mac) from https://www.sas.com/en_us/software/university-edition.html.

The regular edition of SAS is available for Windows and Linux operating systems. It can be downloaded from the UW-Madison Campus Software Library (https://software.wisc.edu/cgi-bin/ssl/csl_download.cgi). (Note that these files are very large. They will take a long time to download even on a wired connection. It is strongly recommended that you not use a wireless connection to download these files.)

If you have not taken Pop Hlth/BMI 451 Introduction to SAS for Population Health, you should review the Introduction to SAS class notes from the UCLA Institute for Digital Research and Education (https://stats.idre.ucla.edu/sas/seminars/notes/).

Course Schedule

Date	Topic	Room (Signe Skott Cooper Hall)
1/22	Course Introduction	2305
1/24	Least Squares Estimation	2309
1/25	Diagnostics	2309
1/28	Inference and Prediction	2309
1/29	Binary Predictor	2305
1/31	Lab 1: Binary Predictor (Sex) Description of Measurement Due	2305
2/4	Categorical Predictor	2305
2/5	Lab 2: Categorical Predictor (Race/Ethnicity)	2305
2/7	Quantitative Predictor (Splines)	2305
2/8	Quantitative Predictor (Linear)	2305
2/11	Lab 3: Quantitative Predictor (Age)	2305
2/12	Lab 4: Quantitative Predictor (Age)	2305
2/18	Multiple Predictors (Additive)	2305
2/19	Multiple Predictors (Interaction)	2305
2/21	Lab 5: Multiple Predictors (Age, Race, Sex)	2305

2/22	Lab 6: Multiple Predictors (Age, Race, Sex)	2305
2/25	Weighted Least Squares	2305
2/26	Robust Standard Errors	2305
2/28	Lab 7: WLS and Robust Standard Errors	2305
3/4	Bootstrap	2305
3/5	Transformations	2305
3/8	Lab 8: Bootstrap and Transformations Preliminary Analyses Due	2305
3/11	Predictor Selection	2305
3/12	Logistic Regression	2305
3/14	Likelihood Inference	2305
3/15	Labs 1-8 Due	
3/28	Lab 9: Logistic Regression Data Analysis Project Due	2305
4/1	Conditional Logistic Regression	2305
4/2	Alternative Link Functions	2305
4/4	Lab 10: Alternative Link Functions	2309
4/8	Nominal Multinomial Outcome	2305
4/9	Ordinal Multinomial Outcome	2305
4/11	Lab 11: Multinomial Regression	2309
4/15	Poisson Regression	2305
4/16	Negative Binomial Regression	2305
4/18	Lab 12: Count Regression	2305
4/22	Survival Data	2305
4/23	Kaplan-Meier Estimate	2305
4/25	Lab 13: Kaplan-Meier Estimate Labs 9-12 Due	2305
4/29	Proportional Hazards Regression	2305
4/30	Lab 14: Proportional Hazards Regression	2305
5/2	Time-Dependent Covariates	2305
5/3	Non-Proportionality Labs 13-14 Due	2305
5/7	Final Exam	1309 HSLC

Lab Sessions

Attendance and participation in the lab sessions is mandatory. In a typical lab session, students will be assigned to small groups to work on two hand written and/or computer-based exercises

related to the material covered in prior lecture(s). If you are unable to attend a lab session due to extenuating circumstances, please notify the course instructor in advance to make alternative arrangements to complete the exercises. You will periodically submit written solutions to the lab exercises for grading. Labs 1-8 will be due on 3/15, Labs 9-12 will be due on 4/25, and Labs 13-14 will be due on 5/3. Each lab exercise will be worth 2% of your final grade.

Data Analysis Project

For this project, you will perform an analysis of data from the Survey of the Health of Wisconsin (SHOW). You will investigate the association between a quantitative outcome variable (anthropometry, blood pressure, laboratory measurements or spirometry) and three predictor variables: sex (a binary predictor), race/ethnicity (a categorical predictor) and age (a quantitative predictor).

You will analyze the data in a full and appropriate fashion using appropriate methods and techniques. You will prepare a complete written report of the analysis. A good report will include the following elements: (1) a statement of the problem and why it is interesting, (2) a description of the data and its source, (3) the research questions of interest, (4) any graphical or statistical methodology used, (5) discussion of the results obtained, including any adjustments to the data or corrective actions that might have been taken, (6) implications of your results for the scientific problem and/or public health, and (6) a summary of what you've learned from the analysis. You should report and summarize the implications of your "best" model even if you ultimately decide that it is deficient for some reason (you should also discuss why you think it's deficient, and what you might be able to do to fix it, even if you don't have the tools yet to do so). Remember "all models are wrong; the practical question is how wrong do they have to be to not be useful."

A brief description of your outcome variable (worth 4% of your final grade) will be due on 1/31. A guided series of preliminary analyses (worth 12% of your final grade) will be due on 3/8.

You will turn in your report on 3/28. The data analysis project will be worth 20% of your final grade.

Final Exam

There will be a cumulative, in-class final examination. For the exam, you will need a calculator, and you can bring two pages of notes. The exam will be held from 4-7pm on 5/7 in 1309 HSLC. The final exam will be worth 20% of your final grade.

Academic Integrity

By enrolling in this course, each student assumes the responsibilities of an active participant in UW-Madison's community of scholars in which everyone's academic work and behavior are held to the highest academic integrity standards. Academic misconduct compromises the integrity of the university. Cheating, fabrication, plagiarism, unauthorized collaboration, and helping others commit these acts are examples of academic misconduct, which can result in disciplinary action. This includes but is not limited to failure on the assignment/course,

disciplinary probation, or suspension. Substantial or repeated cases of misconduct will be forwarded to the Office of Student Conduct & Community Standards for additional review. For more information, refer to https://conduct.students.wisc.edu/academic-integrity/

Accomodations for Students with Disabilities

The University of Wisconsin-Madison supports the right of all enrolled students to a full and equal educational opportunity. The Americans with Disabilities Act (ADA), Wisconsin State Statute (36.12), and UW-Madison policy (Faculty Document 1071) require that students with disabilities be reasonably accommodated in instruction and campus life. Reasonable accommodations for students with disabilities is a shared faculty and student responsibility. Students are expected to inform faculty [me] of their need for instructional accommodations by the end of the third week of the semester, or as soon as possible after a disability has been incurred or recognized. Faculty [I], will work either directly with the student [you] or in coordination with the McBurney Center to identify and provide reasonable instructional accommodations. Disability information, including instructional accommodations as part of a student's educational record, is confidential and protected under FERPA.

Diversity and Inclusion

"Diversity is a source of strength, creativity, and innovation for UW-Madison. We value the contributions of each person and respect the profound ways their identity, culture, background, experience, status, abilities, and opinion enrich the university community. We commit ourselves to the pursuit of excellence in teaching, research, outreach, and diversity as inextricably linked goals.

The University of Wisconsin-Madison fulfills its public mission by creating a welcoming and inclusive community for people from every background – people who as students, faculty, and staff serve Wisconsin and the world." https://diversity.wisc.edu/