BIOS 52103/62133 Spring 2025 Syllabus

Aaron R. Caldwell

2025-01-06

Table of contents

Course Information

Course Title: Biostatistics II: Advanced Linear Models

Section: 001

Semester: Spring 2025

Credit Hours: 3

Prerequisite: Biostatistics I or equivalent. Or my approval.

Meeting Location: Online.

- Material lectures via VODs (Videos On Demand) with links in BlackBoard.
- Supplemental class review meetings via Zoom/Teams
  - **ATTENDANCE IS NOT REQUIRED**
  - Attendance is recommended, when possible
  - Weekly *review* of material
  - Meetings will be approximately one hour or less.
  - Links for Zoom/Teams meetings will be sent via BlackBoard announcements.
  - The day/time of the meetings will be announced weekly and may change at my discretion.
  - Meetings will be recorded and, barring any technical issues, posted to BlackBoard.

Course Webpage: BlackBoard https://uams.blackboard.com

Course Description: This is a course covering the basics of linear regression models. This material includes:

- Multiple regression and linear models for analysis of variance.
- Experimental Designs with factorial arrangement of treatments, and multiple covariates.
- Introduction to logistic regression.

Course Instructor

Name: Aaron Caldwell, PhD

Office: Teams, unless you're visiting UAMS Northwest campus

E-mail: caldwellaaron@uams.edu

Office hours: Meet via Teams. E-mail for an appointment. My time is fairly flexible.

Course Objectives

This course is meant to help you understand (and hopefully do!) some of most common tools in standard statistical analyses:

- Linear regression: Can we please use "lines" to model data? Probably…
  - Simple Linear Regression
  - Multiple Linear Regression
- Analysis of Variance (ANOVA): Applying linear models to groups.
  - One-Way ANOVA
  - Block Design
  - Factorial ANOVA
- Analysis of Covariance (ANCOVA): How categorical variables can be included in linear regression.
- Generalized Linear Models: Other ways to view data as "linear".
  - Logistic Regression: Turning modeling probabilities/odds into linear regression.

Potential Extra Topics

- Repeated measures: What to do when we keep track of things over time. (Briefly… its complicated.)

Online Class Review Meetings

This is an online and *asynchronous* course and all lectures will be recorded and posted onto Blackboard. There are two main components that are incorporated into this course's conceptualization of lectures.

1. Lectures Videos on Demand (VODs).

- Each major topic will have several videos that range from 5 to 20 minutes each on various topics for each chapter of the course's content.
- These videos are meant to be how you learn the details of the statistical methods that are covered in this course.
- The videos will cover **all** the content you are expected to learn.
- Additional videos are provided to supplement this content.

2. Weekly live review meetings via Teams.

- *optional* attendance.

- They are planned to be approximately one hour long.
- The objective will be to take the a week's material, review, and apply it when necessary.
- The idea is we explore data using the techniques learned and see if we can answer questions about the data using the techniques covered during the week's lecture VODs.
- This is also meant to be a session where students can ask questions about the material and see how it relates to what we do with statistics.

Required Materials

- I will be creating a sort of miniature textbook for the course. It will include any and all content you are expected to learn in the course.
  - I make sure to put in enough details where this is all you will need.
  - There is not a required textbook; I provide pdfs of the open textbooks we will read.
- Referenced reading texts will be posted on BlackBoard.
  - I will be pulling from several different resources.
- Lecture notes will be a truncated version of the chapters of this "textbook".

Books and Suggested Texts

None of the texts are required, but they are helpful for this course. The blackboard page for this course will host all the PDFs!

Click book title for link to source.

Introductory Statistics for the Life and Biomedical Sciences by Julie Vu and Dave Harrington

- Here is a free text that covers the basics of some course topics.
  - You can pay $0 for the PDF version or tip them how ever much you want.
  - It can be used to review topics from Biostatistics I.
  - Relevant chapters for what we will cover:
    - Chapter 5, section 6: Comparing means with ANOVA
    - Chapter 6: Simple linear regression
    - Chapter 7: Multiple linear regression

Learning Statistics with R - A tutorial for Psychology Students and other Beginners by Danielle Navarro

- This is course that covers some of the basics and is *very* helpful with learning R

  - Reviews topics for Biostatistics I
  - Introduces the topics of this course in the later chapters
  - Relevant chapters:
    * Chapters 14-16: ANOVA, linear regression, and factorial ANOVA

A Second Course in Statistics: Regression Analysis, by William Mendenhall and Terry Sincich

- Here is the main text that defines what methods we will cover. It is considered an intermediate level text, hence the name I would suppose.
  - It is a good reference text.
  - It would not matter much which edition you have, but more recent is better. Like phones!
  - Relevant chapters:
    * Pretty much all of it.
    * Maybe not Chapter 10: Time and forecasting (a bit dense of a subject in practice)

Regression Modeling Strategies, by Frank Harrell

- Fantastic textbook with some online resources at https://hbiostat.org/rmsc/
- Relevant chapters are 1-2, 4, 5, and 10.
- A *lot* more detail than we need in this course, but very useful nonetheless.

Applied Linear Statistical Models by Michael Kutner, Christopher Nachtsheim, John Neter, and William Li

- Here is another far more extensive text book.
  - Apparently you can get the PDF for free via the link.
  - It is far more technical and detailed.
  - It is great for people that are more mathematically inclined.
  - Relevant Chapters:
  - 1 Through 25
  - Does not include Logistic Regression and Odds analysis.

An Introduction to Statistical Learning, by Gareth James, Daniela Witten, Trevor Hastie and Robert Tibshirani

- Here is another text that is an advanced (the name is deceptive) overview of what we will cover.
- It's another free PDF book. (You can buy hardback if you want.)
- It can be dense and indecipherable because of math sometimes. I suppose this course might be, too.

- It does provide lots of R examples if you want to get into coding.
- Relevant chapters:
  - Chapter 1: Introduction
  - Chapter 2: Statistical learning
  - Chapter 3: Linear regression
  - Chapter 4, Sections 1 - 3: Logistic regression
  - Chapter 6, Section 1: Subset selection

Course Competencies

This course is designed to address specific competencies that students are expected to master during the course of their degree program. No single course is likely to provide you with all the skills you need to be successful in a given area; however, each course contributes to your overall competency development. The list of competencies for each degree program can be found at http://publichealth.uams.edu/students/current-students/competencies/. The list of specific competencies addressed by this course, by degree program, is included in the Appendix.

Software

Statistical analyses of any type require software. I will be using and teaching how to use R/RStudio. It is not required to use this software. **If you have knowledge of other statistical software and want to use that to complete homework and exams, go right ahead!**

- R is the base software/programming "language" that allows you to perform pretty much any statistical analyses in existence.
- RStudio makes R easier to use. R must be installed first.
- They called it R because there used to be this software called S. *R*oss Ihaka and *R*obert Gentleman created the langauge (hence R).

Here is where you can get R and RStudio.

- Mac
  - R: https://cloud.r-project.org/bin/macosx/
  - RStudio: https://rstudio.com/products/rstudio/download/#download
- Windows:
  - R: https://cloud.r-project.org/bin/windows/
  - RStudio: https://rstudio.com/products/rstudio/download/#download

- You can also run R/Rstudio in your web browser using the "cloud" using the Posit Cloud. If you have trouble installing R this might be an option!

I used the term programming language. That may sound intimidating. Try not to worry. I will provide most of the code you will need to do anything. You will just need to add/change a few things to the code to make it work for the particular analysis.

Most the stuff we do in this course requires only a few lines of code!

Alternative Software

If R is too intimidating you can use software that runs R in the background. These use graphical user interfaces (GUIs) to make the experience more user friendly. *I strongly recommend trying to learn R though!*

1. Jamovi (preferred)
2. JASP

Course communication

Announcements will be put out via Blackboard.

To start direct communication with me, e-mail me.

Grades

The total course grade we will come from your performance on Quizzes (25%), Homework (25%) and Tests (50%).

Your final grade will follow a traditional scheme. Cutoffs may be lowered depending on final outcomes of the course.

Grade	Percentage Range
A	90% - 100%
B	80 - 89%
C	70 - 79%
D	60% - 69%
F	0% - 59%

Quizzes (Knowledge Checks)

There will be a number of "knowledge checks" throughout the course available on blackboard. These are multiple choice quizzes checking that you have an adequate understanding of the course material. The quizzes have an unlimited number of attempts. However, the questions are randomly assigned and therefore "memorizing" the questions will not help improve your performance.

Homeworks

There will be 5 homework assignments.

- The amount they contribute to your final grade will vary with difficulty and number of problems.
- They are intended to be a simple assessment of the topics covered.
- If you find you are spending a lot of time on problems in a given assignment, it is probably something that can be resolved a lot easier than you think. Shoot me an e-mail and we can hash things out!

Exams

Exams will be in the form of a report on data you will be provided. You use the techniques learned to answer questions or complete objectives I specify.

- Exam 1 will cover Simple and Multiple Linear Regression models.
- Exam 2 will cover Analysis of Variance and Logistic Regression models.
  - Exam 2 is not a "final" in traditional terms. It is just your last exam.
  - It will not be cumulative, but previous knowledge is required to understand the topics.
- Each exam is equally weighted toward your final grade. I will do my best to ensure length and difficulty of each exam reflects this.

Cheating

- It is pretty easy to cheat in an online course, however it is way easier to catch cheating than you probably think.
- Cheating, in my experience, occurs because students don't understand the material, so the result is poor grade overall in the class mainly because of failed exams even if I do not penalize you via grade or higher echelon's of penalization.

- If you are struggling, I am here to help. No point in taking this class if you don't try to learn.

College of Public Health Policies

This course adheres to all COPH policies as outlined in the Student Handbook https://secure.uams.edu/cophstudent/student-handbook.aspx. Policies that have specific relevance to this course include:

- Attendance: https://secure.uams.edu/cophstudent/student-handbook.aspx#attendance~~
- Students with disability: https://secure.uams.edu/cophstudent/student-handbook.aspx#disabilities
- Academic Integrity: https://secure.uams.edu/cophstudent/student-handbook.aspx#honorcode
- Plagiarism: https://secure.uams.edu/cophstudent/student-handbook.aspx#plagarism

It is your responsibility as a student to familiarize yourself with these policies and to adhere to them. If you have any questions about any of these policies, please contact your instructor, the Associate Dean for Professional Programs, or the Associate Dean for Academic Affairs.

Course Competencies

MPH Foundational Competencies

Competency	Exam #1	Exam #2	HWs
F03. Analyze quantitative and qualitative data using biostatistics, informatics, computer-based programming and software, as appropriate.	x	x	All
F04. Interpret results of data analysis for public health research, policy or practice.	x	x	All

MPH Epidemiology Competencies

Competency	Exam #1	Exam #2	Homeworks
Use statistical software and other informatics tools in support of epidemiologic practice and research.	x	x	All

MPH Biostatistics Competencies

Competency	Exam #1	Exam #2	Homeworks
Evaluate statistical associations based on multivariate methods.	x	x	2, 4-6
Produce computer code for data management and statistical analyses.	x	x	All
Choose the appropriate assumptions of fundamental statistical tests.	x	x	2-6
Examine graphical displays of data that accompany statistical analysis.	x	x	2-6
Evaluate hypothesis tests for comparing two or more groups with respect to a covariate.		x	5

Academic Integrity

As a student enrolled in the COPH, you are subject to the Academic Integrity, Principles, Policy & Procedures. The Honor Pledge you signed at enrollment documents your knowledge that you may not engage in, or facilitate engagement in, the following forms of academic misconduct :

a. Cheating: Cheating is considered to be an attempt to use or provide unauthorized assistance, materials, information, or study aids in any form and in any academic exercise or environment.

   i) A student must not use external assistance on any in-class or take-home examination, unless the instructor specifically has authorized external assistance. This prohibition includes, but is not limited to, the use of tutors, books, notes, calculators, computers, and wireless communication devices.

ii) A student must not use another person as a substitute in the taking of an examination or quiz, nor allow other persons to conduct research or to prepare work, without advanced authorization from the instructor to whom the work is being submitted.
iii) A student must not use materials from a commercial term paper company, files of papers prepared by other persons, or submit documents found on the Internet.
iv) A student must not collaborate with other persons on a particular project and submit a copy of a written report that is represented explicitly or implicitly as the student's individual work.
v) A student must not use any unauthorized assistance in a laboratory, at a computer terminal, or on fieldwork.
vi) A student must not steal examinations or other course materials, including but not limited to, physical copies and photographic or electronic images.
vii) A student must not submit substantial portions of the same academic work for credit or honors more than once without permission of the instructor or program to whom the work is being submitted.
viii) A student must not, without authorization, alter a grade or score in any way, nor alter answers on a returned exam or assignment for credit.

b. Fabrication: A student must not falsify or invent any information or data in an academic exercise including, but not limited to, records or reports, laboratory results, and citation to the sources of information.

c. Plagiarism: Plagiarism is defined as presenting someone else's work, including the work of other students, as one's own. Any ideas or materials taken from another source for either written or oral use must be fully acknowledged, unless the information is common knowledge. What is considered "common knowledge" may differ from course to course. A student must not adopt or reproduce ideas, opinions, theories, formulas, graphics, or pictures of another person without acknowledgment. A student must give credit to the originality of others and acknowledge indebtedness whenever:

i) directly quoting another person's actual words, whether oral or written;
ii) using another person's ideas, opinions, or theories;
iii) paraphrasing the words, ideas, opinions, or theories of others, whether oral or written;
iv) borrowing facts, statistics, or illustrative material; or
v) offering materials assembled or collected by others in the form of projects or collections without acknowledgment

d. Interference: A student must not steal, change, destroy, or impede another student's work, nor should the student unjustly attempt, through a bribe, a promise of favors or threats, to affect any student's grade or the evaluation of academic performance. Impeding another student's work includes, but is not limited to, the theft, defacement, or mutilation of resources so as to deprive others of the information they contain.

e. Violation of Course Rules: A student must not violate course rules established by a department, the course syllabus, verbal or written instructions, or the course materials that are rationally related to the content of the course or to the enhancement of the learning process in the course.

f. Facilitating Academic Dishonesty: A student must not intentionally or knowingly help or attempt to help another student to commit an act of academic misconduct, nor allow another student to use his or her work or resources to commit an act of misconduct.

g. Failing to comply with the terms or conditions of a settlement or reconciliation agreement that has been approved by the Associate Dean for Academic Affairs(ADAA).

h. Acting in any other way that would undermine the academic integrity of the COPH.

These policies are based on Code of Student Rights, Responsibilities, & Conduct. Section 6. Indiana University. Available at https://studentcode.iu.edu/responsibilities/academic-misconduct.html. Used with permission (IU Chief Policy Officer Kip Drew 9/24/2020).