

STA 610: MULTILEVEL AND HIERARCHICAL MODELS

FALL 2021
DUKE UNIVERSITY

| | |
|------------------------|---|
| INSTRUCTOR: | OLANREWAJU MICHAEL AKANDE, PH.D. |
| EMAIL: | ✉ olanrewaju.akande@duke.edu |
| OFFICE: | 223 Gross Hall |
| OFFICE HOURS: | Mondays and Thursdays (9am – 10am). Zoom meeting ID: See Sakai. |
| COURSE PAGE: | https://sta610-f21.olanrewajuakande.com |
| MEETING TIMES: | Mondays and Wednesdays (1:45 – 3:00pm), Perkins Link 071 (Classroom 5). |
| LABS: | Fridays (1:45 – 3:00pm), Perkins Link 071 (Classroom 5). |
| TEACHING ASSISTANTS: | CHENGXIN YANG. Tuesdays (1pm – 2pm) and Thursday (6pm – 7pm). Zoom meeting ID: See Sakai. |
| | ALESSANDRO ZITO. Wednesdays (10am – 11am) and Fridays (8:30am – 9:30am). Zoom meeting ID: See Sakai. |
| RECOMMENDED TEXTBOOKS: | – Data Analysis Using Regression and Multilevel/Hierarchical Models by Gelman A., and Hill, J. – Peter Hoff’s Lecture Notes on Hierarchical Modeling (Sakai) |
| IMPORTANT DATES: | <ul style="list-style-type: none">• Monday, August 23 Fall classes begin• Friday, September 3 Drop/Add ends• Monday, September 6 Labor day. Classes in session• Wednesday, September 27 Exam I (<i>tentative</i>)• Sunday, October 17 Case Study I reports due (<i>tentative</i>)• Monday, November 15 Exam II (<i>tentative</i>)• Tuesday, November 23 End of semester• Thursday, December 9 Case Study II reports due (<i>tentative</i>) |

1 Course Overview

Statistical models are necessary for analyzing the type of multivariate (often large) datasets that are usually encountered in data science and statistical science, and hierarchical models often play a vital role in many of those applications. This is a graduate-level course that introduces students to the building blocks of hierarchical modeling and provides students with the tools needed to build, fit and interpret hierarchical models.

Hierarchical or multilevel models provide a principled way to model naturally grouped or clustered data, in a way that takes advantage of the relationship between observations in the same group, but

also allows for borrowing of information across the groups. In this course, you will be introduced to these models, with particular emphasis on the theoretical and conceptual foundations, as well as implementation, model fitting, and interpretation of the results.

This course emphasizes the mathematical theory behind hierarchical models, as well as real data analyses, including interpretation of results. All students must have the theoretical background covered in the prerequisites (in particular, in the context of Bayesian statistics) to be able to keep up with and understand the materials.

2 Learning Objectives

By the end of this course, students should be able to

- ▣ Understand the foundations and general structure of both classical and Bayesian hierarchical models.
- ▣ Specify and fit hierarchical models to various types of grouped or clustered data.
- ▣ Use the models covered in class to analyze real data sets.
- ▣ Assess the adequacy of hierarchical models to any given data and make a decision on what to do in cases when certain models are not appropriate for a given dataset.

3 Prerequisites

ALL students are expected to be familiar with all the topics covered within the required prerequisites to be in this course. That is, courses covering linear and matrix algebra and STA 360/601/602. Students are also expected to be very familiar with R and are encouraged to have learned \LaTeX or a Markdown language by the end of the course.

4 Team Work

This course also emphasizes the ability to work in teams so that students can learn team productivity and performance. Each student must be ready to contribute to their team's success. During the first or second week of class, we will develop teams of around four to five students - these teams will stay consistent throughout the semester (barring extraordinary circumstances). You will work in these teams during classes, labs and on the group case studies.

5 Class Materials

Lecture notes and slides, lab exercises and assigned readings will be posted on the course website, while lecture and lab videos will be posted on Sakai.

6 Workload

Work hours will include time spent going through the preassigned readings, attending the lectures, attending the lab sessions, and doing all graded work. Please note that the more focused and engaged you are, the quicker you will be able to get through all the materials.

7 Graded Work

Graded work for the course will consist of problem sets, lab exercises, two case studies and two exams. Regrade requests for problem sets, lab exercises and case studies must be done via Gradescope AT MOST **24 hours** after grades are released! Regrade requests for the exams must be done via Gradescope AT MOST **12 hours** after grades are released!

- ▣ Students' final grades will be determined as shown by the breakdown in Table 1.

Table 1: Breakdown of graded work

| Component | Percentage |
|---------------------|------------|
| Participation | 2.5% |
| Lab Exercises | 10 % |
| Group Case Study I | 12.5% |
| Group Case Study II | 15% |
| Problem Sets | 20% |
| Exam I | 20% |
| Exam II | 20% |

- ▣ There are no make-ups for any of the graded work except for medical or familial emergencies or for reasons approved by the instructor BEFORE the due date. Contact the instructor in advance of relevant due dates to discuss possible alternatives.
- ▣ Grades may be curved at the end of the semester. Cumulative averages of 90% – 100% are guaranteed at least an A-, 80% – 89% at least a B-, and 70% – 79% at least a C-, however the exact ranges for letter grades will be determined at the end of the course.

8 Descriptions of graded work

8.1 Problem sets

Four problem sets will be handed out over the course of the semester. These problem sets are to be completed independently without collaborating with other students. Be sure to check the website regularly for the dates and deadlines! **Please note that any work that is not legible by the instructor or TAs will not be graded (given a score of 0). Every write-up must be clearly written in full sentences and clear English. Any assignment that is completely unclear to the instructors and/or TAs, may result in a grade of a 0.** For programming exercises, you are required to use R and you must submit ALL of the code as an appendix.

Each student MUST write up and turn in their own answers. You are encouraged to talk to each other, regarding problem sets or to the instructor/TA. However, the write-up, solutions, and code must be entirely your own work. The assignments must be submitted on Gradescope under “Assignments”. Note that you will not be able to make online submissions after the due date, so be sure to submit before or by the Gradescope-specified deadline.

Solutions to the problem sets will be curated from student solutions with proper attribution. For each problem set, the TAs will select a representative correct solution for the assigned problems and put them together into one solution set with each answer being attributed to the student who wrote it. **If you would like to OPT OUT of having your solutions used for as a representative solution, please let the instructor and TAs know in advance.**

8.2 Lab exercises

The objective of the lab assignments is to give you more hands-on experience with hierarchical data analysis. Come learn a concept or two and some R from the TAs, and then work on the computational part of the problem sets. Each lab assignment should be submitted in timely fashion. You are **REQUIRED** to use R Markdown to type up your lab reports.

8.3 Group Case Studies

There will be detailed analysis of two case studies using current, relevant data. Each case study will involve a final submission per team. Each team will present their findings in class, and also turn in written reports. You must participate in each presentation to receive full credit for that group case study. Individual contributions to each submission will also be assessed. Team members must provide these assessments in order to receive credit for an assignment. An individual team member's grade may be modified if assessments indicate this is appropriate. Additional details on the case studies will be made available later.

8.4 Exams

There will be two exams as specified on the syllabus and schedule. If you know in advance you must miss an examination, please contact the instructor as soon as possible. Detailed instructions will be made available later.

9 Late Submission Policy

- You will lose
 1. 40% of the total points on each problem set or case studies if you submit within the first 24 hours after it is due, and
 2. 100% of the total points if you submit later than that.

In addition, you will lose

1. 50% of the total points on each lab exercise if you submit within the first 24 hours after it is due, and
 2. 100% of the total points if you submit later than that.
- You will lose 100% of the total points on exams if you miss the dates/times.

10 Auditing

Students who audit this course will be expected to complete most of the graded work with the goal of getting an overall score of at least 70%; you will only need to complete enough graded work to get to 70%. You are also expected to attend most lectures, go through the readings, and generally, participate like everyone else. You must contact the instructor in advance if you wish to audit the course.

11 Tentative Course Schedule

We will cover the topics below. We may spend different amounts of time on each topic. For a detailed and updated outline, check on the updated course schedule on the course page regularly.

▣▶ **Building blocks:**

- | One way ANOVA: scalar and matrix formulations
- | MLEs, contrasts, coding schemes, and interactions
- | Random effects ANOVA
- | Bayesian estimation
- | Random effects ANCOVA

▣ Linear mixed effect models:

- | Review of linear model
- | Linear mixed effect models
- | Diagnostics and influence measures
- | Bayesian linear mixed effect models
- | Longitudinal data
- | Crossed/non-nested random effects
- | Measurement error
- | Poststratification and weighting
- | Missing data
- | Meta analysis
- | Multilevel categorical outcomes
- | Introduction to finite mixture models

▣ Wrap up and review

12 Academic Integrity

Duke University is a community dedicated to scholarship, leadership, and service and to the principles of honesty, fairness, respect, and accountability. Citizens of this community commit to reflect upon and uphold these principles in all academic and nonacademic endeavors, and to protect and promote a culture of integrity. To uphold the [Duke Community Standard](#):

- 👉 I will not lie, cheat, or steal in my academic endeavors;
- 👉 I will conduct myself honorably in all my endeavors; and
- 👉 I will act if the Standard is compromised.

Cheating or plagiarism on any graded assessments, lying about an illness or absence and other forms of academic dishonesty are a breach of trust with classmates and faculty, violate the Duke Community Standard, and will not be tolerated. Such incidences will result in a 0 grade for all parties involved. Additionally, there may be penalties to your final class grade along with being reported to the Office of Student Conduct. Review the academic dishonesty policies at <https://studentaffairs.duke.edu/conduct/z-policies/academic-dishonesty>.

13 Diversity & Inclusiveness:

This course is designed so that students from all backgrounds and perspectives all feel welcome both in and out of class. Please feel free to talk to me (in person or via email) if you do not feel well-served by any aspect of this class, or if some aspect of class is not welcoming or accessible to you. My goal is for you to succeed in this course, therefore, let me know immediately if you feel you are struggling with any part of the course more than you know how to manage. Doing so will not affect your grades, but it will allow me to provide the resources to help you succeed in the course.

14 Disability Statement

Students with disabilities who believe that they may need accommodations in the class are encouraged to contact the [Student Disabilities Access Office](#) at 919.668.1267 or disabilities@aes.duke.edu as soon as possible to better ensure that such accommodations are implemented in a timely fashion.

15 Other Information

It can be a lot more pleasant oftentimes to get one-on-one answers and help. Make use of the teaching team's office hours, we're here to help! Do not hesitate to talk to me during office hours or by appointment to discuss a problem set or any aspect of the course. Questions related to course assignments and honesty policy should be directed to me. When the teaching team has announcements for you we will send an email to your Duke email address. Be sure to check your email daily.

If you have any concerns, issues or challenges, let the instructor know as soon as possible. Also, all students are strongly encouraged to rely on Ed Discussion, for interacting among yourself and asking other students questions. You can also ask the instructor or the TAs questions on there and we will try to respond as soon as possible. If you experience any technical issues with joining or using Ed Discussion, let the instructor know.

16 Professionalism

Try as much as possible to refrain from texting or using your computer for anything other than coursework while in class. Again, the more engaged you are, the quicker you will be able to get through the materials. You are responsible for everything covered in class, the lecture notes/slides, and in the assigned readings.