



DATA PM511 *correlated data* ANALYSIS

PM511c: Correlated Data Analysis

Units: 4

Fall 2023

Location: SSB 116

Lecture: 9-10:55am

Lab: 11-12:55pm

Instructor: **Sandy Eckel, PhD**

Associate Professor, Division of Biostatistics

Director, PhD Program in Biostatistics

Office: SSB-202B

Office Hours: By appointment. Email me to schedule.

Contact Info: eckel@usc.edu (expect replies ≤ 1 business day)

Teaching Assistants:

TBA, email TBA

IT Help: George Martinez

Hours of Service: Weekdays, 8am-4pm

Contact Info: georgem@usc.edu

Blackboard help: Call 213-740-5555 and choose option 2 to receive assistance 24 hours a day, 365 days a year.

Course Description

This course introduces regression methods for correlated data, including longitudinal and multilevel data. Correlated data violate the usual independence assumption in regression models. In typical multilevel data, the predictor and the outcome variables occur at multiple levels of aggregation (e.g., at the person, family, town, and/or regional levels). Multilevel models account for the correlation induced by the clusters at each level and can be used to quantify associations of the outcome with factors at each level. Longitudinal data arises from studies with repeated measurements on participants. Depending on research goals, the within-person correlation can be addressed in generalized estimating equation (GEE) models or mixed models. Growth curve models focus on modeling trends over time. The audience for this course includes second year and beyond Biostatistics graduate students as well as graduate students from other Divisions, Departments, or Schools interested in analyzing correlated data for their research.

Learning Objectives

Upon successfully completing this course, students will be able to:

- Interpret parameters of multilevel/longitudinal models
- Translate substantive questions into the form of a multilevel/longitudinal model
- Use the R statistical software package to: prepare graphical and tabular displays of multilevel/longitudinal data that effectively communicate the patterns of scientific interest, and to fit multilevel/longitudinal models
- Critique multilevel/longitudinal data analyses published in the biomedical literature
- Write methods and results sections summarizing multilevel/longitudinal data analyses, as if for a manuscript

Prerequisite(s): PM511b; open to other students by consent of instructor

Recommended Preparation: Coursework in concepts and applications of Generalized Linear Models.

Course Notes

This course is taught in a traditional on ground format with live in-person lecture, followed by live in-person lab. Should you be unable to attend the live in-person course session (e.g., due to COVID), please notify the instructional team (email Dr. Eckel and the TA). Please do not attend class if you are feeling unwell. We aim to make the course accessible by zoom as a backup (not primary) mode of attendance. Recordings from a prior version of this class are also available.

This 4 unit course is taught during the 15 week Fall semester, according to the following [USC rules](#):

- Minimum required weekly contact time (e.g., lectures/labs/discussion sections): 3 hour 20 min
- Weekly out-of-class time (e.g., homework, readings): 8 hours

Technological Proficiency and Hardware/Software Required

You need to have access to a computer to access course materials, complete the lab exercises, HW, midterm, and final project for this course. The statistical software R will be used extensively in this course.

Required Readings and Supplementary Materials

There is no required textbook.

Recommended textbooks:

Multilevel and longitudinal modeling using Stata. Rabe-Hesketh, S. and Skrondal, A., 2008. STATA press. A fantastic reference and a good read for applied data analyses, but uses Stata. I've adopted many examples/datasets from this book. I use the 2nd edition, but a 4th edition (2 books) is available.

Analysis of Longitudinal Data, 2nd Ed. Diggle P, Heagerty P, Liang K-Y, Zeger S. Oxford University Press, 2002.
Has been the required book for this course in the past. A great reference for longitudinal data.

Data analysis using regression and multilevel/hierarchical models. Gelman A, Hill J. 2007
[free USC library link] https://uosc.primo.exlibrisgroup.com/permalink/01USC_INST/mbk0s6/alma991043504543403731
Great applied book. Has a Bayesian/Bugs/R focus and fewer biomedical data examples.

Mixed-effects models in S and S-PLUS. Pinheiro JC, Bates DM. 2000.
[free USC library link] https://uosc.primo.exlibrisgroup.com/permalink/01USC_INST/mbk0s6/alma991043335600403731
Great R code and theory. Note: although title says S/Plus, the code presented generally works in R.

Other relevant books available as free pdf downloads from the USC library:

Mixed Effects Models and Extensions in Ecology with R. Zuur, Alain; Ieno, Elena N; Walker, Neil; Saveliev, Anatoly A; Smith, Graham M. New York, NY : Springer New York : Imprint: Springer; 2009
https://uosc.primo.exlibrisgroup.com/permalink/01USC_INST/mbk0s6/alma991043230758903731
(Great readable text, with practical applications and some theory. Applications are in ecology – not biomedical.)

Linear Mixed-Effects Models Using R A Step-by-Step Approach. Gałdecki, Andrzej; Burzykowski, Tomasz. New York, NY : Springer New York : Imprint: Springer; 2013
https://uosc.primo.exlibrisgroup.com/permalink/01USC_INST/mbk0s6/alma991042454011203731
(R for LMM, focuses on nlme; some lme4, provides theory as well. Not as readable.)

Analysis of Longitudinal Data with Examples. Wang YG, Fu L, Paul S. Chapman and Hall/CRC; 2022.
https://uosc.primo.exlibrisgroup.com/permalink/01USC_INST/mbk0s6/alma991043538628203731
(Theory-forward, uses R, includes chapters on: model selection, robust methods, missing data, high dimensional longitudinal data)

Modeling Binary Correlated Responses using SAS, SPSS and R. Wilson, Jeffrey R; Lorenz, Kent A. Cham : Springer International Publishing : Imprint: Springer; 2015
https://uosc.primo.exlibrisgroup.com/permalink/01USC_INST/mbk0s6/alma991042485319503731

Correlated Data Analysis: Modeling, Analytics, and Applications. Song, Peter X. -K. New York, NY : Springer New York : Imprint: Springer; 2007
https://uosc.primo.exlibrisgroup.com/permalink/01USC_INST/mbk0s6/alma991042454400503731
(SAS focus, chapter on missing data in longitudinal studies, chapter on Bayesian approach)

Linear and Generalized Linear Mixed Models and Their Applications. Jiang, Jiming. New York, NY : Springer New York : Springer; 2007
https://uosc.primo.exlibrisgroup.com/permalink/01USC_INST/mbk0s6/alma991042454405903731

Extending the Linear Model with R: Generalized Linear, Mixed Effects and Nonparametric Regression Models. Faraway, Julian James. Chapman & Hall/CRC Imprint; 2006
https://uosc.primo.exlibrisgroup.com/permalink/01USC_INST/mbk0s6/alma991043219577103731

Applied Bayesian hierarchical methods. Congdon, P. Boca Raton : Chapman & Hall/CRC; 2010
https://uosc.primo.exlibrisgroup.com/permalink/01USC_INST/mbk0s6/alma991043219516803731

Growth curve modeling : theory and applications. Panik, Michael J. Hoboken, New Jersey : John Wiley & Sons, Inc.; 2014
https://uosc.primo.exlibrisgroup.com/permalink/01USC_INST/mbk0s6/alma991042476419603731
Uses SAS

Applied mixed models in medicine. Brown, Helen, 1962.; Prescott, Robin, Chichester, England : Wiley; 2015
https://uosc.primo.exlibrisgroup.com/permalink/01USC_INST/mbk0s6/alma991042479984303731

Linear mixed models : a practical guide using statistical software. West, Brady T.; Welch, Kathleen B.; Gatecki, Andrzej T; Gillespie, Brenda W., Boca Raton ; London ; New York : Chapman & Hall/CRC, Taylor & Francis Group; 2007
https://uosc.primo.exlibrisgroup.com/permalink/01USC_INST/mbk0s6/alma991043219549603731
Examples using R, SAS, SPSS, HLM, Stata

Additional resources for learning R

There are many. Here are a few initial places to look.

- **Quick-R**: brief how-to's for common analyses <https://www.statmethods.net/>
- **R-bloggers**: More in-depth treatment of various topics via blog posts <https://www.r-bloggers.com/>
- **R search**: Dedicated R google search <https://rseek.org/>

Description and Assessment of Assignments

Homework: There will be 5 homework assignments. Homework must be submitted in electronic form on blackboard by 9am on the due date. The lowest HW score will be dropped.

Midterm exam: The midterm will be held from Thursday 9-11am on Week 7. The format will likely be an "open book/notes" in-class multiple choice/short answer and you will need to bring your laptop. There will be no lab on the day of the midterm.

Final project:

The final consists of a **group** project (data analysis, report writeup) and a small **individual** in-class component which will take place during the scheduled final exam period.

Grading Breakdown

Assessment Tool	% of Grade
Homework	50%
Midterm Exam	20%
Final project	30%
TOTAL	100%

Assignment Submission Policy

Assignments are to be submitted electronically through blackboard. No late assignments or final projects will be accepted. For the homework, students may discuss strategies with one another, but must turn in individual write ups. For the midterm, all work must be your own (no discussions with anyone else). The final project is group work, so you must work together within your group but you may not discuss with other groups.

Grading Timeline

We aim to have assignments graded within 1 week.

Additional Policies

Note on academic integrity.

Any violations of the academic integrity policies will result in zero credit for all students involved, submission of the incident to the Academic Integrity Coordinator for the Keck School of Medicine for adjudication and, likely, an F in the course for all students involved.

Policy regarding intellectual content related to the course. Recordings, slides or any other material that you are given access to via Blackboard, email or shared drives are strictly for your learning and should not be shared with anyone who is not enrolled in the course. The University's SCampus policy regarding class notes (<https://policy.usc.edu/scampus-part-c/>) applies to these materials and prohibits misuse, inappropriate dissemination, attempted sale or appropriation of intellectual property. Violation of these policies will be met with appropriate disciplinary actions.

Course Schedule: A Weekly Breakdown

Week	Lecture Topic	Lab Topic	Deliverables
Week 1 8/24/23	Introduction to correlated data Review of Generalized Linear Models	Introduction to R/RStudio Linear & Logistic Regression in R	HW1 out
Week 2 8/31/23	Two-level linear models (Part I): Random intercept/variance components models	Linear random intercept models in R	HW1 due HW2 out
Week 3 9/07/23	Two-level linear models (Part II): Random intercept models with covariates Random coefficient models	Two-level linear mixed models in R	
Week 4 9/14/23	Centering and contextual effects Three-level linear models	Centering and contextual effects & Three-level linear models in R	HW2 due HW3 out
Week 5 9/21/23	Exploratory Data Analysis (EDA) for trends in longitudinal data	EDA for longitudinal data mean trends in R	
Week 6 9/28/23	Linear mixed effects models for longitudinal data Midterm Review	Growth curve modeling and LMM model checking in R	HW3 due
Week 7 10/05/23	MIDTERM exam (no module activities this week)	No lab	MIDTERM
10/12/23	<i>Fall Recess. No class</i>		
Week 8 10/19/23	EDA for correlation in longitudinal data Generalized estimating equations (GEE)	EDA for longitudinal data correlation and GEE in R	HW4 out
Week 9 10/26/23	Binary outcomes	GEE for binary outcomes in R	
Week 10 11/02/23	Ordinal outcomes	Two and three-level logistic regression models in R	HW4 due HW5 out
Week 11 11/09/23	Count outcomes	Correlated count outcomes in R Weekly COVID mortality counts	
Week 12 11/16/23	Distributed lag models (Dr. Jason Niu) Additional pearls of wisdom	TBD	HW5 due Final project out
11/23/23	<i>Thanksgiving Holiday. No class</i>		
Week 13 11/30/23	Review	Final project group work	
FINAL 12/07/23 (Thurs) 11am-1pm	Final project - in class component Refer to the USC final exam schedule: classes.usc.edu .		FINAL PROJECT DUE

Statement on Academic Conduct and Support Systems

Academic Integrity:

The University of Southern California is a learning community committed to developing successful scholars and researchers dedicated to the pursuit of knowledge and the dissemination of ideas. Academic misconduct, which includes any act of dishonesty in the production or submission of academic work, comprises the integrity of the person who commits the act and can impugn the perceived integrity of the entire university community. It stands in opposition to the university's mission to research, educate, and contribute productively to our community and the world.

All students are expected to submit assignments that represent their own original work, and that have been prepared specifically for the course or section for which they have been submitted. You may not submit work written by others or "recycle" work prepared for other courses without obtaining written permission from the instructor(s).

Other violations of academic integrity include, but are not limited to, cheating, plagiarism, fabrication (e.g., falsifying data), collusion, knowingly assisting others in acts of academic dishonesty, and any act that gains or is intended to gain an unfair academic advantage.

The impact of academic dishonesty is far-reaching and is considered a serious offense against the university. All incidences of academic misconduct will be reported to the Office of Academic Integrity and could result in outcomes such as failure on the assignment, failure in the course, suspension, or even expulsion from the university.

For more information about academic integrity see [the student handbook](#) or the [Office of Academic Integrity's website](#), and university policies on [Research and Scholarship Misconduct](#).

Please ask your instructor if you are unsure what constitutes unauthorized assistance on an exam or assignment, or what information requires citation and/or attribution.

Students and Disability Accommodations:

USC welcomes students with disabilities into all of the University's educational programs. The Office of Student Accessibility Services (OSAS) is responsible for the determination of appropriate accommodations for students who encounter disability-related barriers. Once a student has completed the OSAS process (registration, initial appointment, and submitted documentation) and accommodations are determined to be reasonable and appropriate, a Letter of Accommodation (LOA) will be available to generate for each course. The LOA must be given to each course instructor by the student and followed up with a discussion. This should be done as early in the semester as possible as accommodations are not retroactive. More information can be found at osas.usc.edu. You may contact OSAS at (213) 740-0776 or via email at osasfrontdesk@usc.edu.

Support Systems:

[Counseling and Mental Health](#) - (213) 740-9355 – 24/7 on call

Free and confidential mental health treatment for students, including short-term psychotherapy, group counseling, stress fitness workshops, and crisis intervention.

[988 Suicide and Crisis Lifeline](#) - 988 for both calls and text messages – 24/7 on call

The 988 Suicide and Crisis Lifeline (formerly known as the National Suicide Prevention Lifeline) provides free and confidential emotional support to people in suicidal crisis or emotional distress 24 hours a day, 7 days a week, across the United States. The Lifeline is comprised of a national network of over 200 local crisis centers, combining custom local care and resources with national standards and best practices. The new, shorter phone number makes it easier for people to remember and access mental health crisis services (though the previous 1 (800) 273-8255 number will continue to function indefinitely) and represents a continued commitment to those in crisis.

[Relationship and Sexual Violence Prevention Services \(RSVP\)](#) - (213) 740-9355(WELL) – 24/7 on call

Free and confidential therapy services, workshops, and training for situations related to gender- and power-based harm (including sexual assault, intimate partner violence, and stalking).

[Office for Equity, Equal Opportunity, and Title IX \(EEO-TIX\)](#) - (213) 740-5086

Information about how to get help or help someone affected by harassment or discrimination, rights of protected classes, reporting options, and additional resources for students, faculty, staff, visitors, and applicants.

[Reporting Incidents of Bias or Harassment](#) - (213) 740-5086 or (213) 821-8298

Avenue to report incidents of bias, hate crimes, and microaggressions to the Office for Equity, Equal Opportunity, and Title for appropriate investigation, supportive measures, and response.

[The Office of Student Accessibility Services \(OSAS\)](#) - (213) 740-0776

OSAS ensures equal access for students with disabilities through providing academic accommodations and auxiliary aids in accordance with federal laws and university policy.

[USC Campus Support and Intervention](#) - (213) 740-0411

Assists students and families in resolving complex personal, financial, and academic issues adversely affecting their success as a student.

[Diversity, Equity and Inclusion](#) - (213) 740-2101

Information on events, programs and training, the Provost's Diversity and Inclusion Council, Diversity Liaisons for each academic school, chronology, participation, and various resources for students.

[USC Emergency](#) - UPC: (213) 740-4321, HSC: (323) 442-1000 – 24/7 on call

Emergency assistance and avenue to report a crime. Latest updates regarding safety, including ways in which instruction will be continued if an officially declared emergency makes travel to campus infeasible.

[USC Department of Public Safety](#) - UPC: (213) 740-6000, HSC: (323) 442-1200 – 24/7 on call

Non-emergency assistance or information.

[Office of the Ombuds](#) - (213) 821-9556 (UPC) / (323-442-0382 (HSC)

A safe and confidential place to share your USC-related issues with a University Ombuds who will work with you to explore options or paths to manage your concern.

[Occupational Therapy Faculty Practice](#) - (323) 442-2850 or otfp@med.usc.edu

Confidential Lifestyle Redesign services for USC students to support health promoting habits and routines that enhance quality of life and academic performance.