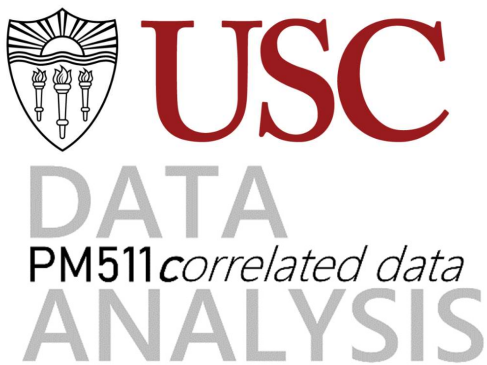




PM511c: Correlated Data Analysis

Units: 4

Fall 2020

Location: Online, see USC Blackboard site

Instructor: **Sandy Eckel, PhD**

Associate Professor, Division of Biostatistics
Director, PhD Program in Biostatistics

Office: Online. Post-COVID, you'll find me back at SSB-202B.

Office Hours: By appointment. Email me to schedule a zoom call.

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

Teaching Assistants:

Lab TA: **Lai Jiang**, jian848@usc.edu

Grading TA: **Jingying Weng**, wengjing@usc.edu

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.

**** Contents of this syllabus may be modified****

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 entirely online. Course materials, including pre-recorded lectures, will be posted on blackboard by noon Mondays. Modules will be completed each week, which will include listening to pre-recorded lectures and participating in live zoom sessions (which will be recorded). The live zoom session will be Thursdays 9-11am.

This 4 unit course will be taught online for the abbreviated **13 week Fall 2020** semester.

Weekly contact time (e.g., lectures/labs/discussion sections; pre-recorded or live zoom): 3 hour 55 min

Weekly out-of-class time (e.g., homework, readings): 9 hour 15 min

Technological Proficiency and Hardware/Software Required

You need to have access to a computer to complete the lab exercises, HW, 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 3rd 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
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.
Great R code and theory. Note: although title says S/Plus, the code presented generally works in R.

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. Gałęcki, 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.)

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 multiple choice/short answer.

Final project:

The final project includes a take-home data analyses and an 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). Policies for the final project will be announced towards the end of the semester.

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.

Guidelines for being an online student. Strive to be engaged, kind, and helpful. During live zoom sessions, please turn on your video camera if you are able. Class dynamics are compromised if participants are not visible by camera. For additional privacy, please feel free to use virtual backgrounds and headsets/earphones. Mute yourself when not actively speaking. Please reach out to the instructor/TAs with any concerns. To ask a question, please use the “raise hand” feature in the participants window.

Policy regarding intellectual content related to the course. Lectures will be recorded and made available to students. The 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

	Topics	Deliverables
Week 1 8/20/20	Introduction to correlated data Introduction to R/RStudio Review of Generalized Linear Models	HW1 out
Week 2 8/27/20	Two-level linear models (Part I): Random intercept/variance components models	HW1 due HW2 out
Week 3 9/3/20	Two-level linear models (Part II): Random intercept models with covariates Random coefficient models	
Week 4 9/10/20	Centering and contextual effects Three-level linear models	HW2 due HW3 out
Week 5 9/17/20	Exploratory Data Analysis for trends in longitudinal data Linear mixed effects models for longitudinal data	
Week 6 9/24/20	Exploratory Data Analysis for correlation in longitudinal data	HW3 due
Week 7 10/1/20	MIDTERM exam (no module activities this week)	MIDTERM
Week 8 10/8/20	Generalized estimating equations: continuous outcomes	HW4 out
Week 9 10/15/20	Binary outcomes	
Week 10 10/22/20	Ordinal outcomes	HW4 due HW5 out
Week 11 10/29/20	Count outcomes	
Week 12 11/5/20	Miscellaneous topics as time permits: Sample size, Missing data, Correlated time-to-event data	HW5 due Final project out
Week 13 11/12/20	Review	
FINAL 11/19/20	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 Conduct:

Plagiarism – presenting someone else’s ideas as your own, either verbatim or recast in your own words – is a serious academic offense with serious consequences. Please familiarize yourself with the discussion of plagiarism in SCampus in Part B, Section 11, “Behavior Violating University Standards” policy.usc.edu/scampus-part-b. Other forms of academic dishonesty are equally unacceptable. See additional information in SCampus and university policies on scientific misconduct, policy.usc.edu/scientific-misconduct.

Support Systems:

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

studenthealth.usc.edu/counseling

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

National Suicide Prevention Lifeline - 1 (800) 273-8255 – 24/7 on call

suicidepreventionlifeline.org

Free and confidential emotional support to people in suicidal crisis or emotional distress 24 hours a day, 7 days a week.

Relationship and Sexual Violence Prevention Services (RSVP) - (213) 740-9355(WELL), press “0” after hours – 24/7 on call

studenthealth.usc.edu/sexual-assault

Free and confidential therapy services, workshops, and training for situations related to gender-based harm.

Office of Equity and Diversity (OED) - (213) 740-5086 | Title IX – (213) 821-8298

equity.usc.edu, titleix.usc.edu

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

usc-advocate.symplicity.com/care_report

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

The Office of Disability Services and Programs - (213) 740-0776

dsp.usc.edu Support and accommodations for students with disabilities. Services include assistance in providing readers/notetakers/interpreters, special accommodations for test taking needs, assistance with architectural barriers, assistive technology, and support for individual needs.

USC Campus Support and Intervention - (213) 821-4710

campussupport.usc.edu

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

Diversity at USC - (213) 740-2101

diversity.usc.edu

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

dps.usc.edu, emergency.usc.edu

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) 442-120 – 24/7 on call

dps.usc.edu Non-emergency assistance or information.