Longitudinal and Dependent Data Analysis – BST513

Spring 2020
SRB 1.404, 9:05-10:55pm TTh

Instructor: Brent A. Johnson, Ph.D.
Office Hours: 10:55-11:55am T and by appt.
Course website, course email list: TBA
Prerequisites: BST 401, 411, 426

Instructional Staff

There is no TA for this course. My office is located in SRB 4.142 and office hours will be held there.

Course Description

The course topic is statistical methods for the analysis of longitudinal data, clustered data, and correlated data. This course is a methods course that will touch on both theoretical and practical aspects of various techniques.

Course Aims and Objectives

We will introduce correlated data through simple exercises and then through some data examples. The first section of the course includes exploratory data analysis, design considerations, and a review of generalized linear models for independent data. The second section of the course is dedicated to mixed effects models and generalized estimating equations. The third and final section of the course covers advanced techniques that either extend or improve the methods of the first two sections as time permits. These topics may include quadratic inference functions, data analyses with time-dependent confounders, variable selection, and the joint analysis of survival and longitudinal data.

By the end of the course, the student should be able to (i) articulate the advantages of longitudinal data as over cross-sectional data, and (ii) recognize correlated data and articulate the pitfalls from ignoring correlation when it exists. The student should be comfortable designing a study with multiple follow-up time points. The student should be comfortable analyzing longitudinal data with mixed effects and marginal models, using existing software and (to a limited extent) writing one's own code from scratch, and write a coherent report that summarizes the results. They should be able to outline the asymptotic arguments for the estimators they implement. Finally, the student should be able to articulate the assumptions of ubiquitous mixed effects and marginal models, recognize the limitations that these assumptions impose in practice, and have the ability to browse the literature for possible solutions to overcome such limitations.

Course Policies and Expectations

Homework: You are encouraged to work in teams with your classroom peers but the write-up of the solution set must be <u>in your own words</u>. Homework assignments will generally be due 14 days after the

day assigned (Due dates will be announced in class). Homework will be due at the beginning of class. Late assignments will be permitted in the case of emergency or with approval of the professor prior to the beginning of class on the due date. Otherwise, late assignments may or may not be accepted and will be completely at the discretion of the professor.

Classroom: Please use common sense and etiquette in the classroom.

Cell phones: Cell phones and pagers should be turned off or set to vibrate mode for emergencies. The classroom is a place of learning and random noises present a distraction to others. If you must take a call, excuse yourself from the classroom. Please refrain from texting during the lecture unless it is an emergency.

Laptops: The use of laptops is discouraged.

Attendance/Participation: There will be no formal attendance sheet but attendance and classroom participation figures into the final grade (See Grading Procedures). Student engagement keeps the learning active and stimulates the classroom environment.

Enrollment: Standard university procedures and dates will govern the successful dropping/adding of courses, changing from a letter grade to pass/fail, and auditing the course.

Finally, if an issue presents itself during class, we will try to address it straight away. If a student wishes to raise an issue privately, they can arrange an appointment with the professor.

Materials and Access
Required Text
1. Diggle, P.J., Heagerty, P.J., Liang, K.-Y., and Zeger, S. L. (2002) *Analysis of Longitudinal Data*, 2nd ed., Oxford University Press: Oxford, UK.
Recommended Text
2. Pinheiro, J.C. and Bates, D.M. (2000) *Mixed-effects models in S and S-Plus*. Spring-Verlag: New York.
3. Veberke, G. and Mohlenberghs, G. (2000) *Linear Mixed Models for Longitudinal Data*. Springer series in statistics. Springer: New York.

Assignments and Grading Procedures
There will be approximately 4-5 homework assignments. The course grade will be based on homework (70%), a final project (20%), attendance and class participation (10%).

Academic Integrity
Academic integrity is a core value of the University of Rochester. Students who violate the University of Rochester University Policy on Academic Honesty are subject to disciplinary penalties, including the possibility of failure in the course and/or dismissal from the University. Since academic dishonesty harms the individual, other students, and the integrity of the University, policies on academic dishonesty are strictly enforced. For further information on the University of Rochester Policy on Academic Honesty, please see the *Jurisdiction and Responsibility for Academic and Nonacademic Misconduct* section in the **Regulations and University Polices Concerning Graduate Studies**
http://www.rochester.edu/GradBulletin/PDFbulletin/Regulations.pdf

Accommodations for Students with Disabilities

Students needing academic adjustments or accommodations because of a documented disability must contact the Access Services Coordinator. For information regarding access services and support at SMD, please refer to our webpage: https://www.urmc.rochester.edu/education/graduate/current-students/disability-supports-services.aspx

Policy Against Discrimination and Harassment
http://www.rochester.edu/president/assets/pdf/attachment-student-policy-against-harrassment-and-discrimination-4.26.16.pdf
Center for Student Conflict Management (CSCM): 585-275-4085 or
Conflict.management@rochester.edu
Anonymous Disclosure: see Bias-related Incident Report Form

Course Schedule

1. Introduction to Longitudinal Data Analysis
2. Exploring Longitudinal Data
3. Considerations in Experimental Design
4. Review of Statistical Methods and Models for Independent Data

5. Random Effects Models
6. Marginal Models, Generalized Estimating Equations

Advanced topics (in no particular order):
6. Quadratic inference functions
7. Data analysis with time-dependent confounding
8. Random observation times, possibly dependent observation times
9. Joint analysis of longitudinal and survival data
10. Variable selection, high-dimensional longitudinal data analysis

Important Dates:
16 January 2020: first day of class.
9-13 March 2020: spring break.
28 April 2018: last day of class.

Other potentially useful information for students

- UCLA statistical consulting website*: http://www.ats.ucla.edu/stat/*
- course website*: none at this time*