Course Title: Applied Longitudinal Data Analysis

Course Number: BIST 0650

Course Location: School of Public Health Building, Room 234

Course Date & Time: Tuesday 10:10 am – 1:00 pm

Course Instructor: Yong Lin, Ph.D., Professor, Biostatistics Department, School of Public Health, Rutgers University, Room 214 in School of Public Health Building, Phone: (732)-235-5513/9621, email: linyo@rutgers.edu

Office Hours: Tuesday 1:00 pm – 1:45 pm or by appointment

Course Assistant: Menghui Chen, Ph.D. Student, Biostatistics Department, School of Public Health, Rutgers University, Phone: 954-882-9152, email: menghui.chen@rutgers.edu


Additional/Supplemental Readings/Resources:


Course Description: Longitudinal data consists of multiple measures over time on a sample of individuals. The analysis of longitudinal data requires much more sophisticated methodologies due to the correlation introduced by repeated measurements. This course covers modern statistical techniques for longitudinal data from an applied perspective. Emphasis will be on data analysis and interpretation. Topics include characteristics of the longitudinal design, graphical exploration of the mean and correlation structure, linear mixed effects models and multilevel modeling, maximum likelihood and restricted maximum likelihood estimation, modeling the variance-covariance structures, inference for random effects, logistic and Poisson mixed effects model for binary and count data, marginal models and generalized estimating equations, and model diagnostics. Analysis of real and substantial data sets using statistical software SAS and R will be integrated throughout.
Selected Department Competencies Addressed: Each Department identifies competencies for each degree offered. The competencies addressed in this course for the Ph.D. for the Department of Biostatistics include:

- Conduct appropriate statistical analysis of data to solve medical and public health problems;
- Reinforce use statistical computer packages to organize, analyze, and report collected data;
- Design experimental and observational studies in biomedical, clinical and public health research;
- Conduct complex statistical analyses for a broad range of applications;
- Communicate the results of statistical studies both in writing and orally to investigators and lay community members.

Please visit the Department webpages on the School of Public Health’s website at http://sp.h rutgers.edu for additional competencies addressed by this course for other degrees and departments.

Course Objectives: By the completion of this course, students will:

a) Identify the special features of longitudinal designs, describe how these features might relate to the analysis. Manipulate the data in a way suitable for longitudinal analysis.

b) Graphically explore and present the longitudinal data.

c) Use SAS Proc mixed and R lme4 package to analyze continuous longitudinal data. Correctly specify fixed and random effects and covariance structure. Interpret the SAS and/or R output.

d) Use SAS Proc glimmix and nlmixed and R lme4 package to perform logistic and Poisson mixed effects modeling for repeated binary or count data.

e) Use SAS genmod and R gee function to analyze repeated binary or count data using generalized estimating equation techniques.

f) Predict the impact of missing data on standard statistical inference and for a particular situation, be able to choose between the common approaches for handling missing values.

g) Plan and design a longitudinal study. Use an appropriate method to analyze a particular study and interpret the result.

Course Requirements and Grading:

- Lectures will be given each week.
- Six to seven homework will be assigned. Students are allowed to work in groups on homework if they like, but no one should copy directly from someone else’s paper.
- Both in-class midterm exam and final examination will be open book.
- The R software can be downloaded free of charge from http://www.r-project.org. The SAS software can be leased through the university’s software service (see class website for more details).
The course grade will be based on homework assignments, the midterm exam, the final exam, a project, and class participation. The relative weight given to each of these components is:

- Attendance, Participation and Class Activities: 10%
- Homework: 40%
- Midterm Exam: 25%
- Final Exam: 25%

Total: 100%

Course Schedule: (Tentative)

<table>
<thead>
<tr>
<th>Date</th>
<th>Class</th>
<th>Topic(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>09/02</td>
<td>1</td>
<td>Definition, features, and objectives of longitudinal studies</td>
</tr>
<tr>
<td>09/09</td>
<td>2</td>
<td>Graphical exploration and presentation of longitudinal data</td>
</tr>
<tr>
<td>09/16</td>
<td>3</td>
<td>Linear Models: –Introduction, Mean models, Estimation SAS Proc mixed</td>
</tr>
<tr>
<td>09/23</td>
<td>4</td>
<td>Linear Models: –Fitting correlation model in R and SAS; R function lm, SAS Proc mixed</td>
</tr>
<tr>
<td>09/30</td>
<td>5</td>
<td>Linear Models: –Inferences in Mean (I): F-test, CI and LRT; Inferences for the Covariance/Correlation Model: ML estimates and CIs for Covariance Parameters, ReML for Covariance Parameters SAS Proc mixed, R function lme</td>
</tr>
<tr>
<td>10/14</td>
<td>7</td>
<td>Linear Mixed Models (II): –Model Building and Model Diagnostics; SAS Proc mixed, glimmix, R function lme</td>
</tr>
<tr>
<td>10/21</td>
<td>8</td>
<td>Midterm exam</td>
</tr>
<tr>
<td>10/28</td>
<td>9</td>
<td>Generalized Linear Model: – Marginal model and generalized estimating equation, SAS Proc genmod, R function gee</td>
</tr>
<tr>
<td>11/04</td>
<td>10</td>
<td>Generalized Linear Model: – Logistic for binary data and Log-Linear model for count data SAS Proc glimmix, nlmixed, R function nlme</td>
</tr>
<tr>
<td>11/18</td>
<td>12</td>
<td>Generalized Linear Mixed Effects Models (subject-specific) SAS Proc glimmix, genmod, nlmixed, R functions nlme, gee</td>
</tr>
<tr>
<td>11/25</td>
<td></td>
<td>Thanksgiving Recess</td>
</tr>
<tr>
<td>12/02</td>
<td>13</td>
<td>Missing data analysis in longitudinal study SAS Proc mi, mianalyze</td>
</tr>
<tr>
<td>12/09</td>
<td>14</td>
<td>Final Review</td>
</tr>
<tr>
<td>12/16</td>
<td>15</td>
<td>Final Exam</td>
</tr>
</tbody>
</table>
School of Public Health Honor Code: The School of Public Health Honor Code is found in the student bulletin (sph.rutgers.edu/academics/catalog/index.html). Each student bears a fundamental responsibility for maintaining academic integrity and intellectual honesty in his or her graduate work. For example, all students are expected to observe the generally accepted principles of scholarly work, to submit their own rather than another’s work, to refrain from falsifying data, and to refrain from receiving and/or giving aid on examinations or other assigned work requiring independent effort. In submitting written material, the writer takes full responsibility for the work as a whole and implies that, except as properly noted by use of quotation marks, footnotes, etc., both the ideas and the works used are his or her own. In addition to maintaining personal academic integrity, each student is expected to contribute to the academic integrity of the school community by not facilitating inappropriate use of her/his own work by others and by reporting acts of academic dishonesty by others to an appropriate school authority. It should be clearly understood that plagiarism, cheating, or other forms of academic dishonesty will not be tolerated and can lead to sanctions up to and including separation from the Rutgers School of Public Health.

Policy Concerning Use of Recording Devices and Other Electronic Communications Systems:
When personally owned communication/recording devices are used by students to record lectures and/or classroom lessons, such use must be authorized by the faculty member or instructor who must give either oral or written permission prior to the start of the semester and identify restrictions, if any, on the use of mobile communications or recording devices.