BIOS 767: Longitudinal Data Analysis
Spring 2017

Instructor: Bahjat Qaqish
Office: 3105-B McGavran-Greenberg
e-mail: qaqish@bios.unc.edu
office hours: Monday, 11am–1pm

Graders:
TBA
e-mail: @unc.edu

Texts:

• Primary Resources
  – Fitzmaurice, Laird, and Ware, Applied Longitudinal Analysis, 2nd edition
  – Relevant papers as assigned

• Supplemental Resources
  – Diggle, Heagerty, Liang, and Zeger, Analysis of Longitudinal Data
  – Background material: Christensen’s book (from bios 762) and the GLM book by McCullagh & Nelder.

Software:

• SAS and R

Location and Time:
Class: 9:05-10:50 Monday and Wednesday, 1304 McGavran-Greenberg Hall

Prerequisite is BIOS 762. Knowledge of an appropriate software package and basic matrix algebra is assumed.

Learning Objectives:

- Gain strong understanding of theory (ideas) of longitudinal data analysis
- Hone skills in fitting and interpreting longitudinal data models for addressing scientific questions that arise in public health and medicine
- Gain competency in study design and power analysis of longitudinal data to complete skills needed to write statistical section of a grant proposal to collect and analyze longitudinal data

Copies of presentations used in class will be available online on Sakai. These notes cover some material not contained in the texts and do not cover all material in the texts, so that the assigned readings are very important. You are encouraged to print the course notes and bring them with you to class. Course credits: I have drawn heavily on the work of professors and colleagues when creating the course, and in particular the work of professors Liang and Zeger at Johns Hopkins, Fitzmaurice, Laird, Ware, and Williams at Harvard; Edwards, Helms, Muller, Stewart and Herring at UNC.
Grades Homework assignments and quizzes: 40%. Midterm Exam: February (tbd) 25%; Final Exam: Monday, May 8, 2017, 35%.

The Graduate School uses the grades H (clear excellence), P (entirely satisfactory), L (low pass), and F (failure). Graduate students are expected to earn “P” grades, with remarkable performances rewarded with other grades from the scale as appropriate. Class participation may modify the association between the numeric average and assigned letter grade.

Honor Code:
Students in BIOS767 are expected to abide by the UNC Honor Code. All suspected Honor Code violations will be reported to the UNC Dean of Students, who will investigate the case. These investigations typically involve lengthy hearings of the Honor Court, and as outlined in the Instrument of Student Judicial Governance, “The usual sanction for a first academic violation is definite suspension for at least one academic semester and a grade penalty of an ‘F’ for the course, a portion of the course, or the assignment.”

Course Content: Note: This course content is (to some extent) tentative. Topics may not be covered in the same order as given. Some topics may be dropped and others added depending on various factors.

1. The BIG picture first
2. What is a model?
3. General Linear Model for Longitudinal Data
4. ANOVA - univariate, multivariate
   - ML and REML estimation
   - EM algorithm
   - General linear mixed-effects model (Laird-Ware)
   - Inference for the random effects: BLUPs, Empirical Bayes, Bayes, Shrinkage
   - Model construction
   - Regression diagnostics
   - Hierarchical (random coefficient) models
   - Multivariate longitudinal data
   - Relaxing parametric assumptions: generalized additive mixed model
   - Design of randomized and observational longitudinal studies
5. Generalized linear models for longitudinal data
• Marginal models for binary, ordinal, and count data
• Random effects models for binary, ordinal, and count data
• Transition models
• Likelihood-based models for categorical data
• GEE
• Models for mixed discrete and continuous responses

6. Dropouts and missing data

• Classification of missing data mechanisms
• Intermittent missing values and dropouts
• Weighted estimating equations
• Modeling the dropout process (selection and pattern mixture models)

7. Time-dependent covariates

• Dangers of time-dependent covariates
• Lagged covariates

8. Joint models for data of different types

9. Special topics