

Longitudinal Data Analysis

The goal of this course is to provide an overview of statistical models and methods that are useful in the analysis of longitudinal data. Note that a defining characteristic of above examples is that the same response is measured repeatedly on each unit. This course will cover modern approaches for analyzing longitudinal data with emphasis on linear mixed effects models for Gaussian response, generalized linear models for non-Gaussian response (such as response variables following binomial or Poisson distributions) and nonlinear mixed effects models for nonlinear profiles. Traditional methods such as repeated measures ANOVA/MANOVA will be discussed in reference to linear mixed effects models. We will discuss graphical data exploration, association structures, parameter estimation / testing /inference, modeling and handling of missing data, advanced tools such as semiparametric inferences and Bayesian methods, model selection, diagnostics and model limitations. Computation will be done using R, SAS, JAGS and WinBugs.

Book: Applied Longitudinal Analysis, 2nd edition (Wiley Series in Probability and Statistics) by Fitzmaurice, Laird and Ware. **ISBN: 978-0-38027-7.**

Website of the book: www.biostat.harvard.edu/~fitzmaur/ala/

Additional reference book: **Analysis of Longitudinal Data** by Diggle, Liang and Zeger.

Course Outline

- Preliminaries: Introduction, Review of matrix algebra, Random vectors, multivariate normal distribution, review of linear regression, Introduction to modeling longitudinal data, exploring covariance structure
- Classical methods for normally distributed, balanced repeated measurements: Univariate repeated measures analysis of variance. Multivariate repeated measures analysis of variance, Drawbacks and limitations of classical methods
- Methods for normally distributed, unbalanced repeated measurements: General linear models and models for correlation, Random effects models, Linear mixed effects models
- Methods for non-normally distributed, unbalanced data: Probability models for discrete and continuous non-Gaussian response and generalized linear models, Generalized estimating equations for population-averaged (marginal) models
- Advanced topics: Generalized linear mixed effects models, Nonlinear mixed effects models, partial linear models, Bayesian models and methods, Semiparametric inference
- Missing data mechanisms and analysis of missing data