ST 790

Longitudinal Data Analysis

Lecture Notes

M. Davidian

Department of Statistics

North Carolina State University

©2018 by Marie Davidian

Contents

1 Introduction and Motivation 1

1.1 Objective of this course . . . 1

1.2 Examples . . . 2

1.3 Statistical models for longitudinal data . . . 16

1.4 Outline of the course . . . 19

2 Modeling Longitudinal Data 21

2.1 Introduction . . . 21

2.2 Data structure and notation . . . 21

2.3 Conceptual framework for continuous response . . . 26

2.4 Population-averaged versus subject-specific modeling . . . 37

2.5 Models for correlation structure . . . 44

2.6 Exploring mean and correlation structure . . . 50

2.7 Considerations for discrete response . . . 59

3 Repeated Measures Analysis of Variance 64

3.1 Introduction . . . 64

3.2 Univariate repeated measures analysis of variance . . . 66

3.3 Specialized within-individual hypotheses and tests . . . 83

3.4 Multivariate repeated measures analysis of variance . . . 94

4 Modern Methods: Preliminaries 102

4.1 Introduction . . . 102

4.2 Drawbacks of classical methods . . . 102

4.3 Large sample theory and estimating equations . . . 105

5 Population-Averaged Linear Models for Continuous Response 115

5.1 Introduction . . . 115

5.2 Model specification . . . 116

5.3 Maximum likelihood estimation under normality . . . 131

5.4 Restricted maximum likelihood . . . 135

5.5 Large sample inference . . . 142

5.6 Missing data 157

6 Linear Mixed Effects Models 169

6.1 Introduction 169

6.2 Model specification 170

6.3 Inference and considerations for missing data 186

6.4 Best linear unbiased prediction and empirical Bayes 190

6.5 Implementation via the EM algorithm 198

6.6 Testing variance components 201

7 Generalized and Nonlinear Models for Univariate Response 207

7.1 Introduction 207

7.2 Nonlinear mean-variance models 208

7.3 Estimation of mean and variance parameters 215

7.4 Large sample results 221

8 Population-Averaged Models and Generalized Estimating Equations 229

8.1 Introduction 229

8.2 Model specification 230

8.3 Linear estimating equations 237

8.4 Quadratic estimating equations 248

8.5 Large sample inference 251

8.6 Modeling issues 255

8.7 Missing data 263

8.8 Examples 272

8.9 Further results for quadratic equations 276

9 Nonlinear and Generalized Linear Mixed Effects Models 280

9.1 Introduction 280

9.2 Model specification 281

9.3 Maximum likelihood 303

9.4 Approximate inference based on individual estimates 307

9.5 Approximate inference based on linearization 315

9.6 “Exact” likelihood inference 325

9.7 Examples 328

10 Additional Topics 333

10.1 Introduction 333

10.2 Bayesian formulation of hierarchical models 333

10.3 Complex nonlinear models 340

10.4 Time-dependent covariates in nonlinear mixed effects models 345

10.5 Multilevel models 349

10.6 Distribution of random effects 357

Appendix A: Fun Matrix Facts 365

Appendix B: Notation and Taylor Series 368

Appendix B: Review of Large Sample Theory 372

Appendix D: Brief Review of Monte Carlo Simulation 384

Appendix E: PROC MIXED Syntax 387

Appendix F: Writing a Data Analysis Report 390

References 396