

ST 732
Applied Longitudinal Data Analysis

Lecture Notes

M. Davidian
Department of Statistics
North Carolina State University

©2005 by Marie Davidian

Contents

1	Introduction and Motivation	1
1.1	Purpose of this course	1
1.2	Examples	2
1.3	Statistical models for longitudinal data	12
1.4	Outline of the course	18
2	Review of matrix algebra	20
2.1	Introduction	20
2.2	Matrix notation	20
2.3	Matrix operations	22
3	Random vectors and multivariate normal distribution	32
3.1	Preliminaries	32
3.2	Random vectors	39
3.3	The multivariate normal distribution	46
3.4	Multiple linear regression	57
4	Introduction to modeling longitudinal data	68
4.1	Basic Statistical Model	68
4.2	Sources of variation in longitudinal data	73
4.3	Exploring mean and covariance structure	76
4.4	Popular models for covariance structure	86
4.5	Diagnostic calculations under stationarity	92
4.6	Implementation with SAS	96
5	Univariate repeated measures analysis of variance	105
5.1	Introduction	105
5.2	Basic situation and statistical model	106
5.3	Questions of interest and statistical hypotheses	121
5.4	Analysis of variance	129
5.5	Violation of covariance matrix assumption	132

5.6	Specialized within-unit hypotheses and tests	134
5.7	Adjusted tests	144
5.8	Implementation with SAS	145
6	Multivariate repeated measures analysis of variance	172
6.1	Introduction	172
6.2	General multivariate problem	175
6.3	Hotelling's T^2	176
6.4	One-way MANOVA	179
6.5	Profile Analysis	184
6.6	Implementation with SAS	188
7	Drawbacks and limitations of classical methods	201
7.1	Introduction	201
7.2	Assumptions and restrictions of classical methods	201
8	General linear models for longitudinal data	208
8.1	Introduction	208
8.2	Simplest case – one group, balanced data	210
8.3	General case – several groups, unbalanced data, covariates	212
8.4	Models for covariance	222
8.5	Inference by maximum likelihood	228
8.6	Restricted maximum likelihood	240
8.7	Discussion	243
8.8	Implementation with SAS	247
8.9	Parameterizing models in SAS: Use of the <code>noint</code> option in SAS model statements in PROC GLM and PROC MIXED	296
8.10	Using SAS <code>model</code> , <code>contrast</code> , and <code>estimate</code> statements	305
9	Random coefficient models for multivariate normal data	308
9.1	Introduction	308
9.2	Random coefficient model	309
9.3	Inference on regression and covariance parameters	323

9.4	Inference on individuals	328
9.5	Discussion	330
9.6	Basic PROC MIXED syntax	332
9.7	Implementation with SAS	335
10	Linear mixed effects models for multivariate normal data	362
10.1	Introduction	362
10.2	Examples	362
10.3	General linear mixed effects model	373
10.4	Inference on regression and covariance parameters	375
10.5	Best linear unbiased prediction	376
10.6	Testing whether a component is random	385
10.7	Time-dependent covariates	387
10.8	Discussion	391
10.9	Implementation with SAS	391
11	Generalized linear models for nonnormal response	422
11.1	Introduction	422
11.2	Probability models for nonnormal data	424
11.3	Generalized linear models	435
11.4	Maximum likelihood and iteratively reweighted least squares	442
11.5	Discussion	446
11.6	Implementation with SAS	446
12	Population-averaged models for nonnormal repeated measurements	463
12.1	Introduction	463
12.2	Population-averaged model	467
12.3	Generalized estimating equations	482
12.4	“Robust” estimator for sampling covariance	486
12.5	Contrasting population-averaged and subject-specific approaches	487
12.6	Discussion	491
12.7	Implementation with SAS	491

13 Advanced topics	517
13.1 Introduction	517
13.2 Generalized linear mixed models	517
13.3 Nonlinear mixed effects models	520
13.4 Issues associated with missing data	522
14 References	527