

Contents

Preface	xv
1 Introduction to Longitudinal Data	1
1.1 What Is Longitudinal Data?	2
1.2 Related Data Types	3
1.3 Inferences from Longitudinal Data	4
1.3.1 The Population Mean	5
1.3.2 Individual Variability	5
1.3.3 Covariance and Correlation	6
1.3.4 Covariates	7
1.3.5 Predictions	8
1.4 Contrasting Longitudinal and Cross-Sectional Data . . .	8
1.5 Benefits of Longitudinal Data Analysis	10
1.5.1 Efficiency and Cost	10
1.5.2 Prediction	10
1.5.3 Time Trends	10
1.6 Time Frames for Observing Longitudinal Data	10
1.7 Complexities in Analyzing Longitudinal Data	12
1.7.1 Correlation	13
1.7.2 Size	13
1.7.3 Profiles Over Time	14
1.7.4 Complex Data Structures	14
1.7.5 Missing Data	15
1.7.6 Non-constant Variance	16

1.7.7	Covariates and Regression	16
1.8	The Language of Longitudinal Data	16
1.8.1	Responses	16
1.8.2	It's About Time	17
1.8.3	Covariates	22
1.8.4	Data Formats	23
1.9	Problems	24
2	Plots	27
2.1	Graphics and Longitudinal Data	28
2.2	Responses Over Time	30
2.2.1	Scatterplots	30
2.2.2	Box Plots	30
2.2.3	Flaws in Scatterplots and Box Plots for Longitudinal Data	32
2.2.4	Profile Plots	33
2.2.5	The Need to Connect-the-Dots in Profiles	34
2.3	Interpreting Profile Plots	37
2.3.1	Sample Means and Standard Deviations	39
2.3.2	Skewness and the Pediatric Pain Data	41
2.3.3	Within-Subject Variability	44
2.4	Elaborations of Profile Plots	46
2.4.1	Covariates	46
2.4.2	Ozone Data	47
2.4.3	Weight Loss Data and Viewing Slopes	50
2.4.4	Empirical Within-Subject Residuals	55
2.4.5	Empirical Population Residuals	56
2.4.6	Too Many Subjects	57
2.5	Inspecting Correlations	57
2.5.1	Scatterplot Matrices	61
2.5.2	Correlation in Profile Plots	63
2.5.3	The Correlogram	65
2.6	Empirical Summary Plots	65
2.7	How Much Data?	69
2.7.1	Cognitive Data: Raven's	70
2.8	Discussion	73
2.9	Problems	73
3	Simple Analyses	85
3.1	Summarizing Longitudinal Data	87
3.1.1	Mean	87
3.1.2	Slope	88
3.2	Paired t -test	91
3.3	Difference of Differences	92
3.3.1	Pediatric Pain Example	94

3.4	Using a Subset of the Data	95
3.4.1	Pediatric Pain Data	96
3.5	Problems	96
4	Critiques of Simple Analyses	99
4.1	Efficiency Loss	100
4.1.1	Omitting Subjects	101
4.1.2	Omitting Observations	101
4.1.3	Omitting Subjects and Observations: Some Special Cases	103
4.2	Bias	104
4.2.1	Bias By Design	104
4.2.2	Bias From Bad Analysis	104
4.3	The Effect of Missing Data on Simple Summaries	105
4.3.1	Summarizing Profiles by the Mean	106
4.3.2	Slopes	107
4.3.3	Weighted Least Squares Analysis	107
4.4	Difference of Differences	107
4.5	The Richness of Longitudinal Data	109
4.6	Discussion	113
4.7	Problems	114
5	The Multivariate Normal Linear Model	117
5.1	Multivariate Normal Model for Balanced Data	118
5.1.1	Estimating μ and Σ	119
5.1.2	Elaborations to the Multivariate Normal Model	122
5.2	Parameterized Covariance Models	123
5.2.1	Compound Symmetry	124
5.2.2	Autoregressive	124
5.2.3	Random Intercept and Slope	124
5.3	Regression Models for Longitudinal Data	125
5.3.1	Covariate Vectors and Covariate Matrices	126
5.3.2	Multivariate Normal Linear Regression Model	127
5.3.3	Parameter Estimates	128
5.3.4	Linear Combinations of Parameters	132
5.3.5	Inference	134
5.3.6	Degrees of Freedom in t -Tests	134
5.4	Graphical Presentation of Inferences	136
5.4.1	Weight Loss Data	137
5.5	Problems	140
6	Tools and Concepts	143
6.1	Likelihood Ratio Tests	144
6.1.1	Issues with Likelihood Ratio Tests for Covariance Parameters	145

6.1.2	Comparing Nested Random Effects Models	145
6.1.3	Fixed Effects	146
6.2	Model Selection	146
6.2.1	Model Selection for the Covariance Model	147
6.2.2	Comparison of AIC and BIC	149
6.2.3	The Probability of a Covariance Model	150
6.2.4	Model Selection for Fixed Effects	151
6.3	Maximum Likelihood and Restricted Maximum Likelihood	154
6.3.1	Residual Maximum Likelihood	155
6.4	Assuming Normality	156
6.5	Computational Issues	157
6.5.1	Maximizing a Likelihood	158
6.5.2	A Function Maximization Analogy	159
6.5.3	Gradient and Hessian	160
6.5.4	Collinearity	161
6.5.5	Dealing with Lack of Convergence	162
6.5.6	Discussion	163
6.6	Back-Transforming a Transformed Response	163
6.6.1	Estimates and Predictions After Transformation .	165
6.6.2	The Log/Exponential Transformation	165
6.6.3	General Back-Transformation	166
6.7	Some Design Considerations	168
6.7.1	Estimating the Population Mean	168
6.7.2	Cost and Efficiency in Longitudinal Data	170
6.7.3	Two Independent Samples versus Paired Comparisons	172
6.8	Problems	173
7	Specifying Covariates	175
7.1	Time-fixed Covariates	176
7.1.1	Categorical Covariates	177
7.1.2	Parameterization	178
7.1.3	Continuous Covariates	182
7.2	Population Means Varying as a Function of Time	184
7.2.1	Polynomials in Time	184
7.2.2	Unstructured Mean	187
7.2.3	Time-varying Group Membership	192
7.2.4	Group Interactions: Pediatric Pain	201
7.3	Groups and Time	206
7.3.1	Raven's Data: Gender and Time Interaction . . .	208
7.3.2	The Four Models Involving One Grouping Variable and Time	211
7.3.3	Three Groups	211
7.3.4	Raven's and Treatment and Time	212

7.4	Defining Time	214
7.4.1	Changing Units	214
7.4.2	Baseline Age and Time in Study	215
7.5	Three-Way Interactions	216
7.5.1	Raven's Example	217
7.6	Step Functions, Bent Lines, and Splines	219
7.6.1	Step Function Parameterizations	221
7.6.2	Bent Line Parameterization	222
7.6.3	Indicator Functions and Positive Parts	223
7.6.4	Knots	223
7.6.5	Higher Order Polynomial Splines	224
7.6.6	Placing Knots	224
7.6.7	BSI Data	225
7.6.8	Big Mice Data	226
7.6.9	Wallaby Tails	227
7.6.10	Other Basis Functions	228
7.7	Adjusting for Baseline; Generalizations of Paired Comparisons	228
7.8	Modeling Strategies	231
7.9	Problems	232
8	Modeling the Covariance Matrix	243
8.1	Parameterized Covariance Models	245
8.1.1	Compound Symmetry	246
8.1.2	Random Intercept Model	249
8.1.3	Autoregressive	250
8.1.4	Independence	254
8.1.5	Random Intercept and Slope	254
8.1.6	Independent Increments	261
8.1.7	Antedependence	264
8.1.8	Autoregressive Moving Average	265
8.1.9	Factor Analytic	269
8.1.10	Toeplitz or Banded Covariance Model	272
8.1.11	Unstructured Covariance Model	273
8.2	Non-constant Variance	273
8.2.1	Non-constant Variance Over Time	273
8.2.2	Constant Variance with Unstructured Correlations	274
8.2.3	Non-constant Variance Component	274
8.3	Examples	275
8.3.1	Small Mice Data	275
8.3.2	Pain Data	278
8.4	Non-constant Variance and Covariates	280
8.4.1	Pain Data Example	282
8.5	Sums of Covariance Matrices	285

8.5.1	Autoregressive Plus Random Intercept Plus Independent Error	287
8.5.2	Local Effect	287
8.6	Parameterized Covariance Matrices and Random Times	288
8.6.1	Models Easily Extensible to Random Times	288
8.6.2	Extensions to Random Times Requiring Additional Specifications	289
8.6.3	Spatial Correlation Models	290
8.7	Family Relationships Among Covariance Models	290
8.8	Reasons to Model the Covariance Matrix	292
8.8.1	The Science	292
8.8.2	Estimating Fixed Effects	292
8.8.3	Prediction of Individual Future Observations	293
8.8.4	Prediction of Missing Observations	293
8.8.5	Variability	293
8.9	Historical and Future Notes	293
8.10	Problems	294
9	Random Effects Models	303
9.1	Random Effects Models as Hierarchical Models	304
9.1.1	Some Examples	308
9.2	Three-Level Hierarchical Data Sets	310
9.2.1	Parent Random Effects in the BSI	313
9.2.2	Weight Loss Data	314
9.3	Marginal Model	316
9.4	Estimation of Random Effects and Shrinkage	317
9.5	A Simple Example of Shrinkage	320
9.6	Problems	324
10	Residuals and Case Diagnostics	327
10.1	Residuals and Outlier Statistics	328
10.2	Influence Statistics	329
10.3	Additional Residuals for Multivariate Regression	330
10.4	Residuals and Residual Plots in Random Effects Models	331
10.4.1	Small Mice Data	333
10.4.2	Pediatric Pain Data	336
10.4.3	Weight Loss Data	338
10.5	Problems	341
11	Discrete Longitudinal Data	343
11.1	The Logistic Random Effects Model	344
11.1.1	Morbidity Data	346
11.1.2	Other Models	350
11.2	Count Data	351

11.2.1	Bolus Count Data	353
11.3	Issues with Discrete Data Models	356
11.4	Problems	357
12	Missing Data	361
12.1	The Effects of Selective Dropout	362
12.2	Dropout versus In-Study Non-response	364
12.3	Missing at Random and Variants	365
12.4	Example: the Difference Between MAR and MNAR . . .	366
12.5	Modeling Missingness	367
12.6	Using Missingness as a Covariate	369
12.6.1	Schizophrenia Severity Data	369
12.7	Missing Data in the Covariates	371
12.8	Problems	372
13	Analyzing Two Longitudinal Variables	375
13.1	Thinking About Bivariate Longitudinal Data	376
13.2	Continuous Time-varying Covariates	377
13.3	Problems with Using a Response as a Predictor	379
13.4	The Bivariate Random Intercept Model	381
13.4.1	Pain Rating and Log Pain Tolerance	382
13.4.2	Cognitive Data	384
13.5	Bivariate Random Intercept and Slope	385
13.5.1	Cognitive Data	385
13.6	Non-simultaneously-Measured Observations	386
13.6.1	Anthropometry and Nutrition	387
13.7	Unstructured Covariance Models for Bivariate Responses	387
13.7.1	Unstructured Within Response, Independent Responses	389
13.7.2	Product Correlation Models	390
13.8	Discussion	392
13.9	Problems	393
14	Further Reading	395
	Bibliography	401
	Appendix: Data Sets	411
A.1	Basic Symptoms Inventory (BSI)	411
A.2	Big Mice and Small Mice	413
A.3	Bolus Counts	413
A.4	Dental	414
A.5	Kenya School Lunch Intervention	414
A.5.1	Anthropometry	414

A.5.2	Cognitive	414
A.5.3	Covariates	415
A.5.4	Morbidity	415
A.5.5	Cognitive	416
A.6	Ozone	417
A.7	Pediatric Pain	417
A.8	Schizophrenia	418
A.9	Small Mice	419
A.10	Vagal Tone	419
A.11	Wallaby	419
A.12	Weight Loss	420
Index		421



<http://www.springer.com/978-0-387-40271-0>

Modeling Longitudinal Data

Weiss, R.E.

2005, XXII, 432 p., Hardcover

ISBN: 978-0-387-40271-0