

Contents

Preface		xv
1 Review of Multiple Linear Regression		1
1.1	Learning Objectives	1
1.2	Introduction to Beyond Multiple Linear Regression	1
1.3	Assumptions for Linear Least Squares Regression	3
1.3.1	Cases Without Assumption Violations	4
1.3.2	Cases With Assumption Violations	6
1.4	Review of Multiple Linear Regression	8
1.4.1	Case Study: Kentucky Derby	8
1.5	Initial Exploratory Analyses	8
1.5.1	Data Organization	8
1.5.2	Univariate Summaries	9
1.5.3	Bivariate Summaries	9
1.6	Multiple Linear Regression Modeling	12
1.6.1	Simple Linear Regression with a Continuous Predictor	12
1.6.2	Linear Regression with a Binary Predictor	17
1.6.3	Multiple Linear Regression with Two Predictors	18
1.6.4	Inference in Multiple Linear Regression: Normal Theory	19
1.6.5	Inference in Multiple Linear Regression: Bootstrapping	20
1.6.6	Multiple Linear Regression with an Interaction Term	22
1.6.7	Building a Multiple Linear Regression Model	24
1.7	Preview of Remaining Chapters	26
1.7.1	Soccer	26
1.7.2	Elephant Mating	27
1.7.3	Parenting and Gang Activity	28
1.7.4	Crime	28
1.8	Exercises	29
1.8.1	Conceptual Exercises	29
1.8.2	Guided Exercises	32
1.8.3	Open-Ended Exercises	36

2	Beyond Least Squares: Using Likelihoods	39
2.1	Learning Objectives	39
2.2	Case Study: Does Sex Run in Families?	40
2.2.1	Research Questions	41
2.3	Model 0: Sex Unconditional, Equal Probabilities	42
2.4	Model 1: Sex Unconditional, Unequal Probabilities	43
2.4.1	What Is a Likelihood?	43
2.4.2	Finding MLEs	46
2.4.3	Summary	49
2.4.4	Is a Likelihood a Probability Function? (optional)	50
2.5	Model 2: Sex Conditional	50
2.5.1	Model Specification	50
2.5.2	Application to Hypothetical Data	51
2.6	Case Study: Analysis of the NLSY Data	53
2.6.1	Model Building Plan	53
2.6.2	Exploratory Data Analysis	54
2.6.3	Likelihood for the Sex Unconditional Model	55
2.6.4	Likelihood for the Sex Conditional Model	56
2.6.5	Model Comparisons	58
2.7	Model 3: Stopping Rule Model (waiting for a boy)	61
2.7.1	Non-nested Models	63
2.8	Summary of Model Building	64
2.9	Likelihood-Based Methods	65
2.10	Likelihoods and This Course	66
2.11	Exercises	67
2.11.1	Conceptual Exercises	67
2.11.2	Guided Exercises	67
2.11.3	Open-Ended Exercises	68
3	Distribution Theory	71
3.1	Learning Objectives	71
3.2	Introduction	71
3.3	Discrete Random Variables	72
3.3.1	Binary Random Variable	72
3.3.2	Binomial Random Variable	73
3.3.3	Geometric Random Variable	74
3.3.4	Negative Binomial Random Variable	75
3.3.5	Hypergeometric Random Variable	77
3.3.6	Poisson Random Variable	79
3.4	Continuous Random Variables	80
3.4.1	Exponential Random Variable	80
3.4.2	Gamma Random Variable	81
3.4.3	Normal (Gaussian) Random Variable	83
3.4.4	Beta Random Variable	84

3.5	Distributions Used in Testing	85
3.5.1	χ^2 Distribution	86
3.5.2	Student's t -Distribution	87
3.5.3	F -Distribution	87
3.6	Additional Resources	88
3.7	Exercises	88
3.7.1	Conceptual Exercises	88
3.7.2	Guided Exercises	90
4	Poisson Regression	93
4.1	Learning Objectives	93
4.2	Introduction to Poisson Regression	94
4.2.1	Poisson Regression Assumptions	95
4.2.2	A Graphical Look at Poisson Regression	95
4.3	Case Studies Overview	96
4.4	Case Study: Household Size in the Philippines	97
4.4.1	Data Organization	98
4.4.2	Exploratory Data Analyses	98
4.4.3	Estimation and Inference	102
4.4.4	Using Deviances to Compare Models	104
4.4.5	Using Likelihoods to Fit Models (optional)	106
4.4.6	Second Order Model	107
4.4.7	Adding a Covariate	109
4.4.8	Residuals for Poisson Models (optional)	110
4.4.9	Goodness-of-Fit	112
4.5	Linear Least Squares vs. Poisson Regression	113
4.6	Case Study: Campus Crime	114
4.6.1	Data Organization	114
4.6.2	Exploratory Data Analysis	115
4.6.3	Accounting for Enrollment	117
4.7	Modeling Assumptions	118
4.8	Initial Models	118
4.8.1	Tukey's Honestly Significant Differences	119
4.9	Overdispersion	121
4.9.1	Dispersion Parameter Adjustment	121
4.9.2	No Dispersion vs. Overdispersion	123
4.9.3	Negative Binomial Modeling	123
4.10	Case Study: Weekend Drinking	125
4.10.1	Research Question	125
4.10.2	Data Organization	126
4.10.3	Exploratory Data Analysis	126
4.10.4	Modeling	127
4.10.5	Fitting a ZIP Model	129
4.10.6	The Vuong Test (optional)	131

4.10.7	Residual Plot	132
4.10.8	Limitations	132
4.11	Exercises	133
4.11.1	Conceptual Exercises	133
4.11.2	Guided Exercises	136
4.11.3	Open-Ended Exercises	142
5	Generalized Linear Models: A Unifying Theory	145
5.1	Learning Objectives	145
5.2	One-Parameter Exponential Families	145
5.2.1	One-Parameter Exponential Family: Poisson	146
5.2.2	One-Parameter Exponential Family: Normal	147
5.3	Generalized Linear Modeling	148
5.4	Exercises	149
6	Logistic Regression	151
6.1	Learning Objectives	151
6.2	Introduction to Logistic Regression	151
6.2.1	Logistic Regression Assumptions	152
6.2.2	A Graphical Look at Logistic Regression	153
6.3	Case Studies Overview	153
6.4	Case Study: Soccer Goalkeepers	154
6.4.1	Modeling Odds	155
6.4.2	Logistic Regression Models for Binomial Responses	155
6.4.3	Theoretical Rationale (optional)	158
6.5	Case Study: Reconstructing Alabama	159
6.5.1	Data Organization	159
6.5.2	Exploratory Analyses	160
6.5.3	Initial Models	161
6.5.4	Tests for Significance of Model Coefficients	162
6.5.5	Confidence Intervals for Model Coefficients	163
6.5.6	Testing for Goodness-of-Fit	164
6.5.7	Residuals for Binomial Regression	166
6.5.8	Overdispersion	167
6.5.9	Summary	170
6.6	Linear Least Squares vs. Binomial Regression	170
6.7	Case Study: Trying to Lose Weight	171
6.7.1	Data Organization	172
6.7.2	Exploratory Data Analysis	173
6.7.3	Initial Models	176
6.7.4	Drop-in-Deviance Tests	179
6.7.5	Model Discussion and Summary	180
6.8	Exercises	181
6.8.1	Conceptual Exercises	181

6.8.2	Guided Exercises	182
6.8.3	Open-Ended Exercises	189
7	Correlated Data	193
7.1	Learning Objectives	193
7.2	Introduction	193
7.3	Recognizing Correlation	194
7.4	Case Study: Dams and Pups	195
7.5	Sources of Variability	195
7.6	Scenario 1: No Covariates	196
7.7	Scenario 2: Dose Effect	199
7.8	Case Study: Tree Growth	203
7.8.1	Format of the Data Set	204
7.8.2	Sources of Variability	205
7.8.3	Analysis Preview: Accounting for Correlation	206
7.9	Summary	207
7.10	Exercises	207
7.10.1	Conceptual Exercises	207
7.10.2	Guided Exercises	209
7.10.3	Note on Correlated Binary Outcomes	210
8	Introduction to Multilevel Models	211
8.1	Learning Objectives	211
8.2	Case Study: Music Performance Anxiety	212
8.3	Initial Exploratory Analyses	213
8.3.1	Data Organization	213
8.3.2	Exploratory Analyses: Univariate Summaries	214
8.3.3	Exploratory Analyses: Bivariate Summaries	216
8.4	Two-Level Modeling: Preliminary Considerations	220
8.4.1	Ignoring the Two-Level Structure (not recommended)	220
8.4.2	A Two-Stage Modeling Approach (better but imperfect)	221
8.5	Two-Level Modeling: A Unified Approach	225
8.5.1	Our Framework	225
8.5.2	Random vs. Fixed Effects	227
8.5.3	Distribution of Errors: Multivariate Normal	227
8.5.4	Technical Issues when Testing Parameters (optional)	229
8.5.5	An Initial Model with Parameter Interpretations	231
8.6	Building a Multilevel Model	234
8.6.1	Model Building Strategy	234
8.6.2	An Initial Model: Random Intercepts	234
8.7	Binary Covariates at Level One and Level Two	236
8.7.1	Random Slopes and Intercepts Model	236
8.7.2	Pseudo R-squared Values	239

8.7.3	Adding a Covariate at Level Two	240
8.8	Adding Further Covariates	242
8.8.1	Interpretation of Parameter Estimates	243
8.8.2	Model Comparisons	245
8.9	Centering Covariates	246
8.10	A Final Model for Music Performance Anxiety	248
8.11	Modeling Multilevel Structure: Is It Necessary?	251
8.12	Notes on Using R (optional)	254
8.13	Exercises	255
8.13.1	Conceptual Exercises	255
8.13.2	Guided Exercises	256
8.13.3	Open-Ended Exercises	258
9	Two-Level Longitudinal Data	263
9.1	Learning Objectives	263
9.2	Case Study: Charter Schools	264
9.3	Initial Exploratory Analyses	265
9.3.1	Data Organization	265
9.3.2	Missing Data	266
9.3.3	Exploratory Analyses for General Multilevel Models	268
9.3.4	Exploratory Analyses for Longitudinal Data	269
9.4	Preliminary Two-Stage Modeling	273
9.4.1	Linear Trends Within Schools	273
9.4.2	Effects of Level Two Covariates on Linear Time Trends	274
9.4.3	Error Structure Within Schools	279
9.5	Initial Models	279
9.5.1	Unconditional Means Model	280
9.5.2	Unconditional Growth Model	281
9.5.3	Modeling Other Trends over Time	284
9.6	Building to a Final Model	286
9.6.1	Uncontrolled Effects of School Type	286
9.6.2	Add Percent Free and Reduced Lunch as a Covariate	289
9.6.3	A Final Model with Three Level Two Covariates	291
9.6.4	Parametric Bootstrap Testing	294
9.7	Covariance Structure among Observations	301
9.7.1	Standard Covariance Structure	302
9.7.2	Alternative Covariance Structures	305
9.7.3	Non-longitudinal Multilevel Models	306
9.7.4	Final Thoughts Regarding Covariance Structures	306
9.7.5	Details of Covariance Structures (optional)	307
9.8	Notes on Using R (optional)	308
9.9	Exercises	309
9.9.1	Conceptual Exercises	309
9.9.2	Guided Exercises	312

9.9.3	Open-Ended Exercises	316
10	Multilevel Data With More Than Two Levels	321
10.1	Learning Objectives	321
10.2	Case Studies: Seed Germination	322
10.3	Initial Exploratory Analyses	323
10.3.1	Data Organization	323
10.3.2	Exploratory Analyses	325
10.4	Initial Models	332
10.4.1	Unconditional Means	333
10.4.2	Unconditional Growth	335
10.5	Encountering Boundary Constraints	337
10.6	Parametric Bootstrap Testing	343
10.7	Exploding Variance Components	349
10.8	Building to a Final Model	352
10.9	Covariance Structure (optional)	358
10.9.1	Details of Covariance Structures	361
10.10	Notes on Using R (optional)	363
10.11	Exercises	364
10.11.1	Conceptual Exercises	364
10.11.2	Guided Exercises	368
10.11.3	Open-Ended Exercises	370
11	Multilevel Generalized Linear Models	373
11.1	Learning Objectives	373
11.2	Case Study: College Basketball Referees	374
11.3	Initial Exploratory Analyses	374
11.3.1	Data Organization	374
11.3.2	Exploratory Analyses	376
11.4	Two-Level Modeling with a Generalized Response	380
11.4.1	A GLM Approach	380
11.4.2	A Two-Stage Modeling Approach	381
11.4.3	A Unified Multilevel Approach	384
11.5	Crossed Random Effects	386
11.6	Parametric Bootstrap for Model Comparisons	390
11.7	A Final Model for Examining Referee Bias	394
11.8	Estimated Random Effects	398
11.9	Notes on Using R (optional)	399
11.10	Exercises	401
11.10.1	Conceptual Exercises	401
11.10.2	Open-Ended Exercises	405
	Bibliography	409
	Index	417