Contents

1 Model Selection Loglinear Analysis 1

Loglinear Modeling Basics 2
A Two-Way Table 2
The Saturated Model 3
Main Effects 4
Interactions 5
Examining Parameters in a Saturated Model 6
Calculating the Missing Parameter Estimates 8
Testing Hypotheses about Parameters 9
Fitting an Independence Model 10
Specifying the Model 11
Checking Convergence 11
Chi-Square Goodness-of-Fit Tests 13
Hierarchical Models 14
Generating Classes 14
Selecting a Model 14
Evaluating Interactions 15
Testing Individual Terms in the Model 19
Model Selection Using Backward Elimination 20

2 Logit Loglinear Analysis 25

Dichotomous Logit Model 26
Loglinear Representation 26
Logit Model 27
Specifying the Model 28
Parameter Estimates for the Saturated Logit Model 28

Unsaturated Logit Model 30
Specifying the Analysis. 30
Goodness-of-Fit Statistics 31
Observed and Expected Cell Counts 31
Parameter Estimates 32
Measures of Dispersion and Association 33
Polychotomous Logit Model. 35
Specifying the Model 36
Goodness of Fit of the Model 36
Interpreting Parameter Estimates. 37
Examining Residuals 41
Covariates 42
Other Logit Models 42

3 Multinomial Logistic Regression 43

The Logit Model. 44
Baseline Logit Example 44
Specifying the Model. 46
Parameter Estimates 46
Likelihood-Ratio Test for Individual Effects 51
Likelihood-Ratio Test for the Overall Model 52
Evaluating the Model. 54
Calculating Predicted Probabilities and Expected Frequencies. . . 54
Classification Table 55
Goodness-of-Fit Tests 56
Examining the Residuals 58
Pseudo-R-square Measures. 58
Correcting for Overdispersion 59
Automated Variable Selection 59
Hierarchical Variable Entry 60
Specifying the Analysis. 61
Step Output 61
Likelihood-Ratio Tests for Individual Effects. 63

Matched Case-Control Studies 63
The Model 64
Creating the Difference Variables 65
The Data File 66
Specifying the Analysis 66
Examining the Results 67

4 Ordinal Regression 69

Fitting an Ordinal Logit Model 70
Modeling Cumulative Counts 70
Specifying the Analysis 73
Parameter Estimates 73
Testing Parallel Lines 74
Does the Model Fit? 75
Comparing Observed and Expected Counts 77
Including Additional Predictor Variables 78
Overall Model Test 79
Measuring Strength of Association 81
Classifying Cases 82
Generalized Linear Models 83
Link Function 83
Fitting a Heteroscedastic Probit Model 85
Modeling Signal Detection 85
Fitting a Location-Only Model 86
Fitting a Scale Parameter 88
Parameter Estimates 88
Model-Fitting Information 89

5 Probit Regression 91

Probit and Logit Response Models 92
Evaluating Insecticides 93
Confidence Intervals for Effective Dosages 96
Comparing Several Groups 97
Comparing Relative Potencies of the Agents 100
Estimating the Natural Response Rate 101
More than One Stimulus Variable 101

6 Kaplan-Meier Survival Analysis 103

PASW Statistics Procedures for Survival Data 103
Background 104
Calculating Length of Time 106
Estimating the Survival Function 107
Estimating the Conditional Probability of Survival 107
Estimating the Cumulative Probability of Survival 108
The PASW Statistics Kaplan-Meier Table 109
Plotting Survival Functions 111
Comparing Survival Functions 112
Specifying the Analysis 112
Comparing Groups 113
Stratified Comparisons of Survival Functions 116

7 Life Tables 121

Background 121
Studying Employment Longevity 122
The Body of a Life Table 123
Calculating Survival Probabilities 125

Assumptions Needed to Use the Life Table 128
Lost to Follow-up 129
Plotting Survival Functions 129
Comparing Survival Functions 131

8 Cox Regression 133

The Cox Regression Model 134
The Hazard Function 135
Proportional Hazards Assumption 136
Modeling Survival Times 136
Coding Categorical Variables 137
Specifying the Analysis 138
Testing Hypotheses about the Coefficient 139
Interpreting the Regression Coefficient 140
Baseline Hazard and Cumulative Survival Rates 141
Including Multiple Covariates 142
Model with Three Covariates 142
Global Tests of the Model 144
Plotting the Estimated Functions 146
Checking the Proportional Hazards Assumption 148
Stratification 148
Log-Minus-Log Survival Plot 149
Identifying Influential Cases 150
Examining Residuals 151
Partial (Schoenfeld) Residuals 152
Martingale Residuals 152
Selecting Predictor Variables 155
Variable Selection Methods 155
An Example of Forward Selection 156
Omnibus Test of the Model At Each Step 161
Time-Dependent Covariates 163
Examining the Data 163

Specifying a Time-Dependent Covariate 165
Calculating Segmented Time-Dependent Covariates. 167
Testing the Proportional Hazard Assumption with a Time-Dependent Covariate 167
Fitting a Conditional Logistic Regression Model 168
Data File Structure 169
Specifying the Analysis. 170
Parameter Estimates 170

9 Variance Components 171

Factors, Effects, and Models 171
Types of Factors. 171
Types of Effects 172
Types of Models. 172
Model for One-Way Classification 173
Estimation Methods. 174
Negative Variance Estimates 179
Nested Design Model for Two-Way Classification. 180
Univariate Repeated Measures Analysis Using a Mixed Model Approach 184
Background Information. 190
Model. 190
Distribution Assumptions. 191
Estimation Methods. 192

10 Linear Mixed Models 195

The Linear Mixed Model. 195
Background 197
Example 1: Unconditional Random-Effects Models. 199
Example 2: Adding a Gender Fixed Effect 206

Example 3: Hierarchical Models 211
Example 4: Random-Coefficient Model 215
Example 5: A Model with School-Level and Individual-Level Covariates 218
Example 6: A Three-Level Hierarchical Model 222
Example 7: Repeated Measurements 227
Example 8: Selecting a Residual Covariance Structure 237

11 Generalized Linear Models 245

The Basics of Generalized Linear Models 246
Selecting the Type of Model 247
Count Data 248
Continuous Data 249
Mixture Distribution 249
Selecting a Link Function 250
Fitting a Logistic Regression 252
Specifying the Analysis 253
Likelihood-Ratio Tests 254
How Well Does the Model Fit? 254
Testing Hypotheses about Coefficients 259
Testing Hypotheses about Effects 260
Model Diagnostics 262
Fitting a Poisson Rate Model 262
Specifying the Analysis 264
The Model 264
Interpreting the Parameter Estimates 266
Fitting a Probit Model 267
Specifying the Model 268
Fitting a Poisson Count Model 269
Overdispersion and Underdispersion 270
Fitting a Loglinear Model 272
Specifying the Analysis 272
Parameter Estimates 273

Fitting a Gamma Distribution to Survival Time 274
Parameter Estimates 274
Estimated Means 276

12 Generalized Estimating Equations 279

The GEE Model 279
Data Organization 280
Specifying the Analysis. 281
Tests of Model Effects 283
Correlation Structure 285
Comparing Correlation Structures 286
Comparing Models 289
Variance Estimators 289
Fitting a Poisson Model 290
Selecting a Model. 292
Diagnostic Statistics 295

13 Nonlinear Regression 297

What Is a Nonlinear Model? 297
Transforming Nonlinear Models 298
Intrinsically Nonlinear Models 299
Fitting a Logistic Population Growth Model 300
Estimating a Nonlinear Model 300
Finding Starting Values 301
Specifying the Analysis. 302
Approximate Confidence Intervals for the Parameters 306
Bootstrapped Estimates 306
Estimating Starting Values 308
Use Starting Values from Previous Analysis. 308
Look for a Linear Approximation 308

Use Properties of the Nonlinear Model 309
Solve a System of Equations 309
Computational Issues 310
Additional Nonlinear Regression Options 310
Nonlinear Regression Common Models 311
Specifying a Segmented Model 312

14 Two-Stage Least-Squares Regression 315

Artichoke Data 315
Demand-Price-Income Economic Model 316
Estimation with Ordinary Least Squares 317
Feedback and Correlated Errors 317
Two-Stage Least Squares 319
Strategy 320
Stage 1: Estimating Price 321
Stage 2: Estimating the Model 322
2-Stage Least Squares Procedure 323

15 Weighted Least-Squares Regression 325

Diagnosing the Problem 325
Estimating the Weights 327
Estimating Weights as Powers 328
Specifying the Analysis 329
Examining the Log-Likelihood Functions 329
WLS Solutions 330
Estimating Weights from Replicates 331
Diagnostics from the Linear Regression Procedure 331

16 Multidimensional Scaling 335

Data, Models, and Analysis of Multidimensional Scaling 336
Example: Flying Mileages 337
Nature of Data Analyzed in MDS 340
Measurement Level of Data 340
Shape of Data 340
Conditionality of Data 341
Missing Data 342
Multivariate Data 342
Classical MDS 343
Example: Flying Mileages Revisited 343
Euclidean Model 347
Details of CMDS 350
Example: Ranked Flying Mileages 354
Repeated CMDS 360
Replicated MDS 361
Details of RMDS 361
Example: Perceived Body-Part Structure 364
Weighted MDS 373
Geometry of the Weighted Euclidean Model 374
Algebra of the Weighted Euclidean Model 380
Matrix Algebra of the Weighted Euclidean Model 382
Details of WMDS 384
Example: Perceived Body-Part Structure 385
Weirdness Index 398
Flattened Weights 402

Bibliography 405

Index 411