dandelon.com
© 2008 AGI-Information Management Consultants
May be used for personal purposes only or by
libraries associated to dandelon.com network.

Ronald Christensen

Log-Linear Models and Logistic Regression

Second Edition

Springer

Contents

Preface to the Second Edition vn

Preface to the First Edition IX

Introduction 1
1.1 Conditional Probability and Independence . . . 2
1.2 Random Variables and Expectations . . . 11
1.3 The Binomial Distribution . . . 13
1.4 The Multinomial Distribution . . . 14
1.5 The Poisson Distribution . . . 18
1.6 Exercises . . . 20

Two-Dimensional Tables and Simple Logistic Regression 23
2.1 Two Independent Binomials . . . 23
2.1.1 The Odds Ratio . . . 29
2.2 Testing Independence in a 2 x 2 Table . . . 30
2.2.1 The Odds Ratio . . . 32
2.3 / x J Tables . . . 33
2.3.1 Response Factors . . . 37
2.3.2 Odds Ratios . . . 38
2.4 Maximum Likelihood Theory for Two-Dimensional Tables . . . 42
2.5 Log-Linear Models for Two-Dimensional Tables . . . 47
2.5.1 Odds Ratios . . . 51

2.6 Simple Logistic Regression . . . 54
2.6.1 Computer Commands . . . 61
2.7 Exercises . . . 61

Three-Dimensional Tables **69**
3.1 Simpson's Paradox and the Need for Higher-Dimensional Tables . . . 70
3.2 Independence and Odds Ratio Models . . . 72
3.2.1 The Model of Complete Independence . . . 72
3.2.2 Models with One Factor Independent of the Other Two . . . 75
3.2.3 Models of Conditional Independence . . . 79
3.2.4 A Final Model for Three-Way Tables . . . 83
3.2.5 Odds Ratios and Independence Models . . . 85
3.3 Iterative Computation of Estimates . . . 87
3.4 Log-Linear Models for Three-Dimensional Tables . . . 89
3.4.1 Estimation . . . 92
3.4.2 Testing Models . . . 94
3.5 Product-Multinomial and Other Sampling Plans . . . 99
3.5.1 Other Sampling Models . . . 102
3.6 Model Selection Criteria . . . 104
3.6.1 R^2 . . . 104
3.6.2 Adjusted $R?$ . . . 105
3.6.3 Akaike's Information Criterion . . . 106
3.7 Higher-Dimensional Tables . . . 108
3.7.1 Computer Commands . . . 110
3.8 Exercises . . . 113

Logistic Regression, Logit Models, and Logistic Discrimination **116**
4.1 Multiple Logistic Regression . . . 120
4.1.1 Informal Model Selection . . . 122
4.2 Measuring Model Fit . . . 127
4.2.1 Checking Lack of Fit . . . 129
4.3 Logistic Regression Diagnostics . . . 130
4.4 Model Selection Methods . . . 136
4.4.1 Computations for Nonbinary Data . . . 138
4.4.2 Computer Commands . . . 139
4.5 ANOVA Type Logit Models . . . 141
4.5.1 Computer Commands . . . 149
4.6 Logit Models for a Multinomial Response . . . 150
4.7 Logistic Discrimination and Allocation . . . 159
4.8 Exercises . . . 170

Independence Relationships and Graphical Models **178**
5.1 Model Interpretations . . . 178

5.2 Graphical and Decomposable Models . . . 182
5.3 Collapsing Tables . . . 192
5.4 Recursive Causal Models . . . 195
5.5 Exercises . . . 209

6 Model Selection Methods and Model Evaluation 211
6.1 Stepwise Procedures for Model Selection . . . 212
6.2 Initial Models for Selection Methods . . . 215
6.2.1 All s-Factor Effects . . . 215
6.2.2 Examining Each Term Individually . . . 217
6.2.3 Tests of Marginal and Partial Association . . . 217
6.2.4 Testing Each Term Last . . . 218
6.3 Example of Stepwise Methods . . . 224
6.3.1 Forward Selection . . . 226
6.3.2 Backward Elimination . . . 230
6.3.3 Comparison of Stepwise Methods . . . 232
6.3.4 Computer Commands . . . 233
6.4 Aitkin's Method of Backward Selection . . . 234
6.5 Model Selection Among Decomposable and Graphical Models . . . 240
6.6 Use of Model Selection Criteria . . . 246
6.7 Residuals and Influential Observations . . . 247
6.7.1 Computations . . . 249
6.7.2 Computing Commands . . . 253
6.8 Drawing Conclusions . . . 254
6.9 Exercises . . . 256

7 Models for Factors with Quantitative Levels 258
7.1 Models for Two-Factor Tables . . . 259
7.1.1 Log-Linear Models with Two Quantitative Factors . . . 260
7.1.2 Models with One Quantitative Factor . . . 262
7.2 Higher-Dimensional Tables . . . 266
7.2.1 Computing Commands . . . 268
7.3 Unknown Factor Scores . . . 269
7.4 Logit Models . . . 275
7.5 Exercises . . . 277

8 Fixed and Random Zeros 279
8.1 Fixed Zeros . . . 279
8.2 Partitioning Polytomous Variables . . . 282
8.3 Random Zeros . . . 286
8.4 Exercises . . . 293

9 Generalized Linear Models **297**
9.1 Distributions for Generalized Linear Models 299
9.2 Estimation of Linear Parameters 304
9.3 Estimation of Dispersion and Model Fitting 306
9.4 Summary and Discussion 311
9.5 Exercises 313

10 The Matrix Approach to Log-Linear Models **314**
10.1 Maximum Likelihood Theory for Multinomial Sampling . 318
10.2 Asymptotic Results 322
10.3 Product-Multinomial Sampling 339
10.4 Inference for Model Parameters 342
10.5 Methods for Finding Maximum Likelihood Estimates . . . 345
10.6 Regression Analysis of Categorical Data 347
10.7 Residual Analysis and Outliers 354
10.8 Exercises 360

11 The **Matrix Approach to Logit Models** **363**
11.1 Estimation and Testing for Logistic Models 363
11.2 Model Selection Criteria for Logistic Regression 371
11.3 Likelihood Equations and Newton-Raphson 372
11.4 Weighted Least Squares for Logit Models 375
11.5 Multinomial Response Models 377
11.6 Asymptotic Results 378
11.7 Discrimination, Allocation, and Retrospective Data 387
11.8 Exercises 394

12 Maximum Likelihood Theory for Log-Linear Models **396**
12.1 Notation 396
12.2 Fixed Sample Size Properties 397
12.3 Asymptotic Properties 402
12.4 Applications 412
12.5 Proofs of Lemma 12.3.2 and Theorem 12.3.8 418

13 Bayesian Binomial Regression **422**
13.1 Introduction 422
13.2 Bayesian Inference 424
13.2.1 Specifying the Prior and Approximating the Posterior 424
13.2.2 Predictive Probabilities 434
13.2.3 Inference for Regression Coefficients 436
13.2.4 Inference for LD_a 438
13.3 Diagnostics 440
13.3.1 Case Deletion Influence Measures 441
13.3.2 Model Checking 446

13.3.3 Link Selection 447
13.3.4 Sensitivity Analysis 448
13.4 Posterior Computations and Sample Size Calculation . . . 449

Appendix: Tables **455**
A.I The Greek Alphabet 455
A.2 Tables of the x^2 Distribution 456

References **458**

Author Index **475**

Subject Index **479**