

# Contents

<b>1</b>	<b>Introduction</b> .....	1
1.1	Categorical Data .....	1
1.1.1	Measurement Scales .....	1
1.1.2	Response and Explanatory Variables .....	2
1.2	Discrete Distributions and Related Inference Problems .....	2
1.2.1	Binomial Distribution .....	3
1.2.2	Multinomial Distribution .....	4
1.2.3	Poisson Distribution .....	6
1.2.4	Hypergeometric Distribution .....	8
1.3	Statistical Inference with Categorical Data .....	9
1.4	Classes of Models for Discrete Data .....	12
1.5	Outline of the Book .....	14
<b>2</b>	<b>Analysis of Two-way Tables</b> .....	17
2.1	Analyzing $2 \times 2$ Tables .....	17
2.1.1	Independence of Two Binary Variables .....	19
2.1.2	Example 2.1(a) .....	21
2.1.3	Comparison of Two Independent Proportions .....	22
2.1.4	Example 2.1(b) .....	24
2.1.5	The Odds Ratio .....	25
2.1.6	Example 2.1 (Continued) .....	28
2.1.7	Fisher's Exact Test .....	29
2.1.8	Example 2.2 .....	31
2.2	Analyzing $I \times J$ Tables .....	33
2.2.1	Possible Sampling Schemes .....	33
2.2.2	Test of Independence .....	35
2.2.3	Example 2.3 .....	36
2.2.4	Analysis of Residuals .....	38
2.2.5	Odds Ratios for $I \times J$ Tables .....	40
2.2.6	Example 2.4 .....	45

- 2.3 Test of Independence for Ordinal Variables ..... 46
  - 2.3.1 The Choice of Scores ..... 47
  - 2.3.2 Example 2.5 ..... 48
  - 2.3.3 The Linear Trend Test in R ..... 49
- 2.4 Graphs for Two-way Tables ..... 50
  - 2.4.1 Barplots ..... 50
  - 2.4.2 Fourfold Plots ..... 53
  - 2.4.3 Sieve Diagrams ..... 54
  - 2.4.4 Mosaic Plots ..... 55
- 2.5 Overview and Further Reading ..... 57
  - 2.5.1 The Continuity Correction ..... 57
  - 2.5.2  $2 \times 2$  Tables and the Odds Ratio ..... 57
  - 2.5.3 Inference for Two-way Tables ..... 59
  - 2.5.4 Partitioning of the  $X^2$  Statistic ..... 60
  - 2.5.5 Ordinal Odds Ratios and Positive Dependencies ..... 60
- 3 Analysis of Multi-way Tables ..... 63**
  - 3.1 Describing Multi-way Contingency Tables ..... 63
    - 3.1.1 Example 3.1 ..... 64
  - 3.2 On Partial and Marginal Tables ..... 65
    - 3.2.1 Joint, Conditional, and Marginal Probabilities ..... 65
    - 3.2.2 Conditional and Marginal Odds Ratios  
for  $2 \times 2 \times K$  Tables ..... 65
    - 3.2.3 Odds Ratios for Tables of Higher Dimension ..... 67
    - 3.2.4 Example 3.2 ..... 67
  - 3.3 Analysis of  $K$   $2 \times 2$  Tables ..... 69
    - 3.3.1 The Mantel–Haenszel Test ..... 71
    - 3.3.2 Homogeneous Association Tests ..... 72
    - 3.3.3 Example 3.1 (Continued) ..... 73
    - 3.3.4 Example 3.3 ..... 75
  - 3.4 Types of Independence for Three-way Tables ..... 76
  - 3.5 Graphs for Multi-way Contingency Tables ..... 78
    - 3.5.1 Fourfold Plots for  $2 \times 2 \times K$  Tables ..... 78
    - 3.5.2 Sieve Diagrams for Multi-way Tables ..... 79
    - 3.5.3 Mosaic Plots for Multi-way Tables ..... 81
  - 3.6 Overview and Further Reading ..... 81
    - 3.6.1 Stratified  $2 \times 2$  Contingency Tables ..... 81
    - 3.6.2 Generalized Mantel–Haenszel Test for  
 $I \times J \times K$  Contingency Tables ..... 82
    - 3.6.3 Visualization of Categorical Data ..... 83
- 4 Log-Linear Models ..... 85**
  - 4.1 Log-Linear Models for Two-way Tables ..... 85
    - 4.1.1 Model of Independence ..... 85
    - 4.1.2 The Saturated Model ..... 87

- 4.2 On Inference and Fit of Log-Linear Models ..... 88
  - 4.2.1 Example 2.4 (Continued)..... 90
  - 4.2.2 Example 2.3 (Continued)..... 93
- 4.3 Log-Linear Models for Three-way Contingency Tables ..... 94
- 4.4 Hierarchical Log-Linear Models for Multi-way Tables ..... 97
- 4.5 Maximum Likelihood Estimation for Log-Linear Models ..... 98
- 4.6 Model Fit and Selection..... 100
  - 4.6.1 Conditional Test of Conditional Independence ..... 102
  - 4.6.2 Log-Linear Model for Example 3.2..... 103
- 4.7 Graphical Models ..... 110
  - 4.7.1 Undirected Graphs ..... 110
  - 4.7.2 Graphical Log-Linear Models ..... 111
- 4.8 Collapsibility in Multi-way Tables ..... 113
  - 4.8.1 Collapsing for Example 3.2 ..... 115
  - 4.8.2 Example 4.1 ..... 116
- 4.9 Overview and Further Reading ..... 118
  - 4.9.1 On Log-Linear Models Analysis..... 118
  - 4.9.2 Residual Analysis: Outlier Detection ..... 120
  - 4.9.3 On Graphical Log-Linear Models ..... 121
  - 4.9.4 On Collapsibility ..... 121
  - 4.9.5 Information-Theoretic Approach in  
Contingency Table Analysis ..... 122
- 5 Generalized Linear Models and Extensions ..... 125**
  - 5.1 The Generalized Linear Model (GLM) in Keywords..... 125
  - 5.2 Log-Linear Model: Member of the GLM Family..... 127
  - 5.3 Inference for GLMs ..... 129
    - 5.3.1 ML Estimation for GLMs ..... 129
    - 5.3.2 Evaluating Model Fit for GLMs ..... 131
    - 5.3.3 Residuals ..... 133
    - 5.3.4 Model Selection in GLMs..... 134
  - 5.4 Software for GLMs..... 135
    - 5.4.1 Example 2.4 by glm..... 136
    - 5.4.2 Example 3.1 (Revisited)..... 138
  - 5.5 Independence for Incomplete Tables ..... 142
    - 5.5.1 Example 5.1 ..... 144
  - 5.6 Models for Joint and Marginal Distributions ..... 145
    - 5.6.1 Example 2.4 by mph..... 147
    - 5.6.2 Example 3.3 by mph..... 149
  - 5.7 Overview and Further Reading ..... 149
    - 5.7.1 Incomplete Contingency Tables..... 150
    - 5.7.2 Marginal Distributions Modeling ..... 151
- 6 Association Models ..... 153**
  - 6.1 Basic Association Models for Two-way Tables..... 153
    - 6.1.1 Linear-by-Linear Association Model ..... 154

6.1.2	Example 6.1 .....	156
6.1.3	Row and Column Effect Models .....	157
6.1.4	Row by Column Effect Model .....	158
6.1.5	Example 6.1 (Revisited) .....	159
6.2	Maximum Likelihood Estimation for Association Models .....	161
6.3	Association Model Selection .....	163
6.3.1	Model Selection for Example 6.1 .....	164
6.4	Features of Association Models .....	164
6.5	Association Models of Higher Order: The $RC(M)$ Model .....	166
6.5.1	Maximum Likelihood Estimation of the $RC(M)$ Model .....	169
6.5.2	Example 6.2 .....	169
6.6	Software Applications for Association Models .....	171
6.6.1	Association Models in R: Example 6.1 .....	172
6.6.2	The $RC(M)$ Model in R: Example 6.2 .....	177
6.6.3	Example 2.4 (Revisited) .....	178
6.6.4	Association Models Fitted on the Local Odds Ratios .....	180
6.7	Association Models for Multi-way Tables .....	181
6.7.1	Example 6.3 .....	183
6.7.2	Homogeneous Uniform Association .....	187
6.8	Overview and Further Reading .....	191
6.8.1	Multi-way Association Models .....	193
6.8.2	Order-Restricted Inference .....	194
6.8.3	Comparison of Two Ordinal Responses .....	194
6.8.4	Cell Frequencies vs. Local Odds Ratios Modeling .....	196
<b>7</b>	<b>More on Association Models and Related Methods .....</b>	<b>197</b>
7.1	Association Models for Global Odds Ratios .....	197
7.1.1	The $U^G$ Model in R: Example 6.1 .....	198
7.2	Correspondence Analysis .....	199
7.2.1	Simple Correspondence Analysis in Steps .....	200
7.2.2	Correspondence Analysis of Example 6.2 .....	203
7.3	Correlation Models .....	206
7.4	Generalized Association Models .....	207
7.5	The Role of Scores in Merging Categories .....	208
7.5.1	Example 6.2 (Continued) .....	210
7.6	Overview and Further Reading .....	211
7.6.1	Homogeneity Analysis .....	212
7.6.2	Canonical Correlation and Correspondence Analysis .....	212
<b>8</b>	<b>Response Variable Analysis in Contingency Tables .....</b>	<b>215</b>
8.1	Logit Models for Binary Response .....	215
8.1.1	Logit Models for Ordinal Explanatory Variables .....	217
8.1.2	Inference for Logit Models .....	218
8.1.3	Logit Models in R .....	219
8.2	Logit Analysis of Stratified $2 \times 2$ Contingency Tables .....	222

- 8.3 Logit Models for Multi-category Response ..... 223
  - 8.3.1 Nominal Response ..... 223
  - 8.3.2 Ordinal Response: The Cumulative Logit Model ..... 223
  - 8.3.3 Alternative Models for Ordinal Response ..... 225
  - 8.3.4 Example 6.1 (Continued)..... 226
- 8.4 Overview and Further Reading ..... 229
  - 8.4.1 Alternative Links to the Logit..... 229
  - 8.4.2 Chain Graph Models and Collapsibility ..... 230
  - 8.4.3 The Rasch Model ..... 230
  - 8.4.4 The Stereotype Model ..... 231
- 9 Analysis of Square Tables** ..... 233
  - 9.1 Comparison of Two Dependent Proportions ..... 233
    - 9.1.1 Example 9.1 ..... 235
  - 9.2 Symmetry Models ..... 236
    - 9.2.1 Complete Symmetry ..... 236
    - 9.2.2 Marginal Homogeneity ..... 237
    - 9.2.3 Quasi Symmetry ..... 238
    - 9.2.4 Conditional (or Triangular) Symmetry ..... 240
    - 9.2.5 Diagonal Symmetry ..... 241
    - 9.2.6 Software for Symmetry Models..... 242
    - 9.2.7 Example 9.2..... 244
  - 9.3 Quasi-Independence Models for Square Tables ..... 246
    - 9.3.1 Example 9.3 ..... 246
    - 9.3.2 Example 9.4..... 247
  - 9.4 Symmetry Models with Scores ..... 248
    - 9.4.1 Homogeneous Association Models ..... 248
    - 9.4.2 Ordinal Quasi Symmetry ..... 249
    - 9.4.3 Example 9.2 (Continued)..... 250
  - 9.5 Rater Agreement ..... 252
    - 9.5.1 Example 9.5..... 253
    - 9.5.2 Agreement on Ordinal Rating Scales ..... 253
    - 9.5.3 Cohen’s Kappa in  $R$  ..... 254
  - 9.6 The Bradley–Terry Model ..... 255
  - 9.7 Overview and Further Reading ..... 256
    - 9.7.1 Mobility Tables and the Mover–Stayer Model ..... 257
    - 9.7.2 Measuring Agreement ..... 257
    - 9.7.3 Symmetry Models for Multi-way Tables ..... 258
    - 9.7.4 Clustered Categorical Data..... 258
- 10 Further Topics** ..... 261
  - 10.1 Overview..... 261
  - 10.2 On Measures of Association ..... 262
  - 10.3 Alternative Approaches in Contingency Table Analysis ..... 264
    - 10.3.1 Latent Class Models ..... 264

- 10.3.2 Graphical Models ..... 264
- 10.3.3 Smoothing Categorical Data ..... 265
- 10.4 Small Sample Inference for Contingency Tables ..... 265
- 10.5 Bayesian Analysis of Contingency Tables ..... 267
- 10.6 Extreme High-Dimensional Categorical Data ..... 269
  
- A Appendix: Contingency Table Analysis in Practice**..... 271
  - A.1 Software for Categorical Data Analysis ..... 271
  - A.2 Contingency Table Analysis with R ..... 272
    - A.2.1 R Packages for Contingency Table Analysis ..... 272
    - A.2.2 Data Input in R..... 272
  - A.3 R Functions Used ..... 273
    - A.3.1 R Functions of Chap. 1 ..... 273
    - A.3.2 R Functions of Chap. 2 ..... 273
    - A.3.3 R Functions of Chap. 3 ..... 273
    - A.3.4 R Functions of Chap. 5 ..... 273
    - A.3.5 R Functions of Chap. 6 ..... 274
    - A.3.6 R Functions of Chap. 9 ..... 274
  - A.4 Contingency Table Analysis with SPSS..... 274
  
- References**..... 275
  
- Index**..... 301