Acknowledgements v
Permission of reproduction v

1 Statistics and the study of language behaviour 1
1.1 The structure of the book 1
1.2 Basic statistical concepts 3
1.3 A few preliminary remarks 6
1.4 Other techniques 7
1.5 Language statistics 10

2 Analysis of variance 13
2.1 Introduction 13
2.2 A simple example 14
2.3 One-way analysis of variance 16
2.4 Testing effects: the F distribution 22
2.5 Multi-factorial designs and interaction 26
2.6 Random and fixed factors 30
2.7 Testing effects in a two-factor design 33
2.8 The interpretation of interactions 35
2.9 Summary of the procedure 40
2.10 Other design types 41
2.11 A hierarchical three-factor design 47
2.12 Simple main effects 49
2.13 Post hoc comparisons and contrasts 51
2.14 Strength of association 58
2.15 Strange F ratios: testing hypotheses made difficult 60
2.16 Pooling and unequal cell frequencies 63
2.17 Overview of the steps in an analysis of variance 64
Basic terms and concepts 66
Exercises 67
References 69

3 Multiple regression 71
3.1 Introduction 71
3.2 Simple regression 73
3.3 Tests of significance 79

3.4 Two independent variables 81
3.5 Selecting a model 90
3.6 Multicollinearity 98
3.7 Coding nominal independent variables 101
3.8 Comparing regression lines 107
3.9 Measurement and specification errors 113
Basic terms and concepts 114
Exercises 115
References 117

4 More ANOVA and MR: assumptions and alternatives 119
4.1 Introduction 119
4.2 Assumptions in analysis of variance 120
4.3 Homoscedasticity 122
4.4 The impact of transformations 124
4.5 Unbalanced designs: unequal cell frequencies 128
4.6 Repeated measures and MANOVA 133
4.7 A nonparametric alternative: randomization tests 142
4.8 Bootstrapping 149
4.9 Assumptions in linear multiple regression 157
4.10 Plots and diagnostics 159
4.11 Outliers and influential observations 163
4.12 Linearizing transformations 168
4.13 Time series analysis 174
4.14 Alternative regression techniques 177
Basic terms and concepts 179
Exercises 180
References 182

5 Reliability and agreement among raters 187
5.1 Introduction 187
5.2 Reliability: true scores and the error component 190
5.3 Interrater reliability 193
5.3.1 Reliability and score models 193
5.3.2 Testing reliability 201
5.3.3 Testing differences between reliability coefficients 204
5.3.4 An example from recent research 205
5.3.5 Low reliability coefficients 207
5.4 Interrater agreement 210
5.4.1 Scale type 210

5.4.2 Ordinal scales 213
5.4.3 Nominal scales 217
5.5 Intrarater reliability and agreement 224
Basic terms and concepts 225
Exercises 226
References 227

6 Introductory matrix algebra 229
6.1 Introduction 229
6.2 Matrices and vectors 229
6.3 Matrix operations 232
6.4 Three applications 234
6.5 Some special matrices 239
6.6 Some key concepts in matrix algebra 240
6.6.1 The transpose 240
6.6.2 The determinant and the rank of matrices 241
6.6.3 The inverse of a matrix 242
6.6.4 The eigenvalues of a matrix 242
6.7 Some matrix operations for statistical data 243
Basic terms and concepts 246
Exercises 247
References 248

7 Factor analysis 251
7.1 Introduction 251
7.2 Dimensionality reduction 252
7.3 Axis rotation and linear tranformation 256
7.4 Criteria for dimensionality reduction 258
7.5 The role of eigenvalues in principal component analysis 259
7.6 Loadings 263
7.7 A PC analysis with SPSS: the two-variable example 265
7.8 Principal component analysis or factor analysis? 267
7.9 How many factors/components are to be retained? 272
7.10 Which factor loadings are significant? 274
7.11 Data problems 275
7.12 Rotations revisited: rotating factors 278
7.13 Oblique factors 280
7.14 Options 282
7.15 An example of factor analysis 284
7.16 Factor scores 289

7.17 Overview of the steps in a factor analysis 291
Basic terms and concepts 292
Exercises 292
References 293

8 Analysis of frequencies 297
8.1 Introduction 297
8.2 Log-linear analysis of a 2×2 table 298
8.3 Multi-dimensional contingency tables 307
8.3.1 A three-dimensional example 307
8.3.2 Carrying out a log-linear analysis with SPSS 310
8.3.3 A less successful model 315
8.3.4 Another example 315
8.3.5 Low frequencies and structural zeros 317
8.4 Logit analysis 317
8.5 Overview of the steps in a log-linear analysis 323
Basic terms and concepts 324
Exercises 324
References 325

9 Logistic regression 327
9.1 Introduction 327
9.2 Regression and the logistic function 328
9.3 Evaluating a model 334
9.4 Polytomous independent variables 339
9.5 Stepwise selection 343
9.6 The model matrix approach and GSK 350
9.7 Comparing different techniques 357
Basic terms and concepts 358
Exercises 360
References 361

Appendices 363
A: Logarithms 365
B: Characteristic root test 367
C: Eigenvalues of random data correlation matrices 369
D: Command lines statistical programs 371
E: Key to the exercises 379

F: Statistical tables 387
A. Normal Distribution 388
B. Critical values of F 389
C. Critical values of χ^2 393

Index 395