Contents

List of Tables xvii

List of Figures xix

List of Listings xxii

Preface xxiii

1 Introduction 1

1.1 Origins and motivation 1

1.2 Notational conventions 3

1.3 Applied or theoretical? 4

1.4 Road map 4

1.5 Installing the support materials 6

I Foundations of Generalized Linear Models 7

2 GLMs 9

2.1 Components 11

2.2 Assumptions 12

2.3 Exponential family 13

2.4 Example: Using an offset in a GLM 15

2.5 Summary 16

3 GLM estimation algorithms 19

3.1 Newton–Raphson (using the observed Hessian) 25

3.2 Starting values for Newton–Raphson 27

3.3 IRLS (using the expected Hessian) 28

3.4 Starting values for IRLS 31

3.5 Goodness of fit 31

3.6 Estimated variance matrices 32

3.6.1 Hessian 34
3.6.2 Outer product of the gradient 35
3.6.3 Sandwich 35
3.6.4 Modified sandwich 36
3.6.5 Unbiased sandwich 37
3.6.6 Modified unbiased sandwich 38
3.6.7 Weighted sandwich: Newey–West 38
3.6.8 Jackknife 40
3.6.8.1 Usual jackknife 40
3.6.8.2 One-step jackknife 41
3.6.8.3 Weighted jackknife 41
3.6.8.4 Variable jackknife 41
3.6.9 Bootstrap 42
3.6.9.1 Usual bootstrap 42
3.6.9.2 Grouped bootstrap 43
3.7 Estimation algorithms 43
3.8 Summary 44

4 Analysis of fit **47**
4.1 Deviance 48
4.2 Diagnostics 49
4.2.1 Cook's distance 49
4.2.2 Overdispersion 49
4.3 Assessing the link function 50
4.4 Checks for systematic departure from the model 51
4.5 Residual analysis 52
4.5.1 Response residuals 53
4.5.2 Working residuals 53
4.5.3 Pearson residuals 54
4.5.4 Partial residuals 54
4.5.5 Anscombe residuals 54

4.5.6 Deviance residuals 55
4.5.7 Adjusted deviance residuals 55
4.5.8 Likelihood residuals 55
4.5.9 Score residuals 55
4.6 Model statistics 55
4.6.1 Criterion measures 56
4.6.1.1 AIC 56
4.6.1.2 BIC 57
4.6.2 The interpretation of R^2 in linear regression 58
4.6.2.1 Percent variance explained 58
4.6.2.2 The ratio of variances 58
4.6.2.3 A transformation of the likelihood ratio 59
4.6.2.4 A transformation of the F test 59
4.6.2.5 Squared correlation 59
4.6.3 Generalizations of linear regression R^2 interpretations 59
4.6.3.1 Efron's pseudo-R^2 60
4.6.3.2 McFadden's likelihood-ratio index 60
4.6.3.3 Ben-Akiva and Lerman adjusted likelihood-ratio index 60
4.6.3.4 McKelvey and Zavoina ratio of variances 61
4.6.3.5 Transformation of likelihood ratio 61
4.6.3.6 Cragg and Uhler normed measnre 61
4.6.4 More R^2 measures 62
4.6.4.1 The count R^2 62
4.6.4.2 The adjusted count R^2 62
4.6.4.3 Veall and Zimmermann R^2 62
4.6.4.4 Cameron–Windmeijer R^2 62

II Continuous-Response Models 65

5 The Gaussian family 67

5.1 Derivation of the GLM Gaussian family 68

5.2 Derivation in terms of the mean . . . 68
5.3 IRLS GLM algorithm (nonbinomial) . . . 70
5.4 Maximum likelihood estimation . . . 73
5.5 GLM log-normal models . . . 74
5.6 Expected versus observed information matrix . . . 75
5.7 Other Gaussian links . . . 77
5.8 Example: Relation to OLS . . . 77
5.9 Example: Beta-carotene . . . 79

6 The gamma family **89**

6.1 Derivation of the gamma model . . . 90
6.2 Example: Reciprocal link . . . 92
6.3 Maximum likelihood estimation . . . 95
6.4 Log-gamma models . . . 96
6.5 Identity-gamma models . . . 100
6.6 Using the gamma model for survival analysis . . . 101

7 The inverse Gaussian family **105**

7.1 Derivation of the inverse Gaussian model . . . 105
7.2 The inverse Gaussian algorithm . . . 107
7.3 Maximum likelihood algorithm . . . 107
7.4 Example: The canonical inverse Gaussian . . . 108
7.5 Noncanonical links . . . 109

8 The power family and link **113**

8.1 Power links . . . 113
8.2 Example: Power link . . . 114
8.3 The power family . . . 115

III Binomial Response Models 117

9 The binomial–logit family **119**

9.1 Derivation of the binomial model . . . 120
9.2 Derivation of the Bernoulli model . . . 123

12.4 Example: Testing overdispersion in the Poisson model 192
12.5 Using the Poisson model for survival analysis 194
12.6 Using offsets to compare models 195
12.7 Interpretation of coefficients 197

13 The negative binomial family **199**

13.1 Constant overdispersion 201
13.2 Variable overdispersion 203
13.2.1 Derivation in terms of a Poisson–gamma mixture 203
13.2.2 Derivation in terms of the negative binomial probability function 206
13.2.3 The canonical link negative binomial parameterization 207
13.3 The log-negative binomial parameterization 209
13.4 Negative binomial examples 211
13.5 The geometric family 215
13.6 Interpretation of coefficients 218

14 Other count data models **221**

14.1 Count response regression models 221
14.2 Zero-truncated models 224
14.3 Zero-inflated models 227
14.4 Hurdle models 232
14.5 Heterogeneous negative binomial models 235
14.6 Generalized Poisson regression models 239
14.7 Censored count response models 241

V Multinomial Response Models **249**

15 The ordered-response family **251**

15.1 Ordered outcomes for general link 252
15.2 Ordered outcomes for specific links 254
15.2.1 Ordered logit 254
15.2.2 Ordered probit 255
15.2.3 Ordered clog-log 255

15.2.4 Ordered log-log 256
15.2.5 Ordered cauchit 256
15.3 Generalized ordered outcome models 257
15.4 Example: Synthetic data 258
15.5 Example: Automobile data 263
15.6 Partial proportional-odds models 269
15.7 Continuation ratio models 273

16 Unordered-response family **279**

16.1 The multinomial logit model 280
16.1.1 Example: Relation to logistic regression 280
16.1.2 Example: Relation to conditional logistic regression 281
16.1.3 Example: Extensions with conditional logistic regression . . . 283
16.1.4 The independence of irrelevant alternatives 284
16.1.5 Example: Assessing the IIA 285
16.1.6 Interpreting coefficients 287
16.1.7 Example: Medical admissions—introduction 287
16.1.8 Example: Medical admissions—summary 289
16.2 The multinomial probit model 295
16.2.1 Example: A comparison of the models 297
16.2.2 Example: Comparing probit and multinomial probit 299
16.2.3 Example: Concluding remarks 302

VI Extensions to the GLM 305

17 Extending the likelihood **307**

17.1 The quasilikelihood 307
17.2 Example: Wedderburn's leaf blotch data 308
17.3 Generalized additive models 316

18 Clustered data **319**

18.1 Generalization from individual to clustered data 319
18.2 Pooled estimators 320

18.3 Fixed effects 321
18.3.1 Unconditional fixed-effects estimators 322
18.3.2 Conditional fixed-effects estimators 323
18.4 Random effects 325
18.4.1 Maximum likelihood estimation 325
18.4.2 Gibbs sampling 329
18.5 GEEs 330
18.6 Other models 333

VII Stata Software **337**

19 Programs for Stata **339**

19.1 The glm command 340
19.1.1 Syntax 340
19.1.2 Description 341
19.1.3 Options 341
19.2 The predict command after glm 345
19.2.1 Syntax 345
19.2.2 Options 345
19.3 User-written programs 347
19.3.1 Global macros available for user-written programs 347
19.3.2 User-written variance functions 348
19.3.3 User-written programs for link functions 350
19.3.4 User-written programs for Newey–West weights 352
19.4 Remarks 353
19.4.1 Equivalent commands 353
19.4.2 Special comments on family(Gaussian) models 353
19.4.3 Special comments on family(binomial) models 353
19.4.4 Special comments on family(nbinomial) models 354
19.4.5 Special comment on family(gamma) link(log) models 354

A Tables **355**

References **369**

Author index **379**

Subject index **383**