

Contents

List of figures	xix
List of tables	xxiii
List of listings	xxviii
Preface	xxix
1 Introduction	1
1.1 Origins and motivation	2
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	17
3 GLM estimation algorithms	19
3.1 Newton–Raphson (using the observed Hessian)	25
3.2 Starting values for Newton–Raphson	26
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	39
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	42
3.6.9	Bootstrap	42
3.6.9.1	Usual bootstrap	43
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	Residual analysis	51
4.4.1	Response residuals	53
4.4.2	Working residuals	53
4.4.3	Pearson residuals	53
4.4.4	Partial residuals	53
4.4.5	Anscombe residuals	54
4.4.6	Deviance residuals	54

4.4.7	Adjusted deviance residuals	54
4.4.8	Likelihood residuals	55
4.4.9	Score residuals	55
4.5	Checks for systematic departure from the model	55
4.6	Model statistics	56
4.6.1	Criterion measures	56
4.6.1.1	AIC	56
4.6.1.2	BIC	58
4.6.2	The interpretation of R^2 in linear regression	59
4.6.2.1	Percentage variance explained	59
4.6.2.2	The ratio of variances	59
4.6.2.3	A transformation of the likelihood ratio	59
4.6.2.4	A transformation of the F test	60
4.6.2.5	Squared correlation	60
4.6.3	Generalizations of linear regression R^2 interpretations	60
4.6.3.1	Efron's pseudo- R^2	61
4.6.3.2	McFadden's likelihood-ratio index	61
4.6.3.3	Ben-Akiva and Lerman adjusted likelihood-ratio index	61
4.6.3.4	McKelvey and Zavoina ratio of variances	62
4.6.3.5	Transformation of likelihood ratio	62
4.6.3.6	Cragg and Uhler normed measure	62
4.6.4	More R^2 measures	63
4.6.4.1	The count R^2	63
4.6.4.2	The adjusted count R^2	63
4.6.4.3	Veall and Zimmermann R^2	63
4.6.4.4	Cameron-Windmeijer R^2	64
4.7	Marginal effects	64
4.7.1	Marginal effects for GLMs	64
4.7.2	Discrete change for GLMs	68

II	Continuous Response Models	71
5	The Gaussian family	73
5.1	Derivation of the GLM Gaussian family	74
5.2	Derivation in terms of the mean	74
5.3	IRLS GLM algorithm (nonbinomial)	77
5.4	ML estimation	79
5.5	GLM log-Gaussian models	80
5.6	Expected versus observed information matrix	81
5.7	Other Gaussian links	83
5.8	Example: Relation to OLS	83
5.9	Example: Beta-carotene	85
6	The gamma family	97
6.1	Derivation of the gamma model	98
6.2	Example: Reciprocal link	100
6.3	ML estimation	103
6.4	Log-gamma models	104
6.5	Identity-gamma models	108
6.6	Using the gamma model for survival analysis	109
7	The inverse Gaussian family	113
7.1	Derivation of the inverse Gaussian model	113
7.2	Shape of the distribution	115
7.3	The inverse Gaussian algorithm	119
7.4	Maximum likelihood algorithm	119
7.5	Example: The canonical inverse Gaussian	120
7.6	Noncanonical links	121
8	The power family and link	127
8.1	Power links	127
8.2	Example: Power link	128
8.3	The power family	129

III	Binomial Response Models	131
9	The binomial-logit family	133
9.1	Derivation of the binomial model	134
9.2	Derivation of the Bernoulli model	137
9.3	The binomial regression algorithm	138
9.4	Example: Logistic regression	140
9.4.1	Model producing logistic coefficients: The heart data	141
9.4.2	Model producing logistic odds ratios	142
9.5	GOF statistics	143
9.6	Grouped data	146
9.7	Interpretation of parameter estimates	147
10	The general binomial family	157
10.1	Noncanonical binomial models	157
10.2	Noncanonical binomial links (binary form)	159
10.3	The probit model	160
10.4	The clog-log and log-log models	164
10.5	Other links	171
10.6	Interpretation of coefficients	172
10.6.1	Identity link	173
10.6.2	Logit link	173
10.6.3	Log link	174
10.6.4	Log complement link	175
10.6.5	Log-log link	175
10.6.6	Complementary log-log link	176
10.6.7	Summary	177
10.7	Generalized binomial regression	177
10.8	Beta binomial regression	182
10.9	Zero-inflated models	184
11	The problem of overdispersion	187
11.1	Overdispersion	187

11.2	Scaling of standard errors	193
11.3	Williams' procedure	200
11.4	Robust standard errors	202
IV	Count Response Models	205
12	The Poisson family	207
12.1	Count response regression models	207
12.2	Derivation of the Poisson algorithm	208
12.3	Poisson regression: Examples	212
12.4	Example: Testing overdispersion in the Poisson model	216
12.5	Using the Poisson model for survival analysis	218
12.6	Using offsets to compare models	219
12.7	Interpretation of coefficients	222
13	The negative binomial family	225
13.1	Constant overdispersion	227
13.2	Variable overdispersion	228
13.2.1	Derivation in terms of a Poisson–gamma mixture	228
13.2.2	Derivation in terms of the negative binomial probability function	231
13.2.3	The canonical link negative binomial parameterization	233
13.3	The log-negative binomial parameterization	235
13.4	Negative binomial examples	238
13.5	The geometric family	244
13.6	Interpretation of coefficients	248
14	Other count-data models	251
14.1	Count response regression models	251
14.2	Zero-truncated models	255
14.3	Zero-inflated models	259
14.4	General truncated models	267
14.5	Hurdle models	272
14.6	Negative binomial(P) models	277

14.7	Negative binomial(Famoye)	284
14.8	Negative binomial(Waring)	285
14.9	Heterogeneous negative binomial models	286
14.10	Generalized Poisson regression models	290
14.11	Poisson inverse Gaussian models	293
14.12	Censored count response models	295
14.13	Finite mixture models	304
14.14	Quantile regression for count outcomes	309
14.15	Heaped data models	311
V	Multinomial Response Models	319
15	Unordered-response family	321
15.1	The multinomial logit model	322
15.1.1	Interpretation of coefficients: Single binary predictor	322
15.1.2	Example: Relation to logistic regression	324
15.1.3	Example: Relation to conditional logistic regression	325
15.1.4	Example: Extensions with conditional logistic regression	327
15.1.5	The independence of irrelevant alternatives	328
15.1.6	Example: Assessing the IIA	329
15.1.7	Interpreting coefficients	331
15.1.8	Example: Medical admissions—introduction	332
15.1.9	Example: Medical admissions—summary	334
15.2	The multinomial probit model	338
15.2.1	Example: A comparison of the models	340
15.2.2	Example: Comparing probit and multinomial probit	343
15.2.3	Example: Concluding remarks	347
16	The ordered-response family	349
16.1	Interpretation of coefficients: Single binary predictor	350
16.2	Ordered outcomes for general link	352
16.3	Ordered outcomes for specific links	355

16.3.1	Ordered logit	355
16.3.2	Ordered probit	356
16.3.3	Ordered clog-log	356
16.3.4	Ordered log-log	357
16.3.5	Ordered cauchit	357
16.4	Generalized ordered outcome models	358
16.5	Example: Synthetic data	359
16.6	Example: Automobile data	365
16.7	Partial proportional-odds models	372
16.8	Continuation-ratio models	377
16.9	Adjacent category model	382
VI	Extensions to the GLM	385
17	Extending the likelihood	387
17.1	The quasilielihood	387
17.2	Example: Wedderburn's leaf blotch data	388
17.3	Example: Tweedie family variance	396
17.4	Generalized additive models	399
18	Clustered data	401
18.1	Generalization from individual to clustered data	401
18.2	Pooled estimators	402
18.3	Fixed effects	404
18.3.1	Unconditional fixed-effects estimators	404
18.3.2	Conditional fixed-effects estimators	405
18.4	Random effects	407
18.4.1	Maximum likelihood estimation	407
18.4.2	Gibbs sampling	412
18.5	Mixed-effect models	413
18.6	GEEs	417
18.7	Other models	420

19	Bivariate and multivariate models	425
19.1	Bivariate and multivariate models for binary outcomes	425
19.2	Copula functions	425
19.3	Using copula functions to calculate bivariate probabilities	426
19.4	Synthetic datasets	427
19.5	Examples of bivariate count models using copula functions	430
19.6	The Famoye bivariate Poisson regression model	436
19.7	The Marshall–Olkin bivariate negative binomial regression model . .	438
19.8	The Famoye bivariate negative binomial regression model	441
20	Bayesian GLMs	447
20.1	Brief overview of Bayesian methodology	447
20.1.1	Specification and estimation	450
20.1.2	Bayesian analysis in Stata	452
20.2	Bayesian logistic regression	458
20.2.1	Bayesian logistic regression—noninformative priors	459
20.2.2	Diagnostic plots	463
20.2.3	Bayesian logistic regression—informative priors	466
20.3	Bayesian probit regression	471
20.4	Bayesian complementary log-log regression	473
20.5	Bayesian binomial logistic regression	474
20.6	Bayesian Poisson regression	477
20.6.1	Bayesian Poisson regression with noninformative priors . . .	478
20.6.2	Bayesian Poisson with informative priors	479
20.7	Bayesian negative binomial likelihood	485
20.7.1	Zero-inflated negative binomial logit	486
20.8	Bayesian normal regression	488
20.9	Writing a custom likelihood	491
20.9.1	Using the llf() option	492
20.9.1.1	Bayesian logistic regression using llf()	492

20.9.1.2	Bayesian zero-inflated negative binomial logit regression using llf()	494
20.9.2	Using the llevaluator() option	496
20.9.2.1	Logistic regression model using llevaluator()	496
20.9.2.2	Bayesian clog-log regression with llevaluator()	498
20.9.2.3	Bayesian Poisson regression with llevaluator()	499
20.9.2.4	Bayesian negative binomial regression using llevaluator()	501
20.9.2.5	Zero-inflated negative binomial logit using llevaluator()	503
20.9.2.6	Bayesian gamma regression using llevaluator()	506
20.9.2.7	Bayesian inverse Gaussian regression using llevaluator()	508
20.9.2.8	Bayesian zero-truncated Poisson using llevaluator()	511
20.9.2.9	Bayesian bivariate Poisson using llevaluator()	513
VII	Stata Software	519
21	Programs for Stata	521
21.1	The glm command	522
21.1.1	Syntax	522
21.1.2	Description	524
21.1.3	Options	524
21.2	The predict command after glm	528
21.2.1	Syntax	528
21.2.2	Options	528
21.3	User-written programs	530
21.3.1	Global macros available for user-written programs	530
21.3.2	User-written variance functions	531
21.3.3	User-written programs for link functions	533
21.3.4	User-written programs for Newey–West weights	535
21.4	Remarks	536
21.4.1	<i>Equivalent commands</i>	<i>536</i>

21.4.2	Special comments on family(Gaussian) models	536
21.4.3	Special comments on family(binomial) models	537
21.4.4	Special comments on family(nbinomial) models	537
21.4.5	Special comment on family(gamma) link(log) models	537
22	Data synthesis	539
22.1	Generating correlated data	539
22.2	Generating data from a specified population	544
22.2.1	Generating data for linear regression	544
22.2.2	Generating data for logistic regression	546
22.2.3	Generating data for probit regression	549
22.2.4	Generating data for complimentary log-log regression	550
22.2.5	Generating data for Gaussian variance and log link	551
22.2.6	Generating underdispersed count data	551
22.3	Simulation	553
22.3.1	Heteroskedasticity in linear regression	554
22.3.2	Power analysis	556
22.3.3	Comparing fit of Poisson and negative binomial	558
22.3.4	Effect of missing covariate on R^2_{Efron} in Poisson regression .	561
A	Tables	563
	References	577
	Author index	589
	Subject index	593
	Anscombe residuals versus log (variance)	102
	Inverse Gaussian ($\mu = 5, \phi = 2$)	115
	Inverse Gaussian ($\mu = 10, \phi = 2$)	116
	Inverse Gaussian ($\mu = 10, \phi = 0.4$)	116
	Inverse Gaussian ($\mu = 3, \phi = 0.1$)	117
	Inverse Gaussian ($\mu = 2, \phi = 3$)	117
	Inverse Gaussian ($\mu = 2, \phi = 10$)	118