

Contents

Preface		xi
1 Introduction		1
2 Binary Response		25
2.1 Heart Disease Example		25
2.2 Logistic Regression		28
2.3 Inference		32
2.4 Diagnostics		34
2.5 Model Selection		38
2.6 Goodness of Fit		40
2.7 Estimation Problems		44
3 Binomial and Proportion Responses		51
3.1 Binomial Regression Model		51
3.2 Inference		53
3.3 Pearson's χ^2 Statistic		54
3.4 Overdispersion		55
3.5 Quasi-Binomial		60
3.6 Beta Regression		64
4 Variations on Logistic Regression		67
4.1 Latent Variables		67
4.2 Link Functions		68
4.3 Prospective and Retrospective Sampling		71
4.4 Prediction and Effective Doses		74
4.5 Matched Case-Control Studies		75
5 Count Regression		83
5.1 Poisson Regression		83
5.2 Dispersed Poisson Model		87
5.3 Rate Models		89
5.4 Negative Binomial		92
5.5 Zero Inflated Count Models		94

6	Contingency Tables	103
6.1	Two-by-Two Tables	103
6.2	Larger Two-Way Tables	108
6.3	Correspondence Analysis	110
6.4	Matched Pairs	112
6.5	Three-Way Contingency Tables	114
6.6	Ordinal Variables	121
7	Multinomial Data	129
7.1	Multinomial Logit Model	129
7.2	Linear Discriminant Analysis	134
7.3	Hierarchical or Nested Responses	136
7.4	Ordinal Multinomial Responses	139
8	Generalized Linear Models	151
8.1	GLM Definition	151
8.2	Fitting a GLM	154
8.3	Hypothesis Tests	156
8.4	GLM Diagnostics	159
8.5	Sandwich Estimation	168
8.6	Robust Estimation	169
9	Other GLMs	175
9.1	Gamma GLM	175
9.2	Inverse Gaussian GLM	181
9.3	Joint Modeling of the Mean and Dispersion	184
9.4	Quasi-Likelihood GLM	186
9.5	Tweedie GLM	188
10	Random Effects	195
10.1	Estimation	196
10.2	Inference	201
10.3	Estimating Random Effects	207
10.4	Prediction	208
10.5	Diagnostics	210
10.6	Blocks as Random Effects	211
10.7	Split Plots	215
10.8	Nested Effects	219
10.9	Crossed Effects	222
10.10	Multilevel Models	224
11	Repeated Measures and Longitudinal Data	237
11.1	Longitudinal Data	238
11.2	Repeated Measures	243
11.3	Multiple Response Multilevel Models	248

12 Bayesian Mixed Effect Models	255
12.1 STAN	256
12.2 INLA	266
12.3 Discussion	271
13 Mixed Effect Models for Nonnormal Responses	275
13.1 Generalized Linear Mixed Models	275
13.2 Inference	275
13.3 Binary Response	277
13.4 Count Response	284
13.5 Generalized Estimating Equations	291
14 Nonparametric Regression	297
14.1 Kernel Estimators	299
14.2 Splines	302
14.3 Local Polynomials	306
14.4 Confidence Bands	307
14.5 Wavelets	308
14.6 Discussion of Methods	313
14.7 Multivariate Predictors	315
15 Additive Models	321
15.1 Modeling Ozone Concentration	323
15.2 Additive Models Using <code>mgcv</code>	324
15.3 Generalized Additive Models	330
15.4 Alternating Conditional Expectations	331
15.5 Additivity and Variance Stabilization	334
15.6 Generalized Additive Mixed Models	336
15.7 Multivariate Adaptive Regression Splines	337
16 Trees	343
16.1 Regression Trees	343
16.2 Tree Pruning	346
16.3 Random Forests	351
16.4 Classification Trees	354
16.5 Classification Using Forests	360
17 Neural Networks	365
17.1 Statistical Models as NNs	366
17.2 Feed-Forward Neural Network with One Hidden Layer	367
17.3 NN Application	368
17.4 Conclusion	371

A Likelihood Theory	375
A.1 Maximum Likelihood	375
A.2 Hypothesis Testing	378
A.3 Model Selection	381
B About R	383
Bibliography	385
Index	395