

Contents

Preface	xi
1 Introduction	1
1.1 Some examples	1
1.2 The scope of this book	14
1.3 Use of statistical software	15
Further reading	16
2 Statistical inference for binary data	17
2.1 The binomial distribution	17
2.2 Inference about the success probability	21
2.3 Comparison of two proportions	29
2.4 Comparison of two or more proportions	37
Further reading	42
3 Models for binary and binomial data	43
3.1 Statistical modelling	43
3.2 Linear models	45
3.3 Methods of estimation	48
3.4 Fitting linear models to binomial data	51
3.5 Models for binomial response data	53
3.6 The linear logistic model	56
3.7 Fitting the linear logistic model to binomial data	56
3.8 Goodness of fit of a linear logistic model	62
3.9 Comparing linear logistic models	67
3.10 Linear trends in proportions	74
3.11 Comparing stimulus–response relationships	77
3.12 Non-convergence and overfitting	81
3.13 A further example on model selection	83
3.14 Predicting a binary response probability	88
Further reading	91
4 Bioassay and some other applications	92
4.1 The tolerance distribution	92

viii Contents

4.2	Estimating an effective dose	96
4.3	Relative potency	101
4.4	Natural response	105
4.5	Non-linear logistic regression models	108
4.6	Applications of the complementary log-log model	112
	Further reading	118
5	Model checking	120
5.1	Definition of residuals	121
5.2	Checking the form of the linear predictor	126
5.3	Checking the adequacy of the link function	138
5.4	Identification of outlying observations	141
5.5	Identification of influential observations	146
5.6	Checking the assumption of a binomial distribution	160
5.7	Model checking for binary data	162
5.8	Summary and recommendations	177
5.9	A further example on the use of diagnostics	179
	Further reading	186
6	Overdispersion	188
6.1	Potential causes of overdispersion	188
6.2	Modelling variability in response probabilities	192
6.3	Modelling correlation between binary responses	194
6.4	Modelling overdispersed data	195
6.5	The special case of equal n_i	199
6.6	The beta-binomial model	204
6.7	Random effects in a linear logistic model	205
6.8	Comparison of methods	215
6.9	A further example on modelling overdispersion	216
	Further reading	221
7	Modelling data from epidemiological studies	223
7.1	Basic designs for aetiological studies	224
7.2	Measures of association between disease and exposure	227
7.3	Confounding and interaction	231
7.4	The linear logistic model for data from cohort studies	234
7.5	Interpreting the parameters in a linear logistic model	238
7.6	The linear logistic model for data from case-control studies	251
7.7	Matched case-control studies	260
7.8	A matched case-control study on sudden infant death syndrome	268
	Further reading	275
8	Some additional topics	277
8.1	Analysis of proportions and percentages	277

8.2	Analysis of rates	278
8.3	Analysis of binary data from cross-over trials	278
8.4	Random effects modelling	279
8.5	Modelling errors in the measurement of explanatory variables	280
8.6	Analysis of binary time series	281
8.7	Multivariate binary data	281
8.8	Experimental design	282
9	Computer software for modelling binary data	284
9.1	Statistical packages for modelling binary data	285
9.2	Computer-based analyses of example data sets	290
9.3	Using packages to perform some non-standard analyses	326
9.4	Summary of the relative merits of packages for modelling binary data	336
	Further reading	338
	Appendix A	340
	Appendix B	342
B.1	An algorithm for fitting a generalized linear model to binomial data	342
B.2	The likelihood function for a matched case-control study	346
	Appendix C	349
C.1	GLIM macro for computing Anscombe residuals	349
C.2	GLIM macro for constructing a half-normal plot of the standardised deviance residuals with simulated envelopes	350
C.3	GLIM macro for constructing a smoothed line	350
C.4	GLIM macro for computing the $\Delta\hat{\beta}$ -statistic	351
C.5	GLIM macro for computing the ΔD -statistic	352
C.6	GLIM macro for computing the C -statistic	353
C.7	GLIM macro for implementing the Williams procedure for modelling overdispersion	353
	References	355
	Index of Examples	362
	Index	364