

# Contents

---

<b>1</b>	<b>Introduction</b>	<b>1</b>
1.1	Some examples	1
1.2	The scope of this book	14
1.3	Use of statistical software	15
1.4	Further reading	16
<b>2</b>	<b>Statistical inference for binary data</b>	<b>19</b>
2.1	The binomial distribution	19
2.2	Inference about the success probability	23
2.3	Comparison of two proportions	31
2.4	Comparison of two or more proportions	38
2.5	Further reading	42
<b>3</b>	<b>Models for binary and binomial data</b>	<b>45</b>
3.1	Statistical modelling	45
3.2	Linear models	47
3.3	Methods of estimation	50
3.4	Fitting linear models to binomial data	53
3.5	Models for binomial response data	56
3.6	The linear logistic model	58
3.7	Fitting the linear logistic model to binomial data	59
3.8	Goodness of fit of a linear logistic model	65
3.9	Comparing linear logistic models	71
3.10	Linear trend in proportions	78
3.11	Comparing stimulus-response relationships	81
3.12	Non-convergence and overfitting	85
3.13	Some other goodness of fit statistics	87
3.14	Strategy for model selection	91
3.15	Predicting a binary response probability	98
3.16	Further reading	101
<b>4</b>	<b>Bioassay and some other applications</b>	<b>103</b>
4.1	The tolerance distribution	103
4.2	Estimating an effective dose	106
4.3	Relative potency	111
4.4	Natural response	114
4.5	Non-linear logistic regression models	118

4.6	Applications of the complementary log-log model	122
4.7	Further reading	128
<b>5</b>	<b>Model checking</b>	<b>129</b>
5.1	Definition of residuals	130
5.2	Checking the form of the linear predictor	135
5.3	Checking the adequacy of the link function	146
5.4	Identification of outlying observations	150
5.5	Identification of influential observations	154
5.6	Checking the assumption of a binomial distribution	168
5.7	Model checking for binary data	169
5.8	Summary and recommendations	185
5.9	Further reading	193
<b>6</b>	<b>Overdispersion</b>	<b>195</b>
6.1	Potential causes of overdispersion	195
6.2	Modelling variability in response probabilities	199
6.3	Modelling correlation between binary responses	201
6.4	Modelling overdispersed data	202
6.5	A model with a constant scale parameter	206
6.6	The beta-binomial model	211
6.7	Discussion	212
6.8	Further reading	213
<b>7</b>	<b>Modelling data from epidemiological studies</b>	<b>215</b>
7.1	Basic designs for aetiological studies	216
7.2	Measures of association between disease and exposure	219
7.3	Confounding and interaction	223
7.4	The linear logistic model for data from cohort studies	226
7.5	Interpreting the parameters in a linear logistic model	230
7.6	The linear logistic model for data from case-control studies	242
7.7	Matched case-control studies	250
7.8	Further reading	264
<b>8</b>	<b>Mixed models for binary data</b>	<b>269</b>
8.1	Fixed and random effects	269
8.2	Mixed models for binary data	270
8.3	Multilevel modelling	277
8.4	Mixed models for longitudinal data analysis	284
8.5	Mixed models in meta-analysis	291
8.6	Modelling overdispersion using mixed models	293
8.7	Further reading	300
<b>9</b>	<b>Exact Methods</b>	<b>303</b>
9.1	Comparison of two proportions using an exact test	303
9.2	Exact logistic regression for a single parameter	307

## CONTENTS

9.3	Exact hypothesis tests	312
9.4	Exact confidence limits for $\beta_k$	317
9.5	Exact logistic regression for a set of parameters	318
9.6	Some examples	319
9.7	Discussion	322
9.8	Further Reading	323
<b>10</b>	<b>Some additional topics</b>	<b>325</b>
10.1	Ordered categorical data	325
10.2	Analysis of proportions and percentages	329
10.3	Analysis of rates	330
10.4	Analysis of binary time series	331
10.5	Modelling errors in the measurement of explanatory variables	331
10.6	Multivariate binary data	332
10.7	Analysis of binary data from cross-over trials	333
10.8	Experimental design	333
<b>11</b>	<b>Computer software for modelling binary data</b>	<b>335</b>
11.1	Statistical packages for modelling binary data	335
11.2	Interpretation of computer output	339
11.3	Using packages to perform some non-standard analyses	341
11.4	Further reading	349
<b>Appendix A</b>	<b>Values of <math>\text{logit}(p)</math> and <math>\text{probit}(p)</math></b>	<b>351</b>
<b>Appendix B</b>	<b>Some derivations</b>	<b>353</b>
B.1	An algorithm for fitting a GLM to binomial data	353
B.2	The likelihood function for a matched case-control study	357
<b>Appendix C</b>	<b>Additional data sets</b>	<b>361</b>
C.1	Tonsil size	361
C.2	Toxicity of rotenone	361
C.3	Food poisoning	362
C.4	Analgesic potency of four compounds	362
C.5	Vasoconstriction of the fingers	363
C.6	Treatment of neuralgia	363
C.7	HIV infection	365
C.8	Aplastic anaemia	365
C.9	Cancer of the cervix	367
C.10	Endometrial cancer	367
<b>References</b>		<b>369</b>
<b>Index of examples</b>		<b>379</b>
<b>Index</b>		<b>381</b>