

Models for Discrete Data

DANIEL ZELTERMAN

Yale Cancer Center, New Haven

CLARENDON PRESS · OXFORD

1999

CONTENTS

1	Introduction	1
1.1	Issues	1
1.2	Some examples	1
2	Sampling distributions	6
2.1	Binomial and multinomial distributions	6
2.2	The Poisson distribution	9
2.3	The hypergeometric distribution	11
2.3.1	The extended hypergeometric distribution	14
2.3.2	Exact inference in a 2×2 table	16
2.3.3	Exact power and sample size estimation	19
2.3.4	The multivariate hypergeometric distribution	24
2.3.5	Distributions in higher dimensions	28
2.4	The negative binomial distribution	30
2.4.1	The negative multinomial distribution	32
	Applied exercises	35
	Theory exercises	40
3	Logistic regression	46
3.1	Three simple examples	46
3.2	Two complex examples	59
3.2.1	Nodal involvement	59
3.2.2	Job satisfaction	64
3.3	Diagnostics for logistic regression	73
	Applied exercises	80
	Theory exercises	86
4	Log-linear models	89
4.1	Models in two dimensions	90
4.2	Models in three dimensions	94
4.3	Models in higher dimensions	100
4.4	Goodness-of-fit tests	108
4.4.1	When the model fits	110
4.4.2	Sample size estimation	111
4.5	Maximum likelihood estimates	114
4.5.1	The Birch criteria	115
4.5.2	Model fitting algorithms	118
4.5.3	Explicit maximum likelihood estimates	122
	Applied exercises	127
	Theory exercises	133

5	Coordinate-free models	138
5.1	Motivating examples	139
5.2	Some linear algebra	143
5.3	The likelihood function	150
5.4	Maximum likelihood estimation	152
5.5	Examples	158
	Applied exercises	169
	Theory exercises	170
6	Additional topics	174
6.1	Longitudinal data	174
6.2	Case-control studies	179
6.3	Advanced goodness-of-fit	186
	6.3.1 Cressie and Read statistics	186
	6.3.2 Sparse data problems	189
	Applied exercises	195
	Theory exercises	201
	Appendix A: Power for chi-squared tests	205
	Appendix B: Program for exact tests	207
	Appendix C: The hypergeometric distribution	213
	References	215
	Selected solutions and hints	221
	Index	227
	Index of examples	227
	Author index	228
	Subject index	231