

Contents

Preface and Acknowledgements to Second Edition	xv
Preface and Acknowledgements	xvii
I The Nature of Spatial Epidemiology	1
1 Definitions, Terminology and Data Sets	3
1.1 Map Hypotheses and Modelling Approaches	5
1.2 Definitions and Data Examples	7
1.2.1 Case event data	7
1.2.2 Count data	8
1.3 Further Definitions	10
1.3.1 Control events and processes	10
1.3.2 Census tract information	10
1.3.3 Clustering definitions	10
1.4 Some Data Examples	11
1.4.1 Case event examples	11
1.4.2 Count data examples	19
2 Scales of Measurement and Data Availability	25
2.1 Small Scale	26
2.2 Large Scale	26
2.3 Rate Dependence	27
2.4 Data Quality and the Ecological Fallacy	27
2.5 Edge Effects	28
3 Geographical Representation and Mapping	31
3.1 Introduction and Definitions	31
3.2 Maps and Mapping	32
3.2.1 Statistical maps and mapping	34
3.2.2 Object process mapping	34
3.2.3 Geostatistical mapping	36

- 3.3 Statistical Accuracy 37
- 3.4 Aggregation 37
- 3.5 Mapping Issues Related to Aggregated Data 37
- 3.6 Conclusions 39

- 4 Basic Models 41**

 - 4.1 Sampling Considerations 41
 - 4.2 Likelihood-Based and Bayesian Approaches 42
 - 4.3 Point Event Models 42
 - 4.3.1 Point process models and applications 43
 - 4.3.2 The basic Poisson process model 44
 - 4.3.3 Hybrid models and regionalisation 49
 - 4.3.4 Bayesian models and random effects 50
 - 4.3.5 MAP estimation, empirical Bayes and full Bayesian analysis 52
 - 4.3.6 Bivariate/multivariate models 53
 - 4.3.7 Hidden structure and mixture models 56
 - 4.3.8 Space-time extensions 56
 - 4.4 Count Models 58
 - 4.4.1 Standard models 60
 - 4.4.2 Approximations 63
 - 4.4.3 Random-effect extensions 63
 - 4.4.4 Hidden structure and mixture models 64
 - 4.4.5 Space-time extensions 65

- 5 Exploratory Approaches, Parametric Estimation and Inference 67**

 - 5.1 Exploratory Methods 68
 - 5.1.1 Cartographic issues 69
 - 5.1.2 Case event mapping 71
 - 5.1.3 Count mapping 75
 - 5.2 Parameter Estimation 80
 - 5.2.1 Case event likelihood models 80
 - 5.2.2 Count event likelihood models 85
 - 5.2.3 Approximations 87
 - 5.2.4 Bayesian models 88
 - 5.3 Residual Diagnostics 96
 - 5.4 Hypothesis Testing 98
 - 5.5 Edge Effects 99
 - 5.5.1 Edge effects in case events 101
 - 5.5.2 Edge effects in counts 101
 - 5.5.3 Edge weighting schemes and MCMC methods 102
 - 5.5.4 Discussion 104
 - 5.5.5 The Tuscan example 105

II Important Problems in Spatial Epidemiology 109

6 Small Scale: Disease Clustering 111

- 6.1 Definition of Clusters and Clustering 112
- 6.2 Modelling Issues 115
- 6.3 Hypothesis Tests for Clustering 118
 - 6.3.1 General non-specific clustering 118
 - 6.3.2 Specific clustering 121
- 6.4 Space-Time Clustering 123
 - 6.4.1 Modelling issues 123
 - 6.4.2 Hypothesis testing 126
- 6.5 Clustering Examples 127
 - 6.5.1 Humberside example 127
 - 6.5.2 Larynx cancer example 131
 - 6.5.3 Count data clustering example 133
 - 6.5.4 Space-time clustering examples 136
- 6.6 Other Methods Related to Clustering 138
 - 6.6.1 Wombling 140

7 Small Scale: Putative Sources of Hazard 143

- 7.1 Introduction 143
- 7.2 Study Design 144
 - 7.2.1 Retrospective and prospective studies 144
 - 7.2.2 Study region design 145
 - 7.2.3 Replication and controls 146
- 7.3 Problems of Inference 147
 - 7.3.1 Exploratory techniques 148
- 7.4 Modelling the Hazard Exposure Risk 153
- 7.5 Models for Case Event Data 162
 - 7.5.1 Estimation 164
 - 7.5.2 Hypothesis tests 164
 - 7.5.3 Diagnostic techniques 166
- 7.6 A Case Event Example 167
- 7.7 Models for Count Data 169
 - 7.7.1 Estimation 171
 - 7.7.2 Hypothesis tests 171
- 7.8 A Count Data Example 172
- 7.9 Other Directions 174
 - 7.9.1 Multiple disease analysis 174
 - 7.9.2 Space-time modelling 184
 - 7.9.3 Space-time exploratory analysis 184
 - 7.9.4 Space-time Bayesian analysis 185

8 Large Scale: Disease Mapping 189

- 8.1 Introduction 189

- 8.2 Simple Statistical Representation 189
 - 8.2.1 Crude rates 190
 - 8.2.2 Standardised mortality/morbidity ratios, standardisation and relative risk surfaces 191
 - 8.2.3 Interpolation 193
 - 8.2.4 Exploratory mapping methods 193
- 8.3 Basic Models 194
 - 8.3.1 Likelihood models 194
 - 8.3.2 Random effects and Bayesian models 197
- 8.4 Advanced Methods 201
 - 8.4.1 Non-parametric methods 202
 - 8.4.2 Incorporating spatially correlated heterogeneity 203
 - 8.4.3 Case event modelling 206
- 8.5 Model Variants and Extensions 209
 - 8.5.1 Semiparametric modelling 209
 - 8.5.2 Geographically weighted regression 210
 - 8.5.3 Mixture models 211
- 8.6 Approximate Methods 212
- 8.7 Multivariate Methods 213
- 8.8 Evaluation of Model Performance 216
- 8.9 Hypothesis Testing in Disease Mapping 219
 - 8.9.1 First-order effects 219
 - 8.9.2 Second-order and variance effects 221
- 8.10 Space-Time Disease Mapping 222
- 8.11 Spatial Survival and Longitudinal Data 229
 - 8.11.1 Spatial survival analysis 229
 - 8.11.2 Spatial longitudinal analysis 231
 - 8.11.3 Spatial multiple event modelling 232
- 8.12 Disease Mapping: Case Studies 232
 - 8.12.1 Eastern Germany 232
 - 8.12.2 Ohio respiratory cancer 239
- 9 Ecological Analysis and Scale Change 247**
 - 9.1 Ecological Analysis: Introduction 247
 - 9.2 Small-Scale Modelling Issues 252
 - 9.2.1 Hypothesis tests 253
 - 9.2.2 Ecological aggregation effects 253
 - 9.3 Changes of Scale and MAUP 255
 - 9.3.1 MAUP: the *modifiable areal unit problem* 255
 - 9.3.2 Large-scale issues 260
 - 9.4 A Simple Example: Sudden Infant Death in North Carolina 261
 - 9.5 A Case Study: Malaria and IDDM 263
- 10 Infectious Disease Modelling 269**
 - 10.1 Introduction 269

- 10.2 General Model Development 270
- 10.3 Spatial Model Development 273
 - 10.3.1 Count data 273
 - 10.3.2 Individual-level data 278
- 10.4 Modelling Special Cases for Individual-Level Data 280
 - 10.4.1 Proportional hazards interpretation 280
 - 10.4.2 Subgroup modifications 281
 - 10.4.3 Cluster function specification 282
- 10.5 Survival Analysis with Spatial Dependence 283
- 10.6 Individual-Level Data Example 284
 - 10.6.1 Distribution of susceptibles $S(x, t)$ 285
 - 10.6.2 The spatial distance function h 285
 - 10.6.3 The function g 285
 - 10.6.4 Fitting the model 286
 - 10.6.5 Revised model 287
- 10.7 Underascertainment and Censoring 288
- 10.8 Conclusions 289

11 Large Scale: Surveillance 293

- 11.1 Process Control Methodology 294
- 11.2 Spatio-Temporal Modelling 295
- 11.3 S-T Monitoring 297
 - 11.3.1 Fixed spatial and temporal frame 297
 - 11.3.2 Fixed spatial frame and dynamic temporal frame 301
- 11.4 Syndromic Surveillance 304
- 11.5 Multivariate–Multifocus Surveillance 305
- 11.6 Bayesian Approaches 308
 - 11.6.1 Bayesian alarm functions, Bayes factors and syndromic analyses 308
- 11.7 Computational Considerations 310
- 11.8 Infectious Diseases 311
- 11.9 Conclusions 312

Appendix A Monte Carlo Testing, Parametric Bootstrap and Simulation

- Envelopes 313**
- A.1 Nuisance Parameters and Test Statistics 313
- A.2 Monte Carlo Tests 314
- A.3 Null Hypothesis Simulation 315
 - A.3.1 Spatial case 316
 - A.3.2 Spatio-temporal case 318
- A.4 Parametric Bootstrap 319
 - A.4.1 Bayesian spatial models 322
 - A.4.2 Spatio-temporal case 323
- A.5 Simulation Envelopes 324

Appendix B Markov Chain Monte Carlo Methods 325

- B.1 Definitions 325
- B.2 Metropolis and Metropolis–Hastings Algorithms 326
 - B.2.1 Metropolis algorithm 326
 - B.2.2 Metropolis–Hastings extension 327
 - B.2.3 The Gibbs sampler 327
 - B.2.4 M–H versus Gibbs algorithms 328
 - B.2.5 Examples 328

Appendix C Algorithms and Code 331

- C.1 Data Exploration 331
- C.2 Likelihood and Bayesian Models 335
- C.3 Likelihood Models 336
 - C.3.1 Case event data 336
 - C.3.2 Count data 340
- C.4 Bayesian Hierarchical Models 341
 - C.4.1 Case event data 341
 - C.4.2 Count data 344
- C.5 Space-Time Analysis 346
 - C.5.1 Data exploration 346
 - C.5.2 Likelihood models 349
 - C.5.3 Bayesian models 351
 - C.5.4 Infectious disease models 357

Appendix D Glossary of Estimators 359

- D.1 Case Event Estimators 359
- D.2 Tract Count Estimators 361

Appendix E Software 363

- E.1 Software 363
 - E.1.1 Spatial statistical tools 363
 - E.1.2 Geographical information systems 365

Bibliography 367

Index 389