Contents

Preface *xiii*
About this book *xv*

1 Data, statistics and geography 1
Chapter overview 1
Learning objectives 2
1.1 Statistics: a brief introduction 2
1.2 Why you should study statistics 3
1.3 Types of statistic 6
1.4 Analysis and error 8
1.5 Geographical data and analysis 9
1.6 Some problems when analysing geographical data 10
Key points 15
References 16

2 Descriptive statistics 17
Chapter overview 17
Learning objectives 18
2.1 Data and variables 18
2.2 Simple ways to make sense of and present data 21
2.3 Some useful notation 24
2.4 A second data set 25
2.5 Measures of central tendency 30
2.6 Some measures of spread and variation 34
2.7 Presenting the centre and spread of data 38
2.8 Types of numeric data 42
2.9 Conclusion 46
Key points 46
References 46

3 The normal curve 48
Chapter overview 48
Learning objectives 49
3.1 Introducing the normal curve 49
3.2 Models, error and uncertainty 53
3.3 Why are the data (approximately) normal? 55
3.4 Important properties of a normal curve 57
3.5 An experiment 58
3.6 Probability and the normal curve 60

3.7 Worked examples 62
3.8 *z* values and the standard normal curve 65
3.9 Signs of non-normality 70
3.10 Moments of a distribution 75
3.11 The quantile plot 76
3.12 Conclusion 80
Key points 80
References 81

4 Sampling 82
Chapter overview 82
Learning objectives 83
4.1 Introduction 83
4.2 The process of sampling, phase 1: scope and scale 86
4.3 The process of sampling, phase 1: issues 89
4.4 The process of sampling, phase 2: sampling method 94
4.5 Sampling methods: issues and practicalities 108
Key points 111
References 112

5 From description to inference 114
Chapter overview 114
Learning objectives 115
5.1 About inference 115
5.2 A measure of unreliability 116
5.3 What is a population? 117
5.4 A geographical example 118
5.5 A thought experiment 120
5.6 Why the thought experiment is useful 122
5.7 Calculating the standard error 123
5.8 Confidence intervals 124
5.9 Consolidation and worked examples 125
5.10 Finding a 99.9% or any other confidence interval 127
5.11 Calculating a confidence interval for a small sample 128
5.12 Conclusion 132
Key points 133
References 133

6 Hypothesis testing 134
Chapter overview 134
Learning objectives 135
6.1 Detecting difference 135
6.2 Hypothesis testing and the one-sample *t* test 137
6.3 One-tailed tests of difference 142
6.4 Power, error and research design 143

6.5 The two-sample *t* test and the *F* test 147
6.6 Additional notes about the *t* test 152
6.7 Analysis of variance 153
6.8 Non-parametric tests 157
6.9 Conclusion: from detection to explanation 159
Key points 160
References 161

7 Relationships and explanations 162
Chapter overview 162
Learning objectives 163
7.1 Looking for relationships 163
7.2 Using scatter plots 164
7.3 Independent and dependent variables 165
7.4 Correlation 167
7.5 Complications with the correlation coefficient 170
7.6 Bivariate (two-variable) regression 172
7.7 Interpreting regression analysis 176
7.8 Assumptions of regression analysis 180
7.9 Checking the residuals 180
7.10 What to do if the residuals do not meet the regression assumptions 184
7.11 Multiple regression 187
7.12 Interpreting multiple regression 188
7.13 Assumptions of multiple regression 190
7.14 Partial regression plots 191
7.15 A strategy for multiple regression 192
7.16 The strength of the effects and the problem of 'too much power' 193
7.17 Using the *F* test to compare regression models 194
7.18 Uses of regression 195
7.19 Other types of regression model 196
Key points 196
References 197

8 Detecting and managing spatial dependency 199
Chapter overview 199
Learning objectives 200
8.1 Introduction 200
8.2 Why does spatial dependency matter in the context of quantitative geographical analyses? 201
8.3 Looking for spatial autocorrelation 207
8.4 Global measures of spatial autocorrelation 209
8.5 Other measures of global autocorrelation 215
8.6 Conclusion 222
Key points 222
References 222

9 Exploring spatial relationships 225
Chapter overview 225
Learning objectives 226
9.1 Introduction 226
9.2 Mapping with cartograms 228
9.3 Spatial analysis and spatial autocorrelation 230
9.4 Who's my neighbour? Defining a weights matrix 232
9.5 Spatial error model 234
9.6 Spatial lagged *y* model 236
9.7 Geographically weighted regression (GWR) 237
9.8 Local indicator of spatial association (LISA) 241
9.9 Modelling at multiple scales 242
9.10 Geography, computation and statistics 246
Key points 247
References 248

Epilogue *249*
Appendix *251*
Index *254*