Contents

Preface xi

1. Planning **1**
1.1 The scientific method 1
1.2 Types of experiment/project 3
1.3 Getting data – using a spreadsheet 4
1.4 Hypothesis testing 4
1.5 Data types 5
1.6 Sampling effort 8
1.7 Tools of the trade 13
1.8 The R program 13
1.9 Excel 15

2. Data recording **19**
2.1 Collecting data – who, what, where, when 19
2.2 How to arrange data 21

3. Beginning data exploration – using software tools **26**
3.1 Beginning to use R 26
3.2 Manipulating data in a spreadsheet 34
3.3 Getting data from Excel into R 50

4. Exploring data – looking at numbers **54**
4.1 Summarizing data 55
4.2 Distribution 63
4.3 A numerical value for the distribution 73
4.4 Statistical tests for normal distribution 82
4.5 Distribution type 83
4.6 Transforming data 88
4.7 When to stop collecting data? The running average 91
4.8 Statistical symbols 96

5. Exploring data – which test is right? **100**
5.1 Types of project 100
5.2 Hypothesis testing 101
5.3 Choosing the correct test 102

6. Exploring data – using graphs **107**
6.1 Introduction to data visualization 107
6.2 Exploratory graphs 117
6.3 Graphs to illustrate differences 123
6.4 Graphs to illustrate correlation and regression 148
6.5 Graphs to illustrate association 162
6.6 Graphs to illustrate similarity 176
6.7 Graphs – a summary 178

7. Tests for differences **182**
7.1 Differences: *t*-test 182
7.2 Differences: *U*-test 191
7.3 Paired tests 197

8. Tests for linking data – correlations **206**
8.1 Correlation: Spearman's rank test 207
8.2 Pearson's product moment 213
8.3 Correlation tests using Excel 216
8.4 Correlation tests using R 222
8.5 Curved linear correlation 226

9. Tests for linking data – associations **230**
9.1 Association: chi-squared test 230
9.2 Goodness of fit test 236
9.3 Using R for chi-squared tests 237
9.4 Using Excel for chi-squared tests 241

10. Differences between more than two samples **247**
10.1 Analysis of variance 248
10.2 Kruskal–Wallis test 272

11. Tests for linking several factors **285**
11.1 Multiple regression 285
11.2 Curved-linear regression 304
11.3 Logistic regression 314

12. Community ecology **329**
12.1 Diversity 329
12.2 Similarity 341

13. Reporting results **359**
13.1 Presenting findings 359
13.2 Publishing 360
13.3 Reporting results of statistical analyses 360
13.4 Graphs 362
13.5 Writing papers 365
13.6 Plagiarism 367
13.7 References 368
13.8 Poster presentations 369
13.9 Giving a talk (PowerPoint) 370

14. Summary **373**

Glossary 375
Appendices 381
Index 394