Contents

Preface to the third edition *page* xv

Preface to the second edition xvii

Preface to the first edition xix

1 Getting started 1

1.1 What is statistical programming? 1
1.2 Outline of this book 2
1.3 The R package 3
1.4 Why use a command line? 3
1.5 Font conventions 4
1.6 Installation of R and RStudio 4
1.7 Getting started in RStudio 5
1.8 Going further 6

2 Introduction to the R language 7

2.1 First steps 7
2.2 Basic features of R 12
2.3 Vectors in R 13
2.4 Data storage in R 22
2.5 Packages, libraries, and repositories 27
2.6 Getting help 29
2.7 Useful R features 32
2.8 Logical vectors and relational operators 37
2.9 Data frames, tibbles, and lists 40
2.10 Data input and output 46

3 Programming statistical graphics 53

3.1 Simple high level plots 54
3.2 Choosing a high level graphic 67
3.3 Low level graphics functions 68
3.4 Graphics as a language: `ggplot2` 70
3.5 Other graphics systems 87

4 Programming with R 93

4.1 Flow control 93
4.2 Managing complexity through functions 108
4.3 The `replicate()` function 114
4.4 Miscellaneous programming tips 115
4.5 Some general programming guidelines 118
4.6 Debugging and maintenance 125
4.7 Efficient programming 132

5 Complex programming in the *tidyverse* 139

5.1 The *tidyverse* principles 140
5.2 The `tibble` package: a data frame improvement 141
5.3 The `readr` package: reading data in the *tidyverse* 143
5.4 The `stringr` package for manipulating strings 144
5.5 The `dplyr` package for manipulating data sets 146
5.6 Other *tidyverse* packages 149

6 Simulation 150

6.1 Monte Carlo simulation 150
6.2 Generation of pseudorandom numbers 151
6.3 Simulation of other random variables 156
6.4 Multivariate random number generation 173
6.5 Markov chain simulation 175
6.6 Monte Carlo integration 177
6.7 Advanced simulation methods 179

7 Computational linear algebra 197

7.1 Vectors and matrices in R 198
7.2 Matrix multiplication and inversion 205
7.3 Eigenvalues and eigenvectors 210
7.4 Other matrix decompositions 211
7.5 Other matrix operations 218

8 Numerical optimization 222
8.1 The golden section search method 222
8.2 Newton–Raphson 225
8.3 The Nelder–Mead simplex method 227
8.4 Built-in functions 231
8.5 Linear programming 231

Appendix A Review of random variables and distributions 248

Appendix B Base graphics details 251
B.1 The plotting region and margins 251
B.2 Adjusting axis tick labels 252
B.3 Setting graphical parameters 255

Index 257