CONTENTS

PREFACE TO THE SECOND EDITION xiii
PREFACE TO THE FIRST EDITION xvii
ACKNOWLEDGMENTS xxi
ABOUT THE COMPANION WEBSITE xxiii

I THE R LANGUAGE 1

1 Basics of R 3

1.1 What Is R? 3
1.2 Installing R 4
1.3 R Documentation 4
1.4 Basics 5
1.5 Getting Help 6
1.6 Data Entry 7
1.7 Missing Values 11
1.8 Editing 12
1.9 Tidying Up 12
1.10 Saving and Retrieving 13
1.11 Packages 13
1.12 Interfaces 14
1.13 Project 16

vii
CONTENTS

2 Summarizing Statistical Data
  2.1 Measures of Central Tendency 17
  2.2 Measures of Dispersion 21
  2.3 Overall Summary Statistics 24
  2.4 Programming in R 25
  2.5 Project 30

3 Graphical Displays
  3.1 Boxplots 31
  3.2 Histograms 36
  3.3 Stem and Leaf 40
  3.4 Scatter Plots 40
  3.5 The Line of Best Fit 43
  3.6 Machine Learning and the Line of Best Fit 44
  3.7 Graphical Displays Versus Summary Statistics 49
  3.8 Projects 53

II FUNDAMENTALS OF PROBABILITY

4 Probability Basics
  4.1 Experiments, Sample Spaces, and Events 58
  4.2 Classical Approach to Probability 61
  4.3 Permutations and Combinations 64
  4.4 The Birthday Problem 71
  4.5 Balls and Bins 76
  4.6 R Functions for Allocation 79
  4.7 Allocation Overload 81
  4.8 Relative Frequency Approach to Probability 83
  4.9 Simulating Probabilities 84
  4.10 Projects 89

5 Rules of Probability
  5.1 Probability and Sets 91
  5.2 Mutually Exclusive Events 92
  5.3 Complementary Events 93
  5.4 Axioms of Probability 94
  5.5 Properties of Probability 96

6 Conditional Probability
  6.1 Multiplication Law of Probability 107
  6.2 Independent Events 108
10.6 Amnesia 217
10.7 Simulating Markov 219
10.8 Projects 224

11 The Binomial Distribution 226
11.1 Binomial Probabilities 227
11.2 Binomial Random Variables 229
11.3 Cumulative Distribution Function 233
11.4 The Quantile Function 235
11.5 Reliability: The General System 238
11.6 Machine Learning 241
11.7 Binomial Expectations 245
11.8 Simulating Binomial Probabilities and Expectations 248
11.9 Projects 254

12 The Hypergeometric Distribution 255
12.1 Hypergeometric Random Variables 257
12.2 Cumulative Distribution Function 260
12.3 The Lottery 262
12.4 Hypergeometric or Binomial? 266
12.5 Projects 273

13 The Poisson Distribution 274
13.1 Death by Horse Kick 274
13.2 Limiting Binomial Distribution 275
13.3 Random Events in Time and Space 281
13.4 Probability Density Function 283
13.5 Cumulative Distribution Function 287
13.6 The Quantile Function 289
13.7 Estimating Software Reliability 290
13.8 Modeling Defects in Integrated Circuits 292
13.9 Simulating Poisson Probabilities 293
13.10 Projects 298

14 Sampling Inspection Schemes 299
14.1 Introduction 299
14.2 Single Sampling Inspection Schemes 300
14.3 Acceptance Probabilities 301
14.4 Simulating Sampling Inspection Schemes 303
14.5 Operating Characteristic Curve 308
14.6 Producer's and Consumer's Risks 310
14.7 Design of Sampling Schemes 311
CONTENTS

14.8 Rectifying Sampling Inspection Schemes 315
14.9 Average Outgoing Quality 316
14.10 Double Sampling Inspection Schemes 318
14.11 Average Sample Size 319
14.12 Single Versus Double Schemes 320
14.13 Projects 324

IV CONTINUOUS DISTRIBUTIONS 325

15 Introduction to Continuous Distributions 327
  15.1 Introduction to Continuous Random Variables 328
  15.2 Probability Density Function 328
  15.3 Cumulative Distribution Function 331
  15.4 The Uniform Distribution 332
  15.5 Expectation of a Continuous Random Variable 336
  15.6 Simulating Continuous Variables 338

16 The Exponential Distribution 341
  16.1 Modeling Waiting Times 341
  16.2 Probability Density Function of Waiting Times 342
  16.3 Cumulative Distribution Function 344
  16.4 Modeling Lifetimes 347
  16.5 Quantiles 349
  16.6 Exponential Expectations 351
  16.7 Simulating Exponential Probabilities and Expectations 353
  16.8 Amnesia 356
  16.9 Simulating Markov 360
  16.10 Project 369

17 Queues 370
  17.1 The Single Server Queue 370
  17.2 Traffic Intensity 371
  17.3 Queue Length 372
  17.4 Average Response Time 376
  17.5 Extensions of the M/M/1 Queue 378
  17.6 Project 382

18 The Normal Distribution 383
  18.1 The Normal Probability Density Function 385
  18.2 The Cumulative Distribution Function 387
18.3 Quantiles
18.4 The Standard Normal Distribution
18.5 Achieving Normality: Limiting Distributions
18.6 Projects

19 Process Control
19.1 Control Charts
19.2 Cusum Charts
19.3 Charts for Defective Rates
19.4 Project

V TAILING OFF

20 The Inequalities of Markov and Chebyshev
20.1 Markov’s Inequality
20.2 Algorithm Runtime
20.3 Chebyshev’s Inequality

Appendix A: Data: Examination Results
Appendix B: The Line of Best Fit: Coefficient Derivations
Appendix C: Variance Derivations
Appendix D: Binomial Approximation to the Hypergeometric
Appendix E: Normal Tables
Appendix F: The Inequalities of Markov and Chebyshev

INDEX TO R COMMANDS
INDEX
POSTFACE