
Contents

Preface	vii
---------------	-----

Part I Preliminaries

Introduction to Part I	3
Mathematical and Statistical Preliminaries	5
1.1 Discovering Structure in Data	6
1.2 Mathematical Tools for Identifying Structure in Data	10
1.3 Data-Generating Processes; Probability Distributions	29
1.4 Statistical Inference	37
1.5 Probability Statements in Statistical Inference	52
1.6 Modeling and Computational Inference	56
1.7 The Role of the Empirical Cumulative Distribution Function ..	59
1.8 The Role of Optimization in Inference	65
Notes and Further Reading	74
Exercises	75

Part II Statistical Computing

Introduction to Part II	83
Computer Storage and Arithmetic	85
2.1 The Fixed-Point Number System	86
2.2 The Floating-Point Number System	88
2.3 Errors	97
Notes and Further Reading	101
Exercises	102

Algorithms and Programming	107
3.1 Error in Numerical Computations	109
3.2 Algorithms and Data	113
3.3 Efficiency	116
3.4 Iterations and Convergence	128
3.5 Programming	134
3.6 Computational Feasibility	137
Notes and Further Reading	138
Exercises	142
Approximation of Functions and Numerical Quadrature	147
4.1 Function Approximation and Smoothing	153
4.2 Basis Sets in Function Spaces	160
4.3 Orthogonal Polynomials	167
4.4 Splines	178
4.5 Kernel Methods	182
4.6 Numerical Quadrature	184
4.7 Monte Carlo Methods for Quadrature	192
Notes and Further Reading	197
Exercises	199
Numerical Linear Algebra	203
5.1 General Computational Considerations for Vectors and Matrices	205
5.2 Gaussian Elimination and Elementary Operator Matrices	209
5.3 Matrix Decompositions	215
5.4 Iterative Methods	221
5.5 Updating a Solution to a Consistent System	227
5.6 Overdetermined Systems; Least Squares	228
5.7 Other Computations with Matrices	235
Notes and Further Reading	236
Exercises	237
Solution of Nonlinear Equations and Optimization	241
6.1 Finding Roots of Equations	244
6.2 Unconstrained Descent Methods in Dense Domains	261
6.3 Unconstrained Combinatorial and Stochastic Optimization	275
6.4 Optimization under Constraints	284
6.5 Computations for Least Squares	291
6.6 Computations for Maximum Likelihood	294
Notes and Further Reading	298
Exercises	301

Generation of Random Numbers 305

- 7.1 Randomness of Pseudorandom Numbers 305
- 7.2 Generation of Nonuniform Random Numbers 307
- 7.3 Acceptance/Rejection Method Using a Markov Chain 313
- 7.4 Generation of Multivariate Random Variates 315
- 7.5 Data-Based Random Number Generation 318
- 7.6 Software for Random Number Generation 320
- Notes and Further Reading 329
- Exercises 329

Part III Methods of Computational Statistics

Introduction to Part III 335

Graphical Methods in Computational Statistics 337

- 8.1 Smoothing and Drawing Lines 341
- 8.2 Viewing One, Two, or Three Variables 344
- 8.3 Viewing Multivariate Data 355
- Notes and Further Reading 365
- Exercises 368

Tools for Identification of Structure in Data 371

- 9.1 Transformations 373
- 9.2 Measures of Similarity and Dissimilarity 383
- Notes and Further Reading 397
- Exercises 397

Estimation of Functions 401

- 10.1 General Approaches to Function Estimation 403
- 10.2 Pointwise Properties of Function Estimators 407
- 10.3 Global Properties of Estimators of Functions 410
- Notes and Further Reading 414
- Exercises 414

Monte Carlo Methods for Statistical Inference 417

- 11.1 Monte Carlo Estimation 418
- 11.2 Simulation of Data from a Hypothesized Model: Monte Carlo Tests 422
- 11.3 Simulation of Data from a Fitted Model: “Parametric Bootstraps” 424
- 11.4 Random Sampling from Data 424
- 11.5 Reducing Variance in Monte Carlo Methods 425
- 11.6 Software for Monte Carlo 429
- Notes and Further Reading 430
- Exercises 431

Data Randomization, Partitioning, and Augmentation	435
12.1 Randomization Methods	436
12.2 Cross Validation for Smoothing and Fitting	440
12.3 Jackknife Methods	442
Notes and Further Reading	448
Exercises	449
Bootstrap Methods	453
13.1 Bootstrap Bias Corrections	454
13.2 Bootstrap Estimation of Variance	456
13.3 Bootstrap Confidence Intervals	457
13.4 Bootstrapping Data with Dependencies	461
13.5 Variance Reduction in Monte Carlo Bootstrap	462
Notes and Further Reading	464
Exercises	465
<hr/>	
Part IV Exploring Data Density and Relationships	
<hr/>	
Introduction to Part IV	471
Estimation of Probability Density Functions Using Parametric Models	475
14.1 Fitting a Parametric Probability Distribution	476
14.2 General Families of Probability Distributions	477
14.3 Mixtures of Parametric Families	480
14.4 Statistical Properties of Density Estimators Based on Parametric Families	482
Notes and Further Reading	483
Exercises	484
Nonparametric Estimation of Probability Density Functions ...	487
15.1 The Likelihood Function	487
15.2 Histogram Estimators	490
15.3 Kernel Estimators	499
15.4 Choice of Window Widths	504
15.5 Orthogonal Series Estimators	505
15.6 Other Methods of Density Estimation	506
Notes and Further Reading	509
Exercises	510
Statistical Learning and Data Mining	515
16.1 Clustering and Classification	519
16.2 Ordering and Ranking Multivariate Data	538
16.3 Linear Principal Components	548
16.4 Variants of Principal Components	560

16.5 Projection Pursuit	564
16.6 Other Methods for Identifying Structure.....	572
16.7 Higher Dimensions	573
Notes and Further Reading.....	578
Exercises	580
Statistical Models of Dependencies	585
17.1 Regression and Classification Models.....	588
17.2 Probability Distributions in Models	597
17.3 Fitting Models to Data.....	600
17.4 Classification	620
17.5 Transformations	628
Notes and Further Reading.....	634
Exercises	636
<hr/>	
Appendices	
<hr/>	
Monte Carlo Studies in Statistics	643
A.1 Simulation as an Experiment.....	644
A.2 Reporting Simulation Experiments.....	645
A.3 An Example	646
A.4 Computer Experiments.....	653
Exercises	655
Some Important Probability Distributions	657
Notation and Definitions	663
C.1 General Notation	663
C.2 Computer Number Systems.....	665
C.3 Notation Relating to Random Variables	666
C.4 General Mathematical Functions and Operators	668
C.5 Models and Data	675
Solutions and Hints for Selected Exercises	677
Bibliography.....	689
E.1 Literature in Computational Statistics	690
E.2 References for Software Packages	693
E.3 References to the Literature	693
Index	715