Numerical Methods of Statistics
Second Edition

This book explains how computer software is designed to perform the tasks required for sophisticated statistical analysis. For statisticians, it examines the nitty-gritty computational problems behind statistical methods. For mathematicians and computer scientists, it looks at the application of mathematical tools to statistical problems. The first half of the book offers a basic background in numerical analysis that emphasizes issues important to statisticians. The next several chapters cover a broad array of statistical tools, such as maximum likelihood and nonlinear regression. The author also treats the application of numerical tools; numerical integration and random number generation are explained in a unified manner reflecting complementary views of Monte Carlo methods. Each chapter contains exercises that range from simple questions to research problems. Most of the examples are accompanied by demonstration and source code available on the author’s Web site. New in this second edition are demonstrations coded in R, as well as new sections on linear programming and the Nelder-Mead search algorithm.

John F. Monahan is a Professor of Statistics at North Carolina State University, where he joined the faculty in 1978 and has been a professor since 1990. His research has appeared in numerous computational as well as statistical journals. He is also the author of A Primer on Linear Models (2008).
CAMBRIDGE SERIES IN STATISTICAL AND PROBABILISTIC MATHEMATICS

Editorial Board
Z. Ghahramani, Department of Engineering, University of Cambridge
R. Gill, Department of Mathematics, Utrecht University
F. Kelly, Statistics Laboratory, University of Cambridge
B. D. Ripley, Department of Statistics, University of Oxford
S. Ross, Epstein Department of Industrial & Systems Engineering, University of Southern California
M. Stein, Department of Statistics, University of Chicago

This series of high-quality upper-division textbooks and expository monographs covers all aspects of stochastic applicable mathematics. The topics range from pure and applied statistics to probability theory, operations research, optimization, and mathematical programming. The books contain clear presentations of new developments in the field and also of the state of the art in classical methods. While emphasizing rigorous treatment of theoretical methods, the books also contain applications and discussions of new techniques made possible by advances in computational practice.

1. Bootstrap Methods and Their Application, by A. C. Davison and D. V. Hinkley
2. Markov Chains, by J. Norris
3. Asymptotic Statistics, by A. W. van der Vaart
4. Wavelet Methods for Time Series Analysis, by Donald B. Percival and Andrew T. Walden
5. Bayesian Methods, by Thomas Leonard and John S. J. Hsu
9. The Estimation and Tracking of Frequency, by B. G. Quinn and E. J. Hannan
11. Statistical Models, by A. C. Davison
13. Exercise in Probability, by Loïc Chaumont and Marc Yor
15. Measure Theory and Filtering, by Lakhdar Aggoun and Robert Elliott
17. Elements of Distribution Theory, by Thomas A. Severini
19. The Coordinate-Free Approach to Linear Models, by Michael J. Wichura
20. Random Graph Dynamics, by Rick Durrett
21. Networks, by Peter Whittle
22. Saddlepoint Approximations with Applications, by Ronald W. Butler
23. Applied Asymptotics, by A. R. Brazzale, A. C. Davison, and N. Reid
24. Random Networks for Communication, by Massimo Franceschetti and Ronald Meester
25. Design of Comparative Experiments, by R. A. Bailey
26. Symmetry Studies, by Marlos A. G. Viana
27. Model Selection and Model Averaging, by Gerda Claeskens and Nils Lid Hjort
28. Bayesian Nonparametrics, by Nils Lid Hjort, Peter Müller, Stephen G. Walker
30. Brownian Motion, by Peter Mörters and Yuval Peres
Numerical Methods of Statistics

Second Edition

JOHN F. MONAHAN
North Carolina State University
## Contents

<table>
<thead>
<tr>
<th>Section</th>
<th>Pages</th>
</tr>
</thead>
<tbody>
<tr>
<td>Preface to the Second Edition</td>
<td>xiii</td>
</tr>
<tr>
<td>Preface to the First Edition</td>
<td>xv</td>
</tr>
<tr>
<td><strong>1 Algorithms and Computers</strong></td>
<td>1</td>
</tr>
<tr>
<td>1.1 Introduction</td>
<td>1</td>
</tr>
<tr>
<td>1.2 Computers</td>
<td>3</td>
</tr>
<tr>
<td>1.3 Software and Computer Languages</td>
<td>5</td>
</tr>
<tr>
<td>1.4 Data Structures</td>
<td>8</td>
</tr>
<tr>
<td>1.5 Programming Practice</td>
<td>9</td>
</tr>
<tr>
<td>1.6 Some Comments on R</td>
<td>10</td>
</tr>
<tr>
<td>References</td>
<td>12</td>
</tr>
<tr>
<td><strong>2 Computer Arithmetic</strong></td>
<td>13</td>
</tr>
<tr>
<td>2.1 Introduction</td>
<td>13</td>
</tr>
<tr>
<td>2.2 Positional Number Systems</td>
<td>14</td>
</tr>
<tr>
<td>2.3 Fixed Point Arithmetic</td>
<td>17</td>
</tr>
<tr>
<td>2.4 Floating Point Representations</td>
<td>20</td>
</tr>
<tr>
<td>2.5 Living with Floating Point Inaccuracies</td>
<td>23</td>
</tr>
<tr>
<td>2.6 The Pale and Beyond</td>
<td>28</td>
</tr>
<tr>
<td>2.7 Conditioned Problems and Stable Algorithms</td>
<td>32</td>
</tr>
<tr>
<td>Programs and Demonstrations</td>
<td>34</td>
</tr>
<tr>
<td>Exercises</td>
<td>35</td>
</tr>
<tr>
<td>References</td>
<td>38</td>
</tr>
<tr>
<td><strong>3 Matrices and Linear Equations</strong></td>
<td>40</td>
</tr>
<tr>
<td>3.1 Introduction</td>
<td>40</td>
</tr>
<tr>
<td>3.2 Matrix Operations</td>
<td>41</td>
</tr>
<tr>
<td>3.3 Solving Triangular Systems</td>
<td>43</td>
</tr>
<tr>
<td>3.4 Gaussian Elimination</td>
<td>44</td>
</tr>
<tr>
<td>3.5 Cholesky Decomposition</td>
<td>50</td>
</tr>
<tr>
<td>3.6 Matrix Norms</td>
<td>54</td>
</tr>
<tr>
<td>3.7 Accuracy and Conditioning</td>
<td>55</td>
</tr>
<tr>
<td>3.8 Matrix Computations in R</td>
<td>60</td>
</tr>
<tr>
<td>Programs and Demonstrations</td>
<td>61</td>
</tr>
<tr>
<td>Exercises</td>
<td>63</td>
</tr>
<tr>
<td>References</td>
<td>65</td>
</tr>
</tbody>
</table>
4 More Methods for Solving Linear Equations 67
  4.1 Introduction 67
  4.2 Full Elimination with Complete Pivoting 67
  4.3 Banded Matrices 71
  4.4 Applications to ARMA Time-Series Models 73
  4.5 Toeplitz Systems 76
  4.6 Sparse Matrices 80
  4.7 Iterative Methods 82
  4.8 Linear Programming 84
    Programs and Demonstrations 87
    Exercises 88
    References 90

5 Regression Computations 91
  5.1 Introduction 91
  5.2 Condition of the Regression Problem 93
  5.3 Solving the Normal Equations 96
  5.4 Gram-Schmidt Orthogonalization 97
  5.5 Householder Transformations 100
  5.6 Householder Transformations for Least Squares 101
  5.7 Givens Transformations 104
  5.8 Givens Transformations for Least Squares 105
  5.9 Regression Diagnostics 107
  5.10 Hypothesis Tests 110
  5.11 Conjugate Gradient Methods 112
  5.12 Doolittle, the Sweep, and All Possible Regressions 115
  5.13 Alternatives to Least Squares 118
  5.14 Comments 120
    Programs and Demonstrations 122
    Exercises 122
    References 125

6 Eigenproblems 128
  6.1 Introduction 128
  6.2 Theory 128
  6.3 Power Methods 130
  6.4 The Symmetric Eigenproblem and Tridiagonalization 133
  6.5 The QR Algorithm 135
  6.6 Singular Value Decomposition 137
  6.7 Applications 140
  6.8 Complex Singular Value Decomposition 144
    Programs and Demonstrations 146
    Exercises 147
    References 150
## Contents

7  **Functions: Interpolation, Smoothing, and Approximation**  
7.1 Introduction  
7.2 Interpolation  
7.3 Interpolating Splines  
7.4 Curve Fitting with Splines: Smoothing and Regression  
7.5 Mathematical Approximation  
7.6 Practical Approximation Techniques  
7.7 Computing Probability Functions  
Programs and Demonstrations  
Exercises  
References  

8  **Introduction to Optimization and Nonlinear Equations**  
8.1 Introduction  
8.2 Safe Univariate Methods: Lattice Search, Golden Section, and Bisection  
8.3 Root Finding  
8.4 First Digression: Stopping and Condition  
8.5 Multivariate Newton’s Methods  
8.6 Second Digression: Numerical Differentiation  
8.7 Minimization and Nonlinear Equations  
8.8 Condition and Scaling  
8.9 Implementation  
8.10 A Non-Newton Method: Nelder-Mead  
Programs and Demonstrations  
Exercises  
References  

9  **Maximum Likelihood and Nonlinear Regression**  
9.1 Introduction  
9.2 Notation and Asymptotic Theory of Maximum Likelihood  
9.3 Information, Scoring, and Variance Estimates  
9.4 An Extended Example  
9.5 Concentration, Iteration, and the EM Algorithm  
9.6 Multiple Regression in the Context of Maximum Likelihood  
9.7 Generalized Linear Models  
9.8 Nonlinear Regression  
9.9 Parameterizations and Constraints  
Programs and Demonstrations  
Exercises  
References  

10  **Numerical Integration and Monte Carlo Methods**  
10.1 Introduction  
10.2 Motivating Problems  
10.3 One-Dimensional Quadrature  
Programs and Demonstrations  
Exercises  
References  

© in this web service Cambridge University Press  www.cambridge.org
### Contents

10.4 Numerical Integration in Two or More Variables 271  
10.5 Uniform Pseudorandom Variables 278  
10.6 Quasi–Monte Carlo Integration 286  
10.7 Strategy and Tactics 291  
Programs and Demonstrations 295  
Exercises 297  
References 299  

11 Generating Random Variables from Other Distributions 303  
11.1 Introduction 303  
11.2 General Methods for Continuous Distributions 304  
11.3 Algorithms for Continuous Distributions 308  
11.4 General Methods for Discrete Distributions 321  
11.5 Algorithms for Discrete Distributions 325  
11.6 Other Randomizations 330  
11.7 Accuracy in Random Number Generation 334  
Programs and Demonstrations 337  
Exercises 338  
References 341  

12 Statistical Methods for Integration and Monte Carlo 343  
12.1 Introduction 343  
12.2 Distribution and Density Estimation 343  
12.3 Distributional Tests 350  
12.4 Importance Sampling and Weighted Observations 353  
12.5 Testing Importance Sampling Weights 359  
12.6 Laplace Approximations 361  
12.7 Randomized Quadrature 363  
12.8 Spherical–Radial Methods 365  
Programs and Demonstrations 370  
Exercises 372  
References 373  

13 Markov Chain Monte Carlo Methods 375  
13.1 Introduction 375  
13.2 Markov Chains 377  
13.3 Gibbs Sampling 378  
13.4 Metropolis–Hastings Algorithm 383  
13.5 Time-Series Analysis 386  
13.6 Adaptive Acceptance/Rejection 390  
13.7 Diagnostics 394  
Programs and Demonstrations 398  
Exercises 398  
References 400
## Contents

14 **Sorting and Fast Algorithms** 403
   14.1 Introduction 403
   14.2 Divide and Conquer 403
   14.3 Sorting Algorithms 405
   14.4 Fast Order Statistics and Related Problems 408
   14.5 Fast Fourier Transform 409
   14.6 Convolutions and the Chirp-$\zeta$ Transform 413
   14.7 Statistical Applications of the FFT 415
   14.8 Combinatorial Problems 425
   Programs and Demonstrations 429
   Exercises 433
   References 436

*Author Index* 439

*Subject Index* 444
Preface to the Second Edition

In the ten years since the first edition of this book went to press, the field of statistical computing has exploded with innovations in many directions. At one time my goal was to write a comprehensive book on the subject. At this moment, however, my goals for a second edition must be more modest. Because the field has grown so much, the scope of this book has now become the core for a subset of this field. To fill in some gaps in this new core, a few sections have been added (e.g., linear programming) and others have been expanded. Many corrections have been made; I can only hope that just a few errors remain.

A second change in this timespan is the rapid widespread adoption of R in the field of statistics. As language and culture shape each other, my own views on computing have changed from teaching this material using R. Small changes scattered throughout reflect this change in viewpoint. Additionally, most of the demonstrations and examples – all that seemed appropriate – have been translated to R and are available on my Web site for this book (http://www4.stat.ncsu.edu/~monahan/nmos2/toc.html).

Thanks are due to Lauren Cowles of Cambridge University Press for encouraging this second edition. Karen Chiswell deserves recognition for finding numerous typos and providing other corrections. I would like to also thank Jerry Davis and Wendy Meiring for pointing out others. Bruce McCullough provided invaluable feedback, comments, questions, and suggestions. Thanks are also due to the many students who, perhaps unknowingly, provided feedback with their questions. And this second edition would not be possible without the love, support, and patience of my wife Carol.
Preface to the First Edition

This book grew out of notes for my Statistical Computing course that I have been teaching for the past 20 years at North Carolina State University. The goal of this course is to prepare doctoral students with the computing tools needed for statistical research, and I have augmented this core with related topics that through the years I have found useful for colleagues and graduate students. As a result, this book covers a wide range of computational issues, from arithmetic, numerical linear algebra, and approximation, which are typical numerical analysis topics, to optimization and non-linear regression, to random number generation, and finally to fast algorithms. I have emphasized numerical techniques but restricted the scope to those regularly employed in the field of statistics and dropped some traditional numerical analysis topics such as differential equations. Many of the exercises in this book arose from questions posed to me by colleagues and students.

Most of the students that I have taught come with a graduate level understanding of statistics, no experience in numerical analysis, and little skill in a programming language. Consequently, I cover only about half of this material in a one-semester course. For those with a background in numerical analysis, a basic understanding of two statistical topics, regression and maximum likelihood, would be necessary.

I would advise any instructor of statistical computing not to shortchange the fundamental topic of arithmetic. I have found that most students resist the idea that computers have limited precision and employ many defense mechanisms to support that denial. Until students are comfortable with finite precision arithmetic, this psychological obstacle will cripple their understanding of scientific computation. As a result, I urge the use of single precision arithmetic in the early part of the course and introduce numerical linear algebra using a low-level language, even though students may eventually use software or languages that completely hide the calculations behind operators and double precision. These operators will continue to be mysterious black boxes until the fundamental concept of finite precision arithmetic is understood and accepted.

Early in this effort, I faced the dilemma of how to describe algorithms. The big picture is easier to present or to understand with pseudocode descriptions of algorithms. But I always felt that skipping over the details was misleading the reader, especially when the details are critical to the success of an implementation. Furthermore, there is no better challenge to one’s understanding of a topic than to take a big-picture description and program it to the smallest detail. On the other hand, writing one’s own implementation of an algorithm often seems like a futile reinvention of the wheel.
Preface to the First Edition

And so my response to this dilemma is to have it both ways: to present algorithms in pseudocode in the text, but also to supplement the pseudocode with Fortran programs and demonstrations on the accompanying disk.

These programs provide the basic tools for extending the realm of statistical techniques beyond the bounds of current statistical software. But my primary goal in providing this code is instructional. Some exercises consist of implementing a particular algorithm, and occasionally I have intentionally included my implementation for the reader to compare with, or, perhaps, improve upon. I encourage the reader to examine the details of the code and to see how the algorithms respond to changes. A secondary goal is to include as many realistic problems as practicable, having endured the frustration of failing to get code to work on anything but toy problems.

I would like to express my appreciation to the many sources of support behind this effort. First of all, three heads of the Department of Statistics have supported my work in statistical computing: Tom Gerig, Dan Solomon, and the late Dave Mason. Some of the work included here is the result of collaborations with many colleagues over the years; especially notable are Al Kinderman on random number generation and Alan Genz on numerical integration. In particular, I would like to thank Sujit Ghosh and Dave Dickey for contributing invaluable advice on Chapter 13. Dennis Boos deserves special acknowledgment as a friend, colleague, and collaborator, and most importantly, for supplying me with many interesting problems over the years. I would like to thank all of the colleagues and students who brought interesting problems to me that have become material in this book. Finally, I appreciate the feedback that students have given me each semester on earlier versions of this manuscript, including their blank stares and yawns, as well as insightful questions.