Computational Probability

Algorithms and Applications in the Mathematical Sciences

Second Edition
Preface

For decades, statisticians have enjoyed the use of “statistical packages” which read in a (potentially) large data set, calculate numerical summaries such as the sample mean and sample variance, calculate more sophisticated statistical quantities such as confidence interval bounds and p-values associated with hypothesis tests, and generate statistical graphics such as histograms, box plots, and multidimensional plots. But pity the poor probabilist, who through all those decades had only paper and pencil for symbolic calculations. The purpose of this monograph is to address the plight of the probabilist by providing algorithms to perform calculations associated with random variables. We refer to a collection of data structures and algorithms that automate probability calculations as “computational probability.” The data structures and algorithms introduced here have been implemented in a language known as APPL (A Probability Programming Language). Several illustrations of problems from the mathematical sciences that can be solved by implementing these algorithms in a computer algebra system are presented in the final chapters of this monograph.

The algorithms for manipulating random variables (e.g., adding, multiplying, transforming, ordering) symbolically result in an entire class of new problems that can now be addressed. APPL is able to perform exact probability calculations for problems that would otherwise be deemed intractable. The work is quite distinct from traditional probability analysis in that a computer algebra system, in this case Maple, is used as a computing platform.

The use of a computer algebra system to solve problems in operations research and probability is increasing. Other researchers also sense the benefits of incorporating a computer algebra system into fields with probabilistic applications, for example, Parlar’s Interactive Operations Research with Maple [98], Karian and Tanis’s second edition of Probability and Statistics: Explorations with Maple [58], Rose and Smith’s Mathematical Statistics and Mathematica [105], and Hasting’s second edition of Introduction to the Mathematics of Operations Research with Mathematica [46].
This monograph significantly differs from the four titles listed above in two ways. First, the four titles listed above are all textbooks, rather than research monographs. They contain exercises and examples geared toward students, rather than researchers. Second, the emphasis in most of these texts is much broader than the emphasis being proposed here. For example, Parlar and Hasting consider all of OR/MS, rather than the probabilistic side of OR/MS proposed here in much more depth. Also, Karian and Tanis emphasize Monte Carlo solutions to probability and statistics problems, as opposed to the exact solutions given in APPL.

The monograph begins with an introductory chapter that contains short examples involving the elementary use of APPL. Chapter 2 reviews the Maple data structures and functions necessary to implement APPL. This is followed by a discussion of the development of the data structures and algorithms (Chaps. 3–6 for continuous random variables and Chaps. 7–9 for discrete random variables) used in APPL. The monograph concludes with Chaps. 10–15 introducing a sampling of various applications in the mathematical sciences. The two most likely audiences for the monograph are researchers in the mathematical sciences with an interest in applied probability and instructors using the monograph for a special topics course in computational probability taught in a mathematics, statistics, operations research, management science, or industrial engineering department. The intended audience for this monograph includes researchers, MS students, PhD students, and advanced practitioners in stochastic operations research, management science, and applied probability.

An indication of the utility of APPL is that the research efforts of the authors and other colleagues have produced many related refereed journal publications, many conference presentations, the ICS Computing Prize with INFORMS, a government patent, the INFORMS Undergraduate Research Prize, a GEM Scholarship based on APPL research, and multiple improvements to pedagogical methods in numerous colleges and universities around the world. We believe that the potential of this field of computational probability in research and education is unlimited. It is our hope that this monograph encourages people to join us in attaining future accomplishments in this field.

We are grateful to Camille Price, Gary Folven, Carolyn Ford, and Matt Amboy from Springer for their support of this project. We also thank Professor Ludolf Meester from TU Delft for class testing and debugging portions of the APPL code. We thank our coauthors Matt Duggan, Kerry Connell, Jeff Mallozzi, and Bruce Schmeiser for their collaboration on the applications given in Chap. 11. We also thank the editors, anonymous referees, and colleagues who have been helpful in the presentation of the material, including John Backes, Donald Barr, Don Campbell, Jacques Carette, Gianfranco Ciardo, Mike Crawford, Mark Eaton, Jerry Ellis, Bob Foote, James Fritz, Winfried Grassmann, Greg Gruver, Carl Harris, Rex Kincaid, Hank Krieger, Marina Kondratovitch, Sid Lawrence, Barry Nelson, David Nicol, Raghu Pasupathy, Steve Roberts, Sheldon Ross, Bruce Schmeiser, Jim Scott, Bob Shumaker,
Paul Stockmeyer, Bill Treadwell, Michael Trosset, Mark Vinson, and Bert Zwart. We thank techie Robert Marmorstein for his LaTeX support, Chris Weld for proofreading the second edition, and Barbara Boyer for her help with the index keywords for the first edition of this monograph. The authors gratefully acknowledge support from the Clare Boothe Luce Foundation, the National Science Foundation (for providing funding for an Educational Innovation Grant CDA9712718 “Undergraduate Modeling and Simulation Analysis,” for scholarship funds provided under the CSEMS Grant 0123022 “Effective Transitions Through Academe to Industry for Computer Scientists and Mathematicians,” for support by an NSF CSUMS grant DMS-0703532), the Army Research Office, Grant 67906-MA, and The College of William & Mary for research leave to support this endeavor.

Finally, our thanks goes to all of our coauthors who have made such significant contributions to this volume and to APPL language. Nearly all of the material in this monograph has originally appeared in journal articles: Chap. 1 is from Glen et al. [38]; Chaps. 2 and 3 contain original material by Diane Evans; Chap. 4 is from Glen et al. [37]; Chap. 5 is from Yang et al. [122]; Chap. 6 is from Glen et al. [39]; Chap. 7 is from Glen et al. [38]; Chap. 8 is from Evans and Leemis [29]; Chap. 9 is from Evans et al. [30]; Chap. 10 is from Leemis [73], Evans et al. [30], and Marks et al. [82]; Chap. 11 is from Webb and Leemis [115]; Chap. 12 is from Duggan et al. [24] and Drew et al. [23]; Chap. 13 is from Kaczynski et al. [56]; Chap. 14 is original work by Matt Robinson; and Chap. 15 is from Leemis [73], Leemis et al. [74], and Glen et al. [38].

Williamsburg, VA, USA
John H. Drew
Terre Haute, IN, USA
Diane L. Evans
Colorado Springs, CO, USA
Andrew G. Glen
Williamsburg, VA, USA
Lawrence M. Leemis
June 2016
# Contents

**Part I Introduction**

1. **Computational Probability** ............................................. 3  
   1.1 Four Simple Examples of the Use of APPL ..................... 3  
   1.2 A Different Way of Thinking .................................. 8  
   1.3 Overview ................................................................ 11

2. **Maple for APPL** .......................................................... 13  
   2.1 Numerical Computations ........................................... 13  
   2.2 Variables .................................................................. 15  
   2.3 Symbolic Computations ............................................ 16  
   2.4 Functions ................................................................... 17  
   2.5 Data Types .............................................................. 19  
   2.6 Solving Equations .................................................... 20  
   2.7 Graphing .................................................................... 22  
   2.8 Calculus ..................................................................... 23  
   2.9 Loops and Conditions .............................................. 26  
   2.10 Procedures ............................................................. 28

**Part II Algorithms for Continuous Random Variables**

3. **Data Structures and Simple Algorithms** ............................ 33  
   3.1 Data Structures ....................................................... 33  
   3.2 Simple Algorithms .................................................... 37  

4. **Transformations of Random Variables** ............................ 47  
   4.1 Theorem ..................................................................... 48  
   4.2 Implementation in APPL ........................................... 50  
   4.3 Examples ................................................................... 52
5 Bivariate Transformations of Random Variables ........... 57
  5.1 Algorithm .............................................. 58
  5.2 Data Structure .......................................... 59
  5.3 Implementation ......................................... 60
  5.4 Examples ............................................... 62

6 Products of Random Variables ............................. 73
  6.1 Theorem ............................................... 74
  6.2 Implementation in APPL ................................. 76
  6.3 Examples ............................................... 79
  6.4 Extensions .............................................. 82
  6.5 Algorithm .............................................. 84

Part III Algorithms for Discrete Random Variables

7 Data Structures and Simple Algorithms ................. 89
  7.1 Data Structures ......................................... 89
  7.2 Simple Algorithms ..................................... 99

8 Sums of Independent Discrete Random Variables .......... 111
  8.1 Preliminary Examples .................................... 111
  8.2 Conceptual Algorithm Development ..................... 116
  8.3 Algorithm .............................................. 126
  8.4 Implementation Issues .................................. 128
  8.5 Examples ............................................... 130

9 Order Statistics for Random Sampling from Discrete Populations ................................................ 139
  9.1 Notation and Taxonomy .................................. 139
  9.2 Sampling Without Replacement ........................... 141
  9.3 Sampling with Replacement ............................... 145
  9.4 Extension .............................................. 150

Part IV Applications

10 Reliability and Survival Analysis ........................ 155
  10.1 Systems Analysis ........................................ 155
  10.2 Lower Confidence Bound on System Reliability .......... 161
  10.3 Survival Analysis ....................................... 164
  10.4 Applying Bootstrap Methods to System Reliability ....... 172
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>11</td>
<td>Symbolic ARMA Model Analysis</td>
<td>191</td>
</tr>
<tr>
<td>11.1</td>
<td>ARMA Model Basics</td>
<td>192</td>
</tr>
<tr>
<td>11.2</td>
<td>Implementation</td>
<td>193</td>
</tr>
<tr>
<td>11.3</td>
<td>Examples</td>
<td>197</td>
</tr>
<tr>
<td>12</td>
<td>Stochastic Simulation</td>
<td>209</td>
</tr>
<tr>
<td>12.1</td>
<td>Tests of Randomness</td>
<td>209</td>
</tr>
<tr>
<td>12.2</td>
<td>Input Modeling</td>
<td>215</td>
</tr>
<tr>
<td>12.3</td>
<td>Kolmogorov–Smirnov Goodness-of-Fit Test</td>
<td>223</td>
</tr>
<tr>
<td>13</td>
<td>Transient Queueing Analysis</td>
<td>241</td>
</tr>
<tr>
<td>13.1</td>
<td>Introduction</td>
<td>241</td>
</tr>
<tr>
<td>13.2</td>
<td>Basics of the $M/M/s$ Queue</td>
<td>243</td>
</tr>
<tr>
<td>13.3</td>
<td>Creating the Sojourn Time Distribution</td>
<td>244</td>
</tr>
<tr>
<td>13.4</td>
<td>Transient Analysis Applications</td>
<td>247</td>
</tr>
<tr>
<td>13.5</td>
<td>Covariance and Correlation in the $M/M/1$ Queue</td>
<td>253</td>
</tr>
<tr>
<td>13.6</td>
<td>Extending Covariance Calculations</td>
<td>260</td>
</tr>
<tr>
<td>13.7</td>
<td>Sojourn Time Covariance with $k$ Customers Initially Present</td>
<td>269</td>
</tr>
<tr>
<td>14</td>
<td>Bayesian Applications</td>
<td>277</td>
</tr>
<tr>
<td>14.1</td>
<td>Introduction</td>
<td>277</td>
</tr>
<tr>
<td>14.2</td>
<td>Bayesian Inference on Single-Parameter Distributions</td>
<td>279</td>
</tr>
<tr>
<td>14.3</td>
<td>Bayesian Inference on Multiple-Parameter Distributions</td>
<td>287</td>
</tr>
<tr>
<td>15</td>
<td>Other Applications</td>
<td>301</td>
</tr>
<tr>
<td>15.1</td>
<td>Stochastic Activity Networks</td>
<td>301</td>
</tr>
<tr>
<td>15.2</td>
<td>Benford’s Law</td>
<td>310</td>
</tr>
<tr>
<td>15.3</td>
<td>Miscellaneous Applications</td>
<td>315</td>
</tr>
<tr>
<td>References</td>
<td></td>
<td>323</td>
</tr>
<tr>
<td>Index</td>
<td></td>
<td>329</td>
</tr>
</tbody>
</table>