An Introduction to Statistical Computing
WILEY SERIES IN COMPUTATIONAL STATISTICS

Consulting Editors:

Paolo Giudici
*University of Pavia, Italy*

Geof H. Givens
*Colorado State University, USA*

Bani K. Mallick
*Texas A & M University, USA*

---

*Wiley Series in Computational Statistics* is comprised of practical guides and cutting edge research books on new developments in computational statistics. It features quality authors with a strong applications focus. The texts in the series provide detailed coverage of statistical concepts, methods and case studies in areas at the interface of statistics, computing, and numerics.

With sound motivation and a wealth of practical examples, the books show in concrete terms how to select and to use appropriate ranges of statistical computing techniques in particular fields of study. Readers are assumed to have a basic understanding of introductory terminology.

The series concentrates on applications of computational methods in statistics to fields of bioinformatics, genomics, epidemiology, business, engineering, finance and applied statistics.

**Titles in the Series**

- Biegler, Biros, Ghattas, Heinkenschloss, Keyes, Mallick, Marzouk, Tenorio, Waanders, Willcox – *Large-Scale Inverse Problems and Quantification of Uncertainty*
- Billard and Diday – *Symbolic Data Analysis: Conceptual Statistics and Data Mining*
- Bolstad – *Understanding Computational Bayesian Statistics*
- Borgelt, Steinbrecher and Kruse – *Graphical Models, 2e*
- Dunne – *A Statistical Approach to Neutral Networks for Pattern Recognition*
- Liang, Liu and Carroll – *Advanced Markov Chain Monte Carlo Methods*
- Ntzoufras – *Bayesian Modeling Using WinBUGS*
- Tuff´ery – *Data Mining and Statistics for Decision Making*
An Introduction to Statistical Computing

A Simulation-based Approach

Jochen Voss

School of Mathematics, University of Leeds, UK

WILEY
Contents

List of algorithms ix

Preface xi

Nomenclature xiii

1 Random number generation 1
1.1 Pseudo random number generators 2
  1.1.1 The linear congruential generator 2
  1.1.2 Quality of pseudo random number generators 4
  1.1.3 Pseudo random number generators in practice 8
1.2 Discrete distributions 8
1.3 The inverse transform method 11
1.4 Rejection sampling 15
  1.4.1 Basic rejection sampling 15
  1.4.2 Envelope rejection sampling 18
  1.4.3 Conditional distributions 22
  1.4.4 Geometric interpretation 26
1.5 Transformation of random variables 30
1.6 Special-purpose methods 36
1.7 Summary and further reading 36
Exercises 37

2 Simulating statistical models 41
2.1 Multivariate normal distributions 41
2.2 Hierarchical models 45
2.3 Markov chains 50
  2.3.1 Discrete state space 51
  2.3.2 Continuous state space 56
2.4 Poisson processes 58
2.5 Summary and further reading 67
Exercises 67
3 Monte Carlo methods 69
  3.1 Studying models via simulation 69
  3.2 Monte Carlo estimates 74
    3.2.1 Computing Monte Carlo estimates 75
    3.2.2 Monte Carlo error 76
    3.2.3 Choice of sample size 80
    3.2.4 Refined error bounds 82
  3.3 Variance reduction methods 84
    3.3.1 Importance sampling 84
    3.3.2 Antithetic variables 88
    3.3.3 Control variates 93
  3.4 Applications to statistical inference 96
    3.4.1 Point estimators 97
    3.4.2 Confidence intervals 100
    3.4.3 Hypothesis tests 103
  3.5 Summary and further reading 106
Exercises 106

4 Markov Chain Monte Carlo methods 109
  4.1 The Metropolis–Hastings method 110
    4.1.1 Continuous state space 110
    4.1.2 Discrete state space 113
    4.1.3 Random walk Metropolis sampling 116
    4.1.4 The independence sampler 119
    4.1.5 Metropolis–Hastings with different move types 120
  4.2 Convergence of Markov Chain Monte Carlo methods 125
    4.2.1 Theoretical results 125
    4.2.2 Practical considerations 129
  4.3 Applications to Bayesian inference 137
  4.4 The Gibbs sampler 141
    4.4.1 Description of the method 141
    4.4.2 Application to parameter estimation 146
    4.4.3 Applications to image processing 151
  4.5 Reversible Jump Markov Chain Monte Carlo 158
    4.5.1 Description of the method 160
    4.5.2 Bayesian inference for mixture distributions 171
  4.6 Summary and further reading 178
  4.6 Exercises 178

5 Beyond Monte Carlo 181
  5.1 Approximate Bayesian Computation 181
    5.1.1 Basic Approximate Bayesian Computation 182
    5.1.2 Approximate Bayesian Computation with regression 188
  5.2 Resampling methods 192
## Contents

5.2.1 Bootstrap estimates 192  
5.2.2 Applications to statistical inference 197  
5.3 Summary and further reading 209  
Exercises 209  

6 Continuous-time models 213  
6.1 Time discretisation 213  
6.2 Brownian motion 214  
6.2.1 Properties 216  
6.2.2 Direct simulation 217  
6.2.3 Interpolation and Brownian bridges 218  
6.3 Geometric Brownian motion 221  
6.4 Stochastic differential equations 224  
6.4.1 Introduction 224  
6.4.2 Stochastic analysis 226  
6.4.3 Discretisation schemes 231  
6.4.4 Discretisation error 236  
6.5 Monte Carlo estimates 243  
6.5.1 Basic Monte Carlo 243  
6.5.2 Variance reduction methods 247  
6.5.3 Multilevel Monte Carlo estimates 250  
6.6 Application to option pricing 255  
6.7 Summary and further reading 259  
Exercises 260  

Appendix A Probability reminders 263  
A.1 Events and probability 263  
A.2 Conditional probability 266  
A.3 Expectation 268  
A.4 Limit theorems 269  
A.5 Further reading 270  

Appendix B Programming in R 271  
B.1 General advice 271  
B.2 R as a Calculator 272  
B.2.1 Mathematical operations 273  
B.2.2 Variables 273  
B.2.3 Data types 275  
B.3 Programming principles 282  
B.3.1 Don’t repeat yourself! 283  
B.3.2 Divide and conquer! 286  
B.3.3 Test your code! 290  
B.4 Random number generation 292  
B.5 Summary and further reading 294  
Exercises 294
Appendix C  Answers to the exercises  299
  C.1  Answers for Chapter 1  299
  C.2  Answers for Chapter 2  315
  C.3  Answers for Chapter 3  319
  C.4  Answers for Chapter 4  328
  C.5  Answers for Chapter 5  342
  C.6  Answers for Chapter 6  350
  C.7  Answers for Appendix B  366

References  375

Index  379
List of algorithms

Random number generation
alg. 1.2 linear congruential generator 2
alg. 1.13 inverse transform method 12
alg. 1.19 basic rejection sampling 15
alg. 1.22 envelope rejection sampling 19
alg. 1.25 rejection sampling for conditional distributions 22

Simulating statistical models
alg. 2.9 mixture distributions 47
alg. 2.11 componentwise simulation 49
alg. 2.22 Markov chains with discrete state space 53
alg. 2.31 Markov chains with continuous state space 58
alg. 2.36 Poisson process 61
alg. 2.41 thinning method for Poisson processes 65

Monte Carlo methods
alg. 3.8 Monte Carlo estimate 75
alg. 3.22 importance sampling 85
alg. 3.26 antithetic variables 89
alg. 3.31 control variates 93

Markov Chain Monte Carlo methods
alg. 4.2 Metropolis–Hastings method for continuous state space 110
alg. 4.4 Metropolis–Hastings method for discrete state space 113
alg. 4.9 random walk Metropolis 117
alg. 4.11 independence sampler 119
alg. 4.12 Metropolis–Hastings method with different move types 121
alg. 4.27 Gibbs sampler 142
alg. 4.31 Gibbs sampler for the Ising model 155
alg. 4.32 Gibbs sampler in image processing 158
alg. 4.36 reversible jump Markov Chain Monte Carlo 165
LIST OF ALGORITHMS

**Beyond Monte Carlo**

alg. 5.1  basic Approximate Bayesian Computation 182
alg. 5.6  Approximate Bayesian Computation with regression 191
alg. 5.11 general bootstrap estimate 196
alg. 5.15 bootstrap estimate of the bias 200
alg. 5.18 bootstrap estimate of the standard error 202
alg. 5.20 simple bootstrap confidence interval 205
alg. 5.21 BCₐ bootstrap confidence interval 207

**Continuous-time models**

alg. 6.6  Brownian motion 217
alg. 6.12 Euler–Maruyama scheme 232
alg. 6.15 Milstein scheme 235
alg. 6.26 multilevel Monte Carlo estimates 251
alg. 6.29 Euler–Maruyama scheme for the Heston model 256
This is a book about exploring random systems using computer simulation and thus, this book combines two different topic areas which have always fascinated me: the mathematical theory of probability and the art of programming computers. The method of using computer simulations to study a system is very different from the more traditional, purely mathematical approach. On the one hand, computer experiments normally can only provide approximate answers to quantitative questions, but on the other hand, results can be obtained for a much wider class of systems, including large and complex systems where a purely theoretical approach becomes difficult.

In this text we will focus on three different types of questions. The first, easiest question is about the normal behaviour of the system: what is a typical state of the system? Such questions can be easily answered using computer experiments: simulating a few random samples of the system gives examples of typical behaviour. The second kind of question is about variability: how large are the random fluctuations? This type of question can be answered statistically by analysing large samples, generated using repeated computer simulations. A final, more complicated class of questions is about exceptional behaviour: how small is the probability of the system behaving in a specified untypical way? Often, advanced methods are required to answer this third type of question. The purpose of this book is to explain how such questions can be answered. My hope is that, after reading this book, the reader will not only be able to confidently use methods from statistical computing for answering such questions, but also to adjust existing methods to the requirements of a given problem and, for use in more complex situations, to develop new specialised variants of the existing methods.

This text originated as a set of handwritten notes which I used for teaching the ‘Statistical Computing’ module at the University of Leeds, but now is greatly extended by the addition of many examples and more advanced topics. The material we managed to cover in the ‘Statistical Computing’ course during one semester is less than half of what is now the contents of the book! This book is aimed at postgraduate students and their lecturers; it can be used both for self-study and as the basis of taught courses. With the inclusion of many examples and exercises, the text should also be accessible to interested undergraduate students and to mathematically inclined researchers from areas outside mathematics.
Only very few prerequisites are required for this book. On the mathematical side, the text assumes that the reader is familiar with basic probability, up to and including the law of large numbers; Appendix A summarises the required results. As a consequence of the decision to require so little mathematical background, some of the finer mathematical subtleties are not discussed in this book. Results are presented in a way which makes them easily accessible to readers with limited mathematical background, but the statements are given in a form which allows the mathematically more knowledgeable reader to easily add the required detail on his/her own. (For example, I often use phrases such as ‘every set $A \subseteq \mathbb{R}^d$’ where full mathematical rigour would require us to write ‘every measurable set $A \subseteq \mathbb{R}^d$’.) On the computational side, basic programming skills are required to make use of the numerical methods introduced in this book. While the text is written independent of any specific programming language, the reader will need to choose a language when implementing methods from this book on a computer. Possible choices of programming language include Python, Matlab and C/C++. For my own implementations, provided as part of the solutions to the exercises in Appendix C, I used the R programming language; a short introduction to programming with R is provided in Appendix B.

Writing this book has been a big adventure for me. When I started this project, more than a year ago, my aim was to cover enough material so that I could discuss the topics of multilevel Monte Carlo and reversible jump Markov Chain Monte Carlo methods. I estimated that 350 pages would be enough to cover this material but it quickly transpired that I had been much too optimistic: my estimates for the final page count kept rising and even after several rounds of throwing out side-topics and generally tightening the text, the book is still stretching this limit! Nevertheless, the text now covers most of the originally planned topics, including multilevel Monte Carlo methods near the very end of the book. Due to my travel during the last year, parts of this book have been written on a laptop in exciting places. For example, the initial draft of section 1.5 was written on a coach travelling through the beautiful island of Kyushu, halfway around the world from where I live! All in all, I greatly enjoyed writing this book and I hope that the result is useful to the reader.

This book contains an accompanying website. Please visit www.wiley.com/go/statistical_computing

Jochen Voss

Leeds, March 2013
# Nomenclature

For reference, the following list summarises some of the notation used throughout this book.

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \emptyset )</td>
<td>the empty set</td>
</tr>
<tr>
<td>( \mathbb{N} )</td>
<td>the natural numbers: ( \mathbb{N} = {1, 2, 3, \ldots } )</td>
</tr>
<tr>
<td>( \mathbb{N}_0 )</td>
<td>the non-negative integers: ( \mathbb{N}_0 = {0, 1, 2, \ldots } )</td>
</tr>
<tr>
<td>( \mathbb{Z} )</td>
<td>the integers: ( \mathbb{Z} = {\ldots, -2, -1, 0, 1, 2, \ldots } )</td>
</tr>
<tr>
<td>( n \mod m )</td>
<td>the remainder of the division of ( n ) by ( m ), in the range ( 0, 1, \ldots, m - 1 )</td>
</tr>
<tr>
<td>( \delta_{kl} )</td>
<td>the Kronecker delta: ( \delta_{kl} = 1 ) if ( k = l ) and ( \delta_{kl} = 0 ) otherwise</td>
</tr>
<tr>
<td>( \mathbb{R} )</td>
<td>the real numbers</td>
</tr>
<tr>
<td>( \lceil x \rceil )</td>
<td>the number ( x \in \mathbb{R} ) ‘rounded up’, that is the smallest integer greater than or equal to ( x )</td>
</tr>
<tr>
<td>( (a_n)_{n \in \mathbb{N}} )</td>
<td>a sequence of (possibly random) numbers: ( (a_n)_{n \in \mathbb{N}} = (a_1, a_2, \ldots) )</td>
</tr>
<tr>
<td>( \mathcal{O}(\cdot) )</td>
<td>the big ( \mathcal{O} ) notation, introduced in definition 3.16</td>
</tr>
<tr>
<td>( [a, b] )</td>
<td>an interval of real numbers: ( [a, b] = {x \in \mathbb{R} \mid a \leq x \leq b} )</td>
</tr>
<tr>
<td>( {a, b} )</td>
<td>the set containing ( a ) and ( b )</td>
</tr>
<tr>
<td>( A^C )</td>
<td>the complement of a set: ( A^C = {x \mid x \notin A} )</td>
</tr>
<tr>
<td>( A \times B )</td>
<td>the Cartesian product of the sets ( A ) and ( B ): ( A \times B = {(a, b) \mid a \in A, b \in B} )</td>
</tr>
<tr>
<td>( \mathbb{I}_A(x) )</td>
<td>the indicator function of the set ( A ): ( \mathbb{I}_A(x) = 1 ) if ( x \in A ) and ( 0 ) otherwise (see section A.3)</td>
</tr>
<tr>
<td>( \mathcal{U}[0, 1] )</td>
<td>the uniform distribution on the interval ( [0, 1] )</td>
</tr>
<tr>
<td>( \mathcal{U}{-1, 1} )</td>
<td>the uniform distribution on the two-element set ( {-1, 1} )</td>
</tr>
<tr>
<td>( \text{Pois}(\lambda) )</td>
<td>the Poisson distribution with parameter ( \lambda )</td>
</tr>
<tr>
<td>( X \sim \mu )</td>
<td>indicates that a random variable ( X ) is distributed according to a probability distribution ( \mu )</td>
</tr>
<tr>
<td>(</td>
<td>S</td>
</tr>
<tr>
<td>( \mathbb{R}^S )</td>
<td>space of vectors where the components are indexed by elements of ( S ) (see section 2.3.2)</td>
</tr>
<tr>
<td>( \mathbb{R}^{S \times S} )</td>
<td>space of matrices where rows and columns are indexed by elements of ( S ) (see section 2.3.2)</td>
</tr>
</tbody>
</table>