

Christian P. Robert George Casella

Monte Carlo Statistical Methods

With 65 Figures



Springer

Christian P. Robert
CREST-INSEE
Laboratoire de Statistique
75675 Paris Cedex 14
France

Dept. de Mathematique
UFR des Sciences
Universite de Rouen
76821 Mont Saint Aignan cedex
France

George Casella
Biometrics Unit
Cornell University
Ithaca, NY 14853-7801
USA

Editorial Board

George Casella
Biometrics Unit
Cornell University
Ithaca, NY 14853-7801
USA

Stephen Fienberg
Department of Statistics
Carnegie Mellon University
Pittsburgh, PA 15213
USA

Ingram Olkin
Department of Statistics
Stanford University
Stanford, CA 94305
USA

Library of Congress Cataloging-in-Publication Data

Robert, Christian P., 1961–

Monte Carlo statistical methods / Christian P. Robert, George Casella.

p. cm. — (Springer texts in statistics)

Includes bibliographical references (p. -) and index.

ISBN 978-1-4757-3073-9 ISBN 978-1-4757-3071-5 (eBook)

DOI 10.1007/978-1-4757-3071-5

1. Mathematical statistics. 2. Monte Carlo method. I. Casella, George. II. Title III. Series.

QA276.R575 1999

519.5—dc21

98-55413

Printed on acid-free paper.

© 1999 Springer Science+Business Media New York

Originally published by Springer-Verlag New York, Inc. in 1999

Softcover reprint of the hardcover 1st edition 1999

All rights reserved. This work may not be translated or copied in whole or in part without the written permission of the publisher Springer Science+Business Media, LLC, except for brief excerpts in connection with reviews or scholarly analysis. Use in connection with any form of information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed is forbidden.

The use of general descriptive names, trade names, trademarks, etc., in this publication, even if the former are not especially identified, is not to be taken as a sign that such names, as understood by the Trade Marks and Merchandise Marks Act, may accordingly be used freely by anyone.

Production managed by Allan Abrams; manufacturing supervised by Jeffrey Taub.

Photocomposed pages prepared from the authors' TeX files.

9 8 7 6 5 4 3 2 1

ISBN 978-1-4757-3073-9 SPIN 10707599

Contents

List of Tables	xvii
List of Figures	xix
1 Introduction	1
1.1 Statistical Models	1
1.2 Likelihood Methods	5
1.3 Bayesian Methods	11
1.4 Deterministic Numerical Methods	17
1.5 Problems	19
1.6 Notes	31
2 Random Variable Generation	35
2.1 Basic Methods	35
2.1.1 Introduction	35
2.1.2 The <i>Kiss</i> Generator	39
2.1.3 Beyond Uniform Distributions	43
2.2 Transformation Methods	45
2.3 Accept–Reject Methods	49
2.3.1 General Principle	49
2.3.2 Envelope Accept–Reject Methods	53
2.3.3 Log-Concave Densities	56
2.4 Problems	60
2.5 Notes	69
3 Monte Carlo Integration	71
3.1 Introduction	71
3.2 Classical Monte Carlo Integration	74
3.3 Importance Sampling	80
3.3.1 Principles	80
3.3.2 Finite Variance Estimators	83
3.3.3 Comparing Importance Sampling with Accept–Reject	92
3.4 Riemann Approximations	96
3.5 Laplace Approximations	103

3.6	The Saddlepoint Approximation	104
3.6.1	An Edgeworth Derivation	105
3.6.2	Tail Areas	108
3.7	Acceleration Methods	109
3.7.1	Antithetic Variables	109
3.7.2	Control Variates	113
3.7.3	Conditional Expectations	116
3.8	Problems	119
3.9	Notes	131
4	Markov Chains	139
4.1	Essentials for MCMC	140
4.2	Basic Notions	141
4.3	Irreducibility, Atoms, and Small Sets	145
4.3.1	Irreducibility	145
4.3.2	Atoms and Small Sets	147
4.3.3	Cycles and Aperiodicity	150
4.4	Transience and Recurrence	151
4.4.1	Classification of Irreducible Chains	151
4.4.2	Criteria for Recurrence	153
4.4.3	Harris Recurrence	154
4.5	Invariant Measures	155
4.6	Ergodicity and Convergence	159
4.6.1	Ergodicity	159
4.6.2	Geometric Convergence	164
4.6.3	Uniform Ergodicity	166
4.7	Limit Theorems	167
4.7.1	Ergodic Theorems	168
4.7.2	Central Limit Theorems	170
4.8	Covariance in Markov Chains	172
4.9	Problems	174
4.10	Notes	185
5	Monte Carlo Optimization	193
5.1	Introduction	193
5.2	Stochastic Exploration	194
5.2.1	A Basic Solution	194
5.2.2	Gradient Methods	196
5.2.3	Simulated Annealing	199
5.2.4	Prior Feedback	202
5.3	Stochastic Approximation	206
5.3.1	Missing Data Models and Demarginalization	207
5.3.2	Monte Carlo Approximation	208
5.3.3	The EM Algorithm	213
5.3.4	Monte Carlo EM	216
5.4	Problems	218

5.5	Notes	227
6	The Metropolis–Hastings Algorithm	231
6.1	Monte Carlo Methods Based on Markov Chains	231
6.2	The Metropolis–Hastings algorithm	233
6.2.1	Definition	233
6.2.2	Convergence Properties	236
6.3	A Collection of Metropolis–Hastings Algorithms	238
6.3.1	The Independent Case	239
6.3.2	Random Walks	245
6.3.3	ARMS: A General Metropolis–Hastings Algorithm	249
6.4	Optimization and Control	251
6.4.1	Optimizing the Acceptance Rate	252
6.4.2	Conditioning and Accelerations	255
6.5	Further Topics	259
6.5.1	Reversible Jumps	259
6.5.2	Langevin Algorithms	264
6.6	Problems	268
6.7	Notes	279
7	The Gibbs Sampler	285
7.1	General Principles	285
7.1.1	Definition	285
7.1.2	Completion	288
7.1.3	Convergence Properties	293
7.1.4	Gibbs Sampling and Metropolis–Hastings	296
7.1.5	The Hammersley–Clifford Theorem	297
7.1.6	Hierarchical Structures	299
7.2	The Two-Stage Gibbs Sampler	303
7.2.1	Dual Probability Structures	303
7.2.2	Reversible and Interleaving Chains	308
7.2.3	Monotone Covariance and Rao–Blackwellization	310
7.2.4	The Duality Principle	315
7.3	Hybrid Gibbs Samplers	317
7.3.1	Comparison with Metropolis–Hastings Algorithms	317
7.3.2	Mixtures and Cycles	319
7.3.3	Metropolizing the Gibbs Sampler	322
7.3.4	Reparameterization	326
7.4	Improper Priors	328
7.5	Problems	332
7.6	Notes	348
8	Diagnosing Convergence	363
8.1	Stopping the Chain	363
8.1.1	Convergence Criteria	363
8.1.2	Multiple Chains	366

8.1.3	Conclusions	367
8.2	Monitoring Convergence to the Stationary Distribution	368
8.2.1	Graphical Methods	368
8.2.2	Nonparametric Tests of Stationarity	370
8.2.3	Renewal Methods	371
8.2.4	Distance Evaluations	375
8.3	Monitoring Convergence of Averages	380
8.3.1	Graphical Methods	380
8.3.2	Multiple Estimates	381
8.3.3	Renewal Theory	388
8.3.4	Within and Between Variances	394
8.4	Simultaneous Monitoring	397
8.4.1	Binary Control	397
8.4.2	Valid Discretization	401
8.5	Problems	402
8.6	Notes	407
9	Implementation in Missing Data Models	415
9.1	Introduction	415
9.2	First Examples	416
9.2.1	Discrete Data Models	416
9.2.2	Data Missing at Random	419
9.3	Finite Mixtures of Distributions	422
9.4	A Reparameterization of Mixtures	425
9.5	Extensions	429
9.5.1	Hidden Markov Chains	430
9.5.2	Changepoint models	432
9.5.3	Stochastic Volatility	435
9.6	Problems	438
9.7	Notes	448
A	Probability Distributions	451
B	Notation	455
B.1	Mathematical	455
B.2	Probability	455
B.3	Distributions	456
B.4	Markov Chains	457
B.5	Statistics	457
B.6	Algorithms	458
C	References	459
	Subject Index	488
	Author Index	495

List of Tables

1.6.1 Some conjugate families	31
3.2.1 Evaluation of some normal quantiles	78
3.3.1 Comparison of instrumental distributions	90
3.3.2 Comparison between Monte Carlo and importance sampling estimators	96
3.5.1 Laplace approximation of a Gamma integral	104
3.6.1 Saddlepoint approximation of a noncentral chi squared integral	109
5.2.1 Stochastic gradient runs	199
5.2.2 Simulated annealing runs	202
5.2.3 Sequence of Bayes estimators for the gamma distribution	205
5.2.4 Average grades of first year students	206
5.2.5 Maximum likelihood estimates of mean grades	206
5.3.1 Estimation result for the factor ARCH model (5.3.10)	213
5.4.1 Tree swallow movements	225
5.4.2 Selected batting average data	226
6.3.1 Monte Carlo saddlepoint approximation of a noncentral chi squared integral	244
6.3.2 Estimators of the mean and the variance of a normal distribution using a random walk Metropolis–Hastings algorithm	246
6.4.1 Estimation in the inverse Gaussian distribution $\mathcal{IN}(\theta_1, \theta_2)$ by the Metropolis–Hastings algorithm	254
6.4.2 Decrease of squared error risk associated with Rao–Blackwellization	257
6.4.3 Improvement in quadratic risk from Rao–Blackwellization	258
6.6.1 Braking distances.	275
6.6.2 Galaxy data	276
6.7.1 Performance of the Metropolis–Hastings algorithm [A.28]	282
7.1.1 Convergence of estimated quantiles simulated from the slice sampler [A.33]	291

7.1.2 Failures of pumps in a nuclear plant	302
7.2.1 Frequencies of passage for 360 consecutive observations.	305
7.3.1 Interquantile ranges for the Gibbs sampling and the modification of Liu	326
7.5.1 Occurrences of clinical mastitis in dairy herds	334
8.3.1 Estimation of the asymptotic variance γ_h^2 for $h(\theta) = \theta$ and renewal control	392
8.3.2 Estimators of γ_h^2 for $h(x) = x$, obtained by renewal in i_0 .	393
8.3.3 Approximations by Gibbs sampling of posterior expectations and evaluation of variances by the renewal method	394
8.4.1 Minimum test sample sizes for the diagnostic of Raftery and Lewis	399
8.4.2 Parameters of the binary control method for three different parameters of the grouped multinomial model	400
8.4.3 Evolution of initializing and convergence times	401
9.2.1 Observation of two characteristics of the habitat of 164 lizards	417
9.2.2 Average incomes and numbers of responses/nonresponses to a survey on the income by age, sex, and marital status	419
9.6.1 Independent observations of $Z = (X, Y) \sim \mathcal{N}_2(0, \Sigma)$ with missing data	441
9.6.2 Yearly number of mining accidents in England, 1851 – 1962	445