Contents

Preface to the Second Edition IX

Preface to the First Edition XIII

1 Introduction 1
1.1 Statistical Models 1
1.2 Likelihood Methods 5
1.3 Bayesian Methods 12
1.4 Deterministic Numerical Methods 19
1.4.1 Optimization 19
1.4.2 Integration 21
1.4.3 Comparison 21
1.5 Problems 23
1.6 Notes 30
1.6.1 Prior Distributions 30
1.6.2 Bootstrap Methods 32

2 Random Variable Generation 35
2.1 Introduction 35
2.1.1 Uniform Simulation 36
2.1.2 The Inverse Transform 38
2.1.3 Alternatives 40
2.1.4 Optimal Algorithms 41
2.2 General Transformation Methods 42
2.3 Accept–Reject Methods 47
2.3.1 The Fundamental Theorem of Simulation 47
2.3.2 The Accept–Reject Algorithm 51
2.4 Envelope Accept–Reject Methods 53
2.4.1 The Squeeze Principle 53
2.4.2 Log-Concave Densities 56
2.5 Problems 62

2.6 Notes 72
2.6.1 The **Kiss** Generator 72
2.6.2 Quasi-Monte Carlo Methods 75
2.6.3 Mixture Representations 77

3 Monte Carlo Integration 79
3.1 Introduction 79
3.2 Classical Monte Carlo Integration 83
3.3 Importance Sampling 90
3.3.1 Principles 90
3.3.2 Finite Variance Estimators 94
3.3.3 Comparing Importance Sampling with Accept–Reject 103
3.4 Laplace Approximations 107
3.5 Problems 110
3.6 Notes 119
3.6.1 Large Deviations Techniques 119
3.6.2 The Saddlepoint Approximation 120

4 Controling Monte Carlo Variance 123
4.1 Monitoring Variation with the CLT 123
4.1.1 Univariate Monitoring 124
4.1.2 Multivariate Monitoring 128
4.2 Rao–Blackwellization 130
4.3 Riemann Approximations 134
4.4 Acceleration Methods 140
4.4.1 Antithetic Variables 140
4.4.2 Control Variates 145
4.5 Problems 147
4.6 Notes 153
4.6.1 Monitoring Importance Sampling Convergence 153
4.6.2 Accept–Reject with Loose Bounds 154
4.6.3 Partitioning 155

5 Monte Carlo Optimization 157
5.1 Introduction 157
5.2 Stochastic Exploration 159
5.2.1 A Basic Solution 159
5.2.2 Gradient Methods 162
5.2.3 Simulated Annealing 163
5.2.4 Prior Feedback 169
5.3 Stochastic Approximation 174
5.3.1 Missing Data Models and Demarginalization 174
5.3.2 The EM Algorithm 176
5.3.3 Monte Carlo EM 183
5.3.4 EM Standard Errors 186

5.4 Problems 188
5.5 Notes 200
5.5.1 Variations on EM 200
5.5.2 Neural Networks 201
5.5.3 The Robbins–Monro procedure 201
5.5.4 Monte Carlo Approximation 203

6 Markov Chains 205
6.1 Essentials for MCMC 206
6.2 Basic Notions 208
6.3 Irreducibility, Atoms, and Small Sets 213
6.3.1 Irreducibility 213
6.3.2 Atoms and Small Sets 214
6.3.3 Cycles and Aperiodicity 217
6.4 Transience and Recurrence 218
6.4.1 Classification of Irreducible Chains 218
6.4.2 Criteria for Recurrence 221
6.4.3 Harris Recurrence 221
6.5 Invariant Measures 223
6.5.1 Stationary Chains 223
6.5.2 Kac's Theorem 224
6.5.3 Reversibility and the Detailed Balance Condition 229
6.6 Ergodicity and Convergence 231
6.6.1 Ergodicity 231
6.6.2 Geometric Convergence 236
6.6.3 Uniform Ergodicity 237
6.7 Limit Theorems 238
6.7.1 Ergodic Theorems 240
6.7.2 Central Limit Theorems 242
6.8 Problems 247
6.9 Notes 258
6.9.1 Drift Conditions 258
6.9.2 Eaton's Admissibility Condition 262
6.9.3 Alternative Convergence Conditions 263
6.9.4 Mixing Conditions and Central Limit Theorems 263
6.9.5 Covariance in Markov Chains 265

7 The Metropolis–Hastings Algorithm 267
7.1 The MCMC Principle 267
7.2 Monte Carlo Methods Based on Markov Chains 269
7.3 The Metropolis–Hastings algorithm 270
7.3.1 Definition 270
7.3.2 Convergence Properties 272
7.4 The Independent Metropolis–Hastings Algorithm 276
7.4.1 Fixed Proposals 276

7.4.2 A Metropolis–Hastings Version of ARS 285
7.5 Random Walks 287
7.6 Optimization and Control 292
7.6.1 Optimizing the Acceptance Rate 292
7.6.2 Conditioning and Accelerations 295
7.6.3 Adaptive Schemes 299
7.7 Problems 302
7.8 Notes 313
7.8.1 Background of the Metropolis Algorithm 313
7.8.2 Geometric Convergence of Metropolis–Hastings Algorithms 315
7.8.3 A Reinterpretation of Simulated Annealing 315
7.8.4 Reference Acceptance Rates 316
7.8.5 Langevin Algorithms 318

8 The Slice Sampler 321
8.1 Another Look at the Fundamental Theorem 321
8.2 The General Slice Sampler 326
8.3 Convergence Properties of the Slice Sampler 329
8.4 Problems 333
8.5 Notes 335
8.5.1 Dealing with Difficult Slices 335

9 The Two-Stage Gibbs Sampler 337
9.1 A General Class of Two-Stage Algorithms 337
9.1.1 From Slice Sampling to Gibbs Sampling 337
9.1.2 Definition 339
9.1.3 Back to the Slice Sampler 343
9.1.4 The Hammersley–Clifford Theorem 343
9.2 Fundamental Properties 344
9.2.1 Probabilistic Structures 344
9.2.2 Reversible and Interleaving Chains 349
9.2.3 The Duality Principle 351
9.3 Monotone Covariance and Rao–Blackwellization 354
9.4 The EM–Gibbs Connection 357
9.5 Transition 360
9.6 Problems 360
9.7 Notes 366
9.7.1 Inference for Mixtures 366
9.7.2 ARCH Models 368

10 The Multi-Stage Gibbs Sampler 371
10.1 Basic Derivations 371
10.1.1 Definition 371
10.1.2 Completion 373

10.1.3 The General Hammersley–Clifford Theorem 376
10.2 Theoretical Justifications 378
10.2.1 Markov Properties of the Gibbs Sampler 378
10.2.2 Gibbs Sampling as Metropolis–Hastings 381
10.2.3 Hierarchical Structures 383
10.3 Hybrid Gibbs Samplers 387
10.3.1 Comparison with Metropolis–Hastings Algorithms 387
10.3.2 Mixtures and Cycles 388
10.3.3 Metropolizing the Gibbs Sampler 392
10.4 Statistical Considerations 396
10.4.1 Reparameterization 396
10.4.2 Rao-Blackwellization 402
10.4.3 Improper Priors 403
10.5 Problems 407
10.6 Notes 419
10.6.1 A Bit of Background 419
10.6.2 The BUGS Software 420
10.6.3 Nonparametric Mixtures 420
10.6.4 Graphical Models 422

11 Variable Dimension Models and Reversible Jump Algorithms 425
11.1 Variable Dimension Models 425
11.1.1 Bayesian Model Choice 426
11.1.2 Difficulties in Model Choice 427
11.2 Reversible Jump Algorithms 429
11.2.1 Green's Algorithm 429
11.2.2 A Fixed Dimension Reassessment 432
11.2.3 The Practice of Reversible Jump MCMC 433
11.3 Alternatives to Reversible Jump MCMC 444
11.3.1 Saturation 444
11.3.2 Continuous-Time Jump Processes 446
11.4 Problems 449
11.5 Notes 458
11.5.1 Occam's Razor 458

12 Diagnosing Convergence 459
12.1 Stopping the Chain 459
12.1.1 Convergence Criteria 461
12.1.2 Multiple Chains 464
12.1.3 Monitoring Reconsidered 465
12.2 Monitoring Convergence to the Stationary Distribution 465
12.2.1 A First Illustration 465
12.2.2 Nonparametric Tests of Stationarity 466
12.2.3 Renewal Methods 470

12.2.4 Missing Mass 474
12.2.5 Distance Evaluations 478
12.3 Monitoring Convergence of Averages 480
12.3.1 A First Illustration 480
12.3.2 Multiple Estimates 483
12.3.3 Renewal Theory 490
12.3.4 Within and Between Variances 497
12.3.5 Effective Sample Size 499
12.4 Simultaneous Monitoring 500
12.4.1 Binary Control 500
12.4.2 Valid Discretization 503
12.5 Problems 504
12.6 Notes 508
12.6.1 Spectral Analysis 508
12.6.2 The CODA Software 509

13 Perfect Sampling 511
13.1 Introduction 511
13.2 Coupling from the Past 513
13.2.1 Random Mappings and Coupling 513
13.2.2 Propp and Wilson's Algorithm 516
13.2.3 Monotonicity and Envelopes 518
13.2.4 Continuous States Spaces 523
13.2.5 Perfect Slice Sampling 526
13.2.6 Perfect Sampling via Automatic Coupling 530
13.3 Forward Coupling 532
13.4 Perfect Sampling in Practice 535
13.5 Problems 536
13.6 Notes 539
13.6.1 History 539
13.6.2 Perfect Sampling and Tempering 540

14 Iterated and Sequential Importance Sampling 545
14.1 Introduction 545
14.2 Generalized Importance Sampling 546
14.3 Particle Systems 547
14.3.1 Sequential Monte Carlo 547
14.3.2 Hidden Markov Models 549
14.3.3 Weight Degeneracy 551
14.3.4 Particle Filters 552
14.3.5 Sampling Strategies 554
14.3.6 Fighting the Degeneracy 556
14.3.7 Convergence of Particle Systems 558
14.4 Population Monte Carlo 559
14.4.1 Sample Simulation 560

14.4.2 General Iterative Importance Sampling . . . 560
14.4.3 Population Monte Carlo . . . 562
14.4.4 An Illustration for the Mixture Model . . . 563
14.4.5 Adaptativity in Sequential Algorithms . . . 565
14.5 Problems . . . 570
14.6 Notes . . . 577
14.6.1 A Brief History of Particle Systems . . . 577
14.6.2 Dynamic Importance Sampling . . . 577
14.6.3 Hidden Markov Models . . . 579

A Probability Distributions . . . 581

B Notation . . . 585
B.1 Mathematical . . . 585
B.2 Probability . . . 586
B.3 Distributions . . . 586
B.4 Markov Chains . . . 587
B.5 Statistics . . . 588
B.6 Algorithms . . . 588

References . . . 591

Index of Names . . . 623

Index of Subjects . . . 631