

Use R!

Jim Albert

Bayesian Computation with R

Second Edition

There has been a dramatic growth in the development and application of Bayesian inferential methods. Some of this growth is due to the availability of powerful simulation-based algorithms to summarize posterior distributions. There has been also a growing interest in the use of the system R for statistical analyses. R's open source nature, free availability, and large number of contributor packages have made R the software of choice for many statisticians in education and industry.

Bayesian Computation with R introduces Bayesian modeling by the use of computation using the R language. The early chapters present the basic tenets of Bayesian thinking by use of familiar one and two-parameter inferential problems. Bayesian computational methods such as Laplace's method, rejection sampling, and the SIR algorithm are illustrated in the context of a random effects model. The construction and implementation of Markov Chain Monte Carlo (MCMC) methods is introduced. These simulation-based algorithms are implemented for a variety of Bayesian applications such as normal and binary response regression, hierarchical modeling, order-restricted inference, and robust modeling. Algorithms written in R are used to develop Bayesian tests and assess Bayesian models by use of the posterior predictive distribution. The use of R to interface with WinBUGS, a popular MCMC computing language, is described with several illustrative examples.

This book is a suitable companion book for an introductory course on Bayesian methods and is valuable to the statistical practitioner who wishes to learn more about the R language and Bayesian methodology. The LearnBayes package, written by the author and available from the CRAN website, contains all of the R functions described in the book.

The second edition contains several new topics such as the use of mixtures of conjugate priors and the use of Zellner's g priors to choose between models in linear regression. There are more illustrations of the construction of informative prior distributions, such as the use of conditional means priors and multivariate normal priors in binary regressions. The new edition contains changes in the R code illustrations according to the latest edition of the LearnBayes package.

Jim Albert is Professor of Statistics at Bowling Green State University. He is Fellow of the American Statistical Association and is past editor of *The American Statistician*. His books include *Ordinal Data Modeling* (with Valen Johnson), *Workshop Statistics: Discovery with Data*, *A Bayesian Approach* (with Allan Rossman), and *Bayesian Computation using Minitab*.

STATISTICS

ISBN 978-0-387-92297-3



springer.com



1	An Introduction to R	1
1.1	Overview	1
1.2	Exploring a Student Dataset	1
1.2.1	Introduction to the Dataset	1
1.2.2	Reading the Data into R	2
1.2.3	R Commands to Summarize and Graph a Single Batch	2
1.2.4	R Commands to Compare Batches	5
1.2.5	R Commands for Studying Relationships	6
1.3	Exploring the Robustness of the t Statistic	8
1.3.1	Introduction	8
1.3.2	Writing a Function to Compute the t Statistic	9
1.3.3	Programming a Monte Carlo Simulation	10
1.3.4	The Behavior of the True Significance Level Under Different Assumptions	11
1.4	Further Reading	13
1.5	Summary of R Functions	14
1.6	Exercises	15
2	Introduction to Bayesian Thinking	19
2.1	Introduction	19
2.2	Learning About the Proportion of Heavy Sleepers	19
2.3	Using a Discrete Prior	20
2.4	Using a Beta Prior	22
2.5	Using a Histogram Prior	26
2.6	Prediction	28
2.7	Further Reading	34
2.8	Summary of R Functions	34
2.9	Exercises	35

3	Single-Parameter Models	39
3.1	Introduction	39
3.2	Normal Distribution with Known Mean but Unknown Variance	39
3.3	Estimating a Heart Transplant Mortality Rate	41
3.4	An Illustration of Bayesian Robustness	44
3.5	Mixtures of Conjugate Priors	49
3.6	A Bayesian Test of the Fairness of a Coin	52
3.7	Further Reading	57
3.8	Summary of R Functions	57
3.9	Exercises	58
4	Multiparameter Models	63
4.1	Introduction	63
4.2	Normal Data with Both Parameters Unknown	63
4.3	A Multinomial Model	66
4.4	A Bioassay Experiment	69
4.5	Comparing Two Proportions	75
4.6	Further Reading	80
4.7	Summary of R Functions	80
4.8	Exercises	81
5	Introduction to Bayesian Computation	87
5.1	Introduction	87
5.2	Computing Integrals	88
5.3	Setting Up a Problem in R	89
5.4	A Beta-Binomial Model for Overdispersion	90
5.5	Approximations Based on Posterior Modes	94
5.6	The Example	95
5.7	Monte Carlo Method for Computing Integrals	97
5.8	Rejection Sampling	98
5.9	Importance Sampling	101
5.9.1	Introduction	101
5.9.2	Using a Multivariate t as a Proposal Density	103
5.10	Sampling Importance Resampling	105
5.11	Further Reading	108
5.12	Summary of R Functions	109
5.13	Exercises	110
6	Markov Chain Monte Carlo Methods	117
6.1	Introduction	117
6.2	Introduction to Discrete Markov Chains	117
6.3	Metropolis-Hastings Algorithms	120
6.4	Gibbs Sampling	122
6.5	MCMC Output Analysis	122

6.6	A Strategy in Bayesian Computing	124
6.7	Learning About a Normal Population from Grouped Data	124
6.8	Example of Output Analysis	129
6.9	Modeling Data with Cauchy Errors	131
6.10	Analysis of the Stanford Heart Transplant Data	140
6.11	Further Reading.....	145
6.12	Summary of R Functions	146
6.13	Exercises	147
7	Hierarchical Modeling	153
7.1	Introduction	153
7.2	Three Examples	153
7.3	Individual and Combined Estimates	155
7.4	Equal Mortality Rates?	157
7.5	Modeling a Prior Belief of Exchangeability.....	161
7.6	Posterior Distribution	163
7.7	Simulating from the Posterior	163
7.8	Posterior Inferences	168
7.8.1	Shrinkage	168
7.8.2	Comparing Hospitals	169
7.9	Bayesian Sensitivity Analysis	171
7.10	Posterior Predictive Model Checking	173
7.11	Further Reading.....	175
7.12	Summary of R Functions	175
7.13	Exercises	176
8	Model Comparison	181
8.1	Introduction	181
8.2	Comparison of Hypotheses.....	181
8.3	A One-Sided Test of a Normal Mean.....	182
8.4	A Two-Sided Test of a Normal Mean	185
8.5	Comparing Two Models	186
8.6	Models for Soccer Goals	187
8.7	Is a Baseball Hitter Really Streaky?	190
8.8	A Test of Independence in a Two-Way Contingency Table	194
8.9	Further Reading.....	199
8.10	Summary of R Functions	199
8.11	Exercises	201
9	Regression Models.....	205
9.1	Introduction	205
9.2	Normal Linear Regression	205
9.2.1	The Model	205
9.2.2	The Posterior Distribution.....	206
9.2.3	Prediction of Future Observations	206

9.2.4	Computation	207
9.2.5	Model Checking	207
9.2.6	An Example	208
9.3	Model Selection Using Zellner's g Prior	217
9.4	Survival Modeling	222
9.5	Further Reading	227
9.6	Summary of R Functions	227
9.7	Exercises	229
10	Gibbs Sampling	235
10.1	Introduction	235
10.2	Robust Modeling	236
10.3	Binary Response Regression with a Probit Link	240
10.3.1	Missing Data and Gibbs Sampling	240
10.3.2	Proper Priors and Model Selection	243
10.4	Estimating a Table of Means	248
10.4.1	Introduction	248
10.4.2	A Flat Prior Over the Restricted Space	250
10.4.3	A Hierarchical Regression Prior	254
10.4.4	Predicting the Success of Future Students	259
10.5	Further Reading	260
10.6	Summary of R Functions	260
10.7	Exercises	261
11	Using R to Interface with WinBUGS	265
11.1	Introduction to WinBUGS	265
11.2	An R Interface to WinBUGS	266
11.3	MCMC Diagnostics Using the coda Package	267
11.4	A Change-Point Model	268
11.5	A Robust Regression Model	272
11.6	Estimating Career Trajectories	276
11.7	Further Reading	281
11.8	Exercises	282
	References	287
	Index	293