Journal of Statistical Software

July 2017, Volume 79, Book Review 1. *doi: 10.18637/jss.v079.b01*

Reviewer: Hakan Demirtas
University of Illinois at Chicago

Bayesian Computation with R (2nd Edition)

Jim Albert
Springer-Verlag, New York, 2009.
ISBN 978-0387922973978. xii + 300 pp. USD 59.95 (P).
http://www.springer.com/9780387922973

The new material added in the 2nd edition include the mixture of conjugate priors, the use of the SIR (sampling importance resampling) to explore the sensitivity of Bayesian inferences with respect to changes in the prior, and Zellner's g priors to choose between models in linear regression. Chapter 1 is allocated to basic data manipulations and rudimentary analysis in R. The highlight of the chapter is a simulation study about the true significance levels computed by Monte Carlo experiments. Chapter 2 introduces the basic elements of the Bayesian inferential approach. Forming the posterior distribution via prior distribution and likelihood, obtaining summaries of this probability distribution as well as predicting the likely outcomes of a new sample are described in an intuitively appealing manner. Chapters 3 and 4 illustrate the use of R for Bayesian inference for standard a few one- and two-parameter models, respectively. These chapters discuss different types of priors and posterior distributions to perform inferences, and the use of predictive distributions. The conjugate priors-posteriors as well as the posteriors that do not have well-known functional forms, requiring the calculations to be done on a grid of points (by brute-force) to elicit meaningful summaries are covered.

Chapter 5 and 6 are the crux of the book, they are concerned with an introduction to modern Bayesian computing. Chapter 5 discusses some of the more complicated computational methods that are employed for the remainder of the book. It covers a multivariate normal approximation to the posterior that serves as a good first approximation in the development of more exact methods. It then provides a general introduction to the use of simulation in computing summaries of the posterior distribution. It describes the rejection sampling algorithm in cases where the posterior distribution does not have a standard functional form. Finally, it presents the importance sampling and SIR algorithms for computing integrals and simulating from a general posterior distribution. Chapter 6 makes readers familiar with the salient ideas of Markov chain Monte Carlo (MCMC) methods for the approximation of the posterior distribution of interest, describes the use of MCMC algorithms in the discrete state space situations, and covers the Metropolis-Hastings algorithms and its two major variants (random walk and independence chains). The popular Gibbs sampling algorithm is also discussed in this chapter. Last but not least, identification of the convergence behavior via graphical tools in MCMC output analysis is explained.

The remaining chapters illustrate the utility of the major computational algorithms for a variety of Bayesian applications. Chapter 7 introduces the use of exchangeable models in the simultaneous estimation of a set of Poisson rates under the hierarchical modeling framework. Chapter 8 characterizes Bayesian tests of simple hypotheses and the use of Bayes factors for model comparisons. Chapter 9 is about Bayesian regression models. Chapter 10 describes several applications such as robust modeling, probit regression, and Gibbs sampling in the presence of missing data. Chapter 11 delineates the use of R to interface with **WinBUGS**.

There are 72 end of chapter exercises, most of which are helpful (some are a little confusing though) to better comprehend Bayesian statistics and applications. At the end of every chapter, a summary of the relevant functions is given, a useful feature for self-learners. The book could be used as a textbook depending on the level. It is definitely not a reference book. However, some functions can be used in a wide range of domains (not only in Bayesian ones). Examples include finding the two shape parameters under the beta distribution assumption, implementing a general-purpose rejection sampling algorithm, the standardization of a distribution so it becomes a legitimate probability density function, the construction of contour plots to determine a valid range, among other things.

It is a noteworthy companion book for an introductory course on Bayesian methods and computing if expectations are set humbly and realistically. The intended audience is quantitatively oriented scientists who seek an entry level exposure to the Bayesian paradigm and want to develop some competence in writing short scripts to define a Bayesian model, using functions to summarize posterior distributions and simulate from them, and constructing graphs to illustrate the posterior inference. Although it is not explicitly stated in the text, typical readers might be late undergraduate or early graduate stage statistics or biostatistics students, practitioners, researchers in social and biomedical sciences, and methodologists in quantitative disciplines, more generally a broad range of scientists who want to understand the basic aspects of Bayesian thinking and computing that involves programming in R.

The book is designed around the fairly popular R package **LearnBayes** which is a software implementation that includes data sets and functions (scripts). The package is utilized throughout the book for illustrative examples. Here are some pivotal points (positive and negative) in no particular order:

- The quality of programming does not feel super professional but is acceptable. It could have been better though, for example the object names could have been chosen more carefully. On the bright side, overall functionality is satisfactory in the sense that one can employ the package functions in many different contexts.

- The writing flow and the sequential structure are generally proper. However, the text and its transitions across topics are choppy at times, nothing is terrible though. Some notation is suboptimal and incomplete, but that is the price to pay for simplification. Some precision is naturally lost when the concentration is on applications and computing-related aspects, not on theory.

- Examples are more or less fine, but some of them are of limited applicability. One often gets the impression that some illustrations are a bit blackboxish. What readers can get out of the book depends on their needs. If they are interested in developing solid comprehension on the underlying theory, it is not the right book. However, if they want to implement some prominent Bayesian techniques, along with a conceptual and computational understanding of the key points, it is insightful. Neither occasional blackboxishness nor notational suboptimalities nor less-than-perfect programming reduces its practical value and intellectual merit. Given the introductory nature of the material, these shortcomings are tolerable, because the goal of this work is to make readers familiar with Bayesian computing without overwhelming them with technical details. Inaccuracies can occur when technical sophistication is avoided.

- Bayesianism is fundamentally concerned with basing statistical inferences on the posterior distribution which is an updated version of a priori beliefs by what the data suggest in the form of likelihood. Although it is easy to grasp this conceptually, what it really means in practice may be mysterious to people who lack statistical background. The text adequately and convincingly explains the rationale and operational characteristics of the Bayesian view.

In a nutshell, is it a good companion book on Bayesian computing? Yes. Do I recommend it? Yes. Should readers have high expectations? Not really. Is it suitable for self-learning? To some extent. It is a nice supplementary book for beginners in statistics and biostatistics as well as for applied people from other disciplines whose primary motivation is to build some proficiency in introductory Bayesianism and intermediate-level statistical programming. Nothing more, nothing less. As long as it is regarded as a complementary text to a heavy-hitter book on Bayesian statistics, it serves a purpose and is advisable.

Reviewer:

Hakan Demirtas
Division of Epidemiology and Biostatistics (MC 923)
University of Illinois at Chicago, School of Public Health
1603 West Taylor Street, Room 950
Chicago, IL, 60612-4336, United States of America
E-mail: demirtas@uic.edu
URL: http://www.healthstats.org/members/hdemirtas.html

Journal of Statistical Software http://www.jstatsoft.org/
published by the Foundation for Open Access Statistics http://www.foastat.org/
July 2017, Volume 79, Book Review 1 *Published:* 2017-07-22
doi:10.18637/jss.v079.b01