Journal of Statistical Software

January 2013, Volume 52, Book Review 1. *http://www.jstatsoft.org/*

Reviewer: Joakim Ekström
University of California, Los Angeles

The BUGS Book: A Practical Introduction to Bayesian Analysis

> David Lunn, Christopher Jackson, Nicky Best, Andrew Thomas, David Spiegelhalter
> Chapman & Hall/CRC, Boca Raton, FL, 2012.
> ISBN 978-1-58488-849-9. 399 pp. USD 49.95 (P).
> http://www.crcpress.com/product/isbn/9781584888499

BUGS is a software project that began in Cambridge, England, and has been actively ongoing for more than twenty years. Its core purpose is to provide a computational companion to Bayesian statistical analyses, and the software makes extensive use of Markov chain Monte Carlo and similar simulation methods. At present the software exists in a number of versions, such as **WinBUGS**, **OpenBUGS**, and **JAGS**, which are available free of charge. The authors of *The **BUGS** Book* are long-standing contributors to the **BUGS** project.

*The **BUGS** Book* has a proper title that alludes to a software manual, and a subtitle, *A Practical Introduction to Bayesian Analysis*, that suggests an introductory Bayesian analysis textbook. As such, the title promises a two-in-one product that provides the reader with both a **BUGS** manual and a Bayesian analysis textbook, a combination that will likely appeal to many potential readers. Books of this kind have become fairly popular in recent years, and besides the obvious convenience a two-in-one product provides, the format holds potential for a number of valuable synergies. In particular, the contents of the chapters can be illustrated through examples with data sets that are bundled to the software, making it easy for readers to replicate the computations themselves, possibly under a variety of assumptions.

At the core, the book discusses estimation of probability distributions of random variables. A random variable, which is a measurable function that maps a probability space into a topological space, cannot be observed directly, but the values it attains, and their distribution, can. The empirical measure is an unbiased and consistent estimator of the probability distribution, and its empirical process is asymptotically normal. If the random variable attains values on the real line, then the probability distribution can be represented by its distribution function, and the empirical distribution function has even stronger asymptotical properties (see, e.g., Dudley 1999).

If the analyst wants to estimate a continuous distribution, a pragmatic method is to smooth the empirical distribution through a mixture of spherical distributions translated by the observed values. If the analyst wants to restrict the set of probability distributions to some subset, the distribution function can given an empirical distribution function be estimated through a basic regression analysis, approximating the error term distributions through simulation or through Donsker's theorem. A simple alternative that works well in some instances but has less than perfect theoretical properties is the method of moments.

The approach of *The **BUGS** Book* is more complicated. The idea is to exploit statistical dependence between the probability distribution of a random variable and the observed value, or sample point. Consider the pair $(\mathcal{L}(X), X(\omega))$ representing the probability distribution of the random variable X and the value of X, respectively, and the expression $\Pr(\mathcal{L}(X) = P|X(\omega) = x)$, where P denotes some probability distribution and x the sample point. In accordance with Bernoulli's fifth axiom, one can then choose as estimate the probability distribution that seems most probable.

The expression $\Pr(\mathcal{L}(X) = P|X(\omega) = x)$, while intuitive, is difficult to evaluate mathematically. Even if a joint distribution is specified, the expression is typically identically zero, making it useless for hypothesis generation purposes. The most common work-around is to apply the Edgeworth-Fisher criterion, after Edgeworth (1883) and Fisher (1912): Suppose P and Q are two probability distributions with density functions f and g, respectively, then if $f(x) > g(x)$ it is deemed that $\Pr(\mathcal{L}(X) = P|X(\omega) = x) > \Pr(\mathcal{L}(X) = Q|X(\omega) = x)$.

Now, the approach of *The **BUGS** Book* is even more complicated. Consider an indexed subset of probability distributions $\{P_\alpha\}_{\alpha \in A}$ with index set A. The index, α, is considered a random variable attaining values in the index set, equipped with some topology, and the end goal of the analysis is to estimate the probability distribution $\mathcal{L}(\alpha)$. Since the sample point, x, is only indirectly related to α, estimation of $\mathcal{L}(\alpha)$ is by construction not possible without, and hence entirely reliant upon, a system of assumptions.

The reliance on assumptions is both the strength and the weakness of this method for estimation of probability distribution. The use of assumptions provides the analyst with ample flexibility, and an ability to produce conclusions in situations where it would otherwise not be possible. An example of the latter is when the sample is not representative of the population, for instance due to the data having been censored, truncated or manipulated. By the same token, the leeway that is afforded to the analyst also makes reverse-engineering of conclusions simple; thus presenting the scientific community with a rather problematic situation.

As an introductory Bayesian analysis textbook, *The **BUGS** Book* is an informal and light read, but is also brief and imprecise. The material is to a significant extent communicated through examples; some sections consist of one or two sentences paired with an illustrative example. The phrase "under broad conditions," is used freely; for instance the likelihood estimator is under broad conditions asymptotically normal (p. 52), and the Bayesian inferences are under broad conditions asymptotically identical to those obtained from a likelihood perspective (p. 53). Spelling out the conditions, or at least providing a reference, seems like a simple way of adding value to the book.

Moreover, there is a general tendency of presenting the best case scenario as being the typical case. As a specific example, the conditions required for Cramér's theorem for asymptotic normality of the likelihood estimator (see, e.g., Ferguson 1996, Theorem 18), whether they are broad or not, are certainly not trivial, and further its conclusion does not even imply asymptotic normality of the estimate. While *The **BUGS** Book* presents the methods of the **BUGS** software in a favorable light, there is also an evident risk of the reader receiving an unbalanced, overly simplistic, and thus inaccurate, view of the methods. A habit of frequently writing important words, terms and phrases within quotation marks adds to the sense of a generally incautious treatment of the material.

The strength of *The **BUGS** Book* is its rich collection of ambitiously constructed and thematically arranged examples, which often come with snippets of code and printouts, as well as illustrative plots and diagrams. The formatting is commendably clear, with examples typeset in Computer Modern sans serif, and snippets of code and printouts typeset in Computer Modern teletype. For readers that want to familiarize themselves with the **BUGS** software and gain a sense of its capabilities, working through the examples of *The **BUGS** Book* will provide a valuable, time-efficient and enjoyable exercise. The examples lend themselves well for incorporation into a teaching environment.

As a software manual, *The **BUGS** Book* provides, in addition to its many examples, one chapter and three appendices dedicated to discussion of features, functions and syntaxes of the **BUGS** software. While most of the material is available in online **BUGS** manuals, *The **BUGS** Book* contains some additional clarifications and explanatory wordings. For example, short discussions on the nature of the probability distributions implemented in **BUGS** are provided. Moreover, many readers feel more comfortable obtaining information through reading a book, rather than scrolling through an electronic document.

All in all, *The **BUGS** Book* has strengths and weaknesses. The numerous ambitiously constructed examples with diagrams and snippets of code are unquestionably of great value to many readers seeking to familiarize themselves with **BUGS** and its capabilities. The introductory text on Bayesian analysis does not meet the standards of the scientific community in terms of clarity and accuracy, and should arguably not be referenced in scientific works. As a result, *The **BUGS** Book* is in view of the two-in-one product that its title promises not entirely successful. For future editions the authors should consider the possibilities of improving the introductory text on Bayesian analysis; or removing it altogether, leaving a high-quality introductory software manual richly illustrated with thematically arranged examples.

References

Dudley RM (1999). *Uniform Central Limit Theorems.* Cambridge University Press, New York.

Edgeworth FY (1883). "The Method of Least Squares." *Philosophical Magazine (5th series)*, **16**, 360–375.

Ferguson TS (1996). *A Course in Large Sample Theory.* Chapman & Hall, New York.

Fisher RA (1912). "On an Absolute Criterion for Fitting Frequency Curves." *Messenger of Mathematics*, **41**, 155–160.

Reviewer:

Joakim Ekström
Department of Statistics
University of California, Los Angeles

Los Angeles, CA, 90025-1554, United States of America
E-mail: joakim.ekstrom@stat.ucla.edu

Journal of Statistical Software http://www.jstatsoft.org/
published by the American Statistical Association http://www.amstat.org/
Volume 52, Book Review 1 *Published:* 2013-01-07
January 2013