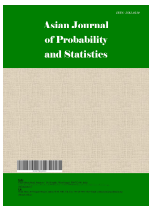

Asian Journal of Probability and Statistics

9(3): 1-4, 2020; Article no.AJPAS.61337
ISSN: 2582-0230

Scholarly Book Review on Bayesian Statistics for Beginners: A Step-by-Step Approach

Eahsan Shahriary[1*] and Amir Hajibabaee[2]

[1]Environmental Science and Engineering Program, University of Texas at El Paso, El Paso, Texas 79968, USA.
[2]Research Engineer, Research & Development Department, Ozinga Ready Mix Concrete, Inc., 2222 S Lumber Street, Chicago, IL 60616, USA.

Authors' contributions

This work was carried out in collaboration between both authors. Authors ES and AH contributed to the book selection, review and writing of the scholarly book review. Both authors read and approved the final manuscript.

Article Information

DOI: 10.9734/AJPAS/2020/v9i330226
Editor(s):
(1) Dr. Manuel Alberto M. Ferreira, Lisbon University, Portugal.
Reviewers:
(1) Zarai Khaireddine, Sciences of Tunis, Tunisia.
(2) Balaji L, Vel Tech Rangarajan Dr. Sagunthala R&D Institute of Science and Technology, India.
(3) Zahraddeen Sufyanu, Federal University Dutse, Nigeria.
(4) Sujata Suvarnapathaki, Ramnarain Ruia Autonomous College, India.
Complete Peer review History: http://www.sdiarticle4.com/review-history/61337

Scholarly Book Article

Received: 25 July 2020
Accepted: 02 October 2020
Published: 25 October 2020

Abstract

This book offers the students and researchers a unique introduction to Bayesian statistics. Authors provide a wonderful journey in the realm of Bayesian Probability and aspire readers to become Bayesian statisticians. The book starts with Introduction to Probability and covers Bayes' Theorem, Probability Mass Functions, Probability Density Functions, The Beta-Binomial Conjugate, Markov chain Monte Carlo (MCMC), and Metropolis-Hastings Algorithm. The book is very well written, and topics are very to the point with real-world applications but does not provide examples for computing using common open-source software.

Keywords: Bayesian probability; PDF; PMF; Gibbs sampling; MCMC.

**Corresponding author: E-mail: eshahriary@miners.utep.edu;*

1 Review

Most of the statistics and probability textbooks for biological and medical sciences try to simplify the complex mathematical aspects of probability, especially Bayesian probability [1,2]. As John Ollason, Scottish ecologist, says, "many of our biology students are refugees from high-school mathematics." These books are similar in starting from basic concepts to advanced and, in general, the difficulty of contents and directions of materials please biology, environmental, engineering, medical, geology, ecology, and wildlife management undergraduate/graduate students as well as researchers [3,4,5,6]. Advanced statistical topics such as Bayesian structural equation modeling, Bayesian network structure learning, nonparametric Bayesian inference, Bayesian reliability, and Bayesian bioinformatics are discussed in special textbooks and handbooks [7,8,9].

Bayesian Statistics for Beginners: A Step-by-Step Approach is one of these, spanning from basics to advanced topics [10]. As the title of the book shows, the authors try to explain and start from the most beginning concepts and move toward more complex topics than commonly found in the mathematical statistics and probability textbooks.

Bayesian Statistics for Beginners: A Step-by-Step Approach has six sections and twenty chapters with overviews, questions, problems, and answers, and five appendices namely, (Beta-Binomial conjugate solution, the Gamma-Poisson conjugate solution, the Normal-Normal conjugate solution, conjugate solutions for simple linear regression, and the standardization of regression data) [10].

Chapter 1, "*Introduction to Probability*" introduces the concept of probability and basic terms in probability such as sample space, outcome, discrete outcome, event, probability, probability distribution, uniform distribution, trial, empirical distribution, and the law of large numbers. Chapter 2, "*Joint, Marginal, and Conditional Probability"* builds an understanding of probability vocabularies, Venn diagram, marginal probability, joint probability, independent events, dependent events, and conditional probability.

Chapter 3, "*Bayes' Theorem*" introduces its derivation and describes the relationship between two inverse conditional probabilities, P(A|B), and P(B|A), degree of belief and interpretation. Chapter 4, "*Bayesian Inference*" provides a complete discussion of Bayesian inference. This chapter covers Bayesian inference, induction, deduction, hypothesis, alternative hypotheses, the prior probability of a hypothesis, likelihood of the observed data, and posterior probability of a hypothesis. Chapter 5, "*The Author Problem: Bayesian Inference with Two Hypotheses*" provides an example of Bayesian inference by using Frederick Mosteller and David Wallace's article about disputed authorship of the Federalist Papers. Bayesian analysis of the unsigned papers by James Madison or Alexander Hamilton shows how Bayesian inference works. Authors in this chapter cover Bayesian inference, hypothesis, alternative hypothesis, the prior probability of a hypothesis, prior probability distribution, the likelihood of the observed data, the posterior probability of a hypothesis, and posterior probability distribution. Chapter 6, "*The Birthday Problem: Bayesian Inference with Multiple Discrete Hypotheses*" describes the prior distribution and concept of likelihood. The authors explain informative prior distribution, non-informative prior distribution, objective priors, subjective priors, and prior sensitivity analysis. In Chapter 7, "*The Portrait Problem: Bayesian Inference with Joint Likelihood*" the combination of data in the Bayesian inference framework is explained. The joint likelihood is the key concept of this chapter.

The probability distribution is discussed thoroughly in two chapters. Chapter 8 "*Probability Mass Functions*" provides a good introduction to probability distribution and Bernoulli and binomial distributions. It introduces function, random variable, probability distribution, parameter, probability mass function (pmf), Binomial pmf, Bernoulli pmf, likelihood, and likelihood profile. In Chapter 9 "*Probability Density Functions*" the authors discuss continuous random variables and probability density functions (pdfs) with examples of uniform and normal distributions. This chapter covers random variables, probability distribution, parameter, probability density, likelihood, and likelihood profile.

Chapters 10,11 and 12 discuss three Bayesian conjugates and parameter estimation. In Chapter 10, "*The White House Problem: The Beta-Binomial Conjugate*" the authors discuss a problem with beta distribution and hypotheses. They explain Beta distribution, binomial data, hyperparameters, conjugate prior, and credible intervals. Chapter 11 "*The Shark Attack Problem: The Gamma-Poisson Conjugate*" covers the use of gamma prior distribution and shark attacks, providing a good base on the Poisson probability mass function, the gamma probability density function, and the gamma-Poisson conjugate. In Chapter 12, titled "*The Maple Syrup Problem: The Normal-Normal Conjugate,*" the authors use the example of maple syrup production in multiple years and show the estimation of μ and σ parameters of a normal distribution using the Bayesian method. The normal distribution pdf in terms of mean (μ), standard deviation (σ), and precision (τ) and the normal-normal conjugate are discussed here in chapter 12.

The theory and application of Markov chain Monte Carlo (MCMC) are discussed in four chapters. In Chapter 13, "*The Shark Attack Problem Revisited: MCMC with the Metropolis Algorithm,*" authors revisit the gamma-Poisson conjugate concept and compare the result of the estimation of the conjugate solution with the result of MCMC and that of the Metropolis algorithm. This chapter provides a good explanation of Monte Carlo, Markov chain, Metropolis algorithm, tuning parameter, MCMC inference, traceplot, and moment matching. Chapter 14 "*MCMC Diagnostic Approaches*" talks about MCMC diagnostic concepts; tuning, acceptance rate, burn in, thinning, and proposal issues that can be used on the posterior distribution.

Chapter 15 "*The White House Problem Revisited: MCMC with the Metropolis-Hastings Algorithm*" compares the results of posterior distribution estimation using the conjugate solution with MCMC and the Metropolis-Hastings algorithm. Monte Carlo, Markov chain, Metropolis-Hastings algorithm, Metropolis-Hastings proposal distribution, and Metropolis-Hastings correction factor are explained. Chapter 16 "*The Maple Syrup Problem Revisited: MCMC with Gibbs Sampling*" readdresses the normal-normal conjugate and estimates the unknown parameters τ and μ using MCMC with the Gibbs sampling.

In the final four chapters 17,18,19 and 20, the authors demonstrate the use of Bayes' theorem in solving important problems. Chapter 17, "*The Survivor Problem: Simple Linear Regression with MCMC*" discusses the estimation of the three parameters using linear regression and Gibbs sampling. This chapter covers linear equation, sums of squares, posterior predictive distribution, and linear regression with MCMC and Gibbs sampling. Chapter 18 "*The Survivor Problem Continued: Introduction to Bayesian Model Selection*" compares two models and introduces Bayesian model selection. Model assessment (model fit) and DIC (Deviance Information Criterion) are also described. In Chapter 19 "*The Lorax Problem: Introduction to Bayesian Networks,*" the authors write about Bayesian belief networks, Netica software, and directed acyclic graph (DAG). They explain the Bayesian network, directed acyclic graph, parent or root node, child node, influence diagram, conditional probability table (CPT), and chain rule for joint probability. The last chapter of the book, Chapter 20 "*The Once-ler Problem: Introduction to Decision Trees*" explains decision trees, and discusses the calculation of expected values, decision tree, decision node, chance node, payoff, utility, and linked decisions.

The book is designed and presented for students and researchers. The complete definition and explanation of each topic and questions, answers, and exercises convey the take-home message of each chapter and make the book one of the best on Bayesian probability. All chapters provide definitions and simple theoretical explanations of probability topics. Readers with no prior experience in probability and Bayes' theorem will find an informative milestone of related topics. Advanced readers also could find interesting topics in sections 5 (chapters 13-16) and 6 (chapters 17-20) and deepen their theoretical and pragmatic knowledge. A useful feature of the book is its appendices, which explain various prior distributions, conjugate solution, and linear regression for voracious readers who are looking for more detailed information. The topics are very to the point and the authors provide a good historical background along with information and photos of leading scientists. Interestingly, the authors provided hyperlinks accessed in addition to the bibliography.

A drawback is that this book does not provide examples for computing using common open-source software such as R and OpenBUGS. As the book title suggests, it is designed for beginners, but future editions should add appendices on introduction to common statistical software (installation, basic commands, and functions)

and solving examples and interpreting the outputs in each chapter with these packages. Knowledge of both the theory and topics of statistics alongside the appropriate software will give mettle to readers to become pragmatists and move from theoretical concepts to practical know-how in the field.

2 Conclusion

Altogether, the book is informative and recommended for undergraduate and graduate students and researchers who are interested to move from frequentist reasoning to a Bayesian approach. The book demonstrates why learning Bayesian statistics is essential. In brief, Bayesian statistics for beginners: A step-by-step approach is a good text for students and provides them a background for advanced Bayesian probability courses.

Competing Interests

Authors have declared that no competing interests exist.

References

[1] Aho KA. Foundational and applied statistics for biologists using R. CRC Press; 2013.

[2] Sullivan LM. Essentials of biostatistics in public health. Jones & Bartlett Learning; 2017.

[3] Lazic ZR. Design of experiments in chemical engineering: a practical guide. Wiley- VCH; 2004.

[4] Dytham C. Choosing and using statistics: A biologist's guide. Wiley-Blackwell; 2011.

[5] Montgomery DC, Runger GC. Applied statistics, and probability for engineers. Wiley; 2013.

[6] Hoffman JI. Biostatistics for medical and biomedical practitioners. Academic Press; 2015.

[7] Darwiche A. Modeling and reasoning with Bayesian networks. Cambridge university press; 2009.

[8] Gelman A, Carlin JB, Stern HS, Dunson DB, Vehtari A, and Rubin DB. Bayesian Data Analysis. Chapman & Hall/CRC; 2013.

[9] Reich BJ, Ghosh SK. Bayesian statistical methods. CRC Press; 2019.

[10] Donovan TM, Mickey RM. Bayesian statistics for beginners: A step-by-step approach. Oxford University Press; 2019.

Peer-review history:
The peer review history for this paper can be accessed here (Please copy paste the total link in your browser address bar)
http://www.sdiarticle4.com/review-history/61337