

Springer Texts in Statistics

Series Editors:

George Casella

Richard DeVeaux

Stephen E. Fienberg

Ingram Olkin

For further volumes:

<http://www.springer.com/series/417>

Jean-Michel Marin • Christian P. Robert

Bayesian Essentials with R

Second Edition

 Springer

Jean-Michel Marin
Université Montpellier 2
Montpellier, France

Christian P. Robert
Université Paris-Dauphine
Paris, France

ISSN 1431-875X ISSN 2197-4136 (electronic)
ISBN 978-1-4614-8686-2 ISBN 978-1-4614-8687-9 (eBook)
DOI 10.1007/978-1-4614-8687-9
Springer New York Heidelberg Dordrecht London

Library of Congress Control Number: 2013950378

© Springer Science+Business Media New York 2014

This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed. Exempted from this legal reservation are brief excerpts in connection with reviews or scholarly analysis or material supplied specifically for the purpose of being entered and executed on a computer system, for exclusive use by the purchaser of the work. Duplication of this publication or parts thereof is permitted only under the provisions of the Copyright Law of the Publisher's location, in its current version, and permission for use must always be obtained from Springer. Permissions for use may be obtained through RightsLink at the Copyright Clearance Center. Violations are liable to prosecution under the respective Copyright Law.

The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

While the advice and information in this book are believed to be true and accurate at the date of publication, neither the authors nor the editors nor the publisher can accept any legal responsibility for any errors or omissions that may be made. The publisher makes no warranty, express or implied, with respect to the material contained herein.

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

*To our most rewarding case studies,
Chloé & Lucas, Joachim & Rachel*

Preface

After that, it was down to attitude.
—Ian Rankin, *Black & Blue*.—

The purpose of this book is to provide a self-contained entry into practical and computational Bayesian statistics using generic examples from the most common models for a class duration of about seven blocks that roughly correspond to 13–15 weeks of teaching (with three hours of lectures per week), depending on the intended level and the prerequisites imposed on the students. (That estimate does not include practice—i.e., R programming labs, writing data reports—since those may have a variable duration, also depending on the students’ involvement and their programming abilities.) The emphasis on *practice* is a strong commitment of this book in that its primary audience consists of graduate students who need to use (Bayesian) statistics as a tool to analyze their experiments and/or datasets. The book should also appeal to scientists in all fields who want to engage into Bayesian statistics, given the versatility of the Bayesian tools. Bayesian essentials can also be used for a more classical statistics audience when aimed at teaching a quick entry to Bayesian statistics at the end of an undergraduate program, for instance. (Obviously, it can supplement another textbook on data analysis at the graduate level.)

This book is an extensive revision of our previous book, *Bayesian Core*, which appeared in 2007, aiming at the same goals. (Glancing at this earlier version will show the filiation to most readers.) However, after publishing *Bayesian Core* and teaching from it to different audiences, we soon realized that the level of mathematics therein was actually more involved than the one expected by those audiences. Students were also asking for more advice and

more R code than what was then available. We thus decided upon a major revision, producing a manual that cut the mathematics and expanded the R code, changing as well some chapters and replacing some datasets. We had at first even larger ambitions in terms of contents, but had eventually to sacrifice new chapters for the sake of completing the book before we came to blows! To stress further the changes from the 2007 version, we also decided on a new title, *Bayesian Essentials*, that was actually suggested by Andrew Gelman during a visit to Paris.

The current format of the book is one of a quick coverage of the topics, always backed by a motivated problem and a corresponding dataset (available in the associated R package, *bayess*), and a detailed resolution of the inference procedures pertaining to this problem, always including commented R programs or relevant parts of R programs. Special attention is paid to the derivation of prior distributions, and operational reference solutions are proposed for each model under study. Additional cases are proposed as exercises. The spirit is not unrelated to that of Nolan and Speed (2000), with more emphasis on the methodological backgrounds. While the datasets are inspired by real cases, we also cut on their description and the motivations for their analysis. The current format thus serves as a unique textbook for a service course for scientists aimed at analyzing data the Bayesian way or as an introductory course on Bayesian statistics.

Note that we have not included any BUGS-oriented hierarchical analysis in this edition. This choice is deliberate: We have instead focussed on the Bayesian processing of mostly standard statistical models, notably in terms of prior specification and of the stochastic algorithms that are required to handle Bayesian estimation and model choice questions. We plainly expect that the readers of our book will have no difficulty in assimilating the BUGS philosophy, relying, for instance, on the highly relevant books by Lunn et al. (2012) and Gelman et al. (2013).

A course corresponding to the book has now been taught by both of us for several years in a second year master's program for students aiming at a professional degree in data processing and statistics (at Université Paris Dauphine, France) as well as in several US and Canadian universities. In Paris Dauphine the first half of the book was used in a 6-week (intensive) program, and students were tested on both the exercises (meaning all exercises) and their (practical) mastery of the datasets, the stated expectation being that they should go beyond a mere reproduction of the R outputs presented in the book. While the students found that the amount of work required by this course was rather beyond their usual standards (!), we observed that their understanding and mastery of Bayesian techniques were much deeper and more ingrained than in the more formal courses their counterparts had in the years before. In short, they started to think about the purpose of a Bayesian statistical analysis rather than on the contents of the final test and they ended up building a true intuition about what the results should look like, intuition

that, for instance, helped them to detect modeling and programming errors! In most subjects, working on Bayesian statistics from this perspective created a genuine interest in the approach and several students continued to use this approach in later courses or, even better, on the job.

Exercises are now focussed on solving problems rather than addressing finer theoretical points. Solutions to about half of the exercises are freely available on our webpages. We insist upon the point that the developments contained in those exercises are often relevant for fully understanding in the chapter.

Thanks

We are immensely grateful to colleagues and friends for their help with this book and its previous version, *Bayesian Core*, in particular, to the following people: François Perron somehow started thinking about this book and did a thorough editing of it during a second visit to Dauphine, helping us to adapt it more closely to North American audiences. He also adopted *Bayesian Core* as a textbook in Montréal as soon as it appeared. George Casella made helpful suggestions on the format of the book. Jérôme Dupuis provided capture–recapture slides that have been recycled in Chap. 5. Arnaud Doucet taught from the book at the University of British Columbia, Vancouver. Jean-Dominique Lebreton provided the European dipper dataset of Chap. 5. Gaelle Lefol pointed out the Eurostox series as a versatile dataset for Chap. 7. Kerrie Mengersen collaborated with both of us on a review paper about mixtures that is related to Chap. 6, Jim Kay introduced us to the Lake of Menteith dataset. Mike Titterington is thanked for collaborative friendship over the years and for a detailed set of comments on the book (quite in tune with his dedicated editorship of *Biometrika*). Jean-Louis Foulley provided us with some dataset and with extensive comments on their Bayesian processing. Even though we did not use those examples in the end, in connection with the strategy not to include BUGS-oriented materials, we are indebted to Jean-Louis for this help. Gilles Celeux carefully read the manuscript of the first edition and made numerous suggestions on both content and style. Darren Wraith, Julyan Arbel, Marco Banterle, Robin Ryder, and Sophie Donnet all reviewed some chapters or some R code and provided highly relevant comments, which clearly contributed to the final output. The picture of the caterpillar nest at the beginning of Chapter 3 was taken by Brigitte Plessis, Christian P. Robert’s spouse, near his great-grand-mother’s house in Brittany.

We are also grateful to the numerous readers who sent us queries about potential typos, as there were indeed many typos and if not unclear statements. Thanks in particular to Jarrett Barber, Hossein Gholami, we thus encourage all new readers of *Bayesian Essentials* to do the same!

The second edition of *Bayesian Core* was started, thanks to the support of the *Centre International de Rencontres Mathématiques* (CIRM), sponsored

by both the Centre National de la Recherche Scientifique (CNRS) and the Société Mathématique de France (SMF), located on the Luminy campus near Marseille. Being able to work “in pair” in the center for 2 weeks was an invaluable opportunity, boosted by the lovely surroundings of the Calanques, where mountain and sea meet! The help provided by the CIRM staff during the stay is also most gratefully acknowledged.

Montpellier, France
Paris, France
September 19, 2013

Jean-Michel Marin
Christian P. Robert

Contents

1	User's Manual	1
1.1	Expectations	2
1.2	Prerequisites and Further Reading	3
1.3	Styles and Fonts	4
1.4	An Introduction to R	5
1.4.1	Getting Started	6
1.4.2	R Objects	8
1.4.3	Probability Distributions in R	15
1.4.4	Graphical Facilities	16
1.4.5	Writing New R Functions	19
1.4.6	Input and Output in R	21
1.4.7	Administration of R Objects	21
1.5	The bayess Package	22
2	Normal Models	25
2.1	Normal Modeling	26
2.2	The Bayesian Toolkit	28
2.2.1	Posterior Distribution	28
2.2.2	Bayesian Estimates	33
2.2.3	Conjugate Prior Distributions	34
2.2.4	Noninformative Priors	35
2.2.5	Bayesian Credible Intervals	37
2.3	Bayesian Model Choice	38
2.3.1	The Model Index as a Parameter	39
2.3.2	The Bayes Factor	41
2.3.3	The Ban on Improper Priors	43
2.4	Monte Carlo Methods	46
2.4.1	An Approximation Based on Simulations	47
2.4.2	Importance Sampling	49
2.4.3	Approximation of Bayes Factors	52

2.5	Outlier Detection	58
2.6	Exercises	61
3	Regression and Variable Selection	65
3.1	Linear Models	66
3.2	Classical Least Squares Estimator	69
3.3	The Jeffreys Prior Analysis	73
3.4	Zellner's G -Prior Analysis	74
3.4.1	A Semi-noninformative Solution	75
3.4.2	The BayesReg R Function	80
3.4.3	Bayes Factors and Model Comparison	81
3.4.4	Prediction	84
3.5	Markov Chain Monte Carlo Methods	85
3.5.1	Conditionals	86
3.5.2	Two-Stage Gibbs Sampler	87
3.5.3	The General Gibbs Sampler	90
3.6	Variable Selection	91
3.6.1	Deciding on Explanatory Variables	91
3.6.2	G -Prior Distributions for Model Choice	93
3.6.3	A Stochastic Search for the Most Likely Model	96
3.7	Exercises	98
4	Generalized Linear Models	103
4.1	A Generalization of the Linear Model	104
4.1.1	Motivation	104
4.1.2	Link Functions	106
4.2	Metropolis–Hastings Algorithms	108
4.2.1	Definition	109
4.2.2	The Independence Sampler	110
4.2.3	The Random Walk Sampler	111
4.2.4	Output Analysis and Proposal Design	111
4.3	The Probit Model	115
4.3.1	Flat Prior	115
4.3.2	Noninformative G -Priors	117
4.3.3	About Informative Prior Analyses	122
4.4	The Logit Model	124
4.5	Log-Linear Models	127
4.5.1	Contingency Tables	127
4.5.2	Inference Under a Flat Prior	131
4.5.3	Model Choice and Significance of the Parameters	133
4.6	Exercises	137

5	Capture–Recapture Experiments	139
5.1	Inference in a Finite Population	140
5.2	Sampling Models	142
5.2.1	The Binomial Capture Model	142
5.2.2	The Two-Stage Capture–Recapture Model	143
5.2.3	The T -Stage Capture–Recapture Model	148
5.3	Open Populations	152
5.4	Accept–Reject Algorithms	156
5.5	The Arnason–Schwarz Capture–Recapture Model	160
5.5.1	Modeling	161
5.5.2	Gibbs Sampler	165
5.6	Exercises	168
6	Mixture Models	173
6.1	Missing Variable Models	174
6.2	Finite Mixture Models	176
6.3	Mixture Likelihoods and Posteriors	177
6.4	MCMC Solutions	182
6.5	Label Switching Difficulty	192
6.6	Prior Selection	198
6.7	Tempering	199
6.8	Mixtures with an Unknown Number of Components	201
6.9	Exercises	206
7	Time Series	209
7.1	Time-Indexed Data	210
7.1.1	Setting	210
7.1.2	Stability of Time Series	212
7.2	Autoregressive (AR) Models	214
7.2.1	The Models	215
7.2.2	Exploring the Parameter Space by MCMC Algorithms	219
7.3	Moving Average (MA) Models	226
7.4	ARMA Models and Other Extensions	232
7.5	Hidden Markov Models	236
7.5.1	Basics	237
7.5.2	Forward–Backward Representation	241
7.6	Exercises	248
8	Image Analysis	251
8.1	Image Analysis as a Statistical Problem	252
8.2	Spatial Dependence	252
8.2.1	Grids and Lattices	252
8.2.2	Markov Random Fields	254
8.2.3	The Ising Model	256
8.2.4	The Potts Model	260

8.3	Handling the Normalizing Constant	262
8.3.1	Path Sampling	264
8.3.2	The ABC Method	267
8.3.3	Inference on Potts Models	270
8.4	Image Segmentation	273
8.5	Exercises	281
About the Authors		285
References		287
Index		291