Chapman & Hall/CRC
Handbooks of Modern Statistical Methods

Handbook of Markov Chain Monte Carlo

Edited by

Steve Brooks
Andrew Gelman
Galin L. Jones
Xiao-Li Meng

CRC Press
Taylor & Francis Group
A CHAPMAN & HALL BOOK

Chapman & Hall/CRC
Handbooks of Modern
Statistical Methods

Handbook of Markov Chain Monte Carlo

Edited by

Steve Brooks
Andrew Gelman
Galin L. Jones
Xiao-Li Meng

CRC Press
Taylor & Francis Group
Boca Raton London New York

CRC Press is an imprint of the
Taylor & Francis Group, an **informa** business
A CHAPMAN & HALL BOOK

Chapman & Hall/CRC
Taylor & Francis Group
6000 Broken Sound Parkway NW, Suite 300
Boca Raton, FL 33487-2742

Printed in the United States of America on acid-free paper
10 9 8 7 6 5 4 3 2

International Standard Book Number: 978-1-4200-7941-8 (Hardback)

Visit the Taylor & Francis Web site at
http://www.taylorandfrancis.com

and the CRC Press Web site at
http://www.crcpress.com

Preface

Over the past 20 years or so, Markov Chain Monte Carlo (MCMC) methods have revolutionized statistical computing. They have impacted the practice of Bayesian statistics profoundly by allowing intricate models to be posited and *used* in an astonishing array of disciplines as diverse as fisheries science and economics. Of course, Bayesians are not the only ones to benefit from using MCMC, and there continues to be increasing use of MCMC in other statistical settings. The practical importance of MCMC has also sparked expansive and deep investigation into fundamental Markov chain theory. As the use of MCMC methods mature, we see deeper theoretical questions addressed, more complex applications undertaken and their use spreading to new fields of study. It seemed to us that it was a good time to try to collect an overview of MCMC research and its applications.

This book is intended to be a reference (not a text) for a broad audience and to be of use both to developers and users of MCMC methodology. There is enough introductory material in the book to help graduate students as well as researchers new to MCMC who wish to become acquainted with the basic theory, algorithms and applications. The book should also be of particular interest to those involved in the development or application of new and advanced MCMC methods. Given the diversity of disciplines that use MCMC, it seemed prudent to have many of the chapters devoted to detailed examples and case studies of realistic scientific problems. Those wanting to see current practice in MCMC will find a wealth of material to choose from here.

Roughly speaking, we can divide the book into two parts. The first part encompasses 12 chapters concerning MCMC foundations, methodology and algorithms. The second part consists of 12 chapters which consider the use of MCMC in practical applications. Within the first part, the authors take such a wide variety of approaches that it seems pointless to try to classify the chapters into subgroups. For example, some chapters attempt to appeal to a broad audience by taking a tutorial approach while other chapters, even if introductory, are either more specialized or present more advanced material. Yet others present original research. In the second part, the focus shifts to applications. Here again, we see a variety of topics, but there are two basic approaches taken by the authors of these chapters. The first is to provide an overview of an application area with the goal of identifying best MCMC practice in the area through extended examples. The second approach is to provide detailed case studies of a given problem while clearly identifying the statistical and MCMC-related issues encountered in the application.

When we were planning this book, we quickly realized that no single source can give a truly comprehensive overview of cutting-edge MCMC research and applications—there is just too much of it and its development is moving too fast. Instead, the editorial goal was to obtain contributions of high quality that may stand the test of time. To this end, all of the contributions (including those written by members of the editorial panel) were submitted to a rigorous peer review process and many underwent several revisions. Some contributions, even after revisions, were deemed unacceptable for publication here, and we certainly welcome constructive feedback on the chapters that did survive our editorial process. We thank all the authors for their efforts and patience in this process, and we ask for understanding from those whose contributions are not included in this book. We believe the breadth and depth of the contributions to this book, including some diverse opinions expressed, imply a continuously bright and dynamic future for MCMC research. We hope this book inspires further work—theoretical, methodological, and applied—in this exciting and rich area.

Finally, no project of this magnitude could be completed with satisfactory outcome without many individuals' help. We especially want to thank Robert Calver of Chapman & Hall/CRC for his encouragements, guidelines, and particularly his patience during the entire process of editing this book. We also offer our heartfelt thanks to the numerous referees for their insightful and rigorous review, often multiple times. Of course, the ultimate appreciation for all individuals involved in this project comes from your satisfaction with the book or at least a part of it. So we thank you for reading it.

Steve Brooks
Andrew Gelman
Galin L. Jones
Xiao-Li Meng

Editors

Steve Brooks is company director of ATASS, a statistical consultancy business based in the United Kingdom. He was formerly professor of Statistics at Cambridge University and received the Royal Statistical Society Guy medal in Bronze in 2005 and the Philip Leverhulme prize in 2004. Like his co-editors, he has served on numerous professional committees both in the United Kingdom and elsewhere, as well as sitting on numerous editorial boards. He is co-author of *Bayesian Analysis for Population Ecology* (Chapman & Hall/CRC, 2009) and co-founder of the National Centre for Statistical Ecology. His research interests include the development and application of computational statistical methodology across a broad range of application areas.

Andrew Gelman is a professor of statistics and political science and director of the Applied Statistics Center at Columbia University. He has received the Outstanding Statistical Application award from the American Statistical Association, the award for best article published in the *American Political Science Review*, and the Committee of Presidents of Statistical Societies award for outstanding contributions by a person under the age of 40. His books include *Bayesian Data Analysis* (with John Carlin, Hal Stern, and Don Rubin), *Teaching Statistics: A Bag of Tricks* (with Deb Nolan), *Data Analysis Using Regression and Multilevel/Hierarchical Models* (with Jennifer Hill), and, most recently, *Red State, Blue State, Rich State, Poor State: Why Americans Vote the Way They Do* (with David Park, Boris Shor, Joe Bafumi, and Jeronimo Cortina).

Andrew has done research on a wide range of topics, including: why it is rational to vote; why campaign polls are so variable when elections are so predictable; why redistricting is good for democracy; reversals of death sentences; police stops in New York City; the statistical challenges of estimating small effects; the probability that your vote will be decisive; seats and votes in Congress; social network structure; arsenic in Bangladesh; radon in your basement; toxicology; medical imaging; and methods in surveys, experimental design, statistical inference, computation, and graphics.

Galin L. Jones is an associate professor in the School of Statistics at the University of Minnesota. He has served on many professional committees and is currently serving on the editorial board for the *Journal of Computational and Graphical Statistics*. His research interests include Markov chain Monte Carlo, Markov chains in decision theory, and applications of statistical methodology in agricultural, biological, and environmental settings.

Xiao-Li Meng is the Whipple V. N. Jones professor of statistics and chair of the Department of Statistics at Harvard University; previously he taught at the University of Chicago (1991–2001). He was the recipient of the 1997–1998 University of Chicago Faculty Award for Excellence in Graduate Teaching, the 2001 Committee of Presidents of Statistical Societies Award, the 2003 Distinguished Achievement Award and the 2008 Distinguished Service Award from the International Chinese Statistical Association, and the 2010 Medallion Lecturer from the Institute of Mathematical Statistics (IMS). He has served on numerous professional committees, including chairing the 2004 Joint Statistical Meetings and the Committee on Meetings of American Statistical Association (ASA) from 2004 until 2010. He is an elected fellow of the ASA and the IMS. He has also served on editorial boards for *The Annals of Statistics*, *Bayesian Analysis*, *Bernoulli*, *Biometrika*, *Journal of the American Statistical*

Association, as well as the coeditor of *Statistica Sinica*. Currently, he is the statistics editor for the IMS Monograph and Textbook Series. He is also a coeditor of *Applied Bayesian Modeling and Causal Inference from Incomplete-Data Perspectives* (Gelman and Meng, 2004, Wiley) and *Strength in Numbers: The Rising of Academic Statistics Departments in the U.S.* (Agresti and Meng, 2012, Springer). His research interests include inference foundations and philosophies, models of all flavors, deterministic and stochastic algorithms, signal extraction in physical, social and medical sciences, and occasionally elegant mathematical statistics.

Contributors

Susan Spear Bassett
Department of Psychiatry and Behavioral Sciences
Johns Hopkins Hospital
Baltimore, Maryland

John T. Behrens
Cisco Systems, Inc.
Mishawaka, Indiana

DuBois Bowman
Center for Biomedical Imaging Statistics (CBIS)
Emory University
Atlanta, Georgia

Steve Brooks
ATASS Ltd
Exeter, United Kingdom

Brian Caffo
Department of Biostatistics
Johns Hopkins University
Baltimore, Maryland

George Casella
Department of Statistics
University of Florida
Gainesville, Florida

Radu V. Craiu
Department of Statistics
University of Toronto
Toronto, Ontario, Canada

Francesca Dominici
Department of Biostatistics
Johns Hopkins University
Baltimore, Maryland

Lynn Eberly
Division of Biostatistics
University of Minnesota
Minneapolis, Minnesota

Yanan Fan
School of Mathematics and Statistics
University of New South Wales
Sydney, Australia

Paul Fearnhead
Department of Mathematics and Statistics
Lancaster University
Lancaster, United Kingdom

James M. Flegal
Department of Statistics
University of California
Riverside, California

Colin Fox
Department of Physics
University of Otago
Dunedin, New Zealand

Filiz Garip
Department of Sociology
Harvard University
Cambridge, Massachusetts

Andrew Gelman
Department of Statistics

and

Department of Political Science
Columbia University
New York, New York

Charles J. Geyer
School of Statistics
University of Minnesota
Minneapolis, Minnesota

Murali Haran
Center for Ecology and Environmental Statistics
Pennsylvania State University
University Park, Pennsylvania

David Higdon
Los Alamos National Laboratory
Los Alamos, New Mexico

James P. Hobert
Department of Statistics
University of Florida
Gainsville, Florida

Mark Huber
Department of Mathematical Sciences
Claremont McKenna College
Claremont, California

Galin L. Jones
School of Statistics
University of Minnesota
Minneapolis, Minnesota

Ruth King
School of Mathematics and Statistics
University of St. Andrews
St. Andrews, United Kingdom

Roy Levy
School of Social and Family Dynamics
Arizona State University
Tempe, Arizona

Thomas A. Louis
Department of Biostatistics
Johns Hopkins University
Baltimore, Maryland

Xiao-Li Meng
Department of Statistics
Harvard University
Cambridge, Massachusetts

Russell B. Millar
Department of Statistics
University of Auckland
Auckland, New Zealand

Robert J. Mislevy
Department of Measurement, Statistics and Evaluation
University of Maryland
Severna Park, Maryland

J. David Moulton
Los Alamos National Laboratory
Los Alamos, New Mexico

Radford M. Neal
Department of Statistics
University of Toronto
Toronto, Ontario, Canada

Jong Hee Park
Department of Political Science
University of Chicago
Chicago, Illinois

Taeyoung Park
Department of Applied Statistic
Yonsei University
Seoul, South Korea

Roger Peng
Department of Biostatistics
Johns Hopkins University
Baltimore, Maryland

C. Shane Reese
Department of Statistics
Brigham Young University
Provo, Utah

Christian Robert
CEREMADE—Université Paris-Dauphine
Paris, France

Jeffrey S. Rosenthal
Department of Statistics
University of Toronto
Toronto, Ontario, Canada

Kenneth Shirley
The Earth Institute
Columbia University
New York, New York

Scott A. Sisson
School of Mathematics and Statistics
University of New South Wales
Sydney, Australia

Elizabeth Thompson
Department of Statistics
University of Washington
Seattle, Washington

David A. van Dyk
Department of Statistics
University of California
Irvine, California

Jasper A. Vrugt
Center for Non-Linear Studies
Irvine, California

Bruce Western
Department of Sociology
Harvard University
Cambridge, Massachusetts

Scott Zeger
Department of Biostatistics
Johns Hopkins University
Baltimore, Maryland

Contents