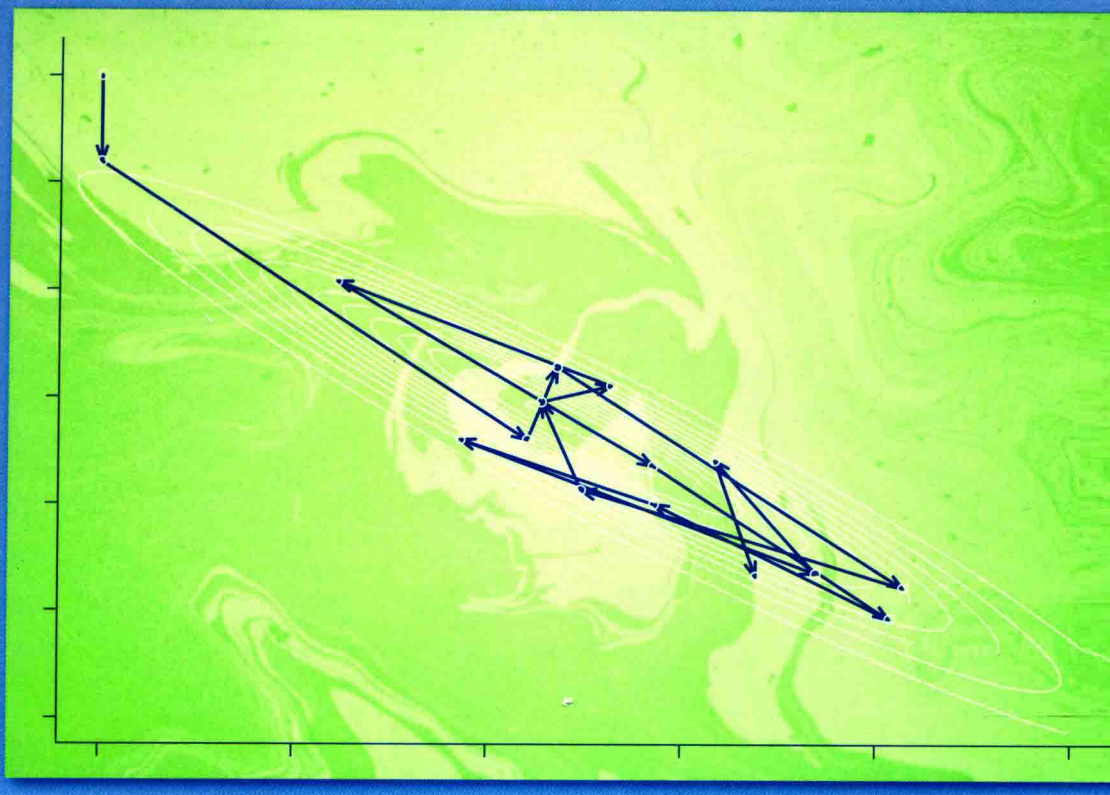
WILEY SERIES IN COMPUTATIONAL STATISTICS



Bayesian Modeling Using WinBUGS

Ioannis Ntzoufras

WILEY

Bayesian Modeling Using WinBUGS

Ioannis Ntzoufras
Department of Statistics
Athens University of Economics and Business
Athens, Greece

WILEY

A JOHN WILEY & SONS, INC., PUBLICATION

Published by John Wiley & Sons, Inc., Hoboken, New Jersey.
Published simultaneously in Canada.

Library of Congress Cataloging-in-Publication Data is available.

Ntzoufras, Ioannis, 1973-
Bayesian modeling using WinBUGS / Ioannis Ntzoufras.
p. cm.
Includes bibliographical references and index.
ISBN 978-0-470-14114-4 (pbk.)
1. Bayesian statistical decision theory. 2. WinBUGS. I. Title.
QA279.5.N89 2009
519.5'42—dc22 2008033316

Printed in the United States of America.

10 9 8 7 6 5 4 3 2 1

Bayesian Modeling Using WinBUGS

WILEY SERIES IN COMPUTATIONAL STATISTICS

Wiley Series in Computational Statistics is comprised of practical guides and cutting edge research books on new developments in computational statistics. It features quality authors with a strong applications focus. The texts in the series provide detailed coverage of statistical concepts, methods and case studies in areas at the interface of statistics, computing, and numerics.

With sound motivation and a wealth of practical examples, the books show in concrete terms how to select and to use appropriate ranges of statistical computing techniques in particular fields of study. Readers are assumed to have a basic understanding of introductory terminology.

The series concentrates on applications of computational methods in statistics to fields of bioinformatics, genomics, epidemiology, business, engineering, finance and applied statistics.

A complete list of titles in this series appears at the end of the volume.

To Ioanna and our baby daughter

PREFACE

Since the mid-1980s, the development of widely accessible powerful computers and the implementation of Markov chain Monte Carlo (MCMC) methods have led to an explosion of interest in Bayesian statistics and modeling. This was followed by an extensive research for new Bayesian methodologies generating the practical application of complicated models used over a wide range of sciences. During the late 1990s, BUGS emerged in the foreground. BUGS was a free software that could fit complicated models in a relatively easy manner, using standard MCMC methods. Since 1998 or so, WinBUGS , the Windows version of BUGS, has earned great popularity among researchers of diverse scientific fields. Therefore, an increased need for an introductory book related to Bayesian models and their implementation via WinBUGS has been realized.

The objective of the present book is to offer an introduction to the principles of Bayesian modeling, with emphasis on model building and model implementation using WinBUGS . Detailed examples are provided, ranging from very simple to more advanced and realistic ones. Generalized linear models (GLMs), which are familiar to most students and researchers, are discussed. Details concerning model building, prior specification, writing the WinBUGS code and the analysis and interpretation of the WinBUGS output are also provided. Because of the introductory character of the book, I focused on elementary models, starting from the normal regression models and moving to generalized linear models. Even more advanced readers, familiar with such models, may benefit from the Bayesian implementation using WinBUGS .

Basic knowledge of probability theory and statistics is assumed. Computations that could not be performed in WinBUGS are illustrated using R. Therefore, a minimum knowledge of R is also required.

This manuscript can be used as the main textbook in a second-level course of Bayesian statistics focusing on modeling and/or computation. Alternatively, it can serve as a companion (to a main textbook) in an introductory course of a Bayesian statistics. Finally, because of its structure, postgraduate students and other researchers can complete a self-taught tutorial course on Bayesian modeling by following the material of this book.

All datasets and code used in the book are available in the book's Webpage: `www.stat-athens.aueb.gr/~jbn/winbugs_book`.

IOANNIS NTZOUFRAS

Athens, Greece
June 29, 2008

ACKNOWLEDGMENTS

I am indebted to the people at Wiley publications for their understanding and assistance during the preparation of the manuscript. Acknowledgments are due to the anonymous referees. Their suggestions and comments led to a substantial improvement of the present book. I would particularly like to thank Dimitris Fouskakis, colleague and good friend, for his valuable comments on an early version of chapters 1–6 and 10–11. I am also grateful to Professor Brani Vidakovic for proposing and motivating this book. Last but not least, I wish to thank my wife Ioanna for her love, support, and patience during the writing of this book as well as for her suggestions on the manuscript.

I. N.

ACRONYMS

ACF	Autocorrelation
AIC	Akaike information criterion
ANOVA	Analysis of variance
ANCOVA	Analysis of covariance
AR	Attributable risk
BF	Bayes factor
BIC	Bayes information criterion
BOA	Bayesian output analysis (R package)
BP	Bivariate Poisson
BOD	Biological oxygen demand (data variable in example 6.3)
BUGS	Bayesian inference using Gibbs (software)
CDF	Cumulative distribution function
COD	Chemical oxygen demand (data variable in example 6.3)
CODA	Convergence diagnostics and output analysis software for Gibbs sampling analysis (R package)
CPO	Conditional Predictive Ordinate
CR	corner (constraint)
CV	Cross-validation

CV-1	Leave-one-out cross-validation
DAG	Directed acyclic graph
DI	Dispersion index
DIBP	Diagonal inflated bivariate Poisson distribution
DIC	Deviance information criterion
GLM	Generalized linear model
GP	Generalized Poisson
GVS	Gibbs variable selection
ICPO	Inverse conditional predictive ordinate
i.i.d.	Independent identically distributed
LS	Logarithmic score
MAP	Maximum a posteriori
MP model	Median probability
MCMC	Markov chain Monte Carlo
MCE	Monte Carlo error
ML	Maximum likelihood
MLE	Maximum-likelihood estimate/estimator
NB	Negative binomial
OR	Odds ratio
PBF	Posterior Bayes factor
PD	Poisson difference
p.d.f.	Probability density function
PO	Posterior model odds
PPO	Posterior predictive ordinate
RJMCMC	Reversible jump Markov chain Monte Carlo
RR	Relative risk
SD	Standard deviation
SE	Standard error
SSVS	Stochastic search variable selection
STZ	sum-to-zero (constraint)
TS	Total solids(data variable in example 6.3)
TVS	Total volatile solids (data variable in example 6.3)
WinBUGS	Windows version of BUGS (software)
ZI	Zero inflated
ZID	Zero inflated distribution
ZIP	Zero inflated Poisson distribution
ZINB	Zero inflated negative binomial distribution

ZIGP	Zero inflated generalized Poisson distribution
ZIBP	Zero inflated bivariate Poisson distribution

CONTENTS