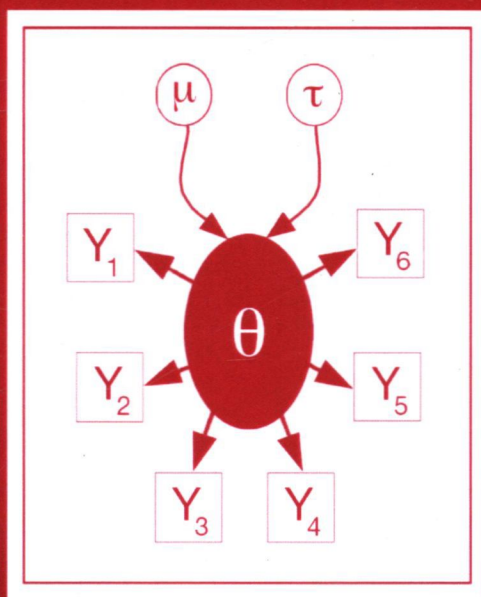


Texts in Statistical Science

The BUGS Book

A Practical Introduction to Bayesian Analysis



David Lunn
Christopher Jackson
Nicky Best
Andrew Thomas
David Spiegelhalter



CRC Press
Taylor & Francis Group

A CHAPMAN & HALL BOOK

Bayesian statistical methods have become widely used for data analysis and modelling in recent years, and the BUGS software has become the most popular software for Bayesian analysis worldwide. Authored by the team that originally developed this software, **The BUGS Book** provides a practical introduction to this program and its use. The text presents complete coverage of all the functionalities of BUGS, including prediction, missing data, model criticism, and prior sensitivity. It also features a large number of worked examples and a wide range of applications from various disciplines.

The book introduces regression models, techniques for criticism and comparison, and a wide range of modelling issues before going into the vital area of hierarchical models, one of the most common applications of Bayesian methods. It deals with essentials of modelling without getting bogged down in complexity. The book emphasises model criticism, model comparison, sensitivity analysis to alternative priors, and thoughtful choice of prior distributions—all those aspects of the “art” of modelling that are easily overlooked in more theoretical expositions.

More pragmatic than ideological, the authors systematically work through the large range of “tricks” that reveal the real power of the BUGS software, for example, dealing with missing data, censoring, grouped data, prediction, ranking, parameter constraints, and so on. Many of the examples are biostatistical, but they do not require domain knowledge and are generalisable to a wide range of other application areas.

Full code and data for examples, exercises, and some solutions can be found on the book's website.



CRC Press

Taylor & Francis Group
an informa business

www.crcpress.com

ISBN 978-1-138-46948-8



9 781138 469488

The BUGS Book

Lunn, Jackson, Best,
Thomas, and Spiegelhalter



Texts in Statistical Science

The BUGS Book

A Practical Introduction to Bayesian Analysis

David Lunn

Christopher Jackson

Nicky Best

Andrew Thomas

David Spiegelhalter



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

MATLAB® is a trademark of The MathWorks, Inc. and is used with permission. The MathWorks does not warrant the accuracy of the text or exercises in this book. This book's use or discussion of MATLAB® software or related products does not constitute endorsement or sponsorship by The MathWorks of a particular pedagogical approach or particular use of the MATLAB® software.

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

First issued in hardback 2017

© 2013 by Taylor & Francis Group, LLC
CRC Press is an imprint of Taylor & Francis Group, an Informa business

No claim to original U.S. Government works

ISBN-13: 978-1-58488-849-9 (pbk)
ISBN-13: 978-1-13846-948-8 (hbk)

This book contains information obtained from authentic and highly regarded sources. While all reasonable efforts have been made to publish reliable data and information, neither the author[s] nor the publisher can accept any legal responsibility or liability for any errors or omissions that may be made. The publishers wish to make clear that any views or opinions expressed in this book by individual editors, authors or contributors are personal to them and do not necessarily reflect the views/opinions of the publishers. The information or guidance contained in this book is intended for use by medical, scientific or health-care professionals and is provided strictly as a supplement to the medical or other professional's own judgement, their knowledge of the patient's medical history, relevant manufacturer's instructions and the appropriate best practice guidelines. Because of the rapid advances in medical science, any information or advice on dosages, procedures or diagnoses should be independently verified. The reader is strongly urged to consult the relevant national drug formulary and the drug companies' and device or material manufacturers' printed instructions, and their websites, before administering or utilizing any of the drugs, devices or materials mentioned in this book. This book does not indicate whether a particular treatment is appropriate or suitable for a particular individual. Ultimately it is the sole responsibility of the medical professional to make his or her own professional judgements, so as to advise and treat patients appropriately. The authors and publishers have also attempted to trace the copyright holders of all material reproduced in this publication and apologize to copyright holders if permission to publish in this form has not been obtained. If any copyright material has not been acknowledged please write and let us know so we may rectify in any future reprint.

Except as permitted under U.S. Copyright Law, no part of this book may be reprinted, reproduced, transmitted, or utilized in any form by any electronic, mechanical, or other means, now known or hereafter invented, including photocopying, microfilming, and recording, or in any information storage or retrieval system, without written permission from the publishers.

For permission to photocopy or use material electronically from this work, please access www.copyright.com (<http://www.copyright.com/>) or contact the Copyright Clearance Center, Inc. (CCC), 222 Rosewood Drive, Danvers, MA 01923, 978-750-8400. CCC is a not-for-profit organization that provides licenses and registration for a variety of users. For organizations that have been granted a photocopy license by the CCC, a separate system of payment has been arranged.

Trademark Notice: Product or corporate names may be trademarks or registered trademarks, and are used only for identification and explanation without intent to infringe.

Visit the Taylor & Francis Web site at
<http://www.taylorandfrancis.com>
and the CRC Press Web site at
<http://www.crcpress.com>

CHAPMAN & HALL/CRC

Texts in Statistical Science Series

Series Editors

Francesca Dominici, *Harvard School of Public Health, USA*

Julian J. Faraway, *University of Bath, UK*

Martin Tanner, *Northwestern University, USA*

Jim Zidek, *University of British Columbia, Canada*

Analysis of Failure and Survival Data

P. J. Smith

**The Analysis of Time Series —
An Introduction, Sixth Edition**

C. Chatfield

**Applied Bayesian Forecasting and Time Series
Analysis**

A. Pole, M. West, and J. Harrison

Applied Categorical and Count Data Analysis

W. Tang, H. He, and X.M. Tu

**Applied Nonparametric Statistical Methods,
Fourth Edition**

P. Sprent and N.C. Smeeton

**Applied Statistics — Handbook of GENSTAT
Analysis**

E.J. Snell and H. Simpson

Applied Statistics — Principles and Examples

D.R. Cox and E.J. Snell

Applied Stochastic Modelling, Second Edition

B.J.T. Morgan

Bayesian Data Analysis, Second Edition

A. Gelman, J.B. Carlin, H.S. Stern,
and D.B. Rubin

**Bayesian Ideas and Data Analysis: An Introduction
for Scientists and Statisticians**

R. Christensen, W. Johnson, A. Branscum,
and T.E. Hanson

**Bayesian Methods for Data Analysis,
Third Edition**

B.P. Carlin and T.A. Louis

Beyond ANOVA — Basics of Applied Statistics

R.G. Miller, Jr.

**The BUGS Book: A Practical Introduction to
Bayesian Analysis**

D. Lunn, C. Jackson, N. Best, A. Thomas, and
D. Spiegelhalter

A Course in Categorical Data Analysis

T. Leonard

A Course in Large Sample Theory

T.S. Ferguson

Data Driven Statistical Methods

P. Sprent

Decision Analysis — A Bayesian Approach

J.Q. Smith

Design and Analysis of Experiments with SAS

J. Lawson

**Elementary Applications of Probability Theory,
Second Edition**

H.C. Tuckwell

Elements of Simulation

B.J.T. Morgan

**Epidemiology — Study Design and
Data Analysis, Second Edition**

M. Woodward

Essential Statistics, Fourth Edition

D.A.G. Rees

Exercises and Solutions in Biostatistical Theory

L.L. Kupper, B.H. Neelon, and S.M. O'Brien

**Extending the Linear Model with R — Generalized
Linear, Mixed Effects and Nonparametric Regression
Models**

J.J. Faraway

A First Course in Linear Model Theory

N. Ravishanker and D.K. Dey

Generalized Additive Models:

An Introduction with R

S. Wood

Generalized Linear Mixed Models:

Modern Concepts, Methods and Applications

W. W. Stroup

Graphics for Statistics and Data Analysis with R

K.J. Keen

**Interpreting Data — A First Course
in Statistics**

A.J.B. Anderson

**Introduction to General and Generalized
Linear Models**

H. Madsen and P. Thyregod

**An Introduction to Generalized
Linear Models, Third Edition**

A.J. Dobson and A.G. Barnett

Introduction to Multivariate Analysis

C. Chatfield and A.J. Collins

**Introduction to Optimization Methods and Their
Applications in Statistics**

B.S. Everitt

Introduction to Probability with R

K. Baclawski

- Introduction to Randomized Controlled Clinical Trials, Second Edition**
J.N.S. Matthews
- Introduction to Statistical Inference and Its Applications with R**
M.W. Trosset
- Introduction to Statistical Limit Theory**
A.M. Polansky
- Introduction to Statistical Methods for Clinical Trials**
T.D. Cook and D.L. DeMets
- Introduction to the Theory of Statistical Inference**
H. Liero and S. Zwanzig
- Large Sample Methods in Statistics**
P.K. Sen and J. da Motta Singer
- Linear Models with R**
J.J. Faraway
- Logistic Regression Models**
J.M. Hilbe
- Markov Chain Monte Carlo — Stochastic Simulation for Bayesian Inference, Second Edition**
D. Gamerman and H.F. Lopes
- Mathematical Statistics**
K. Knight
- Modeling and Analysis of Stochastic Systems, Second Edition**
V.G. Kulkarni
- Modelling Binary Data, Second Edition**
D. Collett
- Modelling Survival Data in Medical Research, Second Edition**
D. Collett
- Multivariate Analysis of Variance and Repeated Measures — A Practical Approach for Behavioural Scientists**
D.J. Hand and C.C. Taylor
- Multivariate Statistics — A Practical Approach**
B. Flury and H. Riedwyl
- Multivariate Survival Analysis and Competing Risks**
M. Crowder
- Pólya Urn Models**
H. Mahmoud
- Practical Data Analysis for Designed Experiments**
B.S. Yandell
- Practical Longitudinal Data Analysis**
D.J. Hand and M. Crowder
- Practical Multivariate Analysis, Fifth Edition**
A. Afifi, S. May, and V.A. Clark
- Practical Statistics for Medical Research**
D.G. Altman
- A Primer on Linear Models**
J.F. Monahan
- Principles of Uncertainty**
J.B. Kadane
- Probability — Methods and Measurement**
A. O'Hagan
- Problem Solving — A Statistician's Guide, Second Edition**
C. Chatfield
- Randomization, Bootstrap and Monte Carlo Methods in Biology, Third Edition**
B.F.J. Manly
- Readings in Decision Analysis**
S. French
- Sampling Methodologies with Applications**
P.S.R.S. Rao
- Statistical Analysis of Reliability Data**
M.J. Crowder, A.C. Kimber, T.J. Sweeting, and R.L. Smith
- Statistical Methods for Spatial Data Analysis**
O. Schabenberger and C.A. Gotway
- Statistical Methods for SPC and TQM**
D. Bissell
- Statistical Methods in Agriculture and Experimental Biology, Second Edition**
R. Mead, R.N. Curnow, and A.M. Hasted
- Statistical Process Control — Theory and Practice, Third Edition**
G.B. Wetherill and D.W. Brown
- Statistical Theory, Fourth Edition**
B.W. Lindgren
- Statistics for Accountants**
S. Letchford
- Statistics for Epidemiology**
N.P. Jewell
- Statistics for Technology — A Course in Applied Statistics, Third Edition**
C. Chatfield
- Statistics in Engineering — A Practical Approach**
A.V. Metcalfe
- Statistics in Research and Development, Second Edition**
R. Caulcutt
- Stochastic Processes: An Introduction, Second Edition**
P.W. Jones and P. Smith
- Survival Analysis Using S — Analysis of Time-to-Event Data**
M. Tableman and J.S. Kim
- The Theory of Linear Models**
B. Jørgensen
- Time Series Analysis**
H. Madsen
- Time Series: Modeling, Computation, and Inference**
R. Prado and M. West

Preface

History Markov chain Monte Carlo (MCMC) methods, in which plausible values for unknown quantities are simulated from their appropriate probability distribution, have revolutionised the practice of statistics. For more than 20 years the BUGS project has been at the forefront of this movement. The BUGS project began in Cambridge, United Kingdom, in 1989, just as Alan Gelfand and Adrian Smith were working 80 miles away in Nottingham on their classic Gibbs sampler paper (Gelfand and Smith, 1990) that kicked off the revolution. But we never communicated (except through the intermediate node of David Clayton) and whereas the Gelfand–Smith approach used image processing as inspiration, the philosophy behind BUGS was rooted more in techniques for handling uncertainty in artificial intelligence using directed graphical models and what came to be called Bayesian networks (Pearl, 1988). Lunn et al. (2009b) lay out all this history in greater detail.

Some people have accused Markov chain Monte Carlo methods of being slow, but nothing could compare with the time it has taken for this book to be written! The first proposal dates from 1995, but things got in the way, as they do, and it needed a vigorous new generation of researchers to finally get it finished. It is slightly galling that much of the current book could have been written in the mid-1990s, since the basic ideas of the software, the language for model description, and indeed some of the examples are unchanged. Nevertheless there have been important developments in the extended gestational period of the book, for example, techniques for model criticism and comparison, implementation of differential equations and nonparametric techniques, and the ability to run BUGS code within a range of alternative programs.

The BUGS project is rooted in the idea of generic reusable components that can be put together as desired, like a child's construction set but not quite as colourful. In this book we typically tackle each of these components one by one using deliberately simplified examples, but hopefully it will be clear that they can be easily assembled into arbitrarily complex models. This flexibility has enabled BUGS to be applied in areas that we had never dreamed about, which is gratifying. But it is also important to note that in many situations BUGS may not be the most efficient method, and there are many things it cannot do. Yet...

What's in the book? Perhaps we should start by saying what is *not* in the book. First, there is minimal statistical theory, neither of statistical infer-

ence nor of Markov chain Monte Carlo methods (although a presumption of some familiarity with probability theory is made). This is partly to keep the book to a manageable length, but also because the very way in which BUGS works removes the need for much of the theory that is taught in standard Bayesian statistics courses. Second, we do not cover decision theory, as BUGS has been designed for handling Bayesian inferences expressed as an appropriate posterior distribution. Finally, we take it for granted that a Bayesian approach is desired, and so barely bother to lay out the reasons why this may be appropriate.

A glance at the chapter contents will reveal that we introduce regression models, techniques for criticism and comparison, and a wide range of modelling issues before going into the vital and traditional Bayesian area of hierarchical models. This decision came after considerable thought and experimentation, and was based on the wish to deal with the essentials of modelling without getting bogged down in complexity. Our aim is to bring to the forefront model criticism, model comparison, sensitivity analysis to alternative priors, and thoughtful choice of prior distributions — all those aspects of the “art” of modelling that are easily overlooked in more theoretical expositions. But we have also really enjoyed going systematically through the large range of “tricks” that reveal the real power of the BUGS software: for example, dealing with missing data, censoring, grouped data, prediction, ranking, parameter constraints, and so on.

Our professional background has meant that many of the examples are biostatistical, but they do not require domain knowledge and hopefully it will be clear that they are generalisable to a wide range of other application areas. Full code and data for the examples, exercises, and some solutions can all be found on the book website: www.mrc-bsu.cam.ac.uk/bugs/thebugsbook.

The BUGS approach clearly separates the model description from the “engine,” or algorithms and software, used to actually do the simulations. A brief introduction to WinBUGS is given in Chapter 2, but fully detailed instructions of how to run WinBUGS and similar software have been deferred to the final chapter, 12, and a reference guide to the modelling language is given in the appendices. Since BUGS now comes in a variety of flavours, we have tried to ensure that the book works for WinBUGS, OpenBUGS, and JAGS, and any differences have been highlighted. Nevertheless the software is constantly improving, and so in some areas the book is not completely prescriptive but tries to communicate possible developments.

Finally, we acknowledge there are many shades of Bayesianism: our own philosophy is more pragmatic than ideological and doubtless there will be some who will continue to spurn our rather informal attitude. An example of this informality is our use of the term ‘likelihood’, which is sometimes used when referring to a sampling distribution. We doubt this will lead to confusion.

How to use the book. Our intended audience comprises anyone who would like to apply Bayesian methods to real-world problems. These might be practising statisticians, or scientists with a good statistical background, say familiarity with classical statistics and some calculus-based probability and mathematical statistics. We do not assume familiarity with Bayesian methods or MCMC. The book could be used for self-learning, for short courses, and for longer courses, either by itself or in combination with a textbook such as Gelman et al. (2004) or Carlin and Louis (2008).

Chapters 1 to 6 provide a basic introduction up to regression modelling, which should be a review for those with some experience with Bayesian methods and BUGS. Beyond that there should be new material, even for experienced users. For a one-semester course we would recommend Chapters 1 to 6, most of Chapter 8 on model criticism and comparison, and Chapter 10 on hierarchical models. A longer course could select from the wide range of issues and models outlined in Chapters 7, 9 and 11, depending on what is most relevant for the audience.

Whether studying on your own or as part of a course, instructions for running the WinBUGS software are given briefly in Chapter 2 and fully in Chapter 12. A full explanation of BUGS model syntax and a list of functions and distributions are given in the appendices. Chapter 12 explains how OpenBUGS and JAGS differ from WinBUGS and gives examples of how all varieties of BUGS can be conveniently run from other software, in particular from R.

Other sources. If an accompanying text on the underlying theory of Bayesian inference is required, possibilities include Gelman et al. (2004), Carlin and Louis (2008) and Lee (2004), with Bernardo and Smith (1994) providing a deeper treatment. Other books focus explicitly on BUGS: Ntzoufras (2009) provides a detailed exposition of WinBUGS with accompanying theory, Gelman and Hill (2007) explore both standard and hierarchical regression models using both R and BUGS, while the texts by Congdon (2003, 2005, 2006, 2010) explore a staggering range of applications of BUGS that we could not hope to match. Jackman (2009) covers both theory and BUGS implementations within social science, ecology applications are covered by Kéry (2010) and Kéry and Schaub (2011), while Kruschke (2010) gives a tutorial in Bayesian analysis and BUGS with applications in psychology. Expositions on MCMC theory include Gamerman and Lopes (2006) and Brooks et al. (2011), while Gilks et al. (1996) is still relevant even after many years.

Finally, there are numerous websites that provide examples and teaching material, and when tackling a new problem we strongly recommend trying to find these using appropriate search terms and adapting someone else's code. We have always been impressed by the great generosity of BUGS users in sharing code and ideas, perhaps helped by the fact that the software has always been freely available.

A suggested strategy for inference and reporting. Rather than leaving it until later in the book, it seems appropriate to lay out at an early stage the approach to modelling and reporting that we have tried to exemplify. Bayesian analysis requires a specification of prior distributions and models for the sampling distribution for the data. For prior distributions, we emphasise that there is no such thing as the “correct” prior, and instead recommend exploring a range of plausible assumptions and conducting sensitivity analysis. Regarding assumptions for the sampling distribution, throughout this book we try to exemplify a reasonably consistent approach to modelling based on an iterative cycle of fitting and checking. We recommend starting with fairly simple assumptions, cross-checking with graphics and informal checks of model fit which can then suggest plausible elaborations. A final list of candidate models can then be compared using more formal methods.

There have been limited “guidelines” for reporting Bayesian analyses, e.g., Spiegelhalter et al. (2004), Sung et al. (2005), and Johnson (2011) in a medical context, and also BaSiS (2001). Naturally the data have to be summarised numerically and graphically. We need to acknowledge that Bayesian methods tend to be inherently more complex than classical analyses, and thus there is an additional need for clarity with the aim that the analysis could be replicated by another investigator who has access to the full data, with perhaps full details of computational methods and code given online.

If “informative” priors are included, then the derivation of the prior from an elicitation process or empirical evidence should be detailed. If the prior assumptions are claimed to be “non-informative,” then this claim should be justified and sensitivity analysis given. The idea of “inference robustness” (Box and Tiao, 1973) is crucial: it would be best if competing models with similar evidential support, or alternative prior distributions, gave similar conclusions, but if this is not the case then the alternative conclusions must be clearly reported. Where possible, full posterior distributions should be given for major conclusions, particularly for skewed distributions.

Finally. We would like to thank, and apologise to, our publishers for being so patient with the repeatedly deferred deadlines. Special thanks are extended to Martyn Plummer for his contributions to the book and for keeping us on our toes with his persistent efforts at doing everything better than us. Special thanks also to Simon White for his contribution, and to four reviewers, whose comments were extremely helpful. Thanks also to our friends, colleagues, and families for their support and words of encouragement, such as “Have you not finished that bloody book yet?” Many thanks to the (tens of) thousands of users out there, whose patience, enthusiasm, and sense of humour are all very much appreciated. And finally, we are deeply grateful to all those who have freely contributed their knowledge and insight to the BUGS project over the years. We shall be thinking of you when we get to share out whatever minimal royalties come our way!

All MATLAB® files found in the book are available for download from the publisher's Web site. MATLAB is a registered trademarks of The Mathworks, Inc. For product information please contact:

The Math Works, Inc.
3 Apple Hill Drive
Natick, MA 01760-2098 USA
Tel: 508-647-7000
Fax: 508-647-7001
E-mail: info@mathworks.com
Web: www.mathworks.com

Contents

Preface	xiii
1 Introduction: Probability and parameters	1
1.1 Probability	1
1.2 Probability distributions	5
1.3 Calculating properties of probability distributions	7
1.4 Monte Carlo integration	8
2 Monte Carlo simulations using BUGS	13
2.1 Introduction to BUGS	13
2.1.1 Background	13
2.1.2 Directed graphical models	13
2.1.3 The BUGS language	15
2.1.4 Running BUGS models	16
2.1.5 Running WinBUGS for a simple example	17
2.2 DoodleBUGS	21
2.3 Using BUGS to simulate from distributions	22
2.4 Transformations of random variables	24
2.5 Complex calculations using Monte Carlo	26
2.6 Multivariate Monte Carlo analysis	27
2.7 Predictions with unknown parameters	29
3 Introduction to Bayesian inference	33
3.1 Bayesian learning	33
3.1.1 Bayes' theorem for observable quantities	33
3.1.2 Bayesian inference for parameters	34
3.2 Posterior predictive distributions	36
3.3 Conjugate Bayesian inference	36
3.3.1 Binomial data	37
3.3.2 Normal data with unknown mean, known variance	41
3.4 Inference about a discrete parameter	45
3.5 Combinations of conjugate analyses	49
3.6 Bayesian and classical methods	51
3.6.1 Likelihood-based inference	52
3.6.2 Exchangeability	52
3.6.3 Long-run properties of Bayesian methods	53

3.6.4	Model-based vs procedural methods	54
3.6.5	The “likelihood principle”	55
4	Introduction to Markov chain Monte Carlo methods	57
4.1	Bayesian computation	57
4.1.1	Single-parameter models	57
4.1.2	Multi-parameter models	59
4.1.3	Monte Carlo integration for evaluating posterior integrals	61
4.2	Markov chain Monte Carlo methods	62
4.2.1	Gibbs sampling	63
4.2.2	Gibbs sampling and directed graphical models	64
4.2.3	Derivation of full conditional distributions in BUGS	68
4.2.4	Other MCMC methods	68
4.3	Initial values	70
4.4	Convergence	71
4.4.1	Detecting convergence/stationarity by eye	72
4.4.2	Formal detection of convergence/stationarity	73
4.5	Efficiency and accuracy	77
4.5.1	Monte Carlo standard error of the posterior mean	77
4.5.2	Accuracy of the whole posterior	78
4.6	Beyond MCMC	79
5	Prior distributions	81
5.1	Different purposes of priors	81
5.2	Vague, “objective,” and “reference” priors	82
5.2.1	Introduction	82
5.2.2	Discrete uniform distributions	83
5.2.3	Continuous uniform distributions and Jeffreys prior	83
5.2.4	Location parameters	84
5.2.5	Proportions	84
5.2.6	Counts and rates	85
5.2.7	Scale parameters	87
5.2.8	Distributions on the positive integers	88
5.2.9	More complex situations	89
5.3	Representation of informative priors	89
5.3.1	Elicitation of pure judgement	90
5.3.2	Discounting previous data	93
5.4	Mixture of prior distributions	95
5.5	Sensitivity analysis	97

6	Regression models	103
6.1	Linear regression with normal errors	103
6.2	Linear regression with non-normal errors	107
6.3	Non-linear regression with normal errors	109
6.4	Multivariate responses	112
6.5	Generalised linear regression models	114
6.6	Inference on functions of parameters	118
6.7	Further reading	119
7	Categorical data	121
7.1	2×2 tables	121
7.1.1	Tables with one margin fixed	122
7.1.2	Case-control studies	125
7.1.3	Tables with both margins fixed	126
7.2	Multinomial models	126
7.2.1	Conjugate analysis	126
7.2.2	Non-conjugate analysis — parameter constraints . . .	128
7.2.3	Categorical data with covariates	129
7.2.4	Multinomial and Poisson regression equivalence . . .	131
7.2.5	Contingency tables	132
7.3	Ordinal regression	132
7.4	Further reading	134
8	Model checking and comparison	137
8.1	Introduction	137
8.2	Deviance	138
8.3	Residuals	140
8.3.1	Standardised Pearson residuals	140
8.3.2	Multivariate residuals	142
8.3.3	Observed p -values for distributional shape	143
8.3.4	Deviance residuals and tests of fit	145
8.4	Predictive checks and Bayesian p -values	147
8.4.1	Interpreting discrepancy statistics — how big is big? .	147
8.4.2	Out-of-sample prediction	148
8.4.3	Checking functions based on data alone	148
8.4.4	Checking functions based on data and parameters . .	152
8.4.5	Goodness of fit for grouped data	155
8.5	Model assessment by embedding in larger models	157
8.6	Model comparison using deviances	159
8.6.1	p_D : The effective number of parameters	159
8.6.2	Issues with p_D	161
8.6.3	Alternative measures of the effective number of pa- rameters	164
8.6.4	DIC for model comparison	165
8.6.5	How and why does WinBUGS partition DIC and p_D ? .	167

8.6.6	Alternatives to DIC	168
8.7	Bayes factors	169
8.7.1	Lindley–Bartlett paradox in model selection	171
8.7.2	Computing marginal likelihoods	172
8.8	Model uncertainty	173
8.8.1	Bayesian model averaging	173
8.8.2	MCMC sampling over a space of models	173
8.8.3	Model averaging when all models are wrong	175
8.8.4	Model expansion	176
8.9	Discussion on model comparison	177
8.10	Prior-data conflict	178
8.10.1	Identification of prior-data conflict	179
8.10.2	Accommodation of prior-data conflict	180
9	Issues in Modelling	185
9.1	Missing data	185
9.1.1	Missing response data	186
9.1.2	Missing covariate data	189
9.2	Prediction	193
9.3	Measurement error	195
9.4	Cutting feedback	201
9.5	New distributions	204
9.5.1	Specifying a new sampling distribution	204
9.5.2	Specifying a new prior distribution	205
9.6	Censored, truncated, and grouped observations	206
9.6.1	Censored observations	206
9.6.2	Truncated sampling distributions	208
9.6.3	Grouped, rounded, or interval-censored data	209
9.7	Constrained parameters	211
9.7.1	Univariate fully specified prior distributions	211
9.7.2	Multivariate fully specified prior distributions	211
9.7.3	Prior distributions with unknown parameters	214
9.8	Bootstrapping	214
9.9	Ranking	215
10	Hierarchical models	219
10.1	Exchangeability	219
10.2	Priors	223
10.2.1	Unit-specific parameters	223
10.2.2	Parameter constraints	223
10.2.3	Priors for variance components	225
10.3	Hierarchical regression models	227
10.3.1	Data formatting	230
10.4	Hierarchical models for variances	237
10.5	Redundant parameterisations	240