Bayesian Modeling
Using WinBUGS

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

Preface xvi
Acknowledgments xix
Acronyms xx

1 Introduction to Bayesian Inference 1

1.1 Introduction: Bayesian modeling in the 21st century 1
1.2 Definition of statistical models 3
1.3 Bayes theorem 3
1.4 Model-based Bayesian inference 4
1.5 Inference using conjugate prior distributions 7
  1.5.1 Inference for the Poisson rate of count data 7
  1.5.2 Inference for the success probability of binomial data 8
  1.5.3 Inference for the mean of normal data with known variance 9
  1.5.4 Inference for the mean and variance of normal data 11
  1.5.5 Inference for normal regression models 12
  1.5.6 Other conjugate prior distributions 14
  1.5.7 Illustrative examples 14
1.6 Nonconjugate analysis 24
Problems 27
2 Markov Chain Monte Carlo Algorithms in Bayesian Inference

2.1 Simulation, Monte Carlo integration, and their implementation in Bayesian inference 31

2.2 Markov chain Monte Carlo methods
   2.2.1 The algorithm 36
   2.2.2 Terminology and implementation details 37

2.3 Popular MCMC algorithms
   2.3.1 The Metropolis–Hastings algorithm 42
   2.3.2 Componentwise Metropolis–Hastings 45
   2.3.3 The Gibbs sampler 71
   2.3.4 Metropolis within Gibbs 76
   2.3.5 The slice Gibbs sampler 76
   2.3.6 A simple example using the slice sampler 77

2.4 Summary and closing remarks 81
Problems 81

3 WinBUGS Software: Introduction, Setup, and Basic Analysis

3.1 Introduction and historical background 83
3.2 The WinBUGS environment
   3.2.1 Downloading and installing WinBUGS 84
   3.2.2 A short description of the menus 85
3.3 Preliminaries on using WinBUGS
   3.3.1 Code structure and type of parameters/nodes 88
   3.3.2 Scalar, vector, matrix, and array nodes 89
3.4 Building Bayesian models in WinBUGS
   3.4.1 Function description 93
   3.4.2 Using the for syntax and array, matrix, and vector calculations 97
   3.4.3 Use of parentheses, brackets and curly braces in WinBUGS 98
   3.4.4 Differences between WinBUGS and R/Sp1us syntax 98
   3.4.5 Model specification in WinBUGS 99
   3.4.6 Data and initial value specification 100
   3.4.7 An example of a complete model specification 107
   3.4.8 Data transformations 108
3.5 Compiling the model and simulating values 108
3.6 Basic output analysis using the sample monitor tool 117
3.7 Summarizing the procedure 120
3.8 Chapter summary and concluding comments 121
Problems 121

4 WinBUGS Software: Illustration, Results, and Further Analysis

4.1 A complete example of running MCMC in WinBUGS for a simple model 125
4.1.1 The model 125
4.1.2 Data and initial values 127
4.1.3 Compiling and running the model 127
4.1.4 MCMC output analysis and results 129

4.2 Further output analysis using the inference menu 132
4.2.1 Comparison of nodes 133
4.2.2 Calculation of correlations 136
4.2.3 Using the summary tool 137
4.2.4 Evaluation and ranking of individuals 138
4.2.5 Calculation of deviance information criterion 140

4.3 Multiple chains 141
4.3.1 Generation of multiple chains 141
4.3.2 Output analysis 142
4.3.3 The Gelman–Rubin convergence diagnostic 143

4.4 Changing the properties of a figure 145
4.4.1 General graphical options 145
4.4.2 Special graphical options 145

4.5 Other tools and menus 148
4.5.1 The node info tool 148
4.5.2 Monitoring the acceptance rate of the Metropolis–Hastings algorithm 148
4.5.3 Saving the current state of the chain 149
4.5.4 Setting the starting seed number 149
4.5.5 Running the model as a script 149

4.6 Summary and concluding remarks 149
Problems 150

5 Introduction to Bayesian Models: Normal Models 151

5.1 General modeling principles 151
5.2 Model specification in normal regression models 152
5.2.1 Specifying the likelihood 153
5.2.2 Specifying a simple independent prior distribution 154
5.2.3 Interpretation of the regression coefficients 154
5.2.4 A regression example using WinBUGS 157
5.3 Using vectors and multivariate priors in normal regression models 161
5.3.1 Defining the model using matrices 161
5.3.2 Prior distributions for normal regression models 162
5.3.3 Multivariate normal priors in WinBUGS 163
5.3.4 Continuation of Example 5.1 164
5.4 Analysis of variance models 167
5.4.1 The one-way ANOVA model 167
5.4.2 Parametrization and parameter interpretation 168
5.4.3 One-way ANOVA model in WinBUGS 169
5.4.4 A one-way ANOVA example using WinBUGS 171
5.4.5 Two-way ANOVA models 173
5.4.6 Multifactor analysis of variance 184
Problems 184

6 Incorporating Categorical Variables in Normal Models and Further Modeling Issues 189

6.1 Analysis of variance models using dummy variables 191
6.2 Analysis of covariance models 195
6.2.1 Models using one quantitative variable and one qualitative variable 197
6.2.2 The parallel lines model 197
6.2.3 The separate lines model 201
6.3 A bioassay example 203
6.3.1 Parallel lines analysis 204
6.3.2 Slope ratio analysis: Models with common intercept and different slope 212
6.3.3 Comparison of the two approaches 217
6.4 Further modeling issues 218
6.4.1 Extending the simple ANCOVA model 218
6.4.2 Using binary indicators to specify models in multiple regression 219
6.4.3 Selection of variables using the deviance information criterion (DIC) 219
6.5 Closing remarks 226
Problems 226

7 Introduction to Generalized Linear Models: Binomial and Poisson Data 229

7.1 Introduction 229
7.1.1 The exponential family 230
7.1.2 Common distributions as members of the exponential family 231
7.1.3 Link functions 234
7.1.4 Common generalized linear models 236
7.1.5 Interpretation of GLM coefficients 238
7.2 Prior distributions 239
7.3 Posterior inference 241
7.3.1 The posterior distribution of a generalized linear model 241
7.3.2 GLM specification in WinBUGS 242
7.4 Poisson regression models 242
7.4.1 Interpretation of Poisson log-linear parameters 242
7.4.2 A simple Poisson regression example 245
### 8 Models for Positive Continuous Data, Count Data, and Other GLM-Based Extensions

<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>8.1</td>
<td>Models with nonstandard distributions</td>
<td>275</td>
</tr>
<tr>
<td>8.1.1</td>
<td>Specification of arbitrary likelihood using the zeros–ones trick</td>
<td>276</td>
</tr>
<tr>
<td>8.1.2</td>
<td>The inverse Gaussian model</td>
<td>277</td>
</tr>
<tr>
<td>8.2</td>
<td>Models for positive continuous response variables</td>
<td>279</td>
</tr>
<tr>
<td>8.2.1</td>
<td>The gamma model</td>
<td>279</td>
</tr>
<tr>
<td>8.2.2</td>
<td>Other models</td>
<td>280</td>
</tr>
<tr>
<td>8.2.3</td>
<td>An example</td>
<td>281</td>
</tr>
<tr>
<td>8.3</td>
<td>Additional models for count data</td>
<td>282</td>
</tr>
<tr>
<td>8.3.1</td>
<td>The negative binomial model</td>
<td>283</td>
</tr>
<tr>
<td>8.3.2</td>
<td>The generalized Poisson model</td>
<td>286</td>
</tr>
<tr>
<td>8.3.3</td>
<td>Zero inflated models</td>
<td>288</td>
</tr>
<tr>
<td>8.3.4</td>
<td>The bivariate Poisson model</td>
<td>291</td>
</tr>
<tr>
<td>8.3.5</td>
<td>The Poisson difference model</td>
<td>293</td>
</tr>
<tr>
<td>8.4</td>
<td>Further GLM-based models and extensions</td>
<td>296</td>
</tr>
<tr>
<td>8.4.1</td>
<td>Survival analysis models</td>
<td>297</td>
</tr>
<tr>
<td>8.4.2</td>
<td>Multinomial models</td>
<td>298</td>
</tr>
<tr>
<td>8.4.3</td>
<td>Additional models and further reading</td>
<td>300</td>
</tr>
<tr>
<td></td>
<td>Problems</td>
<td>301</td>
</tr>
</tbody>
</table>

### 9 Bayesian Hierarchical Models

<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>9.1</td>
<td>Introduction</td>
<td>305</td>
</tr>
<tr>
<td>9.1.1</td>
<td>A simple motivating example</td>
<td>306</td>
</tr>
<tr>
<td>9.1.2</td>
<td>Why use a hierarchical model?</td>
<td>307</td>
</tr>
<tr>
<td>9.1.3</td>
<td>Other advantages and characteristics</td>
<td>308</td>
</tr>
<tr>
<td>9.2</td>
<td>Some simple examples</td>
<td>308</td>
</tr>
<tr>
<td>9.2.1</td>
<td>Repeated measures data</td>
<td>308</td>
</tr>
<tr>
<td>9.2.2</td>
<td>Introducing random effects in performance parameters</td>
<td>313</td>
</tr>
<tr>
<td>9.2.3</td>
<td>Poisson mixture models for count data</td>
<td>315</td>
</tr>
<tr>
<td>9.2.4</td>
<td>The use of hierarchical models in meta-analysis</td>
<td>318</td>
</tr>
<tr>
<td>9.3</td>
<td>The generalized linear mixed model formulation</td>
<td>320</td>
</tr>
<tr>
<td>9.3.1</td>
<td>A hierarchical normal model: A simple crossover trial</td>
<td>321</td>
</tr>
<tr>
<td>9.3.2</td>
<td>Logit GLMM for correlated binary responses</td>
<td>325</td>
</tr>
</tbody>
</table>
9.3.3 Poisson log-linear GLMMs for correlated count data 333
9.4 Discussion, closing remarks, and further reading 338
Problems 340

10 The Predictive Distribution and Model Checking 341
10.1 Introduction 341
    10.1.1 Prediction within Bayesian framework 341
    10.1.2 Using posterior predictive densities for model evaluation and checking 342
    10.1.3 Cross-validation predictive densities 344
10.2 Estimating the predictive distribution for future or missing observations using MCMC 344
    10.2.1 A simple example: Estimating missing observations 345
    10.2.2 An example of Bayesian prediction using a simple model 347
10.3 Using the predictive distribution for model checking 354
    10.3.1 Comparison of actual and predictive frequencies for discrete data 354
    10.3.2 Comparison of cumulative frequencies for predictive and actual values for continuous data 357
    10.3.3 Comparison of ordered predictive and actual values for continuous data 358
    10.3.4 Estimation of the posterior predictive ordinate 359
    10.3.5 Checking individual observations using residuals 362
    10.3.6 Checking structural assumptions of the model 365
    10.3.7 Checking the goodness-of-fit of a model 368
10.4 Using cross-validation predictive densities for model checking, evaluation, and comparison 375
    10.4.1 Estimating the conditional predictive ordinate 375
    10.4.2 Generating values from the leave-one-out cross-validatory predictive distributions 377
10.5 Illustration of a complete predictive analysis: Normal regression models 378
    10.5.1 Checking structural assumptions of the model 378
    10.5.2 Detailed checks based on residual analysis 379
    10.5.3 Overall goodness-of-fit of the model 380
    10.5.4 Implementation using WinBUGS 380
    10.5.5 An Illustrative example 383
    10.5.6 Summary of the model checking procedure 386
10.6 Discussion 387
Problems 387

11 Bayesian Model and Variable Evaluation 389
11.1 Prior predictive distributions as measures of model comparison: Posterior model odds and Bayes factors 389

11.2 Sensitivity of the posterior model probabilities: The Lindley–Bartlett paradox 391

11.3 Computation of the marginal likelihood 392
   11.3.1 Approximations based on the normal distribution 392
   11.3.2 Sampling from the prior: A naive Monte Carlo estimator 392
   11.3.3 Sampling from the posterior: The harmonic mean estimator 393
   11.3.4 Importance sampling estimators 394
   11.3.5 Bridge sampling estimators 394
   11.3.6 Chib’s marginal likelihood estimator 395
   11.3.7 Additional details and further reading 397

11.4 Computation of the marginal likelihood using WinBUGS 397
   11.4.1 A beta–binomial example 399
   11.4.2 A normal regression example with conjugate normal–inverse gamma prior 403

11.5 Bayesian variable selection using Gibbs-based methods 405
   11.5.1 Prior distributions for variable selection in GLM 406
   11.5.2 Gibbs variable selection 409
   11.5.3 Other Gibbs-based methods for variable selection 410

11.6 Posterior inference using the output of Bayesian variable selection samplers 412

11.7 Implementation of Gibbs variable selection in WinBUGS using an illustrative example 414

11.8 The Carlin–Chib method 419

11.9 reversible jump MCMC (RJMCMC) 420

11.10 Using posterior predictive densities for model evaluation 421
   11.10.1 Estimation from an MCMC output 423
   11.10.2 A simple example in WinBUGS 424

11.11 Information criteria 424
   11.11.1 The Bayes information criterion (BIC) 425
   11.11.2 The Akaike information criterion (AIC) 426
   11.11.3 Other criteria 427
   11.11.4 Calculation of penalized deviance measures from the MCMC output 428
   11.11.5 Implementation in WinBUGS 428
   11.11.6 A simple example in WinBUGS 429

11.12 Discussion and further reading 432

Problems 432

Appendix A: Model Specification via Directed Acyclic Graphs: The DOODLE Menu 435

A.1 Introduction: Starting with DOODLE 435
A.2 Nodes 436
A.3 Edges 438
A.4 Panels 438
A.5 A simple example 439

Appendix B: The Batch Mode: Running a Model in the Background Using Scripts 443
B.1 Introduction 443
B.2 Basic commands: Compiling and running the model 444

Appendix C: Checking Convergence Using CODA/BOA 447
C.1 Introduction 447
C.2 A short historical review 448
C.3 Diagnostics implemented by CODA/BOA 448
C.3.1 The Geweke diagnostic 448
C.3.2 The Gelman–Rubin diagnostic 449
C.3.3 The Raftery–Lewis diagnostic 449
C.3.4 The Heidelberger–Welch diagnostic 449
C.3.5 Final remarks 450
C.4 A first look at CODA/BOA 450
C.4.1 CODA 450
C.4.2 BOA 451
C.5 A simple example 453
C.5.1 Illustration in CODA 453
C.5.2 Illustration in BOA 457

Appendix D: Notation Summary 461
D.1 MCMC 461
D.2 Subscripts and indices 462
D.3 Parameters 462
D.4 Random variables and data 463
D.5 Sample estimates 463
D.6 Special functions, vectors, and matrices 464
D.7 Distributions 464
D.8 Distribution-related notation 465
D.9 Notation used in ANOVA and ANCOVA 466
D.10 Variable and model specification 466
D.11 Deviance information criterion (DIC) 466
D.12 Predictive measures 467
<table>
<thead>
<tr>
<th>CONTENTS</th>
<th>XV</th>
</tr>
</thead>
<tbody>
<tr>
<td>References</td>
<td>469</td>
</tr>
<tr>
<td>Index</td>
<td>485</td>
</tr>
</tbody>
</table>