Bayesian Biostatistics

EMMANUEL LESAFFRE
Erasmus MC, Rotterdam, The Netherlands and K.U. Leuven, Belgium

ANDREW B. LAWSON
Medical University of South Carolina, Charleston, USA

An authoritative guide to the complex world of Bayesian biomedical statistics

The growth of biostatistics has been phenomenal in recent years and has been marked by considerable technical innovation in both methodology and computational practicality. One area that has experienced significant growth is Bayesian methods. The growing use of Bayesian methodology has taken place partly due to an increasing number of practitioners valuing the Bayesian paradigm as matching that of scientific discovery. In addition, computational advances have allowed for more complex models to be fitted routinely to realistic data sets.

Through examples, exercises and a combination of introductory and more advanced chapters, this book provides an invaluable understanding of the complex world of biomedical statistics illustrated via a diverse range of applications taken from epidemiology, exploratory clinical studies, health promotion studies, image analysis and clinical trials.

Key Features:

- **Provides an authoritative account of Bayesian methodology, from its most basic elements to its practical implementation, with an emphasis on healthcare techniques.**
- **Contains introductory explanations of Bayesian principles common to all areas of application.**
- **Presents clear and concise examples in biostatistics applications such as clinical trials, longitudinal studies, bioassay, survival, image analysis and bioinformatics.**
- **Illustrated throughout with examples using software including WinBUGS, OpenBUGS, SAS and various dedicated R programs.**
- **Highlights the differences between the Bayesian and classical approaches.**
- **Supported by an accompanying website hosting free software and case study guides.**

Bayesian Biostatistics introduces the reader smoothly into the Bayesian statistical methods with chapters that gradually increase in level of complexity. Master students in biostatistics, applied statisticians and all researchers with a good background in classical statistics who have an interest in Bayesian methods will find this book useful.

STATISTICS IN PRACTICE

A series of practical books outlining the use of statistical techniques in a wide range of applications areas:

- HUMAN AND BIOLOGICAL SCIENCES
- EARTH AND ENVIRONMENTAL SCIENCES
- INDUSTRY, COMMERCE AND FINANCE

Also available as an e-book

WILEY
wiley.com

ISBN 978-0-470-01823-1

Preface xiii

Notation, terminology and some guidance for reading the book xvii

Part I BASIC CONCEPTS IN BAYESIAN METHODS

1 Modes of statistical inference 3
1.1 The frequentist approach: A critical reflection 4
1.1.1 The classical statistical approach 4
1.1.2 The *P*-value as a measure of evidence 5
1.1.3 The confidence interval as a measure of evidence 8
1.1.4 An historical note on the two frequentist paradigms* 8
1.2 Statistical inference based on the likelihood function 10
1.2.1 The likelihood function 10
1.2.2 The likelihood principles 11
1.3 The Bayesian approach: Some basic ideas 14
1.3.1 Introduction 14
1.3.2 Bayes theorem – discrete version for simple events 15
1.4 Outlook 18
Exercises 19

2 Bayes theorem: Computing the posterior distribution 20
2.1 Introduction 20
2.2 Bayes theorem – the binary version 20
2.3 Probability in a Bayesian context 21
2.4 Bayes theorem – the categorical version 22
2.5 Bayes theorem – the continuous version 23
2.6 The binomial case 24
2.7 The Gaussian case 30
2.8 The Poisson case 36
2.9 The prior and posterior distribution of $h(\theta)$ 40
2.10 Bayesian versus likelihood approach 40

2.11 Bayesian versus frequentist approach 41
2.12 The different modes of the Bayesian approach 41
2.13 An historical note on the Bayesian approach 42
2.14 Closing remarks 44
Exercises 44

3 Introduction to Bayesian inference 46
3.1 Introduction 46
3.2 Summarizing the posterior by probabilities 46
3.3 Posterior summary measures 47
3.3.1 Characterizing the location and variability of the posterior distribution 47
3.3.2 Posterior interval estimation 49
3.4 Predictive distributions 51
3.4.1 The frequentist approach to prediction 52
3.4.2 The Bayesian approach to prediction 53
3.4.3 Applications 54
3.5 Exchangeability 58
3.6 A normal approximation to the posterior 60
3.6.1 A Bayesian analysis based on a normal approximation to the likelihood 60
3.6.2 Asymptotic properties of the posterior distribution 62
3.7 Numerical techniques to determine the posterior 63
3.7.1 Numerical integration 63
3.7.2 Sampling from the posterior 65
3.7.3 Choice of posterior summary measures 72
3.8 Bayesian hypothesis testing 72
3.8.1 Inference based on credible intervals 72
3.8.2 The Bayes factor 74
3.8.3 Bayesian versus frequentist hypothesis testing 76
3.9 Closing remarks 78
Exercises 79

4 More than one parameter 82
4.1 Introduction 82
4.2 Joint versus marginal posterior inference 83
4.3 The normal distribution with μ and σ^2 unknown 83
4.3.1 No prior knowledge on μ and σ^2 is available 84
4.3.2 An historical study is available 86
4.3.3 Expert knowledge is available 88
4.4 Multivariate distributions 89
4.4.1 The multivariate normal and related distributions 89
4.4.2 The multinomial distribution 90
4.5 Frequentist properties of Bayesian inference 92
4.6 Sampling from the posterior distribution: The Method of Composition 93
4.7 Bayesian linear regression models 96
4.7.1 The frequentist approach to linear regression 96
4.7.2 A noninformative Bayesian linear regression model 97

4.7.3 Posterior summary measures for the linear regression model 98
4.7.4 Sampling from the posterior distribution 99
4.7.5 An informative Bayesian linear regression model 101
4.8 Bayesian generalized linear models 101
4.9 More complex regression models 102
4.10 Closing remarks 102
Exercises 102

5 Choosing the prior distribution 104
5.1 Introduction 104
5.2 The sequential use of Bayes theorem 104
5.3 Conjugate prior distributions 106
5.3.1 Univariate data distributions 106
5.3.2 Normal distribution – mean and variance unknown 109
5.3.3 Multivariate data distributions 110
5.3.4 Conditional conjugate and semiconjugate distributions 111
5.3.5 Hyperpriors 112
5.4 Noninformative prior distributions 113
5.4.1 Introduction 113
5.4.2 Expressing ignorance 114
5.4.3 General principles to choose noninformative priors 115
5.4.4 Improper prior distributions 119
5.4.5 Weak/vague priors 120
5.5 Informative prior distributions 121
5.5.1 Introduction 121
5.5.2 Data-based prior distributions 121
5.5.3 Elicitation of prior knowledge 122
5.5.4 Archetypal prior distributions 126
5.6 Prior distributions for regression models 129
5.6.1 Normal linear regression 129
5.6.2 Generalized linear models 131
5.6.3 Specification of priors in Bayesian software 134
5.7 Modeling priors 134
5.8 Other regression models 136
5.9 Closing remarks 136
Exercises 137

6 Markov chain Monte Carlo sampling 139
6.1 Introduction 139
6.2 The Gibbs sampler 140
6.2.1 The bivariate Gibbs sampler 140
6.2.2 The general Gibbs sampler 146
6.2.3 Remarks* 150
6.2.4 Review of Gibbs sampling approaches 152
6.2.5 The Slice sampler* 153
6.3 The Metropolis(–Hastings) algorithm 154
6.3.1 The Metropolis algorithm 155
6.3.2 The Metropolis–Hastings algorithm 157

6.3.3 Remarks* 159
6.3.4 Review of Metropolis(–Hastings) approaches 161
6.4 Justification of the MCMC approaches* 162
6.4.1 Properties of the MH algorithm 164
6.4.2 Properties of the Gibbs sampler 165
6.5 Choice of the sampler 165
6.6 The Reversible Jump MCMC algorithm* 168
6.7 Closing remarks 172
Exercises 173

7 Assessing and improving convergence of the Markov chain 175
7.1 Introduction 175
7.2 Assessing convergence of a Markov chain 176
7.2.1 Definition of convergence for a Markov chain 176
7.2.2 Checking convergence of the Markov chain 176
7.2.3 Graphical approaches to assess convergence 177
7.2.4 Formal diagnostic tests 180
7.2.5 Computing the Monte Carlo standard error 186
7.2.6 Practical experience with the formal diagnostic procedures 188
7.3 Accelerating convergence 189
7.3.1 Introduction 189
7.3.2 Acceleration techniques 189
7.4 Practical guidelines for assessing and accelerating convergence 194
7.5 Data augmentation 195
7.6 Closing remarks 200
Exercises 201

8 Software 202
8.1 WinBUGS and related software 202
8.1.1 A first analysis 203
8.1.2 Information on samplers 206
8.1.3 Assessing and accelerating convergence 207
8.1.4 Vector and matrix manipulations 208
8.1.5 Working in batch mode 210
8.1.6 Troubleshooting 212
8.1.7 Directed acyclic graphs 212
8.1.8 Add-on modules: GeoBUGS and PKBUGS 214
8.1.9 Related software 214
8.2 Bayesian analysis using SAS 215
8.2.1 Analysis using procedure GENMOD 215
8.2.2 Analysis using procedure MCMC 217
8.2.3 Other Bayesian programs 220
8.3 Additional Bayesian software and comparisons 221
8.3.1 Additional Bayesian software 221
8.3.2 Comparison of Bayesian software 222
8.4 Closing remarks 222
Exercises 223

Part II BAYESIAN TOOLS FOR STATISTICAL MODELING

9 Hierarchical models 227
9.1 Introduction 227
9.2 The Poisson-gamma hierarchical model 228
9.2.1 Introduction 228
9.2.2 Model specification 229
9.2.3 Posterior distributions 231
9.2.4 Estimating the parameters 232
9.2.5 Posterior predictive distributions 237
9.3 Full versus empirical Bayesian approach 238
9.4 Gaussian hierarchical models 240
9.4.1 Introduction 240
9.4.2 The Gaussian hierarchical model 240
9.4.3 Estimating the parameters 241
9.4.4 Posterior predictive distributions 243
9.4.5 Comparison of FB and EB approach 244
9.5 Mixed models 244
9.5.1 Introduction 244
9.5.2 The linear mixed model 244
9.5.3 The generalized linear mixed model 248
9.5.4 Nonlinear mixed models 253
9.5.5 Some further extensions 256
9.5.6 Estimation of the random effects and posterior predictive distributions 256
9.5.7 Choice of the level-2 variance prior 258
9.6 Propriety of the posterior 260
9.7 Assessing and accelerating convergence 261
9.8 Comparison of Bayesian and frequentist hierarchical models 263
9.8.1 Estimating the level-2 variance 263
9.8.2 ML and REML estimates compared with Bayesian estimates 264
9.9 Closing remarks 265
Exercises 265

10 Model building and assessment 267
10.1 Introduction 267
10.2 Measures for model selection 268
10.2.1 The Bayes factor 268
10.2.2 Information theoretic measures for model selection 274
10.2.3 Model selection based on predictive loss functions 286
10.3 Model checking 288
10.3.1 Introduction 288
10.3.2 Model-checking procedures 289
10.3.3 Sensitivity analysis 295
10.3.4 Posterior predictive checks 300
10.3.5 Model expansion 308
10.4 Closing remarks 316
Exercises 316

11 Variable selection 319
11.1 Introduction 319
11.2 Classical variable selection 320
11.2.1 Variable selection techniques 320
11.2.2 Frequentist regularization 322
11.3 Bayesian variable selection: Concepts and questions 325
11.4 Introduction to Bayesian variable selection 326
11.4.1 Variable selection for K small 326
11.4.2 Variable selection for K large 330
11.5 Variable selection based on Zellner's g-prior 333
11.6 Variable selection based on Reversible Jump Markov chain Monte Carlo 336
11.7 Spike and slab priors 339
11.7.1 Stochastic Search Variable Selection 340
11.7.2 Gibbs Variable Selection 343
11.7.3 Dependent variable selection using SSVS 345
11.8 Bayesian regularization 345
11.8.1 Bayesian LASSO regression 346
11.8.2 Elastic Net and further extensions of the Bayesian LASSO 350
11.9 The many regressors case 351
11.10 Bayesian model selection 355
11.11 Bayesian model averaging 357
11.12 Closing remarks 359
Exercises 360

Part III BAYESIAN METHODS IN PRACTICAL APPLICATIONS

12 Bioassay 365
12.1 Bioassay essentials 365
12.1.1 Cell assays 365
12.1.2 Animal assays 366
12.2 A generic *in vitro* example 369
12.3 Ames/Salmonella mutagenic assay 371
12.4 Mouse lymphoma assay (L5178Y TK+/−) 373
12.5 Closing remarks 374

13 Measurement error 375
13.1 Continuous measurement error 375
13.1.1 Measurement error in a variable 375
13.1.2 Two types of measurement error on the predictor in linear and nonlinear models 376
13.1.3 Accommodation of predictor measurement error 378
13.1.4 Nonadditive errors and other extensions 382
13.2 Discrete measurement error 382
13.2.1 Sources of misclassification 382
13.2.2 Misclassification in the binary predictor 383
13.2.3 Misclassification in a binary response 386
13.3 Closing remarks 389

14 Survival analysis 390
14.1 Basic terminology 390
14.1.1 Endpoint distributions 391
14.1.2 Censoring 392
14.1.3 Random effect specification 393
14.1.4 A general hazard model 393
14.1.5 Proportional hazards 394
14.1.6 The Cox model with random effects 394
14.2 The Bayesian model formulation 394
14.2.1 A Weibull survival model 395
14.2.2 A Bayesian AFT model 397
14.3 Examples 397
14.3.1 The gastric cancer study 397
14.3.2 Prostate cancer in Louisiana: A spatial AFT model 401
14.4 Closing remarks 406

15 Longitudinal analysis 407
15.1 Fixed time periods 407
15.1.1 Introduction 407
15.1.2 A classical growth-curve example 408
15.1.3 Alternate data models 414
15.2 Random event times 417
15.3 Dealing with missing data 420
15.3.1 Introduction 420
15.3.2 Response missingness 421
15.3.3 Missingness mechanisms 422
15.3.4 Bayesian considerations 424
15.3.5 Predictor missingness 424
15.4 Joint modeling of longitudinal and survival responses 424
15.4.1 Introduction 424
15.4.2 An example 425
15.5 Closing remarks 429

16 Spatial applications: Disease mapping and image analysis 430
16.1 Introduction 430
16.2 Disease mapping 430
16.2.1 Some general spatial epidemiological issues 431
16.2.2 Some spatial statistical issues 433
16.2.3 Count data models 433
16.2.4 A special application area: Disease mapping/risk estimation 434
16.2.5 A special application area: Disease clustering 438
16.2.6 A special application area: Ecological analysis 443
16.3 Image analysis 444
16.3.1 fMRI modeling 446
16.3.2 A note on software 455

17 Final chapter 456
17.1 What this book covered 456

17.2 Additional Bayesian developments 456
17.2.1 Medical decision making 456
17.2.2 Clinical trials 457
17.2.3 Bayesian networks 457
17.2.4 Bioinformatics 458
17.2.5 Missing data 458
17.2.6 Mixture models 458
17.2.7 Nonparametric Bayesian methods 459
17.3 Alternative reading 459

Appendix: Distributions **460**

A.1 Introduction 460
A.2 Continuous univariate distributions 461
A.3 Discrete univariate distributions 477
A.4 Multivariate distributions 481

References **484**

Index **509**