# Contents

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>List of models</td>
<td>xv</td>
</tr>
<tr>
<td>List of examples</td>
<td>xvii</td>
</tr>
<tr>
<td>Preface</td>
<td>xix</td>
</tr>
</tbody>
</table>

## Part I: Fundamentals of Bayesian Inference

### 1 Background

1.1 Overview

1.2 General notation for statistical inference

1.3 Bayesian inference

1.4 Example: inference about a genetic probability

1.5 Probability as a measure of uncertainty

1.6 Example of probability assignment: football point spreads

1.7 Example of probability assignment: estimating the accuracy of record linkage

1.8 Some useful results from probability theory

1.9 Summarizing inferences by simulation

1.10 Computation and software

1.11 Bibliographic note

1.12 Exercises

### 2 Single-parameter models

2.1 Estimating a probability from binomial data

2.2 Posterior distribution as compromise between data and prior information

2.3 Summarizing posterior inference

2.4 Informative prior distributions

2.5 Example: estimating the probability of a female birth given placenta previa

2.6 Estimating the mean of a normal distribution with known variance

2.7 Other standard single-parameter models

2.8 Example: informative prior distribution and multilevel structure for estimating cancer rates
## CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.9</td>
<td>Noninformative prior distributions</td>
<td>61</td>
</tr>
<tr>
<td>2.10</td>
<td>Bibliographic note</td>
<td>65</td>
</tr>
<tr>
<td>2.11</td>
<td>Exercises</td>
<td>67</td>
</tr>
<tr>
<td>3</td>
<td>Introduction to multiparameter models</td>
<td>73</td>
</tr>
<tr>
<td>3.1</td>
<td>Averaging over ‘nuisance parameters’</td>
<td>73</td>
</tr>
<tr>
<td>3.2</td>
<td>Normal data with a noninformative prior distribution</td>
<td>74</td>
</tr>
<tr>
<td>3.3</td>
<td>Normal data with a conjugate prior distribution</td>
<td>78</td>
</tr>
<tr>
<td>3.4</td>
<td>Normal data with a semi-conjugate prior distribution</td>
<td>80</td>
</tr>
<tr>
<td>3.5</td>
<td>The multinomial model</td>
<td>83</td>
</tr>
<tr>
<td>3.6</td>
<td>The multivariate normal model</td>
<td>85</td>
</tr>
<tr>
<td>3.7</td>
<td>Example: analysis of a bioassay experiment</td>
<td>88</td>
</tr>
<tr>
<td>3.8</td>
<td>Summary of elementary modeling and computation</td>
<td>93</td>
</tr>
<tr>
<td>3.9</td>
<td>Bibliographic note</td>
<td>94</td>
</tr>
<tr>
<td>3.10</td>
<td>Exercises</td>
<td>95</td>
</tr>
<tr>
<td>4</td>
<td>Large-sample inference and frequency properties of Bayesian inference</td>
<td>101</td>
</tr>
<tr>
<td>4.1</td>
<td>Normal approximations to the posterior distribution</td>
<td>101</td>
</tr>
<tr>
<td>4.2</td>
<td>Large-sample theory</td>
<td>106</td>
</tr>
<tr>
<td>4.3</td>
<td>Counterexamples to the theorems</td>
<td>108</td>
</tr>
<tr>
<td>4.4</td>
<td>Frequency evaluations of Bayesian inferences</td>
<td>111</td>
</tr>
<tr>
<td>4.5</td>
<td>Bibliographic note</td>
<td>113</td>
</tr>
<tr>
<td>4.6</td>
<td>Exercises</td>
<td>113</td>
</tr>
<tr>
<td></td>
<td><strong>Part II: Fundamentals of Bayesian Data Analysis</strong></td>
<td>115</td>
</tr>
<tr>
<td>5</td>
<td>Hierarchical models</td>
<td>117</td>
</tr>
<tr>
<td>5.1</td>
<td>Constructing a parameterized prior distribution</td>
<td>118</td>
</tr>
<tr>
<td>5.2</td>
<td>Exchangeability and setting up hierarchical models</td>
<td>121</td>
</tr>
<tr>
<td>5.3</td>
<td>Computation with hierarchical models</td>
<td>125</td>
</tr>
<tr>
<td>5.4</td>
<td>Estimating an exchangeable set of parameters from a normal model</td>
<td>131</td>
</tr>
<tr>
<td>5.5</td>
<td>Example: combining information from educational testing experiments in eight schools</td>
<td>138</td>
</tr>
<tr>
<td>5.6</td>
<td>Hierarchical modeling applied to a meta-analysis</td>
<td>145</td>
</tr>
<tr>
<td>5.7</td>
<td>Bibliographic note</td>
<td>150</td>
</tr>
<tr>
<td>5.8</td>
<td>Exercises</td>
<td>152</td>
</tr>
<tr>
<td>6</td>
<td>Model checking and improvement</td>
<td>157</td>
</tr>
<tr>
<td>6.1</td>
<td>The place of model checking in applied Bayesian statistics</td>
<td>157</td>
</tr>
<tr>
<td>6.2</td>
<td>Do the inferences from the model make sense?</td>
<td>158</td>
</tr>
<tr>
<td>6.3</td>
<td>Is the model consistent with data? Posterior predictive checking</td>
<td>159</td>
</tr>
<tr>
<td>6.4</td>
<td>Graphical posterior predictive checks</td>
<td>165</td>
</tr>
</tbody>
</table>
### CONTENTS

6.5 Numerical posterior predictive checks 172  
6.6 Model expansion 177  
6.7 Model comparison 179  
6.8 Model checking for the educational testing example 186  
6.9 Bibliographic note 190  
6.10 Exercises 192  

7 Modeling accounting for data collection 197  
7.1 Introduction 197  
7.2 Formal models for data collection 200  
7.3 Ignorability 203  
7.4 Sample surveys 207  
7.5 Designed experiments 218  
7.6 Sensitivity and the role of randomization 223  
7.7 Observational studies 226  
7.8 Censoring and truncation 231  
7.9 Discussion 236  
7.10 Bibliographic note 237  
7.11 Exercises 239  

8 Connections and challenges 247  
8.1 Bayesian interpretations of other statistical methods 247  
8.2 Challenges in Bayesian data analysis 252  
8.3 Bibliographic note 255  
8.4 Exercises 255  

9 General advice 259  
9.1 Setting up probability models 259  
9.2 Posterior inference 264  
9.3 Model evaluation 265  
9.4 Summary 271  
9.5 Bibliographic note 271  

Part III: Advanced Computation 273  

10 Overview of computation 275  
10.1 Crude estimation by ignoring some information 276  
10.2 Use of posterior simulations in Bayesian data analysis 276  
10.3 Practical issues 278  
10.4 Exercises 282  

11 Posterior simulation 283  
11.1 Direct simulation 283  
11.2 Markov chain simulation 285  
11.3 The Gibbs sampler 287
11.4 The Metropolis and Metropolis-Hastings algorithms
11.5 Building Markov chain algorithms using the Gibbs sampler and Metropolis algorithm
11.6 Inference and assessing convergence
11.7 Example: the hierarchical normal model
11.8 Efficient Gibbs samplers
11.9 Efficient Metropolis jumping rules
11.10 Recommended strategy for posterior simulation
11.11 Bibliographic note
11.12 Exercises

12 Approximations based on posterior modes
12.1 Finding posterior modes
12.2 The normal and related mixture approximations
12.3 Finding marginal posterior modes using EM and related algorithms
12.4 Approximating conditional and marginal posterior densities
12.5 Example: the hierarchical normal model (continued)
12.6 Bibliographic note
12.7 Exercises

13 Special topics in computation
13.1 Advanced techniques for Markov chain simulation
13.2 Numerical integration
13.3 Importance sampling
13.4 Computing normalizing factors
13.5 Bibliographic note
13.6 Exercises

Part IV: Regression Models

14 Introduction to regression models
14.1 Introduction and notation
14.2 Bayesian analysis of the classical regression model
14.3 Example: estimating the advantage of incumbency in U.S. Congressional elections
14.4 Goals of regression analysis
14.5 Assembling the matrix of explanatory variables
14.6 Unequal variances and correlations
14.7 Models for unequal variances
14.8 Including prior information
14.9 Bibliographic note
14.10 Exercises
CONTENTS

15 Hierarchical linear models 389
   15.1 Regression coefficients exchangeable in batches 390
   15.2 Example: forecasting U.S. Presidential elections 392
   15.3 General notation for hierarchical linear models 399
   15.4 Computation 400
   15.5 Hierarchical modeling as an alternative to selecting predictors 405
   15.6 Analysis of variance 406
   15.7 Bibliographic note 411
   15.8 Exercises 412

16 Generalized linear models 415
   16.1 Introduction 415
   16.2 Standard generalized linear model likelihoods 416
   16.3 Setting up and interpreting generalized linear models 418
   16.4 Computation 421
   16.5 Example: hierarchical Poisson regression for police stops 425
   16.6 Example: hierarchical logistic regression for political opinions 428
   16.7 Models for multinomial responses 430
   16.8 Loglinear models for multivariate discrete data 433
   16.9 Bibliographic note 439
   16.10 Exercises 440

17 Models for robust inference 443
   17.1 Introduction 443
   17.2 Overdispersed versions of standard probability models 445
   17.3 Posterior inference and computation 448
   17.4 Robust inference and sensitivity analysis for the educational testing example 451
   17.5 Robust regression using Student-t errors 455
   17.6 Bibliographic note 457
   17.7 Exercises 458

Part V: Specific Models and Problems 461

18 Mixture models 463
   18.1 Introduction 463
   18.2 Setting up mixture models 463
   18.3 Computation 467
   18.4 Example: reaction times and schizophrenia 468
   18.5 Bibliographic note 479

19 Multivariate models 481
   19.1 Linear regression with multiple outcomes 481
   19.2 Prior distributions for covariance matrices 483
   19.3 Hierarchical multivariate models 486