# Table of Contents

**Preface**

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Astrostatistics</td>
<td>1</td>
</tr>
<tr>
<td>1.1 The Nature and Scope of Astrostatistics</td>
<td>1</td>
</tr>
<tr>
<td>1.2 The Recent Development of Astrostatistics</td>
<td>4</td>
</tr>
<tr>
<td>1.3 What is a Statistical Model?</td>
<td>6</td>
</tr>
<tr>
<td>1.4 Classification of Statistical Models</td>
<td>7</td>
</tr>
<tr>
<td>2 Prerequisites</td>
<td>9</td>
</tr>
<tr>
<td>2.1 Software</td>
<td>9</td>
</tr>
<tr>
<td>2.2 R</td>
<td>10</td>
</tr>
<tr>
<td>2.3 JAGS</td>
<td>12</td>
</tr>
<tr>
<td>2.4 Python</td>
<td>14</td>
</tr>
<tr>
<td>2.5 Stan</td>
<td>17</td>
</tr>
<tr>
<td>3 Frequentist vs. Bayesian Methods</td>
<td>23</td>
</tr>
<tr>
<td>3.1 Frequentist Statistics</td>
<td>23</td>
</tr>
<tr>
<td>3.1.1 Fitting a Linear Regression in R</td>
<td>25</td>
</tr>
<tr>
<td>3.1.2 Fitting a Linear Regression in Python</td>
<td>26</td>
</tr>
<tr>
<td>3.2 Basic Theory of Bayesian Modeling</td>
<td>27</td>
</tr>
<tr>
<td>3.2.1 Example: Calculating a Beta Prior and Posterior Analytically</td>
<td>32</td>
</tr>
<tr>
<td>3.2.2 Fitting a Simple Bayesian Normal Model using R</td>
<td>38</td>
</tr>
<tr>
<td>3.2.3 Fitting a Simple Bayesian Normal Model using Python</td>
<td>41</td>
</tr>
<tr>
<td>3.3 Selecting Between Frequentist and Bayesian Modeling</td>
<td>43</td>
</tr>
<tr>
<td>4 Normal Linear Models</td>
<td>46</td>
</tr>
<tr>
<td>4.1 The Gaussian or Normal Model</td>
<td>46</td>
</tr>
<tr>
<td>4.1.1 Bayesian Synthetic Normal Model in R using JAGS</td>
<td>48</td>
</tr>
<tr>
<td>4.1.2 Bayesian Synthetic Normal Model in R using JAGS and the Zero Trick</td>
<td>54</td>
</tr>
<tr>
<td>4.1.3 Bayesian Synthetic Normal Model in Python using Stan</td>
<td>56</td>
</tr>
</tbody>
</table>
## Contents

4.1.4 Bayesian Synthetic Normal Model using Stan with a Customized Likelihood 57
4.2 Multivariate Normal Model 58
  4.2.1 Multivariate Linear Regression in R using JAGS 58
  4.2.2 Multivariate Linear Regression in Python using Stan 60
4.3 Bayesian Errors-in-Measurements Modeling 61
  4.3.1 Generating Data with Errors using R 62
  4.3.2 Build Model ignoring Errors in R using JAGS 62
  4.3.3 Build Model including Errors in R using JAGS 63
  4.3.4 Bayesian Errors-in-Measurements Modeling in Python using Stan 65

5 GLMs Part I – Continuous and Binomial Models 68
  5.1 Brief Overview of Generalized Linear Models 68
  5.2 Bayesian Continuous Response Models 74
    5.2.1 Bayesian Lognormal Model 75
    5.2.2 Bayesian Gamma Models 82
    5.2.3 Bayesian Inverse Gaussian Models 87
    5.2.4 Bayesian Beta Model 92
  5.3 Bayesian Binomial Models 98
    5.3.1 Bayesian Bernoulli Logit Models 99
    5.3.2 Bayesian Bernoulli Probit Models 110
    5.3.3 Bayesian Grouped Logit or Binomial Model 117
    5.3.4 Bayesian Grouped Probit Model 124
    5.3.5 Bayesian Beta–Binomial Models 125

6 GLMs Part II – Count Models 135
  6.1 Bayesian Poisson Models 135
    6.1.1 Poisson Models with R 142
    6.1.2 Poisson Models with JAGS 143
    6.1.3 Poisson Models in Python 144
    6.1.4 Poisson Models in Python using Stan 147
  6.2 Bayesian Negative Binomial Models 148
    6.2.1 Modeling the Negative Binomial using JAGS 154
    6.2.2 Negative Binomial Models in Python using pycmc 161
    6.2.3 Modeling the Negative Binomial in Python using Stan 163
  6.3 Bayesian Generalized Poisson Model 164
    6.3.1 Generalized Poisson Model using JAGS 166
    6.3.2 Generalized Poisson Model using Stan 168
  6.4 Bayesian Zero-Truncated Models 169
    6.4.1 Bayesian Zero-Truncated Poisson Model 170
    6.4.2 Zero-Truncated Poisson in Python using Stan 174
    6.4.3 Bayesian Zero-Truncated Negative Binomial Model 176
  6.5 Bayesian Three-Parameter NB Model (NB-P) 179
## 6.5.1 Three-Parameter NB-P Model using JAGS

6.5.2 Three-Parameter NB-P Models in Python using Stan

## 7 GLMs Part III – Zero-Inflated and Hurdle Models

7.1 Bayesian Zero-Inflated Models

7.1.1 Bayesian Zero-Inflated Poisson Model

7.1.2 Bayesian Zero-Inflated Negative Binomial Model

7.2 Bayesian Hurdle Models

7.2.1 Bayesian Poisson–Logit Hurdle Model

7.2.2 Bayesian Negative Binomial–Logit Hurdle Model

7.2.3 Bayesian Gamma–Logit Hurdle Model

7.2.4 Bayesian Lognormal–Logit Hurdle Model

## 8 Hierarchical GLMMs

8.1 Overview of Bayesian Hierarchical Models/GLMMs

8.2 Bayesian Gaussian or Normal GLMMs

8.2.1 Random Intercept Gaussian Data

8.2.2 Bayesian Random Intercept Gaussian Model in R using JAGS

8.2.3 Bayesian Random Intercept Normal Model in R using JAGS

8.2.4 Bayesian Random Intercept Normal Model in Python using Stan

8.3 Bayesian Binary Logistic GLMMs

8.3.1 Random Intercept Binary Logistic Data

8.3.2 Bayesian Random Intercept Binary Logistic Model with R

8.3.3 Bayesian Random Intercept Binary Logistic Model with Python

8.3.4 Bayesian Random Intercept Binary Logistic Model in R using JAGS

8.3.5 Bayesian Random Intercept Binary Logistic Model in Python using Stan

8.4 Bayesian Binomial Logistic GLMMs

8.4.1 Random Intercept Binomial Logistic Data

8.4.2 Bayesian Random Intercept Binomial Logistic Model in R using JAGS

8.4.3 Bayesian Random Intercept Binomial Logistic Model in Python using Stan

8.5 Bayesian Poisson GLMMs

8.5.1 Random Intercept Poisson Data

8.5.2 Bayesian Random Intercept Poisson Model with R

8.5.3 Bayesian Random Intercept Poisson Model in Python

8.5.4 Bayesian Random Intercept Poisson Model in R using JAGS

8.5.5 Bayesian Random Intercept Poisson Model in Python using Stan

8.5.6 Bayesian Random-Intercept–Random-Slopes Poisson Model
## 8.6 Bayesian Negative Binomial GLMMs

- **8.6.1 Random Intercept Negative Binomial Data**
- **8.6.2 Random Intercept Negative Binomial MLE Model using R**
- **8.6.3 Bayesian Random Intercept Negative Binomial Model using Python**
- **8.6.4 Bayesian Random Intercept Negative Binomial Model in R using JAGS**
- **8.6.5 Bayesian Random Intercept Negative Binomial Model in Python using Stan**

## 9 Model Selection

- **9.1 Information Criteria Tests for Model Selection**
  - **9.1.1 Frequentist and Bayesian Information Criteria**
  - **9.1.2 Bayesian Deviance Statistic**
  - **9.1.3 pD and Deviance Information Criteria (DIC)**
- **9.2 Model Selection with Indicator Functions**
- **9.3 Bayesian LASSO**

## 10 Astronomical Applications

- **10.1 Normal Model, Black Hole Mass, and Bulge Velocity Dispersion**
  - **10.1.1 Data**
  - **10.1.2 The Statistical Model Formulation**
  - **10.1.3 Running the Model in R using JAGS**
  - **10.1.4 Running the Model in Python using Stan**
- **10.2 Gaussian Mixed Models, Type Ia Supernovae, and Hubble Residuals**
  - **10.2.1 Data**
  - **10.2.2 Statistical Model Formulation**
  - **10.2.3 Running the Model in R using JAGS**
  - **10.2.4 Running the Model in Python using Stan**
- **10.3 Multivariate Normal Mixed Model and Early-Type Contact Binaries**
  - **10.3.1 Data**
  - **10.3.2 The Statistical Model Formulation**
  - **10.3.3 Running the Model in R using JAGS**
  - **10.3.4 Running the Model in Python using Stan**
- **10.4 Lognormal Distribution and the Initial Mass Function**
  - **10.4.1 Data**
  - **10.4.2 Statistical Model Formulation**
  - **10.4.3 Running the Model in R using JAGS**
  - **10.4.4 Running the Model in Python using Stan**
- **10.5 Beta Model and the Baryon Content of Low Mass Galaxies**
  - **10.5.1 Data**
  - **10.5.2 The Statistical Model Formulation**
  - **10.5.3 Running the Model in R using JAGS**
  - **10.5.4 Running the Model in Python using Stan**
10.6 Bernoulli Model and the Fraction of Red Spirals 307
10.6.1 Data 308
10.6.2 The Statistical Model Formulation 308
10.6.3 Running the Model in R using JAGS 309
10.6.4 Running the Model in Python using Stan 311

10.7 Count Models, Globular Cluster Population, and Host Galaxy Brightness 313
10.7.1 Data 313
10.7.2 The Statistical Poisson Model Formulation 314
10.7.3 Running the Poisson Model in R using JAGS 315
10.7.4 The Statistical Negative Binomial Model Formulation 317
10.7.5 Running the Negative Binomial Model in R using JAGS 318
10.7.6 The Statistical NB-P Model Formulation 320
10.7.7 Running the NB-P Model in R using JAGS 321
10.7.8 Running the NB-P Model in Python using Stan 323

10.8 Bernoulli Mixed Model, AGNs, and Cluster Environment 325
10.8.1 Data 326
10.8.2 Statistical Model Formulation 327
10.8.3 Running the Model in R using JAGS 328
10.8.4 Running the Model in Python using Stan 329

10.9 Lognormal–Logit Hurdle Model and the Halo–Stellar-Mass Relation 332
10.9.1 Data 333
10.9.2 The Statistical Model Formulation 333
10.9.3 Running the Model in R using JAGS 334
10.9.4 Running the Model in Python using Stan 337

10.10 Count Time Series and Sunspot Data 340
10.10.1 Data 341
10.10.2 Running the Normal AR(1) Model in R using JAGS 341
10.10.3 Running the Negative Binomial AR Model in R using JAGS 344
10.10.4 Running the Negative Binomial AR Model in Python using Stan 346

10.11 Gaussian Model, ODEs, and Type Ia Supernova Cosmology 347
10.11.1 Data 348
10.11.2 The Statistical Model Formulation 348
10.11.3 Running the Model in R using Stan 349
10.11.4 Errors in Measurements 353

10.12 Approximate Bayesian Computation 355
10.12.1 Distance 356
10.12.2 Population Monte Carlo ABC 357
10.12.3 Toy Model 357
10.12.4 CosmoABC 359

10.13 Remarks on Applications 363

11 The Future of Astrostatistics 364
## Contents

<table>
<thead>
<tr>
<th>Appendix</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>Bayesian Modeling using INLA</td>
<td>366</td>
</tr>
<tr>
<td>B</td>
<td>Count Models with Offsets</td>
<td>369</td>
</tr>
<tr>
<td>C</td>
<td>Predicted Values, Residuals, and Diagnostics</td>
<td>377</td>
</tr>
</tbody>
</table>

**References**

**Index**

*Color plates are to be found between pages 198 and 199.*