

# Contents

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

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
<b>5</b>	<b>GLMs Part I – Continuous and Binomial Models</b>	<b>68</b>
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
<b>6</b>	<b>GLMs Part II – Count Models</b>	<b>135</b>
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 pymc3	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	180
6.5.2	Three-Parameter NB-P Models in Python using Stan	182
<b>7</b>	<b>GLMs Part III – Zero-Inflated and Hurdle Models</b>	<b>184</b>
7.1	Bayesian Zero-Inflated Models	184
7.1.1	Bayesian Zero-Inflated Poisson Model	184
7.1.2	Bayesian Zero-Inflated Negative Binomial Model	190
7.2	Bayesian Hurdle Models	196
7.2.1	Bayesian Poisson–Logit Hurdle Model	197
7.2.2	Bayesian Negative Binomial–Logit Hurdle Model	202
7.2.3	Bayesian Gamma–Logit Hurdle Model	206
7.2.4	Bayesian Lognormal–Logit Hurdle Model	210
<b>8</b>	<b>Hierarchical GLMMs</b>	<b>215</b>
8.1	Overview of Bayesian Hierarchical Models/GLMMs	215
8.2	Bayesian Gaussian or Normal GLMMs	219
8.2.1	Random Intercept Gaussian Data	219
8.2.2	Bayesian Random Intercept Gaussian Model in R using JAGS	220
8.2.3	Bayesian Random Intercept Normal Model in R using JAGS	222
8.2.4	Bayesian Random Intercept Normal Model in Python using Stan	226
8.3	Bayesian Binary Logistic GLMMs	228
8.3.1	Random Intercept Binary Logistic Data	228
8.3.2	Bayesian Random Intercept Binary Logistic Model with R	229
8.3.3	Bayesian Random Intercept Binary Logistic Model with Python	230
8.3.4	Bayesian Random Intercept Binary Logistic Model in R using JAGS	232
8.3.5	Bayesian Random Intercept Binary Logistic Model in Python using Stan	234
8.4	Bayesian Binomial Logistic GLMMs	235
8.4.1	Random Intercept Binomial Logistic Data	236
8.4.2	Bayesian Random Intercept Binomial Logistic Model in R using JAGS	237
8.4.3	Bayesian Random Intercept Binomial Logistic Model in Python using Stan	238
8.5	Bayesian Poisson GLMMs	240
8.5.1	Random Intercept Poisson Data	241
8.5.2	Bayesian Random Intercept Poisson Model with R	241
8.5.3	Bayesian Random Intercept Poisson Model in Python	242
8.5.4	Bayesian Random Intercept Poisson Model in R using JAGS	244
8.5.5	Bayesian Random Intercept Poisson Model in Python using Stan	246
8.5.6	Bayesian Random-Intercept–Random-Slopes Poisson Model	248

8.6	Bayesian Negative Binomial GLMMs	252
8.6.1	Random Intercept Negative Binomial Data	253
8.6.2	Random Intercept Negative Binomial MLE Model using R	254
8.6.3	Bayesian Random Intercept Negative Binomial Model using Python	255
8.6.4	Bayesian Random Intercept Negative Binomial Model in R using JAGS	257
8.6.5	Bayesian Random Intercept Negative Binomial Model in Python using Stan	259
<b>9</b>	<b>Model Selection</b>	262
9.1	Information Criteria Tests for Model Selection	262
9.1.1	Frequentist and Bayesian Information Criteria	262
9.1.2	Bayesian Deviance Statistic	264
9.1.3	pD and Deviance Information Criteria (DIC)	265
9.2	Model Selection with Indicator Functions	266
9.3	Bayesian LASSO	274
<b>10</b>	<b>Astronomical Applications</b>	276
10.1	Normal Model, Black Hole Mass, and Bulge Velocity Dispersion	277
10.1.1	Data	278
10.1.2	The Statistical Model Formulation	278
10.1.3	Running the Model in R using JAGS	279
10.1.4	Running the Model in Python using Stan	280
10.2	Gaussian Mixed Models, Type Ia Supernovae, and Hubble Residuals	283
10.2.1	Data	284
10.2.2	Statistical Model Formulation	284
10.2.3	Running the Model in R using JAGS	285
10.2.4	Running the Model in Python using Stan	288
10.3	Multivariate Normal Mixed Model and Early-Type Contact Binaries	290
10.3.1	Data	292
10.3.2	The Statistical Model Formulation	292
10.3.3	Running the Model in R using JAGS	293
10.3.4	Running the Model in Python using Stan	296
10.4	Lognormal Distribution and the Initial Mass Function	297
10.4.1	Data	298
10.4.2	Statistical Model Formulation	298
10.4.3	Running the Model in R using JAGS	299
10.4.4	Running the Model in Python using Stan	301
10.5	Beta Model and the Baryon Content of Low Mass Galaxies	302
10.5.1	Data	303
10.5.2	The Statistical Model Formulation	303
10.5.3	Running the Model in R using JAGS	304
10.5.4	Running the Model in Python using Stan	306

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
<b>11</b>	<b>The Future of Astrostatistics</b>	<b>364</b>

---

<b>Appendix A</b>	<b>Bayesian Modeling using INLA</b>	366
<b>Appendix B</b>	<b>Count Models with Offsets</b>	369
<b>Appendix C</b>	<b>Predicted Values, Residuals, and Diagnostics</b>	377
	<i>References</i>	380
	<i>Index</i>	391

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