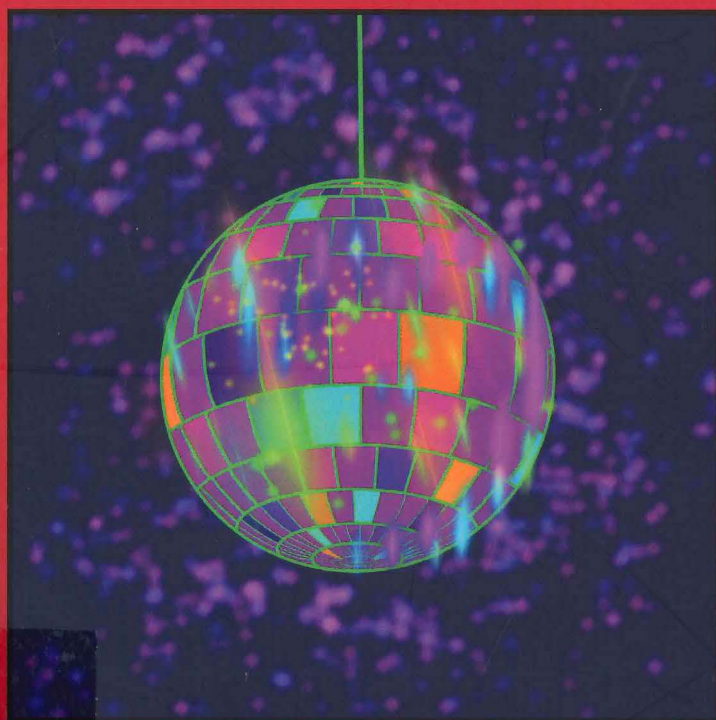


Texts in Statistical Science

Bayes Rules!

An Introduction to
Applied Bayesian Modeling



Alicia A. Johnson

Miles Q. Ott

Mine Dogucu



CRC Press
Taylor & Francis Group

A CHAPMAN & HALL BOOK

สำนักหอสมุด มหาวิทยาลัยเชียงใหม่

616706924

012548563

i 2268749x

Bayes Rules!

An Introduction to Applied Bayesian Modeling

Alicia A. Johnson
Miles Q. Ott
Mine Dogucu



CRC Press

Taylor & Francis Group

Boca Raton London New York

CRC Press is an imprint of the
Taylor & Francis Group, an **informa** business
A CHAPMAN & HALL BOOK

Contents

Foreword	xv
Preface	xvii
About the Authors	xxi
I Bayesian Foundations	1
1 The Big (Bayesian) Picture	3
1.1 Thinking like a Bayesian	4
1.1.1 Quiz yourself	5
1.1.2 The meaning of probability	6
1.1.3 The Bayesian balancing act	6
1.1.4 Asking questions	8
1.2 A quick history lesson	9
1.3 A look ahead	11
1.3.1 Unit 1: Bayesian foundations	11
1.3.2 Unit 2: Posterior simulation & analysis	12
1.3.3 Unit 3: Bayesian regression & classification	12
1.3.4 Unit 4: Hierarchical Bayesian models	13
1.4 Chapter summary	14
1.5 Exercises	14
2 Bayes' Rule	17
2.1 Building a Bayesian model for events	19
2.1.1 Prior probability model	19
2.1.2 Conditional probability & likelihood	20
2.1.3 Normalizing constants	22
2.1.4 Posterior probability model via Bayes' Rule!	24
2.1.5 Posterior simulation	25
2.2 Example: Pop vs soda vs coke	30
2.3 Building a Bayesian model for random variables	31
2.3.1 Prior probability model	32
2.3.2 The Binomial data model	33
2.3.3 The Binomial likelihood function	35
2.3.4 Normalizing constant	36
2.3.5 Posterior probability model	37
2.3.6 Posterior shortcut	38
2.3.7 Posterior simulation	40
2.4 Chapter summary	42
2.5 Exercises	42
2.5.1 Building up to Bayes' Rule	42
2.5.2 Practice Bayes' Rule for events	43

2.5.3	Practice Bayes' Rule for random variables	45
2.5.4	Simulation exercises	47
3	The Beta-Binomial Bayesian Model	49
3.1	The Beta prior model	50
3.1.1	Beta foundations	51
3.1.2	Tuning the Beta prior	54
3.2	The Binomial data model & likelihood function	55
3.3	The Beta posterior model	57
3.4	The Beta-Binomial model	61
3.5	Simulating the Beta-Binomial	63
3.6	Example: Milgram's behavioral study of obedience	64
3.6.1	A Bayesian analysis	65
3.6.2	The role of ethics in statistics and data science	66
3.7	Chapter summary	67
3.8	Exercises	68
3.8.1	Practice: Beta prior models	68
3.8.2	Practice: Beta-Binomial models	71
4	Balance and Sequentiality in Bayesian Analyses	75
4.1	Different priors, different posteriors	77
4.2	Different data, different posteriors	80
4.3	Striking a balance between the prior & data	82
4.3.1	Connecting observations to concepts	82
4.3.2	Connecting concepts to theory	83
4.4	Sequential analysis: Evolving with data	85
4.5	Proving data order invariance	88
4.6	Don't be stubborn	89
4.7	A note on subjectivity	90
4.8	Chapter summary	91
4.9	Exercises	92
4.9.1	Review exercises	92
4.9.2	Practice: Different priors, different posteriors	93
4.9.3	Practice: Balancing the data & prior	93
4.9.4	Practice: Sequentiality	95
5	Conjugate Families	97
5.1	Revisiting choice of prior	97
5.2	Gamma-Poisson conjugate family	100
5.2.1	The Poisson data model	100
5.2.2	Potential priors	103
5.2.3	Gamma prior	104
5.2.4	Gamma-Poisson conjugacy	106
5.3	Normal-Normal conjugate family	109
5.3.1	The Normal data model	109
5.3.2	Normal prior	111
5.3.3	Normal-Normal conjugacy	113
5.3.4	Optional: Proving Normal-Normal conjugacy	116
5.4	Why no simulation in this chapter?	117
5.5	Critiques of conjugate family models	118
5.6	Chapter summary	118

5.7	Exercises	118
5.7.1	Practice: Gamma-Poisson	118
5.7.2	Practice: Normal-Normal	120
5.7.3	General practice exercises	122
II Posterior Simulation & Analysis		125
6	Approximating the Posterior	127
6.1	Grid approximation	129
6.1.1	A Beta-Binomial example	129
6.1.2	A Gamma-Poisson example	134
6.1.3	Limitations	136
6.2	Markov chains via rstan	137
6.2.1	A Beta-Binomial example	139
6.2.2	A Gamma-Poisson example	143
6.3	Markov chain diagnostics	145
6.3.1	Examining trace plots	146
6.3.2	Comparing parallel chains	147
6.3.3	Calculating effective sample size & autocorrelation	148
6.3.4	Calculating R-hat	153
6.4	Chapter summary	155
6.5	Exercises	156
6.5.1	Conceptual exercises	156
6.5.2	Practice: Grid approximation	156
6.5.3	Practice: MCMC	157
7	MCMC under the Hood	159
7.1	The big idea	159
7.2	The Metropolis-Hastings algorithm	164
7.3	Implementing the Metropolis-Hastings	168
7.4	Tuning the Metropolis-Hastings algorithm	170
7.5	A Beta-Binomial example	172
7.6	Why the algorithm works	175
7.7	Variations on the theme	176
7.8	Chapter summary	176
7.9	Exercises	176
7.9.1	Conceptual exercises	177
7.9.2	Practice: Normal-Normal simulation	178
7.9.3	Practice: Simulating more Bayesian models	180
8	Posterior Inference & Prediction	183
8.1	Posterior estimation	184
8.2	Posterior hypothesis testing	187
8.2.1	One-sided tests	187
8.2.2	Two-sided tests	191
8.3	Posterior prediction	192
8.4	Posterior analysis with MCMC	195
8.4.1	Posterior simulation	195
8.4.2	Posterior estimation & hypothesis testing	196

8.4.3	Posterior prediction	199
8.5	Bayesian benefits	200
8.6	Chapter summary	201
8.7	Exercises	202
8.7.1	Conceptual exercises	202
8.7.2	Practice exercises	202
8.7.3	Applied exercises	204
III Bayesian Regression & Classification		209
9	Simple Normal Regression	211
9.1	Building the regression model	213
9.1.1	Specifying the data model	213
9.1.2	Specifying the priors	215
9.1.3	Putting it all together	216
9.2	Tuning prior models for regression parameters	216
9.3	Posterior simulation	219
9.3.1	Simulation via rstanarm	220
9.3.2	Optional: Simulation via rstan	222
9.4	Interpreting the posterior	223
9.5	Posterior prediction	226
9.5.1	Building a posterior predictive model	227
9.5.2	Posterior prediction with rstanarm	229
9.6	Sequential regression modeling	231
9.7	Using default rstanarm priors	232
9.8	You're not done yet!	235
9.9	Chapter summary	236
9.10	Exercises	236
9.10.1	Conceptual exercises	236
9.10.2	Applied exercises	237
10	Evaluating Regression Models	243
10.1	Is the model fair?	243
10.2	How wrong is the model?	245
10.2.1	Checking the model assumptions	245
10.2.2	Dealing with wrong models	248
10.3	How accurate are the posterior predictive models?	250
10.3.1	Posterior predictive summaries	251
10.3.2	Cross-validation	255
10.3.3	Expected log-predictive density	259
10.3.4	Improving posterior predictive accuracy	260
10.4	How good is the MCMC simulation vs how good is the model?	260
10.5	Chapter summary	261
10.6	Exercises	261
10.6.1	Conceptual exercises	261
10.6.2	Applied exercises	263
10.6.3	Open-ended exercises	265
11	Extending the Normal Regression Model	267
11.1	Utilizing a categorical predictor	270
11.1.1	Building the model	271

11.1.2	Simulating the posterior	272
11.2	Utilizing two predictors	274
11.2.1	Building the model	275
11.2.2	Understanding the priors	276
11.2.3	Simulating the posterior	277
11.2.4	Posterior prediction	279
11.3	Optional: Utilizing interaction terms	280
11.3.1	Building the model	280
11.3.2	Simulating the posterior	281
11.3.3	Do you need an interaction term?	283
11.4	Dreaming bigger: Utilizing more than 2 predictors!	286
11.5	Model evaluation & comparison	289
11.5.1	Evaluating predictive accuracy using visualizations	290
11.5.2	Evaluating predictive accuracy using cross-validation	292
11.5.3	Evaluating predictive accuracy using ELPD	293
11.5.4	The bias-variance trade-off	294
11.6	Chapter summary	298
11.7	Exercises	299
11.7.1	Conceptual exercises	299
11.7.2	Applied exercises	300
11.7.3	Open-ended exercises	302
12	Poisson & Negative Binomial Regression	303
12.1	Building the Poisson regression model	306
12.1.1	Specifying the data model	306
12.1.2	Specifying the priors	310
12.2	Simulating the posterior	312
12.3	Interpreting the posterior	313
12.4	Posterior prediction	315
12.5	Model evaluation	317
12.6	Negative Binomial regression for overdispersed counts	319
12.7	Generalized linear models: Building on the theme	324
12.8	Chapter summary	325
12.9	Exercises	326
12.9.1	Conceptual exercises	326
12.9.2	Applied exercises	327
13	Logistic Regression	329
13.1	Pause: Odds & probability	330
13.2	Building the logistic regression model	331
13.2.1	Specifying the data model	331
13.2.2	Specifying the priors	334
13.3	Simulating the posterior	336
13.4	Prediction & classification	339
13.5	Model evaluation	341
13.6	Extending the model	346
13.7	Chapter summary	348
13.8	Exercises	349
13.8.1	Conceptual exercises	349
13.8.2	Applied exercises	350
13.8.3	Open-ended exercises	352

14 Naive Bayes Classification	355
14.1 Classifying one penguin	356
14.1.1 One categorical predictor	357
14.1.2 One quantitative predictor	359
14.1.3 Two predictors	362
14.2 Implementing & evaluating naive Bayes classification	365
14.3 Naive Bayes vs logistic regression	369
14.4 Chapter summary	369
14.5 Exercises	370
14.5.1 Conceptual exercises	370
14.5.2 Applied exercises	370
14.5.3 Open-ended exercises	372
IV Hierarchical Bayesian models	373
15 Hierarchical Models are Exciting	375
15.1 Complete pooling	377
15.2 No pooling	380
15.3 Hierarchical data	382
15.4 Partial pooling with hierarchical models	383
15.5 Chapter summary	384
15.6 Exercises	385
15.6.1 Conceptual exercises	385
15.6.2 Applied exercises	385
16 (Normal) Hierarchical Models without Predictors	387
16.1 Complete pooled model	390
16.2 No pooled model	393
16.3 Building the hierarchical model	397
16.3.1 The hierarchy	397
16.3.2 Another way to think about it	399
16.3.3 Within- vs between-group variability	400
16.4 Posterior analysis	401
16.4.1 Posterior simulation	401
16.4.2 Posterior analysis of global parameters	403
16.4.3 Posterior analysis of group-specific parameters	404
16.5 Posterior prediction	407
16.6 Shrinkage & the bias-variance trade-off	411
16.7 Not everything is hierarchical	414
16.8 Chapter summary	416
16.9 Exercises	417
16.9.1 Conceptual exercises	417
16.9.2 Applied exercises	418
17 (Normal) Hierarchical Models with Predictors	421
17.1 First steps: Complete pooling	422
17.2 Hierarchical model with varying intercepts	423
17.2.1 Model building	423
17.2.2 Another way to think about it	427
17.2.3 Tuning the prior	427
17.2.4 Posterior simulation & analysis	429

17.3	Hierarchical model with varying intercepts & slopes	435
17.3.1	Model building	436
17.3.2	Optional: The decomposition of covariance model	440
17.3.3	Posterior simulation & analysis	442
17.4	Model evaluation & selection	447
17.5	Posterior prediction	450
17.6	Details: Longitudinal data	452
17.7	Example: Danceability	452
17.8	Chapter summary	457
17.9	Exercises	458
17.9.1	Conceptual exercises	458
17.9.2	Applied exercises	459
17.9.3	Open-ended exercises	461
18	Non-Normal Hierarchical Regression & Classification	463
18.1	Hierarchical logistic regression	463
18.1.1	Model building & simulation	466
18.1.2	Posterior analysis	469
18.1.3	Posterior classification	470
18.1.4	Model evaluation	472
18.2	Hierarchical Poisson & Negative Binomial regression	473
18.2.1	Model building & simulation	474
18.2.2	Posterior analysis	477
18.2.3	Model evaluation	479
18.3	Chapter summary	480
18.4	Exercises	480
18.4.1	Applied & conceptual exercises	480
18.4.2	Open-ended exercises	483
19	Adding More Layers	485
19.1	Group-level predictors	485
19.1.1	A model using only individual-level predictors	486
19.1.2	Incorporating group-level predictors	489
19.1.3	Posterior simulation & global analysis	492
19.1.4	Posterior group-level analysis	494
19.1.5	We're just scratching the surface!	497
19.2	Incorporating two (or more!) grouping variables	497
19.2.1	Data with two grouping variables	497
19.2.2	Building a model with two grouping variables	499
19.2.3	Simulating models with two grouping variables	501
19.2.4	Examining the group-specific parameters	503
19.2.5	We're just scratching the surface!	505
19.3	Exercises	505
19.3.1	Conceptual exercises	505
19.3.2	Applied exercises	506
19.4	Goodbye!	509
	Bibliography	511
	Index	517