

Contents

List of Figures	xiii
List of Tables	xix
Preface	xxi
Acknowledgments	xxv
Introduction	xxvii
Part I Introducing Bayesian Analysis	1
1 The foundations of Bayesian inference	3
1.1 What is probability?	3
1.1.1 Probability in classical statistics	4
1.1.2 Subjective probability	5
1.2 Subjective probability in Bayesian statistics	7
1.3 Bayes theorem, discrete case	8
1.4 Bayes theorem, continuous parameter	13
1.4.1 Conjugate priors	15
1.4.2 Bayesian updating with irregular priors	16
1.4.3 Cromwell's Rule	18
1.4.4 Bayesian updating as information accumulation	19
1.5 Parameters as random variables, beliefs as distributions	21
1.6 Communicating the results of a Bayesian analysis	22
1.6.1 Bayesian point estimation	23
1.6.2 Credible regions	26
1.7 Asymptotic properties of posterior distributions	29
1.8 Bayesian hypothesis testing	31
1.8.1 Model choice	36
1.8.2 Bayes factors	37

1.9	From subjective beliefs to parameters and models	38
1.9.1	Exchangeability	39
1.9.2	Implications and extensions of de Finetti's Representation Theorem	42
1.9.3	Finite exchangeability	43
1.9.4	Exchangeability and prediction	43
1.9.5	Conditional exchangeability and multiparameter models	44
1.9.6	Exchangeability of parameters: hierarchical modeling	45
1.10	Historical note	46
2	Getting started: Bayesian analysis for simple models	49
2.1	Learning about probabilities, rates and proportions	49
2.1.1	Conjugate priors for probabilities, rates and proportions	51
2.1.2	Bayes estimates as weighted averages of priors and data	58
2.1.3	Parameterizations and priors	61
2.1.4	The variance of the posterior density	64
2.2	Associations between binary variables	67
2.3	Learning from counts	73
2.3.1	Predictive inference with count data	78
2.4	Learning about a normal mean and variance	80
2.4.1	Variance known	80
2.4.2	Mean and variance unknown	83
2.4.3	Conditionally conjugate prior	92
2.4.4	An improper, reference prior	93
2.4.5	Conflict between likelihood and prior	98
2.4.6	Non-conjugate priors	98
2.5	Regression models	99
2.5.1	Bayesian regression analysis	102
2.5.2	Likelihood function	103
2.5.3	Conjugate prior	104
2.5.4	Improper, reference prior	107
2.6	Further reading	124
Part II	Simulation Based Bayesian Analysis	129
3	Monte Carlo methods	133
3.1	Simulation consistency	134
3.2	Inference for functions of parameters	140
3.3	Marginalization via Monte Carlo integration	142
3.4	Sampling algorithms	153
3.4.1	Inverse-CDF method	153
3.4.2	Importance sampling	156

3.4.3	Accept-reject sampling	159
3.4.4	Adaptive rejection sampling	163
3.5	Further reading	167
4	Markov chains	171
4.1	Notation and definitions	172
4.1.1	State space	173
4.1.2	Transition kernel	173
4.2	Properties of Markov chains	176
4.2.1	Existence of a stationary distribution, discrete case	177
4.2.2	Existence of a stationary distribution, continuous case	178
4.2.3	Irreducibility	179
4.2.4	Recurrence	182
4.2.5	Invariant measure	184
4.2.6	Reversibility	185
4.2.7	Aperiodicity	186
4.3	Convergence of Markov chains	187
4.3.1	Speed of convergence	189
4.4	Limit theorems for Markov chains	191
4.4.1	Simulation inefficiency	191
4.4.2	Central limit theorems for Markov chains	195
4.5	Further reading	196
5	Markov chain Monte Carlo	201
5.1	Metropolis-Hastings algorithm	201
5.1.1	Theory for the Metropolis-Hastings algorithm	202
5.1.2	Choosing the proposal density	204
5.2	Gibbs sampling	214
5.2.1	Theory for the Gibbs sampler	218
5.2.2	Connection to the Metropolis algorithm	221
5.2.3	Deriving conditional densities for the Gibbs sampler: statistical models as conditional independence graphs	225
5.2.4	Pathologies	229
5.2.5	Data augmentation	236
5.2.6	Missing data problems	237
5.2.7	The slice sampler	244
6	Implementing Markov chain Monte Carlo	251
6.1	Software for Markov chain Monte Carlo	251
6.2	Assessing convergence and run-length	252
6.3	Working with BUGS/JAGS from R	256
6.4	Tricks of the trade	261
6.4.1	Thinning	261

6.4.2	Blocking	264
6.4.3	Reparameterization	270
6.5	Other examples	272
6.6	Further reading	292
Part III Advanced Applications in the Social Sciences		299
7	Hierarchical Statistical Models	301
7.1	Data and parameters that vary by groups: the case for hierarchical modeling	301
7.1.1	Exchangeable parameters generate hierarchical models	305
7.1.2	'Borrowing strength' via exchangeability	307
7.1.3	Hierarchical modeling as a 'semi-pooling' estimator	307
7.1.4	Hierarchical modeling as a 'shrinkage' estimator	308
7.1.5	Computation via Markov chain Monte Carlo	310
7.2	ANOVA as a hierarchical model	317
7.2.1	One-way analysis of variance	317
7.2.2	Two-way ANOVA	329
7.3	Hierarchical models for longitudinal data	345
7.4	Hierarchical models for non-normal data	354
7.5	Multi-level models	362
8	Bayesian analysis of choice making	379
8.1	Regression models for binary responses	379
8.1.1	Probit model via data augmentation	380
8.1.2	Probit model via marginal data augmentation	389
8.1.3	Logit model	393
8.1.4	Binomial model for grouped binary data	395
8.2	Ordered outcomes	397
8.2.1	Identification	399
8.3	Multinomial outcomes	415
8.3.1	Multinomial logit (MNL)	415
8.3.2	Independence of irrelevant alternatives	423
8.4	Multinomial probit	424
8.4.1	Bayesian analysis via MCMC	426
9	Bayesian approaches to measurement	435
9.1	Bayesian inference for latent states	435
9.1.1	A formal role for prior information	436
9.1.2	Inference for many parameters	436
9.2	Factor analysis	438
9.2.1	Likelihood and prior densities	439
9.2.2	Identification	440
9.2.3	Posterior density	442
9.2.4	Inference over rank orderings of the latent variable	448
9.2.5	Incorporating additional information via hierarchical modeling	449

9.3	Item-response models	454
9.4	Dynamic measurement models	471
9.4.1	State-space models for ‘pooling the polls’	473
9.4.2	Bayesian inference	474
Part IV Appendices		489
Appendix A: Working with vectors and matrices		491
Appendix B: Probability review		497
B.1	Foundations of probability	497
B.2	Probability densities and mass functions	498
B.2.1	Probability mass functions for discrete random quantities	501
B.2.2	Probability density functions for continuous random quantities	503
B.3	Convergence of sequences of random variables	511
Appendix C: Proofs of selected propositions		513
C.1	Products of normal densities	513
C.2	Conjugate analysis of normal data	516
C.3	Asymptotic normality of the posterior density	533
References		535
Topic index		553
Author index		559

