
Contents

Preface	xiii
Part I: Fundamentals of Bayesian Inference	1
1 Probability and inference	3
1.1 The three steps of Bayesian data analysis	3
1.2 General notation for statistical inference	4
1.3 Bayesian inference	6
1.4 Discrete examples: genetics and spell checking	8
1.5 Probability as a measure of uncertainty	11
1.6 Example: probabilities from football point spreads	13
1.7 Example: calibration for record linkage	16
1.8 Some useful results from probability theory	19
1.9 Computation and software	22
1.10 Bayesian inference in applied statistics	24
1.11 Bibliographic note	25
1.12 Exercises	27
2 Single-parameter models	29
2.1 Estimating a probability from binomial data	29
2.2 Posterior as compromise between data and prior information	32
2.3 Summarizing posterior inference	32
2.4 Informative prior distributions	34
2.5 Normal distribution with known variance	39
2.6 Other standard single-parameter models	42
2.7 Example: informative prior distribution for cancer rates	47
2.8 Noninformative prior distributions	51
2.9 Weakly informative prior distributions	55
2.10 Bibliographic note	56
2.11 Exercises	57
3 Introduction to multiparameter models	63
3.1 Averaging over ‘nuisance parameters’	63
3.2 Normal data with a noninformative prior distribution	64
3.3 Normal data with a conjugate prior distribution	67
3.4 Multinomial model for categorical data	69
3.5 Multivariate normal model with known variance	70
3.6 Multivariate normal with unknown mean and variance	72
3.7 Example: analysis of a bioassay experiment	74
3.8 Summary of elementary modeling and computation	78
3.9 Bibliographic note	78
3.10 Exercises	79

4 Asymptotics and connections to non-Bayesian approaches	83	9 Decision analysis	237
4.1 Normal approximations to the posterior distribution	83	9.1 Bayesian decision theory in different contexts	237
4.2 Large-sample theory	87	9.2 Using regression predictions: survey incentives	239
4.3 Counterexamples to the theorems	89	9.3 Multistage decision making: medical screening	245
4.4 Frequency evaluations of Bayesian inferences	91	9.4 Hierarchical decision analysis for home radon	246
4.5 Bayesian interpretations of other statistical methods	92	9.5 Personal vs. institutional decision analysis	256
4.6 Bibliographic note	97	9.6 Bibliographic note	257
4.7 Exercises	98	9.7 Exercises	257
5 Hierarchical models	101	Part III: Advanced Computation	259
5.1 Constructing a parameterized prior distribution	102	10 Introduction to Bayesian computation	261
5.2 Exchangeability and hierarchical models	104	10.1 Numerical integration	261
5.3 Bayesian analysis of conjugate hierarchical models	108	10.2 Distributional approximations	262
5.4 Normal model with exchangeable parameters	113	10.3 Direct simulation and rejection sampling	263
5.5 Example: parallel experiments in eight schools	119	10.4 Importance sampling	265
5.6 Hierarchical modeling applied to a meta-analysis	124	10.5 How many simulation draws are needed?	267
5.7 Weakly informative priors for variance parameters	128	10.6 Computing environments	268
5.8 Bibliographic note	132	10.7 Debugging Bayesian computing	270
5.9 Exercises	134	10.8 Bibliographic note	271
Part II: Fundamentals of Bayesian Data Analysis	139	10.9 Exercises	272
6 Model checking	141	11 Basics of Markov chain simulation	275
6.1 The place of model checking in applied Bayesian statistics	141	11.1 Gibbs sampler	276
6.2 Do the inferences from the model make sense?	142	11.2 Metropolis and Metropolis-Hastings algorithms	278
6.3 Posterior predictive checking	143	11.3 Using Gibbs and Metropolis as building blocks	280
6.4 Graphical posterior predictive checks	153	11.4 Inference and assessing convergence	281
6.5 Model checking for the educational testing example	159	11.5 Effective number of simulation draws	286
6.6 Bibliographic note	161	11.6 Example: hierarchical normal model	288
6.7 Exercises	163	11.7 Bibliographic note	291
7 Evaluating, comparing, and expanding models	165	11.8 Exercises	291
7.1 Measures of predictive accuracy	166	12 Computationally efficient Markov chain simulation	293
7.2 Information criteria and cross-validation	169	12.1 Efficient Gibbs samplers	293
7.3 Model comparison based on predictive performance	178	12.2 Efficient Metropolis jumping rules	295
7.4 Model comparison using Bayes factors	182	12.3 Further extensions to Gibbs and Metropolis	297
7.5 Continuous model expansion	184	12.4 Hamiltonian Monte Carlo	300
7.6 Implicit assumptions and model expansion: an example	187	12.5 Hamiltonian Monte Carlo for a hierarchical model	305
7.7 Bibliographic note	192	12.6 Stan: developing a computing environment	307
7.8 Exercises	193	12.7 Bibliographic note	308
8 Modeling accounting for data collection	197	12.8 Exercises	309
8.1 Bayesian inference requires a model for data collection	197	13 Modal and distributional approximations	311
8.2 Data-collection models and ignorability	199	13.1 Finding posterior modes	311
8.3 Sample surveys	205	13.2 Boundary-avoiding priors for modal summaries	313
8.4 Designed experiments	214	13.3 Normal and related mixture approximations	318
8.5 Sensitivity and the role of randomization	218	13.4 Finding marginal posterior modes using EM	320
8.6 Observational studies	220	13.5 Conditional and marginal posterior approximations	325
8.7 Censoring and truncation	224	13.6 Example: hierarchical normal model (continued)	326
8.8 Discussion	229	13.7 Variational inference	331
8.9 Bibliographic note	229	13.8 Expectation propagation	338
8.10 Exercises	230	13.9 Other approximations	343

13.10 Unknown normalizing factors	345	18 Models for missing data	449
13.11 Bibliographic note	348	18.1 Notation	449
13.12 Exercises	349	18.2 Multiple imputation	451
Part IV: Regression Models	351	18.3 Missing data in the multivariate normal and t models	454
14 Introduction to regression models	353	18.4 Example: multiple imputation for a series of polls	456
14.1 Conditional modeling	353	18.5 Missing values with counted data	462
14.2 Bayesian analysis of classical regression	354	18.6 Example: an opinion poll in Slovenia	463
14.3 Regression for causal inference: incumbency and voting	358	18.7 Bibliographic note	466
14.4 Goals of regression analysis	364	18.8 Exercises	467
14.5 Assembling the matrix of explanatory variables	365	Part V: Nonlinear and Nonparametric Models	469
14.6 Regularization and dimension reduction	367	19 Parametric nonlinear models	471
14.7 Unequal variances and correlations	369	19.1 Example: serial dilution assay	471
14.8 Including numerical prior information	376	19.2 Example: population toxicokinetics	477
14.9 Bibliographic note	378	19.3 Bibliographic note	485
14.10 Exercises	378	19.4 Exercises	486
15 Hierarchical linear models	381	20 Basis function models	487
15.1 Regression coefficients exchangeable in batches	382	20.1 Splines and weighted sums of basis functions	487
15.2 Example: forecasting U.S. presidential elections	383	20.2 Basis selection and shrinkage of coefficients	490
15.3 Interpreting a normal prior distribution as extra data	388	20.3 Non-normal models and regression surfaces	494
15.4 Varying intercepts and slopes	390	20.4 Bibliographic note	498
15.5 Computation: batching and transformation	392	20.5 Exercises	498
15.6 Analysis of variance and the batching of coefficients	395	21 Gaussian process models	501
15.7 Hierarchical models for batches of variance components	398	21.1 Gaussian process regression	501
15.8 Bibliographic note	400	21.2 Example: birthdays and birthdates	505
15.9 Exercises	402	21.3 Latent Gaussian process models	510
16 Generalized linear models	405	21.4 Functional data analysis	512
16.1 Standard generalized linear model likelihoods	406	21.5 Density estimation and regression	513
16.2 Working with generalized linear models	407	21.6 Bibliographic note	515
16.3 Weakly informative priors for logistic regression	412	21.7 Exercises	516
16.4 Overdispersed Poisson regression for police stops	420	22 Finite mixture models	519
16.5 State-level opinions from national polls	422	22.1 Setting up and interpreting mixture models	519
16.6 Models for multivariate and multinomial responses	423	22.2 Example: reaction times and schizophrenia	524
16.7 Loglinear models for multivariate discrete data	428	22.3 Label switching and posterior computation	533
16.8 Bibliographic note	431	22.4 Unspecified number of mixture components	536
16.9 Exercises	432	22.5 Mixture models for classification and regression	539
17 Models for robust inference	435	22.6 Bibliographic note	542
17.1 Aspects of robustness	435	22.7 Exercises	543
17.2 Overdispersed versions of standard models	437	23 Dirichlet process models	545
17.3 Posterior inference and computation	439	23.1 Bayesian histograms	545
17.4 Robust inference for the eight schools	441	23.2 Dirichlet process prior distributions	546
17.5 Robust regression using t -distributed errors	444	23.3 Dirichlet process mixtures	549
17.6 Bibliographic note	445	23.4 Beyond density estimation	557
17.7 Exercises	446	23.5 Hierarchical dependence	560
		23.6 Density regression	568
		23.7 Bibliographic note	571
		23.8 Exercises	573

Appendixes	575
A Standard probability distributions	577
A.1 Continuous distributions	577
A.2 Discrete distributions	585
A.3 Bibliographic note	586
B Outline of proofs of limit theorems	587
B.1 Bibliographic note	590
C Computation in R and Stan	591
C.1 Getting started with R and Stan	591
C.2 Fitting a hierarchical model in Stan	592
C.3 Direct simulation, Gibbs, and Metropolis in R	596
C.4 Programming Hamiltonian Monte Carlo in R	603
C.5 Further comments on computation	607
C.6 Bibliographic note	608
References	609
Author Index	643
Subject Index	654