Contents

List of models xi

List of examples xiii

Preface xv

Part I: Fundamentals of Bayesian Inference 1

1 Background 3
1.1 Overview 3
1.2 General notation for statistical inference 4
1.3 Bayesian inference 7
1.4 Example: inference about a genetic probability 10
1.5 Probability as a measure of uncertainty 12
1.6 Example of probability assignment: football point spreads 15
1.7 Some useful results from probability theory 18
1.8 Summarizing inferences by simulation 21
1.9 Bibliographic note 24
1.10 Exercises 25

2 Single-parameter models 28
2.1 Estimating a probability from binomial data 28
2.2 Posterior distribution as compromise between data and prior information 32
2.3 Summarizing posterior inference 33
2.4 Informative prior distributions 34
2.5 Example: estimating the probability of a female birth given placenta previa 39
2.6 Estimating the mean of a normal distribution with known variance 42
2.7 Other standard single-parameter models 45
2.8 Noninformative prior distributions 52

2.9 Bibliographic note 57
2.10 Exercises 58

3 Introduction to multiparameter models 65
3.1 Averaging over 'nuisance parameters' 65
3.2 Normal data with a noninformative prior distribution 66
3.3 Normal data with a conjugate prior distribution 71
3.4 Normal data with a semi-conjugate prior distribution 73
3.5 The multinomial model 76
3.6 The multivariate normal model 78
3.7 Example: analysis of a bioassay experiment 82
3.8 Summary of elementary multiparameter modeling and computation 86
3.9 Bibliographic note 87
3.10 Exercises 88

4 Large-sample inference and connections to standard statistical methods 94
4.1 Normal approximations to the posterior distribution 94
4.2 Large-sample theory 100
4.3 Counterexamples to the theorems 101
4.4 Frequency evaluations of Bayesian inferences 104
4.5 Bayesian interpretations of other statistical methods 106
4.6 Bibliographic note 111
4.7 Exercises 112

Part II: Fundamentals of Bayesian Data Analysis 117

5 Hierarchical models 119
5.1 Constructing a parameterized prior distribution 120
5.2 Exchangeability and setting up hierarchical models 123
5.3 Computation with hierarchical models 128
5.4 Estimating an exchangeable set of parameters from a normal model 134
5.5 Example: combining information from educational testing experiments in eight schools 141
5.6 Hierarchical modeling applied to a meta-analysis 148
5.7 Bibliographic note 154
5.8 Exercises 156

6 Model checking and sensitivity analysis 161
6.1 The place of model checking and sensitivity analysis in applied Bayesian statistics 161
6.2 Principles and methods of model checking 162

6.3 Checking a model by comparing data to the posterior predictive distribution 167
6.4 Sensitivity analysis 174
6.5 Comparing a discrete set of models using Bayes factors 175
6.6 Model expansion 177
6.7 Practical advice 179
6.8 Model checking for the educational testing example 179
6.9 Bibliographic note 183
6.10 Exercises 185

7 Study design in Bayesian analysis 190
7.1 Introduction 190
7.2 Relevance of design or data collection: simple examples 192
7.3 Formal models for data collection 194
7.4 Ignorability 199
7.5 Designs that are ignorable and known with no covariates, including simple random sampling and completely randomized experiments 200
7.6 Designs that are ignorable and known given covariates, including stratified sampling and randomized block experiments 205
7.7 Designs that are ignorable and unknown, such as experiments with nonrandom treatment assignments based on fully observed covariates 213
7.8 Designs that are nonignorable and known, such as censoring 215
7.9 Designs that are nonignorable and unknown, including observational studies and unintentional missing data 219
7.10 Sensitivity and the role of randomization 221
7.11 Discussion 224
7.12 Bibliographic note 224
7.13 Exercises 225

8 Introduction to regression models 233
8.1 Introduction and notation 233
8.2 Bayesian justification of conditional modeling 235
8.3 Bayesian analysis of the classical regression model 235
8.4 Example: estimating the advantage of incumbency in U.S. Congressional elections 240
8.5 Goals of regression analysis 248
8.6 Assembling the matrix of explanatory variables 250
8.7 Unequal variances and correlations 253
8.8 Models for unequal variances (heteroscedasticity) 257
8.9 Including prior information 259
8.10 Hierarchical linear models 262

8.11 Bibliographic note 262
8.12 Exercises 263

Part III: Advanced Computation **267**

9 Approximations based on posterior modes **269**
9.1 Introduction 269
9.2 Crude estimation by ignoring some information 270
9.3 Finding posterior modes 271
9.4 The normal and related mixture approximations 274
9.5 Finding marginal posterior modes using EM and related algorithms 276
9.6 Approximating the conditional posterior density, $p(\gamma|\phi, y)$ 283
9.7 Approximating $p(\phi|y)$ using an analytic approximation to $p(\gamma|\phi, y)$ 283
9.8 Example: the hierarchical normal model 284
9.9 Example: a hierarchical logistic regression model for rat tumor rates 291
9.10 Bibliographic note 298
9.11 Exercises 298

10 Posterior simulation and integration **300**
10.1 Posterior inference from simulation 300
10.2 Direct simulation 302
10.3 Numerical integration 305
10.4 Computing normalizing factors 308
10.5 Improving an approximation using importance resampling 312
10.6 Example: hierarchical logistic regression (continued) 313
10.7 Bibliographic note 316
10.8 Exercises 317

11 Markov chain simulation **320**
11.1 Introduction 320
11.2 The Metropolis algorithm and its generalizations 322
11.3 The Gibbs sampler and related methods based on alternating conditional sampling 326
11.4 Inference and assessing convergence from iterative simulation 329
11.5 Constructing efficient simulation algorithms 333
11.6 Example: the hierarchical normal model (continued) 335
11.7 Example: linear regression with several unknown variance parameters 337
11.8 Bibliographic note 343
11.9 Exercises 344

Part IV: Specific Models **345**

12 Models for robust inference and sensitivity analysis **347**
12.1 Introduction 347
12.2 Overdispersed versions of standard probability models 349
12.3 Posterior inference and computation 352
12.4 Robust inference and sensitivity analysis for the educational testing example 354
12.5 Robust regression using Student-t errors 360
12.6 Bibliographic note 362
12.7 Exercises 363

13 Hierarchical linear models **366**
13.1 Regression coefficients exchangeable in batches 367
13.2 Example: forecasting U.S. Presidential elections 369
13.3 General notation and computation for hierarchical linear models 376
13.4 Hierarchical modeling as an alternative to selecting explanatory variables 379
13.5 Bibliographic note 380
13.6 Exercises 381

14 Generalized linear models **384**
14.1 Introduction 384
14.2 Standard generalized linear model likelihoods 386
14.3 Setting up and interpreting generalized linear models 387
14.4 Bayesian nonhierarchical and hierarchical generalized linear models 388
14.5 Computation 389
14.6 Models for multinomial responses 393
14.7 Loglinear models for multivariate discrete data 397
14.8 Bibliographic note 403
14.9 Exercises 404

15 Multivariate models **407**
15.1 Introduction 407
15.2 Linear regression with multiple outcomes 407
15.3 Hierarchical multivariate models 410
15.4 Multivariate models for nonnormal data 412
15.5 Time series and spatial models 415
15.6 Bibliographic note 418
15.7 Exercises 419

16 Mixture models **420**

16.1 Introduction 420
16.2 Setting up the model 421
16.3 Computation 424
16.4 Example: modeling reaction times of schizophrenics and nonschizophrenics 426
16.5 Bibliographic note 438

17 Models for missing data 439
17.1 Introduction 439
17.2 Notation for data collection in the context of missing data problems 439
17.3 Computation and multiple imputation 441
17.4 Missing data in the multivariate normal and t models 443
17.5 Missing values with counted data 447
17.6 Example: an opinion poll in Slovenia 448
17.7 Inference using multiple imputation 453
17.8 Bibliographic note 454
17.9 Exercises 455

18 Concluding advice 456
18.1 Setting up probability models 456
18.2 Posterior inference 461
18.3 Model evaluation 462
18.4 Conclusion 468
18.5 Bibliographic note 469

Appendixes 471

A Standard probability distributions 473
A.1 Introduction 473
A.2 Continuous distributions 473
A.3 Discrete distributions 482
A.4 Bibliographic note 483

B Outline of proofs of asymptotic theorems 484
B.1 Bibliographic note 488

References 489

Author Index 513

Subject Index 518