

Table of Contents

Preface	1
Chapter 1: Thinking Probabilistically	7
Statistics, models, and this book's approach	8
Working with data	9
Bayesian modeling	10
Probability theory	11
Interpreting probabilities	11
Defining probabilities	12
Probability distributions	13
Independently and identically distributed variables	17
Bayes' theorem	19
Single-parameter inference	22
The coin-flipping problem	22
The general model	23
Choosing the likelihood	23
Choosing the prior	25
Getting the posterior	27
Computing and plotting the posterior	28
The influence of the prior and how to choose one	31
Communicating a Bayesian analysis	33
Model notation and visualization	33
Summarizing the posterior	34
Highest-posterior density	34
Posterior predictive checks	35
Summary	37
Exercises	39
Chapter 2: Programming Probabilistically	41
Probabilistic programming	42
PyMC3 primer	43
Flipping coins the PyMC3 way	44
Model specification	44
Pushing the inference button	45
Summarizing the posterior	46
Posterior-based decisions	48
ROPE	48
Loss functions	50
Gaussians all the way down	53
Gaussian inferences	53
Robust inferences	59

Student's t-distribution	60
Groups comparison	65
Cohen's d	67
Probability of superiority	68
The tips dataset	69
Hierarchical models	73
Shrinkage	77
One more example	81
Summary	85
Exercises	86
Chapter 3: Modeling with Linear Regression	89
Simple linear regression	90
The machine learning connection	90
The core of the linear regression models	91
Linear models and high autocorrelation	96
Modifying the data before running	97
Interpreting and visualizing the posterior	99
Pearson correlation coefficient	102
Pearson coefficient from a multivariate Gaussian	103
Robust linear regression	106
Hierarchical linear regression	111
Correlation, causation, and the messiness of life	117
Polynomial regression	118
Interpreting the parameters of a polynomial regression	121
Polynomial regression – the ultimate model?	121
Multiple linear regression	122
Confounding variables and redundant variables	127
Multicollinearity or when the correlation is too high	130
Masking effect variables	135
Adding interactions	137
Variable variance	138
Summary	142
Exercises	142
Chapter 4: Generalizing Linear Models	145
Generalized linear models	146
Logistic regression	147
The logistic model	148
The Iris dataset	149
The logistic model applied to the iris dataset	152
Multiple logistic regression	155
The boundary decision	155
Implementing the model	156
Interpreting the coefficients of a logistic regression	157
Dealing with correlated variables	160

Dealing with unbalanced classes	162
Softmax regression	164
Discriminative and generative models	166
Poisson regression	169
Poisson distribution	169
The zero-inflated Poisson model	171
Poisson regression and ZIP regression	173
Robust logistic regression	176
The GLM module	177
Summary	178
Exercises	179
Chapter 5: Model Comparison	181
Posterior predictive checks	182
Occam's razor – simplicity and accuracy	187
Too many parameters leads to overfitting	189
Too few parameters leads to underfitting	191
The balance between simplicity and accuracy	191
Predictive accuracy measures	192
Cross-validation	193
Information criteria	194
Log-likelihood and deviance	194
Akaike information criterion	195
Widely applicable Information Criterion	196
Pareto smoothed importance sampling leave-one-out cross-validation	196
Other Information Criteria	197
Model comparison with PyMC3	197
A note on the reliability of WAIC and LOO computations	200
Model averaging	200
Bayes factors	204
Some remarks	205
Computing Bayes factors	206
Common problems when computing Bayes factors	209
Using Sequential Monte Carlo to compute Bayes factors	210
Bayes factors and Information Criteria	211
Regularizing priors	214
WAIC in depth	215
Entropy	216
Kullback-Leibler divergence	218
Summary	221
Exercises	222
Chapter 6: Mixture Models	223
Mixture models	224
Finite mixture models	225
The categorical distribution	227

The Dirichlet distribution	228
Non-identifiability of mixture models	232
How to choose K	234
Mixture models and clustering	239
Non-finite mixture model	239
Dirichlet process	240
Continuous mixtures	248
Beta-binomial and negative binomial	248
The Student's t-distribution	249
Summary	250
Exercises	251
Chapter 7: Gaussian Processes	253
Linear models and non-linear data	254
Modeling functions	255
Multivariate Gaussians and functions	257
Covariance functions and kernels	257
Gaussian processes	260
Gaussian process regression	261
Regression with spatial autocorrelation	268
Gaussian process classification	274
Cox processes	282
The coal-mining disasters	282
The redwood dataset	285
Summary	289
Exercises	289
Chapter 8: Inference Engines	291
Inference engines	292
Non-Markovian methods	293
Grid computing	293
Quadratic method	296
Variational methods	298
Automatic differentiation variational inference	301
Markovian methods	301
Monte Carlo	303
Markov chain	305
Metropolis-Hastings	305
Hamiltonian Monte Carlo	310
Sequential Monte Carlo	311
Diagnosing the samples	314
Convergence	315
Monte Carlo error	320
Autocorrelation	320
Effective sample sizes	321

Divergences	322
Non-centered parameterization	325
Summary	326
Exercises	326
Chapter 9: Where To Go Next?	329
Other Books You May Enjoy	333
Index	337
