

# Doing Bayesian Data Analysis

A Tutorial with R, JAGS,  
and Stan

**JOHN K. KRUSCHKE**

Dept. of Psychological and Brain Sciences  
Indiana University, Bloomington

AMSTERDAM • BOSTON • HEIDELBERG • LONDON  
NEW YORK • OXFORD • PARIS • SAN DIEGO  
SAN FRANCISCO • SINGAPORE • SYDNEY • TOKYO

**ELSEVIER**

Academic Press is an imprint of Elsevier

# CONTENT:

<b>1. What's in This Book (Read This First!)</b>	<b>1</b>
1.1 Real people can read this book	1
1.2 What's in this book	3
1.3 What's new in the second edition?	6
1.4 Gimme feedback (be polite)	8
1.5 Thank you!	8
<b>PART I The Basics: Models, Probability, Bayes' Rule, and R</b>	<b>13</b>
<b>2. Introduction: Credibility, Models, and Parameters</b>	<b>15</b>
2.1 Bayesian inference is reallocation of credibility across possibilities	16
2.2 Possibilities are parameter values in descriptive models	22
2.3 The steps of bayesian data analysis	25
2.4 Exercises	31
<b>3. The R Programming Language</b>	<b>33</b>
3.1 Get the software	35
3.2 A simple example of R in action	36
3.3 Basic commands and operators in R	38
3.4 Variable types	42
3.5 Loading and saving data	53
3.6 Some utility functions	56
3.7 Programming in R	61
3.8 Graphical plots: opening and saving	69
3.9 Conclusion	69
3.10 Exercises	70
<b>4. What Is This Stuff Called Probability?</b>	<b>71</b>
4.1 The set of all possible events	72
4.2 Probability: outside or inside the head	73
4.3 Probability distributions	78
4.4 Two-way distributions	89
4.5 Appendix: R code for figure 4.1	93
4.6 Exercises	95

<b>5. Bayes'Rule</b>	<b>99</b>
5.1 Bayes'rule	100
5.2 Applied to parameters and data	105
5.3 Complete examples: estimating bias in a coin	108
5.4 Why Bayesian inference can be difficult	115
5.5 Appendix: R code for figures 5.1,5.2, etc	116
5.6 Exercises	118
<b>PART II All the Fundamentals Applied to Inferring a Binomial Probability</b>	<b>121</b>
<b>6. Inferring a Binomial Probability via Exact Mathematical Analysis</b>	<b>123</b>
6.1 The likelihood function: Bernoulli distribution	124
6.2 A description of credibilities: the beta distribution	126
6.3 The posterior beta	132
6.4 Examples	134
6.5 Summary	138
6.6 Appendix: R code for figure 6.4	138
6.7 Exercises	139
<b>7. Markov Chain Monte Carlo</b>	<b>143</b>
7.1 Approximating a distribution with a large sample	145
7.2 A simple case of the Metropolis algorithm	146
7.3 The Metropolis algorithm more generally	156
7.4 Toward Gibbs sampling: estimating two coin biases	162
7.5 MCMC representativeness, accuracy, and efficiency	178
7.6 Summary	188
7.7 Exercises	189
<b>8. JAGS</b>	<b>193</b>
8.1 JAGS and its relation to R	193
8.2 A complete example	195
8.3 Simplified scripts for frequently used analyses	206
8.4 Example: difference of biases	208
8.5 Sampling from the prior distribution in JAGS	211
8.6 Probability distributions available in JAGS	213
8.7 Faster sampling with parallel processing in RunJAGS	215
8.8 Tips for expanding JAGS models	218
8.9 Exercises	218

<b>9. Hierarchical Models</b>	<b>221</b>
9.1 A single coin from a single mint	223
9.2 Multiple coins from a single mint	230
9.3 Shrinkage in hierarchical models	245
9.4 Speeding up JAGS	249
9.5 Extending the hierarchy: subjects within categories	251
9.6 Exercises	260
<b>10. Model Comparison and Hierarchical Modeling</b>	<b>265</b>
10.1 General formula and the Bayes factor	266
10.2 Example: two factories of coins	268
10.3 Solution by MCMC	274
10.4 Prediction: model averaging	289
10.5 Model complexity naturally accounted for	289
10.6 Extreme sensitivity to prior distribution	292
10.7 Exercises	295
<b>11. Null Hypothesis Significance Testing</b>	<b>297</b>
11.1 Paved with good intentions	300
11.2 Prior knowledge	315
11.3 Confidence interval and highest density interval	317
11.4 Multiple comparisons	325
11.5 What a sampling distribution <i>is</i> good for	329
11.6 Exercises	331
<b>12. Bayesian Approaches to Testing a Point ("Null") Hypothesis</b>	<b>335</b>
12.1 The estimation approach	336
12.2 The model-comparison approach	343
12.3 Relations of parameter estimation and model comparison	352
12.4 Estimation or model comparison?	354
12.5 Exercises	355
<b>13. Goals, Power, and Sample Size</b>	<b>359</b>
13.1 The will to power	360
13.2 Computing power and sample size	366
13.3 Sequential testing and the goal of precision	383
13.4 Discussion	393
13.5 Exercises	396

<b>14. Stan</b>	<b>399</b>
14.1 HMC sampling	400
14.2 Installing Stan	407
14.3 A complete example	407
14.4 Specify models top-down in Stan	414
14.5 Limitations and extras	415
14.6 Exercises	415
<b>PART III The Generalized Linear Model</b>	<b>417</b>
<b>15. Overview of the Generalized Linear Model</b>	<b>419</b>
15.1 Types of variables	420
15.2 Linear combination of predictors	423
15.3 Linking from combined predictors to noisy predicted data	435
15.4 Formal expression of the GLM	444
15.5 Exercises	446
<b>16. Metric-Predicted Variable on One or Two Groups</b>	<b>449</b>
16.1 Estimating the mean and standard deviation of a normal distribution	450
16.2 Outliers and robust estimation: the $t$ distribution	458
16.3 Two groups	468
16.4 Other noise distributions and transforming data	472
16.5 Exercises	473
<b>17. Metric Predicted Variable with One Metric Predictor</b>	<b>477</b>
17.1 Simple linear regression	478
17.2 Robust linear regression	479
17.3 Hierarchical regression on individuals within groups	490
17.4 Quadratic trend and weighted data	495
17.5 Procedure and perils for expanding a model	501
17.6 Exercises	504
<b>18. Metric Predicted Variable with Multiple Metric Predictors</b>	<b>509</b>
18.1 Multiple linear regression	510
18.2 Multiplicative interaction of metric predictors	525
18.3 Shrinkage of regression coefficients	530
18.4 Variable selection	536
18.5 Exercises	549

<b>19. Metric Predicted Variable with One Nominal Predictor</b>	<b>553</b>
19.1 Describing multiple groups of metric data	554
19.2 Traditional analysis of variance	556
19.3 Hierarchical Bayesian approach	557
19.4 Including a metric predictor	568
19.5 Heterogeneous variances and robustness against outliers	573
19.6 Exercises	579
<b>20. Metric Predicted Variable with Multiple Nominal Predictors</b>	<b>583</b>
20.1 Describing groups of metric data with multiple nominal predictors	584
20.2 Hierarchical Bayesian approach	588
20.3 Rescaling can change interactions, homogeneity, and normality	599
20.4 Heterogeneous variances and robustness against outliers	602
20.5 Within-subject designs	606
20.6 Model comparison approach	616
20.7 Exercises	618
<b>21. Dichotomous Predicted Variable</b>	<b>621</b>
21.1 Multiple metric predictors	622
21.2 Interpreting the regression coefficients	629
21.3 Robust logistic regression	635
21.4 Nominal predictors	636
21.5 Exercises	646
<b>22. Nominal Predicted Variable</b>	<b>649</b>
22.1 Softmax regression	650
22.2 Conditional logistic regression	655
22.3 Implementation in JAGS	659
22.4 Generalizations and variations of the models	667
22.5 Exercises	668
<b>23. Ordinal Predicted Variable</b>	<b>671</b>
23.1 Modeling ordinal data with an underlying metric variable	672
23.2 The case of a single group	675
23.3 The case of two groups	682
23.4 The case of metric predictors	685
23.5 Posterior prediction	698
23.6 Generalizations and extensions	699
23.7 Exercises	700

<b>24. Count Predicted Variable</b>	<b>703</b>
24.1 Poisson exponential model	704
24.2 Example: hair eye go again	711
24.3 Example: interaction contrasts, shrinkage, and omnibus test	713
24.4 Log-linear models for contingency tables	715
24.5 Exercises	715
<b>25. Tools in the Trunk</b>	<b>721</b>
25.1 Reporting a Bayesian analysis	721
25.2 Functions for computing highest density intervals	725
25.3 Reparameterization	729
25.4 Censored data in JAGS	732
25.5 What next?	736
<i>Bibliography</i>	737
<i>Index</i>	747