

## Contents

|               |  |           |
|---------------|--|-----------|
|               | <i>List of Illustrations</i>   | page xiii |
|               | <i>List of Tables</i>  | xviii     |
|               | <i>List of Contributors</i>  | xix       |
|               | <i>Preface</i>   | xxi       |
| <b>Part I</b> | <b>Introduction to Modeling</b>  | <b>1</b>  |
| <b>1</b>      | <b>Introduction</b>  | <b>3</b>  |
|               | 1.1 Models and Theories in Science   | 3         |
|               | 1.2 Quantitative Modeling in Cognition   | 6         |
|               | 1.2.1 Models and Data  | 6         |
|               | 1.2.2 Data Description   | 9         |
|               | 1.2.3 Cognitive Process Models   | 13        |
|               | 1.3 Potential Problems: Scope and Falsifiability                               | 17        |
|               | 1.4 Modeling as a “Cognitive Aid” for the Scientist                            | 20        |
|               | 1.5 <i>In Vivo</i>   | 22        |
| <b>2</b>      | <b>From Words to Models</b>  | <b>24</b> |
|               | 2.1 Response Times in Speeded-Choice Tasks                                     | 24        |
|               | 2.2 Building a Simulation  | 26        |
|               | 2.2.1 Getting Started: R and RStudio   | 26        |
|               | 2.2.2 The Random-Walk Model  | 27        |
|               | 2.2.3 Intuition vs. Computation: Exploring the Predictions of a<br>Random Walk | 31        |
|               | 2.2.4 Trial-to-Trial Variability in the Random-Walk Model                      | 33        |
|               | 2.2.5 A Family of Possible Sequential-Sampling Models                          | 37        |
|               | 2.3 The Basic Toolkit  | 38        |
|               | 2.3.1 Parameters   | 38        |
|               | 2.3.2 Connecting Model and Data  | 40        |
|               | 2.4 <i>In Vivo</i>   | 40        |

|                |  |     |
|----------------|--|-----|
| <b>Part II</b> | <b>Parameter Estimation</b>                                | 45  |
| <b>3</b>       | <b>Basic Parameter Estimation Techniques</b>               | 47  |
| 3.1            | Discrepancy Function                                       | 47  |
| 3.1.1          | Root Mean Squared Deviation (RMSD)                         | 48  |
| 3.1.2          | Chi-Squared ( $\chi^2$ )                                   | 49  |
| 3.2            | Fitting Models to Data: Parameter Estimation Techniques    | 50  |
| 3.3            | Least-Squares Estimation in a Familiar Context             | 50  |
| 3.3.1          | Visualizing Modeling                                       | 51  |
| 3.3.2          | Estimating Regression Parameters                           | 53  |
| 3.4            | Inside the Box: Parameter Estimation Techniques            | 57  |
| 3.4.1          | Simplex  | 57  |
| 3.4.2          | Simulated Annealing  | 61  |
| 3.4.3          | Relative Merits of Parameter Estimation Techniques         | 64  |
| 3.5            | Variability in Parameter Estimates                         | 65  |
| 3.5.1          | Bootstrapping  | 65  |
| 3.6            | <i>In Vivo</i>   | 70  |
| <b>4</b>       | <b>Maximum Likelihood Parameter Estimation</b>             | 72  |
| 4.1            | Basics of Probabilities                                    | 72  |
| 4.1.1          | Defining Probability                                       | 72  |
| 4.1.2          | Properties of Probabilities                                | 73  |
| 4.1.3          | Probability Functions                                      | 75  |
| 4.2            | What Is a Likelihood?                                      | 80  |
| 4.3            | Defining a Probability Distribution                        | 85  |
| 4.3.1          | Probability Functions Specified by the Psychological Model | 86  |
| 4.3.2          | Probability Functions via Data Models                      | 86  |
| 4.3.3          | Two Types of Probability Functions                         | 91  |
| 4.3.4          | Extending the Data Model                                   | 92  |
| 4.3.5          | Extension to Multiple Data Points and Multiple Parameters  | 93  |
| 4.4            | Finding the Maximum Likelihood                             | 95  |
| 4.5            | Properties of Maximum Likelihood Estimators                | 101 |
| 4.6            | <i>In Vivo</i>   | 103 |
| <b>5</b>       | <b>Combining Information from Multiple Participants</b>    | 105 |
| 5.1            | It Matters How You Combine Data from Multiple Units        | 105 |
| 5.2            | Implications of Averaging                                  | 106 |
| 5.3            | Fitting Aggregate Data                                     | 109 |
| 5.4            | Fitting Individual Participants                            | 111 |
| 5.5            | Fitting Subgroups of Data and Individual Differences       | 113 |
| 5.5.1          | Mixture Modeling   | 113 |
| 5.5.2          | K-Means Clustering   | 118 |
| 5.5.3          | Modeling Individual Differences                            | 121 |
| 5.6            | <i>In Vivo</i>   | 123 |

|          |   |     |
|----------|---|-----|
| <b>6</b> | <b>Bayesian Parameter Estimation</b>  | 126 |
| 6.1      | What Is Bayesian Inference?   | 126 |
| 6.1.1    | From Conditional Probabilities to Bayes Theorem                             | 126 |
| 6.1.2    | Marginalizing Probabilities   | 129 |
| 6.2      | Analytic Methods for Obtaining Posteriors                                   | 130 |
| 6.2.1    | The Likelihood Function   | 130 |
| 6.2.2    | The Prior Distribution  | 131 |
| 6.2.3    | The Evidence or Marginal Likelihood   | 134 |
| 6.2.4    | The Posterior Distribution  | 135 |
| 6.2.5    | Estimating the Bias of a Coin   | 136 |
| 6.2.6    | Summary   | 139 |
| 6.3      | Determining the Prior Distributions of Parameters                           | 139 |
| 6.3.1    | Non-Informative Priors  | 139 |
| 6.3.2    | Reference Priors  | 142 |
| 6.4      | <i>In Vivo</i>  | 143 |
| <b>7</b> | <b>Bayesian Parameter Estimation</b>  | 146 |
| 7.1      | Markov Chain Monte Carlo Methods  | 146 |
| 7.1.1    | The Metropolis-Hastings Algorithm for MCMC                                  | 147 |
| 7.1.2    | Estimating Multiple Parameters  | 153 |
| 7.2      | Problems Associated with MCMC Sampling                                      | 160 |
| 7.2.1    | Convergence of MCMC Chains  | 161 |
| 7.2.2    | Autocorrelation in MCMC Chains  | 162 |
| 7.2.3    | Outlook   | 162 |
| 7.3      | Approximate Bayesian Computation: A Likelihood-Free Method                  | 163 |
| 7.3.1    | Likelihoods That Cannot be Computed   | 163 |
| 7.3.2    | From Simulations to Estimates of the Posterior                              | 164 |
| 7.3.3    | An Example: ABC in Action   | 166 |
| 7.4      | <i>In Vivo</i>  | 170 |
| <b>8</b> | <b>Bayesian Parameter Estimation</b>  | 172 |
| 8.1      | Gibbs Sampling  | 172 |
| 8.1.1    | A Bivariate Example of Gibbs Sampling                                       | 173 |
| 8.1.2    | Gibbs vs. Metropolis-Hastings Sampling                                      | 176 |
| 8.1.3    | Gibbs Sampling of Multivariate Spaces                                       | 176 |
| 8.2      | JAGS: An Introduction   | 177 |
| 8.2.1    | Installing JAGS   | 177 |
| 8.2.2    | Scripting for JAGS  | 177 |
| 8.3      | JAGS: Revisiting Some Known Models and Pushing Their Boundaries             | 182 |
| 8.3.1    | Bayesian Modeling of Signal-Detection Theory                                | 182 |
| 8.3.2    | A Bayesian Approach to Multinomial Tree Models:<br>The High-Threshold Model | 186 |
| 8.3.3    | A Bayesian Approach to Multinomial Tree Models                              | 190 |
| 8.3.4    | Summary   | 198 |
| 8.4      | <i>In Vivo</i>  | 198 |

|           |  |     |
|-----------|--|-----|
| x         | <b>Contents</b>  |     |
| <b>9</b>  | <b>Multilevel or Hierarchical Modeling</b>                         | 203 |
|           | 9.1 Conceptualizing Hierarchical Modeling                          | 203 |
|           | 9.2 Bayesian Hierarchical Modeling                                 | 204 |
|           | 9.2.1 Graphical Models   | 204 |
|           | 9.2.2 Hierarchical Modeling of Signal-Detection Performance        | 207 |
|           | 9.2.3 Hierarchical Modeling of Forgetting                          | 211 |
|           | 9.2.4 Hierarchical Modeling of Inter-Temporal Preferences          | 218 |
|           | 9.2.5 Summary  | 226 |
|           | 9.3 Hierarchical Maximum Likelihood Modeling                       | 228 |
|           | 9.3.1 Hierarchical Maximum Likelihood Model of Signal Detection    | 228 |
|           | 9.4 Recommendations  | 233 |
|           | 9.5 <i>In Vivo</i>   | 234 |
|           | <b>Part III Model Comparison</b>                                   | 239 |
| <b>10</b> | <b>Model Comparison</b>  | 241 |
|           | 10.1 Psychological Data and the Very Bad Good Fit                  | 241 |
|           | 10.1.1 Model Complexity and Over-Fitting                           | 243 |
|           | 10.2 Model Comparison  | 248 |
|           | 10.3 The Likelihood Ratio Test                                     | 249 |
|           | 10.4 Akaike's Information Criterion                                | 256 |
|           | 10.5 Other Methods for Calculating Complexity and Comparing Models | 261 |
|           | 10.5.1 Cross-Validation  | 262 |
|           | 10.5.2 Minimum Description Length                                  | 262 |
|           | 10.5.3 Normalized Maximum Likelihood                               | 263 |
|           | 10.6 Parameter Identifiability and Model Testability               | 264 |
|           | 10.6.1 Identifiability   | 264 |
|           | 10.6.2 Testability   | 269 |
|           | 10.7 Conclusions   | 270 |
|           | 10.8 <i>In Vivo</i>  | 271 |
| <b>11</b> | <b>Bayesian Model Comparison Using Bayes Factors</b>               | 273 |
|           | 11.1 Marginal Likelihoods and Bayes Factors                        | 273 |
|           | 11.2 Methods for Obtaining the Marginal Likelihood                 | 277 |
|           | 11.2.1 Numerical Integration                                       | 278 |
|           | 11.2.2 Simple Monte Carlo Integration and Importance Sampling      | 280 |
|           | 11.2.3 The Savage-Dickey Ratio                                     | 284 |
|           | 11.2.4 Transdimensional Markov Chain Monte Carlo                   | 287 |
|           | 11.2.5 Laplace Approximation                                       | 294 |
|           | 11.2.6 Bayesian Information Criterion                              | 297 |
|           | 11.3 Bayes Factors for Hierarchical Models                         | 301 |
|           | 11.4 The Importance of Priors                                      | 303 |
|           | 11.5 Conclusions   | 306 |
|           | 11.6 <i>In Vivo</i>  | 306 |

|   |     |
|---|-----|
| <b>Part IV Models in Psychology</b>                                 | 309 |
| <b>12 Using Models in Psychology</b>                                | 311 |
| 12.1 Broad Overview of the Steps in Modeling                        | 311 |
| 12.2 Drawing Conclusions from Models                                | 312 |
| 12.2.1 Model Exploration  | 312 |
| 12.2.2 Analyzing the Model  | 314 |
| 12.2.3 Learning from Parameter Estimates                            | 315 |
| 12.2.4 Sufficiency of a Model                                       | 316 |
| 12.2.5 Model Necessity  | 318 |
| 12.2.6 Verisimilitude vs. Truth                                     | 323 |
| 12.3 Models as Tools for Communication and Shared Understanding     | 324 |
| 12.4 Good Practices to Enhance Understanding and Reproducibility    | 326 |
| 12.4.1 Use Plain Text Wherever Possible                             | 326 |
| 12.4.2 Use Sensible Variable and Function Names                     | 327 |
| 12.4.3 Use the Debugger   | 327 |
| 12.4.4 Commenting   | 328 |
| 12.4.5 Version Control  | 328 |
| 12.4.6 Sharing Code and Reproducibility                             | 329 |
| 12.4.7 Notebooks and Other Tools                                    | 330 |
| 12.4.8 Enhancing Reproducibility and Runnability                    | 331 |
| 12.5 Summary  | 332 |
| 12.6 <i>In Vivo</i>   | 332 |
| <b>13 Neural Network Models</b>                                     | 334 |
| 13.1 Hebbian Models   | 334 |
| 13.1.1 The Hebbian Associator                                       | 334 |
| 13.1.2 Hebbian Models as Matrix Algebra                             | 339 |
| 13.1.3 Describing Networks Using Matrix Algebra                     | 348 |
| 13.1.4 The Auto-Associator  | 349 |
| 13.1.5 Limitations of Hebbian Models                                | 356 |
| 13.2 Backpropagation  | 356 |
| 13.2.1 Learning and the Backpropagation of Error                    | 360 |
| 13.2.2 Applications and Criticisms of Backpropagation in Psychology | 364 |
| 13.3 Final Comments on Neural Networks                              | 365 |
| 13.4 <i>In Vivo</i>   | 366 |
| <b>14 Models of Choice Response Time</b>                            | 369 |
| 14.1 Ratcliff's Diffusion Model                                     | 369 |
| 14.1.1 Fitting the Diffusion Model                                  | 371 |
| 14.1.2 Interpreting the Diffusion Model                             | 383 |
| 14.1.3 Falsifiability of the Diffusion Model                        | 385 |
| 14.2 Ballistic Accumulator Models                                   | 386 |
| 14.2.1 Linear Ballistic Accumulator                                 | 386 |
| 14.2.2 Fitting the LBA  | 388 |

|           |   |     |
|-----------|---|-----|
| xii       | <b>Contents</b>   |     |
|           | 14.3 Summary  | 391 |
|           | 14.4 Current Issues and Outlook                             | 392 |
|           | 14.5 <i>In Vivo</i>   | 393 |
| <b>15</b> | <b>Models in Neuroscience</b>                               | 395 |
|           | 15.1 Methods for Relating Neural and Behavioral Data        | 397 |
|           | 15.2 Reinforcement Learning Models                          | 398 |
|           | 15.2.1 Theories of Reinforcement Learning                   | 398 |
|           | 15.2.2 Neuroscience of Reinforcement Learning               | 404 |
|           | 15.3 Neural Correlates of Decision-Making                   | 410 |
|           | 15.3.1 Rise-to-Threshold Models of Saccadic Decision-Making | 410 |
|           | 15.3.2 Relating Model Parameters to the BOLD Response       | 411 |
|           | 15.3.3 Accounting for Response Time Variability             | 413 |
|           | 15.3.4 Using Spike Trains as Model Input                    | 415 |
|           | 15.3.5 Jointly Fitting Behavioral and Neural Data           | 417 |
|           | 15.4 Conclusions  | 420 |
|           | 15.5 <i>In Vivo</i>   | 421 |
|           | <b>Appendix A Greek Symbols</b>                             | 424 |
|           | <b>Appendix B Mathematical Terminology</b>                  | 425 |
|           | <i>References</i>   | 427 |
|           | <i>Index</i>  | 455 |