Contents

Acknowledgments xv
The Four Steps of Bayesian Modeling xvii
List of Acronyms xix

Introduction 1

1 Uncertainty and Inference 7
1.1 The Goal of Perception 7
1.2 Hypotheses and Their Probabilities 8
1.3 Sensory Noise and Perceptual Ambiguity 12
1.4 Bayesian Inference in Visual Perception 14
1.5 Bayesian Inference in Auditory Perception 18
1.6 Historical Background: Perception as Unconscious Inference 23
1.7 Summary 25
1.8 Suggested Readings 26
1.9 Problems 27

2 Using Bayes' Rule 31
2.1 Steps of Bayesian Modeling 31
2.2 Alternative Form of Bayes' Rule 35
2.3 Areal Representation 35
2.4 The Prosecutor's Fallacy 37
2.5 A Changing Prior: Luggage Carousel Example 38
2.6 A Flat Prior: Gestalt Perception Example 42
2.7 Optimality, Evolution, and Motivations for Bayesian Modeling 44
2.8 Summary 45
2.9 Suggested Readings 45
2.10 Problems 46

3 Bayesian Inference under Measurement Noise 53
3.1 The Steps of Bayesian Modeling 53
3.2 Step 1: The Generative Model 55
3.2.1 The Measurement: An Abstracted Sensory Representation 55
3.2.2 Graphical Model 56
3.2.3 The Stimulus Distribution 57

3.2.4 The Measurement Distribution 58
3.2.5 Joint Distribution 59
3.3 Step 2: Inference 59
3.3.1 The Prior Distribution 60
3.3.2 The Likelihood Function 60
3.3.3 The Posterior Distribution 63
3.3.4 The Posterior Mean 66
3.3.5 Width of the Posterior 67
3.3.6 The Posterior Mean Estimate 68
3.3.7 The MAP Estimate 68
3.4 Uncertainty and Confidence 69
3.4.1 Uncertainty 69
3.4.2 Bayesian Confidence 70
3.5 Model Mismatch in Inference 71
3.5.1 Prior Mismatch 72
3.5.2 Improper Priors 72
3.6 Heteroskedasticity 73
3.7 Magnitude Variables 73
3.8 Applications 75
3.9 Percepts 76
3.10 Summary 76
3.11 Suggested Readings 77
3.12 Problems 78

4 The Response Distribution 83
4.1 Inherited Variability 83
4.2 The Response Distribution 84
4.3 Belief versus Response Distributions 86
4.4 Maximum-Likelihood Estimation 87
4.5 Bias and Mean Squared Error 88
4.5.1 An "Inverted Bias" Perspective 91
4.5.2 All Expenses Paid 92
4.6 Other Estimates 93
4.7 Decision Noise and Response Noise 94
4.8 Misconceptions 95
4.9 Reflections on Bayesian Models 97
4.10 Summary 97
4.11 Suggested Readings 98
4.12 Problems 98

5 Cue Combination and Evidence Accumulation 105
5.1 What Is Cue Combination? 105
5.2 Formulation of the Bayesian Model 108
5.2.1 Step 1: Generative Model 108
5.2.2 Step 2: Inference 110
5.2.3 Step 3: Estimate Distribution 112
5.3 Artificial Cue Conflict 113
5.3.1 Distinguishing the Distributions 114
5.4 Generalizations: Prior, Multiple Cues 114
5.5 Evidence Accumulation 115
5.6 Cue Combination under Ambiguity 116
5.7 Applications 117
5.8 Summary 118

5.9 Suggested Readings 118
5.10 Problems 119

6 Learning as Inference 125
6.1 The Many Forms of Learning 125
6.2 Learning the Probability of a Binary Event 126
6.2.1 Prediction 128
6.2.2 Update Equations 130
6.2.3 Uncertainty 131
6.2.4 Binomial Distribution 131
6.2.5 Nonuniform Prior 131
6.3 Linking Bayesian Learning to Reinforcement Learning 132
6.4 Learning the Precision of a Normal Distribution 134
6.4.1 Why Not Infer Variance? 135
6.4.2 Prediction 136
6.5 Learning the Slope of a Linear Relationship 137
6.6 Learning the Structure of a Causal Model 139
6.7 More Learning 141
6.8 Summary 141
6.9 Suggested Readings 142
6.10 Problems 143

7 Discrimination and Detection 147
7.1 Example Tasks 147
7.2 Discrimination 148
7.2.1 Step 1: Generative Model 148
7.2.2 Step 2: Inference 148
7.2.3 Gaussian Model 152
7.2.4 Decision Rule in Terms of the Measurement 152
7.2.5 Multiple Tasks Can Have the Same Bayesian Decision Rule 153
7.2.6 Step 3: Response Distribution 153
7.3 Detection 156
7.4 Confidence 156
7.5 Further Characterizing the Response Distribution 159
7.5.1 Receiver Operating Characteristic 159
7.5.2 Discriminability 161
7.6 The Relation between Bayesian Inference and Signal Detection Theory 162
7.7 Extensions 163
7.8 Summary 163
7.9 Suggested Readings 163
7.10 Problems 164

8 Binary Classification 169
8.1 Example Tasks 169
8.2 Generative Model 171
8.2.1 Mirror-Symmetric Class-Conditioned Stimulus Distributions 172
8.3 Marginalization 173
8.3.1 The Sum of Two Die Rolls 173
8.3.2 Continuous Variables 174
8.3.3 Conditioned Marginalization 175
8.3.4 Using the Generative Model 176
8.4 Inference 177
8.5 Response Distribution 179

8.6 Beyond Mirror-Image Class-Conditioned Stimulus Distributions 181
8.7 "Following the Arrows" 183
8.8 Summary 184
8.9 Suggested Readings 184
8.10 Problems 185

9 Top-Level Nuisance Variables and Ambiguity 191
9.1 Example Tasks 191
9.2 Size as a Top-Level Nuisance Variable in Depth Perception 192
9.3 Marginalization Formulation 196
9.4 In Color Perception 197
9.5 In Object Recognition 200
9.6 Summary 201
9.7 Suggested Readings 202
9.8 Problems 202

10 Same-Different Judgment 205
10.1 Example Tasks 205
10.2 Binary Stimuli 206
10.2.1 Step 1: Generative Model 206
10.2.2 Step 2: Inference 207
10.2.3 Step 3: Estimate Distribution 208
10.3 Continuous Stimuli 210
10.3.1 Step 1: Generative Model 211
10.3.2 Step 2: Inference 211
10.3.3 Step 3: Response Probabilities 213
10.3.4 Step 2 Revisited: Inferring the Stimuli 213
10.4 Multiple-Item Sameness Judgment 215
10.5 Perceptual Organization 217
10.5.1 Contour Integration 219
10.5.2 Intersecting Lines 220
10.6 Summary 222
10.7 Suggested Readings 223
10.8 Problems 223

11 Search 227
11.1 Many Forms of Visual Search 227
11.2 Target Localization: Camouflage 229
11.3 Target Localization with Measurement Noise 232
11.4 Target Detection: Camouflage 234
11.5 Target Detection with Measurement Noise 237
11.6 Applications 239
11.7 Summary 239
11.8 Suggested Readings 239
11.9 Problems 240

12 Inference in a Changing World 245
12.1 Tracking a Continuously Changing World State 245
12.1.1 Step 1: Generative Model 245
12.1.2 Step 2: Inference 247
12.2 Change Point Detection 249
12.2.1 Single Change Point 250
12.2.2 Random Change Points 252
12.2.3 More Realistic Change Point Detection 253

12.3 Summary 253
12.4 Suggested Readings 254
12.5 Problems 254

13 Combining Inference with Utility 257
13.1 Example Tasks 257
13.2 Deciding between Two Actions 258
13.3 Deciding among Several Actions 260
13.4 Mathematical Formulation: Expected Utility 262
13.4.1 Binary Classification 263
13.4.2 Continuous Estimation 264
13.5 The Cost Functions of "Pure Perception" 265
13.5.1 Discrete Tasks 265
13.5.2 Continuous Estimation 265
13.5.3 Perception and Action 267
13.6 What It Means to Decide Optimally 267
13.7 Complications 268
13.7.1 Cost Functions for Uncertain Outcomes 268
13.7.2 Nonlinearity between Reward and Utility 269
13.7.3 Distortions of Probability 270
13.7.4 Decision Noise 270
13.8 Applications 271
13.8.1 Visual Discrimination 271
13.8.2 Reaching Movements 272
13.8.3 Incentivized Confidence Intervals 272
13.8.4 Inferring Utility Functions 274
13.9 Summary 274
13.10 Suggested Readings 275
13.11 Problems 276

14 The Neural Likelihood Function 281
14.1 Generative Model of the Activity of a Single Neuron 281
14.2 Neural Likelihood Function for a Single Neuron 283
14.2.1 Case 1: A Bell-Shaped Tuning Curve 285
14.2.2 Case 2: A Monotonic Tuning Curve 286
14.3 Neural Likelihood Function Based on a Population of Neurons 287
14.4 Toy Model 290
14.5 Relation between Abstract and Neural Concepts 293
14.6 Using the Neural Likelihood Function for Computation 294
14.7 The Neural Implementation of Bayesian Computation 295
14.8 Applications 295
14.9 Summary 296
14.10 Suggested Readings 297
14.11 Problems 298

15 Bayesian Models in Context 301
15.1 Bayesian versus Optimal Behavior 301
15.2 Overly Strong Claims of Optimality 302
15.3 Understanding Why Some Behaviors Are Optimal and Others Are Not 303
15.4 Bayesian Models Are Not Mechanistic Models 303
15.5 Bayesian Transfer 304
15.6 Probabilistic Computation and Hybrid Models 305
15.7 Real-World Complexity 305

15.8 Learning to Be Bayesian 307
15.9 Suggested Readings 308

Appendices 311

A Notation 313

B Basics of Probability Theory 315
B.1 Objective and Subjective Probability 315
B.2 The Intuitive Notion of Probability 316
B.3 Complementary Event 316
B.4 Venn Diagram Representation 317
B.5 Random Variables and Their Distributions 317
B.5.1 Discrete versus Continuous Random Variables 318
B.5.2 Total Probability = 1 318
B.5.3 Discrete Probability Distributions 319
B.5.4 Continuous Probability Distributions 319
B.5.5 Formal Definition of the Probability Density Function 321
B.5.6 Normalization 322
B.6 Mean, Variance, and Expected Value 322
B.7 The Normal Distribution 323
B.7.1 Definition 323
B.7.2 Central Limit Theorem 323
B.7.3 Multiplying Two Normal Distributions 324
B.7.4 Multiplying Multiple Normal Distributions 325
B.7.5 The Cumulative Normal Distribution 325
B.7.6 The Von Mises Distribution 326
B.8 The Delta Function 326
B.9 The Poisson Distribution 327
B.10 Drawing from a Probability Distribution 327
B.11 Distributions Involving Multiple Variables 327
B.11.1 Joint Probability 328
B.11.2 Marginalization 328
B.11.3 Conditional Probability 329
B.11.4 Independence 331
B.11.5 Bayes' Rule 332
B.12 Functions of Random Variables 333
B.12.1 Functions of One Variable: Changing Variables 333
B.12.2 Apple Example 335
B.12.3 The Commitment of Ignorance 337
B.12.4 Marginalization Formulation 337
B.12.5 Functions of Multiple Variables 338
B.13 Problems 340

C Model Fitting and Model Comparison 343
C.1 What Is a Model? 343
C.2 Free Parameters 343
C.3 The Parameter Likelihood 344
C.4 Maximum-Likelihood Estimation 344
C.5 Fitting Data from an Estimation Task 345
C.5.1 Simple Model 346
C.5.2 A Better Model 350
C.5.3 Model Comparison 350

C.5.4 Likelihood Ratio Test 351
C.5.5 Cross-Validation 352
C.5.6 Comparing Model Comparison Methods 352
C.5.7 Parameter Recovery and Model Recovery 353
C.5.8 Limitations of Model Comparison 353
C.6 Absolute Goodness of Fit 354
C.7 Fitting Data from a Discrimination Task 354
C.7.1 Simple Model 354
C.7.2 A Better Model? 356
C.8 Fitting Data from a Classification Task 357
C.9 Good Experimental Design for Bayesian Modeling 357
C.10 Suggested Readings 358
C.11 Problems 359

Bibliography 361
Index 371