Contents

Preface vii

Glossary xvii

About the Authors xx

1 Introduction 1
- 1.1 Objectives of the Book 1
- 1.2 Background Material 4
  - 1.2.1 Inference, Given a Model 4
  - 1.2.2 The Critical Issue: "What Model To Use?" 5
  - 1.2.3 Science Inputs—Formulation of the Set of Candidate Models 7
  - 1.2.4 Models Versus Full Reality 11
  - 1.2.5 A "Best Approximating Model" 13
- 1.3 Overview of Models in the Biological Sciences 13
  - 1.3.1 Types of Models Commonly Used in the Biological Sciences 14
  - 1.3.2 Likelihood and Least Squares Theory 15
  - 1.3.3 Data Dredging 17
  - 1.3.4 Some Trends 20
- 1.4 Inference and the Principle of Parsimony 21
  - 1.4.1 Over-fitting to Achieve a Good Model Fit 21
  - 1.4.2 The Principle of Parsimony 23
  - 1.4.3 Model Selection Methods 27

1.5 Model Selection Uncertainty 29
1.6 Summary 30

2 Information Theory and Log-Likelihood Models: A Basis for Model Selection and Inference **32**
2.1 The Distance or Discrepancy Between Two Models 33
2.1.1 f, Truth, Full Reality, and "True Models" 36
2.1.2 g, Approximating Models 36
2.1.3 The Kullback–Liebler Distance (or Information) 37
2.1.4 Truth, f, Drops Out as a Constant 41
2.2 Akaike's Information Criterion 43
2.3 Akaike's Predictive Expected Log-Likelihood 49
2.4 Important Refinements to AIC 51
2.4.1 A Second-Order AIC 51
2.4.2 Modification to AIC for Overdispersed Count Data . . 52
2.5 A Useful Analogy 54
2.6 Some History 56
2.6.1 The G-Statistic and K-L Information 56
2.6.2 Further Insights 57
2.6.3 Entropy 58
2.6.4 A Summary 59
2.7 Further Comments 59
2.7.1 A Heuristic Interpretation 60
2.7.2 Interpreting Differences Among AIC Values 60
2.7.3 Nonnested Models 63
2.7.4 Model Selection Uncertainty 64
2.7.5 AIC When Different Data Sets Are to Be Compared . . 64
2.7.6 Order Not Important in Computing AIC Values 64
2.7.7 Hypothesis Testing Is Still Important 65
2.8 Comparisons with Other Criteria 66
2.8.1 Information Criteria That Are Estimates of K-L Information 66
2.8.2 Criteria That Are Consistent for K 68
2.8.3 Contrasts 70
2.9 Return to Flather's Models 71
2.10 Summary 72

3 Practical Use of the Information-Theoretic Approach **75**
3.1 Computation and Interpretation of AIC Values 76
3.2 Example 1—Cement Hardening Data 78
3.2.1 Set of Candidate Models 79
3.2.2 Some Results and Comparisons 79
3.2.3 A Summary 82
3.3 Example 2—Time Distribution of an Insecticide Added to a Simulated Ecosystem 83

3.3.1 Set of Candidate Models 84
3.3.2 Some Results 86
3.4 Example 3—Nestling Starlings 87
3.4.1 Experimental Scenario 87
3.4.2 Monte Carlo Data 88
3.4.3 Set of Candidate Models 88
3.4.4 Data Analysis Results 92
3.4.5 Further Insights into the First 14 Nested Models 94
3.4.6 Hypothesis Testing and Information-Theoretic Approaches Have Different Selection Frequencies . . . 96
3.4.7 Further Insights Following Final Model Selection . . . 99
3.4.8 Why Not Always Use the Global Model for Inference? . 100
3.5 Example 4—Sage Grouse Survival 101
3.5.1 Introduction 101
3.5.2 Set of Candidate Models 102
3.5.3 Model Selection 104
3.5.4 Hypothesis Tests for Year-Dependent Survival Probabilities 106
3.5.5 Hypothesis Testing Versus AIC in Model Selection . . 106
3.5.6 A Class of Intermediate Models 110
3.6 Example 5—Resource Utilization of *Anolis* Lizards 110
3.6.1 Set of Candidate Models 111
3.6.2 Comments on Analytic Method 112
3.6.3 Some Tentative Results 112
3.7 Summary 114

4 Model-Selection Uncertainty with Examples 118
4.1 Introduction 118
4.2 Methods for Assessing Model-Selection Uncertainty 120
4.2.1 AIC Differences and a Confidence Set on the K-L Best Model 122
4.2.2 Likelihood of a Model and Akaike Weights 123
4.2.3 More Options for a Confidence Set for the K-L Best Model 127
4.2.4 Δ_i, Model-Selection Probabilities and the Bootstrap . . 129
4.2.5 Concepts of Parameter-Estimation and Model-Selection Uncertainty 130
4.2.6 Including Model-Selection Uncertainty in Estimator Sampling Variance 133
4.2.7 Unconditional Confidence Intervals 137
4.2.8 Uncertainty of Variable Selection 140
4.2.9 Model Redundancy 141
4.2.10 Recommendations 144
4.3 Examples 145
4.3.1 Cement Data 145

4.3.2 *Anolis* Lizards in Jamaica 151
4.3.3 Simulated Starling Experiment 151
4.3.4 Sage Grouse in North Park 152
4.3.5 Sakamoto et al.'s (1986) Simulated Data 152
4.3.6 Pine Wood Data 153
4.4 Summary 157

5 Monte Carlo and Example-Based Insights 159
5.1 Introduction 159
5.2 Survival Models 160
5.2.1 A Chain Binomial Survival Model 160
5.2.2 An Example 163
5.2.3 An Extended Survival Model 168
5.2.4 Model Selection If Sample Size Is Huge, or Truth Known 171
5.2.5 A Further Chain Binomial Model 173
5.3 Examples and Ideas Illustrated with Linear Regression 176
5.3.1 All-Subsets Selection: A GPA Example 177
5.3.2 A Monte Carlo Extension of the GPA Example 182
5.3.3 An Improved Set of GPA Prediction Models 186
5.3.4 More Monte Carlo Results 189
5.3.5 Linear Regression and Variable Selection 195
5.3.6 Discussion 199
5.4 Estimation of Density from Line Transect Sampling 205
5.4.1 Density Estimation Background 205
5.4.2 Line Transect Sampling of Kangaroos at Wallaby Creek 206
5.4.3 Analysis of Wallaby Creek Data 206
5.4.4 Bootstrap Analysis 208
5.4.5 Confidence interval on D 208
5.4.6 Bootstrap Reps: 1,000 vs. 10,000. 210
5.4.7 Bootstrap vs. Akaike Weights: A lesson on $QAIC_c$ . . . 211
5.5 An Extended Binomial Example 213
5.5.1 The Durban Storm Data 213
5.5.2 Models Considered 214
5.5.3 Consideration of Model Fit 216
5.5.4 Confidence Intervals on Predicted Storm Probability . . 217
5.5.5 Precision Comparisons of Estimators 219
5.6 Lessons from the Literature and Other Matters 221
5.6.1 Use AIC_c, Not AIC, with Small Sample Sizes 221
5.6.2 Use AIC_c, Not AIC, When K Is Large 222
5.6.3 Inference from a Less Than Best Model 224
5.6.4 Are Parameters Real? 226
5.7 Summary 227

6 Statistical Theory **230**
6.1 Useful Preliminaries 230
6.2 A General Derivation of AIC 239
6.3 General K-L–Based Model Selection: TIC 248
6.3.1 Analytical Computation of TIC 248
6.3.2 Bootstrap Estimation of TIC 249
6.4 AIC_c: A Second-Order Improvement 251
6.4.1 Derivation of AIC_c 251
6.4.2 Lack of Uniqueness of AIC_c 255
6.5 Derivation of AIC for the Exponential Family of Distributions . 257
6.6 Evaluation of $\text{tr}(J(\underline{\theta}_o)[I(\underline{\theta}_o)]^{-1})$ and Its Estimator 261
6.6.1 Comparison of AIC vs. TIC in a Very Simple Setting . 261
6.6.2 Evaluation Under Logistic Regression 266
6.6.3 Evaluation Under Multinomially Distributed Count Data 273
6.6.4 Evaluation Under Poisson-Distributed Data 280
6.6.5 Evaluation for Fixed-Effects Normality-Based Linear Models 281
6.7 Additional Results and Considerations 288
6.7.1 Selection Simulation for Nested Models 288
6.7.2 Simulation of the Distribution of Δ_p 290
6.7.3 Does AIC Over-fit? 292
6.7.4 Can Selection Be Improved Based on All the Δ_i? . . . 294
6.7.5 Linear Regression, AIC, and Mean Square Error 296
6.7.6 AIC and Random Coefficient Models 299
6.7.7 AIC_c and Models for Multivariate Data 302
6.7.8 There Is No True TIC_c 303
6.7.9 Kullback–Leibler Information Relationship to the Fisher Information Matrix 304
6.7.10 Entropy and Jaynes Maxent Principle 304
6.7.11 Akaike Weights, w_i, Versus Selection Probabilities, π_i . 305
6.7.12 Model Goodness-of-Fit After Selection 306
6.8 Kullback–Leibler Information Is Always ≥ 0 307
6.9 Summary 312

7 Summary **315**
7.1 The Scientific Question and the Collection of Data 317
7.2 Actual Thinking and A Priori Modeling 317
7.3 The Basis for Objective Model Selection 319
7.4 The Principle of Parsimony 321
7.5 Information Criteria as Estimates of Relative Kullback–Leibler Information 321
7.6 Ranking and Calibrating Alternative Models 324
7.7 Model-Selection Uncertainty 325
7.8 Inference Based on Model Averaging 326

7.9 More on Inferences 327
7.10 Final Thoughts 328

References **329**

Index **351**