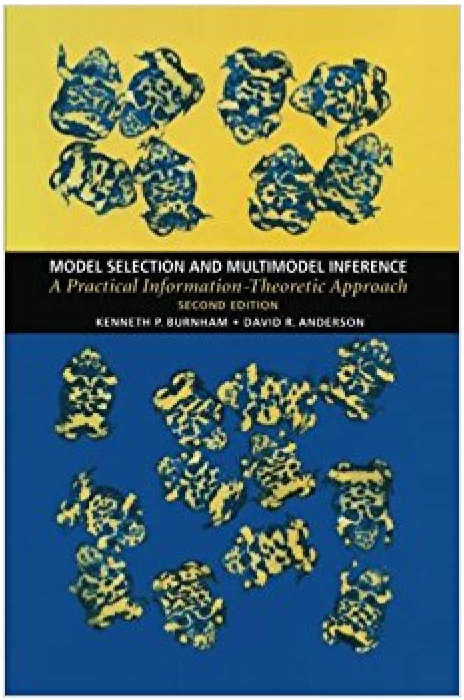
Model Selection And Multimodel Inference: A Practical Information-Theoretic Approach



Synopsis

A unique and comprehensive text on the philosophy of model-based data analysis and strategy for the analysis of empirical data. The book introduces information theoretic approaches and focuses critical attention on a priori modeling and the selection of a good approximating model that best represents the inference supported by the data. It contains several new approaches to estimating model selection uncertainty and incorporating selection uncertainty into estimates of precision. An array of examples is given to illustrate various technical issues. The text has been written for biologists and statisticians using models for making inferences from empirical data.

Book Information

Paperback: 488 pages

Publisher: Springer (December 1, 2010)

Language: English

ISBN-10: 1441929738

ISBN-13: 978-1441929730

Product Dimensions: 6.1 x 1.2 x 9.2 inches

Shipping Weight: 2 pounds (View shipping rates and policies)

Average Customer Review: 4.0 out of 5 stars 10 customer reviews

Best Sellers Rank: #804,947 in Books (See Top 100 in Books) #36 inÂ Books > Science & Math > Mathematics > Applied > Biomathematics #239 inÂ Books > Textbooks > Medicine & Health Sciences > Research > Biostatistics #404 inÂ Books > Medical Books > Basic Sciences > Biostatistics

Customer Reviews

This is an excellent book on model selection and multi-model inference. It covers in great detail the underlying theoretical and philosophical foundations for model selection and provides practical examples with sufficient degree of detail so you could replicate the results on your own. I found it especially useful that this book applies the Aikake Information Criteria to non-stationary time series, as in 3.3 Example 2: Time Distribution of an Insecticide Added to a Simulated Ecosystem.

very good quality, very good price

I admire this book very much for its accessible treatment of AIC, but if were reduced in length by half, it would be twice as good. The authors cannot resist repeating themselves, usually several times, especially when giving advice of the "motherhood and apple pie" variety. Another annoying feature is that many references are given for philosophical points, yet sometimes when a useful result is given without proof, no reference is provided. For example, on page 12 an expression for maximized likelihood is given without a derivation or a reference. Inside this fat book there is a thin book crying to be let out.

Burnham does a tremendous job of explaining both how and why to use information theory to fit statistical models to data. Everything is here with sufficient detail and clarity to guide your own development work.

If you want to learn about model selection techniques and multimodel inference, this is your book. In my opinion, the first few chapters should be required reading for anyone using model selection techniques. The later chapters become quite technical (above my head, I'm not ashamed to say!) but they are undoubtedly important as well, and I'll work through them eventually merely due to the merit I find in the chapters I have read.

These two authors are very well known for their work and opinions on model selection and alternatives to null-hypothesis testing IN THE 20TH CENTURY. However, this book has declined in utility and is not a 21st century view of statistics, as modern methods and synthetic views have outpaced their dated ideas while incorporating some of the best ideas from the early 20th century. More importantly, it is clear now that the strategy where one approach is trashed in favor of another does not contribute to progress in science. Their straw-man characterizations of standard frequentist statistics and null hypothesis approaches are not useful when it is clear that any approach you utilize has strengths and weaknesses and that one should tailor her or his statistical models to the question, the data, and the parameter estimates of interest. I would not recommend this book to anyone at this date (even though I have purchased it, used it, and recommended it in the past), rather I would urge the authors to read the other pieces in the Ecology (2014; Ecology 95: 609-653) special feature to which they contributed but did not really participate in a collaborative or open-minded way. I agree with the other authors in that feature: plot, check assumptions, estimate parameters, examine effect sizes, plot, think about biological significance, and ignore those who yell too loudly about how there is only one approach to science.

This book emphasizes the fundamentals of proper model selection that are pretty hard to master in school. Building a good model is a long process that requires a decent level of qualitative understanding of the data generating mechanism. Based on that knowledge, a sizable amount of work should be done well before the data are plugged into a statistical software package. Of course, if one is in a very data-rich situation, one can get away with "let the computer sort it out" approach, but such cases are rare.After I finished the book, my understanding of the bias-variance tradeoff principle improved substantially. In particular, one should remember that overfitting is not only about including redundant covariates that then cause abysmal out-of-sample performance. Suppose, based on the qualitative information about the data generating process, you are convinced that certain factor(s) "should" be in the model. As the sample size decreases, you may find out that the "compulsory" factor(s) must be dropped to preserve the optimal bias-variance tradeoff. The catch is that, even if you manage to guess exactly what factors constitute the "true" model, the corresponding regression coefficients still have to be estimated from the data. The more coefficients, the greater the complexity of your model pool. If the sample size is low enough, you will be forced to reduce the complexity to avoid overfitting, which entails excluding some "true" factors from consideration.Another major statistical concept the book clarified for me is the elusive distinction between "fixed" and "random" effects. The most popular definition goes as follows: if interested in the particular experimental units per se, then use fixed effects; if interested in the population represented by the experimental units, then use random effects. This is not of much help because it implies that, for a given dataset, one can alternate between fixed and random versions without much justification. To make things worse, there exist four more definitions of what random and fixed effects are (Google Gelman, 2005, "Analysis of variance - why it is more important than ever").Burnham & Anderson provide a very clear explanation of why we may wish to switch from fixed to random effects. Suppose there is a categorical Factor A with a large number of levels. If we treat it as a fixed factor, we have to estimate a large number of parameters, which can result in an overfitted model. If we drop the factor altogether, we obtain an underfitted model, which is not good either. What we need is a model of intermediate complexity, which can be generated by treating Factor A as random.However, this approach has a major pitfall: if we want to do model selection based on the information criteria such as AIC, all of the models in the pool should be obtainable from the "global" model by imposing restrictions on the deterministic parameters. In this example, the model where Factor A is taken as fixed cannot be compared to the intermediate model where Factor A is taken as random. More generally, the models under consideration can differ only w.r.t. fixed factors.I am tempted to get too far into technicalities, so let me stop here by saying that, as long as your goal is to learn the true philosophy of model building art & science, this book is for you.

Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach Multiagent Systems: Algorithmic, Game-Theoretic, and Logical Foundations Insider Secrets From A Model Agent: How To Become A Successful Model (Modeling, Modelling, Model Agency) RCadvisor's Model Airplane Design Made Easy: The Simple Guide to Designing R/C Model Aircraft or Build Your Own Radio Control Flying Model Plane The Selection: The Selection, Book 1 Happily Ever After: Companion to the Selection Series (The Selection Novella) Looking for Information: A Survey of Research on Information Seeking, Needs, and Behavior: 4th Edition (Studies in Information) Looking for Information: A Survey of Research on Information Seeking, Needs, and Behavior (Studies in Information) Fundamentals Of Information Systems Security (Information Systems Security & Assurance) - Standalone book (Jones & Bartlett Learning Information Systems Security & Assurance) Counterfactuals and Causal Inference: Methods and Principles for Social Research (Analytical Methods for Social Research) A Practical Approach to Cardiac Anesthesia (Practical Approach Series) Immunoassays: A Practical Approach (Practical Approach Series) A Practical Approach to Obstetric Anesthesia (A Practical Approach to Anesthesia) HPLC of Macromolecules: A Practical Approach (Practical Approach Series) Transition Metals in Organic Synthesis: A Practical Approach (The Practical Approach in Chemistry Series) Oligonucleotide Synthesis: A Practical Approach (The Practical Approach Series) A Practical Approach to Pediatric Anesthesia (Practical Approach to Anesthesia) Patient Education: A Practical Approach (PATIENT EDUCATION: A PRACTICAL APPROACH (MUMA)) Observation and Experiment: An Introduction to Causal Inference The Elements of Statistical Learning: Data Mining, Inference, and Prediction, Second Edition (Springer Series in Statistics)