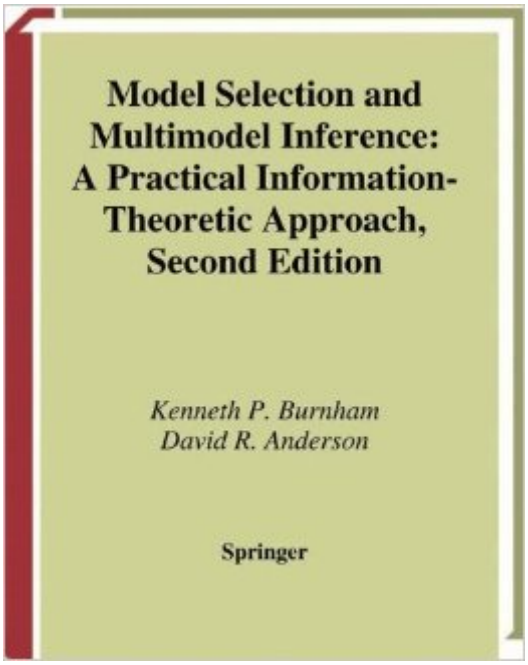
Model Selection And Multimodel Inference: A Practical Information-Theoretic Approach



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

Hardcover: 488 pages

Publisher: Springer; 2nd edition (December 4, 2003)

Language: English

ISBN-10: 0387953647

ISBN-13: 978-0387953649

Product Dimensions: 6.1 x 1.1 x 9.2 inches

Shipping Weight: 1.8 pounds (View shipping rates and policies)

Average Customer Review: 4.2 out of 5 starsÂ Â See all reviewsÂ (9 customer reviews)

Best Sellers Rank: #532,086 in Books (See Top 100 in Books) #19 inÂ Books > Science & Math > Mathematics > Applied > Biomathematics #125 inÂ Books > Textbooks > Medicine & Health Sciences > Research > Biostatistics #225 inÂ Books > Medical Books > Basic Sciences > Biostatistics

Customer Reviews

Burnham and Anderson have put together a scholarly account of the developments in model selection techniques from the information theoretic viewpoint. This is an important practical subject. As computer algorithms become more and more available for fitting models and data mining and exploratory analysis become more popular and used more by novices, problems with overfitting models will again raise their ugly heads. This has been an issue for statisticians for decades. But the problems and the art of model selection has not been commonly covered in elementary courses on statistics and regression. George Box puts proper emphasis on the iterative nature of model selection and the importance of applying the principle of parismony in many of his books. Classic texts on regression like Draper and Smith point out the pitfalls of goodness of ift measures like R-square and explain Mallows Cp and adjusted R-square. There are now also a few good books

devoted to model selection including the book by McQuarrie and Tsai (that I recently reviewed for) and the Chapman and Hall monograph by A. J. Miller.Burnham and Anderson address all these issues and provide the best coverage to date on bootstrap and cross-validation approaches. They also are careful in their historical account and in putting together some coherence to the scattered literature. They are thorough in their references to the literature. Their theme is the information theoretic measures based on the Kullback-Liebler distance measure. The breakthrough in this theory came from Akaike in the 1970s and improvements and refinement came later. The authors provide the theory, but more importantly, they provide many real examples to illustrate the problems and show how the methods work.

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.

Obsolete Diecast Model Cars Basic Painting and Weathering for Model Railroaders, Second Edition (Model Railroader Books: Essentials) Basic Model Railroad Benchwork (Model Railroader Essentials Series) Building N Gauge Model Railway Layouts (Building Model Railway Layouts Book 1) Introduction to Model Railroading (Model Railroader's How-To Guides) Painting Backdrops for Your Model Railroad (Model Railroader's How-To Guides) Designing & Building Multi-Deck Model Railroads (Model Railroader) Model Railroad Electronics: Basic Concepts to Advanced Projects (Model Railroad Handbook) Ho Scale Model Railroading: Getting Started in the Hobby (Model Railroader Books) Building a Ready-To-Run Model Railroad (Model Railroader's How-To Guides) How to Build Model Railroad Benchwork, Second Edition (Model Railroader) Basic Structure Modeling for Model Railroaders (Model Railroader Books) Model Locomotive Boilermaking (Model Engineering Guides) The Jennifer Nicole Lee Fitness Model Diet: JNL's Super Fitness Model Secrets To A Sexy, Strong, Sleek Physique Counterfactuals and Causal Inference: Methods and Principles for Social Research (Analytical Methods for Social Research) Bayesian Methods for Hackers: Probabilistic Programming and Bayesian Inference (Addison-Wesley Data & Analytics)