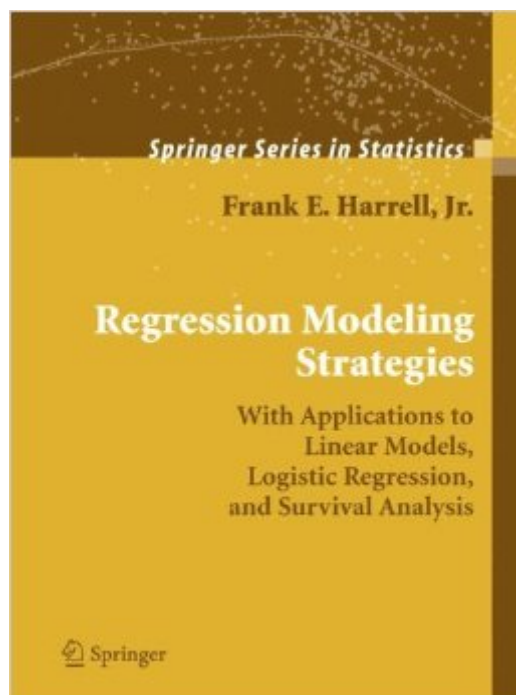


The book was found

# Regression Modeling Strategies: With Applications To Linear Models, Logistic Regression, And Survival Analysis (Springer Series In Statistics)



## Synopsis

Many texts are excellent sources of knowledge about individual statistical tools, but the art of data analysis is about choosing and using multiple tools. Instead of presenting isolated techniques, this text emphasizes problem solving strategies that address the many issues arising when developing multivariable models using real data and not standard textbook examples. It includes imputation methods for dealing with missing data effectively, methods for dealing with nonlinear relationships and for making the estimation of transformations a formal part of the modeling process, methods for dealing with "too many variables to analyze and not enough observations," and powerful model validation techniques based on the bootstrap. This text realistically deals with model uncertainty and its effects on inference to achieve "safe data mining".

## Book Information

Series: Springer Series in Statistics

Paperback: 572 pages

Publisher: Springer (June 15, 2001)

Language: English

ISBN-10: 1441929185

ISBN-13: 978-1441929181

Product Dimensions: 7 x 1.4 x 9.2 inches

Shipping Weight: 2.6 pounds (View shipping rates and policies)

Average Customer Review: 4.4 out of 5 stars [See all reviews](#) (18 customer reviews)

Best Sellers Rank: #703,547 in Books (See Top 100 in Books) #171 in [Books > Textbooks > Medicine & Health Sciences > Research > Biostatistics](#) #309 in [Books > Medical Books > Basic Sciences > Biostatistics](#) #467 in [Books > Computers & Technology > Software > Mathematical & Statistical](#)

## Customer Reviews

Frank Harrell is a Professor who does a lot of consulting in medical research. This book covers a wide variety of topics in regression analysis including many advanced techniques including data reduction, smoothing techniques, variable selection, transformations, shrinkage methods, tree-based methods and resampling. But note the title "Regression Modeling Strategies". Unlike most advanced texts in regression this book emphasizes modeling strategies. So the focus is on things like variable selection and other techniques to avoid overfitting models and diagnostics to look for violations in assumptions such as variance homogeneity or normality and independence of

residuals, or stability problems like colinearity. The book covers an extensive collection of modern techniques for exploratory data analysis. Inferential methods are also considered and he deals appropriately with important issues (particularly for medical research) such as imputation of missing values. Many examples are considered and illustrated in S-PLUS. Harrell also provides many rules of thumb based on his own experience building models. A lot of the techniques are illustrated using data from the Titanic where it is interesting to see which factors affected the probability of survival. My only disappointment was that there is perhaps too much emphasis on this one particular data set. A standard regression text would be expected to include linear and nonlinear regression. Harrell goes much deeper including nonparametric regression, logistic regression and survival models (e.g. the Cox proportional hazards model).

This is a very special statistics book and is unlike any other that I have encountered. Instead of being focused on a specific statistical technique (or family of techniques), Harrell presents a wholistic view of regression modeling for describing real datasets. He starts with the basics of regression assumptions and techniques (splines, shrinkage, etc...), moves on to data management (imputation and reduction), and then addresses the specifics of linear regression, binary logistic regression, ordinal logistic regression, parametric survival regression and Cox regression. Each regression method is approached first with a clear explanation of what the technique is doing and what the critical assumptions are. Then, Harrell demonstrates how to do the analysis in S-Plus/R using a real dataset. Though I lack the advanced mathematical background necessary to fully explore many statistical textbooks, I did not find this to be a problem for this one. The presentation is that of a teacher: clear with developed reasoning. The production of nomograms was a particularly useful exercise and the S-plus code was also very useful. I find his opinions on model building strategies to be well thought out and persuasive...though I suspect that many may find them controversial. Overall, this is one of the best statistics books that I have purchased.

This book covers a lot of advanced regression techniques and is intended for an audience that has been through an introductory regression analysis class. This book sets itself apart by emphasizing modeling strategies rather than teaching just theory or applied regression in the textbook fashion in which all the examples work out perfectly and one doesn't have to worry about dirty data. Instead it talks extensively about topics that occur in the real world, such as handling missing data, a big problem when dealing with real data but rarely mentioned in most regression texts. He also talks about many of the traps students get themselves from what they learn in an intro regression class is

not always exactly the best way of going about doing things and explains better alternatives. Through out the book, use of S-plus and R is liberal which is very nice. Numerous extensive case studies thoroughly analyze data sets using many of the techniques he describes and gives full S-plus/R code for them to recreate on your own. Unfortunately, I really didn't like the data sets he chose to analyze. Many of them were medical related, another used Titanic survivors data, another was about the 2000 election, while very well done, I found the datasets themselves rather uninteresting. This of course leads to a problem, me being an engineer, I'd rather have datasets I can relate to, while of course a social scientist would like sets they could relate to, so I realize the author has a hard time making everyone happy. It would be nice to have perhaps had additional case studies available on the book website, perhaps worked by other individuals from a variety of disciplines.

This is clearly an advanced text that mathematicians and PhD students in statistics would find valuable. It is not for an engineer or novice statistician in industry (like myself) who has to come up with an accurate regression model with quantitative and qualitative data in a short period of time. My rating is four stars: buy this book only if you have the advanced statistical training to understand it, otherwise buy a simpler book if you want to get a basic understanding of the subject.

[Download to continue reading...](#)

Regression Modeling Strategies: With Applications to Linear Models, Logistic Regression, and Survival Analysis (Springer Series in Statistics) Regression Methods in Biostatistics: Linear, Logistic, Survival, and Repeated Measures Models (Statistics for Biology and Health) Deep Learning in Python Prerequisites: Master Data Science and Machine Learning with Linear Regression and Logistic Regression in Python (Machine Learning in Python) Statistics and Data Analysis for Financial Engineering: with R examples (Springer Texts in Statistics) Applied Linear Regression Models- 4th Edition with Student CD (McGraw Hill/Irwin Series: Operations and Decision Sciences) Time Series Analysis: With Applications in R (Springer Texts in Statistics) A Modern Approach to Regression with R (Springer Texts in Statistics) Applied Logistic Regression Introduction to Linear Regression Analysis Time Series Analysis (Springer Texts in Statistics) Applied Regression Analysis: A Second Course in Business and Economic Statistics (Book, CD-ROM & InfoTrac) An Introduction to Statistical Learning: with Applications in R (Springer Texts in Statistics) Regression Analysis: Understanding and Building Business and Economic Models Using Excel The Ada® Generic Library: Linear List Processing Packages (Springer Compass International) Regression to Times and Places (Meditation Regression) Spiritual Progress Through

Regression (Meditation Regression) Regression Through The Mirrors of Time (Meditation Regression) Time Series: Theory and Methods (Springer Series in Statistics) The Elements of Statistical Learning: Data Mining, Inference, and Prediction, Second Edition (Springer Series in Statistics) System Analysis and Modeling: Models and Reusability: 8th International Conference, SAM 2014, Valencia, Spain, September 29-30, 2014. Proceedings (Lecture Notes in Computer Science)

[Dmca](#)