Journal of Statistical Software

July 2004, Volume 11, Book Review 7. http://www.jstatsoft.org/

Reviewer: Thomas H. Scheike
University of Copenhagen

Survival Analysis Using S

Mara Tableman and Jong Sung Kim
Chapman & Hall/CRC, Boca Raton, Florida, 2004.
ISBN 1-58488-408-8. xv + 260 pp. $69.95.

This book is aimed at students having their first encounter with survival analysis. The students that the authors have in mind are not particularly mathematically oriented and the authors aim primarily at introducing the basic concepts and showing how to implement the methods using S. The book contains 8 chapters that I will comment on briefly in the following.

Chapter 1 introduces the reader to the special methods that are needed to deal with censoring by considering some examples.

Chapter 2 introduces the standard non-parametric methods and contains material on the Kaplan-Meier, the log-rank test and some stratified methods.

Chapter 3 discusses parametric methods and the maximum likelihood principle.

Chapter 4 introduces the reader to regression models. The chapter starts off with the exponential and Weibull regression model and moves on to a brief introduction of the Cox regression model and the accelerated failure time model. The material in these sections does not contain a strong applied component but focuses more on the underlying formulae. The final section of this chapter introduces the reader to Akaike's information criterion (AIC) and shows how it can be used for comparing the parametric models of the chapter using S.

Chapter 5 shows how the Cox regression model may be applied to real data using S. One issue that takes up a lot of space is how to use the AIC for automated model selection procedures in S. In biostatistics the search for models will often be guided by subject matter rather than automatic model selection procedures, and then one will investigate specific interactions to validate a simple additive structure among covariates. One section introduces the stratified Cox regression model and shows how it can be used to validate the assumption of constant relative risk over time. Everybody who has tried using these graphical tests will know how difficult it is both in theory and practice to decide if the curves are parallel (on the appropriate scale), and this is perhaps why the authors only have a brief remark saying that the model appears to fit.

Chapter 6 contains some general material on model checking. It introduces Cox's partial likelihood and the various residuals used in survival analysis. S programs illustrate how these

residuals may be used check the fit of Cox's regression model. One final section discusses how to find a maximum likelihood cut point for the Cox model using the profile likelihood principle and suggests a bootstrap to asses the uncertainty.

Chapter 7 moves on to consider the proportional hazards model with time-dependent covariates and shows how to fit the models in S (S makes it possible to deal with piecewise constant covariates). Additional topics covered in this chapter are competing risks models and left-truncation.

Chapter 8 is rather atypical for a first text on survival analyses and contains some new material on censored regression quantiles. This is a nice twist to the book. I think the authors do a good job of pointing out that other methods are needed when the Cox regression model fails to fit the data. Even when the Cox model fits it is often very informative to get hold of other summary statistics than those obtained from the Cox model. Regression quantiles provide such an alternative methodology. I fear that this chapter will be very hard to read for the intended readers, especially because various issues are not quite settled yet. As in previous chapters the methods are implemented in S.

The lacking formulae and mathematics underlying the different analyses makes it hard for the reader to understand all issues in depth, I had quite a number of additional *why*'s, but I am quite sure that the students will be able to do practical analyses on the computer and get the flavour of things.

All in all the book succeeds nicely in getting the reader through the basic methods of survival analysis (Kaplan-Meier, log-rank, Weibull and Cox regression) and how to implement them in S.

Reviewer:

Thomas H. Scheike
University of Copenhagen
Department of Biostatistics
Blegdamsvej 3, DK-2200 Copenhagen N
Telephone: +45/3532-7928
Fax: +45/3532-7907
E-mail: T.Scheike@biostat.ku.dk