Journal of Statistical Software

January 2019, Volume 88, Book Review 1. doi: 10.18637/jss.v088.b01

Reviewer: Carmen Armero
Universitat de València

Data Analysis Using Hierarchical Generalized Linear Models with R

Youngjo Lee, Lars Rönnegård, Maengseok Noh
Chapman & Hall/CRC, Boca Raton, 2017.
ISBN 978-1-138-62782-6. 321 pp. GBP 77.00 Hardback.
https://www.crcpress.com/9781138627826

This book has a very interesting title because it combines three important fields: Data analysis, hierarchical generalized linear models (HGLM), and the R environment. The content of the book is also very attractive: HGLM models, many R packages for dealing with different classes of models, and the analysis of very interesting and motivating examples. The last part of the book is dedicated to survival and joint models for survival and longitudinal data, which are statistical models very useful and powerful for data analysis. Its inclusion in the book is quite valuable because these subjects are not usual in this type of book.

All the statistical material is based on the Lee and Nelder (1996) hierarchical likelihood (h-likelihood) approach. The main objective of the book is helping the reader to understand the practical merits of that approach for data analysis. For that reason, the content of the book is essentially applied: interesting examples, R specific packages, and scripts. As a reader, I appreciated very much that material but I missed more theoretical and statistical information of the statistical models presented in the book. But of course, the book could be understood as a complementary part of a collection of more theoretical books on the subject (Lee, Nelder, and Pawitan 2006, 2017; Ha, Jeong, and Lee 2017), authored by the first author of this book. In any case, an introductory part on the h-likelihood approach would be highly appreciated for non-experts on the subject.

The book starts with a list of notations, a preface, and an introduction. In this part, the authors explain that the majority of data sets used in the book are available in the R package **mdhglm**. The main material of text is organized in ten chapters. Each chapter ends with a small section of exercises, most of them based on real problems.

Chapters 1–5 discuss many and varied motivating examples analysed by means of HGLMs models. In particular, Chapter 1 discusses the strengths of the h-likelihood approach but the definition of this approach appears for the first time in Chapter 3. I think that it would be necessary to introduce it before that discussion. The authors also explain that their book does not focus on the philosophical advantages of using the h-likelihood but a few pages later they consider that the h-likelihood approach is more general than the frequentist and the Bayesian approaches. I would have liked to read the justification of this opinion in the

book. Chapter 2 includes a small revision of generalized linear models (GLMs) with some examples and R code. Chapter 3 introduces the definition of the h-likelihood and the so called extended likelihood principle, Chapter 4 deals with HGLMs models, which incorporate random effects in the linear predictor structure of the GLMs models, and Chapter 5 describes practical analyses of these models by means of the R packages **hglm** and/or **dhglm**.

Chapters 6 and 7 present R packages **dhglm** and **mdhglm** for analysing double HGLM (DHGLM) and multivariate HGLM models, respectively. DHGLMs models are an extension of HGLMs models which allow to consider random effects in the variance components of the model. Multivariate HGLM models extend the case of a single response variable to two or more of them. Chapters 8 is dedicated to survival analysis and mainly discusses the h-likelihood approach to frailty models with some fitted examples analyzed via the **frailtyHL** package. Chapter 9 deals with joint models for longitudinal and survival data by means of the **jointdhglm** package. The last chapter, entitled 'Further Topics', is dedicated to variable selection and hypothesis testing. Chapters generally include the R code of the subsequent examples models as well as comments about the results of the analyses.

The book contains some format or notation limitations that can confuse the reader. For example, $\boldsymbol{y}$ and $\mathbf{y}$ are used interchangeably to represent the data (e.g., p. 29), not always a vector is represented in bold letters (see for instance the line after expression 2.2), and the relationship between the data and the parameters of Bayes' theorem in (2.2) is presented in terms of the stochastic model that will generate the data and not the likelihood function as it should be. In addition, I think that there are some comments on Bayesian Inference that are not clearly explained. I find it is strange to read that "...Bayesian does not make inference of fixed unknown." (p. 14).

In summary, I think that this book could be very interesting for people with an advanced statistical background as well as interest and some prior knowledge on the on the h-likelihood approach. In this sense, all the practical material in the book is an excellent complement for improving the knowledge about the practice of that approach.

References

Ha ID, Jeong J, Lee Y (2017). *Statistical Modelling of Survival Data with Random Effects: H-Likelihood Appoach*. Springer-Verlag, Singapore.

Lee Y, Nelder JA (1996). "Hierarchical Generalized Linear Models." *Journal of the Royal Statistical Society B*, **58**, 619–678.

Lee Y, Nelder JA, Pawitan Y (2006). *Generalized Linear Models with Random Effects*. Chapman and Hall, Boca Raton.

Lee Y, Nelder JA, Pawitan Y (2017). *Generalized Linear Models with Random Effects*. 2nd edition. Chapman and Hall, Boca Raton.

Reviewer:

Carmen Armero
Universitat de València
Department of Statistics and Operations Research
Burjassot 46100, Spain
E-mail: carmen.armero@uv.es
URL: http://www.uv.es/armero/

Journal of Statistical Software http://www.jstatsoft.org/
published by the Foundation for Open Access Statistics http://www.foastat.org/
January 2019, Volume 88, Book Review 1 *Published:* 2019-01-29
doi:10.18637/jss.v088.b01