



Journal of Statistical Software

October 2008, Volume 28, Book Review 1.

<http://www.jstatsoft.org/>

Reviewer: Wolfgang M. Hartmann
Heidelberg, Germany

SAS for Data Analysis: Intermediate Statistical Methods

Mervyn G. Marasinghe and William J. Kennedy
Springer-Verlag, New York, 2008.
ISBN 978-0-387-77371-1. 557 pp. USD 69.95.
<http://mervyn.public.iastate.edu/Book/>

The authors provide an easily readable introduction into the SAS language, some of its basic statistical methods, and many applications for statistical linear modeling. With the first three chapters of this book the authors show their excellence in teaching, with the remaining three chapters they show how to solve some more complicated problems in the statistical analysis of linear models. The book is designed for beginners in both, SAS language and statistics but assumes a steep learning curve or some prior knowledge of linear models. Very helpful are the many exercises in each chapter which make this book valuable for teaching at universities and colleges or just for teaching a very focussed reader. As of today, almost all test examples and data sets are available from the Web page accompanying the book (see above).

The last chapters may also be useful for SAS consultants at colleges, universities, and research facilities to look up a scientifically sound approach to a specific statistical problem. The book does not cover how to write SAS macros. Only the BOXANNO macro by [Friendly \(1991\)](#) is mentioned in book and index. Also, I did not find anything about the very important importing or exporting of data from or into other environments, e.g., SPSS or Microsoft **Excel**.

There is, however, a discrepancy between the general title and the specific content of this book: The book covers only a small part of the SAS/BASE, SAS/GRAPH, and SAS/STAT packages of SAS and not much is said about many other packages like SAS/ETS, SAS/OR, SAS/IML, SAS/QC, SAS Enterprise Miner, or even the financial software of SAS. Even for SAS/STAT only some basic PROCs such as UNIVARIATE and FREQ and the linear modeling PROCs REG, ANOVA, GLM, and MIXED are covered. Many other SAS/STAT PROCs, e.g., LOGISTIC, GENMOD, PHREG, ROBUSTREG, FACTEX, OPTEX, NLIN, PRINCOMP, imputation, PROCs for life and behavioral sciences etc. are not mentioned. A similar selection of SAS software is covered by [Littell, Freund, and Spector \(1991\)](#) and [Littell, Milliken, Stroup, and Wolfinger \(1996\)](#) but assuming more knowledge of SAS and less of statistics by the reader.

The book is divided into the following chapters:

1. *Introduction to the SAS language*: This is written for beginners and should enable every reader to write DATA and PROC steps in the SAS language. The text makes a good

effort to explain the SAS way of record-wise dealing with data which is so difficult to understand for some beginners.

2. *More on SAS programming and some applications:* This chapter describes some necessary tools for data processing and provide a number of examples for basic PROCs from SAS/STAT: UNIVARIATE, FREQ, and from SAS/BASE: PLOT, CHART, and TABULATE.
3. *Statistical graphics using SAS/GRAPH:* The authors provide a thorough introduction into writing those tedious SAS/GRAPH scripts with many examples. On a funny note, two PROCs are considered “useful”: GPLOT and GCHART. Also, a short introduction into some high resolution graphics with SAS is given here.
4. *Statistical analysis of regression models:* The chapter starts with simple and multiple linear regression models and illustrates the use of PROC REG with some examples. Forward, backward, stepwise and all-subsets methods are explained using PROC REG.
5. *Analysis of variance models:* This and the last chapter are probably the most valuable for readers who are already educated in SAS and statistics. Each of these two chapters covers about 100 pages of text. Here, one-way and two-way classification for fixed effects analysis of variance modeling are explained with examples specified for PROCs ANOVA and GLM.
6. *Analysis of variance: Random and mixed effects models:* One-way and two-way crossed and nested random and mixed effects models are estimated using PROCs GLM and MIXED.
7. *Appendix (with SAS/GRAPH statements and tables)*

That means the first three chapters are useful for readers inexperienced in both, the SAS language and statistics. The last three chapters, however, assume some more experience with SAS and a basic statistical education. Since the book contains so many examples, reading the first three chapters may enable the reader to go through the more advanced last chapters.

As a final remark, SAS Institute, the owner and distributor of SAS, does no longer (for several years) relate its name to the term “Statistical Analysis System” as noted by the authors.

References

- Friendly M (1991). *SAS System for Statistical Graphics*. SAS Insitute, Carey, NC.
- Littell RC, Freund RJ, Spector PC (1991). *SAS System for Linear Models*. SAS Insitute, Carey, NC.
- Littell RC, Milliken GA, Stroup WW, Wolfinger RD (1996). *SAS System for Mixed Models*. SAS Insitute, Carey, NC.

Reviewer:

Wolfgang M. Hartmann

Heidelberg, Germany

E-mail: cmat.wolfgang@gmail.com