Aura POPA

PhD Candidate, Department of Statistics and Econometrics,
University of Economics, Bucharest, Romania

E-mail: aura.popa@csie.ase.ro, **Web page:** http://www.aurapopa.ase.ro

Book Review on
REGRESSION MODELING: METHODS, THEORY, AND COMPUTATION WITH SAS[2],

by Michael J. PANIK,
Chapman&Hall/CRC, Taylor&Francis Group,
Boca Raton, FL, USA, 2009

The First Edition of „Modelling Regression: Methods, Theory, and Computation with SAS" describes both the conventional and less common uses of almost all the regression types in the practical context of today's mathematical, economic and scientific research. This book is designed to introduce the reader to the richness and diversity of regression techniques and is particularly well suited for use in a second course in statistics at the undergraduate or first-year graduate level. This book is a robust resource that offers solid methodology for statistical practitioners and professionals in this field and it is also ideal for students of the applied mathematics or statistics, sciences, economics, and engineering who routinely use regression analysis for decision making and problem solving. Scientists and engineers will find the book to be an excellent choice for reference and self-study. This book blends both theory and application to equip the reader with an understanding of the principles necessary to apply regression model-building techniques in the SAS® environment.

JAQM

Vol. 5
No. 1
Spring
2010

176
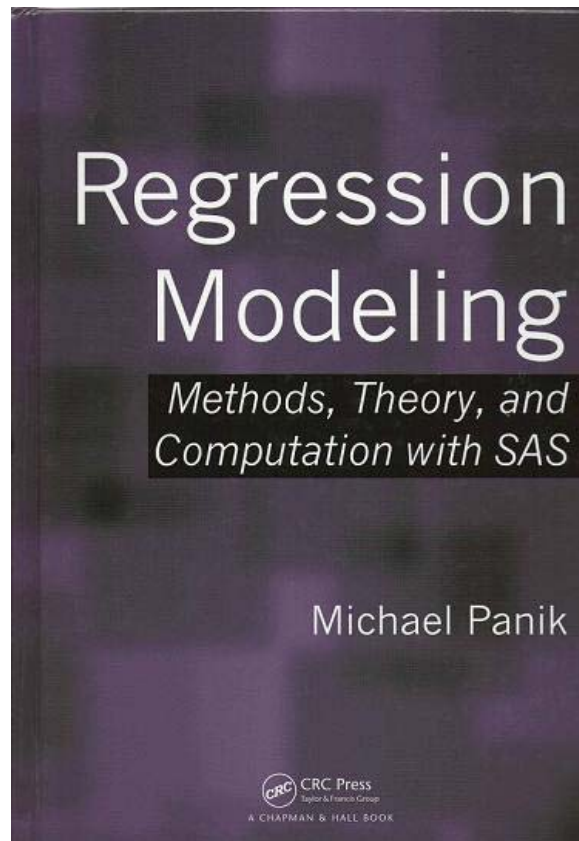
There are a total of nineteen chapters in the book, the first twelve of which form the core, ending in the most well developed chapter, the twelfth one entitled "Multiple Regression". This book has many strengths and important features. It is highly readable, and the material is quite accessible to those enrolled in applied statistics courses or engaged in self-study. In this regard, an objective of this work is to make students aware of the power and density of regression techniques without overwhelming them with calculations. Some common knowledge of matrices and linear algebra is not absolutely essential but, at certain points in the presentation, can be helpful. For convenience, a review of the essentials of random variables, probability distributions, and classical statistical inference is provided in **Chapter 1**. This introductory chapter begins by describing the need for knowing how to apply regression and also reminds the reader the basics of econometric methods.

Illustrating all of the major procedures employed by the contemporary software package SAS, this edition begins with a general introduction to regression modeling, including typical applications. A host of technical tools are outlined, such as basic inference procedures, introductory aspects of model adequacy checking, and polynomial regression models and their variations. The book discusses how transformations and weighted least squares can be used to resolve problems of model inadequacy and also how to deal with influential observations. Although there are many varieties of regression analyses from which to choose, one is most often exposed to ordinary least squares, but this is only a part of the regression story. This text fully exposes OLS and then offers many alternative regression methodologies. Specifically, the regression routines presented here include the following: ordinary least square—along with the method of maximum likelihood, bivariate linear regression and correlation, misspecified disturbance terms, nonparametric regression, logistic regression (including Poisson regression), Bayesian regression, robust regression techniques such as M-estimators, and properties of robust estimators, fuzzy regression, random coefficients regression, L_1 and q-quantile regression, regression in a spatial domain, multiple regression, normal correlation models, ridge regression, indicator variables, polynomial regression, semiparametric regression, nonlinear least squares and some time-series regression issues.

Regression analysis has undoubtedly been one of the most widely used techniques in applied statistics. As a consequence, there are a large number of excellent books written on the topic. And then what makes this book more special than others? Usually authors are bored or avoid mentioning some steps in explaining outputs after running the syntax, but this is not Panik's case! The rigorousness and thoroughness of the evidence in writing this book is appreciated, but for an avid fan of econometrics, many of the chapters can be considered only as a starter in each area, and not a few would expect more, because not only with the `qualitatively` is the reader satisfied, but because of the `quantitatively` that matters.

However, like any book that contains software packages instructions, some mistakes have crept in with the syntax of some programs, and perhaps in the future an errata will be provided to this book that deserves to be cultivated and appreciated, and lead to further reprinting and to also include any further developments because of it's high potential.

"Modeling Regression" is well organized. The chapters are sorted in a logical order, from intermediate theory level, computational algorithms, to advanced applications. Some starred sections are advanced and may be skipped for someone who just wants to apply some regression at some point in the SAS Software.

JAQM

Vol. 5
No. 1
Spring
2010

177

The book is compact, making it easy to extract a feeling of accomplishment and progress and this is making the reader wish to go on, reading more and more, without a large volume of pages being a disarming criteria.

The compactness distinguishes the book from those which try to be too complete and end up being intimidatingly thick. It retains the structure within each chapter. So if the reader is only interested in the example of regression issues related to space, he can jump to the 11th chapter. The presentation of each regression technique is fairly streamlined and designed to offer the reader an unencumbered look into its operation in that proofs and derivations are only supplied in chapter appendices; for those readers who want a more technical treatment, the appendices are a "must read". Moreover, to facilitate the understanding and appreciation of the various regression methods, only the bivariate case is covered in **Chapters 2** through **11**. In these chapters, most of the SAS programs can easily be extended to handle additional explanatory variables once multiple regression is covered. For example, **Chapter 11** provides an introduction to statistical methods for the analysis of spatial data. In a coherent manner, it presents a short classification of spatial data types: geostatistical data, lattice data, and point patterns. Rigorous theory is presented, but not thick, nor boring, but most importantly, bibliographical references are the best, as milestones within each theme separately, and they range from old publications dating from 1935 to date, throughout the entire book and not only in this chapter. Furthermore, regression space syntax can be considered very useful for those who are not good GIS software users, do not use file-type shape, but to reach some notions of space and to analyze the data.

Exercises that are proposed for solution at the end of each chapter are similar to those already solved. For many problems, the data requirements are given which focus more on implementation of syntax, especially running it and interpretation of results afterwards. For most of the regression methods presented, SAS procedure code is included for the convenience of the reader. Although the various sets of SAS code will enable students and practitioners alike to immediately perform their own regression runs, the code given is not all-encompassing, and by no means a substitute for reading the SAS Manuals. It is only intended to give the reader a jumpstart in solving regression problems. Hence, this is not the type of book that only offers theory and proofs; with a modicum of study and effort, one can "hit the ground running", so to speak, and readily generate some fairly sophisticated regression results. Once a regression technique is explained, SAS handles the "how to" portion of the presentation. This is imperative because one can first study a regression method and then, for the most part, directly apply it because numerous example problems are included, with the SAS results explained in considerable detail.

Also, plotting is great, with ample illustrations, explained, closely connected to the theoretical and practical part. Even thought it is printed in black and white, the graphs are so well explained and exposed, that you do not miss the lack of colors in them.

In particular, I appreciated some topics that are covered when we are talking about regression. One involves some highlights over spatial domain. The author might recommend this chapter for further research, because how it was presented: in a short fugitive way in comparison with other chapters, and because this area can not be exposed, even punctual in approximately ten pages. The other involves the excellent description for details of fuzzy theory and how to implement this large area into computer area.

I would appreciate it even more if this book would have had a dedicated website. Being published only at the beginning of last year, the publisher could invest more time and

JAQM

Vol. 5
No. 1
Spring
2010

178

could create a virtual area for the author to give the opportunity to provide for his readers access to several datasets, software, and probably other useful materials, sparing some extra paper, pages and reader's money.

In summary, I can recommend this book as a very good graduate level text-book. I also believe that readers will learn much more not just reading it, but using it with their computer, applying each exercise that the author is proposing.

JAQM

Vol. 5
No. 1
Spring
2010

179