

Review of **Semiparametric Regression**

Reviewer: Michael R. Chernick, Lankenau Institute for Medical Research, review at Amazon.com

David Ruppert and Ray Carroll have been a research team for over 25 years. They have published many articles and books on regression analysis. These articles are always very clearly written and are great at showing the big picture and not just the nitty gritty details of the theorems that they prove. Two of my favorite books that they published are "Transformations and Weighting in Regression" published by Chapman & Hall in 1988 and "Measurement Error in Nonlinear Models " with Stefanski in 1995 and also published by Chapman & Hall.

This book is no exception. It is lucid in exposition and paints a general picture summarizing the area of nonparametric regression models and incorporating them with parametric regression both linear and nonlinear.

Their work has also been motivated by the desire to extend the theory of regression models to practical problems where the standard theory with assumptions such as linearity, normality, and homogeneity of variance don't hold.

In the first chapter, they motivate their methods through a number of examples in the areas of health science and environmental pollution problems. Chapter two goes through the standard linear regression models and the diagnostic checks for those models. They also cover other practical issues including model selection, use of transformations and extensions to nonlinear models. The special case of polynomial regression (a particular example of linear regression) is presented in detail.

Chapter 3 on scatterplot smoothing introduces many of the key ideas to their approach to semiparametric regression. Their approach in its most general form is based on mixed models which are introduced in chapter 4. Chapter 5 deals with automated methods for implementing the scatterplot smoothing techniques. The remaining chapters cover for example, simple semiparametric models, additive models, semiparametric mixed models, and generalized parametric models which include the very useful generalized linear models that they have extended to cover mixed effects.

The generalized additive models of Tibshirani and Hastie are covered in chapter 11, Other important issues including variance function estimation, accounting for measurement error, Bayesian approaches and more are covered in the latter chapters (12-17), Finally in chapter 18 seven examples are introduced to illustrate applications of the various methods. An epilogue, chapter 19, was written to motivate further research.

Many of these chapters are the subjects of whole monographs including some that Ruppert and Carroll have co-authored. In the preface they say that the book is intended for three potential audiences. The first audience is the students and scientists with interest in applying the techniques or learning about them but possess at most a moderate

background in regression. The second audience (the group I would put myself in) are the biostatisticians, econometricians and scientists in other disciplines who have a good working knowledge of regression and want to add the flexibility of semiparametric methods to their arsenal of techniques. The third group is the researchers in nonparametric regression who may not yet know about some of the new advances of Carroll, Ruppert and Wand that are included in this text.

I find in general that their books are masterpieces. As a statistician who has done both applied and theoretical work, I know what it takes to write books that summarize a body of theory or connect the theory and applications, or incorporate new results. These authors do both of these things in this book. They have the rare talent to find a way to unify and simplify existing theory and that is another great feature you will find in this book.

I haven't been able to do that and only a few others that I know can. One example that comes to my mind is the book on extremes by Leadbetter, Lindgren and Rootzen. At the time of publication in 1978, they provided a unified theory for extremes combining the theory for independent and dependent cases. They also provide some examples in the book. But even that landmark book is heavily theoretical. Ruppert, Carroll and Wand emphasize applications and provide a number of examples throughout.