Frédéric Ferraty and Philippe Vieu

Nonparametric Functional Data Analysis

Theory and Practice

April 18, 2006

Springer
Berlin Heidelberg New York
Hong Kong London
Milan Paris Tokyo

Preface

This work is the fruit of recent advances concerning both nonparametric statistical modelling and functional variables and is based on various publications in international statistical reviews, several post-graduate courses and international conferences, which are the result of several years of research. In addition, all these developments have their roots in the recent infatuation for functional statistics. In particular, the synergy around the activities of the working group STAPH is a permanent source of inspiration in these statistical functional fields.

This book presents in a original way new nonparametric statistical methods for functional data analysis. Because we support the idea that statistics should take advantage of interactions between applied and theoretical aspects, we deliberately decided not to privilege one over the other. So, this work proposes two levels of reading. Recent theoretical advances, as far as possible, are presented in self-contained sections while statistical methodology, practical aspects, and elementary mathematics are accessible to a very large public. But, in any case, each part of this book starts with the presentation of general ideas concerning theoretical as well as applied issues.

This book could be useful as well for practitioners as for researchers and students. Non expert researchers and students will find detailed proofs and mathematical tools for theoretical advances presented in this book. For experienced researchers, these advances have been selected to balance the trade-off between comprehensive reading and up-to-date results. Because nonparametric functional statistics is a recent field of research, we discuss the existing bibliography by emphasizing open problems. This could be the starting point for further statistical developments. Practitioners will find short descriptions on how to implement the proposed methods while the companion website (*http://www.lsp.ups-tlse.fr/staph/npfda*) includes large details for codes, guidelines, and examples of use. So, the use of such nonparametric functional procedures will be easy for any users. In this way, we can say that this book is really intended for a large public: practitioners, theoreticians and anybody else who is interested in both aspects.

The novelty of nonparametric functional statistics obliges us to start by clarifying the terminology, by presenting the various statistical problems and by describing the kinds of data (mainly curves). Part I is devoted to these generalities. The remaining parts consist in describing the nonparametric statistical methods for functional data, each of them being basically split into theoretical, applied, and bibliographical issues. Part II focuses on prediction problems involving functional explanatory variables and scalar response. We study regression, conditional mode and conditional quantiles and their kernel nonparametric estimates. Part III concerns the classification of functional data. We focus successively on curve discrimination (prediction of a categorical response corresponding to the class membership) and unsupervised classification (i.e., the class membership is unobserved). Because time series can be viewed as a particular case of functional dataset, we propose in Part IV to extend most of the previous developments to dependent samples of functional data. The dependance structure will be taken into account through some mixing notion. In order to keep the main body of the text clear, theoretical tools are put at the end of this monograph in the appendix.

All the routines are implemented in the R and S+ languages and are available on the companion website (*http://www.lsp.ups-tlse.fr/staph/npfda*). S+ is an object-oriented language intensively used in engineering and applied mathematical sciences. Many universities, intitutions and firms use such a software which proposes just as well a very large number of standard statistical methods as a programming language for implementing and popularizing new ones. In addition, all subroutines are translated into R because many other people work with such software, which is a free-version of S+ developed by academic researchers.

Science finds its source in the collective knowledge which is based on exchanges, collaborations and communications. So, as with any scientific production, this book has taken many benefits from contacts we had along the last few years. We had the opportunity to collaborate with various people including A. Ait-Saidi, G. Aneiros, J. Boularan, C. Camlong, H. Cardot, V. Couallier, S. Dabo-Niang, G. Estévez, W. Gonzalez-Manteiga, L. Györfi, A. Goia, W. Härdle, J. Hart, I. Horova, R. Kassa, A. Laksaci, A. Mas, S. Montcaup, V. Nuñez-Antón, L. Pélégrina, A. Quintela del Rio, M. Rachdi, J. Rodriguez-Poo, P. Sarda, S. Sperlicht and E. Youndjé, and all of them have in some sense indirectly participated to this work. Many other statisticians including J. Antoch, D. Bosq, A. Cuevas, A. Kneip, E. Kontoghiorghes, E. Mammen, J.S. Marron, J. Ramsay and D. Tjostheim have also been useful and fruitful supports for us.

Of course, this book would not have became reality without the permanent encouragements of our colleagues in the working group STAPH in Toulouse. This group acting on functional and operatorial statistics is a source of inspiration and in this sense, A. Boudou, H. Cardot, Y. Romain, P. Sarda and S. Viguier-Pla are also indirectly involved in this monograph. We would also like to express our gratitude to the numerous participants in the activities of

STAPH, with special thanks to J. Barrientos-Marin and L. Delsol for their previous reading of this manuscript and their constructive comments.

Gérard Collomb (1950-1985) was a precursor on nonparametric statistics. His international contribution has been determinant for the development of this discipline, and this is particularly true in Toulouse. Undoubtly, his stamp is on this book and we wish to take this opportunity for honoring his memory.

Frédéric Ferraty
Philippe Vieu
Toulouse, France
January, 2006

Contents

Part I Statistical Background for Nonparametric Statistics and Functional Data

Part IV Nonparametric Methods for Dependent Functional Data

Part V Conclusions