

LA PRÁCTICA DEL ANÁLISIS DE CORRESPONDENCIAS

Michael Greenacre

Fundación BBVA, Rubes Editorial, 2008

ISBN: 978-84-96515-71-0

This book is essentially the translation to Spanish of the recent second edition of the book *Correspondence Analysis in Practice* (2007, Chapman and Hall) that was first published in 1993. The book offers an extremely well structured introduction to this technique, giving special attention to practical issues and didactical aspects. In fact, the author has dedicated more than 30 years to the research, practice and divulgation of correspondence analysis and the book clearly benefits from this, providing real examples in a number of fields, such as medical sciences, linguistics, sociology or biology.

The structure of the book is quite unique in some senses. It contains 25 chapters of eight or nine pages each. This makes the book quite accessible and didactic. A final summary with a list of key concepts is provided at the end of each chapter, which helps to understand the topics presented. The theoretical concepts are always illustrated through motivating examples and a large number of tables and graphs, with very informative captions, are provided.

First chapters (1 to 5) are dedicated to introductory concepts such as plots and distances for categorical data. The basis of correspondence analysis, with its properties and applications, is presented in chapters 6 to 15. Chapters 16 to 22 provide an introduction to some extensions and variants of the correspondence analysis, such as multiple correspondence analysis. Last chapters (23 to 25) offer some technical aspects that can be of interest, especially those regarding inference. Furthermore, a mathematical and complete description of the method is provided in annex A, but which is not at all compulsory, since a general understanding of the technique is easily obtainable from previous chapters. Finally, appendix B is generously dedicated to the computation of correspondence analysis using the statistical package R, through the use of the libraries **ca** for the method itself and **rgl** for visualizing plots in three dimensions. Here, examples used along the chapters are visited again to illustrate the use of some functions implemented by the author and to obtain some of the figures and results in

the book. Indications are also given for the implementation of correspondence analysis using Excel.

First principles of correspondence analysis theory were developed at the beginning of the XXth century. However, the works by Jean Paul Bézecri, Brigitte Escoffier and their French colleagues, during the 1960s, can be considered as the foundation of the present theory of correspondence analysis. Brigitte Escoffier-Cordier thesis (*Analyse des correspondances*, 1963), directed by J.P. Bézecri and the second volume of the book *Analyse des Données* (1973, Dunod, Paris, v. II *L'analyse des correspondances*) offer a complete vision of the pioneers works of the so called “French school” of data analysis. An active community of French researchers participated in the divulgation of this technique, such as Ludovic Lebart and Alain Morineau, by developing the software SPAD. This divulgation was mostly made in France but also at an international level. Furthermore, Michael Greenacre, who initiated his doctoral thesis with Benzécri during the 1970s, has also been an important contributor for the divulgation of correspondence analysis in the English-speaking world. Finally, correspondence analysis was introduced in Spain in the early 1980s and fundamental books in Spanish, such as *Tratamiento estadístico de datos* (Lebart, Morineau y Fénelon, Marcombo, 1984) and *Análisis factorial simple y múltiple. Objetivos, métodos e interpretación* (Escoffier y Pagès, Servicio Editorial de la Universidad del País Vasco, 1990), allowed for a large diffusion of the technique and also for placing it in the context of the multivariate methods.

My first contact with correspondence analysis took place at the end of the 1990s when I was a student of statistics at the Polytechnic University of Catalonia. At that time, Monica Bécue introduced me to these methods, more specially their application to textual data. The book *Análisis estadístico de textos* (Lebart, Salem y Bécue, Ed Milenio, 2000) offers a vision of what is called textual statistics. M Bécue encouraged me to undertake a research stay at Paris with Alain Morineau, at the nowadays disappeared CISIA, which was a little private company, but with an important research activity and vocation. During my stay, I could go into correspondence analysis in more depth and get more conscience of the endless list of applications where this method can be applied. From my posterior own experience as statistical consultant at the Statistical Service of the Autonomous University of Barcelona, I can state that correspondence analysis can be applied in a variety of real contexts, with satisfactory, and quite often surprising, results. Moreover, it is a technique that provides a rapid interpretation and comprehension of the information, sometimes not so evident, in the data. Finally, for the last four years I have been teaching multivariate statistical analysis, including correspondence analysis, at the Universitat Autònoma de Barcelona and, this experience has shown me how motivating this technique can be for intrepid students, who can discover the pleasure of obtaining exciting results from the analysis of data.

In conclusion, I strongly recommend this book to everyone interested in the analysis of categorical data. It can be particularly useful for university students or teachers, researchers and professionals for any field involved with the analysis of categorical data.

Joan Valls Marsal
joan.valls@iconcologia.net
Catalan Cancer Registry, Catalan Institute of Oncology, Barcelona
Departament de Matemàtiques, Universitat Autònoma de Barcelona