

Prólogo

Este libro es una edición revisada y ampliada de *Correspondence Analysis in Practice*, publicado por primera vez en 1993. Creo que hoy sigue siendo válido lo que escribí en el prólogo de la primera edición; era lo siguiente:

El análisis de correspondencias es una técnica estadística útil para estudiantes, investigadores y profesionales que trabajan con datos categóricos, por ejemplo, datos obtenidos en encuestas sociales. El método es especialmente eficaz para analizar las tablas de contingencia con datos de frecuencias numéricas, ya que nos proporciona una representación gráfica elegante y simple que permite una rápida interpretación y comprensión de los datos. A pesar de que los orígenes teóricos de esta técnica tienen más de 50 años, fue el matemático y lingüista francés Jean-Paul Benzécri quien, junto con sus colegas y estudiantes, dio un impulso real a las aplicaciones modernas del análisis de correspondencias, a principios de los años sesenta en la Universidad de Rennes, y posteriormente, en el campus Jussieu de la Universidad de París. En los Países Bajos y Japón, Jan de Leeuw y Chikio Hayashi también fueron pioneros en desarrollos paralelos del análisis de correspondencias. Mi implicación con el análisis de correspondencias empezó en 1973 cuando inicié mi doctorado en el laboratorio de análisis de datos de Benzécri en París. La publicación en 1984 de mi primer libro, *Theory and Applications of Correspondence Analysis*, coincidió con el inicio de una amplia difusión del análisis de correspondencias fuera de Francia. En aquel momento tenía la esperanza de que mi libro pudiera servir para que en el futuro se extendiera la aplicación práctica del análisis de correspondencias. Las evoluciones posteriores y la reciente popularidad del método no pudieron haber sido más gratificantes: centenares de investigadores se introdujeron en el mismo y se familiarizaron con la capacidad de éste para hacer comprender a los no especialistas, mediante representaciones gráficas, la información contenida en complejas tablas de datos numéricos. Los investigadores con los que entré en contacto procedían de disciplinas tan diversas como sociología, ecología, paleontología, arqueología, geología, pedagogía, medicina, bioquímica, microbiología, lingüística, investigación de mercados, publicidad, religión, filosofía, arte y música. En 1989, Jay Magidson de Statistical Innovations Inc., me invitó a colaborar con Leo Goodman y Clifford Clogg en la presentación en un curso de dos días en Nueva York titulado «*Correspondence Analysis and Association Models: Geometric Representation and Beyond*». La mayor parte de los participantes eran profesionales del *marketing* de las principales compañías norteamericanas. Preparé unas notas para este curso que reforzaban la aproximación práctica del análisis de corres-

Extracto del prólogo
de la primera edición de
*Correspondence Analysis
in Practice*

pondencias, la orientación al usuario. La reacción positiva de la audiencia fue contagiosa y me animó posteriormente a impartir cursos sobre análisis de correspondencias en Sudáfrica, Inglaterra y Alemania. Este libro es fruto de las notas preparadas para esos cursos.

Las conferencias de Colonia y Barcelona

En 1991, el profesor Walter Kristof de la Universidad de Hamburgo me propuso que organizáramos, con el apoyo de Jörg Blasius del Archivo Central para Investigación Social Aplicada (*Zentralarchiv für Empirische Sozialforschung*), de la Universidad de Colonia, una conferencia sobre análisis de correspondencias. Este evento fue la primera conferencia internacional de este tipo. Atrajo a Colonia a una gran audiencia procedente de Alemania y países europeos vecinos. Este primer encuentro evolucionó hacia una serie de conferencias cuatrianuales, que tuvieron lugar en 1995 y 1999 en Colonia, en 2003 en la Universidad Pompeu Fabra, en Barcelona, y en 2007 en la Universidad Erasmus de Rotterdam. De la conferencia de 1991 surgió la publicación del libro *Correspondence Analysis in the Social Sciences*, mientras que la de 1995 dio lugar a otro libro, *Visualisation of Categorical Data*. Ambas publicaciones recibieron excelentes críticas. En la conferencia de 1999 sobre análisis de datos a gran escala, los participantes tuvieron que presentar análisis de datos del Programa Internacional de Encuestas Sociales (*Internacional Social Survey Programme*, ISSP). Esta reunión interdisciplinaria incluyó presentaciones no sólo sobre los últimos desarrollos metodológicos del análisis de encuestas, sino también sobre temas tan diversos como la religión, el medio ambiente y la desigualdad social. En 2003, en la conferencia de Barcelona, se retomó el tema original —bautizado con nombre propio de mujer (en catalán), CARME, siglas de *Correspondence Analysis and Related Methods*—, que dio lugar a la creación de la red CARME (www.carme-n.org). Esta conferencia nos llevó a Jörg Blasius y a mí a editar un tercer libro, *Multiple Correspondence Analysis and Related Methods*, publicado por Chapman & Hall en junio de 2006. Como ocurrió con los dos libros anteriores, nuestra idea era editar una obra colectiva con participación de diversos autores. Para ello pedimos la contribución de expertos en el tema y, en nuestra tarea de editores, completamos la obra con una introducción; se relacionaron las distintas contribuciones, se unificó la notación y se compiló una lista común de referencias y el índice. De algún modo, estos libros marcan el desarrollo del tema, al menos en lo que se refiere a las ciencias sociales. Son muy recomendables para cualquier persona interesada en profundizar en el conocimiento del análisis de correspondencias y en los métodos relacionados.

Una segunda edición ampliada

Para mí fue muy gratificante reescribir *Correspondence Analysis in Practice*, 13 años después de la primera edición y después de haber acumulado mucha más experiencia en la investigación social y medioambiental. He impartido cursos en todo el mundo, incluyendo España, Italia, Bélgica, Brasil, Canadá y Australia, al igual que cursos de análisis multivariante para biólogos medioambientales en Noruega, Islandia y España, en los que el análisis de correspondencias era uno de los temas

principales. La experiencia acumulada me ha ido sugiriendo nuevas aplicaciones y renovadas ideas, imprescindibles para actualizar el contenido de la primera edición. Aparte de revisar completamente los capítulos originales y renovar algunos ejemplos, se han añadido cinco nuevos capítulos, sobre las relaciones de transición y de regresión (en los que se exponen los resultados del análisis de correspondencias y donde los datos originales están relacionados mediante funciones lineales), sobre las tablas concatenadas (acerca de cómo podemos analizar conjuntamente varias tablas de contingencia), el análisis de correspondencias de subgrupos (una simple pero efectiva variante del algoritmo del análisis de correspondencias que permite analizar partes seleccionadas de un conjunto de datos), el análisis de tablas cuadradas (en el que se aprende a descomponer las tablas cuadradas —por ejemplo las tablas de movilidad social y las matrices de *brand switching*— en partes que podemos visualizar de forma separada utilizando el análisis de correspondencias) y, por último, el análisis de correspondencias canónico (sobre cómo tener en cuenta las relaciones con variables explicativas externas, una ampliación del análisis de correspondencias muy utilizado en ecología). Después de 33 años de trabajo en este ámbito del conocimiento, puedo aventurar que esta segunda edición contiene casi todo mi saber práctico sobre el tema.

En una conferencia a la que asistí en los años ochenta, me obsequiaron con una insignia de solapa, cuya foto se reproduce en esta página. Dicha insignia contiene una bonita y ambigua máxima, que bien podría ser, en todo el mundo, el lema del análisis de correspondencias.

Comparación entre la primera y la segunda edición



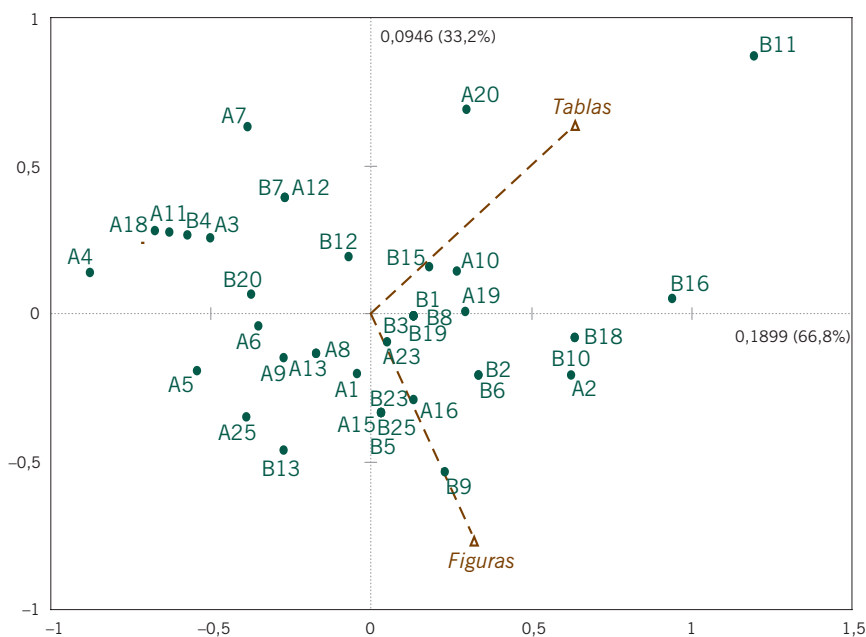
Para ilustrar el significado más obvio de este lema, y para dar un simple ejemplo de análisis de correspondencias, hice un recuento del número de tablas y figuras que contenía cada uno de los capítulos de la primera edición y lo comparé con los capítulos homólogos de la segunda. Para que la comparación entre ambas ediciones fuera válida, expresé estos resultados con relación al número total de páginas de cada capítulo. A continuación llevé a cabo una variante de la técnica llamada

análisis de correspondencias de subgrupos (descrita en el capítulo 21), que me permitió dibujar el mapa que mostramos en la imagen 1 —estas ideas se comprenderán mejor una vez leído el libro, sin embargo, por el momento consideremos este ejemplo como una especie de diagrama de dispersión—. Los dos vectores, *figuras* y *tablas*, apuntan hacia la derecha, por tanto, los capítulos situados a la derecha tienen porcentajes de figuras y tablas superiores a la media. Por ejemplo, los capítulos 9 y 11 de la segunda edición tienen los mayores porcentajes de figuras y tablas, respectivamente. Unas tres cuartas partes de los capítulos de la segunda edición que hemos considerado para el análisis, se hallan situados a la derecha del centro del mapa (el centro representa la media), mientras que solamente un cuarto de los capítulos de la primera edición se hallan situados a la derecha. Este resultado demuestra un incremento sustancial de las figuras y tablas en la segunda edición. En realidad, en ésta, han aumentado sustancialmente la información y las aplicaciones prácticas del libro. La primera edición tenía un cuerpo central (sin incluir los apéndices) de 177 páginas —en la segunda edición en inglés éste ha aumentado a 200; un incremento del 13% que ha consistido en un incremento del 25% del número de capítulos y de temas, un incremento del 18% en el número de conjuntos de datos, un incremento del 44% en el número de figuras y del 63% en el de tablas.

Formato de la segunda edición

Este libro presenta algunas innovaciones que merecen mencionarse. Es nuestra intención que esta edición tenga un enfoque todavía más didáctico que la primera: así cada capítulo representa una cantidad limitada de información con el objetivo

Imagen 1:
Mapa del análisis de correspondencias de subgrupos de los porcentajes de tablas y figuras de los 20 capítulos de la primera edición (simbolizado por A) y de sus homólogos de la segunda edición (simbolizados por B), que hemos presentado como un biplot estándar. Los números de los capítulos de la segunda edición corresponden a los de sus homólogos de la primera edición



de facilitar la lectura y su uso en la enseñanza. No había una mejor manera de conseguirlo que limitar la extensión de los capítulos, por lo que cada uno de ellos [en la versión en inglés] tiene exactamente ocho páginas. [Dicha extensión se ha visto ampliada a nueve páginas en la versión en castellano.] Era uno de los aspectos más interesantes del proyecto, que además fue evolucionando de forma natural a medida que escribía el libro. Uno de mis colegas comparó la labor didáctica con la poética: el símil de escribir sonetos de 14 versos con una rima estricta resultó ser válido en mi caso: definitivamente, el formato contribuyó al proceso creativo. Otra innovación ha sido el amplio uso de apostillas o notas en los márgenes, cuya función de encabezamiento de párrafo se suma a la de resumir al inicio de cada capítulo, los principales contenidos del mismo. También se ha optado por ubicar las leyendas de figuras y tablas en los márgenes, de manera que la información que aporten pueda ser más extensa y visual que con una clásica disposición al pie. Por último, cada capítulo finaliza con un resumen, en el que se esquematizan las ideas clave.

Igual que en la primera edición, el libro se orienta hacia la práctica del análisis de correspondencias, por lo que hemos reunido en un apéndice final los temas más técnicos, así como los aspectos más matemáticos. El apéndice teórico (A) es más extenso que el de la primera edición. Incluye teoría adicional sobre los nuevos temas del libro como, por ejemplo, el análisis de correspondencias canónico.

[Apéndice teórico \(A\)](#)

Una de las principales características de esta edición, que claramente la distingue de la original, y que está en total sintonía con la presente era digital, es la extensión del apéndice de cálculo (B), en el que utilizamos el programa R de dominio público, que se ha convertido de facto en un procedimiento de cálculo estadístico estándar. Casi todos los análisis contenidos en el libro están reflejados en este apéndice. Para ello, se presentan las instrucciones en R necesarias para obtener los correspondientes resultados. Además, se describen tres tecnologías distintas utilizadas para crear las representaciones gráficas del libro. Aunque los mapas puedan parecer simples, obtenerlos no ha sido en absoluto un ejercicio trivial.

[Apéndice de cálculo \(B\)](#)

En los 25 capítulos del libro no se facilitan referencias bibliográficas. Sin embargo, se incluye un capítulo de bibliografía, relativamente breve, para orientar al lector hacia lecturas adicionales que contienen reseñas bibliográficas mucho más extensas. El libro acaba con un glosario de los términos más importantes y con un epílogo con reflexiones finales.

[Apéndice bibliográfico,
glosario y epílogo](#)

La primera edición de este libro la escribí en Sudáfrica, la presente edición la he escrito en Cataluña, España. Mucha gente, y muchas instituciones, han contribuido, de una manera u otra, a este proyecto. En primer lugar, y de forma principal, quiero expresar mi agradecimiento a Rafael Pardo, director de la Fundación BBVA, sin cuyo apoyo, tanto desde el punto de vista humano como económico,

[Agradecimientos](#)

no hubiera podido dejar mis tareas como profesor durante seis meses y, honestamente, puedo afirmar que la segunda edición en inglés de este libro no hubiera visto la luz. La Fundación BBVA, además, hace posible la publicación de la versión en español del libro. Asimismo deseo expresar mi agradecimiento a la Universidad Pompeu Fabra, en Barcelona, en la que trabajo desde 1994, y muy especialmente a Xavier Calsamiglia, director del Departamento de Economía y Empresa, y a todo el equipo humano del departamento, por haberme dado la libertad de dedicar tanto tiempo a este proyecto.

Quiero agradecer a todos mis amigos y colegas de todo el mundo su apoyo moral e intelectual, especialmente a Zerrin Asan, Jörg Blasius, John Gower, Carles Cuadras, Trevor Hastie, Michael Browne, Victor Thiessen, Karl Jöreskog, Lesley Andres, John Aitchison, Paul Lewi, Patrick Groenen, Pieter Kroonenberg, Ludovic Lebart, Michael Friendly, Antoine de Falguerolles, Salve Dahle, Stig Falk-Petersen, Raul Primicerio, Johs Hjellbrekke, Tom Backer Johnsen, Tor Korneliussen, Ümit Senesen, Brian Monteith, Ken Reed, Gillian Heller, Antonella Curci, Gianna Mastroilli, Paola Bordandini, Walter Zucchini, Oleg Nenadić, Thierry Fahmy, Tamara Djermanovic, Volker Hooyberg, Gurdeep Stephens, Rita Lugli y Danilo Guitoli, así como a toda la comunidad de Gréixer —¡todos habéis tenido un papel en esta historia!—. Un agradecimiento especial para Jörg Blasius por una minuciosa corrección del manuscrito original [en inglés] y a Oleg Nenadić por su colaboración en la preparación del paquete **ca** de R. Igual que en la primera edición, he dedicado este libro a mis tres hijas, que nunca dejan de asombrarme por su alegría, sentido del humor y creatividad. También quiero agradecer al editor Rob Calver de Chapman & Hall/CRC Press su confianza y su constante cooperación, lo que ha permitido que la segunda edición en inglés sea una realidad.

Agradecimientos a esta edición

Con respecto a la edición que el lector tiene en sus manos, deseo agradecer la colaboración de Jordi Comas en la traducción de la obra y la experta revisión realizada por Carles Cuadras. Asimismo quiero dar las gracias al equipo humano de Rubes Editorial, en especial a Imma Rullo, Núria Gibert y Jaume Estruch. Las aportaciones de todos ellos han contribuido a dar cuerpo y alma a mi libro.

Por último, deseo expresar, una vez más, mi gratitud al profesor Rafael Pardo, y a Cathrin Scupin, directora editorial de la Fundación BBVA, por el inestimable respaldo que dicha Fundación me ha brindado. Su apoyo convierte este proyecto en una realidad al alcance del público hispanoparlante.

Michael Greenacre
Barcelona, junio de 2008