

M.^a ANGELES CEA D'ANCONA, *Análisis Multivariable. Teoría y práctica en la investigación social*, Madrid: Síntesis, 2002.

Resulta sorprendente la escasa atención que ha prestado la sociología española al análisis multivariable. Aunque existe numerosa literatura en castellano sobre el *análisis de datos* realizada desde la sociología, estas referencias son prácticamente inexistentes en el ámbito del análisis multivariable. Esta situación es radicalmente diferente en otras disciplinas cercanas a la sociología como la psicología, la pedagogía y la investigación comercial; y de hecho es frecuente que los sociólogos citen —en las bibliografías de sus investigaciones— obras realizadas por psicólogos e investigadores comerciales. Basta realizar un breve recorrido por las principales revistas de sociología para comprobar que los textos de Bisquerra (1989), Hair et al (1998), y Luque (2000), entre otros, son muy conocidos —y utilizados— por los sociólogos españoles.

Frente a esta situación, la literatura especializada en otros idiomas se caracteriza por la gran atención prestada —desde la sociología— al análisis de datos y las técnicas de análisis multivariante. Como muestra basta con destacar los volúmenes sobre técnicas multivariantes publicados dentro de la clásica colección *Quantitative Applications in the Social Sciences* de la editorial Sage, y la reciente aparición de una nueva colección dedicada específicamente a las técnicas multivariantes (*Advanced Quantitative Techniques in the Social Sciences*).

En el Congreso Español de Sociología realizado en A Coruña, el profesor Alvira (1998) señalaba la escasa utilización del análisis multivariante por los sociólogos españoles como una explicación de este escaso desarrollo. Una revisión de los artículos publicados en las principales revistas españolas de sociología le lleva a destacar la exigua utilización del análisis multivariante en la sociología española. En el IV Congreso Vasco de Sociología,

realizado en Bilbao a principios de 1998, Oleaga (1998: 544-549) llega a una conclusión similar.

Esta baja utilización de técnicas multivariantes es aún más sorprendente cuando se constatan las mejoras introducidas en los paquetes de análisis estadístico, que en las últimas versiones han facilitado notablemente su uso y cuentan con *ayudas y asesores estadísticos* que explican las técnicas de análisis más complicadas.

Esta es la «situación de partida» que debemos considerar a la hora de analizar el texto de la profesora Cea D'Ancona, si bien el libro supera ampliamente las expectativas creadas al observar el escaso desarrollo del Análisis Multivariante en la investigación social española. Lo primero que llama la atención es la enorme voluminosidad del texto, 638 páginas, en un momento en que los editores apuestan por libros más delgados que presentan una mejor venta. El libro está dividido en seis capítulos, cada uno dedicado a una técnica específica: la *Regresión Múltiple* es explicada en el primer capítulo, el *Análisis Logit* en el segundo, el tercero se dedica al *Análisis de Conglomerados*, el cuarto al *Análisis Discriminante*, el *Análisis Factorial* es explicado en el quinto, y las *Ecuaciones Estructurales* en el sexto —y último— capítulo.

En la medida que cada capítulo está dedicado a una técnica concreta, y la estructura de todos ellos es similar, hablar de la organización de los capítulos implica presentar la estructura expositiva de la obra. Cada capítulo comienza con un apartado donde, además de exponer los rasgos más importantes de cada técnica y sus características más relevantes, se explican sus orígenes y la relación con otras técnicas multivariantes. A continuación se presentan los *supuestos básicos* que deben cumplir los datos antes de aplicar cada técnica. La exposición de estos

supuestos se realiza con un gran nivel de detalle, mostrando ejemplos con datos reales —utilizando una base de datos del CIS— que ayudan notablemente a una mejor comprensión de cada uno.

Desde mi punto de vista se trata de una de las obras que mejor explica este aspecto, que no olvidemos es una condición básica a la hora de utilizar adecuadamente el análisis multivariante. Este hecho presenta una importancia capital en un momento como el actual en el que el gran desarrollo de los paquetes estadísticos, unido a la facilidad de su uso, ha generado una reducción del tiempo dedicado al análisis exploratorio de los datos, aumentando el tiempo empleado en «jugar» con el ordenador «para ver que sale». Un cita de Benjamín González ilustra magistralmente este fenómeno: «se corre el riesgo de eliminar las primeras fases del proyecto de investigación al considerar que los datos proporcionan información por sí mismos: se obvia el análisis del problema y el planteamiento teórico de la investigación, y ésta se convierte en un «cruzar todo con todo para ver que sale» (González, 1994: 257).

Comprobados los *supuestos* se procede con la utilización de cada técnica, para lo cual se muestran unos diagramas que, bajo el título «fases principales de aplicación», presentan gráficamente de las etapas a considerar a la hora de utilizar cada técnica. Los siguientes apartados de cada capítulo siguen el orden expuesto en estos diagramas, que básicamente comprenden cuatro etapas: 1) preliminares del análisis; 2) obtención de resultados; 3) evaluación-ajuste del modelo; y 4) presentación de los datos e interpretación. Analizaremos en detalle cada uno.

La etapa de los preliminares del análisis, definida en algunos capítulos como preparación de los datos para el análisis, es donde se lleva a cabo una primera indagación exploratoria y —cuando procede— se realiza una depuración de los datos; paso previo y fundamental antes de llevar a cabo cualquier análisis multivariante.

En algunos capítulos se utiliza esta primera etapa para explicar los métodos de conglomeración y la elección de variables (capítulo 3), y la especificación del modelo causal y su identificación (capítulo 6).

En la segunda etapa, dedicada a la obtención de resultados, se realiza una explicación de las pruebas y test estadísticos utilizados para la interpretación de cada técnica: el significado de la ecuación de regresión en el caso de la regresión múltiple y logit, la estimación de las funciones discriminantes canónicas en el caso del análisis discriminante, la extracción de factores comunes o componentes principales en el análisis factorial, etc. Esta explicación se lleva a cabo en dos fases: en primer lugar se define y explica cada prueba y test estadístico, para después realizar una interpretación aplicada a un ejemplo concreto.

A continuación se procede con la comprobación de la adecuación del modelo, momento en el que se evalúa el ajuste del modelo y su significatividad (tanto estadística como interpretativa). El ajuste del modelo da paso a la presentación e interpretación de los resultados obtenidos; procediendo de forma similar a la segunda etapa: en primer lugar se definen y explican los resultados obtenidos, para realizar después la explicación utilizando un ejemplo real.

Cada capítulo finaliza con varias lecturas complementarias y una serie de ejercicios a realizar. Considero muy acertada esta forma de terminar, proponiendo la realización de estos ejercicios, en la medida que éstos se han tomado de investigaciones reales realizadas por sociólogos y que han sido publicadas en revistas españolas; lo que permite *consultar* las soluciones acudiendo a las fuentes originales. La gran ventaja de estos ejercicios —a mi juicio— es que muestran aplicaciones concretas de las distintas técnicas explicadas, al tiempo que ilustran cómo se presentan e interpretan los resultados de las técnicas multivariantes en investigaciones concretas.

Este rápido recorrido por la estructura de cada capítulo desvela la exhaustividad del libro; aunque desde nuestro punto de vista los aspectos más positivos de esta obra están relacionados no tanto con la exhaustividad temática —puesto que no son explicadas todas las técnicas multivariantes— sino con la organización interna de cada capítulo.

Definida la estructura de cada capítulo, y considerando que ésta se repite capítulo tras capítulo, pasaremos a analizar los aspectos positivos de esta obra. Dentro de lo difícil que me resulta realizar una reseña de un libro de más de 600 páginas en unos pocos folios, si tuviera que resaltar únicamente algunos aspectos de la obra hay varios elementos que creo que sorprenden enormemente. Además de la detallada explicación de los *supuestos* que deben cumplir los datos a la hora de utilizar una determinada técnica multivariante, ya destacado unos párrafos más arriba, es preciso recalcar la gran profundidad de las explicaciones realizadas y la sucinta utilización de fórmulas matemáticas que tanto asustan a los estudiantes de sociología. Es imposible escribir un libro sobre análisis multivariante sin utilizar fórmulas, pero la autora —sin duda conocedora del bajo conocimiento matemático de los estudiantes de sociología— ha decidido utilizar las imprescindibles con el fin de no *asustar* al lector menos conocedor de esta materia. Digamos pues que es un libro de iniciación, pero que consigue que el lector llegue a dominar las técnicas multivariantes que en él se presentan.

Es de destacar también el empleo de ejemplos reales, utilizando una base de datos del CIS (n.º 2214) que está disponible para los lectores, y que por lo tanto permite la *replicación* de los ejemplos realizados. Sin embargo uno de los aspectos más interesantes es, desde mi punto de vista, que la exposición de cada técnica sigue la estructura de una investigación real; lo que implica una situación muy cercana a la práctica de investigación que

se realiza diariamente en cualquier instituto de investigación. Esta forma de proceder permite solventar una crítica hacia la enseñanza de la investigación, esgrimida recientemente por Jiménez Blanco en el número 96 de la REIS. En la presentación de un *texto clásico* Jiménez Blanco señalaba que «se enseña investigación social *investigando*, es decir, participando en una investigación real. O sea, lo contrario de lo que se hace en nuestras Facultades, en que se enseñan “técnicas” de investigación social casi como ejercicio retórico, en que los ejemplos, cuando se ponen, proceden de investigaciones suecas o inglesas: lo más apartado de la realidad empírica que se pueda» (2001: 228). Este libro constituye un excelente ejemplo de cómo superar esta situación.

Ya hemos señalado la importancia de los ejercicios propuestos al final de cada capítulo, extraídos de investigaciones llevadas a cabo en nuestro país y que permiten consultar cómo otros investigadores presentan los resultados obtenidos al aplicar técnicas multivariantes, algo complicado si se tiene en cuenta el escaso conocimiento de estas técnicas que tienen los propios demandantes de la investigación. De hecho, considero que uno de los factores que explican la escasa utilización de técnicas multivariantes es la dificultad de explicar los resultados a las personas que encargan las investigaciones. La consulta de ejemplos realizados por otros colegas contribuirá sin duda a una mayor utilización de estas técnicas.

Otro aspecto interesante son las numerosas referencias a los resultados y test estadísticos empleados por los paquetes estadísticos más difundidos y utilizados, como el AMOS, LIMDEP, LISREL, SPSS, BMDP y SAS.

Finalizaremos el comentario de esta obra destacando los aspectos que menos nos han gustado; en la medida que una *crítica de un libro* no es tal si no se destacan los aspectos positivos y negativos del mismo. Desde mi punto de vista se echa de menos un capítulo introductorio con

una breve explicación del análisis multivariante, que podría servir para contextualizar la obra e iniciar, al lector menos especializado, en el conocimiento de cada una. De esta forma proceden Bisquerra (1989), Lozares y López (1991), Hair et al (1998), y Luque (2000), y considero que supone una estúpida *toma de contacto* para el lector no especializado. Este capítulo podría utilizarse también para explicar la estructura de la obra y exponer las razones por las que se han elegido unas técnicas y no otras. ¿Se han elegido las más importantes para la investigación social?; ¿aquellas que son más utilizadas?; ¿las que han recibido menos atención?; ¿las que tienen menos *supuestos restrictivos*?, etc.

Suponemos que el elevado tamaño del libro —más de 600 páginas— y los requisitos de las editoriales (que prefieren obras cortas) son los responsables de la falta de otras técnicas de análisis multivariante (análisis de correspondencias, escalamiento multidimensional, etc); si bien consideramos más acertado explicar seis técnicas con gran profundidad, que haber explicado todas con un menor nivel de detalle. De modo que aunque se trata de un libro dedicado a lectores que se inician en el tema, es muy adecuado también para aquellos con un nivel de experiencia medio. En definitiva, un libro FUNDAMENTAL para todos los que realizan investigación social.

REFERENCIAS BIBLIOGRÁFICAS

- ALVIRA, F. (1998): «Ponencia de presentación de la sesión sobre *Técnicas y Análisis Cuantitativos*». *Congreso Español de Sociología*, La Coruña, 24-26 de septiembre.
- BISQUERRA, R. (1989): *Introducción conceptual al Análisis Multivariable*. Barcelona: PPU.
- DÍAZ DE RADA, V. (2002): *Técnicas de Análisis Multivariante para Investigación Social y Comercial*. Madrid: Ra-Ma.
- GONZÁLEZ RODRÍGUEZ, B. (1994): «La utilización de los datos disponibles», en GARCÍA FERRANDO, M.; IBÁÑEZ, J., y ALVIRA, F. (Eds.), *El Análisis de la Realidad Social*. Madrid: Alianza, 2.ª ed., pp. 245-272.
- HAIR, J. F.; ANDERSON, R. E.; TATHAM, R. L., y BLACK, W. C. (1998): *Multivariate data analysis* (5.ª edición). New Jersey: Prentice Hall. Hay traducción española de esta edición publicada por la editorial Prentice Hall.
- LOZARES, C., y LÓPEZ ROLDÁN, P. (1991): «El análisis multivariado: definición, criterios y clasificación», en *Papers*, vol. 37, pp. 9-29.
- LUQUE MARTÍNEZ, T. (Coordinador) (2000): *Técnicas de Análisis de Datos en Investigación de Mercados*. Madrid: Pirámide.
- JIMÉNEZ BLANCO, J. (2001): «Presentación de Texto Clásico», *Revista Española de Investigaciones Sociológicas*, n.º 96, pp. 227-233.
- OLEAGA, P. (1998): «El divorcio entre metodología y práctica, o el ostracismo académico de las técnicas multivariantes de análisis de datos», *Actas del IV Congreso Vasco de Sociología*, vol. II, pp. 544-549.
- PAPERS, revista (1991): *Número monográfico sobre análisis multivariante*, vol. 37.

VIDAL DÍAZ DE RADA