

Aplicación de la investigación de mercados al análisis de problemas de marketing

Autor: Miguel Santesmases Mestre

Catedrático de Comercialización e Investigación de Mercados (Emérito)
Universidad de Alcalá (Madrid)

Resumen

La concepción actual del marketing parte del conocimiento de las necesidades y del comportamiento del consumidor con el fin de establecer relaciones estables, que beneficien tanto al comprador como al vendedor. Para obtener este conocimiento es preciso utilizar los métodos y técnicas de diseño de encuestas y análisis de datos que proporciona la investigación de mercados.

Este artículo pretende exponer los métodos que deben llevarse a cabo para diseñar encuestas y otros estudios empíricos y cómo pueden ser aplicadas las distintas técnicas de análisis de datos que incluye la investigación de mercados. Todo ello con el fin último de obtener un conocimiento del consumidor/comprador que permita llevar a cabo un auténtico marketing de relaciones.

Palabras clave: investigación de mercados, técnicas de análisis de datos, análisis multivariable, marketing, consumidor.

I. Concepción actual del marketing a partir del conocimiento del consumidor/comprador

El marketing parte de las necesidades y deseos del consumidor y tiene como fin su satisfacción del modo más beneficioso, tanto para el comprador o consumidor como para el vendedor. Para llevar a cabo esta relación de intercambio, el marketing diseña estrategias y utiliza un conjunto de métodos y técnicas que permiten identificar las necesidades del consumidor y desarrollar la demanda para los productos que, diseñados de acuerdo con tales necesidades, se ofrecen al mercado.

Para el análisis del mercado y evaluación de estrategias, la dirección de marketing aplica métodos y utiliza técnicas de análisis de la demanda y del comportamiento del consumidor, así como de segmentación de mercados, posicionamiento y, en general, de la investigación de mercados.

El análisis de la demanda implica las tareas de medir, explicar y pronosticar. Medir supone cuantificar el alcance de la demanda, tanto actual como potencial. Explicar la demanda requiere identificar las variables que la determinan y averiguar cómo influyen en su comportamiento. Finalmente, pronosticar supone efectuar una previsión del nivel de demanda en el futuro, partiendo del comportamiento actual y pasado de la demanda y de las variables que influyen en dicho comportamiento.

El estudio del comportamiento del consumidor se deriva del concepto actual del marketing. Dicho estudio constituye el punto de partida para la aplicación de un auténtico marketing de relaciones, puesto que difícilmente se podrán satisfacer de forma efectiva las necesidades del consumidor sin un conocimiento previo de los bienes y servicios que éste desea y de las actividades que el consumidor lleva a cabo para adquirirlos. El estudio del comportamiento del consumidor se refiere, por tanto, al conjunto de actividades que lleva a cabo una persona o una organización desde que tiene una necesidad hasta el momento que efectúa la compra y usa, posteriormente, el producto. El estudio de tal comportamiento incluye también el análisis de todos los factores que influyen en el origen, desarrollo y ejecución de las acciones llevadas a cabo por el consumidor.

Aunque se hace referencia a los consumidores, en general, como si fueran homogéneos, en realidad no lo son. Los consumidores tienen características o atributos diversos, muchos tipos de necesidades diferentes y cuando compran un producto no buscan siempre los mismos beneficios. En consecuencia, no se puede considerar el mercado como una unidad y ofertar a todos sus integrantes los mismos productos o servicios. Si se quieren satisfacer realmente las necesidades del mercado e incrementar la demanda, debe proponerse una oferta diferenciada a cada uno de los grupos de consumidores que presentan características o necesidades distintas. La segmentación pone de manifiesto precisamente estas diferencias en las características y necesidades de los consumidores y permite detectar cuáles de ellas son relevantes, con el fin de llevar a cabo estrategias de marketing diferenciadas que permitan satisfacer de forma más efectiva las necesidades de los distintos grupos de consumidores.

En definitiva, para diseñar las estrategias de marketing es necesario conocer y comprender el mercado. Si se quiere estimular y captar la demanda, es preciso saber cuál es el mercado potencial, a qué segmentos del mismo hay que dirigirse, qué productos necesitan y desean estos mercados objetivo, por qué los compran y cómo, dónde y cuándo los compran.

Conocer estos aspectos del consumidor supone conseguir datos sobre el mismo. Sin embargo, los datos por sí mismos no son información. Los datos deben ser analizados adecuadamente para poder extraer de ellos todo su contenido informativo. La investigación de mercados aporta los métodos y técnicas adecuados para la captura y el análisis de los datos. La investigación de mercados puede proporcionar los estudios necesarios que permitan elaborar estrategias de marketing efectivas.

La investigación de mercados puede ser definida de forma breve y concisa como “la búsqueda y análisis sistemático y objetivo de la información relevante para la identificación y solución de cualquier problema en el campo del marketing” (Green, Tull y Albaum, 1988). De forma más extensa, la American Marketing Association (AMA) aprobó en 1987, y ratificó en 2004, su última definición de la investigación de mercados, que supone una enumeración completa de sus funciones y de las etapas seguidas en su desarrollo. Así, la investigación de mercados “es la función que enlaza al consumidor, cliente y público en general con el comercializador (*marketer*) a través de la información -información que se usa para identificar y definir oportunidades y problemas de marketing; genera, refina y evalúa acciones de marketing; pone de manifiesto el rendimiento del marketing y mejora la comprensión del marketing como un proceso. La investigación de mercados especifica la información requerida para abordar estas cuestiones; diseña el método para recoger la información; dirige y lleva a cabo el proceso de recogida; analiza los resultados y comunica sus hallazgos y sus implicaciones.”

II. Métodos y técnicas de la investigación de mercados para el diseño de encuestas y otros estudios empíricos

En toda investigación hay que establecer un diseño o plan que sirva de guía para la recogida y análisis de los datos (Churchill, 1987, 2001). El diseño de la investigación implica la especificación de los métodos y procedimientos para adquirir la información necesaria para estructurar o solucionar problemas (Green, Tull y Albaum, 1988). El plan de investigación comprende un conjunto de actividades que parten de la identificación del problema que se va a investigar, pasan por la definición, clasificación y medida de las variables y su plasmación en un formulario o cuestionario y finalizan con el análisis de los datos y la interpretación de los resultados. En concreto, incluye las etapas siguientes:

1. Identificación del problema a investigar.

2. Definición de los objetivos de la investigación.
3. Determinación del tipo de diseño de investigación.
4. Especificación de las hipótesis del estudio.
5. Definición, clasificación y medida de las variables del estudio.
6. Selección de las fuentes de información.
7. Diseño del cuestionario o formulario para la captura de los datos.
8. Diseño y selección de la muestra (si procede).
9. Obtención y tratamiento de los datos.
10. Análisis de los datos e interpretación de los resultados.
11. Elaboración del informe de investigación.

La identificación del problema supone partir de una definición clara y precisa del problema que se va a estudiar (“un problema bien definido es un problema medio resuelto”). Para identificar mejor el problema a investigar es necesario revisar previamente los estudios de mercado y científicos que se han realizado sobre el tema de investigación o análogos.

La definición de los objetivos de la investigación implica concretar y enumerar los fenómenos que se quieren investigar y las relaciones que puedan existir entre ellos, para así poder especificar posteriormente las hipótesis que se quieren contrastar (Zikmund, 1998).

La determinación del tipo de diseño de investigación se deriva de la clase de problema a estudiar y de los fines de la investigación. Según su diseño, los estudios pueden ser exploratorios, descriptivos y experimentales. El exploratorio es un estudio preliminar, muy flexible y poco formal, que se basa en el estudio de datos ya existentes, en entrevistas con personas expertas y en el examen de situaciones análogas, mediante casos de estudio y simulaciones. Este tipo de estudios ayuda al investigador a formular hipótesis o a clarificar conceptos y le permiten familiarizarse con el problema (Churchill, 2001). El estudio descriptivo o correlacional es el más usual en investigación de mercados. Tiene como finalidad describir las características de ciertos grupos, determinar la frecuencia con que ocurre algo, estimar la relación entre dos o más variables o efectuar predicciones. Los estudios descriptivos son más formales que los exploratorios y establecen modelos basados en hipótesis. En este tipo de estudios, la encuesta constituye el método más habitual de obtener información. Finalmente, los estudios experimentales son idóneos para contrastar hipótesis y establecer relaciones de causa-efecto, por el control que proporcionan al investigador.

Las hipótesis son una afirmación o proposición no probada sobre un fenómeno, el comportamiento de una o más variables o la relación entre dos o más de ellas. Las hipótesis ponen de manifiesto lo que se está buscando y anticipan las respuestas posibles a las cuestiones planteadas en la investigación. Las hipótesis tienen que ser proposiciones únicas (no múltiples) sobre los resultados esperados de la investigación.

Tienen que estar formuladas de modo afirmativo, no como pregunta con interrogantes. Es una hipótesis, por ejemplo, afirmar que el hábito de fumar aumenta el riesgo de accidente laboral, o que la edad influye positivamente sobre el comportamiento de compra de un determinado producto. Los resultados del estudio tendrán que confirmar o rechazar las hipótesis. Las hipótesis pueden derivarse de la teoría, de investigaciones exploratorias previas o de la experiencia del investigador (Aaker y Day, 1980).

La definición de las variables permite hacer operativos y susceptibles de medida los problemas básicos a investigar (Simon, 1969). Una variable es una magnitud cuyos valores son objeto de estudio. Puede referirse a un individuo, grupo de personas u organización. En función del fenómeno que miden y sin ánimo de ser exhaustivos, las variables pueden clasificarse en comportamientos, atributos, percepciones, preferencias, valores, estilos de vida, opiniones, actitudes, motivaciones, deseos y necesidades. Según la relación causa-efecto analizada, las variables pueden dividirse, de acuerdo con el papel que desempeñen en la relación causal, en dependientes e independientes. En función de los valores que pueden adoptar, las variables se clasifican en continuas, discretas, dicotómicas, binarias, ficticias y tipificadas. La medida de las variables viene determinada por la escala utilizada. El tipo de escala, a su vez, condiciona la elección de la técnica estadística de análisis de los datos. Según la clasificación de Stevens (1946), pueden distinguirse cuatro tipos de escalas: nominales (indican sólo la pertenencia a una clase o categoría), ordinales (implican una ordenación de las categorías de una variable nominal), interválicas (además de ordenación, las distancias entre los valores o intervalos de la escala son iguales, aunque no existe proporcionalidad entre ellos ni el valor 0 de la escala es absoluto) y proporcionales o de razón (además de ordenación e intervalos iguales, existe proporcionalidad entre los valores de la escala y el valor 0 es absoluto).

Las fuentes de información en investigación de mercados son las personas, organización u objetos de los que se obtienen los datos para ser analizados. Las fuentes de información, por su propósito, pueden ser primarias o secundarias, y por su origen, internas o externas (Luque, 1997). Un dato primario es aquel que se obtiene de modo específico para la investigación a efectuar. Un dato secundario es aquel que ya está disponible, que se ha obtenido en estudios anteriores y que sirve subsidiariamente para el propósito de la investigación a realizar. Son internos los datos obtenidos dentro de la propia organización o entidad que desarrolla la investigación. Los datos externos pueden ser públicos o privados (Pedret, Sagnier y Camp, 2000).

El cuestionario es el formulario que contiene las preguntas o variables de la investigación y en el que se registran las respuestas de los encuestados. El diseño del cuestionario no es sencillo y presenta ciertas dificultades. Si bien preguntar es relativamente fácil, hacer buenas preguntas es un arte que requiere imaginación y experiencia (Simon, 1969). Para un diseño apropiado del cuestionario es fundamental

definir correctamente el problema a investigar, formular las hipótesis de forma precisa y especificar adecuadamente las variables y las escalas de medida. El cuestionario debe ser lo más breve posible, pero no puede establecerse una longitud concreta; varía en función del interés que el tema de la investigación tenga para el encuestado (Roscoe, Lang y Sheth, 1975; Bean y Roszkowski, 1995). En el cuestionario pueden formularse preguntas abiertas, sin indicación de posibles respuestas, o preguntas cerradas, con una relación exhaustiva de las respuestas posibles, para que el encuestado elija una o varias de ellas.

Cuando una encuesta se realiza a toda la población o universo de estudio, es un estudio censal. Por el contrario cuando se dirige sólo a una parte de la población, se trata de un estudio muestral. Una muestra es un conjunto de elementos de una población o universo del que se quiere obtener información. A efectos de que la información obtenida de una muestra sea válida, ésta debe ser representativa de la población; es decir, que en su estructura se reproduzcan exactamente las características y comportamientos de la población de la que ha sido obtenida. Aunque la precisión o exactitud de los datos obtenidos a través de una muestra es menor que en un estudio censal, las ventajas de coste y tiempo superan con creces tal inconveniente. El proceso de muestreo implica definir la población objeto de estudio, seleccionar la estructura de la muestra (listas, directorios, etc.), especificar la unidad muestral, seleccionar el método de muestreo (probabilístico o no probabilístico), determinar el tamaño de la muestra, diseñar el plan de muestreo y, por último, seleccionar la muestra. La unidad muestral es el elemento de la población del cual se obtienen los datos. Pueden ser individuos, hogares, tiendas, empresas u objetos (productos, marcas, modelos, etc.). En un muestreo probabilístico todos los elementos de la población tienen igual oportunidad de ser seleccionados para componer la muestra. En un muestreo no probabilístico, en cambio, la selección de los elementos de la muestra se realiza, total o en parte, según criterios fijados por el investigador. La determinación del tamaño de la muestra está en función del presupuesto y tiempo disponible, así como del error que se admita en los datos obtenidos.

Dentro del proceso de obtención de datos, el conjunto de actividades realizadas para la recogida efectiva de los mismos constituye el trabajo de campo. En él se incluyen también la supervisión de los cuestionarios y el control de los errores de falta de respuesta. El tratamiento de los datos es la fase del proceso de investigación posterior a la obtención de los datos. Por lo general, supone un tratamiento informático de los datos, que implica la grabación de los mismos y el almacenamiento en ficheros o bases de datos para su posterior manipulación, tabulación y análisis (Luque, 1997). Tiene como finalidad preparar los datos para que, con el análisis posterior de los mismos, puedan convertirse en información útil para la toma de decisiones. En concreto, incluye las actividades de edición, codificación y grabación de los datos.

III. Técnicas de análisis de datos para detectar y solucionar problemas de marketing

La finalidad del análisis de los datos es su transformación en información relevante para la toma de decisiones. La información es relevante si sirve para identificar y solucionar los problemas objeto de investigación. Esto implica que la información debe contribuir a reducir la incertidumbre, ha de ser susceptible de influir en las decisiones y ha de justificar su coste (Gandz y Whipple, 1977), aunque esto último no implica que su rentabilidad deba medirse únicamente en términos económicos.

La aplicación de técnicas estadísticas de análisis de datos, especialmente las más sofisticadas, ha tenido en los últimos años un crecimiento muy importante en la investigación de mercados, especialmente desde la mayor disponibilidad y abaratamiento de los medios electrónicos de cálculo y el desarrollo de paquetes de programas estadísticos. Esta capacidad creciente de obtener y procesar información que permiten las nuevas tecnologías está produciendo cambios importantes en el papel de la investigación de mercados (Perreault, 1992; Luque, 1997 y 2000).

En función del número de variables analizadas simultáneamente, las técnicas de análisis de datos pueden clasificarse en univariantes, bivariantes y multivariantes, según que analicen, respectivamente, una sola variable, la relación o dependencia entre dos variables y la relación o interdependencia entre más de dos variables.

El análisis estadístico multivariable es el más completo y permite un mejor acercamiento a la realidad de los fenómenos económicos y sociales. Tales fenómenos suelen ser complejos y no se limitan a un número reducido de dimensiones, sino que, por el contrario, son multidimensionales, y su estudio precisa de herramientas adecuadas que no supongan una simplificación de la realidad. Ésta es la principal contribución del análisis multivariable a la investigación de mercados.

Las técnicas que se incluyen en el análisis multivariable suelen clasificarse en dos grandes grupos, según analicen relaciones de dependencia o de interdependencia entre variables u objetos (Morrison, 1976; Sheth, 1977; Green, 1978; Cuadras, 1981; Berenson, Levine y Goldstein, 1983; Luque, 2000). En el primer caso, el análisis multivariable permite estudiar la medida o intensidad con que se da la dependencia entre las variables y determina, por lo general, una relación funcional. En el caso de interdependencias, el análisis multivariable estudia las interrelaciones entre todas las variables o agrupa los sujetos u objetos por su similitud, de acuerdo con los valores mostrados por las variables analizadas. En los apartados siguientes se desarrolla cada uno de estos tipos de técnicas.

III.1. Análisis de relaciones de dependencia (relaciones causa-efecto)

Una dependencia es una relación o asociación entre dos o más variables en la que el comportamiento o el valor de una o más de ellas (variables dependientes, criterio o a explicar) es el efecto (depende) del valor de otra u otras variables causales

(independientes, predictoras o explicativas). Una relación de dependencia es, por ejemplo, la que se da entre el número de periódicos que se lee diariamente (variable a explicar) y la edad, sexo y nivel de estudios del individuo (variables explicativas). El comportamiento de la primera variable puede estar en función o “depender” (por esto se la denomina variable dependiente) de las otras variables consideradas explicativas (identificadas como variables independientes).

Las relaciones de dependencia entre variables pueden referirse a sólo dos grupos de variables (uno de variables dependientes y otro de independientes) o a un mayor número de grupos, en los que las variables que en un momento dado son dependientes de otras pueden, a su vez, actuar posteriormente como variables independientes de otra u otras variables. El primer caso hace referencia a las técnicas clásicas de análisis multivariable, mientras que en el segundo se incluyen los modelos causales de ecuaciones estructurales

III.1.1. Técnicas estadísticas multivariantes de análisis de dependencias entre dos grupos de variables

Estas técnicas se subdividen, a su vez, en dos tipos:

- a. Técnicas que analizan la relación de dependencia entre una sola variable dependiente o a explicar y múltiples independientes o explicativas.
- b. Técnicas que analizan la relación de dependencia entre múltiples variables dependientes y múltiples variables independientes.

Dentro del primer grupo se incluyen las técnicas siguientes:

- Análisis de regresión múltiple.
- Análisis de regresión logística binaria.
- Análisis de clasificación múltiple.
- AID (Automatic Interaction Detection).
- CHAID (Chi Square Automatic Interaction Detection).
- Análisis conjunto categórico.
- Análisis conjunto ordinal (monanova).

El análisis de regresión múltiple es la más conocida de las técnicas de análisis de dependencias. En ella se establece una relación funcional entre una variable dependiente Y y una serie de variables independientes $X_1, X_2, X_3, \dots, X_n$, que se expresa formalmente con la ecuación: $Y = a + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \dots + \beta_n X_n + e$, en la que los coeficientes $a, \beta_1, \beta_2, \dots, \beta_n$ son parámetros a estimar y se denominan coeficientes de regresión. El parámetro a representa el valor que toma la variable dependiente en el supuesto de que el valor de las variables independientes sea igual a cero. Los demás coeficientes determinan el efecto que tienen las variaciones de las varia-

bles independientes en el comportamiento de la variable dependiente. El término e indica el error cometido en la estimación de la variable dependiente (Johnston, 1975). Para estimar los coeficientes se utiliza el método de mínimos cuadrados, por el que se minimiza la suma de cuadrados de las diferencias entre los valores reales y estimados de la variable dependiente. El modelo de regresión requiere que las variables sean métricas, medidas con escalas proporcionales o interválicas. No obstante, pueden utilizarse variables no métricas, medidas con escalas nominales u ordinales, tanto en la variable dependiente como en las independientes, siempre que se transformen en variables ficticias, que toman el valor 1 para indicar la pertenencia a una clase o categoría, o el valor 0 en caso contrario.

La regresión logística binaria, o *logit* binario, es un tipo de análisis de regresión en el que la variable dependiente toma sólo dos valores (1 ó 0), representando, respectivamente, la ocurrencia o no de un fenómeno, la pertenencia o no a un grupo o la posesión o no de un determinado atributo. Las variables independientes han de ser métricas, bien numéricas o categóricas transformadas en ficticias. En este tipo de regresión se ajusta una función logística (que tiene forma de S) a los datos reales mediante el método de máxima verosimilitud, en lugar del método de mínimos cuadrados que se utiliza en regresión lineal.

El análisis de clasificación múltiple (ACM) es una técnica diseñada por Andrews, Morgan y Sonquist (1967, 1973), que es poco conocida, a pesar de sus notables ventajas, tanto en el manejo de los datos, como en la interpretación de los resultados (Santesmases, 1982). El ACM analiza la relación entre una variable dependiente o criterio y varias variables explicativas o predictoras. La primera debe estar medida en una escala métrica o dicotómica, mientras que las variables explicativas o independientes -y ésta es una de las características distintivas del ACM- deben estar medidas con escalas no métricas o categóricas (nominales u ordinales), o convertidas en ellas, si la escala original es de orden superior. El ACM sustituye con ventaja al análisis de regresión múltiple con variables ficticias, por la complejidad que puede suponer en esta última técnica la transformación de las variables categóricas en ficticias y la interpretación de los coeficientes de regresión obtenidos.

El análisis AID, acrónimo de *Automatic Interaction Detection* (Detección automática de interacciones), además de estudiar la relación de dependencia entre una variable dependiente y múltiples independientes o explicativas, detecta el efecto y las interacciones existentes entre las variables explicativas, como su nombre indica (Morgan y Sonquist, 1963; Sonquist, 1970; Sonquist, Baker y Morgan, 1973). No proporciona, sin embargo, una función con coeficientes que determinen la relación existente entre la variable dependiente y las independientes, pero puede utilizarse juntamente con otras técnicas estadísticas multivariantes, como la regresión lineal, para completar el análisis y estimar una relación funcional. La principal aplicación del análisis AID es en segmentación de mercados. Las variables independientes o explicativas utilizadas en el análisis AID han de estar medidas con escalas nominales

u ordinales, y la variable dependiente o a explicar debe estar medida con una escala métrica (proporcional o interválica) o dicotómica (valores 1 ó 0).

El análisis CHAID (*Chi Square Automatic Interaction Detection*) se fundamenta en los trabajos estadísticos de Kass (1980) y es similar al análisis AID, aunque presenta algunas diferencias, entre ellas el que tanto la variable dependiente como las independientes son categóricas. Al igual que el AID identifica las diferencias en la distribución de las respuestas de la variable criterio sobre la base de las características de las variables predictoras. El CHAID realiza también una partición secuencial de los elementos de la muestra, formando grupos, pero la partición no es necesariamente dicotómica, como en el AID, sino que pueden generarse más de dos grupos en cada subdivisión. También se diferencia del AID en el criterio para efectuar la partición de la variable dependiente. En el AID se basa en un análisis de la varianza de las diferencias entre las medias de todos los grupos dicotómicos posibles. En el CHAID, en cambio, al ser la variable dependiente categórica, se efectúan todas las tabulaciones cruzadas posibles con las variables independientes, se calcula en cada tabla cruzada el estadístico χ^2 y se elige la agrupación de las categorías de la variable independiente que presenta el nivel de significación más alto. Una vez obtenido el primer nivel de agrupación de las variables independientes, se procede con el siguiente nivel, subdividiendo los grupos obtenidos de forma similar a la efectuada en el primer nivel.

El análisis conjunto o medición conjunta (*conjoint measurement*) es una técnica de análisis multivariable de dependencias que, como su propio nombre indica, trata de determinar el efecto conjunto de dos o más atributos de las variables estímulo (independientes) sobre la evaluación global de un objeto o sujeto (variable dependiente). El objetivo es descomponer la medida de la evaluación total en los componentes que se pueden imputar a los atributos de las variables estímulo. El criterio para medir la bondad de la descomposición es el grado de acuerdo entre los valores originales de la evaluación y los valores estimados por el análisis conjunto (Sheth, 1977). El análisis conjunto trata de estimar las utilidades para el decisor que se derivan de la valoración de los atributos de un producto u objeto a partir de las preferencias manifestadas a distintas combinaciones de estos atributos. La variable dependiente puede ser categórica, ordinal o numérica, mientras que las variables independientes (dos o más), son siempre categóricas. Esta técnica, por tanto, mide el efecto conjunto de dos o más variables independientes no métricas sobre la clasificación, ordenamiento o valoración de una variable dependiente. Además de ser una técnica de análisis de dependencias, la medición conjunta puede concebirse también como una metodología de la investigación que se apoya en diseños experimentales (Varela y Braña, 1996). Se distingue entre el análisis conjunto categórico, desarrollado por Carroll y Chang (Carroll, 1969; Carroll y Chang, 1993) y el ordinal (MONANOVA), propuesto por Kruskal (1965). En el primero la variable dependiente es categórica, mientras que en el segundo puede ser ordinal o numérica.

Dentro del grupo de técnicas que analizan la relación de dependencia entre múltiples variables dependientes y múltiples variables independientes, se integran las siguientes:

- Análisis discriminante múltiple.
- Análisis de correlaciones canónicas.
- Redes neuronales artificiales.

El análisis discriminante permite explicar la pertenencia de individuos u objetos a grupos preestablecidos (por ejemplo: comprador habitual, comprador ocasional o no comprador). Las variables dependientes están formadas por las categorías o clases de una variable no métrica. Si sólo son dos categorías, se trata de un discriminante simple. Si hay más de dos categorías, es un discriminante múltiple. Las variables independientes son métricas. El resultado del análisis discriminante es un conjunto de funciones que son una combinación lineal de las variables explicativas o independientes de forma que maximizan la distancia o separación entre grupos. Se obtienen tantas funciones discriminantes como grupos menos uno. Los coeficientes de las funciones discriminantes indican la medida en que las variables explicativas determinan la inclusión de los individuos u objetos en los grupos preestablecidos. Los n grupos se pueden transformar en $n-1$ variables ficticias y resolverse el problema aplicando el método general de análisis de correlaciones canónicas (Holbrook y Moore, 1982).

El análisis de correlaciones canónicas es una técnica poco conocida y utilizada, pero que es de gran utilidad para el análisis de relaciones de dependencias entre múltiples variables dependientes e independientes. Supone una extensión del análisis de regresión múltiple, por cuanto permite estimar no sólo la relación funcional de múltiples variables independientes con una variable dependiente, sino con múltiples dependientes. Puede decirse, por tanto, que el análisis de regresión múltiple es un caso particular del análisis de correlaciones canónicas, en el que sólo hay una variable dependiente y se obtiene una única función lineal. En el análisis de correlaciones canónicas, como caso general, hay dos o más variables dependientes y el número de funciones que se estima es igual al número de variables dependientes. El número de variables independientes tiene que ser igual o mayor que el de dependientes (Hair, Anderson, Tatham y Black, 1999). Al igual que en el análisis de regresión múltiple, las relaciones estimadas entre las variables dependientes e independientes son lineales. Asimismo, las variables tienen que ser todas métricas, medidas con escalas proporcionales o interválicas. Para poder analizar variables categóricas tienen que transformarse previamente en ficticias.

Las redes neuronales artificiales tratan de identificar relaciones en estructuras de datos. Se relacionan con los sistemas expertos y utilizan un enfoque estadístico. Los algoritmos de las redes neuronales se basan en el funcionamiento del cerebro humano (Hair *et al.*, 1999; Corchado *et al.*, 2000). Las redes neuronales tratan de reprodu-

cir la estructura cerebral y su forma de trabajar, con un conjunto de neuronas interconectadas por *sinapsis*. Los modelos de redes neuronales emulan el aprendizaje humano, mediante prueba y error, son capaces de aprender de la experiencia (Sánchez, 1999) y pueden identificar patrones de comportamiento y relaciones causa-efecto en estructuras de datos. Las relaciones entre variables que pueden estudiar las redes neuronales pueden ser tanto lineales como no lineales. Las técnicas de redes neuronales se caracterizan también por el método de procesamiento de los datos. Sus algoritmos de cálculo son iterativos, requieren mucho tiempo de procesamiento, no se obtienen siempre los mismos resultados y la interpretación de los resultados es más difícil. Frente a estos inconvenientes, tienen la ventaja de ser más tolerantes en los tipos de variables que pueden tratar y aun cuando falten datos los resultados no son afectados de forma significativa. Además, dado que los modelos de redes neuronales pueden estimar tanto variables cuantitativas (numéricas) como de clasificación (categóricas), pueden utilizarse los mismos algoritmos para fines de previsión y de clasificación (Venugopal y Baets, 1994). Los modelos de redes neuronales pueden ser supervisados y sin supervisión. Los primeros pueden utilizarse como sustitutos de los modelos de regresión y discriminante. La disposición de una red neuronal en varias capas sucesivas de neuronas se denomina perceptrón multicapa (Rumelhart et al., 1986). El perceptrón multicapa es el exponente más típico de las redes neuronales artificiales con aprendizaje supervisado. También se le denomina red de *retropropagación*, por el modo en que se modifican los pesos asociados a las neuronas utilizando el error obtenido en la capa de salida o de variables *output* (Corchado et al., 2000; Barath y Drosen, 1994).

III.1.2. Técnicas de análisis de dependencias entre múltiples grupos de variables

En general, los modelos causales se utilizan para especificar un fenómeno en términos de causa y efecto. La relación causal entre dos variables es directa cuando un cambio en una de ellas produce una variación en la otra sin intervenir una tercera variable. La relación es indirecta, en cambio, cuando una o más variables intermedian entre la causa y el efecto, como sucede en los modelos de redes neuronales. Un modelo es recursivo si todas las relaciones causales son en un solo sentido, es decir, no hay dos variables recíprocamente relacionadas y ninguna variable tiene un efecto *feedback* sobre sí misma, a través de una concatenación indirecta de enlaces causales (Bagozzi, 1980). Todas las técnicas expuestas en los apartados anteriores se basan en modelos recursivos.

Los modelos de ecuaciones estructurales son modelos causales recursivos y con relaciones directas e indirectas. Estos modelos analizan relaciones causales múltiples. En su diseño incluyen construcciones teóricas, denominadas también variables latentes o constructos (que se representan gráficamente por círculos) y variables observadas operativas (representadas por cuadrados). Las variables latentes pueden

ser exógenas o endógenas. Las exógenas son aquellas que no están causadas o no dependen de ninguna otra variable. Son, por tanto, variables independientes cuya variabilidad se atribuye a causas externas. Sus valores se determinan fuera del modelo. Las variables endógenas, en cambio, son aquellas cuyo valor depende de otras variables en el modelo, tanto exógenas como endógenas (Bagozzi, 1980; Luque, 2000).

En un modelo de ecuaciones estructurales se integran simultáneamente un conjunto de ecuaciones de regresión múltiple diferentes pero a la vez interdependientes, ya que las variables que son dependientes en una relación pueden ser independientes en otra relación dentro del mismo modelo (Luque, 2000).

Los modelos de ecuaciones estructurales son adecuados para problemas en los que se quiere analizar una relación compleja entre varios grupos de variables. En el modelo se establece una serie de relaciones simultáneas, en la que, de modo secuencial y en una sola dirección, se plantean asociaciones causa-efecto entre variables latentes (constructos). Estas variables pueden actuar como independientes y explicar otras variables dependientes que, a su vez, pueden explicar otras posteriores.

Para estimar los coeficientes de los modelos de ecuaciones estructurales puede utilizarse diversos programas, como LISREL (Linear Structural RELationship) (Jöreskog y Sörbom, 1993; Kelloway, 1998), EQS (Byrne, 2006), AMOS (Byrne, 2009) y SmartPLS (Ringle, C.M.; Wende, S.; Will, S., 2005). Esta última técnica es especialmente adecuada para el análisis de muestras pequeñas y datos que no muestran necesariamente una distribución normal multivariante (Barroso *et al*, 205; Aldás, 2010).

III.2. Análisis de interdependencias entre variables y detección de dimensiones subyacentes

Dentro de las técnicas de análisis de interdependencias que miden la interrelación entre varias variables, cabe incluir las siguientes:

- Análisis de componentes principales.
- Análisis factorial de correspondencias.
- Análisis multidimensional (métrico y no métrico).

El análisis de componentes principales (ACP) es una técnica de análisis factorial que trata de identificar la estructura de un conjunto de variables observadas. Implica el estudio de las interrelaciones entre las variables (determinadas por las correlaciones o por las covarianzas) con el fin de hallar un nuevo conjunto de variables, menor en número que el de variables originales, que exprese las dimensiones que hay en común en las variables originales (Harman, 1976; Rummel, 1970). Cuando el número de variables originales es muy grande, con el análisis factorial se puede

reducir este conjunto a otro menor de factores, reteniendo la mayor parte de la varianza de las variables originales (Stewart, 1981; Hair *et al.*, 1999; Luque, 2000). Con el ACP se obtienen tantos factores o componentes como variables tiene el conjunto de datos analizado. Un factor es una combinación lineal de las variables originales. Los factores se representan geoméricamente por ejes, que son ortogonales, es decir, perpendiculares, lo que implica que los factores son independientes y no está correlacionados entre sí. Los factores se generan de forma ordenada, según su importancia en la explicación de la varianza de las variables originales consideradas (Green y Carroll, 1976). Con un número de factores reducido, inferior al número de variables originales, se puede conseguir explicar la mayor parte de la varianza de las mismas.

El análisis factorial de correspondencias es una modalidad de análisis factorial desarrollada en 1963 por Benzecri (1980). El análisis factorial de correspondencias (AFC) permite representar de forma simultánea, en una configuración multidimensional, las variables filas y columnas de una tabla de contingencia (Lefebvre, 1983). El método que desarrolla el AFC está basado en la descomposición de la *ji* cuadrado de una tabla de contingencia (Foucart, 1982). El estudio de las interrelaciones entre variables fila y columna se realiza por medio de las representaciones gráficas de las variables en un espacio *n*-dimensional. Estas representaciones gráficas se obtienen a partir de los valores de las coordenadas calculadas por el AFC. A partir de la tabla de contingencia, se calcula una matriz de covarianzas de las variables columna, que luego se factoriza aplicando el análisis de componentes principales (Cuadras, 1981). Las raíces y los vectores característicos que se obtienen permiten calcular las coordenadas de las variables filas y columnas. La correlación de cada variable con cada uno de los ejes factoriales obtenidos depende del valor de la coordenada respecto del eje considerado y las restantes coordenadas con los demás ejes.

El análisis multidimensional parte de una matriz simétrica en la que se expresan las similitudes entre pares de objetos (marcas, empresas, personas, conceptos, etc.). La finalidad es determinar las coordenadas y representar estos *n* objetos en un espacio de *t* dimensiones (normalmente dos o tres), de modo que las distancias entre los puntos en el espacio concuerden lo máximo posible con las similitudes observadas (Luque, 2000; Hair, Anderson, Tatham y Black, 1999). Las similitudes (o disimilitudes) pueden ser calculadas por el programa (por ejemplo, con la distancia euclídea) o estimadas subjetivamente por los encuestados (mediante evaluaciones cuantitativas o cualitativas). Las escalas de medida de las variables pueden ser métricas o no métricas, lo que da lugar al análisis multidimensional métrico y no métrico, respectivamente. Al igual que en el análisis factorial, hay un problema de interpretación subjetiva de las dimensiones, que tendrán que identificarse en función de la posición que ocupan los objetos en el espacio configurado (Schiffman, Reynolds y Young, 1981).

III.3. Técnicas de agrupamiento. Segmentación de mercados y tipología

Las técnicas que analizan la interrelación entre sujetos u objetos suelen agruparse dentro de la denominación genérica de técnicas de *cluster analysis* o análisis de grupos. Existen distintos métodos y algoritmos de agrupamiento, entre los que cabe destacar los siguientes:

- Métodos de agrupamiento ascendente (*building up*), como el algoritmo de Johnson.
- Métodos de agrupamiento descendente (*breaking down*), como el algoritmo de Howard-Harris.
- Métodos de conglomerados de k medias.

El análisis de grupos (*cluster analysis*) es un conjunto de técnicas estadísticas que sirven para determinar grupos internamente homogéneos, pero distintos entre sí, bien por agrupación de unidades más pequeñas o por división de segmentos mayores (Anderberg, 1973; Hartigan, 1975). Para determinar la similitud entre objetos existen diversas medidas de distancia (Sokal, 1977), pero una de las más utilizadas es la distancia euclídea. Las técnicas de análisis de grupos se aplican en investigación de mercados para definir tipologías e identificar segmentos de mercado (Punj y Stewart, 1983; Saunders, 1994).

Existe una gran variedad de técnicas de análisis de grupos. Puede distinguirse entre las ascendentes (*building up*) que construyen los grupos por agregación, a partir de los individuos considerados uno a uno, y las descendentes, que, al contrario de las anteriores, parten del conjunto total de individuos y lo dividen (*breaking down*) en grupos más pequeños (Wind, 1978). Dentro de las técnicas ascendentes, cabe destacar el algoritmo de Johnson (1967), que agrupa los sujetos u objetos a partir de la matriz de distancias euclídeas entre todos los pares de ellos. La agrupación se efectúa de forma secuencial, de menor a mayor distancia. Dentro de las técnicas que siguen el enfoque de partición o descendentes, cabe citar el algoritmo de Howard-Harris, que forma grupos por división de otros de tamaño mayor, de modo también secuencial, utilizando el criterio de la minimización de la varianza intra-grupos en cada nivel de la división (Green y Wind, 1973). En la mayoría de las técnicas de análisis de grupos las variables requeridas son métricas.

Otra técnica de análisis de grupos es el de conglomerados de K -medias, que genera un número fijo de grupos establecido por el usuario del programa, por lo que no puede incluirse en ninguno de los dos tipos anteriores. El objetivo del algoritmo es obtener K grupos (*clusters*) de modo que se minimice la suma de cuadrados intragrupos.

IV. ¿Qué técnicas de análisis de datos hay que utilizar?

La elección de la técnica de análisis depende de la naturaleza del problema a investigar, de los objetivos planteados, de las hipótesis especificadas y de las escalas de medida de los datos.

En primer lugar, el problema a investigar puede suponer el análisis de una relación simple entre dos variables o grupos de variables, o bien un análisis más complejo, que entrañe múltiples relaciones entre múltiples variables o grupos de variables. En el primer supuesto, será suficiente utilizar técnicas de análisis de dependencias, bien con una única variable dependiente o con varias dependientes, como las que se han descrito en el apartado III.1.1. En el caso de que se requiera analizar la interdependencia entre múltiples variables y exista una relación de causalidad, de modo que una o varias variables dependan de otras y a su vez éstas actúen como independientes y expliquen otras variables dependientes, entonces las técnicas de análisis adecuadas son los modelos de ecuaciones estructurales, expuestos en el apartado III.1.2. Si no se supone la existencia de relaciones causales entre las variables y se quiere conocer sólo las interrelaciones o asociaciones existentes entre ellas y, por tanto, comprobar cómo se agrupan esas variables, en este caso es adecuado utilizar las técnicas de análisis de interdependencias entre variables y detección de dimensiones subyacentes, descritas en el apartado III.2. Si en lugar de variables se quiere conocer cómo se agrupan individuos u objetos por su similitud o proximidad, entonces resultan apropiadas las técnicas de análisis de grupos o *cluster analysis*, descritas en el apartado III.3.

Los objetivos del estudio, así como las hipótesis especificadas, determinarán también las técnicas apropiadas a utilizar. Si se quieren contrastar relaciones causa-efecto, las técnicas de análisis de dependencias, descritas en el apartado III.1, serán las adecuadas. Si el objetivo es descubrir dimensiones subyacentes en las interrelaciones entre variables, las técnicas de análisis factorial y multidimensional, descritas en el apartado III.2., serán las apropiadas. Finalmente, si el objetivo es comprobar cómo se agrupan individuos u objetos homogéneos, en función de su similitud o proximidad, las técnicas de análisis de grupos, descritas en el apartado III.3, permitirán alcanzar dicho objetivo.

Dentro de cada grupo de técnicas de dependencias entre variables, interrelaciones entre variables o agrupamiento de individuos u objetos, la escala de medida de las variables condiciona también la técnica adecuada a elegir. Aunque cualquier variable medida en una escala métrica puede convertirse en otra medida en una escala no métrica, y a la inversa, si bien con pérdida de información, hay técnicas que están diseñadas específicamente para determinados tipos de variables. Así, en el análisis de dependencias, el ACM y el AID son técnicas diseñadas para los supuestos en que las variables independientes sean no métricas; el CHAID y el análisis conjunto, para variables dependientes e independientes no métricas y el análisis discriminante múltiple, para variables dependientes no métricas. Dentro de las técnicas de análisis de interdependencias, que en general requieren variables métricas, el análisis factorial de correspondencias opera con variables categóricas, a partir de una tabla de contingencia de frecuencias y el análisis multidimensional no métrico utiliza como input variables medidas con una escala ordinal. También hay técnicas

de *cluster analysis* que pueden determinar grupos a partir de variables cualitativas o no métricas.

V. Conclusiones

Para que el marketing pueda cumplir sus fines de conocer las necesidades y deseos del consumidor y tratar de satisfacerlos del modo más efectivo, debe analizar el mercado y aplicar los métodos y técnicas que proporciona la investigación de mercados.

La investigación de mercados implica buscar y analizar, de modo sistemático y objetivo, la información que es relevante para identificar y solucionar problemas de marketing.

Toda investigación de mercados ha de fundamentarse en un diseño o plan para la recogida y análisis de los datos. El plan de investigación ha de partir de la identificación del problema que se va a investigar e incluir los objetivos a alcanzar, las hipótesis a contrastar, las fuentes de información a utilizar, la forma de definir, clasificar y medir las variables del estudio y su plasmación en un formulario o cuestionario, como pasos previos al análisis de los datos y la interpretación de los resultados.

La finalidad del análisis de los datos es su transformación en información relevante para la toma de decisiones. La complejidad del mismo está en función del número de variables que se analizan simultáneamente. Así, las técnicas de análisis de datos, como se ha expuesto en este artículo, pueden clasificarse en univariadas, bivariadas y multivariadas, según que analicen, respectivamente, el comportamiento de una sola variable, la relación o dependencia entre dos variables y la relación o interdependencia entre más de dos variables.

El análisis estadístico multivariable es el más completo y permite aproximarse mejor a la realidad de los fenómenos económicos y sociales. Las técnicas que incluye suelen clasificarse en dos grandes grupos, según analicen relaciones de dependencia o de interdependencia entre variables u objetos.

A lo largo de este artículo se han descrito las características y aplicaciones de las distintas técnicas de análisis multivariable y se han propuesto criterios para elegir las más adecuadas en función de la naturaleza del problema a investigar, de los objetivos planteados, de las hipótesis especificadas y de las escalas de medida de los datos. Ninguna de las técnicas descritas sirve para analizar todo tipo de problema y es necesario siempre el buen criterio y experiencia del investigador para elegir y combinar adecuadamente las técnicas que permitan descubrir mejor la información que subyace en los datos obtenidos.

Bibliografía

Aaker, David A.; Day, George S. (1980), *Marketing Research: Private and Public Sector Decisions*, John Wiley & Sons, New York.

- Aldas Manzano, Joaquín (2010), “Modelización estructural mediante Partial Least Squares (PLSPM)”, II Seminario AEMARK de modelización estructural, Alcalá de Henares, 12 febrero 2010.
- American Marketing Association (1987), “New Marketing Research Definition Approved”, *Marketing News*, núm. 21, 2 de enero de 1987, pág. 1.
- American Marketing Association (2004), Board of Directors, octubre 2004.
- Anderberg, Michael R. (1982), *Cluster Analysis for Applications*, Academic Press, Inc., New York.
- Andrews, Frank M.; Morgan, James N.; Sonquist, John A. (1967), *Multiple Classification Analysis*, Institute for Social Research, The University of Michigan, Ann Arbor, Michigan.
- Andrews, Frank M.; Morgan, James N.; Sonquist, John A.; Klem, Laura (1973), *Multiple Classification Analysis, 2ª ed.*, Institute for Social Research, The University of Michigan, Ann Arbor, Michigan.
- Bagozzi, Richard P. (1980), *Causal Models in Marketing*, John Wiley & Sons, New York.
- Barroso, Carmen, Gabriel Cepeda y José Luiz Roldán (2005), “Investigar en Economía de la Empresa: ¿Partial Least Squares o Modelos basados en la Covarianza?”, *Best Paper Proceedings 2005: el Comportamiento de la Empresa Ante Entornos Dinámicos*. Vitoria-Gasteiz. AEDEM., vol. 1.
- Bean, Andrew G.; Roszkowski, Michael J. (1995), “The Long and Short of It”, *Marketing Research*, 7, 1, invierno, págs. 21-26.
- Benzécri, J.P. (1980), *Pratique de L'Analyse des Données. I. Analyse des Correspondances*, Dunod, Paris.
- _____ (1980), *L'Analyse des Données, II. L'Analyse des Correspondances*, Dunod, Paris.
- Berenson, Mark L.; Levine, David M.; Goldstein, Matthew (1983), *Intermediate Statistical Methods and Applications*, Prentice-Hall, Inc., Englewood Cliffs, New Jersey.
- Bharath, Ramachandran; Drosen, James (1994), *Neural Network Computing*, Windcrest-McGraw-Hill, Inc., New York.
- Byrne, Barbara.M. (2006), *Structural equation modeling with EQS and EQS/Windows*, 2ª ed.. Sage Publications. Thousand Oaks, California
- _____ (2009), *Structural Equation Modeling With AMOS: Basic Concepts, Applications, and Programming*, 2ª ed., Routledge Academic, New York.
- Carroll, J.Douglas (1969), “Categorical Conjoint Measurement”, en *Meeting of Mathematical Psychology*, Ann Arbor, Michigan.
- Carroll, J.Douglas; Chang, Jih Jie (1993), “Categorical Conjoint Measurement”, programa de ordenador, ATT.
- Churchill, Gilbert A., Jr. (1987), *Marketing Research. Methodological Foundations*, 4ª ed., The Dryden Press, Chicago.

- Churchill, Gilbert A., Jr. (2001), *Basic Marketing Research*, 4ª ed., The Dryden Press, For Worth.
- Corchado, Juan Manuel; Díaz, Fernando; Borrajo, Lourdes; Fernández, Florentino (2000), *Redes Neuronales Artificiales: un enfoque práctico*, Servicio de Publicaciones, Universidad de Vigo.
- Cuadras, C. M. (1981), *Métodos de Análisis Multivariante*, Eunibar, Barcelona.
- Foucart, Thierry (1982), *Analyse Factorielle. Programmation sur microordinateurs*, Masson, París.
- Gandz, Jeffrey; Whipple, Thomas W. (1977), "Making Marketing Research Accountable", *Journal of Marketing Research*, XIV, 2, mayo, págs. 202-208.
- Green, Paul E. (1978), *Analyzing Multivariate Data*, The Dryden Press, Hinsdale, Illinois.
- Green, Paul E.; Carroll, J. Douglas (1976), *Mathematical Tools for Applied Multivariate Analysis*, Academic Press, New York.
- Green, Paul E., Donald S. Tull y Gerald Albaum (1988), *Research for Marketing Decisions*, 5ª ed., Prentice-Hall, Englewood Cliffs, New Jersey.
- Green, Paul E.; Wind, Yoram (1973), *Multiattribute Decisions in Marketing*, The Dryden Press, Hinsdale, Illinois.
- Hair, Joseph F., Jr; Anderson, Rolph E.; Tatham, Ronald L.; Black, William C. (1999), *Análisis Multivariante*, 5ª ed., Prentice-Hall, Madrid.
- Harman, Harry (1976), *Modern Factor Analysis*, 3ª ed., The University of Chicago Press, Chicago.
- Hartigan, J. A. (1975), *Clustering Algorithms*, John Wiley & Sons, New York.
- Holbrook, Morris B.; Moore, William L. (1982), "Using Canonical Correlation to Construct Product Spaces for Objects with Known Features Structures", *Journal of Marketing Research*, XIX, 1, febrero, págs. 87-98.
- Johnson, S. C. (1967), "Hierarchical Clustering Schemes", *Psychometrika*, 32, págs. 241-254.
- Johnston, J. (1975), *Métodos de Econometría*, 3a. ed., Editorial Vicens Vives, Barcelona.
- Jöreskog, K. G. y Sörbom, D. (1993), *LISREL VIII: Structural equation modeling with the SIMPLIS command language*, Hillsdale, N.J. Scientific Software.
- Kass, Gordon V. (1980), "An Exploratory Technique for Investigating Large Quantities of Categorical Data", *Applied Statistics*, 29, 1980, págs. 119-127.
- Kelloway, E. Kevin (1998): *Using LISREL for Structural Equation Modeling*, Sage Publications, Thousand Oaks, California.
- Kruskal, Joseph B. (1965), "Analysis of Factorial Experiments by Estimating Monotone Transformations of the Data", *Journal of the Royal Statistical Society, Series B*, 27, 2, págs. 251-263.
- Lefebvre, Jacques (1983), *Introduction aux Analyses Statistiques Multidimensionnelles*, 3ª ed., Masson, París.

- Luque, Teodoro (1997), *Investigación de marketing. Fundamentos*, Editorial Ariel, Barcelona.
- Luque Martínez, Teodoro (coordinador) (2000), *Técnicas de análisis de datos en investigación de mercados*, Ediciones Pirámide, Madrid.
- Morgan, James N.; Sonquist, John A. (1963), "Problems in the Analysis of Survey Data, and a Proposal", *Journal of the American Statistical Association*, 58, junio, págs. 415-435.
- Morrison, Donald F. (1976), *Multivariate Statistical Methods*, McGraw-Hill Kogakusha, Tokyo.
- Pedret, Ramón; Sagnier, Laura; Camp, Francesc (2000), *La investigación comercial como soporte del marketing*, Ediciones Deusto, Bilbao.
- Perreault, William D., Jr. (1992), "The Shifting Paradigm in Marketing Research", *Journal of the Academy of Marketing Science*, 20, 4, otoño, págs. 367-387.
- Punj, Girish; Stewart, David W. (1983), "Cluster Analysis in Marketing Research: Review and Suggestions for Application", *Journal of Marketing Research*, XX, 2, mayo, págs. 134-148.
- Ringle, C.M.; Wende, S.; Will, S. (2005): *SmartPLS 2.0 (M3) Beta*, Hamburg.
- Roscoe, A. Marvin, Jr.; Lang, Dorothy; Sheth, Jagdish N. (1975), "Follow-up Methods, Questionnaire Length, and Market Differences in Mail Surveys", *Journal of Marketing*, 39, 2, abril, págs. 20-27.
- Rummel, R. J. (1970), *Applied Factor Analysis*, Northwestern University Press, Evanston.
- Rumelhart, D. E.; Hinton, G. E.; Williams, R.J. (1986), "Learning internal representations by error propagation", en D. E. Rumelhart y J. L. McClelland (eds.), *Parallel Distributed Processing: explorations in the microstructure of cognition, vol. 1: Foundations*, Cambridge: The MIT Press.
- Sánchez Herrera, Joaquín (1999), "Las redes neuronales artificiales y su aplicación en marketing", *IPMARK*, 520, abril, págs. 47-49.
- Santesmases Mestre, Miguel (1982), "Aplicación del análisis de clasificación múltiple en investigación comercial", *Esic Market*, 39, septiembre-diciembre, págs. 95-105.
- _____ (1996), *Términos de Marketing, Diccionario-Base de datos*, Ediciones Pirámide, Madrid.
- _____ (2008), *Marketing. Conceptos y estrategias*, 5ª ed. revisada, Ediciones Pirámide, Madrid.
- _____ (2009), *Dyane, Versión 4. Diseño y análisis de encuestas en investigación social y de mercados*, Ediciones Pirámide, Madrid.
- Saunders, John (1994), "Cluster Analysis", en Graham J. Hooley y Michael K. Hussey, *Quantitative Methods in Marketing*, Academic Press, Londres.

- Schiffman, Susan M; Reynolds, Lance; Young, Forrest W. (1981), *Introduction to Multidimensional Scaling*, Academic Press, New York.
- Sheth, Jagdish N. (1977), *Multivariate Methods for Market and Survey Research*, AMA, Chicago.
- Simon, Julian L. (1969), *Basic Research Methods in Social Science. The Art or Empirical Investigation*, Random House, New York.
- Sokal, Robert R. (1977), "Clustering and Classification: Background and Current Directions", en Robert R. J. Van Ryzin, *Classification and Clustering*, Academic Press, New York,...
- Sonquist, John A. (1970), *Multivariate Model Building*, Survey Research Center, ISR, Univ. of Michigan, Ann Arbor, Michigan.
- Sonquist, John A.; Baker, Elizabeth Lauh; Morgan, James N. (1973), *Searching for Structure*, Survey Research Center, ISR, University of Michigan, Ann Arbor, Michigan.
- Stevens, S. S. (1946), "On the Theory of Scales of Measurement", *Science*, 103, págs. 677-680.
- Stewart, David W. (1981), "The Application and Misapplication of Factor Analysis in Marketing Research", *Journal of Marketing Research*, XVIII, 1, febrero, págs. 51-62.
- Varela Mallou, Jesús; Braña Tobío, Teresa (1996), *Análisis Conjunto. Aplicado a la investigación comercial*, Ediciones Pirámide, Madrid.
- Venugopal, V.; Baets, W. (1994), "Neural networks and statistical techniques in marketing research: A conceptual comparison", *Marketing Intelligence & Planning*, 12, 7, págs. 30-38.
- Wind, Yoram (1978), "Issues and Advances in Segmentation Research", *Journal of Marketing Research*, XV, 3, agosto, pág. 31.
- Zikmund, William G. (1998), *Investigación de mercados*, sexta edición, Prentice-Hall Hispanoamericana, México.

