
Análisis de Interrelaciones en las Canastas de Compra en un Supermercado¹

Máximo Bosch
Andrés Musalem
Dpto. de Ing. Industrial
Universidad de Chile

Resumen

El objetivo de este trabajo es presentar el estudio de interrelaciones entre categorías de productos en la canasta de los compradores de un supermercado. Este análisis corresponde a una representación espacial o gráfica de las distintas categorías de un supermercado a través de la cual se pueden detectar fácilmente aquellos conjuntos de productos que exhiben probabilidades altas de ser incluidos en la misma compra.

La metodología desarrollada se basa en el uso de las técnicas de análisis Escalamiento Multidimensional (MDS), Análisis de Conglomerados (Cluster Analysis) y en la estimación de modelos lineales. Estas técnicas utilizan la información transaccional capturada por los scanners en los puntos de venta. A partir de esta información se define y se estima la probabilidad condicional de que productos de dos categorías sean incluidos en la misma compra dado que productos de al menos una de las dos categorías sean comprados. Esta estimación corresponde a la medida de similitud de Jaccard.

Utilizando esta información y la técnica de análisis MDS se genera la representación espacial buscada la cual permite resumir la información contenida en estas probabilidades en un diagrama fácil de ser interpretado. En este diagrama

¹ Los autores agradecen a Marco Antonio Halabi y Andrés Montrone, Gerente de Operaciones y Gerente de Marketing de Economax quienes promovieron el desarrollo de ésta y otras investigaciones en dicha cadena de supermercados. De igual modo, agradecemos a los alumnos de Ingeniería Industrial Lorena Pino y Víctor Ocares quienes trabajaron en el almacenamiento y estructuración de la información necesaria para esta investigación y que han sido los ejecutores de muchas de nuestras ideas. Finalmente, agradecemos también los comentarios de Vicente Cox, Nancy Lacourly y Richard Weber.

cada categoría ocupa una posición en un espacio multidimensional. A partir de estas posiciones y a través de la técnica de Cluster Analysis se procedió a agrupar las distintas categorías en cuatro conglomerados, los cuales corresponden a: productos no-perecibles, productos frescos/consumo inmediato, productos de higiene y otros. Finalmente, utilizando información adicional acerca de las categorías (por ejemplo, ventas de la categoría, número de compras, tamaño promedio compras, etc.) se estimaron modelos lineales que permiten proyectar estas variables descriptoras en el espacio generado a partir de MDS. Esto permite complementar en forma importante las conclusiones de este análisis.

Finalmente, se describen las implicancias de los resultados para la gestión coordinada de las categorías de productos de un supermercado en lo relativo a definición de rol de categorías, decisiones tácticas (precio, surtido y promoción), diseño del lay-out de la sala de ventas y evaluación de desempeño. Adicionalmente, se presentan conclusiones respecto del potencial de aplicación de este análisis dentro y fuera del negocio supermercadista y respecto a las posibles líneas de investigación futura.

1. Introducción

Los supermercadistas enfrentan a diario el desafío de gestionar decenas de miles de productos tomando decisiones de precio, promoción, espacio, ubicación y reaprovisionamiento para cada uno de ellos. Estas decisiones para cada producto no sólo afectan las ventas del mismo sino también las ventas de muchos otros. En la práctica, es muy complejo considerar todas estas interrelaciones en el proceso de toma de decisiones de marketing.

El enfoque de Administración por Categorías (C.M.) ofrece una manera de enfrentar esta situación a través de la descomposición del problema global en un conjunto de sub-problemas cada uno de los cuales es prácticamente independiente de los demás. Esto se consigue agrupando productos altamente interrelacionados en categorías (Bosch, Musalem 2000) de modo tal que el comportamiento de los productos contenidos en una categoría sea prácticamente independiente de las decisiones tomadas sobre productos de otras categorías. Esta propiedad permite a los supermercadistas administrar estas categorías como unidades estratégicas de negocio con bastante autonomía (Nielsen Marketing Research 1992). Esta gestión cuasi-independiente puede ser aún mejorada si se detectan y se tienen en cuenta las principales interacciones y sinergias que existen entre las categorías (Bosch et al 2000).

En la última década la información capturada vía Scanner en el punto de venta (Point of Sale - P.O.S.) ha facilitado la realización de estudios cuantitativos que involucran productos de diversas categorías. Manchanda, Ansari y Gupta (1999) mencionaron tres líneas de investigación en este dominio. Una de ellas está orientada explícitamente al estudio de la dependencia entre productos de

distintas categorías (por ejemplo, ver Mulhern y Leone (1991)). La segunda línea de investigación se refiere al estudio de la relación entre una variable de interés (por ejemplo, elección de sala de ventas) y variables multicatóricas (por ejemplo, ver Bell y Lattin (1998)). La tercera línea de investigación se refiere al estudio transversal de fenómenos de marketing en muchas categorías y las generalizaciones que puedan ser inferidas de éstos (por ejemplo, ver Fader y Lodish (1990)).

En este trabajo se presenta un análisis que usa datos de scanner y las técnicas de análisis Escalamiento Multidimensional (M.D.S.) y Análisis de Conglomerados (Cluster Analysis) para detectar y representar gráficamente interacciones entre diferentes categorías de productos en la canasta de compra. Cabe destacar que la técnica de análisis M.D.S. ha sido ampliamente utilizada en estudios perceptuales (por ejemplo, ver Ghose (1998) y Sinha y DeSarbo (1998)) y también en análisis de elección al interior de una categoría (por ejemplo, ver Andrews y Manrai (1999) y Elrod (1988)). En contraste, en este trabajo se describirá una aplicación de esta técnica al análisis de la composición de la canasta de los compradores de un supermercado. También se presentarán las implicancias de este tipo de análisis para la gestión coordinada de categorías de productos (cross-category management) por parte de los supermercadistas.

2. Definiciones

El estudio de la composición de la canasta de compra se hará a partir del análisis de la probabilidad con la cual productos de diferentes categorías son incluidos en la misma compra. Para plantear la definición y estimación de estas probabilidades, definiremos los siguientes términos:

N = Número total de compras o transacciones del supermercado en un cierto período.

N_A = Número de compras que incluyen productos de la categoría A.

N_B = Número de compras que incluyen productos de la categoría B.

$N_{A \cup B}$ = Número de compras que incluyen productos de la categoría A, productos de la categoría B o productos de ambas categorías.

$N_{A \cap B}$ = Número de compras que incluyen simultáneamente productos de la categoría A y productos de la categoría B.

De esta forma, es posible estimar la probabilidad condicional de que una compra que contiene productos de la categoría B, también contenga productos de la categoría A:

$$p(A | B) = N_{A \cap B} / N_B \quad [2.1]$$

Adicionalmente, podemos estimar la probabilidad condicional de que una compra contenga productos de ambas categorías dado que incluye productos de la categoría A, de la categoría B o de ambas categorías. Esta estimación corresponde a la medida de similaridad de Jaccard también conocida como razón de similaridad:

$$p(A \wedge B \mid A \vee B) = N_{A \cap B} / N_{A \cup B} \quad [2.2]$$

Para ilustrar esta última definición consideremos el siguiente ejemplo: 100 compras incluyeron cerveza, pero no incluyeron bebidas no-alcohólicas (N.A.); 200 compras incluyeron bebidas no-alcohólicas, pero no incluyeron cerveza; y, 150 compras incluyeron simultáneamente bebidas no-alcohólicas y cerveza. Luego, la probabilidad condicional puede ser estimada de la siguiente manera:

$$\begin{aligned} p(\text{N.A.} \wedge \text{Cerveza} \mid \text{N.A.} \vee \text{Cerveza}) &= \text{Card.}(\text{N.A.} \cap \text{Cerveza}) / \text{Card.}(\text{N.A.} \cup \text{Cerveza}) \\ &= 150 / (100+200+150) \\ &= 1/3 \end{aligned}$$

La estimación de estas probabilidades o razones de similaridad entre cada par de categorías nos permite construir una matriz cuadrada. Dado que en este estudio se analizarán 33 categorías de productos, esta matriz contiene 33 filas y 33 columnas y 528 probabilidades relevantes². El análisis de la información contenida en esta matriz puede revelar interesantes conclusiones acerca de que pares de productos suelen ser incluidos en la misma compra con una mayor probabilidad. Obviamente, no es fácil ni eficiente llevar a cabo este análisis a través de la inspección de estas 528 probabilidades. En cambio, sería mucho más conveniente resumir la información de esta matriz representándola a través de un gráfico fácil de ser entendido por los administradores de las categorías.

De acuerdo a este objetivo, representaremos espacialmente las categorías de productos utilizando la información de probabilidades de compra simultánea. La idea es aplicar el procedimiento M.D.S. para generar una representación multidimensional en la cual aquellas categorías con mayor probabilidad condicional de compra simultánea se encuentren más cerca unas de otras. A continuación, se describirá la técnica de análisis (M.D.S.) y su aplicación a este caso particular.

3. Procedimientos para el análisis de interrelaciones

En esta sección se presentan los procedimientos que serán utilizados en este estudio para el análisis de las interrelaciones entre las categorías en la canasta de compra.

² En términos generales, si N es el número total de categorías, la matriz contiene N(N-1)/2 probabilidades relevantes, en vez de N². Esto se debe a que la matriz es simétrica y a que las componentes de su diagonal son iguales a 1.

3.1 Escalamiento Multidimensional (M.D.S.)³

La finalidad básica de los procedimientos de Escalamiento Multidimensional es generar una representación espacial (típicamente euclidiana y de baja dimensionalidad) de un conjunto de individuos a partir de información de distancia o disimilitud entre los individuos. Estos métodos tratan de encontrar el mejor ajuste entre los datos de distancias y las distancias calculadas a partir de la ubicación de los individuos en el espacio generado.

Por ejemplo, si se dispone de la distancia en kilómetros que existe entre cada par de capitales de los países de Sudamérica, se podría generar un mapa de Sudamérica aplicando MDS a estos datos. En efecto, el procedimiento MDS podría determinar la posición de cada ciudad en un espacio de dos dimensiones que permitiera la mayor correspondencia entre las distancias de entrada y aquellas que pueden ser calculadas a partir del mapa generado mediante MDS.

Esta correspondencia o grado de ajuste es medido típicamente mediante el indicador Stress de Kruskal o alguna de sus variantes. Este indicador (S) corresponde a:

$$S = \sqrt{\frac{\sum_{i \neq j}^m (d_{ij} - d'_{ij})^2}{\sum_{i \neq j}^m d_{ij}^2}} \quad [3.1]$$

en que d_{ij} corresponde a la distancia de entrada entre el individuo i y el individuo j ; d'_{ij} corresponde a la distancia calculada a partir de las coordenadas determinadas por el procedimiento de MDS; y, m corresponde al número total de individuos. Otra medida de ajuste corresponde a la correlación al cuadrado entre d_{ij} y d'_{ij} denotada por RSQ.

Un problema que suele enfrentarse en este tipo de análisis es la interpretación de los ejes o dimensiones generadas a través de MDS. En este sentido, los ejes pueden ser rotados manteniendo su ortogonalidad sin que se afecten las distancias entre los objetos. Esto sería equivalente a que en un mapa de una región o un país, se decidiera reemplazar los ejes Norte-Sur y Este-Oeste por los ejes Noroeste-Sureste y Noreste-Suroeste (ver figura 3.1). Esta rotación no modificaría las distancias que existen entre las distintas ciudades, simplemente la orientación de los objetos en el mapa o gráfico sería distinta.

3 Esta descripción se basa en el libro "Analyzing Multivariate Data" de P. Green (1978) y en "Design and Marketing of New Products" de G. Urban y J. Hauser (1993).

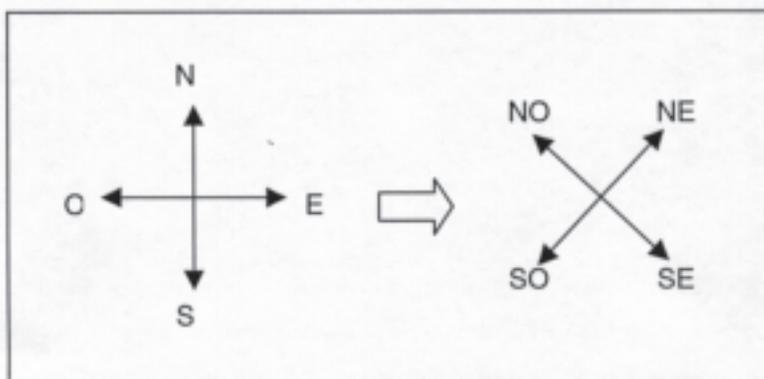


Figura 3.1: Ejemplo de Rotación de ejes.

Para ejecutar esta rotación es necesario encontrar ejes que puedan ser interpretables de alguna manera. Por ejemplo, en el caso de un mapa de una región geográfica de Sudamérica sabemos que, en general, las localidades ubicadas más al Norte exhiben temperaturas más altas que las ubicadas más al Sur. De esta manera, si la variable temperatura es relevante para el analista que está estudiando las distintas ciudades de Sudamérica, puede ser interesante identificar y graficar este eje Norte-Sur en su mapa.

De igual modo, se puede intentar detectar en un mapa generado a través de MDS algún eje tal que los objetos ubicados en distintos extremos de éste presenten diferencias en una o más variables conocidas y relevantes. Formalizando esta idea, si se detecta algún eje tal que las proyecciones de los distintos objetos en él exhiben una correlación significativa con alguna variable de interés, entonces puede resultar conveniente identificar y representar gráficamente este eje en el mapa.

Para esta identificación e interpretación de los ejes se puede recurrir a dos alternativas:

- Utilizar el conocimiento del analista o de un experto acerca de los objetos para detectar posibles ejes con sus respectivas interpretaciones.
- Ocupar información adicional acerca de los objetos. De esta forma, es posible encontrar ejes proyectando variables descriptoras de los objetos en el espacio generado a través de MDS. Para ello se puede estimar el siguiente modelo lineal para una variable descriptora y_i :

$$y_{ij} = c_l + \sum_{k=1}^K D_{lk} x_{kj} \quad [3.2]$$

en que y_{ij} es el valor de la variable descriptora l para el objeto j ; c_l es la constante del modelo; x_{kj} es la posición del objeto j en la dimensión k del mapa generado a través de MDS; y D_{lk} es el coseno direccional entre la variable descriptora l y la dimensión k . Usando las estimaciones de los parámetros D_{lk} es

posible dibujar las proyecciones de la variable descriptora en el mapa de MDS. De esta forma se obtienen ejes que permiten interpretar las diferencias entre las posiciones de los distintos objetos. Es importante mencionar que mientras mejor sea el ajuste de este modelo lineal (R^2), el eje estimado reflejará mejor la variable descriptora⁴.

3.2 Estudio de interrelaciones entre categorías mediante MDS

Como se mencionó anteriormente, en esta aplicación se pretende generar una representación espacial multidimensional en la cual aquellas categorías con mayor probabilidad condicional de compra simultánea se encuentren más cerca unas de otras. Es decir, estas probabilidades serán utilizadas como medidas de similitud entre categorías. Sin embargo, para aplicar el procedimiento MDS necesitaremos una medida de distancia o disimilitud. Esta medida de distancia puede ser simplemente obtenida estimando el complemento de la probabilidad o razón de similitud definida en la ecuación 2.2:

$$p(A \cup B | A \cap B) = 1 - N_{A \cap B} / N_{A \cup B} \quad [3.3]$$

Una vez ejecutado el procedimiento MDS⁵, será posible tener una visión conjunta de qué categorías suelen ser incluidas con mayor frecuencia en la misma compra a través de la representación gráfica de las categorías. Por otro lado, para la interpretación e identificación de los ejes de esta representación, se empleará información adicional referida a las categorías de productos. Esta información proviene de los datos de scanner y se describirá posteriormente.

3.3 Análisis de Conglomerados (Cluster Analysis)⁶

Este procedimiento multivariado permite asignar un conjunto de objetos, en este caso categorías, a un cierto número de segmentos llamados conglomerados. Esta asignación se efectúa a partir de información acerca de los objetos, de modo tal que un par de objetos que pertenecen a un mismo segmento sean más parecidos que un par de objetos que no pertenecen al mismo segmento.

Este parecido entre pares de objetos es medido principalmente de acuerdo a través de medidas de distancia y medidas de coincidencia. Como un ejemplo de las primeras, se puede mencionar la distancia euclídeana (d_{ij}) entre dos objetos de acuerdo a información contenida en r variables que describen dichos objetos:

4 Teniendo en cuenta esta situación puede resultar conveniente que la magnitud del vector que representa el eje en el gráfico sea proporcional al ajuste (R^2) obtenido del modelo lineal.

5 Para ello se empleará la versión 8.0 del software de análisis estadístico SPSS.

6 Esta descripción está basada en el manual "SPSS Base 9.0: Applications Guide", SPSS Inc. (1999) y en el libro "Analyzing Multivariate Data" de Paul Green (1978).

$$d_{ij} = \sqrt{\sum_{t=1}^r (X_{it} - X_{jt})^2} \quad [3.4]$$

en que X_{it} y X_{jt} son los valores de la variable descriptora t para los objetos i y j , respectivamente. Para el cálculo de estas distancias es importante que las r variables se encuentren en una base comparable (típicamente suelen normalizarse) y analizar si existe correlación significativa entre ellas. En este último caso, es posible y recomendable en muchas situaciones aplicar Análisis de Factores para transformar el conjunto de variables correlacionadas en un nuevo conjunto de variables ortogonales entre sí y luego ejecutar algún procedimiento de Cluster Analysis.

Por otro lado, las medidas de coincidencia son típicamente utilizadas cuando se dispone de información escalada nominalmente. En estos casos, una medida simple de coincidencia entre dos objetos corresponde a la razón entre el número de variables para el cual los dos objetos toman el mismo valor y el número total de variables.

Una vez definida la medida de similitud, es necesario seleccionar y ejecutar algún procedimiento para la generación de los conglomerados. El procedimiento utilizado en este trabajo corresponde al método de K-medias. Éste comienza usando los valores de las variables de los primeros k objetos como estimadores temporales de los centros de los k conglomerados, en que k es el número de conglomerados definido por el usuario. Los centros iniciales de los conglomerados se obtienen asignando cada caso al centro más cercano y luego actualizando los centros. Luego, se repite iterativamente este proceso hasta que los centros no cambien significativamente o bien hasta que el número máximo de iteraciones se haya cumplido. De esta forma, se obtienen los centros finales de cada uno de los conglomerados con la correspondiente asignación (estricta) de cada objeto a un conglomerado.

Además de este procedimiento (K-medias), existen métodos jerárquicos aglomerativos los cuales comienzan agrupando el par de objetos más cercanos de acuerdo a alguna medida de distancia combinándolos para formar un conglomerado. Luego, en cada paso sucesivo, se agrupan pares de objetos, pares de conglomerados o un objeto con un conglomerado hasta que todos los objetos sean agrupados en un solo conglomerado. Este agrupamiento suele presentarse a partir de un diagrama de árbol (dendrograma). Nótese que una vez que dos objetos o conglomerados son agrupados en una cierta iteración estos permanecerán juntos durante todas las iteraciones sucesivas.

Finalmente, también existen métodos basados en lógica difusa (Fuzzy C-Means). Estos últimos, a diferencia de los anteriores, permiten que un objeto pueda pertenecer (con grados de pertenencia a estimar) a más de un conglomerado a la vez.

3.4 Agrupación de categorías a partir de sus interrelaciones en la canasta de compra mediante Cluster Analysis

La información que se utilizará para generar la agrupación de categorías en esta aplicación corresponde a las puntuaciones de las categorías en las dimensiones entregadas por el procedimiento de Escalamiento Multidimensional (ver secciones 3.1 y 3.2).

De esta forma se generarán conglomerados de categorías tales que los productos de una categoría perteneciente a un conglomerado tendrán una probabilidad mayor de ser comprados junto con productos de una categoría perteneciente al mismo conglomerado que con productos de categorías de otros conglomerados.

4. Información para el estudio

Los datos para el estudio provienen de una sala de ventas de una cadena mediana de supermercados de Santiago, Chile. Esta sala exhibe un surtido cercano a los 7000 productos y el ticket promedio en dicha sala contiene productos de 7 categorías distintas. Esta información consiste en el detalle de cada compra (transacción), específicamente, productos vendidos, unidades vendidas y precio unitario de los productos vendidos, además de la fecha y hora de la transacción. Solamente un mes (julio 2000) fue considerado para este estudio. Estos datos fueron organizados y almacenados en una base de datos relacional. Esta organización incluyó también la definición de categorías, es decir, determinar qué productos pertenecen a cada una de las 33 categorías de productos analizadas en este estudio. La categorización de ACNielsen fue muy útil para este propósito (ACNielsen 2000).

Usando la base de datos se generó una matriz de disimilaridades (ver ecuación 3.3) que se presenta en la tabla 4.1. En este caso particular, los valores distintos de cero de la matriz son mayoritariamente cercanos a 1 lo cual se explica por ser ésta una sala de compras al paso más que de compras semanales o mensuales.

Esta matriz será utilizada por el algoritmo ALSCAL del módulo MDS del software SPSS. Adicionalmente, un conjunto de variables descriptoras de las categorías fue definido y estimado a partir de los datos de scanner. Este conjunto será utilizado en la etapa de determinación e interpretación de los ejes mediante la estimación de los modelos lineales definidos en la sección 3.1. En definitiva, las variables descriptoras consideradas corresponden a:

- LNT_j : Logaritmo natural del número de tickets que incluyen productos de la categoría j .
- $LCatSales_j$: Logaritmo natural de las ventas totales de productos de la categoría j .
- $LTotExp_j$: Logaritmo natural del gasto total en productos del supermercado que efectúan los compradores de productos de la categoría j .

- $AvgTSize_j$: Monto promedio de los tickets que incluyen productos de la categoría j .
- $LRSCatExp_j$: Logaritmo natural de la razón entre las ventas de la categoría j y el gasto total de los compradores de productos de la categoría j .
- $AvgNCat_j$: Número promedio de categorías diferentes incluidas en los tickets de compradores de productos de la categoría j .

Los valores estimados de estas variables para cada una de las categorías se presentan en la tabla 4.2. Se debe mencionar que también fueron analizadas otras variables estimables a partir de la información de los datos de scanner. Sin embargo, solamente los modelos lineales de las variables presentadas en este trabajo fueron ajustados razonablemente. Además, hubiese sido interesante haber dispuesto de información de margen bruto para poder complementar este análisis, sin embargo, esta información no se encontraba disponible.

Finalmente, es interesante mencionar que con anterioridad a la realización de esta investigación, esta información desagregada al nivel de cada transacción, era borrada semanalmente. En el supermercado sólo se almacenaban las ventas diarias de cada producto. De esta forma, esta investigación constituyó una oportunidad muy interesante de demostrar a los gerentes del supermercado el valor de la información que se solía eliminar.

Categoría	LNT	LCatSales	LTotExp	AvgTSize	LRSCatExp	AvgNCat
ACEITES COMESTIBLES	9,09	15,97	18,41	11156,78	-2,44	8,17
ALIMENTOS INFANTILES	7,94	14,75	17,34	12069,41	-2,59	8,19
ALIMENTOS PARA DESAYUNO	8,61	15,26	17,94	11292,67	-2,69	8,20
ARROCES	8,79	15,39	18,26	12985,70	-2,88	9,23
AZUCAR BLANCA GRANULADA	9,20	16,15	18,48	10736,48	-2,33	7,90
BEBIDAS INSTANTANEAS	8,54	14,59	17,63	8900,72	-3,04	7,73
BEBIDAS NO ALCOHOLICAS	9,57	16,41	18,31	6201,96	-1,90	4,75
CONFITES	8,03	14,60	17,06	8368,03	-2,46	5,77
FIDEOS	9,07	15,79	18,39	11114,27	-2,60	8,60
GALLETAS	8,99	15,40	17,96	7855,45	-2,56	6,10
HARINAS	8,03	14,73	17,59	14198,40	-2,86	9,12
YOGURT	9,07	15,45	18,09	8268,84	-2,64	6,85
QUESOS	8,62	15,42	17,77	9418,97	-2,36	6,84
SALSA DE TOMATES	8,69	14,81	18,12	12439,75	-3,31	9,40
TE	8,95	15,48	18,26	11139,07	-2,78	8,25
HELADOS Y POSTRES CONGELADOS	6,87	13,87	16,03	9423,44	-2,15	6,41

CECINAS	9,31	16,01	18,27	7750,01	-2,25	6,27
MARGARINA	8,86	15,24	18,08	10095,99	-2,84	8,00
SHAMPOO Y BALSAMOS	7,64	14,84	17,16	13538,42	-2,32	7,76
PAÑALES BEBE	7,61	15,59	16,59	7913,13	-1,00	4,47
PANADERIA Y MASAS PROPIAS	10,12	16,36	18,48	4252,16	-2,12	4,21
FRUTAS Y VERDURAS FRESCAS	9,48	15,90	18,16	5882,11	-2,26	5,33
ART. MENAJE Y VAJILLERIA	7,02	14,76	15,42	4452,41	-0,66	1,81
CERVEZAS	7,77	14,48	16,35	5329,54	-1,87	4,08
VINOS	8,83	16,11	17,59	6366,80	-1,48	4,35
CAFÉ	8,59	15,79	17,95	11628,24	-2,16	7,58
MAYONESA	8,39	15,01	17,78	11941,24	-2,78	8,28
TOALLAS HIGIENICAS	7,66	14,23	17,00	11454,56	-2,77	7,59
CARNES	9,17	16,63	18,12	7671,56	-1,48	5,98
PAPEL HIGIENICO	9,42	16,01	18,55	9219,99	-2,55	7,20
DETERGENTES	8,99	16,17	18,30	11030,55	-2,13	7,82
LECHES LIQUIDAS	9,34	16,08	18,36	8284,52	-2,29	6,55
LECHES EN POLVO	8,30	16,01	17,65	11526,25	-1,64	7,40

Tabla 4.2: Descriptores de las categorías en estudio.

5. Resultados

En esta sección se presentarán los resultados de las aplicaciones descritas en la metodología. De esta forma, se presentará en primer lugar la representación espacial generada mediante MDS; a continuación, la agrupación de las categorías a través de Cluster Analysis; luego, la incorporación de información adicional para la representación espacial mediante la estimación de modelos lineales; y, finalmente, la discusión acerca de los resultados obtenidos.

5.1 Representación espacial

Utilizando como datos de entrada los valores mostrados en la tabla 4.1, se obtuvo el escalamiento de las categorías en dos dimensiones que se muestra en la tabla 5.1. El ajuste que se obtuvo entre la medida de distancia original y la distancia escalada se puede observar a través de los indicadores STRESS y RSQ que alcan-

zaron valores de 0,36443 y 0,55527, respectivamente. Estos valores muestran que escalando las categorías en sólo dos dimensiones se puede obtener un ajuste razonable, pero bastante susceptible de ser mejorado. En efecto la adición de cuatro dimensiones más al escalamiento permitiría alcanzar valores de 0,15537 y 0,77163 de los indicadores STRESS y RSQ, respectivamente.

Para efectos de la representación gráfica, se trabajará sólo con dos dimensiones pues la finalidad última de este análisis es proveer una herramienta gráfica a los tomadores de decisión de un supermercado que sea fácil de ser entendida e interpretada. Esta representación se muestra en el gráfico 5.1 y se puede observar a partir de ella qué conjuntos de categorías de productos presentan una probabilidad relativamente más alta de ser incluidos en la misma compra. Para ello basta observar qué categorías se encuentran graficadas en posiciones cercanas.

Dado que en las secciones siguientes este diagrama será enriquecido con los resultados de Cluster Analysis y la estimación de los modelos lineales asociados a las variables descriptoras, se postergará la discusión de estos resultados hasta la sección 5.4.

CATEGORÍA	DIMENSIÓN 1	DIMENSIÓN 2
ACEITES COMESTIBLES	1,00	0,25
ALIMENTOS INFANTILES	-0,64	1,65
ALIMENTOS PARA DESAYUNO	0,95	0,72
ARROCES	0,93	0,48
AZUCAR BLANCA GRANULADA	0,98	0,21
BEBIDAS INSTANTANEAS	1,10	0,65
BEBIDAS NO ALCOHOLICAS	0,13	-1,49
CONFITES	-1,93	-0,29
FIDEOS	0,93	0,19
GALLETAS	0,21	-1,37
HARINAS	0,30	1,48
YOGURT	0,72	-0,92
QUESOS	-0,03	-1,51
SALSA DE TOMATES	0,91	0,52
TE	0,93	0,39
HELADOS Y POSTRES CONGELADOS	-2,08	0,45
CECINAS	0,35	-1,19
MARGARINA	1,00	0,11
SHAMPOO Y BALSAMOS	-0,98	1,46
PAÑALES BEBE	-1,76	0,93
PANADERIA Y MASAS PROPIAS	0,15	-1,37

FRUTAS Y VERDURAS FRESCAS	0,21	-1,29
ART. MENAJE Y VAJILLERIA	-1,92	-0,33
CERVEZAS	-1,81	-0,87
VINOS	-0,97	-1,30
CAFÉ	0,57	1,04
MAYONESA	0,32	1,11
TOALLAS HIGIENICAS	-1,35	1,06
CARNES	0,17	-1,20
PAPEL HIGIENICO	0,80	-0,15
DETERGENTES	0,81	0,31
LECHES LIQUIDAS	0,43	-0,99
LECHES EN POLVO	-0,43	1,27

Tabla 5.1: Solución de MDS en dos dimensiones (medida de distancia para MDS calculada a partir de ecuación 3.3).

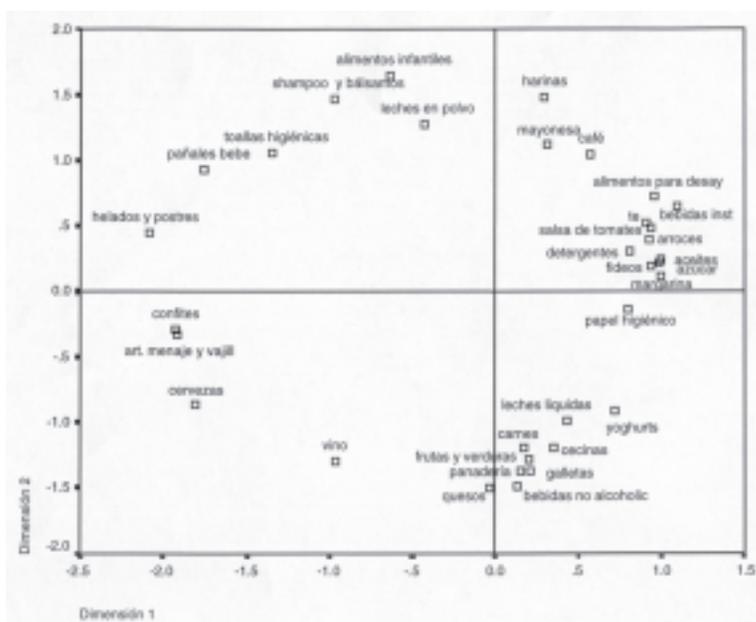


Gráfico 5.1: Solución de MDS en dos dimensiones.

5.2 Agrupación de categorías

En esta sección se agruparán las categorías en conglomerados utilizando MDS y Cluster Analysis. A diferencia de la sección anterior, en la cual se presentó una representación gráfica fácil de ser interpretada, buscaremos una clasificación que

aproveche al máximo la información que el escalamiento proporciona. Para ello recurriremos a la solución de MDS en 6 dimensiones⁷ que, tal como se mencionó anteriormente, ofrece un mejor ajuste entre las distancias de entrada y las distancias escaladas (STRESS = 0,15537; RSQ = 0,77163).

De acuerdo a esto se presenta en primer lugar, la solución de MDS en 6 dimensiones en la tabla 5.2. Esta solución provee 6 variables a partir de las cuales el procedimiento K-Medias de Cluster Analysis puede clasificar las categorías en K conglomerados. Para efectos de este análisis, se puede apreciar por mera inspección del gráfico 5.1 la formación de cuatro conglomerados, prácticamente uno en cada cuadrante. De esta forma se decidió aplicar este procedimiento clasificando las categorías en cuatro conglomerados⁸. Esta clasificación se puede observar en el gráfico 5.2, en el cual se muestra el conjunto de categorías que define cada conglomerado.

Categoría	Dim. 1	Dim. 2	Dim. 3	Dim. 4	Dim. 5	Dim. 6
ACEITES COMESTIBLES	1,43	0,54	0,79	0,25	0,30	0,22
ALIMENTOS INFANTILES	-0,61	1,31	-2,06	1,15	0,19	-0,21
ALIMENTOS PARA DESAYUNO	0,50	0,67	-1,54	0,53	-0,97	0,55
ARROCES	1,37	0,95	0,68	0,18	0,17	0,15
AZUCAR BLANCA GRANULADA	1,44	0,48	0,41	0,43	0,35	0,53
BEBIDAS INSTANTANEAS	0,82	0,53	0,57	-1,23	-0,28	-1,69
BEBIDAS NO ALCOHOLICAS	-0,32	-2,02	0,32	1,05	0,56	0,46
CONFITES	-1,19	-0,08	-0,78	0,76	0,09	-2,57
FIDEOS	1,49	0,46	0,55	0,03	0,31	0,04
GALLETAS	0,07	-1,23	-1,65	0,71	-0,33	-0,69
HARINAS	-0,13	1,42	0,95	0,50	-1,86	0,10
YOGURT	0,60	-1,01	-1,40	-0,76	-0,59	0,15
QUESOS	-0,60	-1,32	-1,02	-0,71	-1,59	0,54
SALSA DE TOMATES	1,38	0,97	0,74	-0,20	0,18	-0,31
TE	1,37	0,74	0,18	0,69	0,23	0,62
HELADOS Y POSTRES CONGELADOS	-2,27	0,31	1,30	-0,03	-2,06	-0,90
CECINAS	0,22	-1,77	-0,38	-0,61	-0,84	0,28
MARGARINA	0,95	0,24	-0,27	-0,72	-1,13	0,95
SHAMPOO Y BALSAMOS	-1,44	1,54	-0,54	-1,96	0,42	0,46
PAÑALES BEBE	-1,52	0,89	-1,06	-0,42	1,97	-1,63
PANADERIA Y MASAS PROPIAS	0,16	-2,32	-0,11	-0,23	0,21	0,11

7 Este fue el número máximo de dimensiones que el software SPSS admitió para el escalamiento.

8 Este procedimiento convergió en 25 iteraciones apreciándose cambios en los centros de los conglomerados inferiores a 10-16.

FRUTAS Y VERDURAS FRESCAS	0,42	-1,95	0,39	-0,66	0,45	-0,63
ART. MENAJE Y VAJILLERIA	-2,31	0,43	0,33	-1,09	1,92	0,49
CERVEZAS	-2,46	-0,44	1,57	1,56	-0,16	0,18
VINOS	-1,24	-0,98	1,23	1,60	1,05	0,49
CAFÉ	0,23	0,99	-0,08	1,61	-0,01	1,33
MAYONESA	0,24	0,48	1,21	-0,53	-1,07	-1,52
TOALLAS HIGIENICAS	-1,72	1,29	-0,24	-1,60	-0,19	1,31
CARNES	0,47	-1,52	0,56	-0,94	0,77	-0,85
PAPEL HIGIENICO	1,33	-0,14	0,39	0,05	0,79	0,40
DETERGENTES	1,19	0,63	0,57	-0,12	0,96	0,49
LECHES LIQUIDAS	0,29	-1,20	0,09	-0,76	0,01	1,56
LECHES EN POLVO	-0,14	1,12	-1,72	1,47	0,13	-0,41

Tabla 5.2: Solución de MDS en seis dimensiones (medida de distancia para MDS calculada a partir de ecuación 3.3).

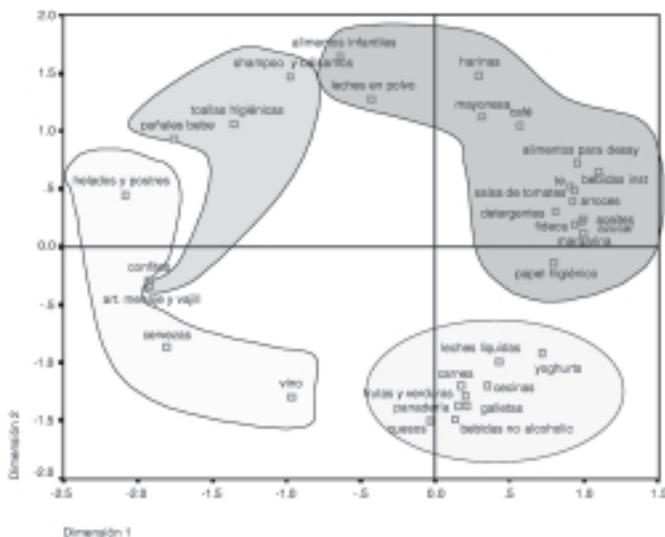


Gráfico 5.2: Solución de MDS en dos dimensiones y agrupamiento de las categorías de acuerdo a procedimiento K-medias de Cluster Analysis utilizando seis dimensiones.

5.3 Información adicional para la representación espacial

En esta sección, se estimarán modelos lineales (ver ecuación 3.2) que permitirán detectar e interpretar dimensiones relevantes del análisis a través de MDS presentado en las secciones anteriores. Como se mencionó previamente, la variable

5.4 Discusión acerca de los resultados

A partir de la representación espacial obtenida (ver gráfico 5.1) es posible detectar una zona central vacía alrededor de la cual las categorías forman una elipse. En otras palabras, no existe una categoría central o equidistante de las demás. Esto muestra que cada categoría tiende a relacionarse con una mayor probabilidad con un determinado subconjunto de categorías que con el resto lo que conduce a la formación de canastas específicas de compra. En otras palabras, debiera ser posible segmentar toda la variedad de tickets en una cantidad reducida de canastas.

A partir del procedimiento K-medias de Cluster Analysis ($k=4$), se clasificaron las categorías en cuatro conglomerados (ver gráfico 5.2). Específicamente, en el primer cuadrante se observa como los productos **no perecibles** (e.g., café, detergente, azúcar, etc.) definen un conglomerado. Lo mismo es válido para el conjunto de productos **frescos o de consumo inmediato** los cuales definen un tercer conglomerado en el cuarto cuadrante. De igual modo, los productos relacionados con **higiene** (además de los productos de menaje y vajilla) definen un tercer conglomerado en el segundo cuadrante. Finalmente, un cuarto segmento (**otros productos**) agrupa a las restantes categorías.

En consecuencia, los productos de alguno de estos conglomerados (no perecibles; frescos/consumo inmediato; higiene) se compran con mayor probabilidad con productos de categorías de su mismo conglomerado que con productos de categorías de otros conglomerados.

A partir de la proyección de las variables descriptoras en la representación espacial (ver gráfico 5.3), es posible apreciar que los compradores de productos de los conglomerados “no perecibles” y “frescos/consumo inmediato” exhiben los niveles más altos de gasto total en productos del supermercado y número de tickets. Más aún, los compradores del primer conglomerado también presentan un alto número promedio de categorías diferentes en el ticket y un alto monto promedio gastado por ticket. Este hecho muestra lo importante que son este tipo de compradores para el supermercado. En efecto, estos compradores realizan una mayor cantidad de compras que son en promedio más grandes y más variadas que las de compradores de otros conglomerados de categorías.

Por otro lado, el conglomerado de productos de higiene se encuentra en una posición opuesta al conglomerado de productos frescos/consumo inmediato. En efecto, un número relativamente pequeño de tickets incluye este tipo de productos, sin embargo, estos tickets son en promedio más grandes en lo que se refiere al monto total de la compra.

Finalmente, las categorías del conglomerado “otros productos” figuran en tickets con poca variedad de productos, es decir, tickets con un bajo número de categorías diferentes. Además, las compras que incluyen este tipo de productos contribuyen en baja medida a las ventas totales del supermercado en análisis.

6. Implicancias del análisis para la gestión coordinada de categorías (Cross Category Management)

Como se mencionó en la introducción, una categoría debe incluir a todos aquellos productos altamente interrelacionados entre sí pues es esta propiedad lo que le permite a un supermercadista administrar sus categorías como unidades estratégicas de negocios. Hasta ahora, esta interrelación ha sido típicamente asociada a sustitución. Sin embargo, existen también interrelaciones asociadas a la composición de la canasta de compra las cuales pueden ser detectadas, por ejemplo, a partir de la estimación de la probabilidad condicional de compra simultánea de un par de categorías o razón de similaridad, tal como se ha mostrado en este artículo.

La existencia de este tipo de interrelaciones sugiere que la coordinación de los esfuerzos de marketing concernientes a las categorías de un mismo conglomerado debiera ser altamente conveniente. De acuerdo a esto, se describirán en esta sección las implicancias del análisis presentado para la gestión coordinada de categorías. Específicamente se discutirán aspectos relativos a definición del rol de una categoría, decisiones tácticas (precio, promoción, surtido, espacio), diseño de la sala de ventas (lay-out) y evaluación de desempeño.

6.1 Definición de rol

La asignación de rol a cada una de las categorías es uno de los pasos más importantes en cualquier proceso de administración de categorías (Blattberg, Purk 1995). Esta asignación permite alinear la gestión de las categorías con la estrategia global del supermercadista. Las consecuencias de esta definición se relacionan directamente con las políticas de surtido, espacio, precio y promoción. Al respecto, los supermercadistas seleccionan un número limitado de categorías para cada rol, como por ejemplo, generador de tráfico, generador de transacciones, generador de utilidades, generador de imagen, generador de caja. Al realizar esta asignación de categorías es conveniente tener en cuenta la existencia de conglomerados de categorías.

En efecto, cada conglomerado puede ser considerado como un negocio con bastante independencia de acuerdo a las probabilidades de compra conjunta. Es interesante notar que esta segmentación se manifiesta también en la existencia de tiendas que se especializan en productos de alguno de estos conglomerados (por ejemplo: tiendas de abarrotes, almacenes, farmacias/perfumerías y panaderías que también comercializan productos lácteos y bebidas). De esta forma, es posible identificar las consecuencias de concentrar un cierto rol en categorías de un mismo conglomerado o negocio. Por ejemplo, si todas las categorías asignadas al rol de generación de tráfico pertenecieran al conglomerado de productos no perecibles, se fortalecería significativamente la afluencia de compradores de este tipo de productos a la sala de ventas mientras que las ventas de las categorías de otros conglomerados o negocios estarían en severo riesgo.

La estrategia alternativa es diseñar un mix de roles en cada conglomerado significativo (no necesariamente el mismo mix) para obtener una adecuada combinación de tráfico, utilidades, caja e imagen para cada uno de ellos. Adicionalmente, el análisis de variables descriptoras (ver sección 5.3) provee una fotografía del rol que cada categoría desempeña desde un punto de vista transaccional y financiero⁹. Esto puede ser usado para determinar prioridades entre los distintos conglomerados y para definir el adecuado balance de roles al interior de cada uno de ellos. Esta definición del balance al interior de cada segmento se puede entender como la asignación de un Super-Rol a cada conglomerado o negocio, lo cual debe estar alineado con el posicionamiento de la sala de ventas respectiva.

6.2 Decisiones tácticas

Las decisiones tácticas de la Administración por Categorías corresponden básicamente a las decisiones de precio, promociones, surtido y asignación de espacio para los distintos productos. En este sentido, el análisis exploratorio presentado puede ser usado por los supermercadistas para detectar oportunidades de promociones cruzadas definiendo combinaciones de alto y bajo margen para conjuntos de productos que se encuentren graficados en posiciones cercanas. Esto es válido, por ejemplo, para pares de categorías como: cereales y jugo en polvo; toallas higiénicas y shampoo y bálsamo; y, panadería y queso (ver gráfico 5.1). Un argumento similar es válido para las demás decisiones tácticas.

6.3 Diseño del lay-out de la sala de ventas

El análisis presentado también entrega lineamientos para el diseño del lay-out de la sala de ventas. Por ejemplo, si el objetivo del supermercadista es ubicar su oferta de productos de modo tal de minimizar el tiempo total que le toma a un comprador completar su canasta, los productos debieran ser ubicados teniendo en cuenta las probabilidades de compra simultánea. Es decir, aquellos conjuntos de productos que aparecen cerca unos de otros en el gráfico 5.1, como por ejemplo, aquellos que pertenecen a un mismo conglomerado, debieran ser ubicados en posiciones cercanas en la sala de ventas. Esto, sin embargo, también es válido para algunos productos que se encuentran en posiciones alejadas en el gráfico, para los cuales se pueda verificar además la existencia de una relación de sustitución entre ellos. En efecto, esto podría ser válido para categorías como bebidas gaseosas y cerveza que se encuentran en posiciones relativamente alejadas en el gráfico mencionado.

Por otro lado, si el objetivo del supermercadista fuese que el comprador de una determinada categoría recorriera un gran número de pasillos y góndolas de

9 Nuevamente, la incorporación de información relativa a margen bruto podría enriquecer en forma importante este análisis y sus implicancias.

la sala de ventas, podrían ubicarse las categorías con mayor probabilidad de ser incluidas en la canasta de este comprador en distintos sectores de la sala “forzando” un recorrido más largo para completar la canasta del comprador en cuestión. Evidentemente, esto puede tener consecuencias negativas en las actitudes del comprador hacia el supermercado como calidad de servicio y satisfacción, por lo cual estamos en presencia de un trade-off que debe ser evaluado. En particular, esta sala de ventas tiene una orientación hacia un formato de tienda de conveniencia, por lo cual, se deberían ubicar los productos privilegiando la comodidad del comprador.

6.4 Evaluación de desempeño

En lo que se refiere a evaluación de desempeño, la contabilidad tradicional no toma en cuenta los efectos que las decisiones y esfuerzos de marketing en una categoría pueden tener en las cifras de otra categoría (Chen et al 1999). La representación espacial presentada permite identificar fácilmente un grupo importante de categorías que pueden resultar afectadas después de la implementación de una decisión de marketing en una categoría. Estas categorías son aquellas que exhiben una probabilidad alta (en términos relativos) de ser incluidas en la misma compra, es decir, aquellas ubicadas en posiciones cercanas en la representación espacial. Además, tal como se mencionó anteriormente, si existe evidencia significativa a favor de una relación de sustitución entre dos categorías ubicadas en posiciones lejanas en la representación espacial, éstas también debieran ser consideradas.

En conclusión, la evaluación de la implementación de una decisión de marketing sobre una categoría debiera considerar no sólo los efectos en la misma categoría sino también los efectos cruzados en las categorías interrelacionadas (ventas inducidas). De acuerdo a esto, tiene sentido definir indicadores del desempeño global del conjunto de categorías de un conglomerado. Por ejemplo, la razón entre el Margen Bruto y el espacio asociado a un conglomerado de categorías es un indicador que permite evaluar la asignación eficiente del espacio a cada negocio o conglomerado de categorías. Esta misma razón estimada al nivel de una categoría no es definitivamente un buen indicador por sí sólo, pues no considera la existencia de interrelaciones entre los productos de distintas categorías. Bajo este enfoque miope no tendría sentido que una categoría no generara un significativo margen al supermercadista, lo cual puede ser beneficioso en la medida en que esto permita sostener las ventas de categorías interrelacionadas. Este problema se aísla en la medida en que el indicador de desempeño se calcule para un conjunto de productos relativamente independientes del resto.

7. Conclusiones del Estudio

El propósito de este trabajo fue realizar un análisis de las interrelaciones entre las categorías de productos en la canasta de compra. Este objetivo se cumplió gracias al aprovechamiento de la información proveniente del punto de venta y a la utilización de herramientas de análisis de datos ampliamente aplicadas a problemas de marketing y de otras especialidades.

Como resultado, se obtuvo una representación espacial de las categorías de acuerdo a las probabilidades condicionales de compra simultánea. Esto permitió identificar cuatro canastas de productos que tienden a ser incluidos en la misma compra con mayor probabilidad. Además, a partir del análisis fue posible caracterizar cada una de estas canastas de acuerdo a variables transaccionales lo cual proporciona una visión de la importancia que cada canasta representa para el negocio del supermercadista.

Finalmente, las interrelaciones detectadas sugieren que la coordinación de actividades de marketing concernientes a las categorías de un mismo conglomerado debiera ser altamente conveniente. De acuerdo a esto, se discutieron las implicancias del análisis presentado para la gestión coordinada de categorías, específicamente, en términos de definición del rol de una categoría, decisiones tácticas, diseño de la sala de ventas (lay-out) y evaluación de desempeño.

8. Investigación Futura y Extensiones

Investigación futura concerniente a este tipo de análisis debiera incluir aspectos de heterogeneidad de clientes. Si bien datos demográficos y socioeconómicos pueden servir para detectar este tipo de heterogeneidad, se debe tener en cuenta que el mismo comprador no se comporta de igual modo en distintas situaciones de compra. Tal como se mostró en este trabajo, los datos de scanner proveen información para segmentar a los compradores de acuerdo a la composición de sus canastas de compra. Esta segmentación junto con información demográfica y socioeconómica puede permitir mejorar los modelos de respuesta de los compradores a diferentes esfuerzos de marketing en distintas ocasiones de compra.

Adicionalmente, un aspecto pendiente que no se alcanzó a cubrir es analizar la estabilidad del análisis. Cabe recordar que en esta investigación se recurrió a información transaccional de solamente un mes de ventas del supermercado. Al respecto, sería interesante analizar si la configuración espacial exhibe estabilidad al extender este análisis a otros meses o incluso a los datos de todo un año. Por otro lado, la aplicación de este análisis en salas de ventas de otros formatos (megamercado, hipermercado, etc.) permitiría comparar el comportamiento de compradores de distintas salas de supermercado.

Finalmente, el tipo de análisis presentado puede ser utilizado más allá de las fronteras del negocio supermercadista. Esto se puede entender fácilmente teniendo en cuenta la existencia de clientes que compran o contratan canastas de productos o servicios en una visita real o virtual a un proveedor o que realizan una serie de actividades (como visitar distintas páginas web, bajar información, entregar información, comprar productos) en una visita a un portal de internet. De esta forma empresas como tiendas por departamento, bancos y portales de internet podrían hacer un uso provechoso de este tipo de análisis. Por ejemplo, un portal podría construir un mapa para representar gráficamente la “canasta” o “conjunto” de páginas o actividades (por ejemplo, compras, archivos descargados, etc.) que tienen probabilidades más altas de ser visitadas en la misma conexión al portal en cuestión. Esto podría ser utilizado para la identificación de segmentos de visitantes de acuerdo a su comportamiento de visita lo cual complementado con información adicional acerca de los clientes (por ejemplo, comportamiento pasado, información socioeconómica) permitiría una adecuación coordinada de las páginas visitadas o servicios requeridos por los clientes de un determinado segmento.

Más aún, el uso de probabilidades condicionales como las mencionadas en las ecuaciones 2.1 y 2.2 podría ser más potente desde un punto de vista interactivo. Es decir, cuando nos referimos a portales dinámicos o inteligentes que se adaptan en su contenido y publicidad de acuerdo a la información que se tiene del visitante. Esta información no necesariamente debe ser determinística (por ejemplo, aunque no se conozca el sexo del visitante con certeza, el sólo hecho de conocer la probabilidad de este evento puede ser valioso. Un ejemplo “bayesiano” interesante en este sentido es mencionado por Montgomery (2000)).

Referencias bibliográficas

- A.C. Nielsen (2000), “Categorización de Productos”, Santiago: A.C. Nielsen (CD-ROM).
- Andrews, R. y A. Manrai (1999), “MDS Maps for Product Attributes and Market Response: An application to Scanner Panel Data”, *Marketing Science*, Vol. 18, No. 4, pp. 584-604.
- Bell, D. y J. Lattin (1998), “Grocery shopping behavior and consumer preference for retail price format: why ‘large basket’ shoppers prefer EDLP”, *Marketing Science*, Vol 17, pp. 66-88.
- Blattberg, R. y M. Purk (1995), “Administración de Categorías: Funciones de Otras Organizaciones en la Administración de Categorías”, Washington, DC: Food Marketing Institute.
- Bosch, M., C.Espinoza, A. Musalem, M. Vega y C.Paz (2000), “Experiencias de Administración por Categorías en Chile”, Documento de Trabajo, Departamento de Ingeniería Industrial, Universidad de Chile.

- Bosch, M. y A. Musalem (2000), "Category Management: An Operations Research View", X Latin-American Conference on Operations Research and Systems, Septiembre, Ciudad de Mexico (CD-ROM).
- Chen, Y., J. Hess, R. Wilcox y Z.J. Zhang (1999), "Accounting Profits versus Marketing Profits: A Relevant Metric for Category Management", *Marketing Science*, Vol. 18, No. 3, pp.
- Elrod, T. (1988), "Choice Map: Inferring a Product-Market Map from Panel Data", *Marketing Science*, Vol. 7, No. 1, pp. 21-40.
- Fader, P. y L. Lodish (1990), "A Cross-Category Analysis of Category Structure and Promotional Activity for Grocery Products", *Journal of Marketing*, Vol. 54, No. 4, pp. 52-65.
- Ghose S. (1998), "Distance Representations of Consumer Perceptions: Evaluating Appropriateness by Using Diagnostics", *Journal of Marketing Research*, Vol. XXXV, May 1998, pp. 137-153.
- Green P. (1978), "Analyzing Multivariate Data", Hinsdale, Illinois: The Dryden Press.
- Manchanda, P., A. Ansari y S. Gupta (1999), "The Shopping Basket: A Model for Multicategory Purchase Incidence Decisions", *Marketing Science*, Vol. 18, No. 2, pp. 95-114.
- Montgomery, A. (2000), "Applying Quantitative Marketing Techniques to the Internet", Working Paper, Graduate School of Industrial Administration, Carnegie Mellon University.
- Mulhern, F. Y R. Leone (1991), "Implicit price bundling of retail products: a multiproduct approach to maximizing store profitability", *Journal of Marketing*, Vol. 55, pp. 63-76.
- Nielsen Marketing Research (1992), "Category Management: positioning your organization to win", Chicago: NTC Business Books.
- Sinha, I. y W. DeSarbo (1998), "An Integrated Approach Toward the Spatial Modelling of Perceived Customer Value", *Journal of Marketing Research*, Vol. XXXV, May 1998, pp. 236-249.
- SPSS Inc. (1999), "SPSS Base 9.0: Applications Guide", Chicago, Illinois: SPSS Inc.
- Urban, G. y J. Hauser (1993), "Design and Marketing of new Products", New Jersey: Prentice Hall.