



UNIVERSIDAD DE CHILE

**FACULTAD DE CIENCIAS ECONÓMICAS Y
ADMINISTRATIVAS**

**MODELOS DE MARKETING CON DATOS DE
PANEL SCANNER**

Seminario para optar al Título de Ingeniero Comercial

Mención Administración

Alumna: Carolina Faviola Contreras Pérez

Profesores Guías: Enrique Manzur Mobarec

Sergio Olavarrieta Soto

Santiago de Chile, Julio de 2003

A mi familia



Sería una ingrata si es que no diera las gracias a todas las personas que hicieron posible que yo realizara este trabajo.

Agradezco a mis profesores guías, al profesor Sergio Olavarrieta, por sus aportes y especialmente al profesor Enrique Manzur, quien siempre estuvo ayudándome y sin quien yo no hubiera podido hacer este trabajo.

Agradezco y dedico este trabajo a quienes son los más importante para mí, a mi pilar, mi mamá y mi fuerza, mi papá.

No puedo olvidar a personas e instituciones que desinteresadamente colaboraron con mi trabajo, a la empresa ACNielsen, quienes me proporcionaron los datos y ayuda a Bárbara Delso, a Marianne Osorno y a Claudio Quiroz.

Finalmente quiero agradecer a mi familia en general y amigos, por su apoyo comprensión y paciencia.

ÍNDICE

Introducción	7
Capítulo I: Modelos matemáticos con datos de panel scanner	9
I. 1 Datos de panel scanner y segmentación.	11
I. 2 Un acercamiento simultáneo para segmentación y estructura de Mercado por Grover y Srinivasan.	13
I. 2. 1 Introducción	12
I. 2. 2 Modelo	13
I. 2. 3 Identificación	17
I. 2. 4 Grados de libertad y condición de identificación	17
I. 2. 5 Bondad de ajuste	18
I. 2. 6 Número de segmentos	20
I. 2. 7 Resumen de pasos a seguir para estimar el modelo (Una ilustración)	21
I. 3 Estimación y uso de errores estándar de parámetros de modelos latentes por Rajiv Grover	26
I. 3. 1 Introducción	26
I. 3. 2 Clase de Modelos latentes	26
I. 3. 3 Estimación de errores estándar en el set básico	27
I. 3. 4 Estimación de errores estándar de los parámetros omitidos del set básico	28
I. 3. 5 Implementación	29
I. 4. Una aproximación para rastrear cambios dentro del segmento en el mercado por Grover y Srinivasan.	30
I. 4. 1 Introducción	30

I. 4. 2 Modelo	31
I. 4. 3 Estimación	33
I. 4. 4 Una ilustración	34
I. 4. 4. 1 Formulación de variables predictivas	35
I. 4. 4. 2 Determinación del segmento de mercado	36
I. 4. 4. 3 Test de consistencia de los tamaños de los segmentos	37
I. 4. 4. 4 Rastreo de Modelos dentro de segmentos de cambio	37
I. 5 Un modelo de elección probabilística para segmentación de mercados y elasticidad de estructura por Kamakura y Rusell	38
I. 5. 1 Introducción	38
I. 5. 2 Un modelo de segmentación preferente	38
I. 5. 3 Utilidad aleatoria	39
I. 5. 4 Segmentación de Preferencias	40
I. 5. 5 Estimación	41
I. 5. 5. 1 Representación de la estructura competitiva	41
I. 5. 5. 2 Derivación de la agregación	41
I. 5. 6 Aplicación del modelo	42
I. 5. 6. 1 Descripción de los datos.	42
I. 5. 6. 2 Definición de lealtad	43
I. 5. 6. 3 Elección del número de segmentos	43
I. 5. 6. 4 Segmentación de preferencias	44
I. 5. 6. 5 Estructura por elasticidad	44
I. 5. 6. 6 Corrección de factores	44
I. 5. 6. 7 Clout Competitivo y vulnerabilidad	45
I. 6 Entendiendo la competencia de marca usando micro y macro datos de panel scanner por Kamakura y Rusell	46

I. 6. 1	Introducción	46
I. 6. 2	Modelo de conducta de compra de hogares	46
I. 6. 3	El macro modelo: porción del volumen semanal	48
I. 6. 4	Ilustración Empírica	49
I. 6. 5	Microanálisis	50
I. 6. 6	Macroanálisis	50
	I. 7 Discusión	52
	Capítulo 2: Metodología de estudio	55
	II. 1 Introducción	56
	II. 2 Datos de Café	57
II. 2. 1	Descripción de los datos	57
II. 2. 2	Distribución de los datos	62
II. 2. 3	Elección de compras	71
II. 2. 4	Tabulación cruzada	72
II. 2. 5	Simetría de la matriz	75
II. 2. 6	Efecto Promociones	77
II. 2. 7	Estacionalidad	82
	II. 2 Datos de Detergentes	86
II. 2. 1	Descripción de los datos	86
II. 2. 2	Distribución de los datos	89
II. 2. 3	Elección de compras	93
II. 2. 4	Tabulación cruzada	94
II. 2. 5	Simetría de la matriz	95
	Capítulo 3: Resultado y análisis	98
	III. 1 Introducción	98

III. 2 Datos café	98
III. 2. 1 Determinación del número de segmentos de cambio	98
III. 2. 1. 1 Desarrollando primera opción	108
III. 2. 2. 1 Análisis de Resultado primera opción	118
III. 2. 1. 2 Desarrollando segunda opción	121
III. 2. 2. 2 Análisis de resultados de la segunda opción	124
III. 2. 3 Grados de libertad	26
III. 2. 4 Decisión de opción	28
III. 2. 5 El mercado en base a los resultados	129
III. 3 Datos de Detergentes	132
III. 3. 1. Determinación del número de segmentos de cambio	132
III. 3. 2. 1 Desarrollando primera opción	133
III. 3. 3. 1 Análisis de Resultado primera opción	142
III. 3. 2. 2 Desarrollando segunda opción	144
III. 3. 2. 3 Desarrollando tercera opción	150
III. 3. 3 Grados de libertad	151
III. 3. 4 Decisión de opción	152
III. 3. 5 El mercado en base a los resultados	156
Capítulo 4: Análisis de Percepción.	159
IV. 1 Percepción	159
IV. 2 Análisis Factorial	159
Capítulo 5: Conclusiones, limitaciones y estudios futuros.	169
Bibliografía.	172
Anexo	175

INTRODUCCIÓN

Durante la última década la riqueza y disponibilidad de datos panel escáner han incentivado un número creciente de estudiosos. Su meta común ha sido desarrollar modelos superiores de conducta del consumidor que habilitan a empresas, vendedores y otros para evaluar la sensibilidad del mercado a un rango de las estrategias de marketing, últimamente se han incluido estudios referentes al valor de marca y efectos de publicidad, extensiones de marca y otros sucesos que podrían afectar a diferentes productos del mercado

El trabajar con datos de panel scanner ha sido un gran avance para el marketing, ya que ha podido superar con ello todos los problemas y trabajo que traen consigo el hacer encuestas, coordinarlas y problemas de errores no muestrales de las mismas, en éste sentido para las empresas, los datos de panel scanner son un verdadero avance, ya que entre los beneficios que tiene el trabajar con ellos está la disminución de costos que podría tener hacer un estudio ocupándolos, por ejemplo de segmentación.

Los datos de panel scanner se consiguen mediante un lector que está en algunos hogares, los hogares son parte de una muestra representativa del mercado, la base de datos que entrega éste lector contiene información del hogar, del lugar de compra, del producto, etc. En fin información que puede ser de gran utilidad.

En éste estudio nuestro objetivo es desarrollar un modelo donde podamos usar los datos de panel scanner para segmentar, sabemos que la segmentación para las empresas son muy importantes, sobre todo hoy en día donde la competencia es muy fuerte y la mayoría de las empresas tienen productos para segmentos específicos, por lo que segmentar es

primordial para enfocar las estrategias de marketing, además como los datos de panel scanner reflejan la conducta real de compra de los consumidores podemos segmentar en base a esto y según el modelo que pretendemos utilizar obtener información sobre la estructura de mercado.

El trabajo se desarrollará de la siguiente forma:

En el Capítulo I, expondremos diversos modelos que utilizan datos de panel scanner y mostraremos además los avances que se han ido desarrollando en el tiempo con la utilización de datos de panel scanner .

En el Capítulo 2, comenzaremos el desarrollo de un modelo de segmentación con datos de panel scanner propuesto por Grover y Srinivasan, aquí mostraremos el trabajo previo que se hizo con los datos antes del modelo, mostraremos el proceso de selección de las marcas que entrarían en el estudio y la selección de compras y panelistas (personas que realizaron la compra).

En el Capítulo 3, desarrollaremos el modelo propuesto, para esto utilizaremos el programa Winmira 2001, con el cuál obtendremos el número óptimo de segmentos para luego desarrollarlos, identificando a los segmento y las marcas representativas de cada uno, es decir las marcas que con mayor probabilidad las personas pertenecientes a esos segmentos de cambio comprarían (segmentos de cambio ya que las personas que son leales a las marcas, compran siempre las mismas).

En el capítulo 4 haremos un análisis de los segmentos obtenidos utilizando características objetivas de los productos y características subjetivas de los mismo, dadas por un estudio de percepción realizado.

Finalmente en el capítulo 5 mostraremos las conclusiones del estudio, limitaciones del mismo y futuras aproximaciones y estudios que se podrían hacer tomando como base a éste seminario.

Capítulo I

Modelos matemáticos con datos de panel scanner

I .1 Datos de panel scanner y segmentación.

Los datos de panel escáner son datos de compra que se obtienen mediante un lector que escáner que lee códigos de barra, muy similar a los que podemos ver en supermercados y tiendas, además los datos entregan información muy valiosa sobre el hogar en el que se encuentra el lector, miembro del hogar que realizó la compra, y características del producto (precio, producto, marca, cantidad en unidades y tamaño del producto), el lugar en donde se realizo la compra, además de la fecha de compra, etc. El segmento con datos de panel scanner tiene la ventaja que se evitan parte de los errores no muestrales en los que se incurre por sesgo de los encuestadores, etc. Cabe destacar que éste sistema no está eximio de errores, éstos se producen más que nada por la voluntad de las personas que tienen que pasar los productos comprados por el scanner, algunos de los problemas pueden ser por ejemplo que no se pasen los productos el día que se efectúo la compra, que no se identifique efectivamente al individuo que realizó la compra, etc.

Otra dificultad importante que presentan los datos de panel scanner es que un cambio en la compra de un hogar de una marca A a una marca B no implica necesariamente que A y B son sustitutos, el cambio de marca puede deberse a que hayan habido dos usuarios en el hogar o que era un producto que el usuario lo ocupaba para usos distintos, por lo que tampoco podían llamarse sustitutos.

Cabe destacar que los datos de panel no solo los podemos encontrar a nivel de hogares (nivel micro), si no que también los podemos encontrar a nivel de tiendas y

supermercados (nivel macro), es decir, el lugar en donde los consumidores adquieren los productos. En éstos lugares la forma de obtener los datos es similar, ya que todos ellos cuentan con scanner.

Por la información que éstos datos proveen existen numerosos estudios que se pueden hacer en base a éste tipo de datos entre los que se cuentan: conducta del consumidor, segmentación y estructura.

Para segmentar hay algunos modelos que utilizan las probabilidades de elección de marca. El utilizar probabilidades de elección de marca es un método de segmentar que utiliza como input de segmentación una compra en vez de utilizar variables geográficas, sicográficas o sociográficas como las que se suelen utilizar normalmente como criterio de segmentación, este tipo de análisis tiene como beneficio el que no clasifica a los consumidores por nada más que la elección de compra, sabemos que dos personas que tengan las mismas características, como ser de determinada edad, vivir en una misma área, podrían tener perfectamente conductas de consumo distintas, por lo tanto podríamos equivocarnos al segmentar por esas características y pretender esas dos personas posean conductas de consumo similares; lo que no ocurre si segmentamos tomando como input de segmentación la elección de producto y marca y luego de allí expliquemos esa conducta en base a análisis que podríamos hacer por ejemplo midiendo la sensibilidad de los segmentos a variaciones en elementos de marketing mix y como se influyen entre ellos, además podríamos evaluar las fuerzas que se crean en un mercado, competencia entre marcas, etc.

En éste sentido hay muchas investigaciones que han desarrollado modelos que ocupan datos de panel scanner para segmentar, estructurar mercados y ver las fuerzas que interactúan en ellos, ahora presentaremos algunos ejemplos de modelos que se han desarrollado ocupando datos de panel scanner provisto por la empresa ACNielsen.

I. 2 “Un acercamiento simultáneo para segmentación y estructuración de mercados”

Grover y Srinivasan

I. 2. 1 Introducción

Los autores definen un segmento como un grupo de consumidores que comparten la misma probabilidad de elección de una marca determinada, en donde cada segmento es homogéneo y entre ellos son heterogéneos. La estructura de competencia de mercado es definida como la superposición de grupos de marcas correspondiente a diferentes segmentos. En éste estudio utilizan los datos de compras de un número determinado de hogares, quienes eligen una marca de un set de marcas de un producto en dos ocasiones de consumo. El mercado es por esto dividido en segmentos leales y múltiples segmentos de cambio. Aplicaciones del modelo muestran que la aproximación aquí propuesta produce una mejora en la validación predictiva en comparación con las aproximaciones tradicionales de tratar al mercado como una agregación singular.

Con los datos de panel scanner los autores obtienen la matriz de clasificación cruzada de elección de marca en dos ocasiones de compra y extraen los segmentos usando el método de máxima verosimilitud para estimar la clase de modelos latentes. Una aplicación en el mercado de café instantáneo indican que la aproximación propuesta tiene validez sustancial y sugiere la presencia de submercados relativos a atributos de productos así como a nombres de marcas.

I. 2. 2 Modelo

Se asume que el proceso de elección de marca es estacionario, es decir de orden cero, es decir que la elección de la primera compra no está influenciado por consumo pasado, por esto el espacio de tiempo en el que se hace el estudio, debe ser muy corto. Además el modelo postula que cuando un consumidor se cambia de marca se va a cambiar a una marca que tenga una probabilidad parecida de elección.

Es importante destacar que estas estimaciones son útiles solo para el género de compras no durables para los cuales tal clasificación está disponible en los datos de panel.

Denotaron “n” como el número de marcas en cada clase de marcas de interés. Se asume que la heterogeneidad está dada entre los segmentos leales ($l=1,2,\dots,n$) y los segmentos de cambio ($M= 1,2 , \dots,m$). Dados por las ecuaciones 1 y 2 respectivas

$$1) s_{ijl} = \begin{cases} 1 & \text{si } i = j = l \\ 0 & \text{de otro modo.} \end{cases}$$

$$2) p_{ik} \geq 0 \quad \text{para } k = 1, 2, \dots, m.$$

Y

$$3) \sum_{i=1}^m p_{ik} = 1 \quad \text{para } k = 1, 2, \dots, m.$$

Los consumidores fieles siempre comprarán la marca “i” con probabilidad igual a 1 durante el periodo de tiempo considerado en el estudio. Así no hay heterogeneidad en el segmento en éste caso. La proporción mostrada en la ecuación 1, s_{ijl} muestra la proporción de consumidores que compra la marca “i” en la primera ocasión y que compra la marca “j” en la siguiente ocasión de compra, según las restricciones a las que se somete a la ecuación 1 representa a los segmentos leales los que poseen probabilidad 1 (compran la misma marca en ambas ocasiones).

En la ecuación 2 tenemos al segmento de cambio k, y tenemos p_{ik} denotando al probabilidad de elegir la marca “i”. Así representa la porción de mercado que elige la marca “i” y pertenece al segmento k. Por lo que la ecuación 4 describe la porción del segmento de cambio que elige “i” en la primera ocasión de compra y “j” en la segunda.

$$4) s_{ijl} = p_{ik} * p_{jk}$$

Además se define a V_l y W_k como las porciones de mercado correspondiente al total de los segmentos leales y al total de los segmentos cambiantes respectivamente.

$$5) V_l \geq 0 \quad \text{para } l = 1, 2, \dots, n.$$

$$6) W_k \geq 0 \quad \text{para } k = 1, 2, \dots, m.$$

$$7) \sum_{l=1}^n V_l + \sum_{k=1}^m W_k = 1$$

En donde uniendo las ecuaciones anteriores obtenemos :

$$8) S_{ij} = \sum_{l=1}^n V_l * s_{ijl} + \sum_{k=1}^m W_k * s_{ijk} \quad \text{para todo } i \neq j$$

En donde S_{ij} representa la proporción teórica de los consumidores en el mercado completo que compran la marca i en una ocasión y una marca diferente j en otra ocasión.

Usando el s_{ijk} y substituyendo la ecuación 4 en la ecuación 9, nosotros tenemos

$$9) S_{ij} = \sum_{k=1}^m W_k * p_{ik} * p_{jk} \quad \text{para todo } i \neq j$$

La ecuación 10 es la que demuestra que la matriz es simétrica.

Por simplicidad crearemos un nuevo subíndice h el cual engloba todos los segmentos del mercado, es decir, los segmentos leales mas los segmentos de cambio, de la forma $h = 1, 2, \dots, n+m$; y el parámetro β_h remplazara a los parámetros anteriores, finalmente tendremos:

$$10) S_{ij} = \sum_{h=1}^{m+n} \beta_h * p_{ih} * p_{jh} \quad i, j = 1, 2, \dots, n.$$

Se extienden las restricciones antes mencionadas para los parámetros V_l y W_k en donde el parámetro debe ser positivo y la sumatoria de éste debe ser igual a 1. Este modelo permite que usemos el método de estimación de modelos de clase latentes.

La idea es que se considere una tabulación cruzada con p_i filas y p_j columnas y las proporciones S_{ij} corresponden a la celda (i, j). La clase de modelos latentes postulan que las observaciones dependientes entre las filas y las columnas en la tabla es debido a que se está bajo clases latentes y están bajo es supuesto de independencia local (de aquí viene el supuesto de orden cero, de que el proceso de elección de los consumidores corresponde al supuesto de independencia local).

En el contexto presente el número de filas es igual al número de columnas todas denotando al mismo número de marcas “n” y β_h representa a la porción de los consumidores que pertenecen al segmento “h”.

Los parámetros β_h y p_{ih} pueden ser estimados por el método de máxima verosimilitud usando el proceso iterativo desarrollado por Goodman. El error estándar de los parámetros puede obtenerse según el procedimiento utilizado por Grover (1985).

I. 2. 3 Identificación

La identificación en modelos latentes es análoga a la identificación en sistemas de ecuaciones simultáneas. La identificación radica en la idea de que no pueden dos números de parámetros simultáneos estimando a un mismo S_{ij}

Para identificar podemos cerrar o hacer tender a cero algunos parámetros p_{ih} (solo para los $i \neq j$), esto se hace principalmente bajo dos criterios.

- 1) Conocimiento previo del producto de mercado, o con un análisis exploratorio.
- 2) Parámetros con valores iguales o bajo 0.01

La segunda alternativa al cerrar los parámetros a 0 presenta luego un chi-cuadrado casi inalterados, por lo que se cree es la mejor alternativa.

Se usa la letra “t” para denotar el total de parámetros restringidos a cero para poder identificar el modelo.

I. 2. 4 Grados de libertad y la condición de identificación

Los grados de libertad para el modelo estará dados por la siguiente ecuación :

11)
$$d.f = n(n - m - 1) + t$$

Una condición necesaria para la identificación es que los grados de libertad no sean negativos. La condición necesaria y suficiente para la identificación es que la matriz formada tomando las derivadas parciales de cada S_{ij} no redundantes (las filas) con

respecto a cada uno de los parámetros estimados (Columnas de la matriz), tenga completo el rango de las columnas.

I. 2. 5 Bondad de ajuste

La bondad de ajuste del modelo puede ser evaluado por el estadístico chi- cuadrado de la siguiente forma.

$$12) \quad \chi^2 = 2N \sum_{i=1}^n \sum_{j=1}^n S \ln(S_{ij} / \hat{S}_{ij})$$

Donde N es el tamaño de la muestra. El estadístico chi-cuadrado es directamente proporcional al tamaño de la muestra. Consecuentemente, un buen modelo puede ser rechazado cuando la muestra es de gran tamaño. Un modelo que no es afectado por el tamaño de la muestra es el R^2

$$13) \quad R^2 = \frac{\chi^2(1) - \chi^2(m+n)}{\chi^2(1)}$$

Donde $\chi(m+n)$ es el chi-cuadrado de los segmentos $m+n$ y $\chi(1)$ es la agregación del mercado. Un problema de ésta ecuación es que aumenta cuando incluimos segmentos, por lo que un mejor ajuste está dado por el R^2 ajustado.

$$14) \quad R^2 = \frac{\frac{\chi^2(1)}{d.f.(1)} - \frac{\chi^2(m+n)}{d.f.(m+n)}}{\frac{\chi^2(1)}{d.f.(1)}}$$

Vemos la ecuación 16 que representa el índice de la desigualdad entre las matrices S y \hat{S} .

$$15) \quad \Delta = (1/2) \sum_{i=1}^n \sum_{j=1}^n |S_{ij} - \hat{S}_{ij}|$$

Como se comentó antes una implicación teórica del modelo propuesto es que la matriz de cambio es simétrica. Consecuentemente, la asimetría en la matriz de cambio \hat{S} impone un más bajo límite adelante sobre Δ . El valor mínimo posible Δ_{\min} . Será obtenido poniendo \hat{S} para ser la matriz simétrica dado por $(S + S')/2$, donde S' es el traspuesto de S de la ecuación anterior. Entonces tenemos:

$$16) \quad \Delta_{\min} = (1/4) \sum_{i=1}^n \sum_{j=1}^n |S_{ij} - S_{ji}|$$

Así una medida (entre 0 y 1) de la mejora obtenida considerando el $(n + m)$ solución de segmentos en comparación con tratar al mercado como una agregación única dada por :

$$17) IMP = \frac{M - \Delta(n + m)}{M - \Delta_{\min}}$$

En R^2 y R^2 ajustado junto con el IMP son medidas que se obtienen por segmentación.

I. 2. 6 Número de segmentos

Normalmente se tiene muy poca información a priori sobre el número de segmentos de cambio. En la ausencia de un número de segmentos previo, proponen que el número de segmentos debe determinarse por un proceso similar al test de Cattell (1966) para determinar el número de factores retenidos en el factor de análisis. En el presente contexto la aproximación envuelve estimaciones del modelo para diferentes números de segmentos de cambio y trazando la bondad de ajuste del R^2 ajustado o el IMP contra el número de segmentos “m”. El número de segmentos es escogido de tal manera que el número de segmentos propuestos de, el mejor R^2 ajustado. El proceso se hace de la siguiente forma, se estima el R^2 comenzando con un segmento de cambio y aumentando de uno en uno hasta llegar a un número determinado. Se tomará como el número óptimo de segmentos de cambio aquel que dé el R^2 ajustado más alto y que la inclusión de nuevos

segmentos de cambio haga aumentar el R^2 ajustado poco significativamente (que el aumento marginal sea mínimo o cero)

I. 2. 7 Resumen de pasos a seguir para estimar el modelo (Una ilustración)

Se ocupa como ejemplo el mercado de café instantáneo. De la muestra total fueron eliminadas las marcas que tenían pocas categorías de productos y poca participación de mercado. Para asegurar la independencia estadística de las observaciones de las tablas de clasificación cruzada, usaron los datos de las primeras dos compras.

La muestra utilizada era de 4657 hogares, los que fueron divididos aleatoriamente en tres submuestras.

La primera muestra fue desarrollada como ejemplo exploratorio para decidir el número de segmentos de cambio “m” y los parámetros a ser restringidos a 0. La segunda muestra fue la elegida para realizar la estimación. Aquí los parámetros restringidos a cero eran los determinados por la muestra desarrollada anteriormente, ahora se puede evaluar la significancia estadística ya que los valores necesarios fueron obtenidos anteriormente. La tercera muestra fue utilizada para ver la validez.

A.- Primero se debe verificar que la matriz S sea simétrica, por lo que se utiliza la ecuación 10, antes descrita, para con ella realizar una matriz de tabulación cruzada. El test chi-cuadrado es usualmente considerado como válido cuando la frecuencia esperada es ≥ 5 .

Tabla 1 : Tabulación cruzada

Marcas	HP-D-R	TC-C-FD	TC-D-FD	FL-C-R	MH-C-R	S-D-R	S-D-FD	MX-C-FD	N-C-R	N-D-R	B-D-FD
HP-D-R	93	7	17	19	18	43	1	4	6	7	10
TC-C-FD	9	80	12	11	24	7	4	2	6	3	3
TC-D-FD	9	14	46	3	7	7	4	2	2	0	9
FL-C-R	19	18	4	82	29	14	0	4	9	2	6
MH-C-R	26	11	6	35	184	24	3	11	18	6	6
S-D-R	15	13	8	13	28	127	4	3	3	8	8
S-D-FD	2	0	3	2	1	7	17	3	0	1	4
MX-C-FD	4	3	4	3	6	5	2	27	1	0	4
N-C-R	5	3	2	4	16	4	0	1	46	9	2
N-D-R	6	4	4	1	5	9	0	0	11	15	2
B-D-FD	10	1	4	4	2	10	2	2	5	2	27
Particip	0,136	0,101	0,069	0,117	0,209	0,157	0,025	0,038	0,064	0,035	0,049

Donde :

Siglas	Marca
HP-D-R	High Point Descafeinado Regular
TC-C-FD	Taster's Choice Cafeinado Helado Seco
TC-D-FD	Taster's Choice Descafeinado Helado Seco
FL-C-R	Folgers Cafeinado Regular
MH-C-R	Maxwell House Cafeinado Regular
S-D-R	Sanka Descafeinado Regular
S-D-FD	Sanka Descafeinado Cafeinado Helado Seco
MX-C-FD	Maxim Cafeinado Helado Seco
N-C-R	Nescafé Cafeinado Regular
N-D-R	Nescafé Descafeinado Regular
B-D-FD	Brim Descafeinado Helado Seco

B .-Luego se estima el número de segmentos , leales y de cambio. Aquí tenemos que decir que los segmentos leales son todos aquellos que tengan una alta probabilidad de elegir la misma marca en dos ocasiones de consumo, es

decir , que la probabilidad de elección sea cercana a 1, ésta definición podría incluir algunos segmentos antes definidos como de cambio. Aquí lo que se debe hacer es cerrar a 1 todo segmento que tenga una probabilidad alta cercana a 1. En cuanto a los segmentos de cambio, éstos se obtendrán por el proceso antes descrito.

C.- Se debe medir la bondad del ajuste.

D.- Una vez obtenido todos los datos requeridos (t , l , m) se debe obtener los grados de libertad, los que deben arrojar un número positivo y obtener el chi-cuadrado.

Una vez obtenido todos los datos y resultados anteriormente mencionados se debe proceder a hacer los análisis correspondientes a los segmentos e identificar así la estructura de competencia de mercado.

Tabla 2: Resultados

Marcas	Particip. Mmdo.	Paramet.(V)	Segmentos de Cambio(W)			
			0,19 Segmento1	0,22 Segmento 2	0,18 Segmento 3	0,06 Segmento 4
HP-D-R	0,01	0,05	0,09	0,20	0,13	0,08
TC-C-FD	0,01	0,04	0,07	0,03	0,18	0,03
TC-D-FD	0,07	0,01	-	-	0,32	0,12
FL-C-R	0,12	0,04	0,20	0,16	0,04	-
MH-C-R	0,21	0,08	0,42	0,15	0,07	0,06
S-D-R	0,16	0,07	0,04	0,22	0,11	0,15
S-D-FD	0,03	0,01	-	0,05	0,00	-
MX-C-FD	0,04	0,01	0,04	0,03	0,04	-
N-C-R	0,06	0,01	0,14	-	-	0,27
N-D-R	0,03	0,01	-	0,07	-	0,27
B-D-FD	0,05	0,02	-	0,09	0,08	0,02
	1,00	1,00	1,00	1,00	1,00	1,00

En la tabla 2 podemos apreciar que el mercado del café en Estados Unidos está compuesto principalmente por 4 segmentos de cambio, los que poseen un 19%, 22%, 18%, y 6% de la participación del mercado, en total suman un 65%. El restante 35% corresponde a las 11 marcas leales.

Cada segmento de cambio puede ser interpretado identificando a las marcas que tienen una alta probabilidad de elección por el segmento.

El segmento 1 parece ser el segmento Cafeinado Regular, donde domina la marca Maxwell House con una participación del 42%, seguido por Folgers (20%) y Nescafé Cafeinado Regular (14%). El segmento 2 parece ser el

segmento Descafeinado Regular, las dos marcas con mayor participación son Sanka Descafeinado Regular (22%) y High Point (20%), ambas son marcas descafeinadas regulares. No obstante, las dos marcas con mayor participación del segmento 1, Maxwell House y Folgers, cada una tiene un 15% del segmento 2 por esto el segundo traslapa al primer segmento. El tercer segmento aparece como el segmento Taster's Choice con las dos categorías que encierra ésta marca sumando un 50% del total del segmento. Similarmente, el segmento 4 aparece como un segmento de Nescafé, con las dos categorías de Nescafé sumando el 54%.

I. 3 “Estimación y uso de errores estándar de parámetros de modelos latentes”

Rajiv Grover

I. 3. 1 Introducción

El uso de los modelos de clase latente se ha hecho muy popular en los últimos tiempos, Goodman (1974) implementó un algoritmo de ajuste iterativo en el programa MLLSA para estimar los parámetros de los modelos latentes. No obstante, el MLLSA no provee los errores estándar de la matriz de verosimilitud estimada.

La utilidad de los errores estándar está dada por que éstos pueden ser utilizados en la identificación de un modelo, es decir, podemos ocupar el valor obtenido de la división de el parámetro con su error estándar respectivo, valor

al que denominaremos “Z”. Así, diremos que si el valor Z es pequeño, es decir, cercano a 0.01, debe ser restringido a 0. En éste sentido uno puede usar el chi-cuadrado u otro modelo que mida el ajuste, para ver si efectivamente éste mejora con la restricción de parámetros.

I. 3. 2 Clase de modelos latentes

Este modelo se explica principalmente por la ecuación principal vista en el estudio anterior, en donde existen parámetros que son probabilidades de ocurrencia de un acontecimiento y que son similares para entes de una misma clase.

$$18) S_{ij} = \sum_{k=1}^m W_k * p_{ik} * p_{jk}$$

Los parámetros p_{ik} y p_{jk} son denominados como “parámetros del set básico” y consiste en los parámetros no redundantes de el modelo. El set básico puede ser reducido por la imposición de restricciones sobre los parámetros. Generalmente, los parámetros pueden ser restringidos a una constante y así no son estimados, luego esto hacer que disminuya el número de parámetros a estimar.

I. 3. 3 Estimación de errores estándar en el set básico

Bishop, Fienberg, y Holland dan la base general para la estimación de la matriz de varianza-covarianza de la estimación de maximaverosimilitud de parámetros para un modelo multidimensional.

Brevemente,, si J es la matriz Jacobiana de las derivadas parciales de Π_{ij} con respecto a los parámetros no redundantes del modelo, define la matriz A como:

$$19) A = D^{-1/2} * J$$

En la ecuación 19, D es la diagonal de la matriz , por lo tanto la matriz de varianza-covarianza para una muestra de tamaño “N” está dada por:

$$20) V = \frac{1}{N} (A^T * A)^{-1}$$

En éstas ecuaciones se encuentra contenida la varianza de los parámetros, por lo que la raíz cuadrada de la varianza son el error estándar con los cuales la significancia de esos parámetros puede ser establecida. No

obstante éste procedimiento no provee información del error estándar de los parámetros restringidos a ser constantes ni sobre W_M , p_{nk} y p_{mk} para todo k .

Como ya antes habíamos dicho, los parámetros restringidos no son estimados y todos son enviados a una columna del Jacobiano la que representa a todos los parámetros restringidos.

I. 3. 4 Estimación de errores estándar de los parámetros omitidos del set básico.

Para obtener éstos parámetros hay que cambiar el modelo y reestimarlos. Sabemos que se estima desde el parámetro 1 hasta el $n-1$, por lo que el parámetro n no es estimado. Lo que debemos hacer en éste caso para que el parámetro n sea estimado y así poder calcular su error estándar es cambiar el input de los datos a estimar, es decir, en vez de ordenar las columnas $1, 2, \dots, n$ y las filas $1, 2, 3, \dots, m$, ordenarlas de la forma $n, n-1, \dots, 1$ las columnas y las filas $m, m-1, \dots, 1$. Así en éste caso los parámetros que no serán estimados son los que anteriormente eran los primeros y serán estimados los que no lo eran. La forma de colocar el input no altera de ninguna manera el resultado obtenido, el resultado de cada celda es exactamente el mismo, por lo que éste método de obtener los errores de los parámetros antes no estimados es válido.

I. 3. 5 Implementación

Para obtener los errores estándar de todos los parámetros uno debe seguir los siguientes pasos. Primero los estimadores de máxima verosimilitud deben ser obtenidos. Luego el error estándar de las estimaciones del set básico deben ser calculados con el proceso antes descrito. Finalmente el set básico es cambiado para ver el resto de los errores estándar.

El cálculo de los errores estándar puede ser implementado fácilmente con el programa MLLSA, el programa desarrollado por Clogg (1977) para estimar los parámetros de un modelo de clase latente. El programa calcula la matriz Jacobiana a chequear para identificar los parámetros. No obstante, el número de filas en la matriz Jacobiana que es calculada por el programa es una menos que el número de filas requerida para calcular el error estándar. Cuando uno obtiene la matriz J puede calcular los errores estándar del set básico. Cambios en el set básico pueden ser hechos fácilmente, el programa trata al último parámetro libre (con W_M , con cada p_{nk} y p_{mk} para todo k) como el parámetro no está incluido en set básico y por esto el error estándar de esos parámetros no está disponible. Por consiguiente, todo lo que necesitamos es cambiar el orden del input en el programa para incluir el parámetro que no está presente en el set básico previo. Así podremos obtener

todos los errores estándar, para luego con ellos sacar todos los valores Z necesario y hacer los análisis correspondientes.

I. 4 “Una aproximación para rastrear cambios dentro del segmento en el mercado”

Grover y Srinivasan

I. 4. 1 Introducción

La mayoría de las aproximaciones para examinar el efecto de las variables de marketing sobre la participación de mercado tratan al mercado como una agregación. Recientemente Grover y Srinivasan propusieron que un segmento de mercado es definido como un grupo de consumidores que son homogéneos en términos de probabilidades de elección de diferentes marcas en una clase de producto. El mercado es por esto dividido en segmentos leales a las marcas y segmentos de cambio. Además las porciones de éstos segmentos de cambio pueden variar a través del tiempo, la razón de ésta variación se explica por variables de marketing que influenciaron a los consumidores. El cambio en el tiempo de las porciones de mercado dentro de los segmentos son relativos a variables del marketing mix, como precio y promoción, estos cambios y éstas variables son modeladas a través de un modelo logit multinomial. Aplicaciones de el modelo a datos de panel escáner

muestran que la aproximación propuesta produce una modesta mejora en validación predictiva en comparación con las validaciones tradicionales de tratar al mercado como una agregación singular.

Evaluar el efecto de acciones de marketing sobre porciones de mercado es obviamente importante para las futuras decisiones de marketing.

Cada análisis estudia el hecho que acciones de marketing de una marca puedan tener diferentes efectos sobre distintos segmentos de mercado.

Una buena línea de investigación en éste sentido ha ido evaluando el sentido de las acciones del marketing sobre las porciones de mercado usando conductas de consumo pasadas como base de segmentación. El modelo logit usa las conductas de compra pasadas de los consumidores a través de una variable fidelidad.

Tomaron la definición de segmentación como aquellos que comparten la probabilidad de compra en diferentes periodos, esto mediante una clasificación cruzada, además asumimos la no estacionalidad. El tomar al mercado solo como una agregación de porciones de mercado es un error, ya que con esto se pueden disimular fuerzas en crecimiento y además debilidades en segmentos distintos.

I. 4. 2 Modelo

Por conveniencia se asume que cada consumidor compra una unidad de producto en muchos periodos de tiempo, es fácil probar que este supuesto no afecta al modelo.

Siguiendo con el modelo, se asume que el proceso de elección es de orden cero, esto no implica que no puedan cambiar de marca producto de algo como una acción de marketing.

Denotando “ n ” como el número de marcas en las clases de producto, existen “ n ” segmentos leales a la marca y “ m ” segmentos de cambio, denotados por $K=1,2,\dots,m$. El valor de “ m ” se obtiene por el método visto anteriormente. Ahora dejamos P denotando la probabilidad con la cual los consumidores del segmento de cambio “ k ” de la porción de mercado de la marca “ i ” en “ t ”. Dejamos V_i denotando la proporción de consumidores del total del mercado quienes son leales a la marca “ i ” y w denotando la proporción de mercado quienes son leales a la marca “ i ”.

Consideramos los distintos periodos de tiempo, “ t ” y T ($t < T$) dentro del horizonte de T periodos. Consideramos la matriz de clasificación cruzada para los periodos de tiempo “ t ” y “ T ”. Dejamos N denotando al número de consumidores que compran la marca “ i ” en el periodo “ t ” y la marca “ j ” en el periodo T , y dejamos N como el total de consumidores en la muestra. Con el supuesto de orden cero, el número esperado de segmentos de cambio “ k ” a los

consumidores que compran “i” y “j” en el periodo “t” y “T”. La agregación de segmentos de cambio está dada por:

$$21) E(N_{ij}^{TT}) = N * \sum_{k=1}^m W_k * P_{ik}^t * P_{jk}^T$$

Similarmente el número esperado de consumidores que repiten la compra está dada por:

$$22) E(N_{ii}^{TT}) = N * \sum_{k=1}^n W_k * P_{ik}^t * P_{ik}^T$$

En la conceptualización del procedimiento todos los consumidores en el segmento de cambio “k” son asumidos a tener la misma probabilidad en ambos periodos.

Así los segmentos de cambio son definidos como en la ecuación 21, pero excluyendo a aquellos que tienen probabilidad cercana a la lealtad.

Se construye la matriz de clasificación cruzada en donde las estimaciones se hacen por el modelo de clase latente.

Se ocupó el modelo logit para la función. El método de análisis es igual para cada segmento de cambio. Consecuentemente ya no se ocupará “k” para

denotar a los segmentos de cambio a los que llamaremos “r”, donde $r=1,2,\dots,R$, denotando las variables de interés de mercado como X_{rit} , así definimos

$$23) p_i^t = \frac{\exp(\sum_{r=1}^R \beta_r * X_{rit})}{\sum_{j=1}^n \exp(\sum_{r=1}^R \beta_r X_{rit})}$$

El supuesto de alternativas irrelevantes no es importante pues el modelo es aplicado a segmentos homogéneos y no a un mercado con consumidores heterogéneos (De hecho para superar éste problema es que segmentamos)

I. 4. 3 Estimación

El coeficiente β_r del modelo logit puede ser estimado por un método análogo o el simple OLS (mínimo cuadrado ordinario MCO). Tomando logaritmos de ambos lados nosotros obtenemos.

$$24) \ln P_i = \sum_{r=1}^R \beta_r X_{rit} + S_t$$

donde δ_t es una constante que depende solo de "t" y está dado por

$$25) \delta_t = - \left[\sum_{j=1}^n \exp \left(\sum_{r=1}^R \beta_r X_{rjt} \right) \right]$$

Luego, para estimar los parámetros B y S , después de la minimización de la suma de los residuos cuadrados, y luego de transformaciones tenemos :

$$26) \ln P_i^t - \frac{1}{n} \sum_{i=1}^n \ln P_i^t = \sum_{r=1}^R \beta_r (X_{rit} - \bar{X}_{rt})$$

La ecuación 26 puede ser estimada por el método OLS.

I. 4. 4 Una ilustración

El modelo es ilustrado con un ejemplo con datos del mercado del café, se eligieron 5 marcas, donde dos de ellas son descafeinados. Todos los diferentes tamaños, nombres y tipos con el mismo nombre de marca fue tratado como una marca singular.

Se ocuparon 1000 datos de hogares (datos de panel scanner) de un periodo de dos años. Se mide las porciones de mercado a nivel semanal.

Probablemente la publicidad fluctúe semana a semana, pero se va a tomar las fluctuaciones solo a nivel mensual se cree que entre más corta sea el tiempo de una compra y otra por parte de los consumidores se pueden hacer mejores análisis.

El modelo propuesto asume que los segmentos son constantes en el horizonte de tiempo del estudio (T). Para que éste supuesto se mantenga, el periodo de tiempo debe ser corto. Si aumentamos el horizonte de tiempo del estudio, los parámetros a estimar aumentarán linealmente, pero las tablas de tabulación cruzada aumentarán en forma cuadrática de forma $T*(T-1)/2$, en donde T es el horizonte de tiempo ya antes definido, en éste caso se eligió un $T = 6$, con lo que nos dará como resultado 15 tablas de tabulación cruzada. En conclusión y tomando lo antes dicho sobre los segmentos, el número y las porciones no cambiarán dentro de los 6 meses, pero sí a través de los 4 periodos en los que se dividieron los datos.

I. 4. 4. 1 Formulación de las variables predictivas

Las variables promocionales disponibles en la base de datos fueron: anuncio (1 si la marca fue anunciada en el periódico en la semana, 0 en otro caso); exhibición (1 si la marca fue exhibida y promocionada en el lugar de venta o 0 en otro caso) ;cupón (1 si es que aparecieron cupones de descuentos

anunciados para la marca durante esa semana o 0 si es que no). Todas las variables fueron promediadas sobre 6 tiendas en el mercado y las cuatro semanas en cada periodo de tiempo. Así pudieron establecer la correlación entre las variables cupón, descuento, exhibición y anuncio. Así se descubrió que las variables anuncio tenían una alta correlación con exhibición y cupón, por lo que las separaron para evitar el problema de multicolinealidad. Por esto es que solo se tomó la variable anuncio para poder hacer mediante el modelo inferencias respecto del mercado, además se tomó como que anuncio era explicada o contenía a las otras dos variables descartadas. Finalmente las variables que fueron incluidas en el modelo fueron precio y anuncio.

Como se desea comparar y analizar el efecto de las variables además del efecto de unas marcas sobre una marca base, para el estudio se capturó la elección de una marca como línea base y una dummie para las otras

I. 4. 4. 2 Determinación del segmento de mercado

Se tienen 15 clasificaciones por cada 6 meses, por lo que en total se tienen sesenta clasificaciones cruzadas de las cuales se eligen aleatoriamente 10, las que se usan para determinar el número de segmentos de cambio y cuales parámetros van a ser restringidos a cero. Esto se hizo aplicando el modelo de Grover y Srinivasan (1987) , se encontraron dos segmentos de

cambio. El primer segmento, las marcas cafeinadas (con tres marcas cafeinadas (MH, F y HB)), tenían una gran participación de mercado, en contraste con las dos marcas descafeinadas (B y S). El segundo segmento , tiene marcas cafeinadas y descafeinadas , en donde estas últimas capturan dos tercios del mercado. Aquí las marcas descafeinadas B y S fueron restringidas a cero en el primer segmento y la marca HB fue restringida en el segundo.

I. 4. 4. 3 Test de consistencia de los tamaños de los segmentos

En éste paso lo que se intenta hacer es verificar el supuesto de que el tamaño de los segmentos no varía dentro de los 6 meses, pero que podría hacerlo entre los cuatro bloques de 6 meses. Por lo tanto los supuestos testeados fueron:

M1: los tamaños de los segmentos variaron en los 6 ,meses.

M2: los tamaños fueron constantes en los 6 meses.

Luego se procede a estimar el tamaño de los segmentos (mediante la estimación de la ecuación 25).

I. 4. 4. 4 Rastreo de modelos dentro de segmentos de cambio.

La ecuación 26 fue estimada por MCO para ambos segmentos , el cafeinado y el descafeinado, en donde en las dos estimaciones ocuparon como marca base a MH, así podemos ver el nivel de ajuste del modelo parta cada uno de los segmentos, además podemos ver el grado de sustitución que existe entre las marcas y el efecto que tienen sobre la elección el precio y los anuncios.

I. 5 “Un modelo de elección probabilístico para segmentación de mercados y elasticidad de estructura”

Kamakura y Rusell

I. 5. 1 Introducción

Un útil estudio fue en éste respecto fue el hecho por Grover y Srinivasan en 1987 (ya antes descrito) supuestos importantes son que el mercado se mantiene estable durante el tiempo de estudio y que las variables de marketing mix se mantienen constantes, estos supuestos se pueden relajar y hacer un análisis mucho mas extenso.

La propuesta de este estudio es proponer y evaluar un nuevo método de segmentación en base a preferencias, para así poder evaluar cuál es el punto en que un cambio de precios afecta a los consumidores y los induce a un cambio de marca, por ejemplo (sensibilidad). El modelo provee un link entre segmentación y elasticidad precio de marca.

I. 5. 2 Un modelo de segmentación preferente

La aproximación versa sobre el supuesto de que los consumidores pueden ser caracterizados por un vector de medias preferentes y un parámetro de sensibilidad de precio

I. 5. 3 Utilidad aleatoria

Este modelo comienza con el supuesto de utilidad aleatoria, cuando los consumidores pasan a una opción de compra, los consumidores le asignan utilidades

aleatorias a cada marca considerando u optando por la que presente utilidad más elevada. Se descompone esta utilidad en componentes determinantes los cuales dependen de características intrínsecas de la marca y su precio y un componente aleatorio.

Así la utilidad queda como:

$$27) U_{jkt} = u_{jk} + \beta_k X_{jkt} + \varepsilon_{jkt}$$

en donde

u_{jk} = utilidad intrínseca de la marca j para el consumidor k

β_k = parámetro del precio

X_{jkt} = precio

ε_{jkt} = error aleatorio (el cual se asume estocástico e independiente, iid.)

Así la probabilidad condicional de elección de marca j para el consumidor k en el tiempo t dado por el modelo “logit multinomial”.

$$28) P_j(u_k, \beta_k, X_{kt}) = \frac{\exp(u_{jk} + \beta_k X_{jkt})}{\sum_j \exp(u_{jk} + \beta_k X_{jkt})}$$

I. 5. 4 Segmentación de preferencias

Más que estimar un parámetro individual para cada elección de un consumidor k de una marca j , se asume que existen segmentos homogéneos y los parámetros de esos segmentos son los que son estimados. Así el tamaño de los segmentos estará dado por la ecuación:

$$29) \quad f_i = \frac{\exp(\alpha_i)}{\sum_{i=1}^n \exp(\alpha_i)}$$

Ahora se podría reescribir la ecuación 29, dejando como subíndice ya no a los consumidores individuales si no que al segmento. Como se ha hecho en estudios anteriores, aquí se relaja el supuesto de homogeneidad para ver como influye en las preferencias variables como precio, es decir analizaremos la elasticidad del cambio de marca dependiendo del precio.

I. 5. 5 Estimación

Se estiman la ecuación 28 y 29, se asume orden cero en las preferencia y se levanta el supuesto de estacionalidad para poder ver el efecto de cambios en precios y otras variables sobre las elecciones de consumo.

Como el número de segmentos es desconocido, este se estima mediante el procedimiento de Akaike.

I. 5. 5. 1 Representación de la estructura competitiva.

El modelo propuesto permite describir y analizar la estructura competitiva de una marca en dos direcciones.

1° Podemos establecer porciones y tamaños de segmentos .

2° Analizar la estructura competitiva en términos de sensibilidad de precios.

Un distintivo es que la estimación de la “preferencia intrínseca por cada marca será visto como reacciones de los consumidores a las diferencias de precios

I. 5. 5. 2 Derivando agregación

Dado la formulación logit para el tamaño de segmentos y el supuesto de homogeneidad dentro del segmento, la participación de la elasticidad en la elección está dada por:

$$30) \quad \eta'_{jj} = \beta_i \left(-S_{ij} \bar{X}_{j^*} \right)$$

$$31) \quad \eta'_{jj^*} = -\beta_i S_{ij} X_{j^*} \quad j \neq j^*$$

donde

$$32) \quad S_{ij} = \frac{\exp(\beta_{ji} + \beta_i X_{j^*})}{\sum_j \exp(\beta_{ji} + \beta_i X_{j^*})}$$

Es la porción agregada de la marca j dentro de i.

La agregación de las elasticidades cruzadas se obtienen combinando las elasticidades de los niveles de segmentos.

$$33) \quad \eta'_{jj^*} = \sum_i \left[f_i S_{ij} \right] S_j \bar{\eta}'_{jj^*}$$

Esta elasticidad agregada aporta la estructura competitiva del espacio de mercado.

I. 5. 6 Aplicación del modelo

Para ilustrar la aproximación se analizarán niveles entre las marcas nacionales y privadas en una categoría de producto. El ejemplo es elegido para permitir explorar las características de competencia entre marcas nacionales y marcas privadas.

I. 5. 6. 1 Descripción de datos

Los datos consisten en 78 semanas de datos scanner para una categoría de producto. Son de 585 hogares que compraron en 9 tiendas. El producto es de comida con tiempo de intercompra de diez semanas, como se trataba de un producto perecible que debía ser refrigerado y consumido una vez abierto, por esto los consumidores compraban un producto por ocasión de compra.

I. 5. 6. 2 Definición de lealtad

Se deben separar a priori las marcas leales, es decir, que durante el tiempo siempre fueron compradas por el mismo hogar. Además se separan las marcas clasificándolas por grupos de precios y cantidad de publicidad,

entonces quedan en A las marcas con altos precios (premium) y publicidad considerable en canales de televisión, en B están las marcas con precios altos con pequeña publicidad, pero específica a su categoría, en C están las marcas que compiten en precio y tienen publicidad para las familias de productos y no específica para esa categoría en especial y en P están las marcas privadas.

I. 5. 6. 3 Elección del número de segmentos

Se utilizan los datos de la historia de compra para calibrar los parámetros del modelo de segmentación de preferencias.

El modelo fue estimado ocupando el procedimiento de máxima verosimilitud. Para especificar el modelo se hizo variar M (el número de segmentos) sintomáticamente y luego se calculó el número de segmentos por el criterio de información Akaike.

$$34) \quad AIC = -2(LL - p)/N$$

Donde LL es el máximo valor del modelo estimado por máxima verosimilitud.

I. 5. 6. 4 Segmentación de preferencias

Los parámetros fueron estimados con el método de máxima verosimilitud , de ahí se procedió a identificar el modelo en donde algunos parámetros fueron restringidos a cero, de ahí se pueden hacer inferencias , ya que se tienen los valores y magnitudes de los parámetros , etc.

I. 5. 6. 5 Estructura por Elasticidad

Usando las relaciones desarrolladas previamente , nosotros podemos ver una representación de competencia de marcas en término de elasticidad precio.

I. 5. 6. 6 Corrección de factores

Como se dijo antes, para poder hacer un análisis más acabado sobre los efectos de variaciones de precios sobre las elecciones de los consumidores, se relajaron algunos supuestos como que los segmentos son homogéneos. Las preferencias para cada segmento son estimadas con una distribución Dirichlet, la corrección de factores fue pequeña, o dejando pequeños grados de heterogeneidad, se hace una tabla en donde se están las elasticidades cruzadas por marca y por segmentos.

I. 5. 6. 7 Clout Competitivo y Vulnerabilidad

El Clout Competitivo y la Vulnerabilidad nos indican lo sensibles o expuestas que están una marca en el mercado, es así como el clout es así como una marca que tenga un clout alto indica que tiene un mayor impacto sobre las participaciones de mercado de las marcas de la competencia, y una marca que tenga un índice de vulnerabilidad alto indica que es muy sensible en cuanto a participación a cambios de precios de los competidores, es así que se definen:

$$35) \text{ Clout competitivo} = \sum_{j \neq i} \eta_{ji}^2$$

$$36) \text{ Vulnerabilidad} = \sum_{j \neq i} \eta_{ij}^2$$

El link entre segmentación por preferencia y estructura por elasticidad puede ser visto como un trade-off competitivo entre clout y vulnerabilidad.

I. 6 “Entendiendo la competencia de marca usando micro y macro datos de panel scanner”

Russell y Kamakura

I. 6. 1 Introducción

Los autores proponen un método que enriquece el estudio de conducta competitiva combinando información obtenida de micro y macro paneles scanner, donde los micro datos de panel, corresponde al consumo de los hogares y los macro datos de panel corresponde a las ventas hechas por tiendas comerciales y supermercados que tengan el scanner.

Primero se usa el modelo de clase latentes en los micro datos para identificar los segmentos y obtener información sobre las preferencias. El resultado de los micro datos serán luego usados en conjunto con los datos de las tiendas para estimar un modelo de porciones de mercado y hacer análisis con los resultados obtenidos.

I. 6. 2 Modelo de conducta de compra de hogares

La primera fase usa las clases latentes.

Dejaron a X_{iht} denotando al número de unidades equivalentes de la marca i comprada por el hogar j en la semana t . Se asume que el hogar pertenece al segmento s , con preferencias homogéneas en él. Para simplificar

el modelo se asume que X_{iht} sigue una distribución Poisson y que posee una media de

$$37) \lambda_{iht} = \lambda_h \lambda_t u_{iht}$$

donde

λ_{iht} = media del número de unidades del producto.

λ_h = medida de compra semanal por volumen por semana h.

λ_t = factor de estacionalidad que afecta a las compras de todos los hogares.

u_{iht} = porción de volumen esperada de compra de la marca i por h en t

Donde u_{iht} se escribe como:

$$38) U_{iht} = \exp(\alpha_{ist}) / \sum \exp(\alpha_{ist})$$

El término α_{ist} representa el atractivo de la marca i para un miembro del segmento s. El modelo de log-atractivo se escribe entonces:

$$39) \alpha_{ist} = \phi_{is} + \beta_i p_{it} + \delta_i d_{it},$$

donde:

ϕ_{is} = atracción intrínseca de la marca i (generalmente se asume 0).

β_i = la sensibilidad al precio.

p = el precio de la marca i .

δ_i = parámetro de sensibilidad a la promoción.

d_{it} = intensidad de la promoción de la marca i .

I. 6. 3 El macro modelo : porción del volúmen semanal.

Primero se mostrará como el macro modelo se levanta de la agregación de compra de los hogares. Se definen $X_{it} = \sum X_{iht}$ como el volumen total de la marca i comprada por los hogares en la semana t , y dejaron:

$$40) Y_{it} = X_{it} / (X_{it} + \dots + X_{Bt})$$

donde Y_{it} es la porción de mercado de cada X_{it} . Además se define:

$$41) fs = \frac{\sum_{(hans)} \lambda_h}{\sum_{(hen1)} \lambda_h + \dots + \sum_{(hens)} \lambda_h}$$

Así f_s es el volumen esperado relativo para cada celda.

$$42) MS_{it} = \sum_s f_s MS_{ist},$$

donde

$$43) MS_{ist} = \frac{\exp(\psi_{ist})}{\sum \exp(\psi_{ist})}$$

I. 6. 4 Ilustración empírica

Datos:

Micro :-Compras de 12 marcas

-1361 hogares.

-42 semanas.

Macro:-27 tiendas.

-42 semanas.

Hay que ver que los datos sean representativos, éste es un supuesto crucial.

1° Se tomaron 22 semanas para la calibración del modelo y los siguientes 20 semanas para la validación propuesta. Eliminaron la combinación tienda-semana en las cuales el set de 10 marcas no estaba disponible, éste proceso dio de 442 a 409 datos de unidades de semanas en tiendas

Por cada tienda y semana se obtuvo:

- 1.-Porción de mercado por cada una de las 10 categorías de marcas.
- 2.-Precio promedio por cada categoría de marca.
- 3.-Intensidad de promoción por cada marca durante la semana.

I. 6. 5 Microanálisis

Aplican clases de modelos latentes para una muestra de 1361 hogares . Usando Akaike se obtiene el número de segmentos y el tamaño de los segmentos.

I. 6. 6 Macroanálisis

1° Se usan las tablas de segmentos fs obtenidas en el microanálisis

2° El parámetro ϕ fue asumido en ser igual a $I_{is} + p$

El modelo fue estimado por el método de máxima verosimilitud por medio de Scale Dirichlet

Luego se saca el U que tiene una interpretación similar al R y mide el ajuste del modelo.

p_i = captura el efecto del marketing.

I_{is} = es la elasticidad de las promociones.

$P_i + I_{is}$ = las preferencias de cada segmento por una marca.

S_{ij} = índice de sustitución y representación de la estructura del mercado.

M_{ij} = medida de la fuerza de una marca.

I. 7 Discusión

De acuerdo a las investigaciones aquí expuestas podemos ver como se ha ido avanzando en cuanto a la utilización de datos de panel scanner como una herramienta del marketing, de hecho, podemos hacer mención de otros estudios que han contribuido a la misma causa.

En éste sentido el primer paper a mencionar sería el de “Segmentación de consumidores vía análisis de clases latentes”¹ . Este paper es uno de los

¹ Paul Green, Frank Carmone y David Wachspress,1976

estudios de base en cuanto a la segmentación por el modelo de clases latentes, es una investigación donde no se utilizan datos de panel scanner, si no más bien, datos producto de un estudio exploratorio en donde se clasifican a los consumidores mediante variables como educación, cambios de casa (en lapsus de tiempo), edad, ingresos y por su respuesta de adoptar o no una innovación, el modelo es el mismo utilizado luego por Gover y Srinivasan y muchos otros.

Si seguimos en el tiempo el siguiente avance sería la inclusión de datos de panel scanner al modelo de segmentación, claro que ahora con una gran diferencia, ya no estaríamos segmentando por variables sociográficas, ni sicográficas, tampoco por datos de opinión, si no más bien por conductas de compra, hecho muy importante y lo que significa un gran avance, Grover y Srinivasan son unos de los primeros en hacerlo, de hecho en su paper “Una simultánea aproximación a la segmentación y estructuración de mercado”², utilizan el modelo no solo para segmentar, si no también para ver como se mueven las fuerzas en el mercado y como se influyen las distintas marcas y categorías de producto.

Con el afán de mejorar el modelo anterior Rajiv Grover con su paper “Estimación y uso de errores estándar de los parámetros del modelo de Clases latentes”³ formalizó la restricción de parámetros para la identificación del modelo de clases latentes. . Guadagni and Little, fueron los primeros en

² Grover y Srinivasan, 1987

calibrar el modelo logit con datos de panel scanner, Neslin, Henderson, and Quelch (1985) examinan y cronometran las decisiones de compra, Krishnamurthi y Raj (1988) modelaron simultáneamente “cuando, que y cuanto cuesta”⁴. En el año 1989 Grover y Srinivasan vuelven a tomar el tema de la segmentación con datos de panel scanner y quieren avanzar en éste sentido haciendo variar algunas variables del marketing mix, que en el modelo inicial se mantenían constantes⁵, por esto se toma un horizonte de tiempo mucho mas largo y ahora se utiliza un modelo logit, de tal forma de detectar como acciones de marketing de una marca , puedan tener diferentes efectos sobre distintos segmentos de mercado. En el mismo año Kamakura y Russell siguiendo los pasos de Grover y Srinivasan en el último estudio mencionado, en su paper “ Un modelo de elección probabilístico para segmentación de mercados y elasticidad de estructura” hacen variar algunos elementos del marketing mix para medir la sensibilidad de la estructura competitiva de mercado ante variaciones en el precio. En 1990 Greg Allenby focaliza su estudio no en modelos como lo habían hecho sus antecesores, si no más bien en el tratamiento de datos de panel scanner, antes de modelarlos, específicamente se preocupa del problema que tienen los datos en los test por ser muestras muy grandes en su paper “Test de Hipótesis con Datos de Panel

³ Rajiv Grover, 1987

⁴ Greg Allenby, 1990.

⁵ “Ceteris paribus”, supuesto impuesto en Grover y Srinivasan 1987.

Scanner: La Ventaja del Método Bayesiano”. En el año 1994 Gary Russell y Wagner Kamakura desarrollaron un modelo en el que utilizaron datos de macro y micro panel scanner simultáneamente en su paper “ Entendiendo la competencia de marca usando micro y macro datos scanner”, tomando de los micro datos de panel scanner información de segmentación y de los macro panel scanner información específica sobre los productos y el lugar de compra, cosas como promociones en el lugar, publicidad en el lugar, etc. con éstas herramientas se desarrollo un modelo donde se explicaba el comportamiento de los segmentos según las variables independientes, así podían ver la estructura y competencias de mercado.

Hasta ahora solo hemos visto como los datos de panel scanner han sido utilizados básicamente para segmentar y ver comportamiento y estructura de mercado, recientes estudios han ido más allá en la utilización de los datos llegando a la valoración de marcas (Kamakura y Russel en “Midiendo el valor de marca con datos de panel scanner” en 1989)⁶ y ver por ejemplo como afectan las extensiones de marca a las otras marcas del porfolio en el paper “Efecto recíproco Spillover : Un beneficio estratégico de las extensiones de marca” de Subramanian Balachander y Sajoy Ghose, donde primero utilizan el modelo desarrollado por Grover y Srinivasan para segmentar y luego controlan por variables como publicidad y esas cosas para ver como afecta al

⁶ Fuente: Tesis “Esquemas de segmentación y valoración de marca con datos de panel scanner”

porfolio la extensión de una marca, ven que una extensión tiene un efecto positivo, dado por las economías de escala por publicidad, el que la nueva marca afecte positivamente la imagen de marca, etc y un efecto negativo, dado por el efecto sustitución que se puede dar entre productos de una misma marca o por dilución de la imagen de marca, efecto no deseado que también se puede dar si es que se incluyen demasiadas categorías de producto y productos distintos a una marca paraguas.

II. 1 Introducción

Para ilustrar de mejor manera el modelamiento de datos de panel escáner, desarrollaremos el modelo propuesto por Grover y Srinivasan (1987), en donde se segmenta mediante la información de compra por parte de los hogares existentes en la muestra. Poseemos los datos de un producto, café, procederemos a desarrollar el modelo propuesto para dos productos.

Los datos fueron proporcionados por la empresa ACNielsen y corresponden a los productos de café.

II. 2. Datos Café

II. 2. 1. Descripción de los datos

Gracias a la colaboración de la empresa ACNielsen tenemos 7767 observaciones correspondientes a compras de alrededor de 1000 hogares, el producto en cuestión es el café, los datos en un principio estaban divididos por semana, que en total sumaban 52. En total la muestra contenía 31 marcas originales las que aumentaron porque categorías de producto como Nescafé Dolca y Monterrey Gold fueron tomadas como marcas independientes por simplicidad y como una forma de estudiar de mejor manera el efectos sustitución entre marcas, otro cambio que sufrieron las marcas fue que Carrefour, Cinco Continentes, Ekono, Lider y Unimarc fueron tomadas como un conjunto llamado “marcas privadas”.

Para poder trabajar en los datos los unimos en una sola tabla y sacamos las participaciones de mercado por marca, luego por marca y categoría de producto. Así podremos elegir las marcas que tengan la mayor participación

de mercado para luego con ellas desarrollar el modelo descrito anteriormente.

Tabla 2-1: Participaciones de mercado por categorías de producto.

Particip. (%)	Ventas	caf	grano	agreg	Marca
0,637800841	9093913	1	0	NO	N
0,086335523	1230992	1	0	MA	M
0,042508072	606090	1	0	NO	ND
0,034116359	486439	1	0	PR	N
0,029775148	424541	2	0	NO	CD
0,025962754	370183	2	0	NO	EC
0,016119178	229831	0	0	NO	N
0,013727718	195733	1	0	NO	TP
0,013013253	185546	0	0	NO	M
0,008821289	125776	1	0	NO	CZ
0,008739863	124615	1	0	CA	N
0,008430077	120198	1	0	MK	H
0,00711666	101471	1	0	CL	N
0,006801264	96974	1	0	NO	MG
0,006433195	91726	0	0	PR	MG
0,004208726	60009	1	0	NO	M
0,003684889	52540	1	0	NO	CR
0,003555279	50692	1	0	NO	U
0,003138608	44751	1	1	NO	N
0,002903165	41394	1	0	NO	MR

La celda agregado posee las siguientes simbologías que tienen los siguientes significados.

- NO: Normal
- MA: Mezclas Arábocas
- MK: Moka
- AM: Amareto
- CR: Cremossimo
- EX: Express
- CL: Con leche
- TS: Tostado Suave
- CA: Capuccino
- PR: Premium

Aquí los símbolos que están en la celda de marca corresponde a:

- 3M: 3 Marías
- B:Braschile
- CR:Caribe
- CL:Colcafé
- CP:Copacabana
- CD:Coronado
- CZ:Cruzeiro
- EC:Ecco
- H:Haití
- HW:Hanse Wappend
- HS:Husco
- J:Jacobs Kronung
- K:Kimbo
- KR:Kruger
- M.A:M.A.S
- MH:Maxwell House
- ML:Melitta
- M:Monterrey

0,002700615	38506	1	0	NO	CP
0,002579843	36784	1	0	EX	N
0,002119477	30220	0	0	NO	K
0,002104048	30000	1	0	NO	N
0,001912439	27268	0	0	NO	CL
0,001904163	27150	1	0	EX	M
0,001890487	26955	1	0	MA	H
0,001879125	26793	1	0	NO	MH
0,001752251	24984	0	0	NO	N
0,001653781	23580	1	0	NO	TC
0,001614997	23027	1	0	NO	L
0,00125289	17864	1	1	NO	CR
0,001248402	17800	1	0	EX	H
0,001127349	16074	1	1	MK	H
0,001100066	15685	1	0	NO	EK
0,000842461	12012	1	0	AM	M
0,000798696	11388	1	0	NO	M
0,000781584	11144	1	0	AM	MG
0,000684447	9759	1	0	NO	HS
0,00058233	8303	1	0	MK	MG
0,000575106	8200	1	0	NO	CR
0,000546561	7793	1	0	NO	CF
0,000536111	7644	1	0	CR	M
0,000414918	5916	1	1	EX	3M
0,000402925	5745	1	0	NO	3M
0,000344643	4914	0	0	NO	MH
0,000318974	4548	0	0	NO	MR
0,000313643	4472	1	0	NO	CL

En la celda de
cafeinado en donde 1
es cafeinado, 0 es
descafeinado y 2 es
cebada tostada.

En la celda de grano,
definimos como 0 a
los cafés molidos y
como 1 a los cafés en
granos.

0,000311189	4437	1	0	MA	PC
0,000293094	4179	1	0	CR	MG
0,000276752	3946	1	0	NO	KR
0,000221626	3160	0	1	NO	CP
0,000210264	2998	1	1	NO	TI
0,000207038	2952	1	0	NO	P
0,000189084	2696	0	0	NO	ML
0,000175688	2505	1	0	NO	CC
0,000136833	1951	1	0	NO	ML
0,000126944	1810	1	0	EX	MG
0,000126173	1799	1	0	NO	MG
0,000123858	1766	1	1	NO	H
0,000111164	1585	1	1	NO	CP
8,87908E-05	1266	1	0	CA	HW
7,85511E-05	1120	1	1	NO	B
6,99245E-05	997	1	0	NO	M.A
4,37642E-05	624	1	0	AM	MG
3,50675E-05	500	1	0	NO	T

La clasificación fue hecha tomando en cuenta los elementos que a nuestro juicio son importantes para los consumidores al momento de elegir un producto, además son características que los diferencian en cuanto a la forma de preparación (grano), a posibles agregados que podrían incluir en sus ingredientes que cambien el sabor del café puro, o un producto de distintas calidades o intensidades de sabor y aroma (agregado), o composición o

materia prima de la que está hecha el producto, descafeinado, cebada tostada, cafeinado (cafeinado).

De acuerdo con la Tabla 2-1 las 10 categorías de producto que tienen mayor participación (más de un 1%) de mercado son:

Tabla 2-2

Marca	Participación
-Nescafe Cafeinado Normal	63.78%
-Monterrey Cafeinado Mezclas Arábicas	8.63%
-Nescafé Dolca Cafeinado Normal	4.25%
.Nescafé Cafeinado Premium	3.41%
-Coronado Cebada tostada Normal	2.97%
-Ecco Cebada tostada Normal	2.59%
-Nescafé Descafeinado Normal	1.61%
-Tempo Cafeinado Tostado Suave	1.37%
-Monterrey Descafeinado Normal	1.30%

-Marcas Privadas Cafeinado Normal	0.88%
-----------------------------------	-------

Una excepción se produce en las llamadas marcas privadas, donde no poseen el lugar 10 en las participaciones de mercado, sin embargo, creemos que es importante agregarla para ver el efecto de las emergentes marcas privadas sobre las marcas tradicionales. Marcas privadas se entiende como las marcas de café que han sacado los Supermercados como Carrefour, Cinco Continentes, Ekono, Lider y Unimarc.; marcas que en total suman una participación de mercado de 0.69%.

Obviamos la clasificación de si los productos son molidos o están en grano porque los 10 elegidos para participar en el modelo son molidos, por lo que esa categorización está, según nuestro criterio, demás.

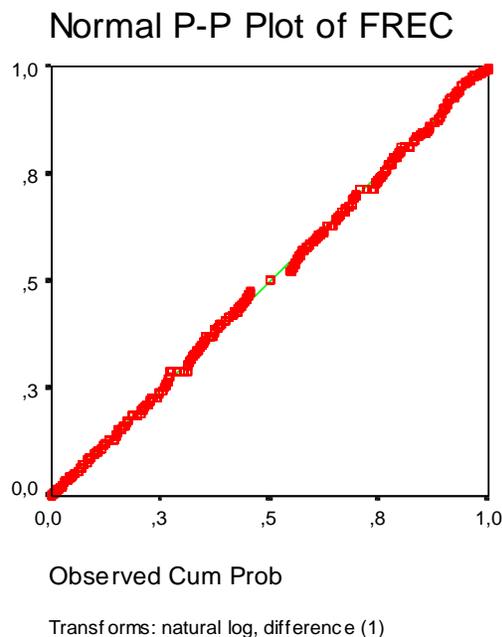
Una vez elegidas las marcas y correspondientes categorías de producto con mayor participación, procedimos a limpiar la muestra dejando en ella solo las compras de éstas marcas, reduciéndose la muestra total a 7102 datos.

II. 2. 2. Distribución de los datos

En los datos existe un elemento importante que no se puede obviar, existen panelistas que presentan solo una compra y otros que incluso

presentan mas de 60 compras. Para ver si es que éstos datos son relevantes o solo ruidos en la muestra, tratamos de encontrar la distribución que permita comprobar la relevancia y peso de los datos, como a priori no se sabe la distribución ocupamos el criterio P-P del programa SPSS 11.0 para ver si sería conveniente analizar la muestra como si siguiera una normal.

Gráfico 2-1



El gráfico 2-1 muestra claramente como los datos se ajustan casi perfectamente a la línea , por lo que podríamos decir que la distribución claramente se

A pesar de que el gráfico 2-1 muestra que el ajuste de los datos a una normal es casi perfecto, decidimos realizar test matemático de ajuste con el fin de comprobar éste resultado, utilizamos para esto contrastes no paramétricos, ya que “en estadística paramétrica se trabaja bajo el supuesto de que las muestras se han originado a partir de poblaciones que poseen

ciertas distribuciones conocidas, donde cada función de distribución teórica depende de uno o más parámetros poblacionales. El uso eficiente de los métodos paramétricos de estimación y de contrastación requiere, por lo general, la normalidad de la población y la estabilidad de la varianza. Sin embargo, en muchas situaciones, es imposible especificar la forma de la distribución poblacional, y solo podemos suponer en forma razonable que es continua. El proceso de sacar conclusiones directamente de las observaciones muestrales, sin formar los supuestos con respecto a la forma matemática de la distribución poblacional se llama teoría no paramétrica o métodos de distribución libre⁷.

Ahora haremos un contraste de ajuste de normalidad, para ver si es que la distribución sigue efectivamente una normal o no. A continuación presentamos dos test, el Kolmogorov-Smirnov y el Shapiro y Willks.

El test Kolmogorov-Smirnov es un test para la bondad de ajuste que es solo aplicable a variables continuas y trata de medir el ajuste entre la función de distribución empírica de una muestra y la función de distribución teórica. Se trata por tanto de un contraste de ajuste de la distribución de una muestra dada a una distribución continua determinada.

La distribución empírica de la muestra x_1, x_2, \dots, x_n , se define como :

⁷ Cesar Perez ,Técnicas estadísticas con SPSS. Pág 273

$F_n(x) = \frac{\text{N}^\circ \text{ de valores del conjunto } \{x_1, x_2, \dots, x_n\} \text{ que son } \leq x}{N}$

N

Para contrastar la hipótesis de que la muestra se ajusta a una distribución teórica $F(x)$, se calcula el estadístico:

$$D_n = \text{Máx } |F_n(x) - F(x)|$$

Cuya distribución es conocida y está tabulada. Si la distancia calculada D_n es mayor que la encontrada en las tablas, para un nivel α , rechazamos la distribución $F(x)$ para la muestra.

El test Shapiro y Wilks ve específicamente los contrastes de normalidad, donde se trata de comprobar si los datos provienen de una distribución normal. Este test mide el ajuste de la muestra a una recta al dibujarla en un papel probabilístico normal. Se rechaza la normalidad cuando el ajuste es bajo, que corresponde a valores pequeños del estadístico del test, por lo tanto se rechaza la normalidad cuando su valor calculado a

partir de la muestra es ,menor que el correspondiente valor crítico dado en las tablas.⁸

Tabla 2-3 : Test de Normalidad

	Kolmogoro v-Smirnov			Shapiro- Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
VAR00002	,168	1042	,000	,787	1042	,000

a Lilliefors Significance Correction

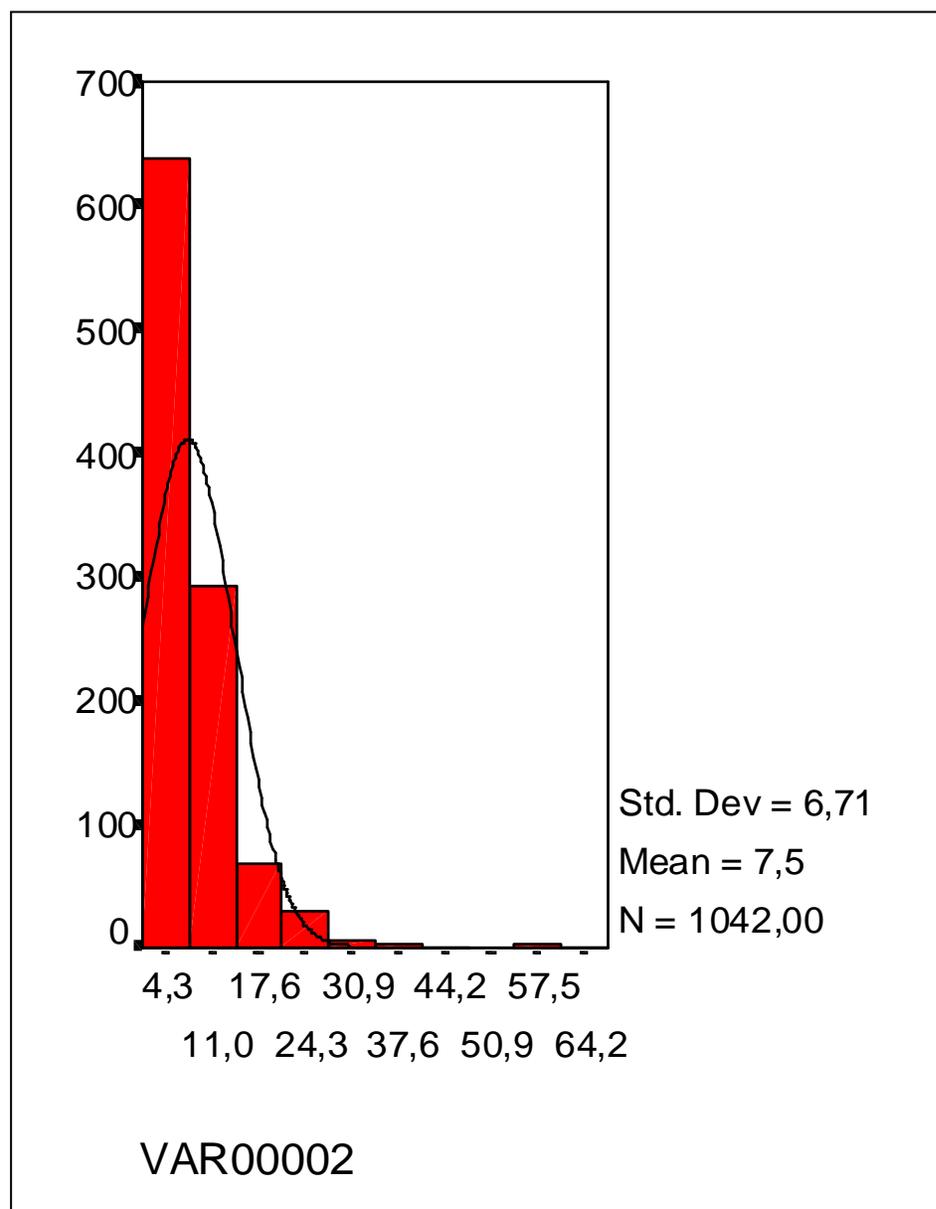
Al ver la tabla 2-3 podríamos decir que los datos no siguen una distribución normal, ya que en los dos test rechazan la hipótesis nula que supone la distribución normal. Pese a esto utilizaremos un histograma que tenga encima de éste una línea que asemeje la muestra a una normal, esto no afecta el análisis posterior de los datos ni los pasos posteriores que se requieren para segmentar, solo lo haremos para ver si es que en la muestra existen datos

Como consecuencia del análisis anterior y con los datos de compra del total de los panelista, hicimos un histograma que contenía el número de

⁸ Cesar Perez ,Técnicas estadísticas con SPSS. Págs. 280 y 290.

ocasiones de compra en el año por panelista, en el gráfico 2-2 podemos ver los resultados obtenidos.

Gráfico 2-2



En el gráfico 2-2 podemos apreciar que existe un número considerable de consumidores que compran café solo una vez al año, que la media de consumo de café esta en 7,5 unidades (bajo el supuesto que por cada ocasión de compra se adquirió solo una unidad de producto)⁹, que la desviación estándar es de 6.71, número bastante alto por lo que podríamos inferir que el consumo de café en los hogares no es homogéneo y que varía bastante de un hogar a otro, además que existen hogares outlier, es decir, que se salen de la distribución de la muestra y no aportan mayormente, si no más bien generan “ruido”. Como consecuencia decidimos eliminar de la muestra a los panelista que forman parte del ruido de la muestra y además a los pertenecientes a las colas de la distribución. Ahora bien, la eliminación de los datos que hacen ruido además de los datos de la primera cola, personas que solo hacen una compra en el año, es necesaria para poder tener una distribución “sana” y además para poder luego hacer la tabulación cruzada necesaria para el modelo. Para poder elegir las opciones en cuanto al número de consumidores a eliminar de la muestra, fuimos eliminando uno por uno para ver cuál era la forma en la que quedaría una mejor distribución,

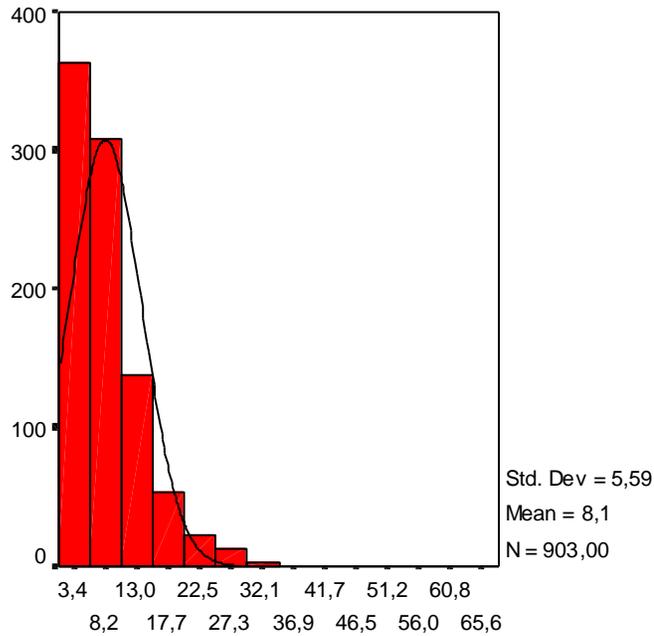
⁹Grover y Srinivasan 1987 propusieron éste supuesto para desarrollar el modelo de clases latentes.

luego de éste procedimiento, concluimos que existen dos opciones una de las cuales se detalló anteriormente y la opción de eliminar no solo los datos de la primera cola si no además los datos de la última cola, datos que no son muy representativos de la población y además que aumentan la variabilidad de los resultados. Ahora analizaremos las dos opciones:

- *La primera opción elimina en total observaciones, donde 132 pertenecían a la primera cola (compraban solo 1 tarro de café al año), y los últimos tres pertenecían al ruido de la muestra (compraron más de 58 tarros), es decir la muestra disminuyó en un 87 % aproximadamente. La distribución resultante es la que muestra el gráfico 2-3.*

Gráfico 2-3

Podemos observar que la desviación estándar de la muestra (antes 6.71) disminuye considerablemente a 5.95, que la media aumenta a 8.2, luego de haber disminuido la muestra inicial en un 12.95%. Además vemos que el efecto “ruido” desapareció por completo.

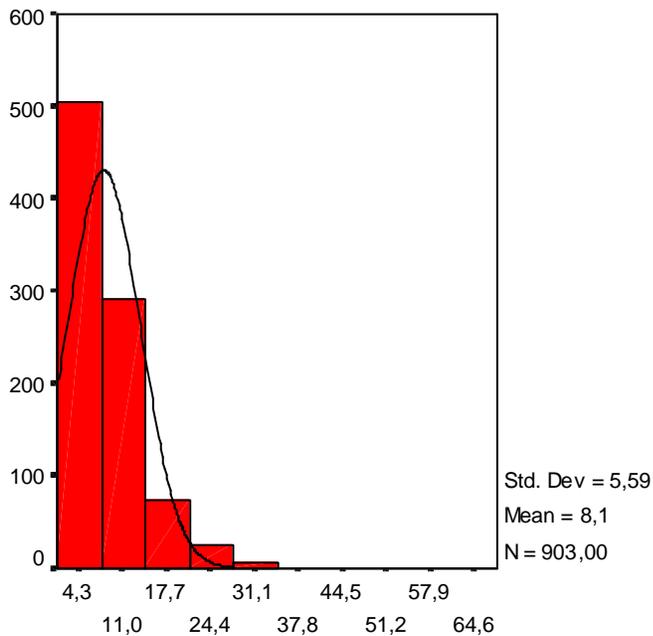


VAR00004

En éste caso la muestra total disminuyó a 907 datos

- *La segunda opción sugiere eliminar de la muestra a los 132 consumidores que pertenecían a la primera cola (compraban solo 1 tarro de café al año), los últimos tres que pertenecían al ruido de la muestra (compraron más de 58 tarros) y además eliminar a los consumidores pertenecientes a la segunda cola de la muestra, es decir, a 4 consumidores (compraban entre 35 y 46 tarros de café al año); por lo tanto la muestra se redujo en un 86.6% aproximadamente. El gráfico 2-4 nos muestra la distribución resultante después de los cambios aquí mencionados.*

Gráfico 2-4



VAR00004

Observamos que la desviación estándar disminuyó a 5.59, que la media bajó a 8.1 y que la muestra total se redujo a 903 observaciones.

Para poder concluir acerca de cuál va a ser la opción a elegir, debemos evaluar los costos y beneficios que traerían ambas opciones. Para hacer éste análisis de mejor manera construimos un cuadro resumen (Tabla 2-4) de los resultados obtenidos por cada una de las alternativas.

Tabla 2-4

	Desv. Estándar	Media	N. de observaciones
Opción Inicial	6.71	7.5	1042
Opción 1	5.95	8.2	907
Opción 2	5.59	8.1	903

De acuerdo con los datos obtenidos podemos ver que por un lado con la primera opción tenemos una variabilidad en las observaciones menor que la distribución original, la cual contenía todos los datos. Si tomamos como supuesto que la disminución de la desviación estándar es igual por cada panelista omitido, es decir, que la variación marginal de la desviación es igual, podremos hacer el siguiente análisis. La variación de observaciones entre la opción inicial y la opción 1, es de 135 panelistas y que la variación de la desviación estándar es de 0.76, esto nos da como resultados que por cada panelista el cambio en la desviación fue de 0.005629, por otro lado en la opción 2 tuvo una variación de la desviación estándar con respecto a la distribución original de 1.12 y la variación de consumidores fue de 139

panelistas, por lo que la disminución de la desviación estándar por cada consumidor omitido fue de 0.00805755.

En conclusión podríamos decir que a pesar de que se pierden más observaciones (un 0.4% más que en la opción 1) es mucho más conveniente trabajar con la distribución entregada por la opción 2, ya que se acerca más a la media de la distribución original y que la desviación estándar es menos, por lo que la dispersión de las observaciones va a ser menor (la opción dos presenta una alternativa mucho mas eficiente en la disminución de la dispersión de los datos).

Algo curioso y destacable es que los panelistas omitidos que pertenecían a la segunda cola de la distribución y al ruido de la muestra eran consumidores catalogables como leales, es decir que en todas las ocasiones de compra (más de 35) elegían siempre la misma marca, las marcas compradas por éstos eran :

- Nescafe Cafeinado Normal
- Monterrey Cafeinado Mezclas Arábicas
- Coronado Cebada tostada Normal
- Ecco Cebada tostada Normal
- Tempo Cafeinado Tostado Suave
- Monterrey Descafeinado Normal

II. 2. 3. Elección de las compras

Como el modelo requiere de la tabulación cruzada de datos de dos compras en un periodo corto de tiempo, por el supuesto de que el proceso de elección de marca es estacionario y de orden cero, de la muestra resultante, se escogieron las dos primeras compras por panelista, compras hechas en dos ocasiones distintas. Para la elección de compras se siguió el siguiente criterio de selección:

- *Dos compras se refiere a compras realizadas en dos ocasiones distintas, eso implica que dos compras hechas en el mismo día no pueden ser elegidas, es decir, se deberá elegir a una de ellas y como segunda compra una hecha en un día posterior. Como no hay forma de saber a priori cual es la verdadera razón por la que aparecen en el panel dos compras en el mismo día, en caso de que las compras hechas en el mismo día sean de distintas marcas, se debe elegir la primera que aparezca, esto bajo el supuesto de que las personas compren en días distintos y las pasen por el escáner en el mismo día, además de que el orden de compra sea igual al orden en el que fueron pasados los productos por el escáner.*

- *Si es que ocurre que un panelista tiene solo dos compras y las dos fueron hechas en el mismo día, éste panelista debe ser eliminado de la muestra ya que no cumple con la condición o requerimientos del modelo. Lo mismo ocurrirá si es que un panelista tiene solo una compra.*

II. 2. 4. Tabulación cruzada

Finalmente quedaron 857 observaciones con las cuales se hizo la tabulación cruzada, para poder llevarla a cabo numeramos las 10 marcas y sus correspondientes categorías elegidas, la tabulación (Anexo 1-Tabla A1) se hizo en el programa SPSS, lo que arrojó la siguiente tabla, donde cada una de las celdas corresponde a la frecuencia con la que algunos panelistas eligieron la misma combinación de marcas, donde las filas representan a la primera ocasión de consumo en donde se realiza la primera elección de marca, las columnas representan a la segunda ocasión de consumo, la diagonal de la tabla representa los segmentos que son leales a las marcas, es decir, personas que en las dos ocasiones de compra prefirieron la misma marca.

Tabla 2-5

	N-NO-C	M-MA-C	ND-NO-C	N-PR-C	CD-NO-S	EC-NO-S	N-NO-D	TP-TS-C	M-NO-D	PR-NO-C	Suma
N-NO-C	405	20	14	9	19	14	6	1	1	1	490
M-MA-C	25	43	3	2	2	3	0	0	0	0	78
ND-NO-C	31	1	29	1	2	2	0	1	0	3	70
N-PR-C	19	2	2	13	3	2	0	0	0	0	41
CD-NO-S	23	3	2	1	32	5	1	1	0	0	68
EC-NO-S	15	4	0	1	7	41	1	0	1	0	70
N-NO-D	2	2	0	0	1	0	6	0	1	0	12
TP-TS-C	4	1	0	0	0	1	0	6	0	0	12
M-NO-D	1	1	0	0	0	0	0	0	5	0	7
PR-NO-C	3	3	0	0	0	1	0	0	0	2	9
Suma	528	80	50	27	66	69	14	9	8	6	857
Particip	0,594	0,092	0,070	0,040	0,078	0,081	0,015	0,012	0,009	0,009	

Podemos observar en la tabla que la marca que tiene mayor participación de mercado es el Nescafé Normal Cafeinado (N NO C) con un 61%, éste es un hecho no trivial, ya que ésta marca es la que se lleva la mayor parte de la participación y el resto de ésta se reparte entre los demás competidores de la industria, por lo que podríamos decir, que aparentemente Nescafé no tiene un competidor importante (en su línea de Nescafé tradición), además que ésta marca es la que posee un mayor número de consumidores fieles, exactamente los consumidores fieles corresponden al 76.65% de los consumidores de Nescafé. Un hecho relevante al momento de hacer un análisis de la industria es que esta marca es la única que es sustituto y es sustituido por absolutamente todas las marcas que se están evaluando, incluyendo otras categorías de productos de Nescafé dándose así canibalismo entre ellas. Es por ésta razón que se puede pensar que Nescafé con su producto

Nescafé Tradición tiene una estrategia general de crecimiento enfocándose así a todos los consumidores y no a un nicho específico.

Otros hechos relevantes que podemos obtener de la tabla de frecuencia es que existen marcas con un gran porcentaje de fidelidad en relación a sus ventas, es decir, en que porcentaje las personas eligen en las dos ocasiones de consumo la misma marca, en relación a las ventas totales de la marca (Anexo 2-Tabla A2), algunas de las marcas destacables en ese sentido son por ejemplo Tempo, quienes poseen una fidelidad en el consumo de un 66%, Ecco y Nescafé Dolca también poseen tasas de fidelidad por sobre un 55%, quizás la fidelidad a las marcas se deba por una estrategia de nicho de éstas marcas, ya que las tres están orientadas a consumidores con un gusto en particular, claro que ésta teoría se contradice con la alta participación que presenta Nescafé Normal Cafeinado, ésa es una pregunta que sería muy interesante poder contestar a medida que se va avanzando en el estudio. Las marcas privadas por otro lado tienen una tasa de fidelidad muy baja, de hecho es la más baja en comparación con los competidores aquí mencionados.

En cuanto a los cambios de consumo de marca podemos decir que los clientes de cambio de Nescafé hacia otras marcas representan un ingreso muy importante para ellas, de hecho para algunas incluso representa el 42% de sus ventas totales, como es el caso de Nescafé Normal Descafeinado.

II. 2. 5. Simetría de la matriz

Para comprobar si la matriz de tabulación cruzada es simétrica, debemos hacer un test chi-cuadrado, donde la hipótesis nula dice que la matriz no es simétrica.

*Considerando que tenemos 10 marcas tendremos entonces $(10 * 9 / 2) = 45$ pares.*

Para sacar el chi-cuadrado acumulado para celdas de cambio se obtiene primero tomando las frecuencias en ambas direcciones, promediándolas, y obteniendo así el número esperado de elecciones de cambio para ese par de marcas, por ejemplo, tomamos el par de marcas N-NO-C y M-MA-C, es éstas marcas tenemos las siguientes frecuencias.

$$\text{-N-NO-C/M-MA-C} = 20$$

$$\text{-M-MA-C/N-NO-C} = 25$$

Aquí el número esperado de cambios es :

$$(20 + 25) / 2 = 22.5$$

El chi-cuadrado calculado para éste par es

$$\frac{\sum \frac{(\text{Observado} - \text{Esperado})^2}{\text{Esperado}}}{2} = \frac{(20 - 22.5)^2}{22.5} + \frac{(25 - 22.5)^2}{22.5} = 0.55$$

El test chi-cuadrado usualmente es considerado válido cuando el valor esperado de la frecuencia es mayor o igual a 5, por lo que es éste caso solo 12 celdas satisfacen ésta condición.

La siguiente tabla resume los valores obtenidos para las 12 celdas.

Tabla 2-6

Celda	Valor esp.	Valor obs.
N-NO-C/M-MA-C	22,5	20
M-MA-C/N-NO-C	22,5	25
ND-NO-C/N-NO-C	22,5	31
N-NO-C/ND-NO-C	22,5	14
N-PR-C/N-NO-C	14	19
N-NO-C/N-PR-C	14	9
CD-NO-S/N-NO-C	21	23

N-NO-C/CD-NO-S	21	19
EC-NO-S/N-NO-C	14,5	15
N-NO-C/EC-NO-S	14,5	14
EC-NO-S/CD-NO-S	6	7
CD-NO-C/EC-NO-S	6	5

Así con los valores obtenidos en la tabla 2-6 y con la fórmula presentada para calcular el valor chi-cuadrado obtenemos los valores que aparecen en la tabla 2-7.

Tabla 2-7

Combinación	Valor chi-cuadrado
N-NO-C/M-MA-C	0,55
N-NO-C/ND-NO-C	6,44
N-NO-C/N-PR-C	3,571
N-NO-C/CD-NO-S	0,3809
EC-NO-S/N-NO-C	0,0344
CD-NO-C/EC-NO-S	0,333
Acumulada	11,3093

Finalmente para ver la simetría de la matriz debemos comprar el valor calculado y el de la tabla chi-cuadrado. Así el valor de la chi de tabla para un nivel de significancia del 0.05% es de 11.0705 por lo que se comprueba que la matriz de tabulación cruzada es simétrica, porque se rechaza la hipótesis nula.

II. 2. 6. Efecto promociones

Dentro de las marcas mencionadas existen algunas que han tenido promociones incluidas al momento de ser elegidas por los consumidores, éste hecho podría influir en la elección de compra, por lo que creemos es interesante hacer un análisis por separado de las marcas que al momento de ser elegidas incluían promociones y las que no lo hacían para así tratar de ver si es que influyen o no en la elección y como cambian las participaciones de mercado.

La tabla 2-8 nos muestra el consumo de productos que no incluían promociones al momento de la elección de consumo.

Tabla 2-8

	N-NO-C	M-MA-C	ND-NO-C	N-PR-C	CD-NO-S	EC-NO-S	N-NO-D	TP-TS-C	M-NO-D	PR-NO-C
N-NO-C	248	6	8	9	8	11	5	1	1	0
M-MA-C	8	20	0	2	2	2	0	0	0	0
ND-NO-C	17	0	17	1	1	1	0	0	0	3
N-PR-C	12	1	1	13	1	2	0	0	0	0
CD-NO-S	10	0	1	0	19	3	1	0	0	0
EC-NO-S	10	1	0	0	7	22	0	0	1	0
N-NO-D	2	2	0	0	1	0	6	0	0	0
TP-TS-C	1	0	0	0	0	1	0	2	0	0
M-NO-D	1	0	0	0	0	0	0	0	5	0
PR-NO-C	3	1	0	0	0	0	0	0	0	2
Suma	312	31	27	25	39	42	12	3	7	5
Particip	0,621	0,062	0,054	0,050	0,077	0,083	0,024	0,006	0,014	0,010

*Aquí las participaciones de mercado en general se mantienen ,
sabiendo la participación de mercado de Nescafé Premium Normal (“Nescafé
Fina Selección”) pasando de un 3.1% a un 5%.*

*La tabla a continuación presenta el porcentaje de las ventas de cada
marca que corresponde a productos que no incluían promoción al momento
de ser elegidas.*

Tabla 2-9

VENTAS	N-NO-C	M-MA-C	ND-NO-C	N-PR-C	CD-NO-S	EC-NO-S	N-NO-D	TP-TS-C	M-NO-D	PR-NO-C
s. prom/total	59%	39%	54%	93%	59%	61%	86%	33%	88%	83%

Podemos observar que existen productos en donde la mayor parte del consumo de éstos se realiza en su forma sin promoción, en el producto N-PR-C se podría explicar por el posicionamiento que tiene éste en el mercado, como un producto de calidad “premium”, en donde el precio es mayor que el precio promedio del mercado, otro hecho importante es que los dos productos descafeinados tienen un alto porcentaje de sus ventas como productos sin promoción, con las marcas privadas también sucede lo mismo, este comportamiento del consumidor se podría explicar por la sensibilidad de éste ante las promociones, se sabe que algunos segmentos de mercado son más sensibles que otros, dependiendo de que es lo que buscan y esperan de un producto.

En la tabla 2-9 queda el cuadro resumen de las compras con promoción ya sea en la primera compra, en la segunda o en ambas. En la tabla no están representadas todas las marcas que forman parte del estudio, si no solo las que tienen incluidas promociones. Aquí las marcas que tenían promociones incluidas serán diferencias con las letras “cp”, colocándolas al final de cada categoría , un ejemplo de esto es M-AM-C con promoción quedará de la forma M-NO-Ccp.

Como es de esperarse vemos en la tabla que los cambios de un producto con promoción a uno sin promoción son mínimos ¹⁰. En éste sentido el mayor cambio de

¹⁰ Según la teoría económica, los humanos somos entes racionales, por lo que siempre un individuo debiera preferir más que menos.

consumo, sustitución, se produce de N-NO-Ccp a N-NO-C, el cual representa alrededor del 14% del total de ventas de Nescafé relacionadas a una promoción. Ahora esta “irracionalidad en el consumo” se podría explicar por la disponibilidad del producto, es decir, al momento de la primera compra, el producto en promoción estaba disponible y en la segunda compra probablemente no, y por esa razón se produce el cambio de marca, a uno más barato u otra explicación podría ser que se trata de un consumidor sensible a las promociones. Un caso extraño en éste sentido se da en dos consumidores que eligieron en su primera compra CD-NO-S y EC-NO-s, luego en la segunda compra ambos se cambiaron a N-PR-C, donde las primeras marcas son las dos más baratas que hay en el mercado y N-PR-C es una de las más caras.

Tabla 2-10

	N	M	ND	N	CD	EC	N	TP	PR	N	M	ND	CD	EC	TP	M
	NO	MA	NO	PR	NO	NO	NO	TS	NO	NO	MA	NO	NO	NO	TS	NO
	C	C	C	C	S	S	D	C	C	Ccp	Ccp	Ccp	Scp	Scp	Ccp	Dcp
N-NO-C	0	0	0	0	0	0	0	0	0	62	8	4	7	1	0	0
M-AM-C	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
M-MA-C	0	0	0	0	0	0	0	0	0	3	5	0	0	1	0	0
ND-NO-C	0	0	0	0	0	0	0	0	0	6	1	2	1	1	0	0
N-PR-C	0	0	0	0	0	0	0	0	0	7	1	1	2	0	0	0
CD-NO-S	0	0	0	0	0	0	0	0	0	5	3	1	3	1	0	0
EC-NO-S	0	0	0	0	0	0	0	0	0	3	1	0	1	4	0	0
N-NO-D	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
TP-TS-C	0	0	0	0	0	0	0	0	0	1	0	0	0	0	1	0
M-NO-D	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
PR-NO-C	0	0	0	0	0	0	0	0	0	0	2	0	0	1	0	0
N-NO-Ccp	30	5	1	0	3	1	1	0	1	64	1	1	1	1	0	0
M-MA-Ccp	11	3	2	0	0	0	0	0	0	3	15	1	0	0	0	0
ND-NO-Ccp	6	0	4	0	0	0	0	1	0	2	0	6	0	0	0	0
CD-NO-Scp	7	0	0	1	1	1	0	1	0	1	0	0	10	0	0	0
EC-NO-Scp	1	1	0	1	0	7	1	0	0	1	1	0	0	7	0	0
TP-TS-Ccp	1	1	0	0	0	0	0	1	0	1	0	0	0	0	2	0
M-NO-Dcp	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Suma	56	11	7	2	4	9	2	3	1	159	38	16	25	17	3	1
Particip	0,158	0,031	0,020	0,006	0,011	0,028	0,006	0,008	0,003	0,449	0,107	0,045	0,068	0,048	0,006	0,003

En cuanto de los cambios de una marca cualquiera de café a un producto con promoción, nuevamente el gran ganador es N-NO-C, quienes capturan el 61% del total de los cambios que se producen a un producto con promoción. En este mismo sentido, es M-NO-Dcp la marca que obtiene el menor rendimiento, capturando solo el 0.3% de los cambios.

II. 2. 7. Estacionalidad

Hasta el momento hemos analizados solo las dos primera compras hechas por un consumidor en el año, pero existe un elemento que no se está controlando en el modelo, la estacionalidad en el consumo. Puede que las compras de café sean estacionales, es decir, que existen algún factor que determine que el comportamiento del consumidor sea distinto al principio de año que a fin, para clarificar este punto es que tomamos las dos últimas compras y las sometimos al mismo tratamiento que los datos anteriores, para ver lo resultados obtenidos y para poder comparar la matrices de tabulación, que muestran la frecuencia de compra, fidelidad y sustitución de consumo es que insertamos la tabla 2-11 de tabulación cruzada de las ultimas dos compras de los mismos panelistas que pertenecen a la muestra de las tablas 2-8, 2-9, 2-10.

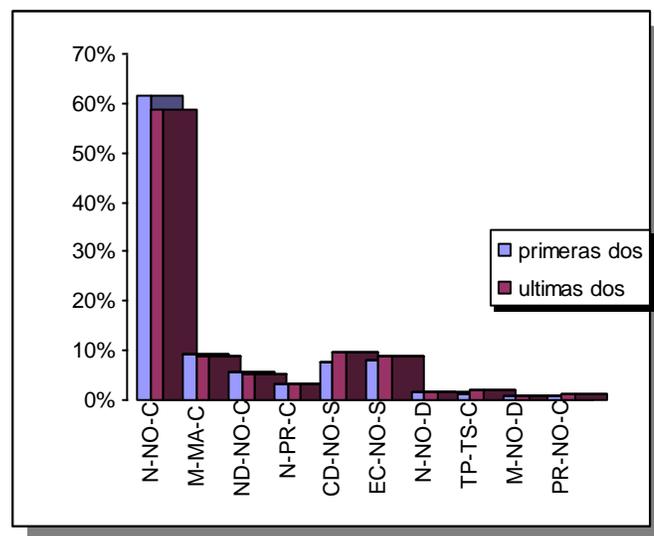
Tabla 2-11

	N-NO-C	M-MA-C	ND-NO-C	N-PR-C	CD-NO-S	EC-NO-S	N-NO-D	TP-TS-C	M-NO-D	PR-NO-C
N-NO-C	409	17	19	8	25	23	5	2	3	5
M-MA-C	24	46	0	2	2	5	0	0	0	2
ND-NO-C	18	4	20	0	2	2	0	1	0	0
N-PR-C	10	3	0	12	0	1	0	0	0	0
CD-NO-S	19	2	1	2	44	6	0	1	0	0
EC-NO-S	14	3	3	0	11	38	1	1	0	0
N-NO-D	7	0	0	1	0	0	5	0	0	0
TP-TS-C	1	0	1	1	0	0	0	11	0	0
M-NO-D	2	1	0	1	1	2	2	0	3	0
PR-NO-C	4	1	0	0	0	0	0	1	0	3
Suma	508	77	44	27	85	77	13	17	6	10
Particip	58,8%	8,9%	5,1%	3,1%	9,8%	8,9%	1,5%	2,0%	0,7%	1,2%

Para comparar los resultados de ambas tablas es que presentaremos un gráfico que ,muestra una superposición de las participaciones de mercado de ambos periodos, es decir, de las primeras dos elecciones y las últimas dos elecciones, hechas por los mismos panelistas.

En el gráfico 2-5 vemos claramente que no existe una diferencia significativa en las participaciones de mercado que pudieran hacer pensar que existen factores estacionales importantes que afecten la conducta del consumidor.

Gráfico 2-5



Capítulo 2

Metodología de estudio

II. 2. Datos de Detergente

II. 2. 1. Descripción de los datos

Esta nueva muestra posee en total 20.833 observaciones, las que corresponden a compras realizadas por alrededor de 1119 panelistas, los que deben elegir entre más de 78 productos el que finalmente comprarán. El periodo de estudio comprende 52 semanas del año 2002.

Para poder desarrollar el modelo propuesto por Grover y Srinivasan debimos sacar la participación de mercado de cada uno de los productos que se encuentran en la muestra.

La tabla 2-12 muestra los resultados obtenidos.

Tabla 2-12

partic.	marca	matic	agreg	polvo	comb	descripción
16,830%	1	1	6	1	1	OMO MAT.PVO,
13,150%	1	0	6	1	2	OMO PVO.
12,338%	2	1	6	1	3	ACE POLVO
6,983%	1	1	18	1	58	OMO MATIC ACCION CONT.
6,287%	1	1	3	1	8	OMO MAT.FRE.PVO.

6,267%	3	1	1	1	4	DRIVE MAT.PR.PVO
6,152%	3	0	1	1	9	DRIVE PRO.PVO.
5,797%	5	1	6	1	7	ARIEL PVO.
4,208%	6	1	6	1	68	RINSO MATIC PVO.
3,139%	5	1	5	0	11	ARIEL FUTUR LIQ.
2,924%	6	0	6	1	69	RINSO PVO.
1,675%	4	1	6	1	5	DETERGENTE
1,374%	39	1	6	1	64	POPEYE MATIC LIQ.
1,348%	3	1	10	1	30	DRIVE MATIC COLOR LIQ.
1,164%	1	0	3	1	57	OMO FRESH PVO.
0,955%	28	0	6	1	47	LAR PVO.
0,880%	6	1	7	1	67	RINSO MATIC BLANCURA
0,858%	3	1	11	0	32	DRIVE ULTRAFLUID LIQ.
0,798%	3	1	9	1	29	DRIVE MATIC BIO-ACTIVO PVO.
0,554%	7	1	7	1	12	ACUENTA ALBO MATIC PVO.
0,537%	12	0	6	1	19	BOLD 3 PVO.
0,514%	6	0	7	1	66	RINSO BLANCURA PVO.
0,441%	42	0	6	1	76	TISBE PVO.BL.
0,426%	39	0	6	1	65	POPEYE PVO.
0,421%	40	1	16	1	70	SKIP INT.ACCION ANTISARRO
0,403%	27	0	10	1	46	KORALL COLOR PLUS PVO
0,281%	30	0	6	1	53	LINA PVO.
0,259%	44	0	7	1	77	UNIMARC BLANQUISIMO PVO.
0,249%	21	0	6	1	39	FUZOL PVO.
0,201%	44	1	11	1	78	UNIMARC ULT.MAT.PVO.
0,188%	13	1	6	1	20	BOLD MATIC PVO.
0,148%	5	1	7	0	14	ARIEL ALPES LIQ.
0,141%	11	1	11	1	18	BISSEL FULL MATIC
0,140%	37	0	6	1	61	PERSIL PVO.
0,140%	40	1	19	1	72	SKIP INT.SPRING LIQ.
0,140%	24	0	6	1	43	KEL PVO.CTN.
0,135%	40	1	6	1	73	SKIP INTELLIGENT
0,125%	5	1	17	1	15	ARIEL MAQ.FRONTAL PVO.
0,113%	19	0	7	1	33	EKONO BLANCO
0,113%	3	1	6	1	31	DRIVE MATIC PVO.
0,097%	1	1	11	1	59	OMO ULT.MAT.
0,093%	29	1	11	1	52	LIDER ULT.MAT.PVO.
0,078%	25	1	11	1	44	KLAREN PLUS ULTRA PVO
0,077%	51	0	6	1	28	DIAMANTE PVO.
0,073%	7	1	6	0	13	ACUENTA MATIC LIQ.
0,057%	53	0	6	1	54	LORE PVO.
0,056%	21	1	13	0	38	FUZOL PREN.DEL.LIQ.
0,054%	14	1	6	1	21	BRISA 3 MATIC PVO.
0,053%	46	0	6	1	81	ZIG ZAG PVO.
0,049%	50	1	6	1	27	DETTER MATIC PVO.
0,047%	41	0	11	1	75	SUPER-TI PVO.
0,043%	29	0	6	1	49	LIDER PVO.
0,034%	31	0	6	1	56	MONTECARLO PVO.
0,033%	29	1	7	1	51	LIDER ULT.CONCEN.EXT.BLANCO PV O.

0,031%	38	0	13	1	62	PERWOLL LANA Y ROPA DELICADA P VO.
0,028%	29	1	11	1	48	LIDER MAT.ULT.PVO.
0,026%	47	0	6	1	0	ABRIL PVO.
0,025%	19	1	11	1	37	EKONO ULT.MAT.PVO.
0,025%	15	1	6	1	25	CINCO CONTINENTES ULT.MAT.PVO
0,023%	14	0	6	1	22	BRISA 3 PVO.
0,022%	31	1	11	1	55	MONTECARLO MAT.ULT.CONC.PVO.
0,022%	19	0	6	1	35	EKONO PVO
0,021%	39	1	6	1	63	POPEYE LIQ.
0,021%	23	0	6	1	40	IBERIA PVO.
0,019%	15	0	6	1	24	CINCO CONTINENTES PVO.
0,017%	52	0	6	1	41	KAS PVO
0,015%	24	0	12	1	42	KEL DOBLE ACCION PVO.
0,013%	54	1	6	0	79	WOOLITE LIQ.
0,012%	9	0	6	1	16	BILORE PVO.
0,009%	26	1	6	1	45	KOP MATIC PVO.
0,008%	10	1	6	1	17	BIO WASH MATIC PVO.
0,006%	48	1	6	1	23	CASABLANCA MATIC PVO.
0,005%	19	1	6	1	34	EKONO LIQ.
0,004%	41	1	11	1	74	SUPER TI ULT.MAT.PVO.
0,004%	55	0	11	1	80	ZAS ULTRA LIMPIEZA PVO
0,003%	29	0	13	1	50	LIDER ROPA DELICADA LIQ.
0,003%	49	1	6	1	26	DEMAX MATIC PVO.
0,002%	19	0	13	1	36	EKONO ROPA DELICADA PVO.

La tabla muestra la participación de cada producto y la descripción de los mismos. Podemos ver que existen variables que están definidas en la tabla, cada variable representa a alguna característica del producto, como marca (que identifica a que marca pertenece el producto),

matic (nos dice si se refiere a un detergente para lavado a mano o lavadora automática), agregado (se refiere a las características especiales que pudiera tener como olor, para ropa delicada, etc.), polvo (nos indica la condición física, es decir, si se trata de un producto líquido, en polvo, etc), el detalle de éstas variables las presentamos en el anexo 5.

De acuerdo a los resultados obtenidos, las marcas que posee la mayor participación de mercado son:

Tabla 2-13

partic.	comb	Descripción
16,83%	1	OMO MAT.PVO,
13,15%	2	OMO PVO.
12,34%	3	ACE POLVO
6,98%	58	OMO MATIC ACCION CONT.
6,29%	8	OMO MAT.FRE.PVO.

6,27%	4	DRIVE MAT.PR.PVO
6,15%	9	DRIVE PRO.PVO.
5,80%	7	ARIEL PVO.
4,21%	68	RINSO MATIC PVO.
3,14%	11	ARIEL FUTUR LIQ.

Por simplificación, el nombre de los productos va a ser reemplazado para algunas partes de nuestro estudio por la variable combinación, la cual resume a las características por las que éstos productos fueron descritos anteriormente.

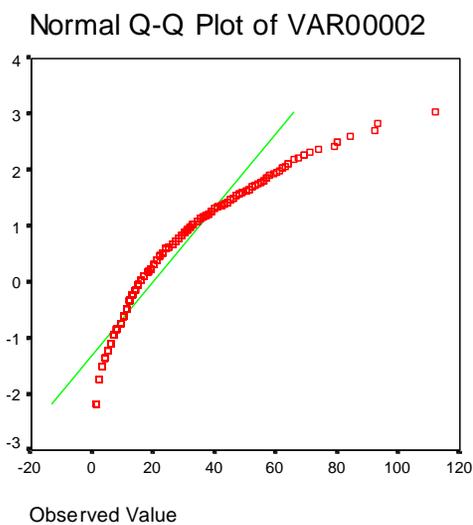
Como dijimos anteriormente para poder desarrollar el modelo de segmentación propuesto por Grover y Srinivasan en el año 1987, debemos limpiar la muestra dejando los datos correspondientes solo a compras de las marcas con mayor participación y que entraran en el estudio, en este caso fueron las 10 marcas que aparecen en la tabla 2-13. De ésta, manera se redujeron los datos a 17329.

III. 2. 2. Distribución de los datos

Sabemos que el conocer la distribución de los datos, es importante ya que si es que siguieran una distribución conocida, se podría hacer inferencia estadística de los datos, además puede ser muy útil para trabajar con los datos.

Para ver si es que la distribución de los datos se ajusta a una normal, haremos algunos test, los

mismos ocupados para ver la distribución de los datos del café.



Lo que vemos a priori en éste gráfico, es que los datos no se ajustan perfectamente a la línea, si los datos (línea roja) se ajustaran de mejor forma a la línea que estandariza la normal (línea verde), podríamos inferir que se acerca la distribución a una normal.

Gráfico 2-6

Para corroborar el resultado anterior es que haremos un test de normalidad, el test llamado Kolmogorov- Smirnov y el Shapiro-Wilk, dos test que fueron explicados anteriormente y que tienen como hipótesis nula el que los datos se ajusten a una normal, si así fuera, la celda sig, que quiere decir, nivel de significancia, debiera tener un valor mayor que 0.005.

Test de normalidad

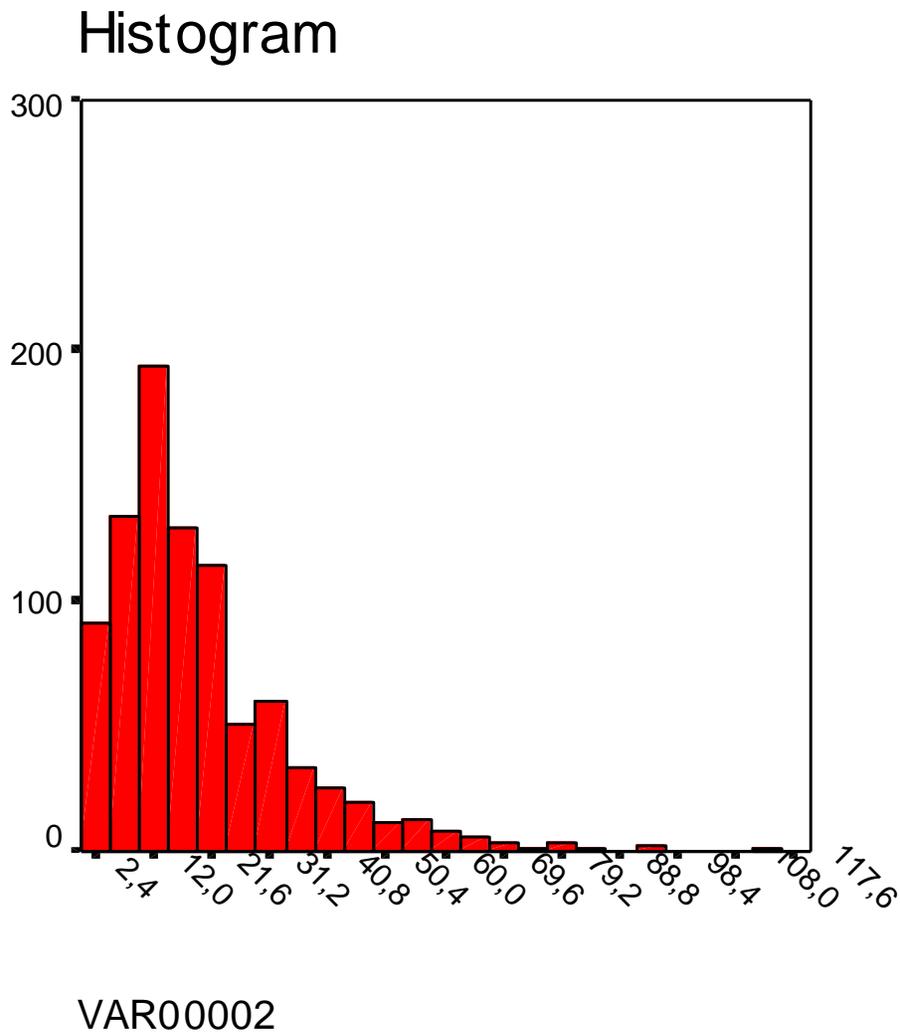
	Kolmogoro v-Smirnov Statistic	df	Sig.	Shapiro- Wilk Statistic	Df	Sig.
VAR00002	,128	902	,000	,867	902	,000

a. Lilliefors Significance Correction

Como vemos el test rechaza la hipótesis nula, por lo que podemos decir con propiedad que los datos no siguen una distribución normal.

De igual forma hicimos un histograma con los datos para ver si es que existen datos innecesarios, es decir, que sean como un tipo de outlier de la muestra (que la frecuencia de compra de algunos consumidores sea demasiado inusual y por lo tanto, no reflejen al general de los consumidores).

Gráfico 2-7



El gráfico 2-7, nos muestra que efectivamente existen datos innecesarios en la muestra, todos los panelistas cuyas ocasiones de compra exceden las 74 debiéramos sacarlas.

Es así como existen 8 panelistas que poseen éstas características de consumo en el año en estudio. La tabla 2-13 muestra cuales son los panelistas y el número de ocasiones de compra .

Tabla 2-13

<i>panelista</i>	<i>frecuencia</i>
11342	74
426	79
11077	80
11098	80
218	84
11015	92
821	93
11530	112

II. 2. 3. Elección de las compras

Como el modelo de Grover y Srinivasan requiere la tabulación cruzada de los datos, se deben seleccionar dos compras por panelistas, las cuales deben cumplir con las mismas condiciones que en el proceso de elección de compras de los datos del café, éstas son:

- *Dos compras se refiere a compras realizadas en dos ocasiones distintas, eso implica que dos compras hechas en el mismo día no pueden ser elegidas, es decir, se deberá elegir a una de ellas y como segunda compra una hecha en un día posterior. Como no hay forma de saber a priori cual es la verdadera razón por la que aparecen en el panel dos compras en el mismo día, en caso de que las compras hechas en el mismo día sean de distintas marcas, se debe elegir la primera que aparezca, esto bajo el supuesto de que las personas compren en días distintos y las pasen por el escáner en el mismo día, además de que el orden de compra sea igual al orden en el que fueron pasados los productos por el escáner.*
- *Si es que ocurre que un panelista tiene solo dos compras y las dos fueron hechas en el mismo día, éste panelista debe ser eliminado de la muestra ya que no cumple con la condición o requerimientos del modelo. Lo mismo ocurrirá si es que un panelista tiene solo una compra.*

De acuerdo a éstas condiciones, se realizó la selección de las dos compras por panelistas que representarán al mercado del detergente. Es así como finalmente en la muestra quedaron solo 2024 observaciones.

II. 2. 4. Tabulación cruzada

Con la muestra final y ocupando el programa SPSS, obtuvimos la tabla de contingencia que muestra cuales las compras cruzadas hechas en dos ocasiones distintas.

Tabla 2-14

Comb	1	2	3	4	7	8	9	11	58	68	Total
1	56	18	23	4	6	6	7	1	13	6	140
2	14	174	31	3	9	3	29	3	17	6	289
3	10	20	96	0	7	4	4	3	8	7	159
4	10	2	2	21	2	5	8	3	1	2	56
7	3	6	12	2	27	0	6	1	0	1	58
8	11	5	4	3	2	19	2	1	6	1	54
9	10	17	7	4	9	4	62	0	3	0	116
11	0	0	3	1	2	0	1	10	0	0	17
58	7	11	3	5	2	1	6	0	34	3	72
68	7	4	6	0	0	3	1	2	3	25	51
Total	128	257	187	43	66	45	126	24	85	51	
Particip.	13,24%	26,98%	17,09%	4,89%	6,13%	4,89%	11,96%	2,03%	7,76%	5,04%	

Recordemos que en ésta ocasión para nominar a cada una de las marcas usamos números llamados combinación, las que fueron descritas en tablas anteriores, éstas combinaciones resumen las características de cada producto.

A diferencia de los datos del café, las participaciones de mercado en el mercado de detergentes no es tan concentrada, de hecho la marca que posee la mayor participación de mercado es la combinación 2 (Omo en polvo) y la marca que posee la menor participación de mercado es la combinación 11 (Ariel Futur).

II. 2. 5. Simetría de la matriz

De acuerdo a lo expuesto en el paper de Grover y Srinivasan para constatar que la matriz de tabulación cruzada es simétrica debemos realizar el test de chi-cuadrado, para esto ocupamos el programa SPSS.

La tabla 2-15 muestra las frecuencias obtenidas en la tabulación cruzada, versus la frecuencia esperada.

Tabla 2-15

	<i>Observed N</i>	<i>Expected N</i>	<i>Residual</i>
1	268	202,4	65,6
2	546	202,4	343,6
3	346	202,4	143,6
4	99	202,4	-103,4
7	124	202,4	-78,4
8	99	202,4	-103,4
9	242	202,4	39,6
11	41	202,4	-161,4
58	157	202,4	-45,4
68	102	202,4	-100,4

Finalmente, en la tabla 2-16 mostramos el resultado obtenido del test chi-cuadrado, además muestra el nivel de significancia, según el cual

deberíamos rechazar la hipótesis nula, por lo que podemos inferir que la tabla de contingencia es simétrica.

Tabla 2-16

<i>Chi-Square</i>	1038,905
<i>Df</i>	9
<i>Asymp. Sig.</i>	0

Capítulo 3

Resultado y análisis

III:1 Introducción

Una vez limpiado los datos de café y detergentes, procederemos a sacar los resultados, es decir, el número de segmentos óptimos por producto, interpretarlos y posteriormente analizarlos.

III. 2 Datos de café

III. 2. 1. Determinando el número de segmentos de cambio

Como se tiene muy poca información a priori del número de segmentos de cambio, se deben usar distintos criterios que nos acerquen a ellos. Uno de los mencionados criterios es el R^2 ajustado¹¹. Para el caso de éste estudio utilizaremos el criterio AIC, presente en el programa Winmira 2001, el proceso se hace de la siguiente forma, se hace correr el programa primero para

¹¹ Grover y Srinivasan 1987

un segmento, y luego se va aumentando el número de segmentos uno en uno, hasta llegar a un número suficiente de segmentos que nos permita elegir el número adecuado. Se tomará como número óptimo de segmentos de cambio aquél que de el AIC¹² que sea como un punto de inflexión, donde el criterio AIC es una medida que nos indica la bondad de ajuste del modelo, cuya fórmula es la siguiente:

$$24) \quad AIC = \ln \frac{e'e}{n} + 2k$$

En éste caso se busca minimizar la suma del cuadrado de los errores, pero el criterio agrega una penalización, la cual crece con el número de regresores.

Para poder trabajar con el programa Winmira 2001 debimos cambiar la forma de los datos, ya que el imput requerido por el programa es de variables discretas, es así que tomamos todas las elecciones y mediante la utilización de funciones lógicas existentes en el programa Excel, cambiamos los datos con los cuales se logró la tabulación cruzada que estaban contenidos en la tabla A1 (Anexo 1) a variables discretas, finalmente quedaron 10 columnas correspondientes a cada una de las marcas que conformaban la muestra, las elecciones eran marcadas con un 1, las elecciones en dos ocasiones, es decir,

¹² Kamakura and Russell 1989

las elecciones leales eran marcadas con un 2, y las marcas que no habían sido elegidas por ese panelista eran marcadas con un 0 (Figura 3-1).

Figura 3-1

Case	V1	V2	V3	V4	V5	V6	V7	V8	V9	V10	N.OBS
554	0.00	2.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	554.00
555	0.00	2.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	555.00
556	0.00	2.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	556.00
557	0.00	2.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	557.00
558	0.00	2.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	558.00
559	0.00	2.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	559.00
560	0.00	2.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	560.00
561	0.00	2.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	561.00
562	0.00	1.00	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	562.00
563	0.00	1.00	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	563.00
564	0.00	1.00	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	564.00
565	0.00	1.00	0.00	1.00	0.00	0.00	0.00	0.00	0.00	0.00	565.00
566	0.00	1.00	0.00	1.00	0.00	0.00	0.00	0.00	0.00	0.00	566.00
567	0.00	1.00	0.00	0.00	1.00	0.00	0.00	0.00	0.00	0.00	567.00
568	0.00	1.00	0.00	0.00	1.00	0.00	0.00	0.00	0.00	0.00	568.00
569	0.00	1.00	0.00	0.00	0.00	1.00	0.00	0.00	0.00	0.00	569.00
570	0.00	1.00	0.00	0.00	0.00	1.00	0.00	0.00	0.00	0.00	570.00

Además se incluyó una variable que indicara el número de observación, esto porque el programa requería que al momento de estimar el modelo, se dejara al menos una de las variables que pertenecían al imput, fuera¹³.

Como necesitábamos sacar el número de segmentos de cambio óptimo, separamos los datos pertenecientes a compras leales y los pertenecientes a compras de cambio. Así quedaron del total de 857 solo 276 compras, es decir, un 32%.

Con esos datos fuimos haciendo estimaciones que partieron con un segmento hasta 12 (Tabla 3-1)

Tabla 3-1

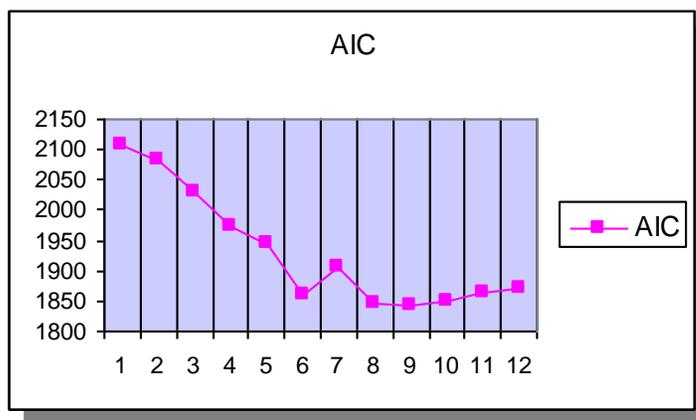
<i>Nº de segm.</i>	<i>AIC</i>
1	2106
2	2082,15
3	2028,68
4	1974,7
5	1943,88
6	1861,66
7	1904,68

¹³ Ayuda Winmira 2001.

8	1847,58
9	1843,8
10	1850,06
11	1863,04
12	1869,17

El gráfico 3-1, muestra que el número de segmentos óptimo es de 6 segmentos de cambio.

Gráfico 3-1



Así, el output obtenido en el programa fue el siguiente

Tabla 3-2

no.	label	cats	0	1	N
1	VAR00001	2	68	208	276
2	VAR00002	2	204	72	276
3	VAR00003	2	214	62	276
4	VAR00004	2	234	42	276
5	VAR00005	2	205	71	276
6	VAR00006	2	218	58	276
7	VAR00007	2	262	14	276
8	VAR00008	2	267	9	276
9	VAR00009	2	271	5	276
10	VAR00010	2	265	11	276

Aquí se ve el número de elecciones por marca, donde cada marca está representada por una variable de la siguiente forma (Tabla 3-3).

Tabla 3-3

<i>Variable</i>	<i>Marca</i>
Var 1	N-NO-C
Var 2	M-MA-C
Var 3	ND-NO-C
Var 4	N-PR-C
Var 5	CD-NO-S
Var 6	EC-NO-D
Var 7	N-NO-D
Var 8	TP-TS-C
Var 9	M-NO-D

Var 10	PR-NO-C
--------	---------

A continuación presentaremos las tablas que muestran la probabilidad de elegir una marca por cada uno de los segmentos, cabe destacar que estas tablas nos muestran la probabilidad de realizar una elección, como sabemos, el modelo presenta dos elecciones por panelista, entonces esto explica que la suma de las probabilidades no sea igual a uno, si no más bien que sea igual a dos o un poco más.

En éstas tablas nosotros podemos ver si es que algunas elecciones podrían clasificarse como con una alta lealtad, es decir la probabilidad de elegir esa marca es muy alta (cercana a uno), por lo que se clasificaría como marca casi leal¹⁴ y debiéramos sacarla de la muestra, ya que como antes se había dicho, aquí estamos analizando los segmentos de cambio.

Al mostrarnos el output probabilidades de que una marca sea comprada por una persona perteneciente a cada segmento, nos sirve para la identificación del modelo, en otras palabras, si existen marcas que tienen una probabilidad muy baja de ser elegidas, es decir, probabilidad menor que 0.01, debería esa marca ser restringida a cero¹⁵.

Tabla 3-4.a

¹⁴ Gover y Srinivasan 1987 (Segmentos casi leales)

¹⁵ Gover y Srinivasan 1987 (Identificación del modelo)

Final estimates in CLASS 1 of 6 with size 0.24481

=====

expected category frequencies and item scores:

Item label	Item`s		relative category frequencies	
	Score	Stdev	0	1
VAR00001	0.01	0.07	0.995	0.005
VAR00002	0.40	0.49	0.603	0.397
VAR00003	0.25	0.43	0.749	0.251
VAR00004	0.21	0.41	0.792	0.208
VAR00005	0.43	0.49	0.573	0.427
VAR00006	0.43	0.49	0.573	0.427
VAR00007	0.09	0.29	0.909	0.091
VAR00008	0.06	0.24	0.938	0.062
VAR00009	0.05	0.21	0.953	0.047
VAR00010	0.11	0.31	0.894	0.106

Sum: | 2.02

Tabla 3-4.b

Final estimates in CLASS 2 of 6 with size 0.17089

=====

expected category frequencies and item scores:

Item label	Item`s		relative category frequencies	
	Score	Stdev	0	1
VAR00001	1.00	0.06	0.004	0.996
VAR00002	0.00	0.07	0.995	0.005
VAR00003	0.00	0.07	0.995	0.005
VAR00004	0.59	0.49	0.408	0.592
VAR00005	0.00	0.07	0.995	0.005
VAR00006	0.00	0.07	0.995	0.005
VAR00007	0.17	0.38	0.828	0.172
VAR00008	0.11	0.31	0.892	0.108
VAR00009	0.05	0.21	0.954	0.046
VAR00010	0.09	0.28	0.913	0.087

Sum: | 2.02

Tabla 3-4.c

Final estimates in CLASS 3 of 6 with size 0.16298

=====

expected category frequencies and item scores:

Item label	Item`s		relative category frequencies	
	Sco	Stdev	0	1
VAR00001	1.00	0.06	0.004	0.996
VAR00002	0.00	0.06	0.996	0.004
VAR00003	1.00	0.06	0.004	0.996
VAR00004	0.00	0.06	0.996	0.004
VAR00005	0.00	0.06	0.996	0.004
VAR00006	0.00	0.06	0.996	0.004
VAR00007	0.00	0.06	0.996	0.004
VAR00008	0.00	0.06	0.996	0.004
VAR00009	0.00	0.06	0.996	0.004
VAR00010	0.00	0.06	0.996	0.004

Sum: | 2.02

Tabla 3-4.d

Final estimates in CLASS 4 of 6 with size 0.16293

=====

expected category frequencies and item scores:

Item label	Item`s		relative category frequencies	
	Score	Stdev	0	1
VAR00001	1.00	0.06	0.004	0.996
VAR00002	1.00	0.06	0.004	0.996
VAR00003	0.00	0.06	0.996	0.004
VAR00004	0.00	0.06	0.996	0.004
VAR00005	0.00	0.06	0.996	0.004
VAR00006	0.00	0.06	0.996	0.004
VAR00007	0.00	0.06	0.996	0.004
VAR00008	0.00	0.06	0.996	0.004
VAR00009	0.00	0.06	0.996	0.004

VAR00010| 0.00 | 0.06 | 0.996 | 0.004

Sum: | 2.02

Tabla 3-4.e

Final estimates in CLASS 5 of 6 with size 0.15225

=====

expected category frequencies and item scores:

Item Label	Item`s Score	Stdev	relative category frequencies	
			0	1
VAR00001	1.00	0.06	0.004	0.996
VAR00002	0.00	0.06	0.996	0.004
VAR00003	0.00	0.06	0.996	0.004
VAR00004	0.00	0.06	0.996	0.004
VAR00005	1.00	0.06	0.004	0.996
VAR00006	0.00	0.06	0.996	0.004
VAR00007	0.00	0.06	0.996	0.004
VAR00008	0.00	0.06	0.996	0.004
VAR00009	0.00	0.06	0.996	0.004
VAR00010	0.00	0.06	0.996	0.004

Sum: | 2.02

Tabla 3-4.f

Final estimates in CLASS 6 of 6 with size 0.10614

=====

expected category frequencies and item scores:

Item label	Item`s Score	Stdev	relative category frequencies	
			0	1
VAR00001	1.00	0.06	0.004	0.996
VAR00002	0.00	0.06	0.996	0.004
VAR00003	0.00	0.06	0.996	0.004
VAR00004	0.00	0.06	0.996	0.004
VAR00005	0.00	0.06	0.996	0.004
VAR00006	1.00	0.06	0.004	0.996
VAR00007	0.00	0.06	0.996	0.004
VAR00008	0.00	0.06	0.996	0.004
VAR00009	0.00	0.06	0.996	0.004
VAR00010	0.00	0.06	0.996	0.004

Cada una de las tablas aquí presentadas representa a un segmento en particular. Si vamos analizando cada segmento, podemos decir que el segmento 1 es bastante normal, donde tenemos a N-NO-C marca que debería ser restringida a cero¹⁶ (marca representada por la Var 1), en el segmento dos, vemos que Nescafé Normal Cafeinado (N-NO-C), tiene una probabilidad muy alta (0.996) de ser elegida por lo que también debe ser restringida, para luego ser agregada a los segmentos leales, además podemos apreciar que existen cuatro marcas que deben ser restringidas por identificación , las marcas son Monterrey Mezclas Arábicas Cafeinado, Nescafé Dolca Cafeinado, Coronado Normal Cebada Tostada y Ecco Normal Cebada Tostada (M-MA-C, ND-NO-C, CD-NO-S y EC-NO-S, respectivamente), representadas por las variables (Var 2, Var 3, Var 5 y Var 6 respectivamente).

Un caso distinto lo vemos en los segmentos 3, 4, 5 y 6, donde en todos se da un hecho bastante particular, dos marcas con un alto grado de lealtad y el resto de las marcas con probabilidades menores que 0.01. Como resultado del procedimiento seguido anteriormente, estos cuatro segmentos debieran

¹⁶ Según criterio para identificación del modelo ,Grover y Srinivasan 1987.

ser eliminados de la muestra para pasar a formar parte de los segmentos leales. Luego de éste procedimiento tenemos dos alternativas.

1. Sacar a los elementos pertenecientes a los segmentos 3, 4, 5 y 6, para pasarlos a los segmentos leales y con los datos restantes sacar un nuevo AIC, ver cuál es el nuevo número de segmentos óptimos y hacer una análisis con ellos.

2. Sacar a los elementos pertenecientes a los segmentos 3, 4, 5 y 6 y hacer una análisis con los dos segmentos de cambio restantes.

En función de tomar la mejor decisión, llevaremos a cabo las dos alternativas, eligiendo finalmente la que a nuestro juicio, nos de un resultado mas óptimo y coherente. Comenzaremos analizando la primera opción.

III. 2. 1. 1 Desarrollando primera opción

Luego de la eliminación de la muestra de segmentos de cambio a los integrantes de los segmentos antes mencionados (3, 4, 5 y 6), procedimos a sacar un nuevo AIC con los datos restantes. Finalmente quedaron solo 126 compras, las cuales representaban a los segmentos de cambio.

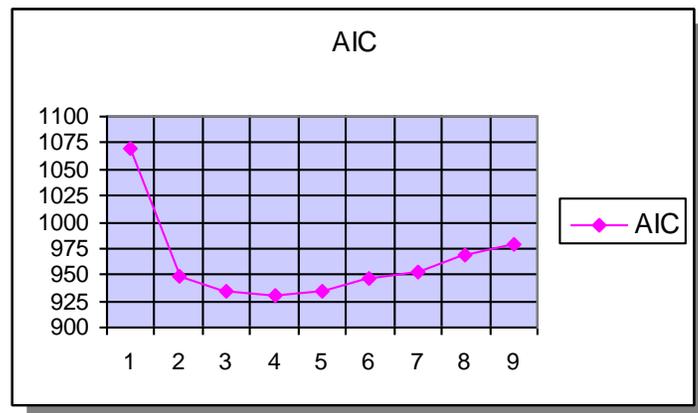
La tabla 3-5 muestra los resultados obtenidos

Tabla 3-5

<i>Nº de segm.</i>	AIC
1	1070,41
2	948,54
3	934,6
4	929,72
5	934,26
6	946,08
7	953,39
8	968,79
9	978,44

Podemos apreciar en el gráfico 3-2, que el número de segmentos óptimo es de 4.

Gráfico 3-2



Con éste resultado volvimos a sacar un output. Cabe destacar que el programa Winmira 2001, no solo entrega el output aquí mostrado, si no que además, acompañando al input aparecen la probabilidad máxima de pertenecer a un segmento, al segmento al que pertenece la compra de ése panelista y también aparecen los ajustes que se van dando a medida que se van haciendo estimaciones (Anexo 3-Tabla B1).

En la tabla 3-6 vemos la cantidad de elecciones por marca, elecciones que pertenecen a los segmentos de cambio. Podemos ver que las marcas con mayor número de elecciones son N-NO-C y N-PR-C.

Tabla 3-6

no.	label	cats	0	1	N
1	VAR00001	2	68	47	115
2	VAR00002	2	88	27	115
3	VAR00003	2	98	17	115
4	VAR00004	2	73	42	115
5	VAR00005	2	86	29	115
6	VAR00006	2	86	29	115
7	VAR00007	2	101	14	115
8	VAR00008	2	106	9	115
9	VAR00009	2	110	5	115
10	VAR00010	2	104	11	115

Ahora presentamos las tablas que nos muestran las probabilidades de que una persona perteneciente a un segmento compre la marca.

Tabla 3-7.a

Final estimates in CLASS 1 of 4 with size 0.33161

=====

expected category frequencies and item scores:

Item Label	Item`s Score	Stde	relative category frequencies	
			0	1
VAR00001	0.01	0.10	0.990	0.010
VAR00002	0.57	0.49	0.428	0.572
VAR00003	0.34	0.47	0.661	0.339
VAR00004	0.26	0.44	0.740	0.260
VAR00005	0.02	0.15	0.978	0.022
VAR00006	0.41	0.49	0.592	0.408
VAR00007	0.10	0.29	0.904	0.096
VAR00008	0.08	0.28	0.916	0.084
VAR00009	0.07	0.26	0.928	0.072
VAR00010	0.19	0.39	0.813	0.187

Sum: | 2.05

Tabla 3-7.b

Final estimates in CLASS 2 of 4 with size 0.25236

=====

expected category frequencies and item scores:

Item label	Item`s Score	Stdev	relative category frequencies	
			0	1
VAR00001	0.01	0.09	0.991	0.009
VAR00002	0.18	0.38	0.824	0.176
VAR00003	0.15	0.35	0.855	0.145
VAR00004	0.15	0.36	0.852	0.148
VAR00005	0.97	0.16	0.028	0.972
VAR00006	0.46	0.50	0.539	0.461
VAR00007	0.08	0.27	0.921	0.079
VAR00008	0.04	0.21	0.956	0.044
VAR00009	0.01	0.09	0.991	0.009
VAR00010	0.01	0.09	0.991	0.009

Sum: | 2.05

Tabla 3-7.c

Final estimates in CLASS 3 of 4 with size 0.24349

=====

expected category frequencies and item scores:

Item label	Item`s		relative category frequencies	
	Score	Stdev	0	1
VAR00001	0.99	0.10	0.010	0.990
VAR00002	0.01	0.09	0.991	0.009
VAR00003	0.01	0.09	0.991	0.009
VAR00004	0.99	0.10	0.009	0.991
VAR00005	0.01	0.09	0.991	0.009
VAR00006	0.01	0.09	0.991	0.009
VAR00007	0.01	0.09	0.991	0.009
VAR00008	0.01	0.09	0.991	0.009
VAR00009	0.01	0.09	0.991	0.009
VAR00010	0.01	0.09	0.991	0.009

Sum: | 2.05

Tabla 3-7.d

Final estimates in CLASS 4 of 4 with size 0.17254

=====

expected category frequencies and item scores:

Item Label	Item`s		relative category frequencies	
	Score	Stdev	0	1
VAR00001	0.97	0.18	0.035	0.965
VAR00002	0.01	0.10	0.991	0.009
VAR00003	0.01	0.10	0.991	0.009
VAR00004	0.01	0.11	0.988	0.012
VAR00005	0.01	0.09	0.991	0.009
VAR00006	0.01	0.10	0.991	0.009
VAR00007	0.44	0.50	0.564	0.436
VAR00008	0.26	0.44	0.740	0.260
VAR00009	0.13	0.34	0.867	0.133

VAR00010| 0.21 | 0.41 | 0.791 | 0.209

Sum: | 2.05

En la tabla 3-7.a vemos las probabilidades representativas del segmento 1, aquí restringimos a cero dos marcas (N-NO-C y CD-NO-S), representadas por las Var 1 y la Var 5.

En la tabla 3-7.b tenemos las probabilidades correspondientes al segmento 2, vemos que existen tres marcas que debieran ser restringidas a cero, por no tener una alta probabilidad de elección, en éste caso las marcas son N-NO-C, M-NO-C y PR-NO-C, representadas por las variables Var 1, Var 9 y Var 10. Además existe una marca (CD-NO-S) que tiene una probabilidad muy alta de ser elegidas por personas que pertenezcan a este segmento, por lo que éstas compras podrían ser clasificadas como casi leales. Las compras correspondientes a esas marcas serán separadas de la muestra para adherirse a los segmentos leales.

La tabla 3-7.c muestra las probabilidades correspondientes al segmento 3, aquí vemos un caso especial, existen dos marcas que tienen altas probabilidades y el resto de las marcas tienen probabilidades menores que 0.01. Por las características de este segmento lo sacaremos de los segmentos de cambio y haremos que forme parte de los segmentos leales.

En la tabla 3-7.d tenemos las probabilidades correspondientes al segmento 4, aquí tenemos una marca (N-NO-C) con alta probabilidad de ser elegida (casi leal), además tenemos cinco marcas que deben ser restringidas a

cero pues poseen una probabilidad menor de 0.01 de ser elegida, éstas marcas son las siguientes M-MA-C, ND-NO-C, N-PR-C, CD-NO-S y EC-NO-S.

En la siguiente tabla mostraremos cuales fueron los elementos de la muestra de segmentos de cambio (Anexo 5-Tabla B1) que fueron sacados y mandados a ser parte de los segmentos leales.

Tabla 3-8

<i>Observación</i>	<i>Segmento</i>	<i>Variables</i>
190-208	4	Var1
249-252	2	Var 5
217-221	2	Var 5
239-242	2	Var 5
249-232	2	Var 5
236-239	2	Var 5
260-271	2	Var 5
91-113	3	Var 1 y Var 4

Luego de tener definidos a los integrantes de los segmentos de cambio procedimos a obtener el output deseado (para esto utilizamos un método que aparece en las tablas 3-9), una tabla que nos muestre las probabilidades que las persona perteneciente a un segmento compre una marca, pero que la sumatoria de las probabilidades sea igual a uno. Para obtener éstos resultados utilizamos las tablas 3-9.

Hay que recordar que algunas probabilidades fueron restringidas porque tenían un valor menor que 0.01 y otras porque pasaron a formar parte de los segmentos leales, donde a los primeros los marcamos con un color rojo y a los últimos de color azul.

Tabla 3-9.a

Segmento 1

Marcas	Var 1	Var 2	Var 3	Var 4	Var 5	Var 6	Var 7	Var 8	Var 9	Var 10
	0	1	1	0	0	0	0	0	0	0
	0	1	1	0	0	0	0	0	0	0
	0	1	1	0	0	0	0	0	0	0
	0	1	1	0	0	0	0	0	0	0
	0	1	0	1	0	0	0	0	0	0
	0	1	0	1	0	0	0	0	0	0
	0	1	0	1	0	0	0	0	0	0
	0	1	0	1	0	0	0	0	0	0
	0	1	0	0	0	1	0	0	0	0
	0	1	0	0	0	1	0	0	0	0
	0	1	0	0	0	1	0	0	0	0
	0	1	0	0	0	1	0	0	0	0
	0	1	0	0	0	1	0	0	0	0
	0	1	0	0	0	0	1	0	0	0
	0	1	0	0	0	0	1	0	0	0
	0	1	0	0	0	0	0	1	0	0
	0	1	0	0	0	0	0	0	0	1
	0	1	0	0	0	0	0	0	0	1
	0	1	0	0	0	0	0	0	1	0
	0	0	1	1	0	0	0	0	0	0
	0	0	1	1	0	0	0	0	0	0
	0	0	1	1	0	0	0	0	0	0

	0	0	0	0	1	1	0	0	0	0
	0	0	0	0	1	1	0	0	0	0
	0	0	0	0	1	1	0	0	0	0
	0	0	0	0	1	1	0	0	0	0
	0	0	0	0	1	1	0	0	0	0
	0	0	0	0	1	0	1	0	0	0
	0	0	0	0	1	0	1	0	0	0
	0	0	0	0	1	0	0	1	0	0
Total	0	5	4	4	0	13	2	1	0	0
Prob.	0,00	0,17	0,14	0,14	0,00	0,45	0,07	0,03	0,00	0,00

Tabla 3-9.c

Segmento 3

Marcas	Var 1	Var 2	Var 3	Var 4	Var 5	Var 6	Var 7	Var 8	Var 9	Var 10
	1	0	0	0	0	0	1	0	0	0
	1	0	0	0	0	0	1	0	0	0
	1	0	0	0	0	0	1	0	0	0
	1	0	0	0	0	0	1	0	0	0
	1	0	0	0	0	0	1	0	0	0
	1	0	0	0	0	0	1	0	0	0
	1	0	0	0	0	0	1	0	0	0
	1	0	0	0	0	0	1	0	0	0
	1	0	0	0	0	0	0	1	0	0
	1	0	0	0	0	0	0	1	0	0
	1	0	0	0	0	0	0	1	0	0
	1	0	0	0	0	0	0	1	0	0
	1	0	0	0	0	0	0	0	0	1
	1	0	0	0	0	0	0	0	0	1
	1	0	0	0	0	0	0	0	0	1
	1	0	0	0	0	0	0	0	0	1
	1	0	0	0	0	0	0	0	1	0
	1	0	0	0	0	0	0	0	1	0
Total	0	0	0	0	0	0	8	5	2	4
Prob.	0,00	0,00	0,00	0,00	0,00	0,00	0,42	0,26	0,11	0,21

Finalmente los resultados obtenidos los resumimos en la siguiente tabla.

Tabla 3-10

Marcas	Segmento 1	Segmento 2	Segmento 4
N-NO-C	-	-	-
M-MA-C	0,28	0,17	-
ND-NO-C	0,17	0,14	-
N-PR-C	0,13	0,14	-
CD-NO-S	-	-	-
EC-NO-D	0,21	0,45	-
N-NO-D	0,05	0,07	0,42
TP-TS-C	0,04	0,03	0,26
M-NO-D	0,04	-	0,11
PR-NO-C	0,09	-	0,21
	1,00	1,00	1,00

Ahora procederemos a sacar las participaciones de mercado por marca y por cada uno de los segmentos, los leales y los de cambio, recordemos que las participaciones de mercado deben satisfacer las siguientes ecuaciones¹⁷ :

$$5) V_l \geq 0 \quad \text{para } l= 1,2 , \dots,n.$$

$$6) W_k \geq 0 \quad \text{para } k = 1,2 , \dots,m.$$

$$7) \sum_{l=1}^n V_l + \sum_{k=1}^m W_k = 1$$

¹⁷ Grover y Srinivasan 1987

Donde V_i corresponde a la participación de mercado de cada una de las marcas leales, y W_k corresponde a la participación de mercados de cada una de las marcas de cambio.

Para conseguir las participaciones de mercado de las marcas leales, usamos un método similar al que usamos para conseguir las probabilidades de elección de los segmentos cambiantes, la tabla B2 (anexo 2) contiene los datos.

III. 2. 2. 1 Análisis de los resultados (Primera Opción)

Finalmente la tabla 3-11 contiene los resultados obtenidos.

Tabla 3-11

Marcas	Particip.	Lealtad (V)	Participación (W)		
			0,046	0,016	0,011
			Segmento 1	Segmento 2	segmento 3
N-NO-C	0,59	0,59	-	-	-
M-MA-C	0,09	0,08	0,28	0,17	-
ND-NO-C	0,07	0,06	0,17	0,14	-
N-PR-C	0,04	0,03	0,13	0,14	-
CD-NO-S	0,07	0,06	-	-	-
EC-NO-D	0,08	0,08	0,21	0,45	-
N-NO-D	0,01	0,01	0,05	0,07	0,42
TP-TS-C	0,01	0,01	0,04	0,03	0,26
M-NO-D	0,01	0,01	0,04	-	0,11
PR-NO-C	0,01	-	0,09	-	0,21
	1,00	1,00	1,00	1,00	1,00

Según los resultados obtenidos podemos decir que los segmentos de cambio en total suman un 7.3% del total del mercado, es decir, de 857 compras, 63 pertenecen a compras hechas por personas que no son leales a una marca, el 92.7% restante de las compras corresponden a compras hechas por personas que son habitualmente leales a una marca, por lo que podríamos decir que la conducta del consumidor en éste producto es principalmente de lealtad.

Si hacemos una análisis por separado de los segmentos leales, podremos decir que en cuanto a la lealtad, Nescafé Normal Cafeinado (N-NO-C), la marca que se lleva la mayor parte de la participación de mercado por marca, es además la que posee una mayor participación por segmento (V), un 59%, de hecho ésta marca no tiene presencia en segmentos de cambio. También existe otra marca en la que sus consumidores tienen un comportamiento similar, Coronado Normal Cebada Tostada (CD-NO-S), ésta marca tiene una participación de mercado menor, un 7% y la participación de su segmento es de un 6%, pero al igual que los consumidores de Nescafé Normal Cafeinado. Las marcas restantes no poseen un alto valor de participación por segmento, incluso las marcas privadas tienen un 0% de lealtad de marca.

En cuanto a los segmentos de cambio, debemos definir cuales son las marcas que los definen, es decir, debemos ver cuales son las marcas que

prefieren las personas que forman parte de éste segmento, esto lo haremos viendo las probabilidades de elección, es decir, tomaremos a las marcas que poseen una mayor probabilidad de ser elegidas por las personas pertenecientes a ese segmento y esas serán las marcas que lo definan. En el segmento 1, que posee una participación o un peso en la muestra (parámetro W) de 4.6%, existen tres marcas que son las que tienen la mayor participación de mercado, éstas marcas son Monterrey Mezclas Arábicas Cafeinado (M-MA-C), con una probabilidad del 28 %, seguida por Ecco Normal Cebada tostada (EC-NO-S), con una probabilidad del 21% y Nescafé Dolca Normal Cafeinado (ND-N-C), con una probabilidad del 17 %. El segmento 2, el cual tiene un peso en la muestra de 1.6%, es descrito principalmente por Ecco Normal Cebada tostada (EC-NO-S), con una probabilidad del 45 %, seguido por Monterrey Mezclas Arábicas Cafeinado (M-MA-C), con una probabilidad del 17 % y Nescafé Premium Cafeinado, con una probabilidad del 14% y finalmente por Nescafé Dolca Normal Cafeinado, quien tiene también una probabilidad del 14 %. Finalmente el segmento 3 que tiene una participación por segmento de 1.1% y es explicado principalmente por Nescafé Normal Descafeinado, con una probabilidad del 42%, Tempo Tostado Suave Cafeinado, con una probabilidad del 26%, las Marcas Privadas Normal Cafeinadas, con una Probabilidad del 21% y Monterrey Normal Descafeinado, con una probabilidad del 11%.

II. 2. 1. 2 Desarrollando segunda opción

Como antes habíamos dicho en la opción dos nosotros tomaremos los datos de los segmentos 3, 4, 5 y 6 y los convertiremos en datos de compras leales, los segmentos de cambio que poseen panelistas con probabilidades muy altas de elegir una marca en particular, es llamado “segmento casi leal”. Luego de esto tomaremos los segmentos restantes, el 1 y el 2 y los analizaremos de la misma forma que lo hicimos en la primera opción.

En la tabla 3-12, mostramos como obtuvimos la probabilidad de elección de cada una marca tomando en cuenta todas las elecciones, recordemos que son dos por panelista, este es justamente el hecho que marca la diferencia entre la tabla que presenta el output del programa Winmira, donde éste output presenta la probabilidad de que un panelista elija una marca en alguna de sus dos elecciones, sin embargo necesitamos la probabilidad de elección de una marca en todas las elecciones que el panelista haga, en éste caso dos.

Tabla 3-12

Segmento 1

Marca	v1	v2	v3	v4	v5	v6	v7	v8	v9	v10
	0	1	1	0	0	0	0	0	0	0
	0	1	1	0	0	0	0	0	0	0
	0	1	1	0	0	0	0	0	0	0
	0	1	1	0	0	0	0	0	0	0
	0	1	0	1	0	0	0	0	0	0
	0	1	0	1	0	0	0	0	0	0
	0	1	0	1	0	0	0	0	0	0
	0	1	0	0	1	0	0	0	0	0
	0	1	0	0	1	0	0	0	0	0
	0	1	0	0	1	0	0	0	0	0
	0	1	0	0	1	0	0	0	0	0
	0	1	0	0	0	1	0	0	0	0
	0	1	0	0	0	1	0	0	0	0
	0	1	0	0	0	1	0	0	0	0
	0	1	0	0	0	1	0	0	0	0
	0	1	0	0	0	1	0	0	0	0
	0	1	0	0	0	1	0	0	0	0
	0	1	0	0	0	0	1	0	0	0
	0	1	0	0	0	0	1	0	0	0
	0	1	0	0	0	0	0	1	0	0
	0	1	0	0	0	0	0	0	0	1
	0	1	0	0	0	0	0	0	0	1
	0	1	0	0	0	0	0	0	0	1
	0	1	0	0	0	0	0	1	0	0
	0	0	1	1	0	0	0	0	0	0
	0	0	1	1	0	0	0	0	0	0
	0	0	1	0	1	0	0	0	0	0
	0	0	1	0	1	0	0	0	0	0
	0	0	1	0	1	0	0	0	0	0
	0	0	1	0	0	1	0	0	0	0
	0	0	1	0	0	0	1	0	0	0
	0	0	1	0	0	0	0	0	0	1
	0	0	1	0	0	0	0	0	0	1
	0	0	1	0	0	0	0	0	0	1
	0	0	0	1	1	0	0	0	0	0
	0	0	0	1	1	0	0	0	0	0
	0	0	0	1	1	0	0	0	0	0
	0	0	0	1	0	1	0	0	0	0
	0	0	0	1	0	1	0	0	0	0

1	0	0	1	0	0	0	0	0	0	
1	0	0	1	0	0	0	0	0	0	
1	0	0	1	0	0	0	0	0	0	
1	0	0	1	0	0	0	0	0	0	
1	0	0	1	0	0	0	0	0	0	
1	0	0	1	0	0	0	0	0	0	
1	0	0	1	0	0	0	0	0	0	
1	0	0	1	0	0	0	0	0	0	
1	0	0	1	0	0	0	0	0	0	
1	0	0	1	0	0	0	0	0	0	
1	0	0	1	0	0	0	0	0	0	
1	0	0	1	0	0	0	0	0	0	
1	0	0	1	0	0	0	0	0	0	
1	0	0	1	0	0	0	0	0	0	
1	0	0	1	0	0	0	0	0	0	
1	0	0	1	0	0	0	0	0	0	
1	0	0	0	0	0	1	0	0	0	
1	0	0	0	0	0	1	0	0	0	
1	0	0	0	0	0	1	0	0	0	
1	0	0	0	0	0	1	0	0	0	
1	0	0	0	0	0	1	0	0	0	
1	0	0	0	0	0	1	0	0	0	
1	0	0	0	0	0	0	1	0	0	
1	0	0	0	0	0	0	1	0	0	
1	0	0	0	0	0	0	1	0	0	
1	0	0	0	0	0	0	0	1	0	
1	0	0	0	0	0	0	0	0	1	
1	0	0	0	0	0	0	0	0	1	
1	0	0	0	0	0	0	0	1	0	
1	0	0	0	0	0	0	0	1	0	
Total	0	0	0	28	0	0	8	5	2	4
Partic	0,00	0,00	0,00	0,60	0,00	0,00	0,17	0,11	0,04	0,09

De acuerdo a éstas tablas, fueron restringidas a cero en el segmento 1, N-NO-C (representada por V1), porque no tenían ninguna elección,. Y en el segmento dos fueron restringidas por la misma razón las marcas M-MA-C,

ND-NO-C, CD-NO-S y EC-NO-S (representadas por V2, V3, V5, V6, respectivamente), en el mismo segmento, fueron restringidos a cero y llevados a formar parte de los segmentos leales. Los datos correspondientes a compras hechas para la marca N-NO-C (representada por la V1).

III. 2. 2. 2 Análisis de los resultados (Segunda Opción)

De acuerdo a los resultados aquí obtenidos y a un tratamiento de datos igual que el efectuado en la opción 1, obtuvimos el siguiente resultado.

Tabla 3-14

Marcas	Particip.	Lealtad (V)	Participación (W)	
			0,079	0,027
			Segmento 1	Segmento 2
N-NO-C	0,59	0,57	-	-
M-MA-C	0,09	0,08	0,20	-
ND-NO-C	0,07	0,06	0,13	-
N-PR-C	0,04	0,02	0,10	0,60
CD-NO-S	0,07	0,06	0,21	-
EC-NO-S	0,08	0,06	0,21	-
N-NO-D	0,01	0,01	0,04	0,17
TP-TS-C	0,01	0,01	0,03	0,11
M-NO-D	0,01	0,01	0,02	0,04
PR-NO-C	0,01	0,00	0,05	0,09
	1,00	1,00	1,00	1,00

Obviamente las participaciones de mercado, en comparación con los resultados obtenidos en la opción 1, no cambian, pero si cambia todo lo demás.

En cuanto a los segmentos de cambio, podemos decir que en el segmento 1 existen dos marcas que poseen las más altas probabilidades y podría decirse que son las marcas que identifican a éste segmento, Ecco Normal Cebada Tostada y Coronado Normal Cebada Tostada con una probabilidad de un 21%, en éste sentido a éste segmento lo podríamos llamar como el *segmento de Cebada Tostada*, ya que ésta es la característica que poseen ambas marcas. En el segmento dos podemos ver que las marcas que poseen una mayor probabilidad de elección y las que podríamos decir definen al segmento serían, Nescafé Premium Cafeinado y Nescafé Normal Descafeinado, por las características que presentan éstos segmentos podríamos denominarlo como el *segmento de Nescafé premium*, ya que los dos productos corresponden a distintas categorías de Nescafé, además son las dos marcas mas caras de la muestra (Anexo 4-Tabla B3).

III. 2. 3 Grados de libertad y condición de identificación

Los grados de libertad para el modelo estará dados por la siguiente ecuación :

$$d.f = n(n - m - 1) + t$$

Una condición necesaria para la identificación es que los grados de libertad no sean negativos. La condición necesaria y suficiente para la identificación es que la matriz formada tomando las derivadas parciales de cada S_{ij} no redundantes (las filas) con respecto a cada uno de los parámetros estimados (Columnas de la matriz), tenga completo el rango de las columnas.

Grados de libertad primera opción

En éste caso en particular tenemos los siguientes valores:

- $n= 10$ (número de segmentos leales).
- $m= 3$ (número de segmentos de cambio).
- $t= 0$ (segmentos restringidos a cero por identificación)¹⁸.

Finalmente, la ecuación arrojó el siguiente resultado.

$$\mathbf{d.f = 10(10-3-1)+0}$$

$$\mathbf{d.f = 60}$$

¹⁸ A pesar de que en las tablas aparecen parámetros marcados con color rojo, que indica que fueron restringidos a cero por tener una probabilidad menor que 0.01, en realidad esas marcas no tenían compras en ese segmento, por lo que los parámetros no fueron restringidos a priori.

Al ser los grados de libertad positivos, podemos decir que se cumple con la condición necesaria y suficiente para la identificación del modelo.

Grados de libertad segunda opción

En éste caso en particular tenemos los siguientes valores:

- $n = 10$ (número de segmentos leales).
- $m = 2$ (número de segmentos de cambio).
- $t = 0$ (segmentos restringidos a cero por identificación)¹⁹.

Finalmente, la ecuación arrojó el siguiente resultado.

$$\mathbf{d.f = 10(10-2-1)+0}$$

$$\mathbf{d.f = 70}$$

Al ser los grados de libertad positivos, podemos decir que se cumple con la condición necesaria y suficiente para la identificación del modelo.

III. 2. 4 Decisión de opción

Si vemos las tablas que contienen los segmentos de cambio de cada opción, y tratamos de interpretar a los segmentos que ahí aparecen, nos

¹⁹ Ocurre lo mismo que con el “t” de la primera opción.

*daremos cuenta de que la tarea se hace mucho más difícil en la opción 1, que en la opción 2, de hecho, los segmentos que aparecen en la opción 1, no tienen interpretación lógica aparente. Ya que las marcas que poseen una mayor probabilidad de elección no tienen factores comunes con los cuales pudiéramos nominar al segmento. No ocurre lo mismo con la opción dos, los segmentos de cambio de la opción dos son fácilmente nominables, de hecho ahí se identifican dos segmentos de cambio, uno perteneciente al **segmento de Cebada tostada**, donde las marcas con mayor probabilidad de elección son Ecco y Coronado, el segundo segmento de cambio también es fácilmente interpretable, de hecho aquí son dos categorías de la marca Nescafé las que posee una mayor participación de mercado, los productos son Nescafé Premium (Fina Selección) y Nescafé descafeinado, otra característica común en éstos dos productos es que tienen el primer y segundo lugar, respectivamente, del ranking de precios, es decir, son las marcas más caras de la muestra, por lo que lo denominamos, **segmento Nescafé Premium**.*

De acuerdo a lo aquí expuesto, es fácil concluir cual debiera ser la opción a tomar, obviamente nos quedaremos con la opción 2, básicamente seguidos por el criterio de lo razonable, siempre será mejor una opción que tenga en si una interpretación lógica de los resultados.

II. 2. 5 El mercado en base a los resultados

En éste sentido, podríamos incluir algunas características de los productos para poder hacer una análisis un poco más acabado del mercado.

En al tabla siguiente presentamos la marca corporativa de cada uno de los productos del estudio, se ve fácilmente que son dos los grandes colosos del mercado del café.

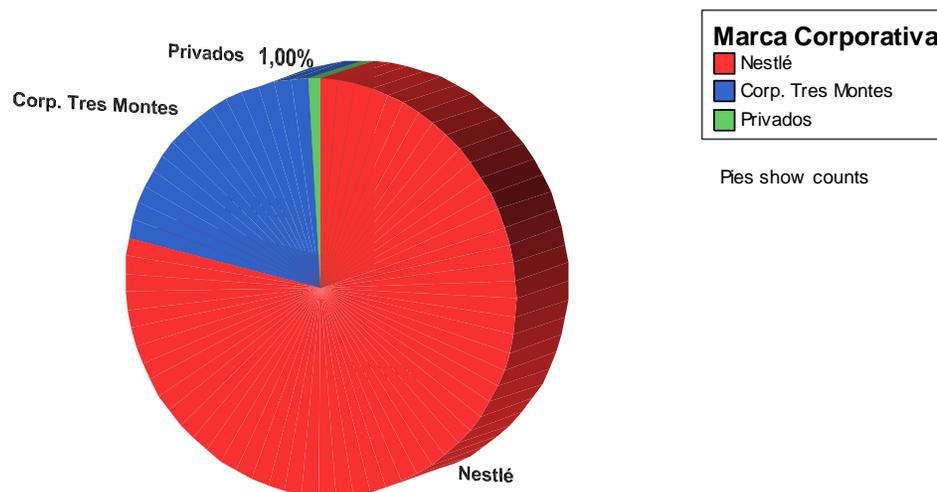
Tabla 3-15

Marcas	Fábrica
N-NO-C	Nestlé
M-MA-C	C. Tres Montes
ND-NO-C	Nestlé
N-PR-C	Nestlé
CD-NO-S	Nestlé
EC-NO-D	C. Tres Montes
N-NO-D	Nestlé
TP-TS-C	C. Tres Montes
M-NO-D	C. Tres Montes
PR-NO-C	Privados

En el gráfico 3-3 presentaremos las participaciones de mercado por marca corporativa (aquí solo se incluyen las marcas contenidas en éste estudio). Ahí vemos que lejos la marca corporativa que posee la mayor

participación es Nestlé, que sumando las marcas y categorías de producto que aquí están dan un 78%, por otro lado Tres Montes se queda con un 19% de la participación, las marcas privadas no tienen mucha presencia, capturando solo un 1% del mercado.

Gráfico 3-3



La marca perteneciente a Nestlé quien posee una mayor participación es Nescafé Normal Cafeinado, existe un hecho relevante en cuanto a la marca que posee una mayor participación de mercado perteneciente a Corp. Tres Montes, Monterrey Mezclas Arábicas Cafeinado la marca con mayor

presencia superando a las tradicionales y más antiguas marcas que pertenecían a categorías de producto de Monterrey, como son Monterrey Descafeinado y Monterrey Normal Cafeinado, ha superado a éstas con creces logrando el segundo lugar en las participaciones de mercado. Ahora, esto es relevante no solo porque ésta nueva categoría desplazó a las anteriores si no además porque en el mercado común o masivo del café, los nuevos productos que han aparecido, que traen café con algún agregado como Moka, Amaretto, Leche, Capuccino y el mismo Mezclas Arábicas han ido tomando fuerza, la inclusión de ésta categoría de producto por parte de Monterrey ha sido una buena apuesta, ahora ¿Esto querrá decir algo?, ¿Será que los conservadores Chilenos están cambiando sus gustos y preferencias, volviéndose más audaces en el consumo y llegando a consumir en grandes cantidades productos innovadores?.

Si miramos la tabla 3-14 veremos que Monterrey Mezclas Arábicas no solo posee una alta participación de mercado, si no que además la fidelidad al producto es alta y posee una fuerte presencia en el primer segmento de cambio.

En cuanto a los segmentos leales, Nestlé con su marca de familia Nescafé, es quien domina al mercado, con un 72%, seguida por Tres Montes quienes tienen un 16%, en este caso, las marcas privadas no tienen presencia,

lo que era de esperar, ya que son productos que son percibidos como de no muy buena calidad y precios bajos, por lo que son de bajo compromiso.

Si analizamos a los segmentos de cambio el primer segmento de cambio denominado como Segmento Cebada Tostada, podríamos decir que es un segmento mixto, ya que las marcas que lo lideran pertenecen a Nestlé y a Corp. Tres Montes, de hecho si tomamos el total de las marcas que tienen probabilidad mayor que cero en éste segmento, veremos que las marcas de Nestlé poseen una probabilidad de un 48% de ser elegidas por personas de ese segmento y que Corp. Tres Montes posee un 46%, por lo tanto aquí la probabilidad de elección está repartida equitativamente, por lo que la competencia entre los dos productos, Coronado y Ecco, es muy fuerte; un dato relevante es que ambos productos poseen precios muy similares, la diferencia de precio por gramo entre ambos es de 0.18 pesos (Anexo 4 -Tabla B3).

Si nos trasladamos al segundo segmento de cambio, denominado Segmento de Nescafé Premium, es fácil notar que éste segmento está dominado por las marcas de Nestlé, donde éstas aquí poseen una probabilidad de un 77% de ser elegidas por personas pertenecientes a ese segmento, la presencia de Corp. Tres Montes se limita al 15%.

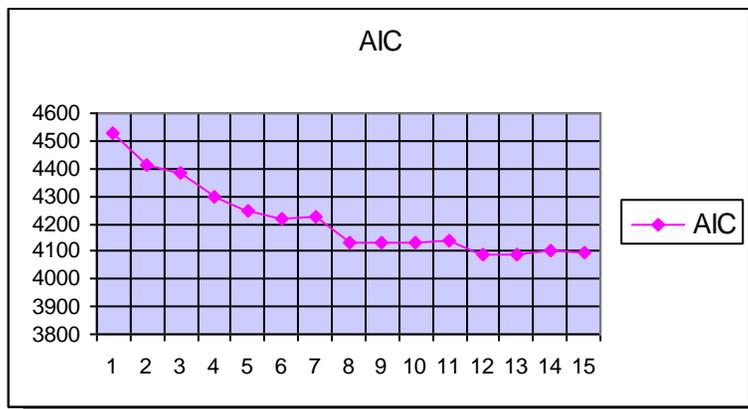
III. 3 Datos de detergentes

III. 3. 1. Determinando el número de segmentos de cambio

Para determinar el número óptimo de segmentos de cambio, debemos primero convertir los datos de la tabulación cruzada a variables dicotómicas, pero como lo que necesitamos es obtener el número óptimo de segmentos de cambio, tomamos solo las elecciones cambiantes, es decir, solo a aquellos panelistas que, en la segunda elección de compra, elegían una marca distinta que en la primera.

Una vez transformados los datos en dummies, obtuvimos con el programa Winmira 2001 el índice de ajuste AIC, el cual varía dependiendo del número de segmento de cambio que nosotros le exijamos al programa. La tabla 3-16 muestra los resultados obtenidos al hacer variar el número de segmentos de cambio, desde uno a quince.

Gráfico 3-4



En éste caso tenemos un hecho curioso, según se ve en el gráfico, podríamos tener tres puntos donde el AIC muestra que podrían ser el número óptimo de segmentos de cambio, la lógica nos indica queelijamos como número óptimo de segmentos el que de un mejor ajuste al modelo, que en éste caso sería elegir como número óptimo 12 segmentos de cambio, pero por la experiencia con los datos del café de igual forma analizaremos las otras alternativas, por lo menos en forma preliminar y así poder tomar la mejor decisión con la mayor cantidad de información posible.

Cabe destacar que cada producto ahora será representado por una variable, esto se hará de la siguiente forma

III. 3. 2. 1 Desarrollando segmentos de cambio primera opción

Para llevar acabo la denominada primera opción debemos pedir al programa Winmira 2001 que desarrolle los 12 segmentos, así el output requerido lo mostraremos en las siguientes tablas.

Variable	Comb	Descripción
Var 1	1	OMO MAT.PVO,
Var 2	2	OMO PVO.
Var 3	3	ACE POLVO

Var 4	58	OMO MATIC ACCION CONT.
Var 5	8	OMO MAT.FRE.PVO.
Var 6	4	DRIVE MAT.PRO.PVO
Var 7	9	DRIVE PRO.PVO.
Var 8	7	ARIEL PVO.
Var 9	68	RINSO MATIC PVO.
Var 10	11	ARIEL FUTUR LIQ.

Tabla 3-16.a.

Final estimates in CLASS 1 of 12 with size 0.11913
 =====

expected category frequencies and item scores:

Item label	Item's Score	Item's Stdev	relative category frequencies	
			0	1
VAR00001	0.00	0.07	0.996	0.004
VAR2	0.00	0.06	0.997	0.003
VAR3	0.65	0.48	0.348	0.652
VAR4	0.41	0.49	0.590	0.410
VAR5	0.34	0.47	0.659	0.341
VAR6	0.01	0.08	0.993	0.007
VAR7	0.01	0.08	0.994	0.006
VAR8	0.01	0.08	0.994	0.006
VAR9	0.43	0.49	0.572	0.428
VAR10	0.16	0.36	0.845	0.155
Sum:	2.01			

Tabla 3-16.b.

Final estimates in CLASS 2 of 12 with size 0.11104
 =====

expected category frequencies and item scores:

Item label	Item's		relative category frequencies	
	Score	Stdev	0	1
VAR00001	0.00	0.05	0.997	0.003
VAR2	0.00	0.05	0.997	0.003
VAR3	0.20	0.40	0.798	0.202
VAR4	0.17	0.37	0.833	0.167
VAR5	0.11	0.31	0.888	0.112
VAR6	0.22	0.41	0.784	0.216
VAR7	1.00	0.07	0.005	0.995
VAR8	0.28	0.45	0.725	0.275
VAR9	0.02	0.14	0.980	0.020
VAR10	0.02	0.14	0.980	0.020
Sum:	2.01			

Tabla 3-16.c.

Final estimates in CLASS 3 of 12 with size 0.10985

=====

expected category frequencies and item scores:

Item label	Item's		relative category frequencies	
	Score	Stdev	0	1
VAR00001	1.00	0.07	0.004	0.996
VAR2	0.00	0.05	0.997	0.003
VAR3	0.61	0.49	0.388	0.612
VAR4	0.37	0.48	0.630	0.370
VAR5	0.00	0.06	0.996	0.004
VAR6	0.00	0.06	0.996	0.004
VAR7	0.00	0.05	0.997	0.003
VAR8	0.01	0.07	0.994	0.006
VAR9	0.00	0.07	0.996	0.004
VAR10	0.01	0.11	0.988	0.012
Sum:	2.01			

Tabla 3-16.d.

Final estimates in CLASS 4 of 12 with size 0.10390

=====

expected category frequencies and item scores:

Item label	Item's Score	Item's Stdev	relative category frequencies	
------------	--------------	--------------	-------------------------------	--

			0	1
VAR00001	0.00	0.05	0.998	0.002
VAR2	1.00	0.05	0.002	0.998
VAR3	1.00	0.05	0.002	0.998
VAR4	0.00	0.05	0.998	0.002
VAR5	0.00	0.05	0.998	0.002
VAR6	0.00	0.05	0.998	0.002
VAR7	0.00	0.05	0.998	0.002
VAR8	0.00	0.05	0.998	0.002
VAR9	0.00	0.05	0.998	0.002
VAR10	0.00	0.05	0.998	0.002
Sum:	2.01			

Tabla 3-16.e.

Final estimates in CLASS 5 of 12 with size 0.09384

=====

expected category frequencies and item scores:

Item label	Item`s		relative category frequencies	
	Score	Stdev	0	1
VAR00001	0.00	0.05	0.998	0.002
VAR2	1.00	0.05	0.002	0.998
VAR3	0.00	0.05	0.998	0.002
VAR4	0.00	0.05	0.998	0.002
VAR5	0.00	0.05	0.998	0.002
VAR6	0.00	0.05	0.998	0.002
VAR7	1.00	0.05	0.002	0.998
VAR8	0.00	0.05	0.998	0.002
VAR9	0.00	0.05	0.998	0.002
VAR10	0.00	0.05	0.998	0.002
Sum:	2.01			

Tabla 3-16.f.

Final estimates in CLASS 6 of 12 with size 0.09066

=====

expected category frequencies and item scores:

Item label	Item`s		relative category frequencies	
	Score	Stdev		

			0	1
VAR00001	0.31	0.46	0.686	0.314
VAR2	0.11	0.31	0.893	0.107
VAR3	0.04	0.20	0.956	0.044
VAR4	0.14	0.34	0.864	0.136
VAR5	0.18	0.39	0.818	0.182
VAR6	1.00	0.06	0.004	0.996
VAR7	0.01	0.10	0.990	0.010
VAR8	0.09	0.28	0.914	0.086
VAR9	0.05	0.21	0.954	0.046
VAR10	0.09	0.29	0.908	0.092
Sum:	2.01			

Tabla 3-16.g.

Final estimates in CLASS 7 of 12 with size 0.07495

expected category frequencies and item scores:

Item label	Item`s		relative category frequencies	
	Score	Stdev	0	1
VAR00001	0.00	0.05	0.997	0.003
VAR2	0.99	0.08	0.007	0.993
VAR3	0.00	0.05	0.997	0.003
VAR4	0.00	0.05	0.997	0.003
VAR5	0.22	0.42	0.778	0.222
VAR6	0.01	0.11	0.989	0.011
VAR7	0.00	0.05	0.997	0.003
VAR8	0.41	0.49	0.587	0.413
VAR9	0.28	0.45	0.723	0.277
VAR10	0.08	0.28	0.916	0.084
Sum:	2.01			

Tabla 3-16.h.

Final estimates in CLASS 8 of 12 with size 0.07358

expected category frequencies and item scores:

Item label	Item`s		relative category frequencies	
	Score	Stdev		

			0	1
VAR00001	0.24	0.43	0.756	0.244
VAR2	0.00	0.07	0.995	0.005
VAR3	0.53	0.50	0.470	0.530
VAR4	0.06	0.23	0.944	0.056
VAR5	0.06	0.23	0.945	0.055
VAR6	0.01	0.09	0.991	0.009
VAR7	0.01	0.07	0.994	0.006
VAR8	1.00	0.06	0.004	0.996
VAR9	0.03	0.16	0.973	0.027
VAR10	0.09	0.28	0.915	0.085
Sum:	2.01			

Tabla 3-16.i.

Final estimates in CLASS 9 of 12 with size 0.06583

=====

expected category frequencies and item scores:

Item label	Item`s		relative category frequencies	
	Score	Stdev	0	1
VAR00001	1.00	0.05	0.002	0.998
VAR2	1.00	0.05	0.002	0.998
VAR3	0.00	0.05	0.998	0.002
VAR4	0.00	0.05	0.998	0.002
VAR5	0.00	0.05	0.998	0.002
VAR6	0.00	0.05	0.998	0.002
VAR7	0.00	0.05	0.998	0.002
VAR8	0.00	0.05	0.998	0.002
VAR9	0.00	0.05	0.998	0.002
VAR10	0.00	0.05	0.998	0.002
Sum:	2.01			

Tabla 3-16.j.

Final estimates in CLASS 10 of 12 with size 0.06349

=====

expected category frequencies and item scores:

Item	Item`s	relative category
------	--------	-------------------

label	Score	Stdev	frequencies	
			0	1
VAR00001	0.99	0.08	0.006	0.994
VAR2	0.00	0.05	0.997	0.003
VAR3	0.00	0.06	0.996	0.004
VAR4	0.01	0.08	0.994	0.006
VAR5	0.55	0.50	0.447	0.553
VAR6	0.00	0.07	0.996	0.004
VAR7	0.00	0.05	0.997	0.003
VAR8	0.01	0.08	0.994	0.006
VAR9	0.42	0.49	0.578	0.422
VAR10	0.02	0.13	0.983	0.017
Sum:	2.01			

Tabla 3-16.k.

Final estimates in CLASS 11 of 12 with size 0.05793

=====

expected category frequencies and item scores:

Item label	Item`s Score	Stdev	relative category frequencies	
			0	1
VAR00001	0.00	0.05	0.998	0.002
VAR2	1.00	0.05	0.002	0.998
VAR3	0.00	0.05	0.998	0.002
VAR4	1.00	0.05	0.002	0.998
VAR5	0.00	0.05	0.998	0.002
VAR6	0.00	0.05	0.998	0.002
VAR7	0.00	0.05	0.998	0.002
VAR8	0.00	0.05	0.998	0.002
VAR9	0.00	0.05	0.998	0.002
VAR10	0.00	0.05	0.998	0.002
Sum:	2.01			

Tabla 3-16.l.

Final estimates in CLASS 12 of 12 with size 0.03580

=====

expected category frequencies and item scores:

Item label	Item`s		relative category frequencies	
	Score	Stdev	0	1
VAR00001	1.00	0.05	0.002	0.998
VAR2	0.00	0.05	0.998	0.002
VAR3	0.00	0.05	0.998	0.002
VAR4	0.00	0.05	0.998	0.002
VAR5	0.00	0.05	0.998	0.002
VAR6	0.00	0.05	0.998	0.002
VAR7	1.00	0.05	0.002	0.998
VAR8	0.00	0.05	0.998	0.002
VAR9	0.00	0.05	0.998	0.002
VAR10	0.00	0.05	0.998	0.002

Sum: | 2.01

De acuerdo al output obtenido podemos decir que en el segmento 1 existen 5 marcas que tienen probabilidad menor que 0.01, por lo que esas marcas serán restringidas a cero.

El segmento dos presenta una marca (Drive Progress) con probabilidad de elección cercana a 1, ésta sería una de las llamadas elecciones casi leales, por lo que la elección de ésta marca debiera pasar a formar parte de los segmentos leales. Existen además dos marcas (Omo Matic polvo y Omo polvo) con probabilidades muy bajas, por lo que esas observaciones serán restringidas a cero.

En el segmento tres podemos observar que la Var 1 (Omo matic) tiene una probabilidad demasiado alta de elección, por lo que ésta marca pasará a ser parte de los segmentos leales, por otro lado tenemos 6 marcas que deben ser restringidas porque no presentan observaciones suficientes (si es que presentan observaciones en éste segmento), las marcas a ser restringidas por identificación son Var 2 (Omo pvo.) Var 5 (Omo Matic Fresh Pvo.), Var 6

(Drive Matic Progress pvo.), Var 7 (Drive Progress pvo.), Var 8 (Ariel Pvo.), Var 9 (Rinso Matic).

El segmento 4 es un caso especial, ya en todo el segmento se presentan solo dos marcas con observaciones, la Var 2 (Omo pvo.) y la Var. 3 (Ace pvo.), el problema es que esto hace que la probabilidad de elección en las dos marcas sea demasiado alta, por lo que éstas marcas pasan a formar parte de los segmentos leales.

Con el segmento 5 ocurre lo mismo que con el anterior, aquí las marcas que son casi leales son Var 2 (Omo pvo.), Var 7 (Drive Progress pvo.)

El segmento 6, presenta una marca casi leal, que es la Var 6 (Drive Matic Progress), y una marca con probabilidad muy baja que es la Var 7 (Drive Progress).

En el segmento 7, la Var 2 (Omo pvo.) es una marca que en éste caso se presenta como casi leal, y se restringen por identificación a la Var 1 (Omo Matic pvo.), Var 3 (Ace pvo.), Var 4 (Omo Matic Fresh) y la Var7 (Drive Progress).

La marca Var 8 (Ariel pvo.) presenta casi lealtad en el segmento 8, y se restringen en el mismo por identificación a la Var 2 (Omo pvo.), Var 6 (Drive Matic Progress Pvo) y a la Var 7 (Drive Progress).

Como elección casi leal se presenta la Var 1 (Omo Martiv pvo.) en el segmento 10, y como marcas a restringir se presentan Var 2 (Omo pvo.), Var

3 (Ace pvo.), Var 4 (Omo Matic Acción Cont.), Var 6 (Drive Matic Progress), Var 7 (Drive Progress) y la Var 8 (Ariel pvo.).

Finalmente los segmentos 9, 11, 12 quedan como segmentos con elecciones casi leales con las marcas Omo Matic y Omo pvo. en el segmento 9, Omo pvo y Omo Acción Cont. en el segmento 11 y por último Omo Matic pvo. y Drive Progress pvo en el segmento 12.

El siguiente paso fue separar las elecciones por segmento para así poder sacar las probabilidades resultantes después de los cambios antes mencionados, en el Anexo 6 mostramos con más detalles el procedimiento, se utilizó el mismo que con los datos del café.

En la tabla siguiente resumimos el resultado obtenido, ahí mostramos las probabilidades por segmento de cambio.

Tabla 3-17

Marcas	Segmento 1	Segmento 2	Segmento 3	Segmento 6	Segmento 7	Segmento 8	Segmento 10
Var 1	-	-	-	0,31	-	0,26	-
Var 2	-	-	-	0,11	-	-	-
Var 3	0,33	0,20	0,61	0,04	-	0,51	-
Var 4	0,21	0,16	0,37	0,13	-	0,06	-
Var 5	0,17	0,11	-	0,18	0,22	0,06	0,57
Var 6	-	0,22	-	-	-	-	-

Var 7	-	-	-	-	-	-	-
Var 8	-	0,27	-	0,09	0,42	-	-
Var 9	0,22	0,02	-	0,04	0,28	0,03	0,43
Var 10	0,08	0,02	0,02	0,09	0,08	0,09	-
	1,00	1,00	1,00	1,00	1,00	1,00	1,00

Como vemos finalmente resultaron 7 segmentos de cambio.

III. 3. 3. 1 Análisis de los resultados primera opción.

Luego a los segmentos leales les unimos la información y datos de los segmentos casi leales, sacamos la participación de mercado de todos los segmentos, es decir, cuanto pesa en la muestra cada segmento perteneciente a las elecciones de cambio, segmentos cambiantes, a las elecciones leales y casi leales, segmentos leales. En la tabla 3-18, mostramos los resultados obtenidos, además de las participaciones de mercado por marca y las probabilidades de elección que cada segmento tiene de cada marca.

En la misma tabla con color fucsia resaltamos a las marcas que tienen mayor probabilidad de ser elegidas por las personas pertenecientes a ese segmento.

Tabla 3-18

Marcas	Particip.	Lealtad (V)	Participación (W)						
			0,06	0,03	0,03	0,02	0,02	0,02	0,01
			Segmento 1	Segmento 2	Segmento 3	Segmento 6	Segmento 7	Segmento 8	Segmento 10
Omo Matic Pvo.	0,13	0,08	-	-	-	0,31	-	0,26	-
Omo Pvo.	0,27	0,25	-	-	-	0,11	-	-	-
Ace Pvo	0,17	0,12	0,33	0,20	0,61	0,04	-	0,51	-
Omo Matic Acc. C.	0,05	0,05	0,21	0,16	0,37	0,13	-	0,06	-
Omo Matic Fres.	0,06	0,02	0,17	0,11	-	0,18	0,22	0,06	0,57
Drive Matic Prog.	0,05	0,02	-	0,22	-	-	-	-	-
Drive Prog.	0,12	0,09	-	-	-	-	-	-	-
Ariel Pvo.	0,02	0,03	-	0,27	-	0,09	0,42	-	-
Rinso	0,08	0,02	0,22	0,02	-	0,04	0,28	0,03	0,43
Ariel Futur. Liq..	0,05	0,01	0,08	0,02	0,02	0,09	0,08	0,09	-
	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00

Vemos que en el segmento 1 las marcas que tienen mayor probabilidad de lección son Ace y Rinso Matic, en el segmento dos son Drive Matic progress y Ariel, en el segmento 3 son Ace y Omo Matic Acción Cont., en el segmento

6 son Omo Matic y Omo Matic Fresh, en el 7 son Omo Matic Fresh y Rinso Matic, el segmento 8 tiene a Omo Matic y Ace, finalmente el segmento 10 posee a Omo Matic Fresh y Rinso.

Es fácil darse cuenta que con ésta opción de ocupar como número óptimo de segmentos óptimos el que diera el mejor ajuste, sin analizar opciones alternativas, quizás no era lo mejor, ya que por lo menos en ésta ocasión el número de segmentos de cambio era muy alto, 12 al principio, y aunque disminuyeron bastante al sacar a los segmentos casi leales, aún así el número sigue siendo alto. Por ésta razón, además porque el interpretar a los segmentos, cuando el número de éstos es muy elevado, es muy difícil, vamos a estudiar opciones alternativas, en un principio habíamos dicho que existían tres puntos en el gráfico de los AIC, que indicaban que podían ser candidatos a ser el número óptimo de segmentos de cambio. A continuación analizaremos las otras opciones, con 6 y 8 segmentos de cambio.

III. 3. 2. 2 Desarrollando segunda opción

En las siguientes tablas vemos el output del programa Winmira 2001 con 8 segmentos. Cada tabla representa a un segmento, con ellas podremos ver igual que en los casos anteriores, cuales de ellos, más que un segmento de cambio son segmentos casi leales y además ver que marcas deberían ser restringidas por identificación.

Tabla 3-19. a

Final estimates in CLASS 1 of 8 with size 0.20331

=====

expected category frequencies and item scores:

Item label	Item`s		relative category frequencies	
	Score	Stdev	0	1
VAR00001	0.32	0.47	0.679	0.321
VAR2	1.00	0.07	0.005	0.995
VAR3	0.00	0.05	0.997	0.003
VAR4	0.28	0.45	0.718	0.282
VAR5	0.08	0.27	0.918	0.082
VAR6	0.04	0.19	0.961	0.039
VAR7	0.00	0.05	0.997	0.003
VAR8	0.15	0.36	0.848	0.152
VAR9	0.10	0.30	0.898	0.102
VAR10	0.03	0.18	0.968	0.032
Sum:	2.01			

Tabla 3-19. b

Final estimates in CLASS 2 of 8 with size 0.17891

=====

expected category frequencies and item scores:

Item label	Item`s		relative category frequencies	
	Score	Stdev	0	1
VAR00001	0.69	0.46	0.315	0.685
VAR2	0.01	0.07	0.995	0.005
VAR3	0.00	0.06	0.996	0.004
VAR4	0.40	0.49	0.603	0.397
VAR5	0.35	0.48	0.647	0.353
VAR6	0.01	0.07	0.995	0.005
VAR7	0.01	0.07	0.994	0.006
VAR8	0.19	0.39	0.813	0.187
VAR9	0.29	0.46	0.705	0.295
VAR10	0.08	0.26	0.925	0.075
Sum:	2.01			

Tabla 3-19. c

Final estimates in CLASS 3 of 8 with size 0.14241
 =====

expected category frequencies and item scores:

Item label	Item`s		relative category frequencies	
	Score	Stdev	0	1
VAR00001	0.00	0.06	0.997	0.003
VAR2	0.00	0.06	0.997	0.003
VAR3	0.98	0.13	0.016	0.984
VAR4	0.16	0.37	0.838	0.162
VAR5	0.12	0.32	0.882	0.118
VAR6	0.02	0.14	0.979	0.021
VAR7	0.15	0.36	0.848	0.152
VAR8	0.28	0.45	0.718	0.282
VAR9	0.19	0.39	0.808	0.192
VAR10	0.09	0.29	0.905	0.095
Sum:	2.01			

Tabla 3-19. d

Final estimates in CLASS 4 of 8 with size 0.12514
 =====

expected category frequencies and item scores:

Item label	Item`s		relative category frequencies	
	Score	Stdev	0	1
VAR00001	0.28	0.45	0.722	0.278
VAR2	0.00	0.06	0.997	0.003
VAR3	0.01	0.12	0.987	0.013
VAR4	0.15	0.36	0.852	0.148
VAR5	0.10	0.30	0.901	0.099
VAR6	0.19	0.39	0.807	0.193
VAR7	1.00	0.07	0.005	0.995
VAR8	0.25	0.43	0.753	0.247
VAR9	0.02	0.13	0.983	0.017
VAR10	0.02	0.13	0.982	0.018
Sum:	2.01			

Tabla 3-19. e

Final estimates in CLASS 5 of 8 with size 0.10453

=====

expected category frequencies and item scores:

Item label	Item`s		relative category frequencies	
	Score	Stdev	0	1
VAR00001	0.00	0.05	0.998	0.002
VAR2	1.00	0.05	0.002	0.998
VAR3	1.00	0.05	0.002	0.998
VAR4	0.00	0.05	0.998	0.002
VAR5	0.00	0.05	0.998	0.002
VAR6	0.00	0.05	0.998	0.002
VAR7	0.00	0.05	0.998	0.002
VAR8	0.00	0.05	0.998	0.002
VAR9	0.00	0.05	0.998	0.002
VAR10	0.00	0.05	0.998	0.002
Sum:	2.01			

Tabla 3-19. f

Final estimates in CLASS 6 of 8 with size 0.09447

=====

expected category frequencies and item scores:

Item label	Item`s Score	Stdev	relative category frequencies
------------	--------------	-------	-------------------------------

			0	1
VAR00001	0.00	0.05	0.998	0.002
VAR2	1.00	0.05	0.002	0.998
VAR3	0.00	0.05	0.998	0.002
VAR4	0.00	0.05	0.998	0.002
VAR5	0.00	0.05	0.998	0.002
VAR6	0.00	0.05	0.998	0.002
VAR7	1.00	0.05	0.002	0.998
VAR8	0.00	0.05	0.998	0.002
VAR9	0.00	0.05	0.998	0.002
VAR10	0.00	0.05	0.998	0.002
Sum:	2.01			

Tabla 3-19. g

Final estimates in CLASS 7 of 8 with size 0.08290

expected category frequencies and item scores:

Item label	Item`s		relative category frequencies	
	Score	Stdev	0	1
VAR00001	0.35	0.48	0.653	0.347
VAR2	0.04	0.18	0.965	0.035
VAR3	0.02	0.14	0.981	0.019
VAR4	0.15	0.36	0.850	0.150
VAR5	0.20	0.40	0.799	0.201
VAR6	0.99	0.07	0.005	0.995
VAR7	0.01	0.10	0.989	0.011
VAR8	0.10	0.30	0.899	0.101
VAR9	0.05	0.22	0.949	0.051
VAR10	0.10	0.30	0.898	0.102
Sum:	2.01			

Tabla 3-19. h

Final estimates in CLASS 8 of 8 with size 0.06833

expected category frequencies and item scores:

Item label	Item`s		relative category frequencies	
	Score	Stdev	0	1

VAR00001		1.00		0.05		0.002		0.998
VAR2		0.00		0.05		0.998		0.002
VAR3		1.00		0.05		0.002		0.998
VAR4		0.00		0.05		0.998		0.002
VAR5		0.00		0.05		0.998		0.002
VAR6		0.00		0.05		0.998		0.002
VAR7		0.00		0.05		0.998		0.002
VAR8		0.00		0.05		0.998		0.002
VAR9		0.00		0.05		0.998		0.002
VAR10		0.00		0.05		0.998		0.002

Sum: | 2.01

En las tablas, podemos ver que existen tres segmentos que debieran a pasar a formar parte de los segmentos leales, por ser marcas casi leales, son los segmentos 5, 6 y 8. En los segmentos 1, 3, 4 y 7, tenemos marcas que debieran pasar a formar parte de los segmentos leales, marcas que poseen probabilidad cercana a uno de ser elegidos por personas de ese segmentos, las marcas son Omo Pvo, Ace Pvo., Drive Progress Pvo., Drive Matic Progress Pvo., respectivamente. En todos los segmentos tenemos marcas cuya probabilidad de elecci3n debiera ser restringida a cero, por el bajo valor de 3stas.

Finalmente nos quedan solo 5 segmentos de cambio, para obtener el resultado deseado debimos realizar el mismo proceso que en los casos anteriores, el cu3l nos arroj3 la siguiente tabla.

Tabla 3-20

	Segmento 1	Segmento 2	Segmento 3	Segmento 4	Segmento 7
Omo Matic Pvo.	0,32	0,34	-	0,28	0,37
Omo Pvo.	-	-	-	-	-
Ace Pvo	-	-	-	-	-
Omo Matic Acc. C.	0,28	0,20	0,16	0,15	0,16
Omo Matic Fres.	0,08	0,18	0,11	0,10	0,21
Drive Matic Prog.	0,05	-	0,03	0,20	-
Drive Prog.	-	-	0,16	-	-
Ariel Pvo.	0,15	0,10	0,27	0,25	0,11
Rinso	0,10	0,15	0,19	0,02	0,05
Ariel Futur. Liq..	0,03	0,04	0,09	0,02	0,11

1,00	1,00	1,00	1,00	1,00
------	------	------	------	------

Las marcas que debieran tener mayor probabilidad de ser elegidas por personas pertenecientes a un segmento, vemos que en el segmento 1 son Omo Matic y Omo Matic Accion Cont., en el segmento 2 son Omo Matic y Omo Matic Acción Cont. y Omo Matic Fresh, en el segmento 3 son Omo Matic Acción Cont y Ariel en pvo., en el segmento 4 tenemos, Omo Matic, Ariel y finalmente el segmento 5 las marcas son Omo Matic y Omo Matic Fresh.

III. 3. 2. 3 Desarrollando tercera opción

La denominada tercera opción consta de 6 segmentos de cambio, los que se desarrollan de la forma mostrada en las siguientes tablas

Tabla 3-21.a

Final estimates in CLASS 1 of 6 with size 0.20764
 =====

expected category frequencies and item scores:

Item label	Item`s		relative category frequencies	
	Score	Stdev	0	1
VAR00001	0.33	0.47	0.674	0.326
VAR2	0.00	0.06	0.997	0.003
VAR3	0.99	0.08	0.006	0.994
VAR4	0.11	0.31	0.890	0.110
VAR5	0.08	0.27	0.919	0.081
VAR6	0.01	0.12	0.986	0.014
VAR7	0.10	0.30	0.897	0.103
VAR8	0.19	0.39	0.810	0.190
VAR9	0.13	0.34	0.870	0.130
VAR10	0.06	0.24	0.938	0.062
Sum:	2.01			

Tabla 3-21.b

Final estimates in CLASS 2 of 6 with size 0.19938
=====

expected category frequencies and item scores:

Item label	Item`s		relative category frequencies	
	Score	Stdev	0	1
VAR00001	0.33	0.47	0.671	0.329
VAR2	0.99	0.07	0.005	0.995
VAR3	0.00	0.05	0.997	0.003
VAR4	0.29	0.45	0.711	0.289
VAR5	0.08	0.28	0.916	0.084
VAR6	0.02	0.12	0.984	0.016
VAR7	0.00	0.07	0.995	0.005
VAR8	0.16	0.36	0.844	0.156
VAR9	0.10	0.31	0.896	0.104
VAR10	0.03	0.18	0.967	0.033
Sum:	2.01			

Tabla 3-21.c

Final estimates in CLASS 3 of 6 with size 0.19533
=====

expected category frequencies and item scores:

Item label	Item`s		relative category frequencies	
	Score	Stdev	0	1
VAR00001	0.18	0.38	0.823	0.177
VAR2	0.48	0.50	0.519	0.481
VAR3	0.01	0.10	0.989	0.011
VAR4	0.10	0.29	0.905	0.095
VAR5	0.06	0.24	0.936	0.064
VAR6	0.01	0.08	0.993	0.007
VAR7	1.00	0.07	0.004	0.996
VAR8	0.16	0.37	0.841	0.159
VAR9	0.01	0.11	0.988	0.012
VAR10	0.01	0.11	0.988	0.012
Sum:	2.01			

Tabla 3-21.d

Final estimates in CLASS 4 of 6 with size 0.18173

expected category frequencies and item scores:

Item label	Item`s		relative category frequencies	
	Score	Stdev	0	1
VAR00001	0.68	0.47	0.321	0.679
VAR2	0.01	0.07	0.995	0.005
VAR3	0.01	0.07	0.994	0.006
VAR4	0.39	0.49	0.605	0.395
VAR5	0.35	0.48	0.649	0.351
VAR6	0.01	0.08	0.994	0.006
VAR7	0.01	0.09	0.991	0.009
VAR8	0.19	0.39	0.810	0.190
VAR9	0.29	0.46	0.706	0.294
VAR10	0.08	0.27	0.921	0.079
Sum:	2.01			

Tabla 3-21.e

Final estimates in CLASS 5 of 6 with size 0.11102

expected category frequencies and item scores:

Item label	Item`s		relative category frequencies	
	Score	Stdev	0	1
VAR00001	0.26	0.44	0.741	0.259
VAR2	0.07	0.26	0.930	0.070
VAR3	0.02	0.13	0.983	0.017
VAR4	0.11	0.32	0.887	0.113
VAR5	0.15	0.36	0.850	0.150
VAR6	0.99	0.07	0.005	0.995
VAR7	0.22	0.41	0.783	0.217
VAR8	0.08	0.27	0.924	0.076
VAR9	0.04	0.19	0.961	0.039
VAR10	0.08	0.27	0.924	0.076
Sum:	2.01			

Tabla 3-21.f

Final estimates in CLASS 6 of 6 with size 0.10490
 =====

expected category frequencies and item scores:

Item label	Item`s		relative category frequencies	
	Score	Stdev	0	1
VAR00001	0.00	0.05	0.998	0.002
VAR2	1.00	0.05	0.002	0.998
VAR3	1.00	0.05	0.002	0.998
VAR4	0.00	0.05	0.998	0.002
VAR5	0.00	0.05	0.998	0.002
VAR6	0.00	0.05	0.998	0.002
VAR7	0.00	0.05	0.998	0.002
VAR8	0.00	0.05	0.998	0.002
VAR9	0.00	0.05	0.998	0.002
VAR10	0.00	0.05	0.998	0.002

Sum: | 2.01

Al igual que en los casos anteriores, éstas tablas nos muestran cuales segmentos y elecciones de marca pasan a ser casi leales y que marcas son restringidas a cero por identificación, en éste caso tenemos que el segmento 6 es casi leal, podemos ver además que los segmentos que poseen elecciones casi leales son, el segmento 1 con la marca Ace, el segmento dos con Omo en pvo., el segmento 3 con Drive Matic Pro., el segmento5 con Drive Matic. Pro. Al igual que en los casos anteriores, debemos restringir algunas elecciones a cero por tener muy baja probabilidad. En base a esto, los resultados los mostramos en la tabla siguiente..

Tabla 3-22

	Segmento 1	Segmento 2	Segmento 3	Segmento 4	Segmento 7
Omo Matic Pvo.	0,32	0,33	0,18	0,34	0,25
Omo Pvo.	-	-	0,48	-	0,09
Ace Pvo	-	-	-	-	-
Omo Matic Acc. C.	0,11	0,29	0,09	0,20	0,11
Omo Matic Fres.	0,08	0,08	0,06	0,18	0,15
Drive Matic Prog.	0,02	-	-	-	-
Drive Prog.	0,11	-	-	-	0,22
Ariel Pvo.	0,18	0,16	0,16	0,10	0,07
Rinso	0,13	0,10	0,01	0,15	0,04
Ariel Futur. Liq..	0,06	0,03	0,01	0,04	0,07
	1,00	1,00	1,00	1,00	1,00

En cuanto a las marcas que debieran tener mayor probabilidad de ser elegidas por personas pertenecientes a un segmento y que por lo tanto debieran definirlo, vemos que en el segmento 1 son Omo Matic y Ariel en Polvo, en el segmento 2 son Omo Matic y Omo Matic Acción Cont., en el segmento 3 son Omo Matic y Omo pvo., en el segmento 4 tenemos, Omo Matic , Omo Matic Accion Cont., Omo Matic Fresh. Y finalmente el segmento 5 las marcas son Omo Matic y Drive Progress.

III. 3. 4 Grados de libertad y condición de identificación

Los grados de libertad para los datos de detergentes se hará de la misma forma como se obtuvieron los grados de libertad para los datos del café.

Como sabemos una condición necesaria para la identificación es que los grados de libertad no sean negativos.

Grados de libertad primera opción

En éste caso en particular tenemos los siguientes valores:

- $n = 10$ (número de segmentos leales).
- $m = 7$ (número de segmentos de cambio).
- $t = 0$ (segmentos restringidos a cero por identificación)²⁰.

Finalmente, la ecuación arrojó el siguiente resultado.

$$d.f = 10(10-7-1)+0$$

$$d.f = 20$$

²⁰ A pesar de que en las tablas aparecen parámetros marcados con color rojo, que indica que fueron restringidos a cero por tener una probabilidad menor que 0.01, en realidad esas marcas no tenían compras en ese segmento, por lo que los parámetros no fueron restringidos a priori.

Al ser los grados de libertad positivos, podemos decir que se cumple con la condición necesaria y suficiente para la identificación del modelo.

Grados de libertad segunda opción

En éste caso en particular tenemos los siguientes valores:

- $n = 10$ (número de segmentos leales).
- $m = 5$ (número de segmentos de cambio).
- $t = 0$ (segmentos restringidos a cero por identificación)²¹.

Finalmente, la ecuación arrojó el siguiente resultado.

$$\mathbf{d.f = 10(10-5-1)+0}$$

$$\mathbf{d.f = 40}$$

Al ser los grados de libertad positivos, podemos decir que se cumple con la condición necesaria y suficiente para la identificación del modelo.

Grados de libertad tercera opción

En éste caso en particular tenemos los siguientes valores:

²¹ Ocurre lo mismo que con el “t” de la primera opción.

- $n= 10$ (número de segmentos leales).
- $m= 5$ (número de segmentos de cambio).
- $t= 0$ (segmentos restringidos a cero por identificación)²².

Finalmente, la ecuación arrojó el siguiente resultado.

$$\mathbf{d.f = 10(10-5-1)+0}$$

$$\mathbf{d.f = 40}$$

Al ser los grados de libertad positivos, podemos decir que se cumple con la condición necesaria y suficiente para la identificación del modelo.

III. 3. 5 Decisión de opción

Para poder elegir la mejor opción, tenemos que ver la solución que de el mejor ajuste para el modelo y además que la solución sea la más lógica y racional en cuanto a la interpretación de los segmentos.

En éste sentido resumiremos las interpretaciones antes expuestas por cada una de las alternativas, para así tomar una mejor decisión.

En éstas tablas mostramos el resumen de las marcas que describen a cada segmento por opción, es éste momento los instrumentos que tenemos

²² Ocurre lo mismo que con el “t” de la primera opción.

para nominar a los segmentos son solo variables descriptivas de los productos, características físicas, como precio, alguna característica especial que tendrían cada producto y marca.

Tabla 3-22a

Primera Opción	
Segmento 1	Ace Rinso Matic
Segmento 2	Drive Matic Ariel Pvo.
Segmento 3	Ace Omo Matic Acc.
Segmento 6	Omo Matic Omo Matic Fresh
Segmento 7	Rinso Matic Ariel Pvo.
Segmento 8	Omo Matic Ace

Segmento 10	Omo Matic Fresh Rinso Matic
--------------------	--------------------------------

*Si tratamos de analizar por segmento para ver como nominarlos, podemos ver en el segmento 1 que Ace y Rinso Matic, en cuanto a características claves que compartan, podríamos decir que Rinso Matic, es un detergente claramente que sirve para lavadoras automáticas y que Ace, aunque en teoría sirve también para lavadoras automáticas, no es tan claro, de hecho la palabra Matic no está en su nombre de marca y según las instrucciones de uso de éste detergentes dice que sirve para lavado a mano y para lavado en lavadoras con carga por arriba, por lo que no podríamos denominarlo como un segmento Matic, en cuanto a la similitud que podrían tener en precios (Anexo 7) éstos dos productos, podemos ver que Ace ocupa el séptimo lugar en el ranking de precios y que Rinso Matic ocupa el último lugar. De acuerdo a lo que aquí expuesto no queda muy claro como nominar a éste segmento, excepto que podemos decir que se trata de un **segmento relativamente económico**. En cuanto al segmento 2 Vemos a dos marcas Drive Matic y Ariel en Polvo, la primera obviamente es un producto para lavadoras automáticas, pero Ariel en polvo es un producto solo para lavar a mano y para lavadoras con carga superior, por lo que al igual que con Ace, no podríamos decir que se trata de un segmento matic, Ariel está en el lugar*

número dos del ranking de precios está en el 6 lugar, pero la diferencia con las marcas que ocupan los dos lugares anteriores, es mínima, es decir, está dentro de la denominada banda de precios medio (Anexo 7), por lo que se trata de un **segmento relativamente alto**. En el segmento 3 vemos una marca matic y otra no, y las dos marcas pertenecen a las marcas cuyo precio está en la banda media del ranking, por lo que lo denominaremos, **segmento medio**. El segmento 6 posee dos categorías de las marcas Omo Matic y en cuanto a precios, Omo Matic pertenece a la banda de precios medios y Omo Matic Fresh, corresponde a los precios bajos, por lo que el segmento será denominado, **segmento Omo Matic medio bajo**. El segmento 7, está compuesto por dos marcas, una es completamente matic y la otra no, Rinso Matic pertenece a la primera banda de precios del ranking, es decir a la de precios premium, la otra marca pertenece a la banda de precios bajos, por lo que vemos, en cuanto a las dos características estas dos marcas son diametralmente opuestas, por lo que se nos hace difícil el definir a este segmento, al menos con la información disponible. El segmento 8 está compuesta por una marca matic y la otra que no sirve para todo tipo de lavadoras, en cuanto a los precios ambas marcas pertenecen a la banda de precios medios, por lo que este segmento es un **segmento medio**. El segmento 10 posee dos marcas matic, y en cuanto a precios, las dos pertenecen a la

última banda, por lo que podríamos denominarlo **segmento económico matic**.

Tabla 3-22b

Segunda Opción	
Segmento 1	Omo Matic Omo Matic Accion
Segmento 2	Omo Matic Omo Matic Fresh Omo Matic Accion
Segmento 3	Omo Matic Acción Ariel Pvo.
Segmento 4	Omo Matic Ariel Pvo.
Segmento 7	Omo Matic Omo Matic Fresh

*Esta opción presenta a 5 segmentos, las dos marcas que deberían definir al segmento 1 son variaciones de la marca Omo Matic, y ambas marcas pertenecen a la banda de precios media, por lo que será definido éste segmento como **segmento Omo Matic medio**, el segmento dos posee tres marcas pertenecientes a la marca Omo Matic, además dos de ellas pertenecen a la banda de precios media y Omo Matic Fresh pertenece la banda de precios bajos, por lo que éste segmento será denominado **segmento Omo Matic medio bajo**. El segmento 3, posee una marca matic y otra que no sirve*

para todas las lavadoras, Ariel, pertenece a la banda de precios lata y Omo Matic Acción C., pertenece a la banda de precios media, según éstas características podríamos decir que es un **segmento medio alto**, el segmento 4 posee las mismas características del anterior, por lo que también es un **segmento medio alto**, el segmento 7 posee las mismas marcas identificándolo que el segmento 6 de la primera opción por lo que será denominado de igual forma, es decir, **segmento Omo Matic medio bajo**.

Tabla 3-22c

Tercera Opción	
Segmento 1	Omo Matic Ariel Pvo.
Segmento 2	Omo Matic Omo Matic Acción
Segmento 3	Omo Matic Omo
Segmento 4	Omo Matic Omo Matic Fresh Omo Matic Acción
Segmento 5	Omo Matic Drive Progress

Esta opción posee 5 segmentos , el primero posee las mismas marcas que identifican al segmento 4 de la segunda opción, al que habíamos denominado *segmento medio alto*, el segmento número es también identificado por marcas que identificaron a segmentos de la

segunda opción, es éste caso al segmento 1 el que había sido denominado como *segmento Omo Matic medio*, el segmento siguiente es también identificado principalmente por dos productos de Omo, en éste caso Omo en polvo y Omo Matic, esto es algo interesante, ya que éstas marcas más que posibles sustitutos, podrían ser productos complementarios, ya que sirven para cosas distintas, uno netamente para lavados a máquina (las personas no utilizan para lavados a mano, ya que san muy poca espuma y maltratan las manos), y el otro netamente para lavados a mano (no sirve para las máquinas, ya que dan mucha espuma y pueden llegar a dañar a una lavadora automática), por lo tanto podríamos decir que es un segmento mixto, en cuanto a precios, Omo Matic pertenece a la banda de precios medio y Omo a la banda de precios bajos, en resumen éste segmento será denominado como segmento *Omo mixto medio bajo*, el segmento siguiente es igual segmento 2, descrito en la opción dos, segmento *Omo Matic medio bajo*, el último segmento es un segmento mixto ya que una marca es para lavado a mano y la otra es para lavado a máquina, una de ellas pertenece a la banda de precios media y la otra a la banda de precios altos, por lo que éste segmento sería *segmento mixto medio alto*.

Vemos que con la información existente, la solución óptima estaría entre la segunda o la tercera opción , ya que la primera opción tiene un segmento no definido, en la segunda opción tenemos que se repite una definición en dos segmentos, ya que no hay forma de diferenciarlos, según las características objetivas que poseen. En canto al ajuste del modelo, sabemos que la primera opción tiene el mejor ajuste seguido por la segunda opción y finalmente la tercera, pero la diferencia en cuanto a ajuste no es tan alta, por lo que el criterio que se debe aplicar en éste caso para saber cuál es la mejor opción es la interpretación de segmentos, y en éste caso la mejor opción sería la opción 3, con segmentos óptimos (los que se reducen a 5).

III. 3. 6 El mercado en base a los resultados

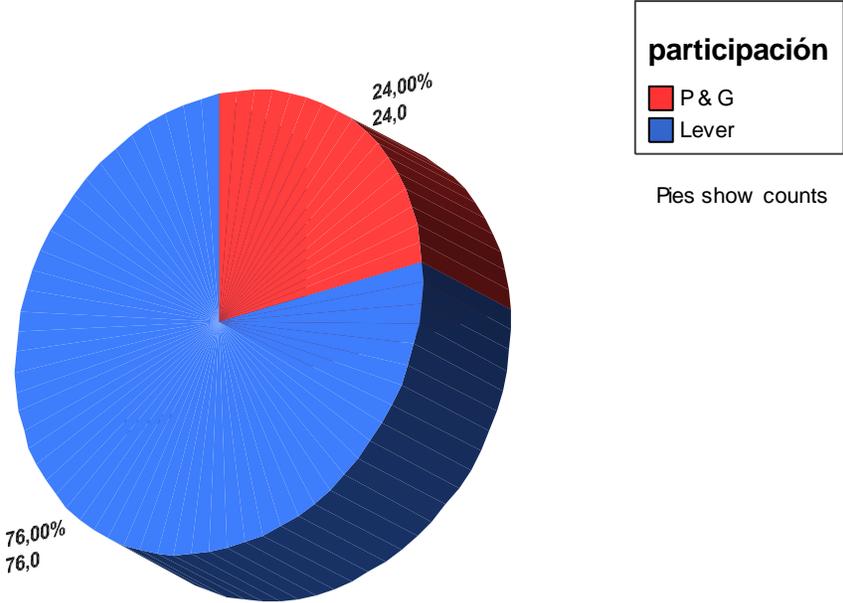
No podemos hablar del mercado en base a los resultados sin antes describir completamente a los productos en estudio, un elemento importante y que no ha sido mencionado son las marcas corporativas que hay detrás de éstos productos, en éste sentido, en la siguiente tabla mostraremos cuáles son y haremos un pequeño comentario con respecto a ello.

Tabla 3-23

Marcas	Marca Corporativa
Ariel Futur. Liq..	P & G
Ariel Pvo.	P & G
Drive Prog.	Lever
Omo Matic Acc. C.	Lever
Omo Matic Pvo.	Lever
Drive Matic Prog.	Lever
Ace Pvo	P & G
Omo Pvo.	Lever
Omo Matic Fres.	Lever
Rinso Matic	Lever

Podemos ver que la mayor parte del mercado la tiene Lever, de hecho P&G con sus productos Ariel y Ace, son relativamente nuevos en el mercado, para ver en detalle la participación de mercado de cada una de éstas marcas tenemos al gráfico siguiente.

Gráfico 3-5



Capítulo 4

Análisis de Percepción.

IV. 1 Percepción

Como sabemos la conducta de consumo se basa tanto en variables objetivas, como el precio, la marca, experiencias anteriores, etc. y en variables subjetivas como la percepción, como en éste estudio estamos analizando datos de compras efectivas, las que fueron influenciadas por las variables antes mencionadas, creemos que es importante tener alguna idea de cómo perciben los consumidores a los productos que estamos analizando.

IV. 2 Análisis Factorial

Nuestro objetivo es determinar los factores relevantes que buscan los consumidores en la compra de detergente y café, además determinar el posicionamiento de las marcas en estudio, de los mismos productos. Se asume que existe una estructura “abstracta” subyacente de factores. En base a esto realizamos una encuesta para saber cuáles son los factores que los consumidores consideran relevantes al momento de elegir una marca (la encuesta se encuentra en el anexo 8). Realizamos la encuesta a 15 personas, la cual tuvo los siguientes resultados.

Tabla 4-1

Cualidades del Detergente	Datos
Efectivo (elimine las manchas)	11
Precio	10
Para lavadora automática	6

En ésta tabla vemos los resultados de la primera encuesta, como era la etapa exploratoria del estudio y las preguntas eran abiertas, las

Marca	6
Aromatico	5
Suavisante	4
Que no destiña los colores	4
Rendidor	3
Liquido	2
Productos de regalo	2
Blancura	2
Que después sea fácil planchar	2
Gramos de regalo	1
Que no dañe la lavadora	1
Que no dañe las manos	1
Recomendaciones	1

estandarizamos para así establecer hacer algún tipo de frecuencia y así poder ver cuales son las características que irán en la segunda encuesta.

Tabla 4-2

Cualidades del Café	Datos
Marca	12
Sabor	12
Aroma	10
Costumbre	5
Cafeinado	4
Precio	3
Gramos	2
Origen del café	1
Envase	1
Cuerpo	1
Promociones	1

Aquí tenemos las frecuencias como resultado de la estandarizar las respuestas que obtuvimos.

Elegimos de éstas dos tablas las observaciones que tenían las mayores frecuencias y con ellas realizamos la encuesta (anexo 8) con la que luego haremos el análisis factorial.

Finalmente, para los datos de detergentes dejamos a 10 características y para los datos de café dejamos a 7. Las variables están medidas en base a una escala de intervalo, en donde le pedimos a los consumidores que evaluara los atributos para cada marca en una escala de 1 a 5.

La encuesta fue realizada para un total de 50 personas, ya que el tamaño de la muestra debe al menos exceder cinco veces el número de características a analizar, el muestreo fue hecho por conveniencia.

Resultados de la encuesta para datos de detergentes

Tabla 4-2

KMO and Bartlett's Test			
Kaiser-Meyer-Olkin Measure of Sampling Adequacy.			,862
Bartlett's Test of Sphericity	Approx. Chi-Square	1270,424	
	df	45	
	Sig.	,000	

a Based on correlations

En este recuadro tenemos dos test, donde verificamos la aplicabilidad del análisis factorial. El test de Bartlett 's es mayor a 1, y el de KMO es mayor a 0.5, condiciones necesarias y suficientes para poder aplicar el análisis.

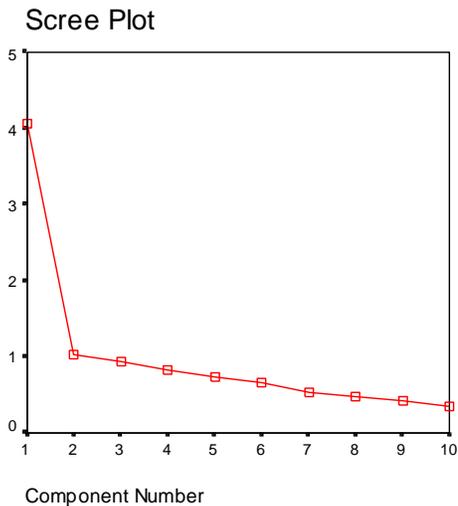
Luego de obtener la matriz de correlación y comprobar la alta correlación entre las variables, obtuvimos la siguiente tabla, en donde se indica la varianza explicada por cada componente, y además se define la cantidad de componentes importantes según el criterio de eigenvalue. Podemos ver que según éste análisis son dos los componentes que poseen atributos correlacionados entre ellos (recordemos que el criterio para elegir al número de factores es que los eigenvalue sean mayores que 1).

Tabla 4-2

Total Variance Explained				Extraction			Rotation		
Component	Initial Eigenvalues			Sums of Squared Loadings			Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	4,056	40,558	40,558	4,056	40,558	40,558	4,037	40,368	40,368
2	1,029	10,290	50,848	1,029	10,290	50,848	1,048	10,480	50,848
3	,931	9,305	60,153						
4	,819	8,189	68,343						
5	,727	7,267	75,610						
6	,657	6,565	82,175						
7	,528	5,276	87,451						
8	,483	4,827	92,279						
9	,420	4,204	96,483						
10	,352	3,517	100,000						

Extraction Method: Principal Component Analysis.

Para corroborar los resultados obtenidos, utilizaremos el gráfico Scree Plot



Podemos ver claramente en este gráfico que son dos factores los relevantes, ya que la inclusión de más factores no aumenta mayormente la varianza acumulada.

Gráfico 4-1

Utilizamos el planeamiento de análisis de componentes principales, es decir tomamos en cuenta la varianza total en los datos, para obtener la tabla que viene a continuación, donde aparecen los componentes y las cargas factoriales para cada uno

Tabla 4-3

Rotated Component Matrix
Component

	1	2
EFFECTIVO	,692	,190
PRECIO	,346	,612
AUTOMÁTI	,615	-,104
MARCA	,723	-3,114E-02
AROMÁTIC	,678	,336
SUAVIZAN	,637	,140
COLORES	,647	-5,068E-02
RENDIDOR	,752	9,569E-02
REGALOS	,417	-,685
BLANCURA	,724	-,116

Extraction Method: Principal Component Analysis. Rotation Method: Varimax with Kaiser Normalization.

a Rotation converged in 3 iterations.

El criterio para ver que factores son los relevantes, es tomar solo los atributos que tengan un valor absoluto mayor que 0.5, donde describen al componente positiva o negativamente, por simplicidad, mencionaremos solo los primero cinco factores con mayor valor.

Realizamos la rotación de factores con la metodología Varimax.

Interpretación de los factores

Para cada factor encontramos los siguientes atributos relevantes.

Factor 1: Rendidor, que deje la ropa blanca, marca, que saque bien las marcas (efectivo), que sea aromático.

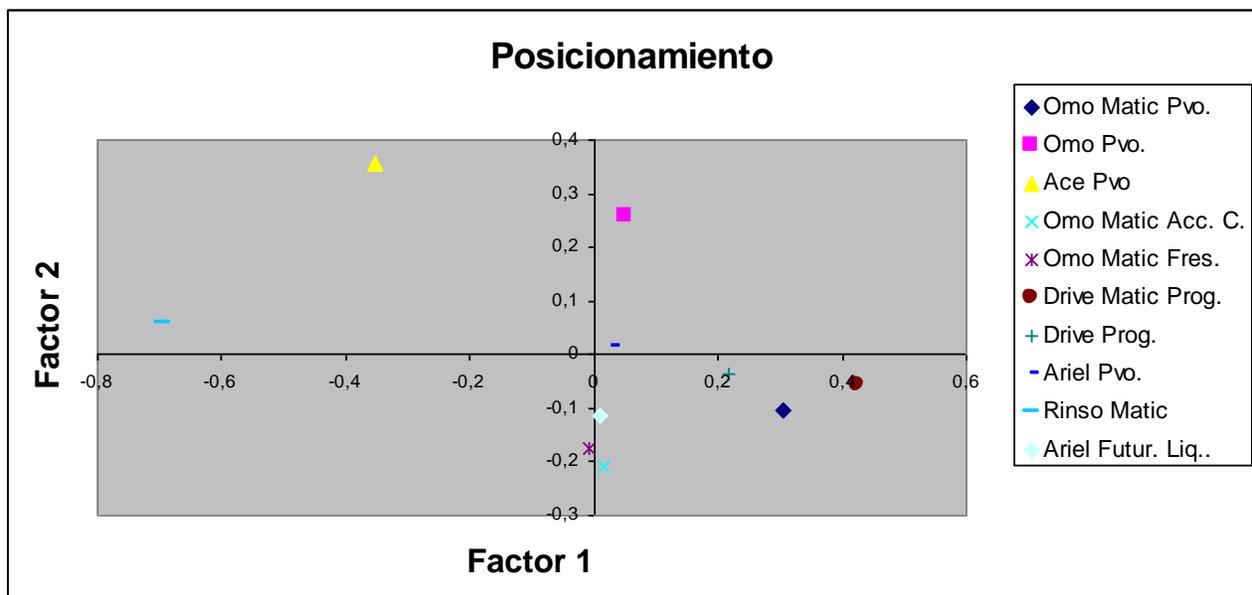
Factor 2: Precio (que sea barato) y negativamente que tenga regalos.

Tabla 4-4

	Factor 1	Factor 2
Omo Matic Pvo.	0,30441587	-0,10330674
Omo Pvo.	0,05018457	0,25926326
Ace Pvo	-0,35246891	0,35572587
Omo Matic Acc. C.	0,01580733	-0,21006467
Omo Matic Fres.	-0,00713767	-0,17411209
Drive Matic Prog.	0,42522727	-0,05581386
Drive Prog.	0,21800682	-0,03540159
Ariel Pvo.	0,03004159	0,01408864
Rinso	-0,69389889	0,05952956
Ariel Futur. Liq..	0,01030136	-0,1129424

Aquí vemos el cuadro que resume las cargas factoriales para cada marca, éste cuadro lo obtuvimos sumando las cargas de cada atributo relacionado a cada marca, por factor.

Con esto obtuvimos el siguiente mapa de posicionamiento.



Vemos que Omo Matic es percibido como una marca de buena calidad y no muy económico, Omo en polvo en cambio es percibida como una marca económica y de una calidad media, Ace en polvo es percibido como la marca más económica del mercado también como una marcas de baja calidad Omo Matic Acción C. posee una percepción de calidad ambigua, esto puede deberse por desconocimiento de la marca en cuanto a su calidad, pero si creen que no es un producto barato(esto puede ser influenciado por la marca Omo Matic), en cuanto a Omo Matic Fresh existe un desconocimiento también de su calidad y al igual que con la marca anterior, no es percibido como una marca económica, Drive Matic Progress es la marca que es percibida como de buena calidad, no es vista como una marca económica, Drive Progress para los consumidores es una marca de buena calidad y de precio medio, con Ariel en polvo ocurre algo especial, la marca es percibida como de precio medio y calidad media, esto puede ser porque esa es la percepción real o porque existe absoluto desconocimiento del producto, Rinso Matic es percibido como un producto de pésima calidad y relativamente económico, finalmente Ariel Futur es percibido como u producto de calidad media o desconocido para los consumidores y de precios altos.

Podemos distinguir en el mapa de posicionamiento algunas marcas agrupadas con características similares, tenemos a Drive Matic Progress junto con Omo Matic y Drive Progress en polvo, como productos de buena calidad y precios medios, aparentemente compiten por el mismo segmento, Omo Matic Acción junto con Omo Matic Fresh y Ariel Futur forman un grupo de productos percibidos como de calidad media y desconocida, Ariel en polvo, Ace, Rinso y Omo están cada una en un lugar distinto y aislado del mapa de posicionamiento.

Análisis de segmentos de cambio con resultados del estudio de percepción

En el capítulo anterior teníamos tres alternativas de número óptimo de segmentos de cambio, tomamos la decisión en ese momento, pero solo con la información existente,

volveremos a cuestionarnos sobre el número óptimo de segmentos de cambio, pero ésta vez, tomando en cuenta la información adicional sobre la percepción de los consumidores.

La primera opción había sido desechada, ya que existía un segmento que no pudo ser identificado con la información disponible, se trata del segmento 7, compuesto por Rinso y Ariel Matic, tomando en cuenta la información que nos entrega el mapa de posicionamiento, podríamos decir que ambos productos pertenecen, según la percepción colectiva, a un rango de precios medios, pero ésta apreciación no es concluyente, ya que también habíamos dicho que probablemente la posición de la marca Ariel en polvo sea producto del desconocimiento del mismo.

En cuanto a la segunda opción, había sido desechada porque existían dos segmentos que habían sido nominados de la misma forma, ya que como solo se disponía de información física y objetiva de los productos, ésta coincidía en los productos que describían a ambos segmentos. Los segmentos en cuestión eran el segmento 3 y el segmento 4, ambos nominados, **segmento medio alto**. Si vemos el mapa de posicionamiento, nos daremos cuenta que en el segmento 3, las dos marcas tienen en común que son percibidas como de calidad media o desconocida y en el segmento 4, la similitud radica en que los dos productos que lo caracterizarían son percibidos como de precios medios, por lo que podríamos renombrar al segmento 3 como **segmento de calidad media** y al segmento 4 como **segmento de precios medio altos**.

Ahora la situación ha cambiado, pudimos nominar a todos los segmentos de la opción dos y de forma distinta, al igual que los del segmento 6, entonces tenemos ahora que elegir entre éstas dos soluciones, donde las dos poseen el mismo número de segmentos de cambio y todos están identificados y correctamente nominados, un dato importante es que el ajuste que el número de segmentos de cambio da al modelo cambia de una opción a otra, donde es la segunda opción la que entrega el mejor ajuste. En conclusión y considerando la nueva información, el número óptimo de segmentos de cambio debiera ser 8, es decir, la opción dos.

Capítulo 5

Conclusiones, limitaciones y estudios futuros.

En éste seminario tratamos de mostrar la relevancia y utilidad de los datos de panel scanner, los avances que se han hecho en éste sentido en cuanto a estudios y como se pueden utilizar para segmentar utilizando solo probabilidades de elección de compra.

A pesar de que el trabajar con datos de panel scanner tiene muchas ventajas, también existen algunos hechos que son una desventaja, principalmente estos están dados por la voluntad de las personas que tienen que pasar los productos por el escáner, no existe forma de verificar de que la persona que pasó el producto por el escáner y que aparece como panelista fue efectivamente la persona que realizó la compra o si lo hizo en el día preciso que se hizo la compra, éste hecho lo pudimos constatar en nuestro estudio, ya que en un primer momento, al comenzar a trabajar con los datos, nos encontramos como por ejemplo con hogares donde en un mismo día, supuestamente se realizaron dos compras de café, como un supuesto del modelo es que la elección de las compras que entrarían en la tabulación cruzada debían ser hechas en ocasiones de compra distintas, no podíamos ocupar los datos de dos compras hechas en un mismo día.

Si de desventaja se trata, el modelo utilizado también tiene las suyas, recordemos que para modelar utilizamos solo las dos primeras compras de un año, con ellas hicimos una tabulación cruzada, es realmente difícil que solo dos compras puedan reflejar a un

mercado, creemos que sería pertinente hacer más de una tabulación cruzada, y tener un horizonte de tiempo más largo, hay modelos en donde se trata de ver el efecto que tiene las variables de marketing que utilizan más tabulaciones y un horizonte de tiempo más largo.

El tiempo también fue una limitante en éste estudio, el tiempo y la disponibilidad de información, se podría ver como afectan variables como las promociones y publicidad en el lugar de compra, además de relacionar la publicidad en la televisión cómo influencia a la conducta del consumidor.

A pesar de las limitaciones, se hizo un gran avance en el modelo desarrollado por Grover y Srinivasan, logramos obtener los segmentos de cambio, participaciones de mercado por marca y participación por segmento, logramos nominar a los segmentos de cambio según las marcas que tienen mayor probabilidad de elección por parte de la personas pertenecientes al segmento, además pudimos posicionar a todos los panelistas en sus segmentos correspondiente y hacer un análisis conjunto de conducta de compra con percepción. Además en uno de los productos realizamos un análisis factorial que mostró cuales son las características que los consumidores consideran como más importantes al momento de elegir un producto, de ésta forma enriquecimos con éstos datos el estudio de segmentación. Recordemos que los datos de panel scanner son hechos en base a la conducta de compra de los consumidores y que las decisiones de compra por parte de los consumidores se hace en base a variables objetivas como el precio y características del producto y principalmente a variables subjetivas como la percepción, por lo que sería importante hacer una relación de percepción con los segmentos obtenidos.

Finalmente, en cuanto a los estudios futuros, podríamos partir diciendo que en base al estudio aquí desarrollado se podría hacer una relación en base a información sociográfica de los panelistas con los resultados obtenidos de los segmentos, también se podrían tratar de hacer una regresión de los segmentos con datos del producto como promociones, precio, etc.

BIBLIOGRAFÍA

- Allenby, Greg. (1990), “Hypothesis Testing with Scanner Data: The Advantage of Bayesian Methods.
- Balachander and Ghose.(2003), “ Reciprocal Spillover Effects: A Strategic Benefit of Brand Extensions” , Journal of Marketing. Vol 67 Enero.
- Delso, Bárbara. Galindo, Carolina, Soto, Regina. (2003), “Esquemas de segmentación y valoración de marca con datos de panel scanner”, Seminario para optar al Título de ingeniero Comercial. Profesor Guía : Sergio Olavarrieta Soto.
- Green, Paul E., Carmone, Frank J. y Wachspres, David P. (1976) , “Consumer Segmentation Via Latent Class Analysis”, Journal of Consumer Research , Vol 3 Diciembre.
- Grover, Rajiv y Srinivasan. (1987), “A simultaneous Approach to Market Segmentation and Market Structuring”, Journal of Marketing Research, Vol 24 Mayo.
- Grover, Rajiv y Srinivasan. (1992), “Reflections on “A simultaneous Approach to Market Segmentation and Market Structuring”, Journal of Marketing Research. Vol. 29 Noviembre.
- Grover, Rajiv. (1987), “Estimation and Use of Standard Errors of Latent Class Model Parameters”, Journal of Marketing Research. Vol. 24 Agosto.

- Kamakura and Russell. (1989), “A Probabilistic Choice Model for Market Segmentation and Elasticity Structure” , Journal of Marketing Research. Vol. 26
Noviembre.
- Kamakura and Russell. (1994), “Understanding Brand Competition Using Micro and Macro Scanner Data” , Journal of Marketing Research. Vol. 31 Mayo
- Perez Cesar (2001), “ Técnicas Estadísticas con SPSS”

ANEXOS

ANEXO 1

Tabla A1

Panelista	comb 1	Comb 2
3	1	1
5	6	6
6	4	4
8	6	6
9	1	1
10	1	1
12	1	9
13	1	5
14	1	4
17	1	1
18	1	1
22	1	1
24	1	1
26	4	1
27	4	4
32	1	1
39	1	1
41	3	1
43	7	7
44	1	1
50	1	1
54	4	1
59	1	1
60	5	7
61	1	1
66	1	3
68	7	7
73	11	3
74	1	6
76	9	1
83	6	4
84	6	6
89	1	6
92	4	11
96	6	6
98	1	1
99	1	1
102	5	4
107	1	1
113	1	1
118	7	6

120	7	1
123	1	6
124	7	7
136	1	8
141	3	3
144	1	1
148	4	1
150	1	1
160	1	1
163	1	1
165	6	6
169	3	3
171	1	1
175	3	3
181	1	1
183	1	1
185	3	1
188	6	7
193	1	1
204	1	7
207	4	4
212	1	1
215	1	1
218	3	4
220	1	3
223	1	1
226	1	1
228	1	3
229	1	1
231	1	4
233	1	1
236	1	1
238	1	1
242	3	3
245	1	1
246	1	1
249	6	1
252	1	8
255	1	1
256	1	1
259	1	1
262	1	1
265	1	1
267	6	1
269	6	6
277	3	3
278	1	10
282	1	1
283	5	7

284	1	1
288	1	1
290	1	1
292	4	1
293	4	6
294	4	4
295	3	3
296	1	4
299	1	1
301	3	3
302	6	6
306	1	1
308	7	3
310	6	3
318	1	6
319	6	1
322	1	1
326	1	1
328	1	1
329	3	3
330	1	3
331	1	1
333	1	1
340	3	6
343	1	1
345	7	7
347	1	1
349	1	1
350	4	4
355	6	6
357	4	4
358	7	1
366	4	4
371	4	11
374	1	7
376	1	1
377	1	1
380	8	8
382	1	1
387	1	1
391	7	7
394	7	1
396	4	4
400	5	1
402	4	1
403	1	1
409	1	1
414	1	1
422	1	1

423	1	1
424	1	7
426	1	1
441	1	1
443	1	1
444	1	7
445	1	1
447	7	6
451	1	1
452	3	7
458	1	1
461	6	7
462	1	1
465	1	5
466	3	3
473	6	6
474	1	1
480	4	4
493	8	6
496	6	3
501	3	1
509	3	5
510	4	4
513	1	4
514	1	1
515	5	1
519	1	1
521	4	4
523	1	1
524	1	1
525	1	5
526	1	1
530	1	1
534	1	1
535	1	1
536	1	1
538	1	1
539	6	6
540	4	1
541	3	1
546	3	3
547	1	1
549	7	7
553	1	1
561	4	1
563	1	1
567	4	4
570	4	4
572	7	7

582	1	1
583	7	7
585	7	7
589	1	1
590	1	1
593	1	1
601	4	4
602	7	1
603	4	4
612	8	3
615	1	1
619	1	1
620	3	3
622	1	1
627	1	7
628	1	1
630	1	1
631	1	4
632	1	1
635	7	7
637	1	1
639	1	1
643	1	1
646	1	1
654	3	1
657	3	3
661	1	1
663	10	1
665	1	1
668	4	1
669	7	7
670	1	1
673	7	7
683	1	1
684	1	1
688	1	7
693	1	1
694	1	1
701	1	1
703	1	1
705	4	7
708	5	1
709	1	1
711	6	6
713	1	1
714	1	1
716	4	1
717	6	9
718	6	1

720	1	1
727	5	5
728	6	1
730	1	4
731	5	1
732	4	4
739	3	1
741	1	1
744	3	3
756	1	1
758	7	7
759	1	1
764	6	1
765	1	6
767	4	4
771	1	6
773	4	4
776	11	11
777	1	1
778	1	1
779	1	1
781	1	1
786	1	1
787	1	1
789	1	1
791	5	6
792	1	1
793	1	1
794	3	1
795	1	1
797	5	1
799	6	6
802	1	1
809	7	6
810	1	1
815	1	1
820	7	1
821	3	3
831	7	1
833	1	1
841	1	1
850	3	1
851	1	1
859	6	7
861	1	1
864	7	7
865	1	1
866	3	3
870	1	1

871	1	1
872	1	1
875	1	1
876	1	1
879	5	5
889	1	1
894	5	1
896	1	1
898	1	1
899	1	1
905	1	3
912	7	7
913	3	6
914	8	8
918	1	1
921	1	6
922	6	1
924	1	1
925	1	1
930	1	1
931	1	1
932	3	1
933	6	1
942	1	1
949	1	1
953	7	1
955	4	1
957	4	3
960	6	1
967	1	1
968	6	1
973	1	1
974	3	3
983	1	1
984	1	1
1004	10	10
1012	1	4
1013	1	1
1021	1	6
1029	1	1
1033	1	1
1034	5	5
1035	1	1
1039	1	1
1043	1	1
1046	1	1
1048	1	1
1050	1	1
1051	1	6

1052	4	6
1054	9	9
1055	11	7
1059	5	1
1061	4	1
1071	1	1
1074	5	1
1079	7	7
1094	6	8
1095	4	1
1097	11	11
1101	5	5
1105	3	3
1114	1	1
1115	4	1
1116	1	1
1122	6	6
1125	3	3
1126	1	1
1127	1	1
1135	1	1
1140	7	7
1141	4	1
1142	1	1
1144	1	1
1145	7	7
1146	1	1
1149	1	1
1152	6	5
1153	6	6
1160	5	1
1167	3	1
1169	5	5
1170	6	1
1171	1	1
1173	10	3
1174	1	1
1175	1	1
1178	3	3
1179	6	3
1180	1	1
1181	1	1
1182	3	3
1185	1	1
1191	3	3
1197	3	3
1202	1	6
1208	1	1
1217	1	1

1227	1	1
1230	1	1
1232	5	1
1233	6	1
1235	1	1
1238	6	6
1249	7	1
1250	1	1
1257	1	1
1259	10	10
1260	4	4
1263	1	1
1265	3	1
1273	3	1
1276	1	1
1281	1	8
1282	1	1
1283	1	1
1292	3	1
1299	5	5
1301	1	6
1303	1	1
1305	1	1
1306	5	5
1307	6	6
1314	11	3
1315	1	5
1316	1	1
1317	1	1
1320	7	1
1321	1	1
1324	1	1
1325	1	3
1330	10	10
1333	1	1
1343	8	8
1348	3	7
1349	7	7
1351	3	3
1356	1	1
1357	1	1
1358	5	1
1359	7	10
1360	3	3
1361	1	1
1362	1	1
1365	1	1
1367	4	1
1369	1	3

1373	11	1
1378	1	1
1380	1	1
1381	8	10
1383	1	1
1390	3	1
1391	1	1
1397	1	4
1400	1	1
1405	1	1
1406	7	7
1407	1	1
1413	1	1
1414	1	1
1417	7	7
1419	5	1
1422	4	4
1425	4	4
1428	1	6
1430	1	1
1434	7	3
1435	7	7
1436	6	6
1443	1	1
1444	3	3
1446	1	1
1448	6	1
1452	3	3
1459	1	6
1460	1	1
1464	3	1
1465	5	4
1473	3	3
1477	1	1
1478	1	1
1491	1	4
1492	4	4
1494	1	4
1495	1	1
1496	1	1
1500	1	4
1501	1	7
1503	8	1
1504	1	1
1508	4	7
1510	1	1
1514	6	1
1515	1	6
1517	4	4

1518	6	6
1520	7	7
1530	1	6
1559	1	1
1560	1	1
1561	1	1
1564	1	1
1571	1	1
1574	1	1
1577	3	4
1578	1	1
1582	3	1
1584	1	1
1587	6	6
1588	6	6
1589	7	7
1593	5	1
1597	1	1
1601	3	3
1611	1	1
1612	3	3
1616	9	1
1618	7	7
1619	1	3
1624	5	1
1631	7	3
1632	7	7
1641	8	8
1642	1	1
1645	1	7
1652	3	1
1654	1	1
1661	7	8
1662	1	4
1674	1	1
1678	4	1
1681	1	1
11001	3	3
11002	1	1
11003	1	3
11005	1	4
11006	3	3
11007	1	1
11008	1	1
11010	9	9
11011	1	1
11015	1	1
11017	7	6
11021	7	6

11026	3	3
11028	1	1
11029	1	1
11032	1	1
11034	4	1
11037	3	1
11040	1	1
11041	1	1
11046	1	1
11048	1	1
11049	3	7
11051	1	1
11054	9	3
11057	1	1
11058	1	1
11059	6	6
11061	1	7
11066	11	3
11069	1	5
11074	6	7
11076	1	1
11077	1	1
11078	6	6
11081	1	1
11082	1	1
11083	1	1
11084	1	1
11086	1	1
11089	1	1
11092	4	11
11095	5	5
11097	6	6
11098	3	1
11101	7	5
11102	6	1
11103	1	1
11106	1	1
11108	7	7
11110	1	1
11113	7	7
11114	1	1
11115	4	4
11116	1	5
11117	1	1
11123	1	3
11124	1	1
11129	1	3
11134	1	1
11135	6	6

11139	3	1
11140	1	7
11144	1	3
11145	1	1
11146	1	1
11147	1	1
11151	1	1
11152	6	6
11153	1	1
11156	1	1
11158	5	3
11159	7	1
11160	6	6
11162	7	7
11164	5	6
11170	1	1
11173	1	3
11176	1	1
11177	3	1
11178	6	4
11179	1	4
11180	1	1
11182	6	6
11190	1	1
11193	1	1
11197	7	7
11198	1	1
11201	3	1
11202	7	3
11203	1	1
11208	1	1
11211	1	1
11213	1	5
11215	3	3
11216	1	1
11217	8	1
11218	1	7
11219	1	1
11224	1	1
11225	6	1
11230	1	1
11231	9	1
11232	7	1
11233	8	3
11236	1	1
11237	1	1
11238	1	8
11243	1	1
11246	1	1

11247	1	1
11252	3	1
11254	1	1
11256	1	1
11257	3	3
11258	7	7
11259	1	3
11262	1	7
11263	6	6
11265	1	1
11266	4	1
11267	5	1
11268	6	1
11269	3	3
11272	1	1
11273	1	6
11276	10	10
11280	9	7
11282	9	1
11284	3	4
11286	3	3
11293	1	1
11294	1	1
11298	1	1
11302	4	5
11305	4	1
11311	1	1
11312	1	7
11320	4	1
11321	1	1
11322	8	8
11331	1	1
11333	9	9
11336	1	1
11338	1	1
11340	3	3
11341	1	1
11342	3	3
11349	4	9
11350	1	1
11355	6	1
11357	6	6
11358	6	6
11360	1	1
11363	1	1
11364	4	1
11365	1	1
11366	4	1
11367	1	1

11368	1	1
11372	1	1
11373	7	6
11375	11	1
11380	7	6
11381	6	1
11386	1	1
11387	6	1
11389	1	1
11390	1	1
11391	1	1
11392	3	3
11393	4	1
11396	6	7
11398	5	1
11399	1	1
11400	1	8
11401	1	1
11402	1	3
11405	5	1
11406	1	1
11407	1	1
11411	1	1
11412	7	7
11414	1	1
11416	1	1
11418	7	7
11420	1	1
11424	1	1
11426	4	1
11427	1	11
11428	4	1
11429	1	5
11432	7	1
11433	1	5
11434	1	1
11436	1	1
11438	1	1
11441	6	1
11442	1	1
11443	8	8
11446	4	1
11447	7	1
11449	1	3
11450	1	1
11453	1	1
11454	1	1
11458	1	1
11459	1	1

11460	4	1
11461	1	1
11462	1	1
11467	1	1
11469	1	1
11470	1	1
11471	7	7
11473	4	1
11475	7	7
11476	1	1
11477	1	3
11480	7	7
11481	4	1
11484	1	1
11485	1	6
11488	1	7
11489	9	9
11490	7	7
11491	5	5
11493	1	1
11498	1	1
11500	1	1
11501	1	1
11502	1	1
11504	1	1
11505	7	7
11506	1	1
11510	4	1
11511	6	1
11512	5	5
11522	6	1
11525	6	6
11526	1	6
11528	1	1
11530	1	1
11533	1	1
11536	1	1
11542	1	1
11545	1	1
11549	11	1
11550	4	4
11551	1	1
11557	1	1
11558	3	1
11559	1	1
11560	4	4
11561	6	6
11563	3	1
11564	1	1

11566	1	1
11567	1	1
11568	1	6
11569	4	4
11570	1	1
11571	1	1
11572	3	3
11574	1	1
11577	4	4
11578	1	1
11579	1	1
11581	1	1
11582	4	1
11583	1	1
11584	1	1
11585	3	3
11588	5	3
11589	5	5
11594	3	5
11595	5	1
11596	1	1
11602	1	1
11605	1	1
11606	1	8
11607	1	1
11608	1	1
11611	7	7
11612	4	4
11613	1	1
11614	1	1
11620	10	10
11622	1	1
11624	1	1
11625	1	1
11627	1	1
11629	1	1
11630	1	1
11632	1	1
11634	1	1
11636	7	1
11637	1	1
11638	1	3
11640	1	3
11641	1	1
11643	1	1
11647	1	1
11648	3	3
11649	1	1
11653	5	5

11656	1	1
11661	7	6
11662	1	1
11665	5	6
11666	1	3
11667	5	1
11669	1	1
11670	1	1
11671	1	1
11673	1	1
11676	1	1
11677	7	1
11683	7	7
11684	1	1
11686	1	1
11689	1	1
11693	1	1
11694	1	1
11695	1	1
11697	1	1
11710	7	7
11716	1	1
11725	1	1
11730	1	1
13006	1	1
13405	3	3
14141	5	5
14269	9	9
14350	1	1
14352	1	1
14353	1	1
14356	1	1

ANEXO 2

Tabla A2

	N-NO-C	M-MA-C	ND-NO-C	N-PR-C	CD-NO-S	EC-NO-S	N-NO-D	TP-TS-C	M-NO-D	PR-NO-C
N-NO-C	79,6%	25,3%	23,3%	26,5%	28,4%	20,1%	46,2%	9,5%	13,3%	13,3%
M-MA-C	4,9%	54,4%	5,0%	5,9%	3,0%	4,3%	0,0%	0,0%	0,0%	0,0%
ND-NO-C	6,1%	1,3%	48,3%	2,9%	3,0%	2,9%	0,0%	9,5%	0,0%	40,0%
N-PR-C	3,7%	2,5%	3,3%	38,2%	4,5%	2,9%	0,0%	0,0%	0,0%	0,0%
CD-NO-S	4,5%	3,8%	3,3%	2,9%	47,8%	7,2%	7,7%	9,5%	0,0%	0,0%
EC-NO-S	2,9%	5,1%	0,0%	2,9%	10,4%	59,0%	7,7%	0,0%	13,3%	0,0%
N-NO-D	0,4%	2,5%	0,0%	0,0%	1,5%	0,0%	46,2%	0,0%	13,3%	0,0%
TP-TS-C	0,8%	1,3%	0,0%	0,0%	0,0%	1,4%	0,0%	57,1%	0,0%	0,0%
M-NO-D	0,2%	1,3%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	66,7%	0,0%
PR-NO-C	0,6%	3,8%	0,0%	0,0%	0,0%	1,4%	0,0%	0,0%	0,0%	26,7%

La tabla A2 muestra el porcentaje de ventas que representa cada celda con respecto a las ventas totales por marca.

ANEXO 3

Tabla B1

N-NO-C	M-MA-C	ND-NO-C	N-PR-C	CD-NO-S	EC-NO-D	N-NO-D	TP-TS-C	M-NO-D	PR-NO-C	obs		OldFit	NewFit1	NewFit2	Max.Prob	Segmento
1	0	0	1	0	0	0	0	0	0	91	. .	0,2935	-0,1346	-0,1345	0,9971	3
1	0	0	1	0	0	0	0	0	0	92	. .	0,2935	-0,1346	-0,1345	0,9971	3
1	0	0	1	0	0	0	0	0	0	93	. .	0,2935	-0,1346	-0,1345	0,9971	3
1	0	0	1	0	0	0	0	0	0	94	. .	0,2935	-0,1346	-0,1345	0,9971	3
1	0	0	1	0	0	0	0	0	0	95	. .	0,2935	-0,1346	-0,1345	0,9971	3
1	0	0	1	0	0	0	0	0	0	96	. .	0,2935	-0,1346	-0,1345	0,9971	3
1	0	0	1	0	0	0	0	0	0	97	. .	0,2935	-0,1346	-0,1345	0,9971	3
1	0	0	1	0	0	0	0	0	0	98	. .	0,2935	-0,1346	-0,1345	0,9971	3
1	0	0	1	0	0	0	0	0	0	99	. .	0,2935	-0,1346	-0,1345	0,9971	3
1	0	0	1	0	0	0	0	0	0	100	. .	0,2935	-0,1346	-0,1345	0,9971	3
1	0	0	1	0	0	0	0	0	0	101	. .	0,2935	-0,1346	-0,1345	0,9971	3
1	0	0	1	0	0	0	0	0	0	102	. .	0,2935	-0,1346	-0,1345	0,9971	3
1	0	0	1	0	0	0	0	0	0	103	. .	0,2935	-0,1346	-0,1345	0,9971	3
1	0	0	1	0	0	0	0	0	0	104	. .	0,2935	-0,1346	-0,1345	0,9971	3
1	0	0	1	0	0	0	0	0	0	105	. .	0,2935	-0,1346	-0,1345	0,9971	3
1	0	0	1	0	0	0	0	0	0	106	. .	0,2935	-0,1346	-0,1345	0,9971	3
1	0	0	1	0	0	0	0	0	0	107	. .	0,2935	-0,1346	-0,1345	0,9971	3
1	0	0	1	0	0	0	0	0	0	108	. .	0,2935	-0,1346	-0,1345	0,9971	3
1	0	0	1	0	0	0	0	0	0	109	. .	0,2935	-0,1346	-0,1345	0,9971	3
1	0	0	1	0	0	0	0	0	0	110	. .	0,2935	-0,1346	-0,1345	0,9971	3
1	0	0	1	0	0	0	0	0	0	111	. .	0,2935	-0,1346	-0,1345	0,9971	3
1	0	0	1	0	0	0	0	0	0	112	. .	0,2935	-0,1346	-0,1345	0,9971	3
1	0	0	1	0	0	0	0	0	0	113	. .	0,2935	-0,1346	-0,1345	0,9971	3
1	0	0	1	0	0	0	0	0	0	114	. .	0,2935	-0,1346	-0,1345	0,9971	3
1	0	0	1	0	0	0	0	0	0	115	. .	0,2935	-0,1346	-0,1345	0,9971	3
1	0	0	1	0	0	0	0	0	0	116	. .	0,2935	-0,1346	-0,1345	0,9971	3
1	0	0	1	0	0	0	0	0	0	117	. .	0,2935	-0,1346	-0,1345	0,9971	3
1	0	0	1	0	0	0	0	0	0	118	. .	0,2935	-0,1346	-0,1345	0,9971	3
1	0	0	0	0	0	1	0	0	0	190	. .	0,6456	0,5391	0,5388	0,9987	4
1	0	0	0	0	0	1	0	0	0	191	. .	0,6456	0,5391	0,5388	0,9987	4
1	0	0	0	0	0	1	0	0	0	192	. .	0,6456	0,5391	0,5388	0,9987	4
1	0	0	0	0	0	1	0	0	0	193	. .	0,6456	0,5391	0,5388	0,9987	4
1	0	0	0	0	0	1	0	0	0	194	. .	0,6456	0,5391	0,5388	0,9987	4

1	0	0	0	0	0	1	0	0	0	195	.	.	0,6456	0,5391	0,5388	0,9987	4
1	0	0	0	0	0	1	0	0	0	196	.	.	0,6456	0,5391	0,5388	0,9987	4
1	0	0	0	0	0	1	0	0	0	197	.	.	0,6456	0,5391	0,5388	0,9987	4
1	0	0	0	0	0	0	1	0	0	198	.	.	0,1235	-0,2044	-0,2041	0,9974	4
1	0	0	0	0	0	0	1	0	0	199	.	.	0,1235	-0,2044	-0,2041	0,9974	4
1	0	0	0	0	0	0	1	0	0	200	.	.	0,1235	-0,2044	-0,2041	0,9974	4
1	0	0	0	0	0	0	1	0	0	201	.	.	0,1235	-0,2044	-0,2041	0,9974	4
1	0	0	0	0	0	0	1	0	0	202	.	.	0,1235	-0,2044	-0,2041	0,9974	4
1	0	0	0	0	0	0	0	0	1	203	.	.	-0,074	-0,3998	-0,3984	0,9938	4
1	0	0	0	0	0	0	0	0	1	204	.	.	-0,074	-0,3998	-0,3984	0,9938	4
1	0	0	0	0	0	0	0	0	1	205	.	.	-0,074	-0,3998	-0,3984	0,9938	4
1	0	0	0	0	0	0	0	0	1	206	.	.	-0,074	-0,3998	-0,3984	0,9938	4
1	0	0	0	0	0	0	0	1	0	207	.	.	-0,4309	-0,703	-0,7012	0,9946	4
1	0	0	0	0	0	0	0	1	0	208	.	.	-0,4309	-0,703	-0,7012	0,9946	4
0	1	1	0	0	0	0	0	0	0	209	.	.	0,6842	0,6089	0,6079	0,9959	1
0	1	1	0	0	0	0	0	0	0	210	.	.	0,6842	0,6089	0,6079	0,9959	1
0	1	1	0	0	0	0	0	0	0	211	.	.	0,6842	0,6089	0,6079	0,9959	1
0	1	1	0	0	0	0	0	0	0	212	.	.	0,6842	0,6089	0,6079	0,9959	1
0	1	0	1	0	0	0	0	0	0	213	.	.	0,4279	0,26	0,2591	0,9922	1
0	1	0	1	0	0	0	0	0	0	214	.	.	0,4279	0,26	0,2591	0,9922	1
0	1	0	1	0	0	0	0	0	0	215	.	.	0,4279	0,26	0,2591	0,9922	1
0	1	0	1	0	0	0	0	0	0	216	.	.	0,4279	0,26	0,2591	0,9922	1
0	1	0	0	1	0	0	0	0	0	217	.	.	-0,1365	-0,4011	-0,3937	0,9511	2
0	1	0	0	1	0	0	0	0	0	218	.	.	-0,1365	-0,4011	-0,3937	0,9511	2
0	1	0	0	1	0	0	0	0	0	219	.	.	-0,1365	-0,4011	-0,3937	0,9511	2
0	1	0	0	1	0	0	0	0	0	220	.	.	-0,1365	-0,4011	-0,3937	0,9511	2
0	1	0	0	1	0	0	0	0	0	221	.	.	-0,1365	-0,4011	-0,3937	0,9511	2
0	1	0	0	0	1	0	0	0	0	222	.	.	0,8372	0,9073	0,902	0,9847	1
0	1	0	0	0	1	0	0	0	0	223	.	.	0,8372	0,9073	0,902	0,9847	1
0	1	0	0	0	1	0	0	0	0	224	.	.	0,8372	0,9073	0,902	0,9847	1
0	1	0	0	0	1	0	0	0	0	225	.	.	0,8372	0,9073	0,902	0,9847	1
0	1	0	0	0	1	0	0	0	0	226	.	.	0,8372	0,9073	0,902	0,9847	1
0	1	0	0	0	1	0	0	0	0	227	.	.	0,8372	0,9073	0,902	0,9847	1
0	1	0	0	0	1	0	0	0	0	228	.	.	0,8372	0,9073	0,902	0,9847	1
0	1	0	0	0	0	1	0	0	0	229	.	.	-0,3377	-0,5363	-0,5331	0,9868	1
0	1	0	0	0	0	1	0	0	0	230	.	.	-0,3377	-0,5363	-0,5331	0,9868	1
0	1	0	0	0	0	0	1	0	0	231	.	.	-0,4172	-0,6077	-0,6056	0,992	1
0	1	0	0	0	0	0	0	0	1	232	.	.	0,1911	-0,037	-0,037	0,999	1
0	1	0	0	0	0	0	0	0	1	233	.	.	0,1911	-0,037	-0,037	0,999	1
0	1	0	0	0	0	0	0	0	1	234	.	.	0,1911	-0,037	-0,037	0,999	1
0	1	0	0	0	0	0	0	1	0	235	.	.	-0,4972	-0,6778	-0,677	0,9976	1
0	0	1	1	0	0	0	0	0	0	236	.	.	-0,2089	-0,4116	-0,4084	0,9829	1
0	0	1	1	0	0	0	0	0	0	237	.	.	-0,2089	-0,4116	-0,4084	0,9829	1
0	0	1	1	0	0	0	0	0	0	238	.	.	-0,2089	-0,4116	-0,4084	0,9829	1
0	0	1	0	1	0	0	0	0	0	239	.	.	-0,2513	-0,5037	-0,499	0,9759	2
0	0	1	0	1	0	0	0	0	0	240	.	.	-0,2513	-0,5037	-0,499	0,9759	2
0	0	1	0	1	0	0	0	0	0	241	.	.	-0,2513	-0,5037	-0,499	0,9759	2
0	0	1	0	1	0	0	0	0	0	242	.	.	-0,2513	-0,5037	-0,499	0,9759	2
0	0	1	0	0	1	0	0	0	0	243	.	.	0,1988	0,0019	0,0019	0,9688	1

0	0	1	0	0	1	0	0	0	0	244	.	.	0,1988	0,0019	0,0019	0,9688	1
0	0	1	0	0	0	0	1	0	0	245	.	.	-1,0478	-1,0867	-1,0783	0,9827	1
0	0	1	0	0	0	0	0	0	1	246	.	.	-0,4177	-0,6121	-0,6114	0,9976	1
0	0	1	0	0	0	0	0	0	1	247	.	.	-0,4177	-0,6121	-0,6114	0,9976	1
0	0	1	0	0	0	0	0	0	1	248	.	.	-0,4177	-0,6121	-0,6114	0,9976	1
0	0	0	1	1	0	0	0	0	0	249	.	.	-0,2374	-0,487	-0,4834	0,9821	2
0	0	0	1	1	0	0	0	0	0	250	.	.	-0,2374	-0,487	-0,4834	0,9821	2
0	0	0	1	1	0	0	0	0	0	251	.	.	-0,2374	-0,487	-0,4834	0,9821	2
0	0	0	1	1	0	0	0	0	0	252	.	.	-0,2374	-0,487	-0,4834	0,9821	2
0	0	0	1	0	1	0	0	0	0	253	.	.	-0,0875	-0,2846	-0,2791	0,951	1
0	0	0	1	0	1	0	0	0	0	254	.	.	-0,0875	-0,2846	-0,2791	0,951	1
0	0	0	1	0	1	0	0	0	0	255	.	.	-0,0875	-0,2846	-0,2791	0,951	1
0	0	0	0	1	1	0	0	0	0	256	.	.	0,7228	0,6786	0,6769	0,9934	2
0	0	0	0	1	1	0	0	0	0	257	.	.	0,7228	0,6786	0,6769	0,9934	2
0	0	0	0	1	1	0	0	0	0	258	.	.	0,7228	0,6786	0,6769	0,9934	2
0	0	0	0	1	1	0	0	0	0	259	.	.	0,7228	0,6786	0,6769	0,9934	2
0	0	0	0	1	1	0	0	0	0	260	.	.	0,7228	0,6786	0,6769	0,9934	2
0	0	0	0	1	1	0	0	0	0	261	.	.	0,7228	0,6786	0,6769	0,9934	2
0	0	0	0	1	1	0	0	0	0	262	.	.	0,7228	0,6786	0,6769	0,9934	2
0	0	0	0	1	1	0	0	0	0	263	.	.	0,7228	0,6786	0,6769	0,9934	2
0	0	0	0	1	1	0	0	0	0	264	.	.	0,7228	0,6786	0,6769	0,9934	2
0	0	0	0	1	1	0	0	0	0	265	.	.	0,7228	0,6786	0,6769	0,9934	2
0	0	0	0	1	1	0	0	0	0	266	.	.	0,7228	0,6786	0,6769	0,9934	2
0	0	0	0	1	1	0	0	0	0	267	.	.	0,7228	0,6786	0,6769	0,9934	2
0	0	0	0	1	1	0	0	0	0	268	.	.	0,7228	0,6786	0,6769	0,9934	2
0	0	0	0	1	0	1	0	0	0	269	.	.	-0,6457	-0,8216	-0,8176	0,988	2
0	0	0	0	1	0	1	0	0	0	270	.	.	-0,6457	-0,8216	-0,8176	0,988	2
0	0	0	0	1	0	0	1	0	0	271	.	.	-1,0167	-1,0865	-1,0788	0,9824	2
0	0	0	0	0	1	1	0	0	0	272	.	.	-0,8823	-0,9711	-0,9419	0,922	1
0	0	0	0	0	1	0	1	0	0	273	.	.	-0,9283	-1,0034	-0,9838	0,9501	1
0	0	0	0	0	1	0	0	0	1	274	.	.	-0,2413	-0,4576	-0,4566	0,9952	1
0	0	0	0	0	1	0	0	1	0	275	.	.	-0,9486	-1,0173	-1,0118	0,9869	1
0	0	0	0	0	0	1	0	1	0	276	.	.	-2,4183	-2,2078	-1,8372	0,5129	1

	0	0	0	0	0	0	0	0	0	0
	0	0	0	0	0	0	0	0	2	0
	0	0	0	0	0	0	0	0	2	0
	0	0	0	0	0	0	0	0	2	0
	0	0	0	0	0	0	0	0	2	0
	0	0	0	0	0	0	0	0	2	0
	0	0	0	0	0	0	0	0	0	0
	0	0	0	0	0	0	0	0	0	2
Total	1018	133	101	54	108	140	10	12	10	2

ANEXO 4

Tabla B3

Marca	Precio
N-PR-C	15,07
N-NO-D	14,85
M-NO-D	14,62
M-MA-C	10,58
N-NO-C	10,24
PR-NO-C	10,05
ND-NO-C	9,34
TP-TS-C	8,13
EC-NO-S	3,93
CD-NO-S	3,75

En la tabla que aquí aparece, presentamos una especie de ranking de precios, donde se muestran los precios promedios que tenían éstos productos al

Como sabemos, los productos de la muestra no son homogéneos, es decir de una misma marca y categoría de producto tenemos distintos tamaños, distintos precios, distintas promociones, además la mayoría fueron comprados en lugares distintos. Otro hecho relevante en cuanto a los productos es que las preferencias en cuanto a tamaño varia dependiendo del producto, es así como hay productos que son preferidos en tamaño de 170 gramos y otros que son preferidos en tamaño de 250 o incluso en sobre. Es por esto que para sacar el

precio promedio por marca decidimos homogeneizar las elecciones, es decir, tomamos el tamaño más comprado en casi todas las marcas, en éste caso fueron los tarros de 170 gramos, tomamos solo las compras correspondientes a éste tamaño, en la muestra que tenía solo las compras de las marcas elegidas para el estudio, de ahí sacamos los precios promedios por marca.

ANEXO 5

Ahora describiremos a las variables mencionadas:

-Marca: En la siguiente tabla mostramos como identificamos a cada marca con un número.

Tabla B4

Nº	marca
1	OMO
2	ACE
3	DRIVE
4	DETERGENTE
5	ARIEL
6	RINSO
7	ACUENTA
8	BILLOW
9	BILORE
10	BIO WASH
11	BISSEL
12	BOLD 3

13	BOLD
14	BRISA 3
15	CINCO CONTINENTES
19	EKONO
21	FUZOL
23	IBERIA
24	KEL
25	KLAREN
26	KOP
27	KORALL
28	LAR
29	LIDER
30	LINA

31	MONTECARLO
36	PERS
37	PERSIL
38	PERWOLL
39	POPEYE
40	SKIP
41	SUPER
42	TISBE
44	UNIMARC
46	ZIG ZAG
47	ABRIL

48	CASABLANCA
49	DEMAX
50	DETER
51	DIAMANTE
52	KAS
53	LORE
54	WOOLITE
55	ZAS

- **Matic:** Aquí mostramos que producto sirve para lavado a mano o que producto sirve para lavadora automática, la forma de hacerlo fue utilizando variables dummies donde

