

# Material de consulta

Estructurado en siete temas  
Este corresponde al Tema 6

Elaborado por:  
Econ. Milton Oroche Carbajal

**TEMA 1:  
INVESTIGACIÓN DE MERCADOS.**

**TEMA 2:  
MÉTODOS CUALITATIVOS DE INVESTIGACIÓN DE MERCADOS.**

**TEMA 3:  
MÉTODOS CUANTITATIVOS DE INVESTIGACIÓN DE MERCADOS.**

**TEMA 4:  
MÉTODOS DE ANÁLISIS DE LA INFORMACIÓN: ANÁLISIS  
UNIVARIABLE Y BIVARIABLE.**

**TEMA 5:  
ANÁLISIS CAUSAL: LA EXPERIMENTACIÓN COMERCIAL.**

**TEMA 6:  
EL ANÁLISIS MULTIVARIABLE Y SUS APLICACIONES AL  
MARKETING.**

6.1 Métodos de Análisis Multivariable: Dependencia versus Interdependencia.

6.2 Aplicación del Análisis Factorial a la Investigación de Mercados.

6.3 Análisis Cluster.

6.4 Análisis Discriminante.

**TEMA 7:  
DISEÑO Y PRESENTACIÓN DE INFORMES DE INVESTIGACIÓN.**

## TEMA 6: EL ANÁLISIS MULTIVARIABLE Y SUS APLICACIONES AL MARKETING.

### 6.1. MÉTODOS DE ANÁLISIS MULTIVARIABLE: DEPENDENCIA VERSUS INTERDEPENDENCIA.

Las técnicas multivariantes se pueden definir como el conjunto de métodos estadísticos que analizan de forma simultánea dos o más variables observadas, es decir, permiten tener una visión de conjunto de los fenómenos y estudiar las posibles interacciones que puedan existir entre los diversos factores. Las tres razones por las que se aplica el análisis multivariable en marketing son las siguientes:

- 1- Permite analizar simultáneamente toda la información relativa a un determinado fenómeno.
- 2- Permite analizar un gran número de encuestas.
- 3- Por su capacidad para hacer comprensible para la mente humana grandes cantidades de datos que de otra manera serían difíciles de manejar, aceptando eso si una mínima pérdida de información y proporcionando en muchos casos una representación gráfica del fenómeno.

#### ➤ **CLASIFICACIÓN DE LOS MÉTODOS MULTIVARIABLES**

Se pueden distinguir dos grandes bloques de métodos o técnicas multivariantes: métodos descriptivos y métodos explicativos.

- 1- **Métodos descriptivos (o de interdependencia):** En estos métodos no se va a distinguir entre variables dependientes o a explicar y variables independientes o explicativas. Esto quiere decir que todas las variables están relacionadas unas con otras, es decir, están interrelacionadas. Dentro de los métodos descriptivos se pueden distinguir los siguientes grupos de técnicas:

- a) Métodos de clasificación, los cuales permiten agrupar individuos, empresas, etc. a partir de un grupo de variables de partida. En concreto, en este tema, de este grupo de métodos se va a ver el análisis cluster.
- b) Otro bloque de métodos descriptivos serían las técnicas que permiten resumir información. Estos serían los métodos factoriales, dentro de los cuales se va a estudiar el análisis de componentes principales.

2- **Métodos explicativos (o de dependencia):** Son aquellos que explican unas variables en función de otras, es decir, hay variables dependientes o a explicar y variables independientes o explicativas. Dentro de este grupo, los métodos o técnicas concretas que se van a estudiar en este tema son las siguientes:

- a) Métodos de segmentación.
- b) Análisis del discriminante.

## **6.2. APLICACIÓN DEL ANÁLISIS FACTORIAL A LA INVESTIGACIÓN DE MERCADOS.**

Los métodos factoriales son técnicas de análisis multivariable cuyo objetivo principal es resumir la información de partida para que sea más fácil de interpretar. Una definición de los métodos factoriales es la siguiente: son un conjunto de técnicas multivariadas que proporcionan una visión simplificada de un determinado fenómeno considerando todas las variables que influyen sobre dicho fenómeno a la vez.

Concretamente, los métodos factoriales crean, a partir de un gran conjunto inicial de variables, un conjunto nuevo y menor con variables que son combinaciones lineales de las variables de partida.

La esencia de los métodos factoriales es que a través de consentir una pequeña pérdida de información se gana en significación, de tal manera que el resumen de variables que resulta de la aplicación de estas técnicas ofrece una representación simplificada de la realidad, es decir, da una visión de un fenómeno mucho más fácil de comprender y de ver por los individuos.

Existen muchos métodos factoriales, entre los cuales se puede citar el análisis de componentes principales, que es en el que nos vamos a centrar a continuación.

➤ **ANÁLISIS DE COMPONENTES PRINCIPALES**

El análisis de componentes principales es un método factorial, es decir, es una técnica multivariable descriptiva que tiene como objetivo principal resumir la información de partida. Las aplicaciones fundamentales del análisis de componentes principales son las dos siguientes:

- 1) Resumir el número de variables originales de partida, eliminando redundancias (las variables iniciales pueden estar midiendo la misma característica).
- 2) Identificar estructuras latentes, es decir, descubrir fenómenos que no son directamente observables y que resultan de la interacción de varios atributos o variables.

Algo muy importante es saber que datos son los idóneos para la aplicación de esta técnica. Los datos de partida para poder aplicar el análisis de componentes principales tienen que venir medidos con escalas de razón o con escalas de intervalo, siendo las siguientes escalas de intervalos las comúnmente utilizadas: escalas de Likert y escalas de diferencial-semántico. Con otro tipo de datos no se podría aplicar esta técnica o análisis.

Los pasos que se han de seguir para realizar un análisis de componentes principales son los siguientes:

1- Verificar la idoneidad de los datos de partida o iniciales para la aplicación de un análisis de componentes principales. Para comprobar la idoneidad de los datos se pueden examinar varios indicadores o índices:

a) La matriz de correlaciones de las variables iniciales. Por ejemplo, una matriz de correlaciones tendría la siguiente forma:

	A <sub>1</sub>	A <sub>2</sub>	A <sub>3</sub>	A <sub>4</sub>	A <sub>5</sub>	A <sub>6</sub>	A <sub>7</sub>	A <sub>8</sub>	A <sub>9</sub>	A <sub>10</sub>
A <sub>1</sub>	1									
A <sub>2</sub>		1								
A <sub>3</sub>			1							
A <sub>4</sub>				1						
A <sub>5</sub>					1					
A <sub>6</sub>						1				
A <sub>7</sub>							1			
A <sub>8</sub>								1		
A <sub>9</sub>									1	
A <sub>10</sub>										1

Una prueba de la idoneidad de las variables iniciales para poder aplicar el análisis de componentes principales es que las correlaciones entre las variables iniciales sean altas. Para ello miraríamos los datos de la matriz de correlaciones, de forma que cuanto más próximos a uno estén los datos esto significaría que las variables están altamente intercorrelacionadas.

Otro indicador del grado de asociación de los atributos sino se tuviera la matriz de correlaciones sería el determinante de la matriz de correlaciones. Si este determinante es muy bajo significa que hay variables con intercorrelaciones muy altas y, por lo tanto, se podría aplicar el análisis de componentes principales. En resumen, lo que interesa es que el determinante sea lo más bajo posible ya que esto implica un elevado grado o nivel de interrelación.

b) Un segundo indicador sería el índice de Kaiser-Meyer-Olkin, que oscila entre 0 y 1. Si el valor está próximo a uno indica que los datos iniciales son adecuados para aplicar el análisis de componentes principales. Una clasificación, aunque hay que dejar claro que existen muchas otras, sería la siguiente.

- 0,9 < Índice de K-M-O ≤ 1 : Muy bueno
- 0,8 < Índice de K-M-O ≤ 0,9 : Meritorio
- 0,7 < Índice de K-M-O ≤ 0,8 : Bueno
- 0,5 < Índice de K-M-O ≤ 0,7 : Mediano
- Índice de K-M-O ≤ 0,5 : Inaceptable

c) Un tercer indicador sería el test de esfericidad de Bartlett, el cual somete a comprobación la hipótesis de que la matriz de

correlaciones es una matriz identidad. Lo ideal es que salga un valor muy alto y un nivel de significación muy bajo ( $< 0,05$ ).

2- Extracción de los factores necesarios para representar los datos.

a) El paso siguiente es estudiar los factores obtenidos o derivados del análisis de componentes principales. Lo primero que se tiene que hacer es determinar el número de factores retenidos, es decir, el número de factores con los cuales se van a trabajar (número de factores en los que se resumen las variables iniciales). Para determinar este número hay muchas reglas, pero la más utilizada es que el valor propio o autovalor del factor sea mayor que 1. El valor propio o autovalor es la cantidad de información explicada por un factor.

b) Estudiar la composición de los factores retenidos. Se trata de ver de que son combinación lineal las nuevas variables, para lo cual se tiene que mirar la matriz factorial, que es una matriz que reproduce la primera de forma sencilla. Los números que están dentro de esta matriz se denominan cargas o pesos factoriales y son los coeficientes de las combinaciones lineales que se han creado para cada uno de los factores. Estas cargas indican el peso que cada variable inicial tiene en los factores obtenidos. Puede haber cargas factoriales con signo positivo y negativo, pero lo que nos interesa a nosotros es el valor absoluto. Las cargas factoriales oscilan en valor absoluto entre 0 y 1.

3- Interpretación y representación gráfica de los factores.

A partir de los resultados obtenidos hay que proceder a interpretar los factores. A efectos prácticos, en la interpretación de los factores se sugieren los pasos siguientes:

a) Estudiar la composición de las cargas factoriales significativas de cada factor. Para ello se podría utilizar una matriz factorial, pero sin embargo se va a emplear una matriz factorial rotada, que se elabora a partir de la matriz factorial antes dicha.

b) Esta nueva matriz a partir de un proceso concreto de rotación permite obtener factores ortogonales (independientes totalmente), de tal manera que las variables iniciales que tienen un elevado peso en un determinado factor tengan poco peso en los demás factores. La matriz de componentes rotados es muy parecida a la anterior pero con diferencias más claras.

c) A partir de las cargas factoriales que aparecen en la matriz de componentes rotados se puede poner un nombre a los factores obtenidos, de tal manera que la denominación de un factor estará relacionada con las variables que más peso tengan en el factor.

Por último, habría que realizar la representación gráfica de los resultados. La representación gráfica se hace tomando los factores de dos en dos. Cada factor representa un eje de coordenadas, es decir, se representará un factor sobre un eje horizontal y el otro factor sobre un eje vertical.

Sobre este mapa se pueden representar diversas cosas como las cargas factoriales, las puntuaciones otorgadas por los individuos a cada uno de los factores y a través de ellas se permite analizar la imagen de una empresa, la imagen de un producto o la existencia de huecos en el mercado a través de los cuales la empresa pueda diversificarse. Estos mapas se denominan mapas de perfección y posicionamiento.

### **6.3. ANÁLISIS CLUSTER.**

La segmentación de mercados consiste en la agregación de consumidores en grupos homogéneos, cada uno de los cuales puede ser seleccionado como un segmento-objetivo al cual aplicar un Marketing-mix diferenciado. Hay dos grandes grupos de clasificación o agrupación de elementos o, lo que es lo mismo, de división del mercado:

a) Técnicas de segmentación: los métodos clásicos de segmentación (Belson, Sonquist y Morgan, Chi-cuadrado, etc.) son ejemplos de técnicas de segmentación. Estas técnicas o métodos se caracterizan por tener una variable dependiente a explicar y una serie de variables explicativas, que son los criterios de segmentación y que se establecen a priori por el investigador.

b) Técnicas de tipología: en las técnicas de tipología, entre las que se incluye el análisis cluster, no hay variables a explicar ni variables explicativas sino que los grupos se forman o los elementos se clasifican atendiendo a la similitud de los perfiles multivariables de los individuos o de las empresas. En estas técnicas no se seleccionan o se forman los grupos a priori.

Las etapas o pasos que se han de seguir en la aplicación del análisis cluster son los siguientes:

1) Hay que especificar o definir las variables que se van a utilizar en la segmentación del mercado, es decir, escoger las variables en función de las cuales se van a clasificar los individuos, para lo cual:

- Tomar como referencia bases de segmentación.
- Ser creativos en la propuesta de variables.

2) Seleccionar las medidas de proximidad que se van a utilizar en la segmentación. Las medidas de proximidad reflejan y cuantifican el grado de semejanza o diferencia que existe entre los elementos comparados. Hay distintos tipos de medidas de proximidad que se pueden agrupar en tres grandes grupos:

- Índices de Semejanza.
- Medidas de Distancia.
- Coeficientes de Correlación.

De estos tres grupos se va a trabajar principalmente con las del grupos dos, es decir, con las medidas de distancia.

a- Los índices de semejanza, que se llaman también coeficientes de asociación, miden el grado de similitud existente entre dos

elementos (dos empresas, dos marcas, etc.). A medida que aumenta el valor de este índice aumenta también la semejanza entre los elementos comparados, o lo que es igual, disminuye la diferencia. A medida que aumenta la diferencia disminuye el valor de este índice.

b- Las medidas de distancia son medidas de no similitud de tal manera que cuanto mayor sea su valor menos semejanza existe entre los elementos o individuos comparados. El significado de las medidas de distancia es el inverso a el de los índices de semejanza. La medida de distancia más popular y por defecto es la euclídea y se calcula como la raíz cuadrada de la suma de los cuadrado de las diferencias existentes entre los valores de los atributos para los individuos comparados.

$$D_{ij} = \sqrt{\sum (X_{ik} - X_{jk})^2}$$

$D_{ij}$  = Distancia euclídea entre el consumidor i y el consumidor j.

$X_{ik}$  = Valor del atributo k para el consumidor i.

$X_{jk}$  = Valor del atributo k para el consumidor j.

Las medidas de distancia, en concreto las euclídeas, tienen los siguientes dos inconvenientes:

1- Si los atributos están expresados en diferentes unidades de medida la solución más usual es normalizar los datos, lo que conlleva alisar las diferencias que existen entre los elementos, lo que perjudica al cálculo de la distancia.

2- Si los atributos están muy correlacionados las medidas de distancia sesgan los resultados.

c- Los coeficientes de correlación miden el grado de dependencia mutua que existe entre los elementos comparados. Los principales son el coeficiente de Pearson (se emplea para datos de tipo cuantitativo), el coeficiente de Kendall y el coeficiente de Spearman (se emplean para datos ordinales).

3) La tercera etapa del análisis cluster consiste en seleccionar el algoritmo de clasificación o agrupación que se va a utilizar o aplicar en la segmentación, es decir, determinar la manera en que se van a formar los grupos.

Hay tres grandes grupos de métodos o algoritmos de clasificación:

I) Métodos disjuntos: cada individuo analizado es asignado exclusivamente a un grupo, es decir, ningún individuo puede pertenecer simultáneamente a más de un grupo. Estos métodos pueden ser clasificados en tres categorías.

a) Los métodos jerárquicos constituyen la familia de técnicas más numerosa y difundida entre los investigadores. Los métodos ascendentes, también denominados asociativos-aglomerativos o tipología por concentración, responden al siguiente esquema: inicialmente los individuos se supone que son heterogéneos y que, por tanto, cada uno conforma un grupo; en cada etapa se trata de asociar o agrupar individuos o grupos según el proceso que realiza cada algoritmo hasta quedar un solo grupo. Los métodos descendentes, también denominados disociativos o divisivos, siguen el esquema opuesto: inicialmente los individuos se supone que son homogéneos y que, por tanto, conforman un único grupo; en cada etapa se trata de dividir cada grupo en subgrupos.

b) Los métodos de reasignación o técnicas de participación iterativa tratan de solventar el problema de una inadecuada asignación de individuos a los cluster correspondientes y pueden trabajar con muestras de gran tamaño. Se caracterizan porque en estos métodos el investigador determina a priori el número de segmentos a obtener. Se plantea una asignación inicial de los elementos o individuos a los grupos y a partir de dicha asignación los individuos son clasificados reiteradas veces en los diferentes cluster hasta alcanzar una cierta norma preespecificada. La técnica de participación iterativa más utilizada es el método k-means.

c) Métodos de búsqueda de densidad.

II) Métodos no disjuntos: permiten que un individuo pueda pertenecer a más de un grupo. Consideran que la pertenencia exclusiva no es realista en la práctica empresarial.

III) Métodos borrosos: mientras que los dos primeros establecen que un elemento pertenece o no a un cluster determinado (con o sin carácter exclusivo), los métodos borrosos reemplazan este

concepto por otro más gradual o tenue que indica la mayor o menor cercanía de un individuo a un conglomerado concreto.

4) La etapa final a la que se enfrenta el investigador que desea aplicar la metodología del análisis cluster, consiste en validar la existencia de los grupos obtenidos. El método más sencillo de validar la agrupación realizada es el método de reproducción que consiste en tomar muestras sucesivas y comprobar si se mantiene la estructura inicialmente propuesta.

#### **6.4. ANÁLISIS DISCRIMINANTE.**

El análisis discriminante, al igual que el análisis cluster, es una técnica de clasificación que es útil para explicar diferencias en el comportamiento de grupos, elementos, segmentos o conglomerados. En este análisis se va a trabajar con una variable dependiente y una serie de variables independientes o explicativas.

En el análisis discriminante la variable dependiente es la pertenencia a grupos de elementos (consumidores, empresas, individuos, etc.). La variable dependiente tiene dos categorías: 1 cuando el elemento si pertenece al grupo y 2 cuando no pertenece. En cuanto a las variables independientes o predictores son un conjunto de variables relevantes para el estudio que se está realizando y sobre las cuales se poseen observaciones para cada elemento de la muestra.

El objetivo del análisis discriminante es, en primer lugar, determinar si los grupos establecidos quedan suficientemente discriminados o diferenciados en función de las variables independientes disponibles. El segundo objetivo es analizar que variables son las que más contribuyen a discriminar entre los grupos formados.

El procedimiento a seguir en un análisis discriminante es el siguiente: a partir de las variables independientes o predictores se crean funciones discriminantes que son ecuaciones lineales en las que la variable dependiente es la pertenencia al grupo y las

variables independientes son los predictores ya mencionados combinados linealmente.

**Función discriminante = Combinación lineal de los predictores iniciales u originales**

Si se establecen a priori dos segmentos el análisis discriminante se llama simple y el número de funciones discriminantes es igual a uno. Si se trabajan con n segmentos el análisis discriminante se llama múltiple y el número de funciones discriminantes va a ser igual al número de segmentos menos uno.

$$FD = a \cdot P_1 + b \cdot P_2 + \dots + n \cdot P_n$$

Los pasos a seguir para interpretar los resultados de un análisis discriminante son los siguientes:

- 1- Análisis descriptivo de los predictores o variables independientes y estudio de las correlaciones que hay entre ellos. Para ello se estudian los valores de las medias y de las desviaciones típicas para cada una de las variables independientes y, a continuación, se calcula la matriz de correlaciones.
- 2- Análisis de los ratios F-Univariables, que consiste en hacer una comparación entre las medias de los grupos para cada una de las variables independientes consideradas individualmente.

Después de estos dos pasos se puede decir que las variables independientes o predictores discriminan o diferencian bastante entre el comportamiento de los individuos del grupo 1 y los del grupo 2. Pero no todos explican por igual esas diferencias.

3- El tercer paso es el estudio de los méritos de la función discriminante. Para estudiar los méritos existen varios indicadores.

- Valor propio, que es el indicador de la eficiencia de la función discriminante, de forma que cuanto más alto sea el valor propio más eficaz será la función.
- Varianza explicada tanto la intergrupo como la varianza total.

- Nivel de significación que se alcanza para una prueba Chi-Cuadrado en la que la hipótesis nula es que las medias de las funciones discriminantes son iguales para los grupos (lo que se pretende es rechazar la hipótesis nula porque lo que se quiere es que sean valores muy distintos).
- 4- Estudiar la contribución de cada variable a la función discriminante, para ello se van a tratar todas las variables conjuntamente. El indicador de la contribución de una variable a la función discriminante es el coeficiente estandarizado que le corresponde a dicha variable en la función discriminante y/o la correlación que existe entre el valor de dicha variable y el valor de la función. Cuanto más alto sea el coeficiente estandarizado mayor es la contribución. Por otra parte, cuanto más alta sea la correlación mayor es la contribución de la variable a la función discriminante.
- 5- La quinta etapa es la validación de los resultados. La herramienta que se suele utilizar es la matriz de confusión. En dicha matriz, por filas se recoge el grupo al que realmente pertenecen los elementos de la muestra y por columnas aparece recogido el grupo de pertenencia pronosticado a partir del valor que toma la función discriminante.

### **Fuente:**

- Artículo de revista: Pérez Díaz, J.L. (2000): "Título del Artículo", Revista, Volumen 54, Número 2, Páginas 41 – 53.
- Libro: Pérez Díaz, J.L. (2000): "Título del Libro", Editorial Díaz de Santos, Madrid (7ª edición).
- STANTON, William. Fundamentos de Marketing. McGraw-Hill. Décima Primera Edición. Méjico 2000 (únicamente el capítulo 4)
- POPE, Jeffry. Investigación de Mercados. Editorial Norma. 1986
- LOPEZ ALTAMIRANO, Alfredo. ¿ Qué son, para qué sirven y cómo se hacen las Investigaciones de Mercado? Editorial CECSA. Primera Edición. Méjico 2001.
- Kotler, P "Dirección de la Mercadotecnia, Análisis, Planeación, Implementación y Control" sexta edición 1992.
- - Lambin, JJ "Marketing Estrategico" Ed. Mc Graw-Hill
- Ceavens, Hills, Woodruff Administración en mercadotecnia Ed. CECSA
- Andrés Quijano Ponce de León, Elementos básicos de la mercadotecnia
- Gist, Ronald R., Principio de Mercadotecnia, Editora Interamericana, México, 1973,448 págs.
- Thomas, Kinnear, Investigación de Mercado, Mcgraw-Hill, Colombia, 1994, 760 págs.