

CURSO ACADÉMICO 1999 – 2000

TERCER CURSO

ÍNDICE

TEMA 1: INVESTIGACIÓN DE MERCADOS.

- 1.1. El Sistema de Información y la Investigación Comercial.**
- 1.2. Concepto, Contenido y Aplicaciones de la Investigación de Mercados.**
- 1.3. Metodología para Realizar un Estudio de Investigación de Mercados.**
- 1.4. Fuentes de Información y Análisis Documental.

TEMA 2: MÉTODOS CUALITATIVOS DE INVESTIGACIÓN DE MERCADOS.

- 2.1. Finalidad de la Investigación Cualitativa.
- 2.2. Reuniones de Grupo y Entrevista en Profundidad.
- 2.3. Pseudocompra y Técnicas Proyectivas.
- 2.4. Otras Técnicas Cualitativas de Investigación de Mercados.

TEMA 3: MÉTODOS CUANTITATIVOS DE INVESTIGACIÓN DE MERCADOS.

- 3.1. Finalidad de la Investigación Cuantitativa.
- 3.2. Concepción y Organización General de la Encuesta.
- 3.3. Determinación del Tamaño de la Muestra y Métodos de Muestreo.
- 3.4. Diseño de Cuestionario, Trabajo de Campo, Plan de Escrutinio y Análisis de Resultados.
- 3.5. Estudios de Mercado mediante Paneles y Observación.

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

- 4.1. Introducción. Codificación y Tabulación de Datos en el Análisis de la Información.
- 4.2. Métodos de Análisis de la Información Obtenida en la Investigación Comercial.
- 4.3. Métodos de Análisis Univariantes y Bivariantes.
- 4.4. Métodos de Inferencia Estadística.

TEMA 5: ANÁLISIS CAUSAL: LA EXPERIMENTACIÓN COMERCIAL.

5.1. Fundamentos y Planificación de la Experimentación de Mercado o Comercial.

5.2. Tipos de Experimentos.

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.

7.1. Tipos de Informes de Investigación de Mercados.

7.2. Normas de Redacción y Estructura del Informe.

7.3. Presentación Oral del Informe.

TEMA 1: INVESTIGACIÓN DE MERCADOS.

1.1. EL SISTEMA DE INFORMACIÓN Y LA INVESTIGACIÓN COMERCIAL.

Los directivos de las empresas llevan a cabo la gestión de muy variados recursos como: recursos financieros, recursos humanos, información, recursos tecnológicos y recursos físicos, siendo hoy en día uno de los más importantes la información. Por eso, actualmente las empresas tienen herramientas para tratar esa información. Hoy en día, la información se ha convertido en un aspecto fundamental ante los cambios que se han producido en el entorno de la empresa. El entorno es cada vez más complejo, más competitivo y sobre todo más incierto. Debido a estos cambios del entorno, las necesidades de información en las empresas han crecido de tal manera que muchas empresas han pensado que las actividades de marketing iban a ser las que permitirían conseguir toda esa información. Esto trajo consigo un desarrollo de las técnicas de marketing que son cada vez más sofisticadas, es decir, se han ido perfeccionando. El hecho de que las técnicas de marketing sean más sofisticadas se traduce en un incremento de las necesidades de información.

El encargado de suministrar toda esta información (entorno, departamento de marketing) va a ser el Sistema de Información de Marketing (SIM) que es el que permite recoger información del entorno y analizarla para tomar decisiones de manera eficiente. Por una parte, para conocer las oportunidades y amenazas que existan en el entorno tratando de aprovechar las oportunidades, y de neutralizar y solucionar las amenazas. Por otra parte, el SIM debe permitir diseñar políticas comerciales, planes de marketing concretos, etc. El análisis del entorno está relacionado con el marketing estratégico y el diseño de las políticas comerciales está relacionado con el marketing operativo.

Una definición sobre el Sistema de Información de Marketing dada por Kotler en el año 1992 sería la siguiente: Un Sistema de Información de Marketing es un conjunto de personas, equipos y procedimientos diseñados para recoger, clasificar, analizar, valorar y distribuir a tiempo la información demandada por los gestores de marketing. Las tres características principales de un Sistema de Información de Marketing son:

1- El Sistema de Información de Marketing está formado o integrado por cuatro subsistemas, que están relacionados unos con otros, es decir, están interrelacionados. Los cuatro subsistemas son: datos internos,

inteligencia de marketing, apoyo de marketing e investigación de mercados.

1)

2)

1) Valorar las necesidades de información.

2) Distribuir la información a tiempo.

- **DATOS INTERNOS:** Es el subsistema más básico o elemental y es el que recoge la información que se genera dentro de la empresa. Ejemplos: datos sobre costes, ventas, precios, producción, beneficios, personal de la empresa, niveles de productividad, clientes, inventarios, etc.
- **INTELIGENCIA DE MARKETING:** Este subsistema recoge información del entorno de la empresa utilizando fuentes de información secundarias, es decir, se trataría de información que ya está elaborada. Ejemplo: libros, periódicos, revistas, censos, estadísticas que elabora el INE, base de datos, paneles, anuarios, etc.
- **APOYO A LAS DECISIONES DE MARKETING:** Este subsistema está integrado por técnicas de análisis de datos que, con el apoyo de los equipos informáticos de las empresas, ayudan a los directivos de las empresas a analizar la información y tomar decisiones.
- **INVESTIGACIÓN DE MERCADOS:** Este subsistema es el encargado de realizar estudios relacionados con un problema determinado o una decisión concreta. Ejemplo: efectos del lanzamiento de un producto, de una subida del 10% del precio, eficacia de un anuncio, consecuencias de una promoción.

2- El Sistema de Información de Marketing contribuye a gestionar eficientemente la información. El Sistema de Información de Marketing contribuye a la toma de decisiones dentro de la empresa, es decir, un buen Sistema de Información de Marketing debe permitir a los responsables de marketing decidir que información es relevante para la empresa, clasificar esa información y analizarla, y ante todo distribuir esa información a tiempo.

3- El funcionamiento del Sistema de Información de Marketing requiere la utilización de tecnologías de información. Para poder analizar y evaluar con precisión y en un tiempo o plazo de tiempo razonable el enorme volumen de datos que se maneja hoy en día en las empresas es imprescindible utilizar tecnologías de información, tales como: ordenadores, scanners, faxes, televisión por cable, conexiones a la red, etc.

1.2. CONCEPTO, CONTENIDO Y APLICACIONES DE LA INVESTIGACIÓN DE MERCADOS.

La investigación de mercados es una parte integrante del Sistema de Información de Marketing, es decir, es un componente del Sistema de Información de Marketing que cumple la función de recoger y analizar información para resolver un problema puntual y concreto. En teoría, la investigación de mercados debería cumplir esa función pero en la práctica no siempre se limita a la solución de problemas concretos sino que hay una tendencia a utilizarla más de lo debido. Lo que caracteriza a la investigación de mercados es:

1- APOYAR EL PROCESO DE TOMA DE DECISIONES: La investigación de mercados sirve de apoyo al proceso de toma de decisiones, sobre todo cuando se trata de situaciones nuevas que no pueden ser resueltas en función de la experiencia que tenga la empresa.

2- INFORMACIÓN RELEVANTE: La investigación de mercados debe proporcionar información útil y

relevante para la empresa.

- a) Cuando reduce la incertidumbre que hay entorno a un problema.
- b) Cuando influye realmente sobre la toma de decisiones, es decir, no cuando se limita a confirmar algo que ya sospechamos sino que aporta algo nuevo.
- c) Cuando el coste de obtener esa información es menor que el valor de esa información.

Las aplicaciones de la investigación de mercados son las siguientes:

- 1- Para detectar la presencia de problemas que deben ser neutralizados o solucionados.
- 2- Estudiar o analizar el comportamiento de elección de los consumidores. Ejemplo: ver cuales son los niveles de respuesta de los compradores potenciales ante cambios en las variables.
- 3- Estudiar o investigar los factores del entorno que pueden influir sobre la actividad de la empresa.
- 4- Analizar las consecuencias de la actuación de la empresa sobre las variables de marketing.

La función de la investigación de mercados implica el diagnóstico de las necesidades de información, su búsqueda sistemática y objetiva mediante el diseño de métodos para su obtención, la recogida de datos, el análisis e interpretación de la información relevante para identificar y solucionar cualquier problema u oportunidad en el campo del Marketing y su distribución a tiempo entre los usuarios que toman decisiones.

1.3. METODOLOGÍA PARA REALIZAR UN ESTUDIO DE INVESTIGACIÓN DE MERCADOS.

I. PLANIFICACIÓN DE LA INVESTIGACIÓN

El primer paso o fase en la investigación de mercados es **especificar los objetivos de la investigación**. Esta fase consistiría en tratar de plantear el problema de toma de decisiones como una pregunta que ha de ser contestada al final de la investigación. La segunda fase sería **establecer las exigencias de información**. En esta segunda fase se suele elaborar un listado de la información relevante necesaria para cumplir o satisfacer los objetivos que se han establecido en la etapa anterior. La tercera fase es **determinar el valor estimado de la información a obtener**, y decidir si el coste que tiene la información es menor o no que el valor que ésta nos aporta.

II. DISEÑO DE LA INVESTIGACIÓN

La cuarta fase sería **definir el enfoque de la investigación**, es decir, a que fuentes de información se van a acudir: primaria o secundaria. La quinta fase es el **desarrollo específico del enfoque de investigación**. En esta quinta fase se realizan las siguientes tareas o actividades:

- a) Determinar o diseñar los procedimientos que nos permitirán obtener los datos necesarios para llevar a cabo la investigación. Estos procedimientos son, fundamentalmente, tres:
 - *Investigación exploratoria*: Se diseña para producir ideas, formular de forma más precisa y familiarizarse con los problemas, plantear hipótesis, identificar información relevante para detallar objetivos. Son estudios preliminares, cualitativos, flexibles, vinculados a la creatividad, al sentido común y a la intuición del investigador.
 - *Investigación descriptiva*: Se trataría de describir las características (demográficas, socioeconómicas, etc.)

de un problema, cuantificar los comportamientos, explicar actitudes. El procedimiento usual de la investigación descriptiva es el de realización de encuestas o paneles.

– *Investigación causal*: Su propósito es conseguir establecer relaciones causa–efecto. Aunque la investigación descriptiva puede ser utilizada en estos casos, el mejor método para analizar causalidades es la experimentación comercial.

b) Elaborar o desarrollar un Plan de Muestreo que contempla tres decisiones: a quién entrevistar, a cuántos y la forma de seleccionarlos. Se trataría de hacer una especie de ficha técnica.

c) Establecer los Métodos de Contacto, es decir, como se va a contactar con el objeto de estudio.

d) Seleccionar a los investigadores y elaborar un Presupuesto del Estudio de Mercado.

III. EJECUCIÓN Y PUESTA EN PRÁCTICA

La sexta fase es la **recogida y análisis de la información** que comprende actividades como la recogida de datos, el trabajo de campo, la codificación y tabulación de los datos, el procesamiento de los datos y el análisis e interpretación de la información obtenida. La séptima y última fase es **la interpretación de resultados y presentación de conclusiones**. Se debe presentar un resumen comprensible de los principales resultados derivados del estudio que se ha llevado a cabo.

1.4. FUENTES DE INFORMACIÓN Y ANÁLISIS DOCUMENTAL.

Todo estudio de investigación de mercados ha de fundamentarse en fuentes de información de las que pueden obtenerse los conocimientos o datos necesarios para la adopción de decisiones comerciales en el seno de una organización. Las fuentes de información pueden clasificarse de acuerdo con diferentes criterios siendo el más generalizado el que hace referencia a su disposición por la entidad que busca la información. De esta forma, al examinar el proceso de investigación conviene dividir la información en dos tipos: información primaria y secundaria.

La información secundaria es la que está disponible en el momento de realizar la investigación de mercados, y la información primaria es la que se realiza por primera vez para cubrir las necesidades específicas que se necesiten para realizar la investigación de mercados.

La información secundaria puede estar dentro de la empresa y se llamaría entonces datos internos, o puede estar fuera, en cuyo caso se llamará datos externos. Las ventajas de la información secundaria interna son las siguientes: el coste de obtención es muy bajo desde el punto de vista monetario y es fácilmente disponible desde el punto de vista del tiempo. Los inconvenientes de la información secundaria interna serían que este tipo de información es una aportación bastante parcial, es decir, se trata de una visión muy parcial ya que sólo nos aporta datos sobre lo que sucede dentro de nuestra empresa. Sin embargo, no nos aporta datos sobre las empresas del entorno, es decir, sobre las empresas competidoras. En cuanto a la información secundaria externa (revistas, censos, catálogos), estos datos externos pueden haber sido elaborados tanto por entidades privadas como públicas. Las ventajas de la información secundaria externa son que su coste es reducido y puede servir de punto de referencia para el estudio que se va a realizar porque aporta ideas, metodología. Los inconvenientes son que esta información no tiene porque ajustarse perfectamente a las necesidades del estudio, otro inconveniente sería la fiabilidad de la fuente.

Si vemos que con estos datos no es suficiente para cubrir todas las necesidades para el estudio se pasará a recopilar información primaria.

SECUENCIA DE INFORMACIÓN PARA LA INVESTIGACIÓN DE MERCADOS

La información primaria puede ser de dos tipos: cualitativa y cuantitativa. La diferencia entre estas dos sería que el objetivo de la cualitativa es describir hechos y no persigue conseguir demostrar algo que sea representativo de la población objeto de estudio, es decir, su objetivo no es cuantificar. Algunos ejemplos de métodos cualitativos son: reuniones de grupo, entrevistas en profundidad, pseudocompra, técnicas proyectivas, tormenta de ideas, método Delphi, método Phillips. Por otra parte, algunos ejemplos de métodos cuantitativos son: encuestas, paneles, etc. La observación se puede utilizar en cualquiera de los dos métodos, cualitativos y cuantitativos, es decir, tanto para describir como para cuantificar.

A continuación, vamos a centrarnos en la información secundaria. Lo primero que habrá que hacer será realizar un análisis documental, también llamado investigación de despacho o gabinete. El análisis documental tiene como punto de partida la información que ya está disponible a la hora de realizar la investigación de mercados, es decir, la información secundaria, y esta técnica se suele utilizar conjuntamente con técnicas de obtención de datos primarios, es decir, se utiliza como complemento de otras técnicas. El análisis documental no suele permitir completar una investigación de mercados.

Las herramientas más útiles a la hora de realizar el análisis documental es la creatividad y el sentido común del individuo que realiza la investigación de despacho. Las áreas de estudio donde puede ser más útil o adecuado el análisis documental son las siguientes:

- Conocer o estudiar las características de un producto a través de catálogos, anuncios, folletos, revistas, prensa diaria, visitando escaparates, etc.
- Estudio o análisis del marco económico y legal aplicable a un determinado producto.
- Evaluar la estructura y el tamaño del mercado al que se va a dirigir un producto.
- Recogida de información sobre empresas del sector: su sistema de financiación, los canales de distribución que utilizan, su organigrama, etc.
- Selección de una muestra a través de listados de personas, empresas que nos proporcione la cámara de comercio, y a partir de esos listados se podrá determinar el tamaño de la muestra que se va a utilizar.

TEMA 2: MÉTODOS CUALITATIVOS DE INVESTIGACIÓN DE MERCADOS.

2.1. FINALIDAD DE LA INVESTIGACIÓN CUALITATIVA.

La investigación cualitativa tiene como objetivo facilitar o proporcionar información de grupos reducidos de personas que no son representativos de la población objeto de estudio. Por tanto, no se trata de realizar análisis numéricos que sean estadísticamente significativos, sino que la finalidad principal es conocer y comprender las actividades, las opiniones, los hábitos y motivaciones de las personas (consumidores, personas que influyen en el consumo, etc.). No se pretende conocer cuantas personas hacen cada cosa si no conocer el cómo y el qué de las cosas.

La investigación de mercados cualitativa, en definitiva, se orienta a describir los hechos, a comprender el mercado de referencia, etc. Las técnicas cualitativas más utilizadas son: reuniones de grupo, entrevista en profundidad, pseudocompra, técnicas proyectivas, etc.

2.2. REUNIONES DE GRUPO Y ENTREVISTA EN PROFUNDIDAD.

• REUNIONES DE GRUPO

Las reuniones de grupo es una técnica de investigación cualitativa en la que se reúne a varios individuos para que discutan libremente sobre un determinado tema. Generalmente, se suele reunir a un número de personas que suele oscilar entre ocho y quince. El investigador en esta técnica actúa como moderador de la reunión y su papel es clave ya que tiene que lograr la participación de todos los asistentes a la reunión. La duración media oscila entre una y dos horas, y el contenido de la reunión de grupo debería ser grabado, ya sea en una cinta de

audio o de video, para poder analizar dicho contenido posteriormente con calma. Normalmente, se suele realizar varias reuniones de grupo por estudio. Los pasos o fases que hay que seguir para la preparación de una reunión de grupo son los siguientes:

1ª- Fijar o establecer los objetivos. En esta etapa se trataría de definir con claridad y precisión el tema central que se va a tratar en la reunión de grupo.

2ª- Planificación de la reunión de grupo.

A) Decidir quienes van a integrar la reunión. Las variables más utilizadas como criterio de selección son las variables demográficas y las variables socioeconómicas (nivel de estudios, etc.). Pero hay otras variables que se pueden utilizar como criterio de selección, como son: los consumidores del producto, la actitud positiva o negativa, etc. En los mercados industriales a veces se suele utilizar como criterio de selección el cargo que tenga dentro de la empresa cada persona.

B) Habrà que decidir también el número de personas que va a haber en cada reunión, que suele oscilar entre 8 y 15, y también habrá que decidir el número de reuniones que se van a realizar por estudio. Estas decisiones van a depender del presupuesto y del plazo temporal que tenga el investigador para realizar el estudio. En último término, lo importante no es el número exacto de personas que van a formar el grupo sino que haya variedad y heterogeneidad entre los componentes del grupo para que pueda haber discusión entre ellos.

C) Habrà que elaborar un guión con los puntos a tratar en la reunión. Es conveniente elaborarlo para no apartarse del tema central.

3ª- Reclutamiento de los asistentes. El primer contacto se suele hacer por teléfono y se propone a esa persona que se llama por teléfono la participación en la reunión, si acepta se le envía una invitación formal donde se especifica la fecha, la hora y el lugar donde se va a celebrar la reunión.

4ª- Inicio de la reunión. Una vez que empieza la reunión uno de los aspectos más importantes de esa reunión va a ser la moderación porque el investigador tiene que ser muy sensible para lograr la participación de todos los asistentes pero a la vez tiene que ser firme.

- El moderador tiene que procurar que no haya protagonistas en la reunión.
- El moderador tiene que evitar un enfrentamiento personal entre los miembros de la reunión.
- También tiene que hacer ver a los participantes que todas las opiniones son válidas (no reírse ni despreciar alguna de las opiniones, etc.).
- El moderador tiene que procurar que la reunión no se aleje del tema central.
- El moderador no debe influir en las opiniones de los asistentes.

5ª- Análisis de los resultados de la reunión. La última fase o etapa es el análisis e interpretación de los resultados de la reunión. Se trataría de transcribir el contenido de la cinta y presentar un informe que agrupe los distintos temas tratados en las distintas reuniones de grupo. En ese informe tiene que aparecer recogidos las comprensiones y valoraciones personales del propio investigador.

♦ ENTREVISTA O ENCUESTA EN PROFUNDIDAD

En esta técnica el investigador interactúa con un solo individuo y le anima o invita a que se manifieste con toda libertad acerca de un producto o un determinado tema en concreto. Un aspecto fundamental de la entrevista en profundidad es conseguir o crear un ambiente relajado, agradable, no tenso para conseguir una respuesta lo más completa posible por parte del entrevistado.

El entrevistador tiene que dominar el tema para ser capaz de aclarar cualquier duda que se le presente al

entrevistado y tiene que ser capaz de retomar la entrevista cuando se desvíe del tema principal o central. Esta técnica, normalmente, constituye la primera aproximación a cualquier problema o tema a analizar, especialmente cuando no se tiene un conocimiento previo del mismo, que permita establecer las bases teóricas requeridas para la posterior aplicación a cualquier otra metodología o investigación, ya sea cuantitativa o cualitativa, para formular las hipótesis de trabajo. El número de entrevistas dependerá del presupuesto y del tiempo.

2.3. PSEUDOCOMPRA Y TÉCNICAS PROYECTIVAS.

◆ **PSEUDOCOMPRA**

La pseudocompra, también llamada cliente oculto o cliente fantasma, es una técnica en la que el investigador se presenta en una empresa como un cliente potencial y se comporta como un comprador normal, aunque en realidad está actuando de forma premeditada.

El objetivo de la pseudocompra es analizar como reacciona normalmente el vendedor de una empresa. El informe se suele realizar a la salida del establecimiento ya que es el momento en que la información está más fresca y en ese informe se refleja:

- La actitud del vendedor.
- Los argumentos de venta que ha utilizado el vendedor.
- Las marcas ofrecidas al cliente.
- Las soluciones dadas a los problemas planteados por el falso comprador.
- El aspecto interior y exterior del local, así como las características personales del vendedor y su apariencia.
- El movimiento de clientes en ese local.

Un rasgo importante de la pseudocompra es que no existe cuestionario ni guión, sino que el entrevistador tiene que estar altamente cualificado para saber que es lo realmente importante. En esta técnica no se graba nada en cinta, sino que sólo se utiliza para grabar la memoria. La pseudocompra es una técnica de interacción uno a uno.

◆ **TÉCNICAS PROYECTIVAS**

Las técnicas proyectivas son técnicas indirectas que están muy relacionadas con la psicología. Su objetivo principal es descubrir las motivaciones auténticas, también llamadas las razones ocultas, de los individuos presentándoles una serie de estímulos ambiguos que provoquen en ellos una respuesta espontánea. Dentro de estas técnicas, que son todas de aproximación indirecta, al individuo se le pide siempre que interprete la conducta de otros y nunca la suya propia. Existen varias modalidades de técnicas proyectivas, siendo alguna de las más conocidas y utilizadas en el campo del Marketing las siguientes: técnicas de asociación de palabras, tests de frases incompletas, tests de respuesta a imágenes.

1– Técnicas de asociación de palabras: Estas técnicas tienen como objetivo que el individuo responda lo primero que se le venga a la mente en relación con una determinada palabra. Algunas variantes de esta técnica son las siguientes:

- Solicitar al individuo entrevistado que cite *sinónimos o antónimos* de las palabras que se le ha leído.
- Otra técnica sería la *asociación libre* que consiste en presentar una lista de palabras y que el entrevistado agrupe aquellas palabras que tengan un significado, para él, similar dentro del conjunto de términos escogidos.
- En la *asociación controlada* se trata de agrupar palabras que estén relacionadas no entre si, sino con un término determinado,
- Las *asociaciones* en ocasiones hacen *referencia a objetos, empresas, marcas o personas* con resultados de

gran interés.

2- Tests de frases incompletas: Consiste en presentar al entrevistado una serie de frases sin acabar, o en las que faltan algunas palabras intermedias, y se pide al entrevistado que las complete. Algunas alternativas o variantes dentro de las frases incompletas son las siguientes:

- Escribir frases relacionadas con productos, marcas, personas, empresas nombrados por el investigador.
- Presentar al individuo el inicio de una historia y luego pedir al encuestado que la termine o desarrollar un relato de hechos y al final preguntar: con qué personaje de la historia se quedaría; qué hubiera hecho él si hubiera estado en una situación similar; cuál cree que ha sido la motivación de los personajes para actuar de la forma expuesta.
- Pedir al entrevistado que dibuje personajes, en ocasiones se le suele dar al entrevistado algún personaje, en una situación ambigua que revista interés para el investigador y solicitar que termine los diálogos de las caricaturas.

3- Tests de respuesta a imágenes: Esta técnica consiste en presentar al entrevistado fotografías, dibujos, videos, diapositivas, transparencias y dejar que éste exprese lo que le sugieren esas imágenes. Ejemplo: anuncios en televisión, envases y etiquetas de los productos, elección de los logotipos para una empresa.

2.4. OTRAS TÉCNICAS CUALITATIVAS DE INVESTIGACIÓN DE MERCADOS.

Otras técnicas cualitativas de investigación de mercados son las tres siguientes: Tormenta de ideas, Phillips 66, Método Delphi.

♦ **TORMENTA DE IDEAS**

La tormenta de ideas es un tipo particular de reunión de grupo cuyo único fin es crear ideas. La tormenta de ideas se distingue, también, de una reunión de grupo porque en este tipo de reuniones sólo pueden participar expertos, es decir, grandes conocedores del tema a tratar. Esta técnica suele dividirse en dos fases:

- Fase de generación de ideas.
- Fase de selección de ideas.

En la primera fase se trataría de producir el mayor número posible de ideas no haciendo ningún tipo de crítica, mientras que en la segunda fase se realizaría un análisis y selección de las ideas anteriores. En esta fase, si se suelen poner a las ideas posibles pegas, comentarios para quedarnos con las ideas más relevantes.

♦ **PHILLIPS 66**

La principal ventaja de esta técnica es que permite agrupar o reunir a un número de personas comprendido entre 50 y 100, es decir, permite trabajar con colectivos de tamaño medio. El procedimiento que se sigue a la hora de aplicar esta técnica distingue las siguientes fases:

- Inicialmente se reúne a todos los participantes en una sala y se les propone un determinado tema a debatir. A continuación, se les invita a separarse en pequeñas comisiones, de 6 a 10 personas, para discutir sobre ese tema en salas aisladas.
- En cada comisión se elige un coordinador y los participantes nombran un portavoz. En esa comisión, los participantes debaten sobre el tema y extraen las principales conclusiones de su discusión.
- En esta tercera fase, se vuelve a reunir de nuevo a todos los participantes y los portavoces de cada comisión exponen los resultados obtenidos, para que posteriormente se debatan sobre estas conclusiones. Hay que dejar bien claro que lo que se pone en común son las conclusiones y no el tema a tratar, esto se realiza para no tener que partir de cero.

♦ MÉTODO DELPHI

En el método Delphi se trabaja con un grupo de expertos independientes que no tengan nada que ver con la empresa que encarga la investigación. En este caso no va a ver comunicación oral, sino que la comunicación va a ser por escrito. Las fases o etapas del método Delphi son las siguientes:

- En la primera fase se define el problema a estudiar.
- En la segunda fase se realizaría la selección de los expertos, en un número de 20 a 25, siendo todos especialistas en el tema a tratar.
- En la tercera fase se les envía a los expertos una carta con la descripción del problema objeto de estudio y con una pregunta muy general acerca del estudio.
- Una vez que los expertos seleccionados han respondido a esa pregunta se extraen los principales aspectos mencionados por ellos, y se elabora a partir de estos aspectos mencionados un cuestionario.
- La quinta fase consistiría en el envío de los cuestionarios a los expertos en los que se le piden estimar algunos aspectos concretos de tipo cuantitativo.
- Una vez que los expertos responden al cuestionario se trata la información recibida y se calcula medidas de tendencia, como la mediana, y se calcula también la dispersión de los datos. Todos estos datos se calculan para ver hasta que punto hay consenso entre los expertos seleccionados.
- En la séptima fase se envía los resultados obtenidos a los expertos y se les pide que revisen su posición. Si el experto dice que permanece en su posición tiene que explicar o argumentar las razones que le mueven a dar esa opinión, y lo mismo para el caso en el que diga que no mantiene su primera respuesta.
- Repetir las dos últimas fases para intentar que la dispersión de la información sea lo más pequeña posible. El objetivo es alcanzar una evaluación consensuada sobre el tema objeto de estudio.

TEMA 3: MÉTODOS CUANTITATIVOS DE INVESTIGACIÓN DE MERCADOS.

3.1. FINALIDAD DE LA INVESTIGACIÓN CUANTITATIVA.

El objetivo de la investigación cuantitativa es recoger información primaria para estudiar cuantos elementos de una población poseen una determinada característica, cuantos son consumidores, cuantos establecimientos siguen unas determinadas estrategias, etc.

Estas técnicas van a trabajar con grupos relativamente grandes de elementos, y persiguen en todo momento extraer datos que sean representativos estadísticamente de la población objeto de estudio. Hay dos técnicas cuantitativas dentro de este grupo que son fundamentales: encuestas y paneles.

3.2. CONCEPCIÓN Y ORGANIZACIÓN GENERAL DE LA ENCUESTA.

Las encuestas son una de las herramientas cuantitativas más utilizadas a la hora de obtener información primaria, pero ello no significa que sea adecuado recurrir siempre a ellas. Antes de hacer una encuesta hay que asegurarse de que realmente esa encuesta es necesaria, habrá que comprobar también que este totalmente agotada la información secundaria sobre el tema que se va a estudiar.

También, se ha de verificar o demostrar que la información que se necesita no se puede obtener a través de técnicas cualitativas, ya que éstas son más baratas que las cuantitativas. Habrá también que estudiar o analizar el coste de la encuesta y los recursos disponibles para realizarla. También habrá que demostrar que el tema es de interés para la empresa y que ciertamente existe información disponible, accesible. Una vez que vemos que es necesaria la encuesta, el segundo paso es diseñar el plan de sondeo, es decir, desarrollar la encuesta. Esto se realizará a través de las siguientes fases o etapas:

- Hay que definir la población a estudiar o universo de estudio.
- Determinar la unidad muestral que puede ser individuos, familias, empresas, etc.

- Delimitar o especificar el marco muestral, es decir, habría que escoger el censo o listado, de las unidades muestrales disponibles, de personas que van a formar parte de la población o universo de estudio, y determinar el alcance geográfico del estudio y el tiempo o plazo temporal al que se refiere el comportamiento u opiniones analizadas.
- Seleccionar el método o procedimiento de muestreo, pudiendo utilizar métodos probabilísticos y no probabilísticos.
- Determinar el tamaño de la muestra que tiene que ser representativa. Su tamaño va a venir condicionado por el procedimiento de muestreo que hayamos escogido.
- Diseñar y preparar el cuestionario. En esta fase habrá que determinar el contenido y formato del cuestionario.
- Decidir el tipo de encuesta que vamos a realizar. Hay tres tipos de encuestas:

- Encuestas por correo.
- Encuestas por teléfono.
- Encuestas personales.

• **TIPOS DE ENCUESTAS**

La elección del tipo de encuesta va a depender de los recursos disponibles, como el presupuesto siendo las más caras las personales, de la longitud de la encuesta, de la complejidad del tema, de la dispersión geográfica de los individuos, etc.

• *ENCUESTA POR CORREO*

A las encuestas por correo también se las llama encuestas postales. En las encuestas por correo se envía un cuestionario al encuestado y junto a él una carta explicativa de la utilidad e importancia de la encuesta, así como de la forma de cubrir el cuestionario y también será indicado el agradecimiento por participar en la encuesta. Generalmente, se suele incluir un sobre franqueado para la respuesta con la dirección de la empresa.

En aquellos casos en los que el ámbito objeto de estudio lo permita, como por ejemplo en los mercados industriales, conviene que las cartas vayan dirigidas nominalmente. Es aconsejable también que los cuestionarios no sean excesivamente largos y que la carta de presentación sea de uno o dos folios como máximo. Las ventajas de una encuesta por correo son:

- El coste de una encuesta por correo es bajo.
- Esta modalidad de encuestas permite llegar a un número grande de personas.

Mientras que los inconvenientes de una encuesta por correo son los cuatro siguientes:

- Baja tasa de respuesta ya que hay pocas personas que responden. Se suele considerar como aceptable un nivel entre el 15% y el 20% de respuestas.
- Es imposible controlar quien responde la encuesta ya que pueden contestarla personas que no son realmente interesantes para nuestro estudio.
- En las encuestas por correo es imposible controlar la veracidad de todas las respuestas.
- Tampoco se puede controlar en las encuestas por correo la influencia de terceras personas.

Este método es un procedimiento lento ya que la gente tarda en contestar. Un problema adicional es que estas encuestas no se pueden enviar a todo tipo de personas, ya que hay que tener un nivel cultural mínimo.

• *ENCUESTAS TELEFÓNICAS*

Las encuestas telefónicas se suelen utilizar en los estudios de audiencia de televisión y radio, y también en los de intención de voto. Las ventajas de las encuestas telefónicas son las siguientes:

- La principal ventaja es la rapidez en la obtención de información.
- Otra ventaja es que la comunicación es directa y el rechazo es menor que en las encuestas por correo.
- Son más caras que las encuestas por correo pero más baratas que las encuestas personales, es decir, son más económicas comparadas con las encuestas personales.

Por otra parte, los principales inconvenientes de las encuestas telefónicas son los siguientes:

- No se puede controlar la veracidad de las respuestas, y tampoco se puede controlar quien responde y la influencia de terceros.
- Otro inconveniente es la desconfianza del encuestado.
- Este tipo de encuestas sólo son útiles con cuestionarios muy breves.
- No nos permite utilizar material auxiliar o de apoyo como fotografías, tarjetas, dibujos, etc.

• *ENCUESTAS PERSONALES*

Las encuestas personales son el método más directo de obtención de información primaria cuantitativa y entre las principales ventajas que tienen podemos señalar las siguientes:

- Fiabilidad de la información porque controlamos quien responde, como responde y podemos evitar la influencia de terceras personas.
- Son las más flexibles ya que permiten resolver dudas o hacer aclaraciones al encuestado sobre la marcha.
- No generan desconfianza en el encuestado.
- Permite la utilización de materiales auxiliares o de apoyo como material gráfico (fotografías, dibujos, etc.).
- La tasa de respuesta suele ser mayor que la de las encuestas por correo y telefónicas.
- La información de cada una de las encuestas se obtiene con bastante rapidez.

Por otra parte, los principales inconvenientes de las encuestas personales son los tres siguientes:

- El coste de este tipo de encuestas es muy alto.
- Error o sesgos introducidos por el propio entrevistador, por eso es fundamental para que la información obtenida sea válida tener encuestadores bien formados.
- Limitaciones en cuanto a los temas a tratar.

• *ENCUESTAS ELECTRÓNICAS*

Se considera que este tipo de encuesta pueden llegar a sustituir a las encuestas por correo y a las encuestas por teléfono. Las ventajas de este tipo son las siguientes:

- Entre las ventajas la rapidez es una de las más importantes no sólo en el envío sino también en la consecución de la información.
- Este tipo de encuesta es relativamente barata.
- El encuestador está claramente identificado con lo que se reduce la desconfianza por parte del entrevistado.
- Este tipo de encuesta permite utilizar material auxiliar o de apoyo.
- Este tipo de encuesta es bastante directa.

Otra ventaja es que si se diseña bien el cuestionario y lo responden se puede ir creando una serie de datos a medida que se va recibiendo la información. Entre los inconvenientes de este tipo de encuesta se pueden señalar los cinco siguientes:

- Hay un porcentaje de la población que no tiene ordenador conectado a la red.
- La muestra es voluntaria.
- No se puede controlar bien la veracidad de las repuestas ni la influencia de terceras personas.
- No se puede garantizar el anonimato de quien es entrevistado.
- La longitud del cuestionario es un inconveniente porque no puede ser muy largo.

3.3. DETERMINACIÓN DEL TAMAÑO DE LA MUESTRA Y MÉTODOS DE MUESTREO.

• **MÉTODOS DE MUESTREO**

♦ **MUESTRA**

Una muestra es un conjunto de elementos de una población o de un universo del que se quiere obtener o extraer información. Se trabaja con muestras para realizar una investigación de mercado por el ahorro tan grande que suponen en tiempo y en dinero. Ahorro que es mucho mayor que la pérdida de precisión en la información conseguida. Para que la información sea válida esa muestra tiene que ser representativa de la población objeto de estudio.

• **PROCEDIMIENTOS**

Hay dos procedimientos o tipos de muestreo básicos: probabilísticos y no probabilísticos. La diferencia entre los dos sería la siguiente: Los procedimientos no probabilísticos son aquellos en los que no conocemos la probabilidad de que un elemento de la población pase a formar parte de la muestra ya que la selección de los elementos muestrales dependen en gran medida del criterio o juicio del investigador. La muestra, en este caso, se selecciona mediante procedimientos no aleatorios. Hay tres tipos: muestreo de conveniencia, muestreo discrecional, muestreo por cuotas.

Los métodos probabilísticos son procedimientos de muestreo en los que todos los elementos de la población tienen la misma probabilidad de pasar a formar parte de la muestra y además nosotros conocemos esa probabilidad. La selección de los elementos muestrales se realiza mediante procedimientos aleatorios. Estos métodos no son mejores que los anteriores sino que simplemente nos permiten calcular el error muestral que se está cometiendo. Dentro de los métodos probabilísticos podemos señalar los siguientes: muestreo aleatorio simple, muestreo aleatorio sistemático, muestreo estratificado, muestreo por conglomerados.

• *PROCEDIMIENTOS NO PROBABILÍSTICOS*

– **Muestreo de conveniencia.**

El investigador decide que elementos de la población pasan a formar parte de la muestra en función de la disponibilidad de los mismos (proximidad con el investigador, amistad, etc.). Este tipo de muestreo sólo es adecuado en investigaciones de tipo exploratorio, es decir, para obtener una idea general sobre el tema objeto de estudio.

– **Muestreo discrecional.**

En el muestreo discrecional la selección de los elementos de la muestra la realiza un experto que indica al investigador que elementos de la población son los que más pueden contribuir al estudio. Este muestreo es adecuado cuando existen líderes de opinión dentro de la población objeto de estudio, sabemos quienes son esos líderes y no queremos que se nos escapen por utilizar un método totalmente aleatorio o de conveniencia.

– **Muestreo por cuotas.**

Lo primero que tiene que hacer el investigador en este caso es estudiar las características del universo de estudio. A continuación, se determina el tamaño de la muestra y se deja que el investigador elija libremente los elementos de la muestra respetando siempre ciertas cuotas por edad, sexo, nivel de renta, profesión, zona geográfica en la que vive, etc.

• *PROCEDIMIENTOS PROBABILÍSTICOS*

– **Muestreo aleatorio simple.**

En este tipo de muestreo se asigna un número a cada elemento de la población y se eligen aleatoriamente tantos elementos como indique el tamaño de la muestra. En principio, este es el método más perfecto a la hora de obtener una muestra representativa pero tiene algunos problemas o inconvenientes.

- El método es impracticable con muestras muy grandes.
- Su aplicación requiere siempre de un listado completo que incluya todos los elementos de la población.
- Pueden existir líderes de opinión o elementos de la población más representativos que otros pero que con este procedimiento queden fuera de la muestra. Para solucionar este problema existe un caso particular que es el denominado muestreo aleatorio opinático.

Ejemplo: Considerar una población de tamaño (N_p) igual a 250.000 personas y se elige una muestra (n) de 500 personas. En esa población existe un líder, por tanto este se escogería directamente para la muestra y los 499 restantes se eligen mediante un muestreo aleatorio simple entre las 249.999 personas restantes de la población.

– **Muestreo aleatorio sistemático.**

El investigador calcula un coeficiente de elevación o avance que es igual al cociente entre el número total de elementos de la población y el número total de elementos de la muestra. De esta manera, el investigador determina cada cuantos elementos de la población hay que elegir uno para que componga la muestra.

Ejemplo:

N_p n° elementos de la población 1.000

$= = 10$

n n° elementos de la muestra 100

De cada 10 elementos de la población se escoge 1 para que pase a formar parte de la muestra, es decir, 10 es el intervalo entre dos elementos muestrales. El problema de este método es que existan comportamientos cíclicos en la población que distorsionen la representatividad de la muestra.

– **Muestreo estratificado.**

Este tipo de muestreo consiste en dividir los elementos de la población en distintos segmentos o estratos formados por elementos que sean lo más homogéneos posibles entre si. Pero, eso si, tiene que haber heterogeneidad entre los distintos segmentos. Una vez que se han establecido esos grupos, se reparte la muestra entre los distintos estratos siguiendo un determinado criterio de afijación. Vamos a suponer que se ha de entrevistar a 100 personas y la población se encuentra dividida en cuatro estratos, siendo el número de elementos de la población igual a 1.000.000. Los posibles criterios de afijación son los tres siguientes:

- Afijación simple: supone repartir la muestra total en partes iguales para cada estrato, independientemente

del tamaño de la población, es decir, este criterio determina que el número de encuestas que se realizarán será igual para cada estrato. Este criterio es muy sencillo pero poco recomendable. Ejemplo: 25 elementos por estrato.

– Afijación proporcional: este criterio supone la división de la muestra en partes proporcionales a la población de cada estrato, es decir, el reparto de las encuestas se hace teniendo en cuenta el tamaño del estrato.

Ejemplo:

$$N_i = 100.000$$

=

$$N_p = 1.000.000$$

$$400.000$$

=

$$1.000.000$$

– Afijación óptima: el reparto de la muestra se hace no solo atendiendo al tamaño del estrato sino también a la dispersión de los datos dentro de los estratos, medida ésta a través de la desviación típica.

$$n_i = N_i \cdot i$$

=

$$n = N \cdot i$$

n_i = número de elementos del segmento i que van a pertenecer a la muestra.

n = tamaño de la muestra.

N_i = tamaño del segmento i .

i = desviación típica de las respuestas a obtener en el estrato i .

– **Muestreo por conglomerados.**

Consiste en extraer dentro de una población grupos de elementos al azar y después dentro de esos grupos se eligen al azar grupos más pequeños, y a su vez dentro de este grupo se elige al azar grupos más pequeños y así sucesivamente hasta que el investigador decida parar. Al muestreo por conglomerados también se le llama muestreo por áreas.

Un tipo particular de procedimiento probabilístico es el muestreo por rutas aleatorias o RANDOM ROUTE. Este tipo de procedimiento se utiliza para encuestas a domicilio no personalizadas donde se proporciona al entrevistador un sistema para la elección del individuo o entidad a entrevistar. En este procedimiento se utilizan tres elementos materiales fundamentales.

- La *hoja de ruta*: al encuestador o entrevistador se le entrega un mapa con un punto de partida en la que empieza su trabajo y la ruta que debe seguir para localizar a los encuestados, por ejemplo, caminar por

calles en zig-zag y realizar encuestas en todos los portales o edificios pares.

- La *tablilla de selección*: es un cuadro o tabla de números aleatorios de doble entrada, donde por filas se indica el número de plantas del edificio o de viviendas en cada planta y por columnas el número de orden de la encuesta a realizar.
- Las *normas de sustitución*: son criterios a seguir cuando la persona selecciona no se encuentra en el domicilio o se niega a contestar, hay plantas con oficinas y otras con viviendas.

• **TAMAÑO DE LA MUESTRA**

El tamaño de la muestra está condicionado por el método de muestreo utilizado, pero en este caso para simplificar el cálculo del tamaño de la muestra vamos a determinar dicho tamaño para un muestreo aleatorio simple ya que, generalmente, ese muestreo aleatorio simple exige muestras superiores (para un mismo grado de fiabilidad o nivel de confianza) al resto de procedimientos.

La expresión a utilizar para calcular el tamaño de la muestra es diferente según sea la población finita o infinita. Se considera que una población es infinita cuando es mayor o igual a 100.000 individuos. Mientras que se considera que una población es finita cuando la población es menor de 100.000 individuos. Para el caso en que la población es infinita la expresión a utilizar es la siguiente:

$$Z^2 \cdot P \cdot (1 - P)$$

n =

K²

Por el contrario, para el caso en que la población es finita la expresión a utilizar para determinar el tamaño de la muestra es la siguiente:

$$Z^2 \cdot N_p \cdot P \cdot (1 - P)$$

n =

$$(N_p - 1) \cdot K^2 + Z^2 \cdot P \cdot (1 - P)$$

donde,

- n: Tamaño de la muestra.
- N_p: Tamaño de la población objeto de estudio.
- Z: Número de unidades de desviación típica en una distribución normal que va a producir el grado de confianza deseado. Cuando el grado de confianza con que trabajamos es del 95% Z va a ser igual a 2 y cuando el grado de confianza es del 99% Z va a ser igual a 3.
- P: Sería la proporción de individuos de la población que cumplen una determinada característica, por ejemplo, proporción de consumidores del bolígrafo X, etc.
- 1 - P: Sería la proporción de individuos de la población que no cumplen una determinada característica.

Podemos conocer P por estudios anteriores o también por pruebas piloto. Si no tenemos nada de nada podemos ponernos en el peor de los casos y considerar que la probabilidad de que cumpla la característica es igual a la probabilidad de que no la cumpla, es decir, P sería igual a 0,5.

- K: es el error muestral o máxima diferencia que estamos dispuestos a admitir entre la proporción de la población y la proporción muestral para el nivel de confianza que se ha fijado. Este error es debido a

trabajar con una muestra y no con el total de la población, es decir, es un error que se debe al método de muestreo. Este es el error muestral pero a la hora de realizar una encuesta puede haber muchos otros tipos de errores. Por ejemplo, otros tipos de errores que no son muestrales son los siguientes:

- Definición defectuosa del universo de estudio.
- Definición incorrecta del marco muestral.
- Mal diseño del cuestionario.
- Errores debidos al entrevistador.
- Errores de no respuesta.
- Errores de medición, que se deben al uso de escalas no adecuadas para un determinado tema.
- Errores en el tratamiento de la información.

Unos mayores niveles de confianza exigen unos mayores niveles de muestra, así como cuanto menores errores se quieran cometer.

3.4. DISEÑO DE CUESTIONARIO, TRABAJO DE CAMPO, PLAN DE ESCRUTINIO Y ANÁLISIS DE RESULTADOS.

◇ **DISEÑO DEL CUESTIONARIO**

Un cuestionario es el instrumento material de recogida de la información y, además, es un medio de garantizar que se van a plantear las mismas preguntas a todos los encuestados. A pesar de lo importante que es para que la información recogida para una encuesta sea válida hay que decir que para elaborar un cuestionario no existen recetas universales o comúnmente aceptadas. Pero si se pueden señalar una serie de orientaciones generales para la elaboración del cuestionario. Concretamente, antes del diseño se suele recomendar tres pasos para garantizar que el cuestionario sea adecuado:

- Tener muy claro cual es el problema a investigar.
- Formular de forma precisa y clara las hipótesis de trabajo.
- Especificar adecuadamente las variables y las escalas de medida que se van a utilizar.

Una vez que se comienza a diseñar el cuestionario hay otra serie de normas básicas que es conveniente seguir para el diseño del cuestionario:

- Un cuestionario tiene que ser lo más breve posible.
- Tiene que ser comprensible PARA TODOS, de forma que las preguntas tienen que ser sencillas y claras y, además, las preguntas tienen que tener o mantener un orden lógico.

En un cuestionario tienen que aparecer las siguientes cosas y, además, en el siguiente orden:

- Poner los datos de identificación del encuestador o investigador.
- Petición de colaboración en el estudio.
- Preguntas de introducción, por ejemplo, realizar preguntas de opinión.
- Preguntas que sean sencillas e interesantes y que estén relacionadas con el tema que se va a tratar.
- Preguntas que constituyan el núcleo de la investigación. Dentro del núcleo de la investigación aquellas preguntas que sean más difíciles o las más difícil de configurar se dejarán para el final del cuestionario.
- Variables de clasificación, que pueden ser de tipo demográfico, socioeconómicas como, por ejemplo, el nivel de estudios, el nivel de ingresos, la edad, etc.
- El agradecimiento por haber participado en el estudio y si se puede unas líneas para posibles observaciones.

Los tipos de preguntas que se pueden realizar en un cuestionario, según la clasificación más tradicional, son las siguientes:

a) Preguntas abiertas: Son aquellas preguntas que dejan opinar libremente al encuestado y son, por ello, adecuadas para investigaciones exploratorias o situaciones en las que no se conoce bien las posibles respuestas. La ventaja de este tipo de pregunta es que va a haber una alta gama de respuestas, es decir, se va a tener una información muy rica en contenido. Por el contrario, el inconveniente es que los datos obtenidos con las preguntas abiertas son difíciles de tratar estadísticamente. Otro inconveniente es que en este tipo de preguntas es más fácil que el encuestador influya en los encuestados, es decir, que introduzca sesgos.

b) Preguntas cerradas: Son aquellas preguntas que van acompañadas de una relación exhaustiva de las posibles respuestas, luego este tipo de preguntas son adecuadas cuando se ha realizado un estudio previo o cuando se ha hecho un pretest de la encuesta. Las principales ventajas de las preguntas cerradas son que exigen un menor esfuerzo por parte del encuestado y el tratamiento estadístico de los datos obtenidos es sencillo. El inconveniente es que las preguntas cerradas son mucho más difíciles de elaborar.

c) Preguntas semiabiertas o semicerradas: Son aquellas preguntas en las que se recoge una lista de posibles respuestas y se deja una opción abierta, que sería la opción otros/as. Este tipo de preguntas tendría una combinación de las ventajas e inconvenientes de las dos anteriores tipos.

Otros tipos de preguntas que se podrían realizar en un cuestionario serían las siguientes

1- Preguntas de introducción: Son preguntas muy sencillas que solicitan la opinión del encuestado sobre un tema y que tratan de disponerlo favorablemente para que conteste a las preguntas. Es decir, este tipo de preguntas se utilizan para crear un clima de confianza.

2- Preguntas en batería: Son preguntas que se realizan de manera concatenada, es decir, unas detrás de otras, para obtener mucha información sobre un determinado tema.

3- Preguntas filtro: Son aquellas preguntas que permiten seleccionar al entrevistado según se manifieste a favor o en contra de un determinado tema, problema o situación.

4- Preguntas de relleno o de cambio de tema: Son preguntas muy sencillas cuyo objetivo es que el encuestado se relaje después de haber respondido a un núcleo importante de preguntas dentro de la investigación.

5- Preguntas de tarjeta: Son preguntas cerradas en las que se le entrega al encuestado las posibles respuestas por escrito.

6- Preguntas de control: Son preguntas que tienen como objetivo verificar la veracidad de las respuestas.

Los errores más comunes a la hora de redactar las preguntas de un cuestionario son los siguientes:

- Preguntas cargadas o embarazosas, es decir, redactar preguntas que sugieren una respuesta concreta.
- Hacer preguntas y/o respuestas dobles: Consiste o bien en hacer dos o más preguntas a las que se solicita una sola respuesta, o bien proponer dos o más alternativas dentro de una misma respuesta.
- Hacer preguntas ambiguas: Consiste en formular preguntas que tienen dos o más interpretaciones, es decir, que son poco concretas. Por ejemplo, utilizar palabras en las preguntas como habitualmente, frecuentemente.
- Utilizar vocabulario inadecuado: Realizar preguntas que utilizan términos poco corrientes o muy técnicos que resultan desconocidos para el entrevistado.
- Solicitar al encuestado que haga cálculos complicados.
- Hacer preguntas tendenciosas o con suposiciones implícitas.

◇ TRABAJO DE CAMPO

El trabajo de campo no es más que la ejecución material de las encuestas. El encuestador en esta parte de la investigación de mercados va a tener un papel clave o fundamental. Por ello, hay que prestar especial atención en la selección de los encuestadores, en su formación, en las instrucciones que se les da al encuestador y en el control o seguimiento de su actividad. En cuanto a la selección se pide a los encuestadores:

- ◊ Dotes de observación.
- ◊ Buena memoria para recoger la información vinculada no sólo a las preguntas del cuestionario sino también a otros detalles, como la actitud del encuestado, la sinceridad con que responde, etc.
- ◊ Se valora mucho la seriedad, la responsabilidad, la honestidad, etc.
- ◊ Se valora también la capacidad de relacionarse con las personas, es decir, que el encuestador sea un buen relaciones públicas.

En cuanto a la formación, va a estar vinculada a la selección de los encuestadores. Se va a pedir a los encuestadores que tengan un nivel intelectual mínimo. En cuanto a las instrucciones, son importantes para que todos los entrevistadores se comporten de manera homogénea. Estas instrucciones tratan sobre:

- ◊ Como interpretar el cuestionario y las respuestas del encuestado.
- ◊ Como recoger la información.
- ◊ Como elegir a los entrevistados.

Se recomienda siempre hacer un ensayo general o pretest para que los encuestadores vayan cogiendo práctica de forma que tengan tablas a la hora de hacer las preguntas y, por otra parte, para ver si esas preguntas son comprensibles por los encuestados. En cuanto al control, al final del cuestionario se suele pedir al encuestado el nombre de pila y un número de teléfono o dirección para contactar con él. Esto se pide para comprobar, posteriormente, que efectivamente se ha realizado la encuesta a esa persona.

• **PLAN DE ESCRUTINIO**

Una vez hecho el trabajo de campo pasamos a lo que se denomina plan de escrutinio, que va a consistir en hacer lo siguiente:

- Se va a determinar la nomenclatura de las variables que se van a utilizar en el estudio.
- Se van a codificar las respuestas, es decir, asignar números, códigos o claves a las respuestas de los encuestados.
- Tabular los datos, es decir, introducir los datos en tablas en el ordenador.

Una forma de comprobar que se han introducido bien todos los datos es pedir la frecuencia, para limpiar la base de datos.

• **ANÁLISIS DE RESULTADOS**

En esta etapa de análisis de resultados se realizarán las siguientes actividades:

- Resumir toda la información obtenida de la muestra con que hemos trabajado.
- Cálculos de estadística descriptiva (medidas de posición y de dispersión).
- Desarrollar inferencias estadísticas y contrastes de hipótesis que orientan en la tarea de extrapolar los datos de la encuesta a la población.

En esta etapa es muy importante extraer conclusiones relevantes después de analizar los datos.

3.5. ESTUDIOS DE MERCADO MEDIANTE PANELES Y OBSERVACIÓN.

• **PANELES**

Los paneles son un medio de obtener información cuantitativa, al igual que las encuestas, que se caracterizan porque la información se recoge de una manera periódica para hacer, fundamentalmente, estudios sobre comportamientos históricos. Los paneles más utilizados son el de consumidores y el de detallistas, y dentro del panel de consumidores hay un tipo particular, muy importante, que es el panel de audiencias.

– **Panel de consumidores.**

En el panel de consumidores la unidad o elemento muestral es el hogar, familia o individuo. Es decir, se pueden hacer preguntas sobre el consumo familiar o sobre el consumo de un individuo. En España, la empresa con mayor reconocimiento en el desarrollo de paneles de consumidores es DYMPANEL, que cuenta con varios paneles:

- Un Panel de Hogares con un tamaño aproximado de 4.000 hogares.
- Baby Panel, sobre productos consumidos por bebés.
- Paneles de Gastos Relacionados con el Automóvil, sobre consumos realizados en Estaciones de Servicios.
- Panel Individual de Compras, que es un panel formado sólo por mujeres con un tamaño aproximado de 5.500 mujeres.

De estos cuatro paneles, los más demandados son el Panel de Hogares y el Panel Individual de Compras. Los sectores que estudia DYMPANEL son: alimentación y bebidas, droguería y hogar, perfumería y cosmética, textil, compras realizadas para el uso y consumo del bebé, compras llevadas a cabo en Estaciones de Servicio (carburantes, lubricantes). La información sobre productos para bebés se utiliza en el Baby Panel y la información sobre las compras llevadas a cabo en Estaciones de Servicio se utiliza en el Panel del Automóvil. Todos los demás sectores se utilizan para el Panel de Hogares y para el Panel Individual de Compras.

El ámbito objeto de estudio es todo el territorio nacional dividido en áreas geográficas, que coinciden exactamente con la división empleada por la empresa NIELSEN, que son seis u ocho en el caso de que se considere separadamente Madrid y Barcelona.

La empresa DYMPANEL tiene panelistas en todos los municipios con más de 30.000 habitantes. Estos panelistas se eligen mediante criterios socioeconómicos y demográficos. Una parte de la muestra rota aleatoriamente y siempre existe una reserva de panelistas. Al cabo de un tiempo se pregunta a los panelistas si quieren seguir formando parte del panel. La colaboración de los panelistas supone otorgar compensaciones, bien de tipo económico o bien un regalo personal.

Las herramientas para recoger los datos pueden ser: cuestionarios, libretas o diarios de gastos, lector de códigos de barras. Los datos que más se solicitan a los panelistas son: marca de los productos consumidos, precios, tipos de envases, número de unidades adquiridas, promociones existentes, lugar, fecha, hora de la compra, etc. La información que se obtiene con todos estos datos sirve para calcular el porcentaje de hogares compradores de un producto, la cantidad comprada por hogar, el gasto por hogar, precio medio pagado por los consumidores, volumen y cuotas de mercado, participación de las marcas, distribución del consumo por zonas geográficas y canales de ventas, perfil del consumidor de cada producto. Existen también análisis más complejos, como distinguir entre usuarios fuertes, medios y débiles; nuevos compradores y repetidores; intercambio de marcas; segmentos de mercado; previsiones de ventas; relaciones causa-efecto.

– **Panel de audiencias.**

Son un tipo particular del panel de consumidores que lo que estudian son las audiencias de televisión. La unidad muestral son individuos que permanecen en el anonimato. La herramienta que se utiliza en los paneles de audiencias es el audímetro que es un aparato que se coloca al lado del televisor y que recoge señales

cuando la televisión está encendida. Actualmente, también existen los audímetros pasivos. La empresa más famosa a la hora de estudiar los niveles de audiencia en España es SOFRES. Esta empresa recoge los niveles de audiencia de cadenas de televisión nacionales, tanto públicas como privadas, autonómicas y locales. También recoge información sobre otros usos que se da a la televisión, por ejemplo, videos, juegos, teletexto, etc.

El tamaño de la muestra con la que trabaja SOFRES es de 8.448 individuos, aunque periódicamente se amplía puesto que crecen el número de cadenas de televisión. Estos 8.448 individuos suponen una representación de 2.500 hogares y unos 4.000 audímetros (uno por televisión) aproximadamente. La muestra está estratificada por regiones. Los paneles de audiencia recogen información sobre:

- ◊ El tiempo en que está encendida la televisión.
- ◊ Que canal se está viendo.
- ◊ El programa que se está viendo.
- ◊ El número de personas que están viendo la televisión en un momento dado.
- ◊ La calificación de los programas que se están viendo.

Esta información queda recogida en el audímetro y cada madrugada SOFRES contacta vía telefónica con todos los audímetros desde su ordenador central. La información que SOFRES ofrece a las cadenas de televisión es la siguiente:

- ◊ Chrono TV: ofrece información sobre emisiones y audiencias de la televisión.
- ◊ Telescope: es más completo y ofrece, por ejemplo, información sobre audiencias medias por cuartos de hora o incluso por minutos, el ranking de programas, el perfil de la audiencia, es decir, el tipo de público que ve cada programa, las audiencias de los anuncios, simulación y evaluación, diseño de programas de optimización.

– Panel de detallistas.

La unidad muestral va a ser los establecimientos minoristas, que son los que venden al consumidor final, clasificados según unas determinadas características y categorías de productos.

La empresa más conocida en la elaboración de paneles de detallistas es NIELSEN que trabaja con todo el territorio nacional dividido en seis zonas, que pueden ser ocho si se considera Madrid y Barcelona separadamente. Dentro de los paneles de detallistas el sector del que se suele pedir más información es el sector de la alimentación. Las zonas o elementos de la muestra se seleccionan de acuerdo con un muestreo aleatorio con afijación proporcional al censo de establecimientos de la zona. A su vez, dentro de cada zona, hay que tener en cuenta los tipos y las ventas.

Las herramientas utilizadas son las visitas a los minoristas. En una primera visita se hace un recuento de todas sus existencias (productos, marcas, formatos, etc.) y en una nueva visita se revisa y contabiliza, al menos, el stock existente y las compras realizadas, ya que de esta forma se pueden conocer las ventas. Las recompensas pueden ser o bien de tipo económico, o bien un regalo o bien una ayuda en la gestión del negocio, que suele ser lo más común.

En España se realizan una gran multitud de paneles, aunque los más importantes son los que se han tratado anteriormente, es decir, el panel de consumidores, de detallistas y de audiencias.

• OBSERVACIÓN

Con la observación se recoge información primaria, tanto cualitativa como cuantitativa, pero sin entrevistar a nadie. La observación es un método de campo que se suele utilizar como complemento de otros métodos de

recogida de información. Entre las principales ventajas de la observación es que no se influye sobre el individuo observado y gracias a los avances tecnológicos (cámaras, etc.) se pueden tener muchos detalles sobre el elemento observado. Los inconvenientes más importantes son los siguientes:

- A veces es muy difícil que el objeto de estudio no se percate de que esta siendo observado o investigado.
- El coste puede ser prohibitivo.
- La información puede ser poco profunda. Esta sería la mayor limitación que tiene, puesto que la observación no nos permite conocer sentimientos, creencias de los individuos investigados.

En cuanto a la clasificación de los métodos de observación podemos distinguir los cinco siguientes:

- Observación Natural frente a la observación Manipulada o Artificial. En la observación natural el investigador no modifica ningún elemento del entorno, mientras que en la manipulada si se altera el entorno.
- Según que los individuos que son investigados u observados son conscientes o no que están participando en una investigación. Si se conoce la presencia del investigador hay posibilidades de introducir un sesgo ya que los individuos pueden comportarse de forma distinta, aunque esto también puede ayudar a que el individuo colabore.
- Observación Estructurada frente a la Observación No Estructurada. En la observación estructurada el investigador conoce perfectamente la situación que va a investigar y los comportamientos que pueden darse de forma que solo anota ciertos datos que considera relevantes para el estudio. En la observación no estructurada el investigador no sabe bien lo que puede ocurrir y anota en este caso todo lo que observa.
- Observación Directa frente a la Observación Indirecta. La observación directa se hace en el mismo momento en que se está dando un comportamiento o se produce una acción, mientras que la observación indirecta estudia comportamientos pasados.
- La Observación realizada por Personas frente a la Observación realizada mediante Aparatos Mecánicos como las cámaras ocultas, etc.

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

4.1. INTRODUCCIÓN. CODIFICACIÓN Y TABULACIÓN DE DATOS EN EL ANÁLISIS DE LA INFORMACIÓN.

• INTRODUCCIÓN

En un proceso de investigación de mercados una vez que la información está recogida, el paso siguiente es adoptar una metodología concreta o específica de análisis de datos con el fin de garantizar que las decisiones que se toman, se realicen con una base que sea racional.

Adoptar una metodología de análisis de datos supone trabajar o apoyarse en la estadística. Dentro de la estadística se pueden distinguir dos partes bien diferenciadas: la estadística descriptiva y la inferencia estadística. La estadística descriptiva tiene como objetivo sintetizar la información contenida en los elementos de una muestra. Por lo tanto, la estadística descriptiva sólo nos proporciona lo que se denomina medidas resumen como por ejemplo: la media, la moda, la mediana, la desviación típica siendo estas medidas de posición y dispersión. Es muy importante dejar claro que la información sólo es válida para la muestra y no para toda la población. El objetivo de la inferencia estadística es generalizar o inferir los datos de la muestra al conjunto de la población investigada o estudiada. Para poder aplicar la técnica de la inferencia estadística es imprescindible contar con una muestra que sea representativa de la población objeto de estudio.

• CODIFICACIÓN Y TABULACIÓN DE DATOS EN EL ANÁLISIS DE LA INFORMACIÓN

Para poder aplicar los métodos de la estadística descriptiva o de la inferencia estadística, lo primero que hay que hacer es asignar valores concretos a las opiniones o comportamientos de las personas, empresas, marcas que se estén estudiando, dejando bien claro que asignar valores no es más que realizar mediciones. A continuación, se van a ver distintos tipos de escalas o medidas, concretamente cuatro, que se pueden utilizar para asignar valores:

1– Escalas nominales: Este tipo de escala consiste en clasificar en categorías a las personas, empresas, marcas, es decir, a los elementos que se estén estudiando o que son objeto de medida. Posteriormente, se asigna un número a cada una de las categorías y se considera que todos los elementos a los que se ha asignado el mismo número son cualitativamente idénticos en la variable de medida. Con las escalas nominales sólo se puede saber si un elemento es igual o distinto a otros elementos de la muestra. Las categorías tienen que cumplir dos requisitos para que estén bien diseñadas:

- Que sean exhaustivas: esto significa que no puede haber ningún elemento que quede sin clasificar en alguna de las categorías que se han fijado.
- Que sean exclusivas: lo que significa que sólo se puede estar en una y sólo una de las categorías.

Las escalas nominales son, especialmente, útiles en el caso de las variables cualitativas, funcionando los números asignados únicamente como etiquetas identificativas. Algunos ejemplos de variables cualitativas a las que se podría aplicar esta escala son las siguientes: profesión, nivel de estudios, sexo, edad, ingresos, comunidad autónoma del que procede el encuestado, estado civil, si una persona conoce o no una determinada marca, etc.

2– Escalas ordinales: Este tipo de escala consiste en asignar a los elementos medidos un número que permita ordenarlos según la cantidad de variable que poseen desde el punto de vista del encuestado. Las escalas ordinales son útiles principalmente para variables cualitativas y aquí los números permiten afirmar si la cantidad de variable que posee un elemento es mayor o menor que la de otro, pero no dice cuanto mayor o cuanto menor, además de saber si un elemento es igual o distinto a otros elementos de la muestra.

Un ejemplo típico es el de indicar el orden de preferencia de las marcas, de unos productos, etc. Con los números se va a poder decir que una marca es más o menos preferible a otra, pero no cuanto más o menos preferida. Otro ejemplo son los rankings de preferencia.

3– Escalas de intervalo: Las escalas de intervalo consisten en definir una unidad de medida y después asignar a cada elemento medido un número indicativo de la cantidad de variable que posee según la unidad de medida establecida. En las escalas de intervalo no existe el cero absoluto de forma que la diferencia entre los elementos no son constantes. Esta escala se utiliza para medir el grado de satisfacción y el grado de acuerdo o desacuerdo con unas determinadas afirmaciones. Este tipo de escala es adecuada para variables cuantitativas y se utiliza en gran medida para la evaluación de actitudes. En esta escala los números permiten afirmar si un elemento es igual o distinto a otros elementos de la muestra, si la cantidad de variable que posee un elemento es mayor o menor que la de otro, pero en este caso si nos dice cuanto mayor o cuanto menor.

4– Escalas de razón: Las escalas de razón son como las escalas de intervalos pero añadiéndoles el cero absoluto, es decir, la ausencia total de cantidad de variable. Este tipo de escalas están indicadas para variables cuantitativas. En esta escala los números permiten afirmar si un elemento es igual o distinto a otros elementos de la muestra, si la cantidad de variable que posee un elemento es mayor o menor que la de otro, pudiendo decir cuanto mayor o cuanto menor. La diferencia entre los elementos medidos son constantes y, además, se puede afirmar si la cantidad de uno es el doble, el tripe, etc. que la cantidad de otro elemento. Algunos ejemplos de elementos a los que se podría aplicar escalas de razón son: los ingresos, la edad, los volúmenes de venta, etc.

La información sobre una misma variable se puede recoger mediante cualquiera de las cuatro escalas vistas,

aunque eso si, unas son más adecuadas para unos casos que para otros.

Escalas concretas para la medición de actitudes:

Nos detenemos en estos tipos de escalas porque la tarea de medición de actitudes es muy compleja ya que es algo que no se puede observar directamente. Las actitudes se suelen medir, generalmente, a través de las escalas de intervalo pero con algunas características particulares.

a) Escala de clasificación por sumas constantes: Esta escala consiste en asignar un número de puntos a distintos atributos según su importancia. Por ejemplo, 100 puntos a repartir entre cuatro atributos por orden de importancia. El inconveniente es que este tipo de escala es sólo útil cuando se trabaja con un número pequeño de atributos.

b) Escala diferencial-semántico: Esta escala consiste en utilizar palabras o frases bipolares, es decir, términos o extremos opuestos para que el entrevistado exprese sus sentimientos indicando su posición entre esos extremos.

Ejemplo. Los atributos a considerar son los tres siguientes: X, Y, Z.

-3 -2 -1 0 +1 +2 +3

X: Muy bien decorado x Muy mal decorado

Y: Soluciona mis problemas x No soluciona mis problemas

Z: Me inspira confianza x No me inspira confianza

Este tipo de escala permiten diseñar perfiles y realizar comparaciones entre los distintos elementos de la investigación, además de analizar los puntos débiles y fuertes de cada elemento. La escala que se utiliza va desde el -3 hasta el +3. Por último, hay que decir que este tipo de escala se utiliza mucho para estudiar la imagen de las marcas.

c) Escala de Stapel: En este tipo de escala se utiliza solo una palabra, adjetivo o frase. Es una variación de la escala diferencial-semántico, ya que ahora las puntuaciones van de -5 a +5, de forma que cuanto más se aproxime el término a la realidad más positiva será la puntuación que reciba.

Ejemplo: -5 -4 -3 -2 -1 1 2 3 4 5

x (suave) Sabor Fuerte

x (más barato) Precio alto

Color claro x (bastante claro)

Este tipo de escala sirve para lo mismo que la escala diferencial-semántico, es decir, para el diseño de perfiles, la comparación de elementos, etc.

d) Escala Likert: Este tipo de escala consiste en pedirle al entrevistado que indique su grado de acuerdo o desacuerdo en relación a una serie de afirmaciones en una escala que va de 1 a 5, donde el 1 significa que se está totalmente en desacuerdo con la afirmación, el 2 significa que se está en desacuerdo, el 3 ni en acuerdo ni en desacuerdo, el 4 significa que se está de acuerdo con la afirmación y el 5 totalmente en acuerdo. En la escala Likert los atributos más interesantes son los que discriminan mucho entre los encuestados.

4.2. MÉTODOS DE ANÁLISIS DE LA INFORMACIÓN OBTENIDA EN LA INVESTIGACIÓN COMERCIAL.

Los métodos de análisis de la información obtenida en la investigación comercial se pueden dividir entre los de la estadística descriptiva y los de la estadística inferencial. Dentro de la estadística descriptiva hay dos grupos de métodos: univariantes y bivalentes. Los métodos univariantes analizan información para una sola variable mientras que los métodos bivalentes analizan información simultáneamente para dos variables. Dentro de los métodos univariantes hay que distinguir entre medidas de posición y medidas de dispersión. Dentro de las medidas de posición se puede distinguir entre: medidas de tendencia central, medidas de tendencia no central, cálculo de frecuencias. Dentro las medidas de dispersión se pueden señalar: la varianza, la desviación típica, el rango. En los métodos bivalentes la herramienta más utilizada son las tablas de contingencia o tabulaciones cruzadas. El contraste que se utiliza para ver si dos variables están relacionadas estadísticamente es la Chi- Cuadrado ().

Dentro de la inferencia estadística se va a ver: la estimación de parámetros tanto puntual como por intervalos y el contraste de hipótesis, donde nos vamos a centrar en los contrastes paramétricos sobre medias.

4.3. MÉTODOS DE ANÁLISIS UNIVARIANTES Y BIVARIANTES.

• **MÉTODOS UNIVARIANTES**

Entre los métodos univariantes se pueden citar los siguientes, los cuales se encuentran divididos entre medidas de posición y medidas de dispersión.

Medidas de posición

1) Medidas representativas de la tendencia central de una distribución:

- Media aritmética.
- Media armónica.
- Media geométrica.
- Media ponderada.
- Mediana y moda.

2) Medidas de posición no centrales:

- Cuartiles.
- Deciles.
- Centiles.
- Percentiles.

3) Otras medidas de posición:

- Frecuencia absoluta y relativa.

- Frecuencia absoluta y relativa acumuladas.

Medidas de dispersión

1) Medidas absolutas:

- Recorrido o rango.
- Rango intercuartílico.
- Varianza.
- Desviación típica.
- Desviaciones absolutas.
- Desviaciones cuadráticas.

2) Medidas relativas:

- Recorrido relativo.
- Coeficiente de variación:
- Basados en desviación absolutas.
- Basados en desviación cuadráticas (coeficiente de Pearson).

• MÉTODOS BIVARIABLES

Entre los métodos bivariados las tablas de contingencia permiten analizar la información de dos variables simultáneamente. Pero en las tablas de contingencia esas variables tienen que estar medidas con escalas nominales y tener un número limitado de categorías. Las tablas de contingencia van a medir la asociación que existe entre dos variables, pero en ningún caso sirven para detectar relaciones causa-efecto entre las variables.

Ejemplo: Analizar si existe algún tipo de relación entre ir o no de vacaciones y el nivel de ingresos. A continuación, se da la tabla de frecuencias observadas:

Ingresos (A)				
Irse de vacaciones (B)	REDUCIDOS (A1)	MEDIOS (A2)	ELEVADOS (A3)	<u>Total</u>
SI (B1)	34	45	55	134
NO (B2)	53	52	27	132
<u>Total</u>	87	97	82	266

Con observar esta tabla se puede ver que de los 87 con ingresos reducidos más de la mitad no van de vacaciones y que de los 82 con ingresos elevados más de la mitad si van de vacaciones. Por lo tanto, a priori se puede decir que si existe o parecer existir una relación entre ir o no de vacaciones y el nivel de ingresos.

Para ver si esa relación es estadísticamente significativa se va a calcular la Chi-Cuadrado () la cual permite comparar dos distribuciones de frecuencia, por una parte, la distribución de frecuencias observadas y, por otra parte, la distribución de frecuencias esperadas que se han de calcular en el caso de que las variables fuesen independientes.

Si A y B son independientes (hipótesis nula) $\Rightarrow P(A \cap B) = P(A) \cdot P(B)$

87 134

$P(A) \cdot P(B) =$

266 266

La frecuencia esperada de que suceda $(A \cap B)$ es igual a: $n \cdot P(A \cap B)$.

87 134 $87 \cdot 134$

$n \cdot P(A) \cdot P(B) = 266 \cdot =$

266 266 266

En general, la frecuencia esperada se calculará mediante la siguiente expresión:

$n_{Ai} \cdot n_{Bj}$

$E_{ij} =$

n

REDUCIDOS MEDIOS ELEVADOS

$87 \cdot 134 / 266 = 43,83$	$97 \cdot 134 / 266 = 48,86$	$82 \cdot 134 / 266 = 41,31$
$87 \cdot 132 / 266 = 43,17$	$97 \cdot 132 / 266 = 48,14$	$82 \cdot 132 / 266 = 40,69$

$(\text{Frecuencias observadas} - \text{Frecuencias esperadas})^2$

$=$

Frecuencias esperadas

$(34 - 43,83)^2 (27 - 40,69)^2$

$= + + = 14,201$

43,83 40,69

A continuación, se compara el valor de la Chi-Cuadrado con el valor crítico o valor en tablas de una Chi-Cuadrado () con:

– Grados de libertad : $(n_C - 1) \cdot (n_F - 1) = (3 - 1) \cdot (2 - 1) = 2$ siendo n_C el número de categorías de la variable columna y n_F el número de categorías de la variable fila.

– Nivel de confianza : 95%

En este caso el valor crítico o valor en tablas será igual a: $(2, 0.05) = 5,991$. Por lo tanto, se rechaza la hipótesis nula que era que no existía asociación entre las variables, es decir, que las variables A (ingresos) y B (irse de vacaciones) eran independientes. Hay, por lo tanto, una asociación significativa entre las variables.

Para la interpretación de la asociación existente entre dos variables hay que acudir a la tabla de frecuencias observadas, pero también hay que fijarse en las frecuencias relativas, es decir, habrá que calcular porcentajes. En este caso, interesa los porcentajes por columnas porque son los ingresos los que pueden determinar el que una persona se vaya de vacaciones o no, lo que significa que el ingreso es el que condiciona. Un truco es poner por columnas la variable independiente y luego buscar siempre el porcentaje por columnas.

Pero, ¿cómo se profundiza en el estudio de la asociación existente?. Habrá que ver si la asociación se mantiene cuando se considera variables adicionales. Por ejemplo, en el caso anterior se podría utilizar como variable de control el sexo.

Hombres

Variable control : Sexo

Mujeres

A la asociación original se la conoce como Asociación Total o de Orden Cero. La asociación incluyendo variables de control se llama asociación condicional. Si la asociación desaparece al introducir las variables de control entonces se dice que la asociación inicial era una asociación espúrea (hay una tercera variable, que es la variable de control, que explica a las dos de partida, eso si, explica el comportamiento de las variables iniciales pero no de la asociación inicial).

4.4. MÉTODOS DE INFERENCIA ESTADÍSTICA.

• ESTIMACIÓN DE PARÁMETROS

Dentro de la estimación de parámetros se pueden distinguir dos tipos, que serían la estimación de parámetros puntual y la estimación de parámetros por intervalos.

1– Puntual: En la estimación puntual se asigna un valor concreto al parámetro poblacional que se desea estimar. Por ejemplo, utilizar la media muestral como estimador de la media poblacional. Los requisitos que se suelen pedir a un estimador son los siguientes:

- Que sea insesgado, lo que quiere decir que las estimaciones que se hagan sean correctas (que estén libres de error).
- Que sea consistente, es decir, que a medida que aumenta el tamaño de la muestra (n) mayor sea la probabilidad de que el estimador coincida con el parámetro poblacional (que la estimación sea correcta).
- Que tenga la menor varianza posible.
- Que sea suficiente, es decir, que utilice el mayor número posible de los datos muestrales disponibles para estimar el parámetro.

2– Por intervalos: Consiste en determinar un intervalo de confianza o un rango de valores entre los que se crea que puede estar el parámetro poblacional a estimar con una probabilidad determinada y alta. Para ello, lo que se hace es, simplemente, sumar y restar al estimador una cantidad concreta que se denomina error máximo. Ese error máximo es igual a:

S1

Error máximo = $Z \cdot$

n

Z: número de unidades de desviación típica asociadas a un determinado nivel de confianza.

S1 / n : es la desviación típica de la distribución muestral del estimador.

El estimador más el error máximo da el límite superior mientras que el estimador menos el error máximo da el límite inferior.

Ejemplo: Una muestra aleatoria de 1.000 personas responde a una cuestión sobre el gasto mensual en periódicos y revistas observándose una media de 3.200 pesetas y una cuasidesviación típica de 320 pesetas. ¿Entre que límites se encuentra el verdadero gasto en periódicos de la población con un nivel deseado de confianza del 95%?. La pregunta que se nos realiza es sobre cual es la media poblacional, es decir, .

n = 1.000 personas

X = 3.200 pesetas

S1 = 320 pesetas

S1 320

$3.200 \pm Z \cdot = 3.200 \pm 1,96 = 19,833 ! 3.200 \pm 19,833$

n 1.000

El intervalo que se nos pide será el siguiente: (3180.167, 3219.833). En este intervalo se considera que puede estar la media poblacional con un nivel de confianza del 95%.

• CONTRASTE DE HIPÓTESIS

El contraste de hipótesis, también denominado prueba de significación, es un método de toma de decisiones que parte de un enunciado o proposición acerca del valor o conjunto de valores que toma un parámetro de la población, y trata de ver o averiguar si se sostiene o se rechaza según los datos muestrales disponibles. Al enunciado o proposición de la que se parte se le denomina hipótesis nula. Como norma general en marketing a la hora de redactar el enunciado del que se parte se propone el valor del parámetro que entraña un menor riesgo para la empresa que va a tomar la decisión.

La hipótesis alternativa son los enunciados que recogen lo que interesa aceptar si se rechaza la hipótesis nula. El error tipo I sería rechazar una hipótesis nula siendo ésta verdadera y el error tipo II sería aceptar una hipótesis nula siendo esta falsa. Estos errores se suelen representar, respectivamente, con α y β . En la investigación de mercados el procedimiento más común es fijar el nivel de error tipo I que el investigador está dispuesto a aceptar, es decir, α es la probabilidad de cometer el error tipo I. El nivel de error con el que se suele trabajar es de 0.05, por lo tanto, $1 - \alpha$ es igual a 0,95. Al $1 - \alpha$ también se le denomina nivel de significación y a $1 - \beta$ se le llama nivel de confianza.

Nos vamos a centrar en los contrastes paramétricos sobre medias, en concreto en tres casos, que son:

- Contraste de hipótesis para la media de una muestra.
- Contraste de hipótesis para las medias de dos muestras independientes.
- Contraste de hipótesis para las medias de muestras relacionadas.

Se va a asumir que hay normalidad y que se está trabajando con escalas de intervalo o de razón.

a) Contraste de hipótesis para la media de una muestra: Permite comparar la media muestral con una hipotética media poblacional, y ver si hay diferencias significativas entre ellas. Hay dos tipos de pruebas que se pueden realizar para llevar a cabo la comparación.

– Prueba z: esta prueba se utiliza cuando la desviación típica poblacional es conocida o cuando no se conoce la desviación típica poblacional y el tamaño de la muestra es mayor de 30.

– Prueba t: esta prueba se utiliza cuando no se conoce la desviación típica poblacional y el tamaño de la muestra es menor de 30.

La prueba que se utiliza con una mayor frecuencia es la prueba z ya que no se suele conocer la desviación típica poblacional y el tamaño de la muestra suele ser mayor de 30. La expresión a utilizar sería la siguiente:

$$X -$$

$$z =$$

$$S1 / n$$

Luego se compara el valor de z con el valor crítico o valor en tablas para un determinado nivel de confianza. Si z es mayor que el valor crítico para un nivel de confianza se rechaza la hipótesis nula.

Ejemplo: Una empresa industrial está interesada en conocer si una nueva máquina de embotellado recibe una valoración superior a 7 en una escala de intención de compra de diez puntos. Para ello contacta con una muestra de 40 ingenieros responsables de compra de otras tantas plantas embotelladoras a quienes se les explica el funcionamiento del nuevo producto. Los resultados obtenidos son: media muestral igual a 7.9, cuasidesviación típica de 1,6. ¿Será introducido ese nuevo producto en el mercado para un nivel de confianza del 95%?

La hipótesis nula para este problema es que $\mu \leq 7$ y la hipótesis alternativa sería que $\mu > 7$. Para un nivel de confianza del 95% el valor crítico será igual a 1,64. En este caso se utiliza la prueba z porque el tamaño de la muestra (n) es igual a 40 que es mayor de 30 y porque la desviación típica poblacional es desconocida.

$$X - 7,9 - 7$$

$$z = \frac{7,9 - 7}{1,6 / \sqrt{40}} = 3,5576$$

$$S1 / n 1,6 / 40$$

Se rechaza la hipótesis nula porque z es mayor que el valor crítico ($3,55 > 1,64$) y, por lo tanto, se lanza el nuevo producto al mercado porque la intención de compra va a estar por encima de 7 en una escala de 10 puntos.

b) Contraste de hipótesis para las medias de dos muestras independientes: Nos permite comparar las medias muestrales de dos grupos de individuos distintos que han sido seleccionados aleatoriamente. Como en el caso anterior hay dos posibles pruebas que son:

- Prueba z: esta prueba se utilizará si la desviación típica poblacional es conocida o si ésta no es conocida y el tamaño de la muestra es mayor de 30.
- Prueba t: se utiliza esta prueba si la desviación típica poblacional no es conocida y el tamaño de la muestra es menor de 30.

Además de fijarnos si el tamaño de la muestra es mayor o menor de 30 y si la desviación típica poblacional es conocida o desconocida habrá que tener en cuenta la homogeneidad de varianzas. En el caso de que haya homogeneidad de varianzas se utilizará una expresión determinada y si no hay homogeneidad se utilizará otra expresión que sería el test de Levene. La expresión a utilizar en el caso de que la desviación típica poblacional sea desconocida, el tamaño de la muestra sea mayor de 30 y haya homogeneidad de varianzas será la siguiente:

$$\bar{X}_1 - \bar{X}_2 - (\quad)$$

$$z =$$

$$\frac{(n_1 - 1) \cdot S_1^2 + (n_2 - 1) \cdot S_2^2}{n_1 + n_2}$$

.

$$n_1 + n_2 - 2 \quad n_1 \cdot n_2$$

La hipótesis nula (H0) suele ser la siguiente: $\mu_1 = \mu_2$. Mientras que la hipótesis alternativa (H1) suele ser la siguiente: $\mu_1 \neq \mu_2$.

El resultado obtenido mediante la anterior expresión se compara con el valor crítico para un determinado nivel de confianza y si z es mayor que el valor crítico se rechaza la hipótesis nula.

Ejemplo: Un fabricante de patatas fritas decide realizar un test de prueba para un nuevo producto que son las patatas fritas al ajo. Una muestra de 100 personas menores de 30 años le otorga una valoración media de 4,9 con una cuasidesviación típica de 1,7 en un test de sabor que compara dicho nuevo producto con la elaboración de patatas fritas tradicionales. El valor medio de otra muestra de 100 personas mayores de 30 años es de 4,3 con una cuasidesviación típica de 1,8. ¿Se puede considerar diferente la opinión de ambos segmentos de mercado para un coeficiente o nivel de confianza del 95%?

En este caso se utiliza la prueba Z porque el tamaño de la muestra es mayor de 30 y no se conoce la desviación típica poblacional y se va a asumir la homogeneidad de varianzas. La hipótesis nula será que no hay diferencias entre las opiniones de los menores de 30 años y las de los mayores de 30 años ($\mu_1 = \mu_2$). La hipótesis alternativa será que si hay diferencias entre las opiniones de los dos segmentos de mercado ($\mu_1 \neq \mu_2$).

$$4,9 - 4,3$$

$$z = \frac{4,9 - 4,3}{\sqrt{\frac{(100 - 1) \cdot 1,7^2 + (100 - 1) \cdot 1,8^2}{100 + 100}}}$$

$$= \frac{0,6}{\sqrt{\frac{170,7 + 324,0}{200}}} = \frac{0,6}{\sqrt{2,4735}} = \frac{0,6}{1,5727} = 0,3815$$

.

$$100 + 100 - 2 \quad 100 \cdot 100$$

Para un nivel de confianza del 95% el valor crítico es igual a 1,96. El valor que se obtiene mediante la anterior expresión es mayor que el valor crítico de forma que se rechaza la hipótesis nula, lo que quiere decir que hay diferencias entre las opiniones, lo que llevaría a segmentar el mercado y adaptar la política de comunicación para dirigirla hacia los jóvenes.

c) Contraste de hipótesis para las medias de muestras relacionadas: Tenemos un mismo grupo de entrevistados que valoran un producto o servicio en dos momentos distintos del tiempo, que valoran dos productos competidores, o que valoran la importancia de dos atributos o aspectos distintos de un servicio o producto.

En estas situaciones se crea una variable nula igual a la diferencia de opiniones. La hipótesis nula va a ser que no existe diferencia entre las opiniones, valoraciones o puntuaciones. Las pruebas que se pueden utilizar son las mismas que se utilizan en los dos casos anteriores. La expresión a utilizar, en el caso de que el tamaño de la muestra sea mayor de 30 y la desviación típica poblacional sea desconocida, será la siguiente:

D

$z =$

SD / n

A continuación, se compara el valor obtenido con esta expresión con el valor crítico para un determinado nivel de confianza, y si z es mayor que el valor crítico entonces se rechaza la hipótesis nula.

Ejemplo: Un investigador de mercados selecciona una muestra aleatoria de 270 personas quienes valoran la importancia de dos atributos a la hora de seleccionar entidades bancarias, siendo dicho atributos los siguientes:

- La amabilidad o atención personal.
- La rentabilidad ofrecida por las inversiones.

Los resultados obtenidos son los siguientes:

	n	D	SD
Amabilidad	270	5,63	0,67
Rentabilidad	270	4,19	1,62
Diferencias en las opiniones para cada individuo	270	1,44	1,61

¿Hay evidencia suficiente para afirmar, con un nivel deseado de confianza del 95%, que la importancia asignada a ambos atributos es diferente?. Se utiliza la prueba z porque el tamaño de la muestra es mayor de 30 y la desviación típica poblacional es desconocida. La hipótesis nula (H_0) sería que no existen diferencias en las opiniones, es decir, la diferencia entre la importancia de la amabilidad y la rentabilidad es nula.

1,44

$z = 14,6967$

En este caso se rechaza la hipótesis nula porque Z es mayor que el valor crítico para un nivel de confianza del 95% ($14,69 > 1,96$). Desde el punto de vista del marketing esto supone que si hay diferencia en la importancia que se concede a los distintos atributos, amabilidad y rentabilidad, asignando una mayor importancia a la amabilidad.

TEMA 5: ANÁLISIS CAUSAL: LA EXPERIMENTACIÓN COMERCIAL.

5.1. FUNDAMENTOS Y PLANIFICACIÓN DE LA EXPERIMENTACIÓN DE MERCADO O COMERCIAL.

• INTRODUCCIÓN

A lo largo de este quinto tema se va a ver el análisis de relaciones causa-efecto. Desde el punto de vista del marketing, se va a tratar de analizar las consecuencias de las actuaciones en la empresa sobre el precio, el producto, la publicidad, etc. Es decir, las consecuencias de las actuaciones sobre las variables del marketing.

• DEFINICIÓN

La experimentación de mercado o experimentación comercial es una prueba o serie de pruebas en las cuales se inducen cambios deliberados en una o más variables independientes (precio, publicidad, promoción, productos, fuerza de ventas, etc.), de manera que sea posible observar, identificar y determinar su efecto sobre una o más variables dependientes (generalmente las ventas).

• CONCEPTOS CLAVES

Hay una serie de elementos o conceptos que se deben especificar claramente para que la experimentación de mercado o comercial esté perfectamente diseñada.

1- Tratamiento experimental. Se refiere o corresponde a cada una de las variables independientes (modificación en los precios, presentación de un nuevo producto, cambios en los medios publicitarios, un nuevo canal de distribución, etc.) cuyos efectos se someterán a prueba en un experimento. Al tratamiento también se le denomina, en ocasiones, factor experimental. Cada uno de estos tratamientos tendrá varios niveles o categorías que se denominarán nivel i del tratamiento siendo $i = 1, 2, 3, \dots$

2- Unidades experimentales. Son las unidades que se observan como se comportan ante las manipulaciones en las variables independientes. Hay dos tipos de unidades experimentales:

- *De prueba*: se refieren a todas aquellas personas o entidades a las que se les presentan los tratamientos y cuya respuesta se debe medir, es decir, es objeto de estudio.
- *De control*: son unidades observadas durante el período de prueba, es decir, cuya respuesta es medida, pero sin ser sometidas al tratamiento.

Las unidades de control se utilizan para verificar, por lo que no son imprescindibles en un experimento de mercado, mientras que las unidades de prueba si son imprescindibles.

3- Variables respuesta o variables dependientes. Son las medidas que se toman de las unidades de prueba (unidades experimentales). Por ejemplo, las ventas, la cuota de mercado, el grado de recuerdo, la notoriedad, etc.

4- Variables externas. Son todas aquellas variables diferentes a las independientes o tratamientos que también

pueden afectar a las unidades de prueba, es decir, pueden influir sobre la respuesta de las unidades de prueba.

5- Diseño experimental. Se refiere al método de experimentación que se va a utilizar en la investigación comercial. Dentro del diseño experimental hay que especificar los siguientes datos:

- Tratamiento o variables independientes.
- Unidades de prueba elegidas.
- Variables dependientes a medir.
- Modo en que se van a controlar las variables externas.
- El enfoque estadístico que se adoptará para analizar los datos y garantizar la validez interna y externa de los resultados obtenidos. La validez interna permite asegurar que los efectos observados en las unidades de prueba se deben a las variables independientes o tratamientos. La validez externa permite generalizar los resultados de la muestra al total de la población.

La técnica estadística que se suele utilizar es el análisis de la varianza (ANOVA). Los tres requisitos que se deben cumplir para aplicar esta técnica son los siguientes:

- Normalidad de la distribución.
- Homogeneidad de las varianzas.
- Independencia de los resultados, para lo cual hay que hacer una asignación aleatoria de los tratamientos.

5.2. TIPOS DE EXPERIMENTOS.

Entre los tipos de diseño que se pueden utilizar, en esta pregunta se van a ver los cuatro siguientes:

- Diseño completamente aleatorio.
- Diseño en bloques aleatorios.
- Diseño cuadrado latino.
- Diseño factorial.

• **DISEÑO COMPLETAMENTE ALEATORIO**

En el diseño completamente aleatorio se analiza o manipula una variable independiente con varias categorías o niveles, cada uno de los cuales es asignado de modo aleatorio a las unidades de prueba, suponiendo que las variables externas no afectan de diferente manera a cada uno de los grupos. Esto es lo mismo que decir que este diseño o este tipo de experimento comercial es adecuado para aquellas situaciones en las que las unidades de prueba son homogéneas, es decir, las unidades de prueba no difieren unas de otras en variables externas que sean importantes.

El objetivo de un diseño completamente aleatorio es estudiar si existen diferencias significativas en la variable dependiente para cada nivel de la variable independiente o tratamiento. Específicamente, lo que nos interesa es hacer un Test de Singularidad de Medias. El procedimiento a utilizar para alcanzar este objetivo (la singularidad de medias) es el análisis de la varianza (ANOVA), técnica muy útil en el campo de la inferencia estadística, que toma como referencia el modelo lineal estadístico. En concreto, en este diseño se toma como referencia el siguiente modelo lineal estadístico:

$$Y_{ij} = EG + ET_j + EA$$

Y_{ij} : Valor observado de la variable dependiente en la unidad de prueba i con el tratamiento j .

EG: Es el efecto global, que se refiere al promedio de la variable dependiente para el total o conjunto de las unidades de prueba.

ETj: Es el efecto del tratamiento j, es decir, la variación en la variable dependiente con relación al promedio debido a la influencia del tratamiento j.

EA: Es el efecto aleatorio, el cual recoge el efecto de todas las restantes causas posibles de variabilidad del experimento no debidas al tratamiento, es decir, al efecto aleatorio se le podría llamar también otras causas de variabilidad del experimento.

ETj + EA: La suma de estos dos efectos refleja las fuentes de dispersión. Concretamente, el análisis de la varianza (ANOVA) va a consistir en el análisis de esa dispersión, es decir, en el estudio de los componentes de la dispersión total:

1– Dispersión por los tratamientos.

2– Dispersión residual debido a los errores.

Dispersión Total o Global (DG) = Dispersión Tratamiento j (DTj) + Dispersión Residual (DR)

: Valor observado de la variable dependiente en la unidad de prueba i con el tratamiento j.

Gran media.

: Número de mediciones efectuadas para cada tratamiento.

Comportamiento medio de un tratamiento.

Generalmente, se tendrá un cuadro como el siguiente, a partir del cual se deberán calcular las anteriores medidas citadas.

	Nivel 1 del tratamiento	Nivel 2 del tratamiento	Nivel 3 del tratamiento
Unidad de prueba	Y11	Y12	Y13
Unidad de prueba	Y21	Y22	Y23
Unidad de prueba	Y31	Y32	Y33

En concreto, en esta tabla se tienen nueve unidades de prueba, aunque podrían ser muchas más, siendo todas ellas homogéneas entre si. Las medidas que hay que calcular son las siguientes:

$Y_{11} + Y_{21} + Y_{31}$

3

$Y_{12} + Y_{22} + Y_{32}$

3

$Y_{13} + Y_{23} + Y_{33}$

3

Por último, para saber si el efecto de la variable independiente sobre la variable dependiente es o no significativo se ha de construir un estadístico F que es igual a la siguiente expresión:

DTj / Grados de libertad del tratamiento

F =

DR / Grados de libertad residuales

- Grados de libertad totales = N (número de mediciones) – 1.
- Grados de libertad del tratamiento = n (número de mediciones realizadas para cada tratamiento) – 1.
- Grados de libertad residuales = (N – 1) – (n – 1) = N – n.

Este estadístico, una vez construido, se tendrá que comparar con el valor crítico o en tablas para un determinado nivel de confianza. Si el valor del estadístico F es mayor que el valor crítico se dirá que el efecto de la variable independiente sobre la dependiente es significativo. Por el contrario, si el valor del estadístico es menor que el valor crítico entonces se dirá que el efecto de la variable independiente sobre la dependiente no es significativo. La hipótesis nula será, por lo tanto, que el efecto de la variable independiente sobre la dependiente no es significativo.

EJERCICIO: Una empresa se propone fijar el precio de un nuevo producto que va a lanzar al mercado. Para ello decide realizar un experimento de mercado en 15 tiendas detallistas de características similares. En dicho experimento se contempla tres posibles precios: 60 pesetas, 80 pesetas, 100 pesetas. Los resultados obtenidos en el experimento son los que se muestran en la siguiente tabla. ¿Tiene el precio un efecto significativo sobre las ventas?

	TRATAMIENTOS		
	PRECIO 60	PRECIO 80	PRECIO 100
UNIDADES DE PRUEBA (TIENDAS DETALLISTAS)	66	48	36
	67	52	35
	59	50	34
	53	52	45
	51	51	49
TOTAL TRATAMIENTOS	296	253	199

El diseño a aplicar es el completamente aleatorio porque se ha de analizar el efecto de una variable independiente o tratamiento (precio) sobre una variable dependiente (ventas). El tratamiento tiene varios niveles (3) que se asignan de modo aleatorio a las unidades de prueba (cada nivel a un grupo de unidades de prueba). Las unidades de prueba son esencialmente similares y representativas del tipo de establecimiento que vende el producto en el mercado, es decir, no difieren en una variable externa importante. Lo primero que se hace es calcular las siguientes medidas, utilizando para ello los datos dados en la anterior tabla.

$$66 + 67 + 59 + 53 + 51$$

$$= 296$$

$$5$$

$$48 + 52 + 50 + 52 + 51$$

$$= = 50,6$$

5

$$36 + 35 + 34 + 45 + 49$$

$$= = 39,8$$

5

$$66 + 67 + 59 + 53 + 51 + 48 + 52 + 50 + 52 + 51 + 36 + 35 + 34 + 45 + 49$$

$$= = 49,87$$

15

Una vez calculadas estas medidas se podrá comenzar a calcular las distintas dispersiones.

DISPERSIÓN TOTAL = DISPERSIÓN TRATAMIENTO + DISPERSIÓN RESIDUAL

$$\text{DISPERSIÓN TRATAMIENTO (precio)} = 5 \cdot [(59,2 - 49,87)^2 + (50,6 - 49,87)^2 + (39,8 - 49,87)^2] = 944,9$$

$$\text{DISPERSIÓN RESIDUAL} = (66 - 59,2)^2 + (67 - 59,2)^2 + (59 - 59,2)^2 + (53 - 59,2)^2 + (51 - 59,2)^2 + (48 - 50,6)^2 + (52 - 50,6)^2 + (50 - 50,6)^2 + (52 - 50,6)^2 + (51 - 50,6)^2 + (36 - 39,8)^2 + (35 - 39,8)^2 + (34 - 39,8)^2 + (45 - 39,8)^2 + (49 - 39,8)^2 = 406,8$$

A continuación, habría que construir el estadístico F. Para que sea más sencillo calcularlo se construirá el siguiente cuadro:

FUENTES DE DISPERSIÓN	SUMA DE CUADRADOS	GRADOS DE LIBERTAD	MEDIAS CUADRÁTICAS	VALOR DEL ESTADÍSTICO F
• Tratamiento (precio)	944,9	3 - 1 = 2	944,9 / 2 = 472,5	472,5 / 33,9 = 13,937
• Residual	406,8	14 - 2 = 12	406,8 / 12 = 33,9	
TOTAL	1.351,7	15 - 1 = 14		

El valor crítico, que se representaría por t_{α} , para un nivel de confianza del 95% es igual a 3,89. Como el valor del estadístico es mayor que el valor crítico esto quiere decir que la política de precios tiene un efecto significativo sobre la ventas, es decir, se rechaza la hipótesis nula que era que el efecto de la variable independiente (precio) sobre la variable dependiente (ventas) no era significativo.

• DISEÑO EN BLOQUES ALEATORIOS

En el diseño en bloques aleatorios el investigador puede controlar una variable externa, la cual podría confundir los resultados del experimento. Este diseño recibe el nombre de bloques aleatorios porque estratifica las unidades de prueba o experimentales en función de la variable externa (tipo de establecimiento,

tamaño del área geográfica, edad, ingresos del cliente) que se controla. En cada bloque o estrato los tratamientos son asignados de manera aleatoria a las unidades de prueba. El modelo en el que se apoya este método es el análisis de la varianza y el modelo estadístico lineal para este diseño es igual a:

$$Y_{ij} = EG + ET_j + EBi \text{ (efecto del bloque)} + EA$$

$$\text{Dispersión Global} = \text{Dispersión Tratamiento } j + \text{Dispersión Bloque } i + \text{Dispersión Residual}$$

La dispersión residual es igual a la dispersión global menos la dispersión del tratamiento j menos la dispersión del bloque i . A continuación, se construirían dos estadísticos, uno para el tratamiento y otro para el bloque, siendo las expresiones a utilizar las siguientes:

DT_j / Grados de libertad del tratamiento

F (para el tratamiento) =

DR / Grados de libertad residuales

DB_i / Grados de libertad del bloque

F (para el bloque) =

DR / Grados de libertad residuales

Por último, para aceptar o rechazar la hipótesis nula se compararán estos estadísticos con los respectivos valores críticos para un determinado nivel de confianza. Si el valor del estadístico es mayor que el valor crítico se dirá que el efecto de la variable independiente sobre la dependiente es significativo. Por el contrario, si el valor del estadístico es menor que el valor crítico entonces se dirá que el efecto de la variable independiente sobre la dependiente no es significativo. Esto mismo se haría con el estadístico para el bloque.

El principal inconveniente de este tipo de diseño es que no tiene en cuenta el posible efecto interacción entre la variable externa (bloque) y el tratamiento o variable independiente. Es por tanto, adecuado cuando no haya relación alguna entre la variable externa y la variable independiente.

EJERCICIO: Una empresa de productos de belleza acaba de crear una nueva crema cuyo precio está todavía por determinar. Para tomar esta decisión, ensaya tres precios diferentes (900, 1.100 y 1.500), ensayo que realiza en una determinada área geográfica durante un mes, a través de diferentes perfumerías y mediante una campaña de publicidad en prensa y radio local. Las perfumerías son bastante diferentes entre sí, debido a su localización, forma de operar y cifra de ventas. Por ello, la empresa establece cinco grupos de perfumerías con cierta homogeneidad, donde va a experimentar la influencia del precio en las ventas.

Si en cada grupo el reparto de los tres tratamientos del precio se realiza al azar y las ventas medias por perfumería / mes para cada uno de los precios ensayados se indica en el cuadro siguiente. ¿Han tenido influencia los precios en la venta del producto? ¿Qué precio parece el más adecuado? ¿Por qué?

	Precio de 900	Precio de 1.100	Precio de 1.500	Media bloques
GRUPO I	540	420	300	1 . = 420
GRUPO II	570	450	390	2 . = 470
GRUPO III	540	480	408	3 . = 476
GRUPO IV	510	450	330	4 . = 430
GRUPO V	528	432	288	5 . = 416

Media tratamientos	. 1 = 537,6	. 2 = 446,4	. 3 = 343,2	= 442,4
--------------------	-------------	-------------	-------------	---------

Se aplica el diseño en bloques aleatorios como consecuencia de que hay una variable independiente que es el precio cuyo efecto sobre la variable dependiente (ventas) se quiere analizar. El precio tiene tres niveles o categorías (900, 1.100, 1.500). Además, se tiene control sobre una variable externa que es el tipo de perfumería.

$$\text{DISPERSIÓN GLOBAL O TOTAL} = (540 - 442,4)^2 + (570 - 442,4)^2 + (540 - 442,4)^2 + (510 - 442,4)^2 + (528 - 442,4)^2 + (420 - 442,4)^2 + (450 - 442,4)^2 + (480 - 442,4)^2 + (450 - 442,4)^2 + (432 - 442,4)^2 + (300 - 442,4)^2 + (390 - 442,4)^2 + (408 - 442,4)^2 + (330 - 442,4)^2 + (288 - 442,4)^2 = 110.049,6$$

$$\text{DISPERSIÓN TRATAMIENTO (precio)} = 5 \cdot [(537,6 - 442,4)^2 + (446,4 - 442,4)^2 + (343,2 - 442,4)^2] = 94.598,4$$

$$\text{DISPERSIÓN BLOQUE (tipo de perfumería)} = 3 \cdot [(420 - 442,4)^2 + (470 - 442,4)^2 + (476 - 442,4)^2 + (430 - 442,4)^2 + (416 - 442,4)^2] = 9.729,6$$

$$\text{DISPERSIÓN RESIDUAL} = 110.049,6 - 94.598,4 - 9.729,6 = 5.721,6$$

A continuación, se habrá de calcular los valores de los estadísticos, para lo cual se construye el siguiente cuadro:

FUENTES DE DISPERSIÓN	SUMA DE CUADRADOS	GRADOS DE LIBERTAD	MEDIAS CUADRÁTICAS	VALOR DEL ESTADÍSTICO F
• Tratamiento (precio)	94.598,4	3 - 1 = 2	94.598,4 / 2 = 47.299,2	47.299,2 / 715,2 = 66,13 2.432,4 / 715,2 = 3,4
• Bloque (tipo de perfumería)	9.729,6	5 - 1 = 4	9.729,6 / 4 = 2.432,4	
• Residual	5.721,6	14 - 2 - 4 = 8	5.721,6 / 8 = 715,2	

Los valores críticos son los siguientes: $t_{0,05} = 4,46$ y $t_{0,01} = 3,84$. Como el estadístico F para el tratamiento es mayor que el valor crítico para éste, significa que se rechaza la hipótesis nula, es decir, el precio influye significativamente sobre las ventas. Por otro lado, como el valor crítico para el bloque es mayor que el estadístico F para el bloque no se puede rechazar la hipótesis nula, de forma que no se puede afirmar que haya una influencia significativa del tipo de perfumería sobre las ventas. A simple vista, parece ser que el precio más conveniente es el de 900, donde el volumen medio de ventas es mayor. Sin embargo, con un precio de 1.500 los ingresos son los más altos de las tres alternativas posibles. Es decir, un aumento del precio no hace que las ventas caigan tanto como para ganar más con un precio de 900. Dicho de otra forma, un precio de 900 no trae consigo un incremento de las ventas tan grande como para compensar una disminución del precio de 1.500 a 900.

- Con un precio de 900 los ingresos serían igual a: $900 \cdot 537,6 = 483.840$
- Con un precio de 1.100 los ingresos serían igual a: $1.100 \cdot 446,4 = 491.040$
- Con un precio de 1.500 los ingresos serían igual a: $1.500 \cdot 343,2 = 514.800$

• DISEÑO CUADRADO LATINO

El diseño cuadrado latino permite al investigador controlar dos variables externas. En este diseño, las unidades de prueba se van a estratificar en función de dos criterios o variables: Bi y Bk. En cada bloque o estrato, las unidades experimentales se asignan de forma aleatoria a los tratamientos, aunque el procedimiento debe garantizar que un tratamiento determinado aparezca una sola vez en cada bloque o estrato.

Para poder aplicar el diseño cuadrado latino todas las variables, tanto las externas que se controlan como las independientes, tienen que tener el mismo número de niveles. El inconveniente del diseño cuadrado latino es que no tiene en cuenta las interacciones que pueden haber entre las variables independientes y las variables externas ni tampoco la interacción entre las propias variables externas, así que sólo será válido cuando los efectos interacción sean insignificantes o nulos. El modelo estadístico lineal para este tipo de diseño es el siguiente:

$$Y_{ij} = EG + ET_j + EBi + EBk + EA$$

$$\text{Dispersión Global} = \text{Dispersión Tratamiento } j + \text{Dispersión Bloque } i + \text{Dispersión bloque } k \\ + \text{Dispersión Residual}$$

La dispersión residual es igual a la dispersión global menos la dispersión del tratamiento j menos la dispersión del bloque i menos la dispersión del bloque k.

En el diseño cuadrado latino se han de calcular tres estadísticos, que luego habrá que comparar con los respectivos valores críticos o en tablas. Las expresiones para calcular los diferentes estadísticos son las siguientes:

DTj / Grados de libertad del tratamiento

F (para el tratamiento) =

DR / Grados de libertad residuales

DBi / Grados de libertad del bloque i

F (para el bloque i) =

DR / Grados de libertad residuales

DBk / Grados de libertad del bloque k

F (para el bloque k) =

DR / Grados de libertad residuales

Si el valor del estadístico F para el tratamiento es mayor que el valor crítico para un determinado nivel de confianza se rechazará la hipótesis nula, es decir, la variable independiente tendrá un efecto significativo sobre la variable dependiente. Lo mismo se haría con los otros dos estadísticos, es decir, con el estadístico para el bloque i y con el estadístico para el bloque k.

EJERCICIO: Una empresa desea experimentar políticas de formación de vendedores. Dado que los territorios de venta donde se ubican sus clientes tienen diferente potencial de ventas, la empresa piensa que

este hecho puede tener influencia sobre los resultados de venta alcanzados por los vendedores. La edad de los vendedores también puede influir sobre los resultados. Los programas de formación sometidos a experimento son: (A) ningún entrenamiento; (B) cursillo por empresas de consulting; (C) formación mediante visitas a clientes y acompañados por otros vendedores con experiencia. Si los resultados obtenidos son los que se exponen en el siguiente cuadro, ¿Qué opinión le merecen la influencia de las variables sometidas a prueba sobre las ventas?

	POTENCIAL DE VENTAS			
EDAD DE LOS VENDEDORES	200 – 500	501 – 1.000	1.001 – 1.500	MEDIA BLOQUE
18 – 30	314 (A)	599 (B)	703 (C)	538,7
31 – 45	424 (B)	896 (C)	496 (A)	605,3
+ 45	238 (C)	314 (A)	312 (B)	288
MEDIA BLOQUE	325,3	603	503,7	= 477,3

En este ejercicio se está estudiando el efecto de tres programas de formación de los vendedores (variable independiente) sobre el volumen de ventas (variable dependiente). La variable independiente tiene tres niveles. Hay dos variables externas, distintas al tratamiento, que son la edad y el potencial de ventas del territorio que se asigna a los vendedores. El diseño apropiado consiste en aplicar cada programa de formación exactamente una vez para cada intervalo de edad de los vendedores y en cada territorio. Por todos estos motivos se va aplicar el diseño cuadrado latino. En la anterior tabla se tienen las variables externas y el nivel de tratamiento, que vendría representado por las letras A, B y C. Lo primero que hay que hacer es calcular los valores totales del tratamiento, de modo que se pueda construir una tabla con los valores para cada uno de los tratamientos.

	A	B	C
TOTAL			
TRATAMIENTOS	314 + 314 + 496 = 1.124	599 + 424 + 312 = 1.335	238 + 896 + 703 = 1.837
MEDIA			
TRATAMIENTOS	1.124 / 3 = 374,7	1.335 / 3 = 445	1.837 / 3 = 612,3

$$\text{DISPERSIÓN GLOBAL} = (314 - 477,3)^2 + (424 - 477,3)^2 + (238 - 477,3)^2 + (599 - 477,3)^2 + (896 - 477,3)^2 + (314 - 477,3)^2 + (703 - 477,3)^2 + (496 - 477,3)^2 + (312 - 477,3)^2 = 382.174,01$$

$$\text{DISPERSIÓN TRATAMIENTO (programas de formación)} = 3 \cdot [(374,7 - 477,3)^2 + (445 - 477,3)^2 + (612,3 - 477,3)^2] = 89.385,15$$

$$\text{DISPERSIÓN BLOQUE (edad)} = 3 \cdot [538,7 - 477,3)^2 + (605,3 - 477,3)^2 + (288 - 477,3)^2] = 167.965,35$$

$$\text{DISPERSIÓN BLOQUE (potencial de ventas)} = 3 \cdot [325,3 - 477,3)^2 + (603 - 477,3)^2 + (503,7 - 477,3)^2] = 118.804,35$$

$$\text{DISPERSIÓN RESIDUAL} = 382.174,01 - 89.385,15 - 167.965,35 - 118.804,35 = 6.019,16$$

A continuación, se ha de construir los estadísticos F, para lo cual se realizará la siguiente tabla:

FUENTES DE	SUMA DE	GRADOS	MEDIAS	VALOR DEL
-------------------	----------------	---------------	---------------	------------------

DISPERSIÓN	CUADRADOS	DE LIBERTAD	CUADRÁTICAS	ESTADÍSTICO F
• TRATAMIENTO (Programa de formación)				
• BLOQUE i	89.385,15	3 - 1 = 2	89.385,15 / 2 = 44.692,6	
(Edad)	167.965,35	3 - 1 = 2	167.965,35 / 2 = 83.982,7	44.692,6 / 3.009,6 = 14,85
♦ BLOQUE	118.804,35	3 - 1 = 2	118.804,35 / 2 = 59.402,2	83.982,7 / 3.009,6 = 27,9
k	6.019,16	8 - 6 = 2	6.019,16 / 2 = 3.009,6	59.402,2 / 3.009,6 = 19,74
(Potencial de ventas)	382.174,01	9 - 1 = 8		
• RESIDUAL				
• GLOBAL				

El valor crítico es el siguiente: = 19, siendo éste el valor crítico tanto para el tratamiento, como para la edad, como para el potencial de ventas del territorio. Como el estadístico F para el tratamiento es menor que el valor crítico para éste, significa que no se puede rechazar la hipótesis nula, es decir, no podemos afirmar que haya una influencia significativa del programa de formación sobre las ventas. Por otro lado, como el valor crítico para el bloque i es menor que el estadístico F para este bloque se rechaza la hipótesis nula, de forma que se puede afirmar que hay una influencia significativa de la edad sobre las ventas. Por último, como el valor crítico para el bloque k es menor que el estadístico F para este bloque se rechaza la hipótesis nula, lo que supone que el potencial de ventas del territorio influye significativamente sobre las ventas. En resumen, las únicas variables con efecto sobre las ventas son las variables externas.

• DISEÑO FACTORIAL

En el diseño factorial se van a manipular dos o más variables independientes. Un diseño factorial trata de medir o cuantificar los efectos simultáneos de dos o más variables independientes sobre una variable dependiente, lo que da lugar a dos tipos de mediciones: efectos principales y efectos interacción.

Por efecto principal se entiende la influencia de cada una de las variables por separado y el efecto interacción es el efecto conjunto de dos o más variables. Es muy importante señalar que siempre en el diseño factorial tiene que haber el mismo número de observaciones para cada combinación de niveles. El modelo estadístico lineal para este tipo de diseño es el siguiente:

$$Y_{ij} = \mu + \alpha_i + \beta_j + \gamma_{ij} + \epsilon_{ij}$$

$$\text{Dispersión Global} = \text{Dispersión Tratamiento i} + \text{Dispersión Tratamiento j} + \text{Dispersión interacción ij} + \text{Dispersión Residual}$$

La dispersión residual es igual a la dispersión global menos la dispersión del tratamiento i menos la dispersión del tratamiento j menos la dispersión de la interacción ij.

En el diseño factorial se han de calcular tres estadísticos, que luego habrá que comparar con los respectivos valores críticos o en tablas. Las expresiones para calcular los diferentes estadísticos son las siguientes:

DT_i / Grados de libertad del tratamiento i

F (para el tratamiento i) =

DR / Grados de libertad residuales

DT_j / Grados de libertad del tratamiento j

F (para el tratamiento j) =

DR / Grados de libertad residuales

DI_{ij} / Grados de libertad de la interacción ij

F (para la interacción ij) =

DR / Grados de libertad residuales

- Los grados de libertad para la interacción se calculan como: número de elementos que aparecen sumando en la expresión de la dispersión de la interacción menos el número de elementos que aparecen restando en dicha expresión.

Si el valor del estadístico F para el tratamiento i es mayor que el valor crítico para un determinado nivel de confianza se rechazará la hipótesis nula, es decir, la variable independiente tendrá un efecto significativo sobre la variable dependiente. Lo mismo se haría con los otros dos estadísticos, es decir, con el estadístico para el tratamiento j y con el estadístico para la interacción ij.

EJERCICIO: Un investigador está interesado en conocer los efectos de la altura de los estantes y la ocupación de una góndola en las ventas de un producto de consumo enlatado. Para ello realiza un experimento eligiendo al azar 3 supermercados con características muy homogéneas en 6 ciudades de igual importancia. Cada supermercado en cada una de las ciudades pone a la venta el producto combinando dos niveles del frontal con tres tipos de altura de los estantes. Los resultados obtenidos son los siguientes:

	A1 (Pies)	A2 (Manos)	A3 (Ojos)	
Nivel N1	70	85	77	80,7
	75	88	81	
	79	93	78	
Nivel N2	91	94	87	91
	90	97	90	
	87	93	90	
	82	91,7	83,83	85,83

Determinar la influencia de cada una de las variables contempladas en el estudio sobre la ventas.

El diseño a aplicar sería el diseño factorial porque se tiene una variable dependiente (ventas) y dos variables independientes, que son la altura y el frontal. La variable altura tiene 3 niveles (pies: A1, manos: A2, ojos: A3) y la otra variable independiente, el frontal, tiene dos niveles (nivel 1: N1, nivel 2: N2). Además, se nos dice que no difieren en ninguna variable externa importante (supermercados con características muy homogéneas en 6 ciudades de igual importancia). Todos estos motivos nos hacen aplicar el diseño factorial. Lo primero que hay que hacer es construir una tabla que recoja el comportamiento medio de cada interacción.

COMBINACIONES DE LOS TRATAMIENTOS	N1A1	N1A2	N1A3	N2A1	N2A2	N2A3
TOTAL	224	266	236	268	284	267
MEDIA	74,7	88,7	78,7	89,3	94,7	89

$$\text{DISPERSIÓN GLOBAL} = (70 - 85,83)^2 + (75 - 85,83)^2 + (79 - 85,83)^2 + (91 - 85,83)^2 + (90 - 85,83)^2 + (87 - 85,83)^2 + (85 - 85,83)^2 + (88 - 85,83)^2 + (93 - 85,83)^2 + (94 - 85,83)^2 + (97 - 85,83)^2 + (93 - 85,83)^2 + (77 - 85,83)^2 + (81 - 85,83)^2 + (78 - 85,83)^2 + (87 - 85,83)^2 + (90 - 85,83)^2 + (90 - 85,83)^2 = 958,08$$

$$\text{DISPESIÓN TRATAMIENTO i (altura)} = 6 \cdot [(82 - 85,83)^2 + (91,7 - 85,83)^2 + (83,83 - 85,83)^2] = 318,75$$

$$\text{DISPERSIÓN TRATAMIENTO j (frontal)} = 9 \cdot [(80,7 - 85,83)^2 + (91 - 85,83)^2] = 477,41$$

$$\text{DISPERSIÓN INTERACCIÓN ij} = 3 \cdot [(85,83 + 74,7 - 82 - 80,7)^2 + (85,83 + 88,7 - 91,7 - 80,7)^2 + (85,83 + 78,7 - 83,83 - 80,7)^2 + (85,83 + 89,3 - 82 - 91)^2 + (85,83 + 94,7 - 91,7 - 91)^2 + (85,83 + 89 - 83,83 - 91)^2] = 55,47$$

$$\text{DISPESIÓN RESIDUAL} = 958,08 - 318,75 - 477,41 - 55,47 = 106,45$$

FUENTES DE DISPERSIÓN	SUMA DE CUADRADOS	GRADOS DE LIBERTAD	MEDIAS CUADRÁTICAS	VALOR DEL ESTADÍSTICO F
• ALTURA	318,75	3 - 1 = 2	318,75 / 2 = 159,375	
• FRONTAL	477,41	2 - 1 = 1	477,41 / 1 = 477,41	159,375 / 8,871 = 17,97
• INTERACCIÓN	55,47	7 - 5 = 2	55,47 / 2 = 27,735	477,41 / 8,871 = 53,82
• RESIDUAL	106,45	17 - 2 - 1 - 2 = 12	106,45 / 12 = 8,871	27,735 / 8,871 = 3,13
• GLOBAL	958,08	18 - 1 = 17		

Los valores críticos son los siguientes: = 3,89; = 4,75; = 3,89. Como el valor de F para la altura es mayor que el valor crítico (17,97 > 3,89) se rechaza la hipótesis nula, es decir, la variable independiente altura influye significativamente sobre la variable dependiente (ventas). Para el caso del frontal ocurre lo mismo, ya que el valor de F es mayor que el valor crítico (53,82 > 4,75) lo que significa que la variable independiente frontal tiene un efecto significativo sobre las ventas (variable dependiente). Por último, el valor de F para la interacción es menor que el valor crítico (3,13 < 3,89), de forma que no podemos rechazar la hipótesis nula, lo

que quiere decir que no se puede afirmar que el efecto conjunto de las dos variables independientes (altura y frontal) sobre las ventas sea significativo.

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:

- Permite analizar simultáneamente toda la información relativa a un determinado fenómeno.
- Permite analizar un gran número de encuestas.
- 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:

- 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.
- 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:

- Métodos de segmentación.
- 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 multivariantes 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:

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.

$0,9 < \text{Índice de K-M-O} \leq 1$: Muy bueno

$0,8 < \text{Índice de K-M-O} \leq 0,9$: Meritorio

$0,7 < \text{Índice de K-M-O} \leq 0,8$: Bueno

$0,5 < \text{Índice de K-M-O} \leq 0,7$: 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:

- 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. 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.
- 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.

1 F2

-1 1 F1

-1

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 multivariantes 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 grupo 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 cuadrados de las diferencias existentes entre los valores de los atributos para los individuos comparados.

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.

- 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.
- 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.
- 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–Univariados, 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.

TEMA 7: DISEÑO Y PRESENTACIÓN DE INFORMES DE INVESTIGACIÓN.

7.1. TIPOS DE INFORMES DE INVESTIGACIÓN DE MERCADOS.

A la hora de clasificar los tipos de informes se pueden considerar, fundamentalmente, dos criterios:

- Informe escrito.
- Informe oral.

Otra posible clasificación de los tipos de informes de investigación de mercados sería la siguiente:

- Informe técnico: Este tipo de informe va dirigido a un público muy concreto, es decir, va dirigido a personas entendidas en el tema del que se haya realizado el informe.
- Informe divulgativo: En este caso, este tipo de informe iría dirigido al público en general.

Por lo tanto, la diferencia básica entre un informe técnico y un informe divulgativo sería el público al cual van dirigidos cada uno de los dos tipos de informes. Por otra parte, es muy importante señalar que en el caso de que en un informe divulgativo haya que utilizar términos muy técnicos, lo mejor será definirlos en un pie de página, entre paréntesis, etc.

7.2. NORMAS DE REDACCIÓN Y ESTRUCTURA DEL INFORME.

Tanto las normas de redacción como la estructura de un informe están relacionadas con la presentación de informes escritos.

♦ NORMAS

En un informe escrito la normas básicas a tener en cuenta a la hora de su realización son las siguientes:

- La audiencia, es decir, el público al que va dirigido el informe, ya que los términos que se empleen pueden variar según el tipo de personas a quien vaya dirigido el informe. Es fundamental, que se empleen palabras familiares para los lectores del informe y si fuera necesario habría que definir los términos técnicos.
- Tener presente en todo momento cuales eran las necesidades de información. A la hora de realizar un informe escrito se han relacionar en todo momento los resultados obtenidos con los objetivos de quienes van a tomar decisiones a partir de la información elaborada.
- Ser objetivo, es decir, hay que evitar la tentación de reflejar sólo los resultados que nos parecen más aceptables.

♦ ESTRUCTURA O FORMATO DE UN INFORME ESCRITO

- Todo informe escrito debe tener portada en la cual se han de incluir los siguientes datos: título del trabajo; nombre del investigador; fecha y lugar donde se realiza el trabajo. También se podrían incluir, además de los anteriores datos citados, el nombre de la empresa, los logotipos de quien realiza la investigación y de quien la manda, etc. En el caso de que el informe sea confidencial tiene que aparecer esta característica en la portada o en una primera hoja, incluyendo el nombre de aquellas personas a las que se va a entregar el trabajo.
- Índice de contenidos: temas, epígrafes, señalando el número de página en el que se encuentran.
- Resumen del trabajo, que ha de ocupar 1 ó 2 páginas, como máximo 3 y en el que se han de señalar cuales eran los objetivos del estudio, como se diseñó el estudio, la metodología empleada y los resultados obtenidos que sean más relevantes.
- Cuerpo del informe.

– Introducción: naturaleza del problema de decisión o a investigar.

– Metodología: diseño de la investigación, fuentes de información, técnicas de análisis de datos, selección y tamaño de la muestra.

– Resultados: no incluir sólo tablas.

– Conclusiones: valoración crítica y no un simple resumen de los resultados.

– Recomendaciones y limitaciones.

- Referencias bibliográficas.

Ejemplos:

- 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).

Para la presentación de la información en un informe escrito se pueden utilizar gráficos. En las tablas y figuras se han de señalar los siguientes datos:

– Número de la gráfica.

– Título; subtítulos.

– Notas a pie del gráfico (fuentes, aclaraciones, etc.)

– Colocación: lo más cerca posible del sitio en el que se menciona en el texto.

– No saturar el gráfico con datos.

Para la realización de las figuras se pueden utilizar alguno de los tres siguientes diagramas:

1– Diagrama circular: la magnitud de los datos viene dada por el tamaño de la sección (número de divisiones, colores).

2– Diagrama de barras: la magnitud de los datos viene dada por la longitud de la barra (comparaciones

múltiples).

3- Diagramas de líneas.

7.3. PRESENTACIÓN ORAL DEL INFORME.

- Saber a quien va dirigido el informe (analizar su audiencia).
- Verificar todo el equipo y realizar un posible plan de emergencia.
- Practicar la presentación varias veces.
- Vistazo general al empezar.
- Dar la cara a la audiencia en todo momento.
- Hablar a la audiencia.
- Utilizar ayudas visuales.
- Preguntar a la audiencia si tienen preguntas.

ENTORNO DE

MARKETING

MACROENTORNO

MICROENTORNO

DIRECCIÓN DE

MARKETING

FUNCIONES DE MARKETING ESTRATÉGICO

FUNCIONES DE MARKETING OPERATIVO

SISTEMA DE INFORMACIÓN DE MARKETING

APOYO DE

MARKETING

INVESTIGACIÓN DE MERCADOS

DATOS

INTERNOS

INTELIGENCIA

DE MARKETING

INFORMACIÓN SECUNDARIA

- Interna
- Externa

INFORMACIÓN PRIMARIA

- Cualitativa
- Cuantitativa

5

7

8

9

13

13

15

17

21

21

25

30

34

39

42

43

47

53

55

81

81

83

71

72

75

78

1.000.000

SI

NO

100.000

400.000

250.000

250.000