

TÉCNICAS DE ANÁLISIS DE DATOS PARA INVESTIGADORES SOCIALES

Presentamos una obra en la que se exponen algunas de las técnicas de análisis de datos más utilizadas en la investigación social mediante encuesta, contextualizadas dentro de un diseño de investigación concreto. El objetivo de este trabajo no se centra tanto en una exposición de distintas técnicas de análisis de datos, sino que se explica cómo solucionar problemas de una investigación utilizando determinadas técnicas de análisis de datos. Con este fin se plantea una investigación completa, desde la formulación de hipótesis hasta la exposición de resultados y generalización de los mismos. Este proceso de investigación requerirá utilizar técnicas de análisis de datos, oportunidad que aprovecharemos para exponer el uso y la interpretación de cada una de éstas.

El elevado número de técnicas de análisis de datos utilizadas en la investigación social aconseja realizar una selección a fin de explicar algunas de ellas en profundidad. Uno de los criterios más importantes para diferenciar entre las técnicas de análisis de datos hace referencia al tipo de métrica de las variables utilizadas; cualitativas y cuantitativas. El hecho que la mayor parte de la investigación que se realiza en ciencias sociales utilice variables cualitativas, unido a la escasa atención que han recibido éstas en los manuales de investigación social nos anima a centrarnos en las técnicas de análisis de datos para variables cualitativas.

Antes de continuar deseamos realizar algunas consideraciones sobre el contenido de esta obra. Señalar en primer lugar que es un libro de *Técnicas de Investigación Social* ya que nuestro interés es el planteamiento y posterior desarrollo de una investigación completa, investigación que requiere utilizar procedimientos de análisis de datos, como también precisa la elaboración de unas hipótesis de trabajo, un análisis de los datos a recoger, construcción de un cuestionario y prueba del mismo, etc. Dentro del análisis de datos nuestro interés se centrará en razonar porqué se utiliza una determinada técnica y como interpretar los resultados que proporciona. No nos preocupa la comprensión de fórmulas complejas, que además trataremos de evitar en la medida de lo posible, sino que deseamos profundizar en la utilización de determinadas técnicas y en la interpretación de los resultados que proporcionan.

El enorme desarrollo de los paquetes estadísticos, unida a la amplia difusión de ordenadores personales requiere prestar una mayor importancia a los criterios de utilización de las técnicas, y a la interpretación de los resultados proporcionados. En este trabajo se ha elegido uno de los programas estadísticos más utilizados por los investigadores sociales de todo el mundo, el SPSS (versión 8) para windows. Las razones que han motivado esta elección son, entre otras, la facilidad de su uso, la explicación de determinados procesos estadísticos en sus menús de ayuda, y la existencia de una versión para estudiantes que posibilita -a un precio muy asequible- que éstos puedan adquirirlo y familiarizarse con él. Deseamos volver a incidir que no es la explicación del programa estadístico la que guiará la exposición, sino que serán las necesidades de la investigación la que aconseje la utilización -y explicación- de determinados procesos de análisis de datos. Respecto a la utilización del programa estadístico, es preciso indicar que todas las menciones a cuadros de diálogo del SPSS, recuadros, etc. aparecerán entrecomillados y en letra cursiva.

La novedad de este texto, a nuestro juicio, está en la carencia en el mercado de una obra que explique a fondo el proceso completo de una investigación social. Pese a que en los últimos años se ha producido una proliferación de publicaciones sobre la utilización de programas estadísticos, así como otros textos que utilizan ejemplos donde se exponen los primeros pasos en el diseño de la investigación (entre otros Soler y Perdiguer 1992, Abascal y Grande 1994, etc.), el texto que aquí presentamos unifica ambos planteamientos al exponer problemas concretos del diseño de la investigación social y las dificultades del análisis de datos. Creemos que hay una importante carencia en el mercado español de un texto que considere conjuntamente estos planteamientos: por un lado los criterios metodológicos de la investigación social, por otro el análisis de datos reales de una investigación, y por último la capacidad de realizar conclusiones y análisis "especiales" utilizando las propiedades de los paquetes estadísticos.

En cuanto a su estructura este libro está dividido en dos partes; 1) *Del problema social al planteamiento metodológico* y 2) *De la metodología al análisis de datos*. La primera parte comienza con una breve exposición del proceso de investigación social, proceso que será utilizado a la hora de plantear los diferentes pasos de nuestra investigación. Una justificación metodológica y la contextualización teórica del fenómeno a investigar permitirá formular las hipótesis de trabajo que guiarán la investigación. Este primer capítulo termina con una exposición de la lógica del análisis de datos en la investigación social.

Planteada la investigación, será necesario elaborar un sistema de medición que permita *verificar* la adecuación de las hipótesis planteadas en la realidad objeto de estudio. La aplicación de un sistema de medición en una investigación social requiere tener en cuenta los problemas unidos a la medición de hechos sociales, así como una evaluación de estas mediciones analizando la fiabilidad y validez de las mismas. Al final del segundo capítulo se expondrán los criterios que han guiado la construcción del cuestionario utilizado en la presente investigación.

El capítulo tres introduce al lector en la tarea de seleccionar un programa estadístico de análisis de datos, al tiempo que se realizan unas pequeñas explicaciones del programa estadístico elegido. El último capítulo de la primera parte está dedicado a la preparación de los datos para ser procesados con un paquete estadístico. Desde que los datos son recogidos hasta que es posible trabajar con ellos es preciso realizar una serie de verificaciones y comprobaciones de la calidad de los mismos. Este capítulo termina explicando las características de los datos recogidos con el cuestionario utilizado.

Estos cuatro capítulos preparan al lector para la exposición de la segunda parte de esta obra, *de la metodología al análisis de datos*, en la que se exponen diversas técnicas de análisis de datos según el número de variables analizadas: técnicas univariantes, bivariantes y multivariantes. El primer capítulo está dedicado a la descripción univariante de cada una de las variables que forman las hipótesis del trabajo. Aprovecharemos determinadas características de las hipótesis para explicar diversas estrategias de transformación de datos, utilización de gráficos, etc.

En el capítulo seis se exponen diversas técnicas que se ocupan de la relación entre variables, aspecto crucial para solucionar determinadas hipótesis. La relación entre dos variables es también el objetivo del capítulo siete, si bien aquí el interés se centra en la localización y cuantificación de la relación entre las variables.

El análisis de dos variables puede generar, en ocasiones, que se interpreten relaciones entre variables que no existen, al estar generadas por la influencia de terceras variables (influencias espúreas). La localización y eliminación de estas influencias permitirá definir con precisión los factores que inciden en cada una de las conductas planteadas. Más de dos variables utiliza también la técnica expuesta en el penúltimo capítulo, si bien el objetivo de ésta es conocer como determinadas variables permiten "predecir" los resultados de una variable dependiente. El capítulo diez pondrá fin a este libro, exponiendo las conclusiones del trabajo y animando a continuar la investigación utilizando otras técnicas de análisis de datos.

En cuanto al público para el que se ha escrito de este libro, y aunque en un primer momento fue realizado pensando en utilizarlo como manual en un curso de Técnicas de Investigación Social, las sucesivas modificaciones realizadas en los últimos años –tratando de hacer más fácil la comprensión del mismo- lo hacen muy apropiado para todo tipo de personas interesadas en la investigación mediante encuesta: sociólogos, trabajadores sociales, investigadores de mercado, psicólogos, economistas, etc. El libro está escrito en un lenguaje claro y sencillo, y considera que el lector tiene algunos conocimientos previos sobre investigación social; principalmente sobre muestreo y estadística básica. No obstante, dejar claro que este libro es para gente que se inicia, no para especialistas en la materia.

TÉCNICAS DE ANÁLISIS DE DATOS PARA INVESTIGADORES SOCIALES

PARTE PRIMERA:

DEL PROBLEMA SOCIAL AL PLANTEAMIENTO METODOLÓGICO.

1.- PLANTEAMIENTO DE LA INVESTIGACIÓN.

- 1.- El proceso de investigación social.
- 2.- Justificación metodológica.
- 3.- Contextualización teórica del fenómeno a investigar.
- 4.- Elaboración de las Hipótesis de trabajo.
- 5.- La lógica del análisis de datos.

2.- LA MEDICIÓN DE HECHOS Y FENÓMENOS SOCIALES.

- 1.- Problemas inherentes al hecho de medir. Los hechos sociales y su medición.
- 2.- Variables y escalas de medida. Tipos de escalas.
- 3.- Fiabilidad y validez de instrumentos de medida.
- 4.- Elaboración de un cuestionario sobre hábitos de consumo.

3.- INTRODUCCIÓN A LA ESTADÍSTICA MEDIANTE ORDENADOR.

- 1.- Tipos de programas de análisis de datos.
- 2.- Justificación del paquete estadístico utilizado.
- 3.- Normas y rutinas elementales de funcionamiento del programa SPSS para Windows.
 - A.- Operaciones con archivos de datos.
 - B.- Ventana del menú principal.
 - C.- Ventana de sintaxis.
 - D.- Ventana de resultados.

4.- DE LAS OPINIONES SOCIALES AL PROCESAMIENTO DE DATOS.

- 1.- "Preparando" los datos.
- 2.- Codificación, grabación y edición de datos.
 - A.- Definición de variables.
 - B.- Introducción de datos.
 - C.- Edición de datos y modificación de variables.
- 3.- Procesos de depuración de la información recogida.
 - A.- Estrategias de depuración de datos.
 - B.- Imputación.
- 4.- Características del fichero de datos utilizado.
 - A.- La población a estudiar, método de muestreo y características muestrales
 - B.- Ponderación y errores de muestreo.

PARTE SEGUNDA: DE LA METODOLOGÍA AL ANÁLISIS DE DATOS.

5.- CÓMO REALIZAR UN PRIMER ACERCAMIENTO A LOS DATOS: ANÁLISIS UNIVARIANTE.

- 1.- La compra lúdica y sus componentes.
- 2.- El proceso de compra fruto de una necesidad.
 - 2.A.- Estadísticos de una variable.
 - 2.B.- Reagrupación de categorías.
- 3.- ¿Se compran objetos que no son utilizados?. ¿Cuál es la razón por la que se dejan de usar los productos adquiridos?.
 - 3.A.- Utilización de gráficas: opción 1.
 - 3.B.- Utilización de gráficas: opción 2.
 - 3.C.- Selección de determinados sujetos.
- 4.- Colectivos con una mayor presencia del "nuevo" consumidor.
 - 4.A.- Características sociodemográficas de los entrevistados: edad y nivel educativo.
 - 4.B.- Construcción de índices y elaboración de variables complejas.
- 5.- Sumario de los procedimientos de análisis de Datos utilizados.

6.- RELACIONES ENTRE VARIABLES: ANÁLISIS BIVARIANTE (PARTE I).

- 1.- De la muestra a la población: Inferencia y significación estadística.
 - 1.A.- Distribuciones de probabilidad.
 - 1.B.- Prueba de significación para variables de intervalo.
 - 1.C.- Pruebas de significación para variables nominales y ordinales.

- 2.- El proceso de compra producto de una necesidad: Hipótesis II.
 - 2.A.- Interpretación de los valores de la tabla utilizando los residuos.
 - 2.B.- Odds ratio, Razón de Razones.
 - 2.C.- Cálculo y diferencia de porcentajes.
 - 2.D.- Otras medidas para analizar la relación entre variables.
 - 2.E.- Representaciones gráficas de tablas de contingencia.
 - 2.G.- Conclusión de la hipótesis.
- 3.- Sumario de los procedimientos de análisis de Datos utilizados.

7.- ANÁLISIS BIVARIANTE (PARTE II).

- 1.- Rasgos de los consumidores que "disfrutan" en el acto de compra.
 - 1.A.- Relación entre variables nominales utilizando el Ji-Cuadrado.
 - 1.B.- Limitaciones en la utilización del Ji-Cuadrado.
 - 1.C.- Estadísticos basados en el Ji-Cuadrado.
 - 1.D.- Medidas basadas en la reducción del error de predicción.
 - 1.E.- Resolución de la hipótesis.
- 2.- Probarse modelos en la compras de ropa y calzado como un indicador del "disfrute" en el acto de compra.
- 3.- Sumario de los procedimientos de análisis de datos utilizados.

Anexo al capítulo 7: Comprobación de los índices elaborados: Posición Social y Nivel de Equipamientos.

- A.- Relación entre variables ordinales.
- B.- Evaluación de los índices construidos.
- C.- Introducción al cálculo de pares.

8.- ANÁLISIS MULTIVARIANTE DE TABLAS DE CONTINGENCIA: TABLAS DE CONTINGENCIA CON TRES y MÁS VARIABLES

- 1.- Análisis de tres variables para detectar nuevas influencias entre ellas.
- 2.- Crítica a las relaciones entre dos tablas expuestas en el capítulo anterior: Colectivos con una mayor presencia del "nuevo" consumidor.
- 3.- Construcción de tipologías mediante la segmentación de variables discretas.
- 4.- Sumario de los procedimientos de análisis de datos utilizados.

9.- REGRESIÓN LOGÍSTICA.

- 1.- Razones por las que se dejan de usar determinados productos.
 - 2.- Regresión Simple.
 - 2.A.- Ajuste del modelo.
 - 2.B.- Interpretación de los coeficientes.
 - 3.- Regresión Múltiple.
 - 3.A.- Especificación del modelo de regresión.
 - 3.B.- Interpretación de los coeficientes.
 - 4.- Resolución de la hipótesis.
- Sumario de los procedimientos de análisis de datos utilizados.
- Anexo al capítulo 9: Regresión Logística "por etapas".

10.- CONCLUSIÓN.

- 1.- La emergencia del nuevo consumidor
- 2.- Técnicas de análisis de datos utilizadas.
- 3.- Seguir investigando: otras técnicas de análisis de datos.