

**METODOLOGÍA CUANTITATIVA:
ESTRATEGIAS Y TÉCNICAS
DE INVESTIGACIÓN SOCIAL**

M.^a Ángeles Cea D'Ancona

LA ORGANIZACIÓN DE LA INVESTIGACIÓN

Una vez definido el *método científico* y sus condicionantes, corresponde ahora pasar del proceso de investigación “ideal” a la práctica habitual. Esta se halla metodológicamente determinada por los *objetivos* y el *marco teórico* del estudio. Pero también intervienen aspectos económicos (los *recursos* a disposición del investigador) y temporales (el plazo de *tiempo* concedido para su materialización).

Todos estos aspectos marcan el desarrollo de la investigación, desde su configuración inicial en un *proyecto de investigación*. Concretamente, afectan al *diseño*: a la selección de *estrategias*, con las correspondientes *técnicas* para la recogida y el análisis de la información.

Pero, ante todo, el *proceso de investigación* no ha de concebirse como algo fijo e inmutable. Aunque el *proyecto* haga una expresa relación de las actuaciones para el cumplimiento de cada fase de la indagación, su puesta en práctica puede llevar a la modificación –a veces sustancial– del diseño original. Razón por la cual se recomienda la previsión (en el *proyecto de investigación*), de posibles incidencias en su ejercicio, así como *sesgos* previsibles que pudieran invalidar los resultados del estudio. Éste es el *desideratum*.

3.1. El proyecto de investigación

Dos conceptos parejos suelen, a veces, confundirse: el de *diseño* y el de *proyecto*. Si bien, este último constituye un concepto más amplio que el anterior.

En el ámbito de la *investigación social*, el concepto de *diseño* se restringe a la *planificación* de la *estrategia* para alcanzar los *objetivos* de la investigación. Alvira (1989: 85) lo define como un “plan global de la investigación que integra, de modo co-

herente y adecuadamente correcto, técnicas de recogida de datos a utilizar, análisis previstos y objetivos”.

Igualmente, Yin (1989) lo describe como una secuencia lógica que relaciona los datos empíricos del estudio con las preguntas formuladas al inicio de la investigación y, en último término, con las conclusiones. “Se trata de un problema lógico, no logístico, y por tanto es más que un plan de trabajo” (Yin, 1989: 29).

La función esencial del *diseño* es guiar al investigador en la obtención y posterior análisis de la información. El *proyecto* se configura, en cambio, como un conjunto de elementos entre los cuales se encuentra el mismo *diseño*.

Además de englobar al *diseño*, el *proyecto* abarca otros aspectos varios: desde la idea originaria de la investigación y su concreción en objetivos específicos, hasta el detalle económico, organizativo y temporal de las diferentes fases del proceso de investigación.

El origen de la confusión entre estos dos conceptos se halla –como bien explica Alvira (1989: 17)– en el hecho de que la elección y elaboración de un *diseño de investigación* “necesita contemplar todos los extremos que un *proyecto* presenta”. Razón que justifica que ambos términos con frecuencia se “confundan a efectos prácticos”.

En resumen, el *proyecto de investigación* se hace en referencia a tres elementos claves:

- a) Los *objetivos* de la investigación.
- b) Los *recursos* o medios materiales, económicos y humanos de que disponga el investigador para hacer viable el estudio.
- c) El *tiempo* concedido para su realización.

La conjunción de estos elementos marcará el alcance y desarrollo de la investigación, por su incidencia directa en el *diseño de la investigación*.

En el Cuadro 3.1 se resumen los componentes esenciales de un *proyecto de investigación*. Asimismo, en la Figura 3.1 (al final de la presente sección) se representa, de forma gráfica, la configuración del *proyecto de investigación*. A continuación se incluye el detalle de cada uno de estos integrantes.

3.1.1. La formulación del problema de investigación

“No importa en qué problema quieras trabajar y no importa qué método eventualmente vayas a utilizar, tu trabajo empírico deberá comenzar con una consideración cuidadosa del problema de investigación” (Simon, 1978: 98).

En concreto, ha de exponerse:

- a) Qué es exactamente lo que se pretende analizar: cuáles son los *objetivos generales* y *específicos* de la investigación.

- b) Qué *razones* motivaron la elección de ese *problema* concreto.
- c) La *justificación* de dicha elección, relacionándola con la relevancia y trascendencia de la indagación.

Esta exposición de *razones* y *justificaciones* ha de hacerse de forma convincente, en aras de lograr la aprobación del *proyecto de investigación*, y su consiguiente financiación económica (en caso de solicitarse).

Al principio, la *idea* o *problema de investigación* se presenta de forma vaga y genérica. Pero, paulatinamente, ira concretándose, conforme el investigador vaya ahondando en ella.

La *fente originaria de la idea* puede ser muy dispar. A veces es la propia entidad o institución que financia la investigación quien propone el *problema* o *tema de estudio*. En otras ocasiones, es el azar quien suscita la *idea*: en el curso de una conversación; durante la lectura de un libro o artículo de prensa (periódico, revista); o tras la observación de un acontecimiento, de forma directa o indirecta (por medio de la televisión, la radio), entre otras posibles *fuentes originarias de ideas*.

Pero la *idea* también puede ser el resultado de una reflexión profunda, presente durante mucho tiempo en la mente del investigador, aunque hasta entonces no haya sido concretada ni operacionalizada en un *proyecto de investigación* específico.

Al principio, sólo es eso: una mera *idea*. El investigador, posteriormente, tendrá que ir precisándola y configurándola. Para ello deberá adentrarse en el campo de conocimiento determinado donde se ubique la *idea* en cuestión.

CUADRO 3.1. Componentes esenciales de un *proyecto de investigación*.

<ol style="list-style-type: none"> 1. La <i>formulación del problema de investigación</i>: <ul style="list-style-type: none"> • Definición de objetivos. • Elaboración del <i>marco teórico</i> de la investigación: revisión bibliográfica y demás indagaciones exploratorias. 2. La <i>operacionalización del problema</i>: <ul style="list-style-type: none"> • Formulación de hipótesis. • Operacionalización de conceptos teóricos. • Delimitación de las unidades de análisis. 3. El <i>diseño de la investigación</i>: <ul style="list-style-type: none"> • Selección de <i>estrategias</i>. • Diseño de la <i>muestra</i>. • Elección de <i>técnicas</i> de recogida y de análisis de datos. 4. La <i>factibilidad de la investigación</i>: <ul style="list-style-type: none"> • Fuentes bibliográficas. • Recursos disponibles (materiales y humanos). • Recursos necesarios (económicos, materiales, humanos). • Planificación del tiempo de realización.

A tal fin se recomienda efectuar una *revisión bibliográfica* exhaustiva sobre el *tema* concreto de interés y otros similares que, aunque no se refieran al *problema* específico del estudio, sí puedan orientar en su indagación. Esta última *revisión de temas* relacionados con el *problema de investigación* adquiere una mayor relevancia cuando apenas se encuentra *bibliografía* (teórica y empírica) específica sobre dicho tema, bien debido a su carácter novedoso, o bien a su escasa repercusión en la tradición investigadora anterior.

La importancia de llevar a cabo una *revisión bibliográfica* exhaustiva es destacada por la mayoría de los autores. Incluso llega a afirmarse que “todo proyecto de investigación en las ciencias sociales debería implicar la búsqueda de indagaciones previas” (Cooper, 1984: 7).

Con la *revisión bibliográfica* se busca la familiarización con el *tema de estudio* escogido: hallando, leyendo, evaluando y sintetizando indagaciones realizadas con anterioridad. De ellas se extraerá información concerniente a:

- a) La *metodología* seguida y los *resultados* que, mediante ella, se lograron: sus potencialidades y limitaciones concretas.
- b) Aspectos del *problema* tratados, aquellos no abordados y los suscitados tras la conclusión de la investigación, como relevantes para una futura indagación.

El propósito de dicha *revisión bibliográfica* Dankhe (1989) lo resume en uno fundamental: obtener información para, a partir de ella, poder analizar y discernir si la *teoría* existente y la *investigación* previa sugieren una respuesta (aunque sea parcial) a las *preguntas de investigación*, o una dirección a seguir dentro del estudio. Más concretamente, la *revisión bibliográfica* contribuye a:

- a) La estructuración de la *idea* originaria del *proyecto de investigación*, contextualizándola en una *perspectiva teórica y metodológica* determinada.
- b) *Inspirar cómo diseñar la investigación*, a partir de la experiencia obtenida en la utilización de determinadas *estrategias de investigación*. Se trataría, por tanto, de:
 - 1) *Prevenir* posibles *errores* cometidos en estudios anteriores al presente.
 - 2) *No ser redundantes*. La nueva *investigación* debería dar un enfoque distinto al proporcionado en indagaciones anteriores. Por ejemplo, escogiendo una *estrategia de investigación* diferente, con el propósito de comprobar si con ella se alcanzan los mismos resultados.

También, se trataría de plantear otros interrogantes de investigación, con objeto de cubrir aspectos menos observados, y no aquéllos continuamente analizados.

Además de la necesaria *revisión bibliográfica*, la estructuración de la *idea de investigación* requiere el complemento de otras *indagaciones exploratorias*. Éstas pueden materializarse en:

- a) *Discutir el tema con otros investigadores* que puedan aportar *ideas* de cómo llevar a cabo la investigación y qué pasos seguir.
- b) *Entrevistar a algunos de los implicados* en el *problema* a investigar, con el propósito de recabar, de ellos, información que ayude al *diseño de la investigación*.

Todas estas tareas previas en el *proyecto de investigación* (desde la *revisión bibliográfica* a las *entrevistas abiertas*) se convierten en cruciales, cuanto más vaga se muestre la *idea* originaria del estudio, y más desconocedor sea el investigador del área de conocimiento específica donde ésta se ubique.

A partir de la *revisión bibliográfica* y otras *indagaciones exploratorias* (que el investigador opte por seguir), se configurará el *marco teórico* de la investigación. Éste se basará en la integración de toda la información previa que el investigador recopile. Además, le proporcionará un *marco de referencia* al que acudir en la interpretación de los resultados que obtenga tras la conclusión de la investigación.

A veces se partirá de una *teoría*; otras veces, se decidirá imbricar el *problema* de estudio en varias *perspectivas teóricas* para, de este modo, poder proporcionar cobertura a los distintos aspectos en que el *problema* se manifiesta (máxime cuanto más complejo sea el *tema de la investigación*). Pero, en más ocasiones de las deseables, el *marco teórico* que configura el *proyecto de investigación* se fundamenta en *generalizaciones empíricas: proposiciones* que han sido comprobadas en anteriores investigaciones, si bien no han logrado cristalizar en una *perspectiva teórica* determinada.

EJEMPLO DE CONFIGURACIÓN DE UN PROBLEMA DE INVESTIGACIÓN

Como ilustración de lo expuesto hasta ahora, piénsese en una *idea* que se desee desarrollar en un *proyecto de investigación*. Cualquier hecho o acción social puede convertirse en un *problema de investigación*.

La “violencia callejera juvenil” constituye uno de los fenómenos sociales más presentes en la cotidianidad de la sociedad actual. Al menos es, junto con el problema de la corrupción, uno de los más comentados en los informativos (de radio, televisión, prensa escrita), y frecuente tema de coloquio.

El conocimiento, de manera directa o indirecta, de algún acto nuevo de violencia juvenil puede despertar el interés de algún investigador deseoso de encontrar respuestas a *interrogantes* suscitados ante dichos sucesos: ¿por qué los jóvenes actuaron de la forma que lo hicieron?; ¿qué factores inciden en la comisión de actos de violencia?; ¿incita la sociedad actual a la violencia?; ¿qué características diferencian a los jóvenes violentos de los no violentos?; ¿promueve la frecuente aparición en televisión de actos violentos conductas de imitación?

Lo primero que el investigador ha de hacer es formular el *problema de investigación*, inicialmente titulable “la violencia callejera juvenil”. Para ello, deberá centrarse en *tres tareas principales*:

a) *Establecer los objetivos generales y específicos del estudio:*

— *Objetivo general:* analizar las causas de la violencia juvenil.

— *Objetivos específicos:*

- Describir las características sociodemográficas y psicosociológicas del joven violento.
- Indagar en la trayectoria vital del joven (afectiva, laboral, ocupación del tiempo).
- Analizar su ambiente sociofamiliar (familia, amigos, barrio).

b) *Delimitar el tema de estudio* (qué actos de violencia se analizarán) y las *unidades de observación* (qué tipo de jóvenes, de qué edades, de qué ámbito territorial).

c) *Elaboración del marco teórico de la investigación:* qué enfoque teórico y metodológico se dará al análisis de la violencia juvenil.

Para que esta fase inicial y esencial en el *proyecto de investigación* se efectúe de forma rigurosa, el investigador deberá precederla de una *indagación exploratoria*. Ésta se convertirá en referente del *diseño de investigación*. Concretamente:

a) Una *revisión bibliográfica* de las *perspectivas teóricas* existentes e *investigaciones empíricas* llevadas a cabo en España y en otros países (que el investigador especificará) acerca de "la violencia juvenil".

Esta tarea ardua en el pasado se ve, en el momento presente, bastante facilitada por el uso generalizado de la *búsqueda bibliográfica por ordenador mediante CD-ROM*. Este proporciona información existente en bibliotecas y centros de datos de distintos países, en un breve período de tiempo.

b) *Entrevistas abiertas a expertos* en la materia y a *protagonistas* de acciones de violencia (jóvenes implicados en actos de violencia callejera y personas que han sido víctimas de dichos actos), que constituyen la *población de estudio*.

Con la *indagación exploratoria* se busca familiarizarse con el *tema de la investigación* y con la *metodología* a seguir en su análisis. De ahí que estas primeras tareas resulten esenciales para realizar un buen *diseño de investigación*. De lo exhaustivo que se sea en esta fase previa del *proyecto de investigación* dependerá el buen desarrollo de la misma.

3.1.2. La operacionalización del problema de investigación

Después de la definición y delimitación del *objeto de estudio*, procede su concreción, su operacionalización, tanto en su *vertiente teórico-analítica* (principales *conceptos*, *categorías analíticas* y *variables* a analizar) como *poblacional* (qué *población* se estudiará y, en su caso, qué *diseño muestral* se efectuará).

Del *marco teórico* de la investigación se extraen, mediante un proceso deductivo, unas *hipótesis* que representen respuestas probables que el investigador avanza a las *preguntas* iniciales de la investigación. Las *hipótesis* vienen expresadas en forma de *proposiciones*,

en las que se afirma la existencia o inexistencia de asociación esperada, al menos entre dos variables (dependiente e independiente, generalmente), en una situación determinada.

Se recomienda que las *hipótesis* se encuentren relacionadas con los *objetivos de la investigación* y sean lo más concretas y precisas posibles.

EJEMPLOS DE FORMULACIÓN DE HIPÓTESIS

- Los jóvenes que en su infancia han sido víctimas de malos tratos presentan, en su juventud, actitudes más violentas que aquéllos que no los han padecido.
- El consumo de bebidas alcohólicas aumenta la probabilidad de cometer actos de violencia.

Las *hipótesis* contienen *conceptos* que determinan lo que debe analizarse. Para ello las *hipótesis* han de traducirse en *variables empíricas* o *indicadores*. A este proceso se le denomina, normalmente, *operacionalización de conceptos teóricos*. Ésta constituye una fase intermedia que vincula la *formulación teórica del problema* con la *medición de las variables* relevantes. En el Capítulo 4 se tratarán con más detalle el *proceso de medición* y de *operacionalización de conceptos*.

EJEMPLO DE OPERACIÓN DE CONCEPTOS TEÓRICOS

Se parte de la *hipótesis* siguiente:

“Los actos de violencia juvenil callejera crecen, conforme aumenta la insatisfacción social en los jóvenes.”

Esta *hipótesis* contiene *conceptos teóricos* que deberán operacionalizarse (traducirse en *variables empíricas* o *indicadores*), antes de proceder a la recogida de información.

Uno de los *conceptos* es el de “insatisfacción”. Habría que seccionarlo en las diferentes *dimensiones* que comprende: insatisfacción laboral, educativa, afectiva, económica, familiar, entre otras posibles.

Para cada una de estas *dimensiones* se buscarán *indicadores* que ayuden en su concreción. Por ejemplo, la “satisfacción laboral” podrá medirse por los siguientes *indicadores*:

- Situación laboral: activo, en paro.
- Tipo de contrato: fijo, eventual, a tiempo completo, a tiempo parcial.
- Trabajo que realiza en la empresa: si corresponde o no a su formación académica y técnica.
- Remuneración que percibe.

- Relación con los compañeros.
 - Relación con el jefe o personal a cuyas ordenes se halle.
 - Ubicación de la empresa: entorno físico dentro y fuera de la empresa.
-

Además de las *variables*, en las *hipótesis* también se hace mención (aunque genérica) de las *unidades de análisis*: la *población* o *universo* objeto de estudio.

La elección de una *población* u otra puede tener implicaciones significativas, tanto para el contenido sustantivo de la investigación como para sus costes (económicos y de tiempo). Por esta razón, autores como Hakim (1994) recomiendan que la *población de estudio* se establezca en los *objetivos de la investigación*, debido a que ésta determina la información a extraer y la naturaleza de los resultados de la investigación.

Las *unidades de análisis* no sólo incluyen el *ente* individual o colectivo que se observará, sino también el *espacio* donde se ubica y el *tiempo*. En los *estudios longitudinales* el *tiempo* figura (explícitamente) como una de las *variables* en la investigación.

3.1.3. El diseño de la investigación

Una vez que se ha formulado y operacionalizado el *problema de investigación* (concreción de los *objetivos*, las *hipótesis*, las *variables* y las *unidades de análisis*), corresponde decidir el *diseño de la investigación*: especificar cómo va a realizarse la investigación.

Esta fase es clave dentro del *proyecto de investigación* y, al igual que las precedentes, vendrá determinada por los *objetivos del estudio* (exploratorios, explicativos, predictivos, evaluativos), los *recursos* (materiales, económicos, humanos) y el plazo de *tiempo* disponible para su materialización. De ahí que el investigador deba siempre tener en mente –al ocuparse del *diseño*– no sólo los *objetivos* del estudio, sino también la *factibilidad del diseño* elegido. Ha de procurar que el *diseño* sea viable, con los recursos disponibles y dentro de los límites de tiempo fijados.

Esta triple consideración de *objetivos*, *recursos* y *tiempo* incide, directamente, en la selección de una o varias *estrategias de investigación*: uso de *fuentes documentales* y *estadísticas*, *estudio de casos*, *encuesta*, *experimento*.

La *estrategia* (o *estrategias*) finalmente elegida influirá en:

- a) El *diseño muestral*: la *muestra* del estudio (individuos, viviendas, entidades sociales, acontecimientos, documentos), su volumen y forma de selección.
- b) Las *técnicas de recogida de información* (de documentación, observación y entrevista). Concretamente:
 - 1) *Revisión de fuentes de observación secundaria* (estadísticas y documentos).
 - 2) *Observación sistemática, participante y no participante*.
 - 3) *Entrevistas abiertas, semi o no estructuradas* (individuales y/o grupales).

- 4) *Relatos biográficos* (múltiples, paralelos, cruzados) y *documentos personales*.
 - 5) *Cuestionario estandarizado* (sondeos en serie, tipo panel, macroencuestas).
- c) *Las técnicas de análisis de datos:*
- 1) *Documental*.
 - 2) *Estadístico* (univariable, bivariable, multivariable).
 - 3) *Estructural* (análisis del discurso, etnografía, fenomenografía).
 - 4) *Interpretacional* (construcción de teorías, descriptivo/interpretativo).
 - 5) *De contenido* (cuantitativo y cualitativo).

Las *técnicas cuantitativas de recogida y de análisis de datos* constituyen el eje de este manual. En capítulos sucesivos se expondrá cada una de ellas. Para el conocimiento, en cambio, de las *técnicas cualitativas* remito al lector interesado al manual de Miguel S. Valles Martínez, publicado en esta misma colección.

En suma, en la selección de la *estrategia* y de las *técnicas* (de recogida y de análisis de datos) correspondientes intervienen los siguientes elementos:

- a) Los *objetivos de la investigación* y si ésta se ceñirá a un momento temporal concreto (*diseños seccionales o transversales*) o si, por el contrario, incluirá diferentes períodos de tiempo (*diseños longitudinales*).
- b) El *grado de conocimiento*, por el investigador, de la diversidad metodológica existente.
- c) El *grado de rigor y precisión* que el investigador desee para su indagación.
- d) La *factibilidad de la investigación*. Ésta se concreta, siguiendo a Pons (1993: 24), en los factores siguientes:
 - 1) “La capacidad del personal participante en la investigación.
 - 2) La asequibilidad y corrección de las fuentes de información necesarias.
 - 3) El nivel de operacionalidad.
 - 4) El tiempo que se requiere.
 - 5) El coste: humano, social y económico.
 - 6) Las limitaciones deontológicas.”

Pero la práctica real de la investigación suele ser muy dispar. De acuerdo con Ibáñez (1989), el proceso de selección se simplifica bastante. El investigador social suele elegir, sin pensar demasiado en la elección, la *técnica* que tiene más a mano, bien por:

- a) *Razones personales*: uno es experto en esa técnica.
- b) *Razones organizativas*: uno trabaja en una organización constituida para trabajar con esa técnica.
- c) *Razones institucionales*: uno pertenece a una institución interesada en vender esa técnica.

Sea como fuere el proceso de selección, el investigador deberá justificar (en su *proyecto de investigación*) por qué eligió unas *estrategias y técnicas* con preferencias a otras. La exposición de razones suele fundamentarse en tres aspectos esenciales: los *objetivos*, las *características*, y las *condiciones de realización* de la investigación.

Además, téngase presente que, en el curso de la investigación, todo *diseño* inicial puede verse alterado, ante las dificultades que pueda suscitar su puesta en práctica. De ahí la conveniencia de que el investigador prevea, en la medida de lo posible, alternativas al *diseño* original de la investigación.

3.1.4. La factibilidad de la investigación

La fase final del *proyecto* incluye la exposición de las condiciones mínimas requeridas para hacer viable, o factible, la investigación diseñada. Estas condiciones o requisitos cabe resumirlos en los apartados siguientes:

a) Fuentes

En el *proyecto de investigación* suele adjuntarse una *bibliografía* básica, especializada en el tema que constituye el eje de la investigación, tanto en su vertiente teórica como en la empírica y metodológica.

Consistirá en una selección de *obras clave* y de publicaciones actualizadas.

b) Recursos (materiales y humanos) disponibles

El equipo investigador acredita su *capacidad* (experiencia) y *medios* para cumplir el *proyecto de investigación*.

Al *proyecto* se añade, al menos, el *currículum* del investigador principal o de los integrantes del equipo investigador.

En el *currículum* ha de incluirse el detalle de la *experiencia* tenida en el área de conocimiento donde se inserta el *problema* o *tema de investigación*.

c) Otros recursos (materiales, económicos y humanos) necesarios

Si se está solicitando financiación económica de la investigación, en el *proyecto* también habrá de especificarse la *cantidad* necesaria para cada una de las *partidas* siguientes:

- 1) *Personal* (disponible y el que habría de contratarse al efecto).
- 2) *Equipo material inventariable*.

- 3) *Material fungible*: fotocopias, teléfono, correos.
- 4) *Dietas y desplazamientos*.

d) *Planificación del tiempo de ejecución de la investigación*

El investigador también ha de concretar la *cronología de tareas*. Ello supone delimitar la duración de cada fase de la investigación. En función de esto, habrá que precisar los *plazos de entrega parciales* (si los hubiese), y fijar la *fecha de entrega final*.

Existen diversas *técnicas* para el *cálculo de los tiempos de duración* de cada una de las etapas del *proyecto de investigación*, su *coste*, y las posibles *variaciones* que puedan acontecer.

Una de las técnicas más populares es el *método PERT* (*Program Evaluation and Review Technique* o *Program Evaluation/Research Task*). Básicamente, consiste en un *grafo* en el que se especifica cada una de las *tareas*, cómo dependen unas de otras, en qué *tiempo* se realizarán y qué *probabilidad de retraso* existe en su ejecución. Para su cálculo normalmente se pregunta a los responsables de cada una de las tareas el tiempo de duración habitual de su trabajo.

Todos estos detalles que conforman la última fase del *proyecto de investigación* son esenciales, en cuanto que ayudan al buen desarrollo de la investigación proyectada. Por lo que no hay que menosvalorar el esfuerzo que se dedique a su concreción. Como bien apunta Hakim (1994: 157), “quizás el error práctico más común consiste en subestimar el presupuesto, en tiempo y dinero, requerido para un proyecto”.

En la Figura 3.1 se resume cada una de las *fases* que forman un *proyecto de investigación*. Como puede verse, cada *fase* viene determinada por la etapa que la precede. Asimismo, los *objetivos* de la investigación, junto con los *recursos* y el *tiempo disponible* para su realización, se hallan presentes en cada una de las *fases del proyecto de investigación*, marcando su configuración final.

Por último, hay que insistir en la idea de que todo *proyecto de investigación* representa, como su nombre indica, un “proyecto”. Su puesta en práctica puede, por tanto, llevar consigo algunas modificaciones en una o en varias de las partes que lo conforman. De la pericia e ingenio del investigador (además de los medios a su disposición) dependerá, en buena medida, que la investigación finalice con éxito.

3.2. Diseños y estrategias de investigación

En el campo específico de los *métodos* y las *técnicas* de investigación social, a menudo se habla, indistintamente, de *métodos* o *técnicas*. Sin embargo, son cada vez más los autores que reclaman la distinción entre estos dos términos; algunos desde hace tiempo, como Greenwood (1973) o Bulmer (1984), otros en fechas más recientes, como Bryman (1995).

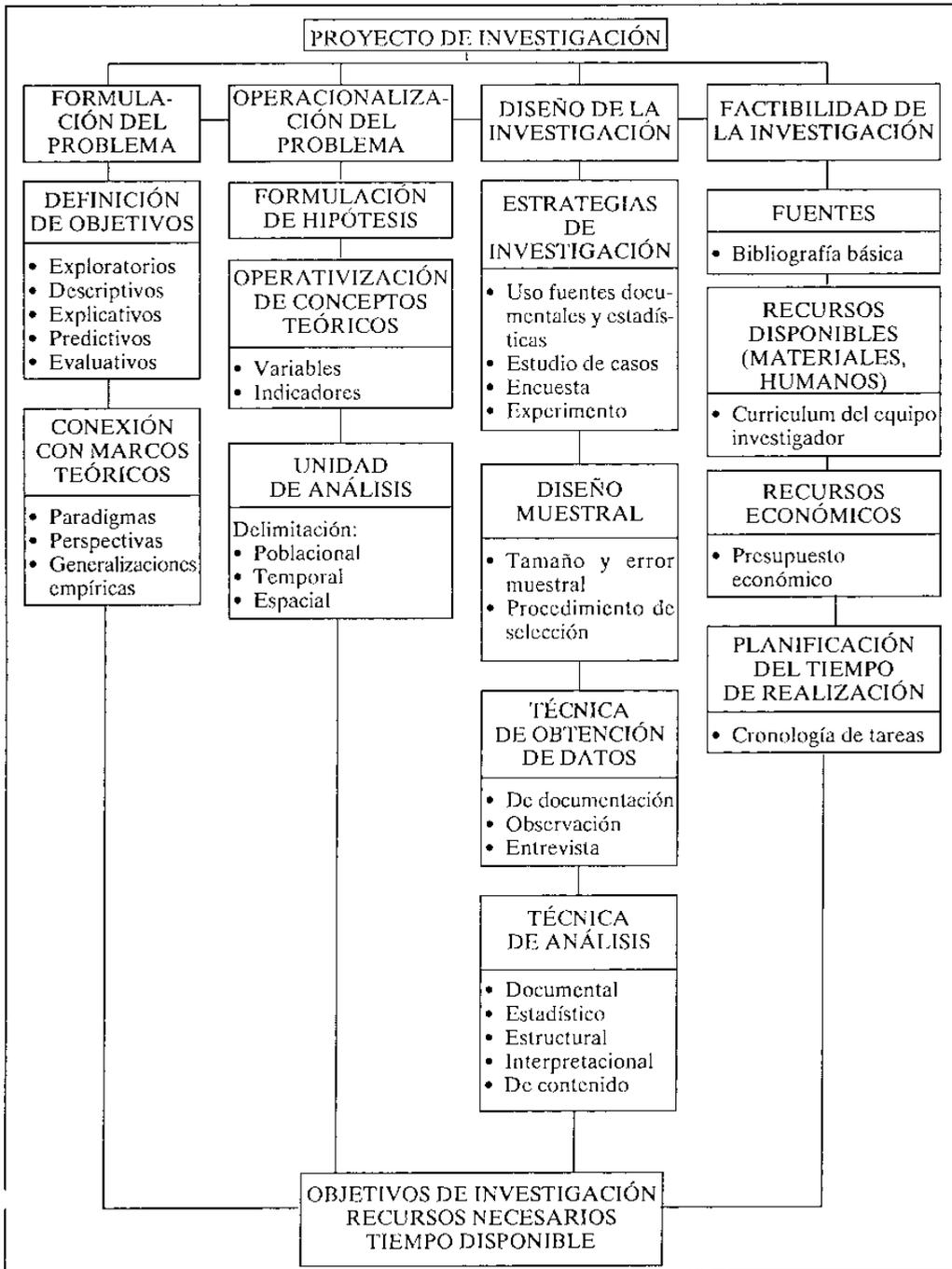


Figura 3.1. El proyecto de investigación.

Greenwood (1973), por ejemplo, emplea una analogía que coincide con el uso que en este manual se da al vocablo *estrategia*: “el método es a la técnica lo que la estrategia a la táctica. Es decir, que la técnica está subordinada al método, es un auxiliar de éste” (Greenwood, 1973: 107).

Bulmer (1984, 1992: 4-5) va más allá y propone diferenciar entre:

- a) *Metodología general*: “el estudio lógico y sistemático de los principios generales que guían la investigación sociológica”.
- b) *Estrategia (o procedimiento) de investigación*: “manera en que un estudio empírico particular es diseñado y ejecutado”.
Incluye el *enfoque* que se dará a la investigación, el *tipo de diseño* que se utilizará, y la combinación particular de las técnicas de investigación que se empleará.
- c) *Técnicas de investigación*: “operaciones manipulativas específicas de búsqueda de datos”.

Partiendo de esta clasificación, conviene, no obstante, barajar un concepto que ocupe un orden semántico superior al de *estrategia*. Se trata del concepto de *diseño*. En él se concentran –como ya se ha expuesto– las tareas directivas (o de ingeniería) de una investigación. Concretamente, la selección de la *estrategia* o *estrategias* a seguir y, en función de ellas, la elección de las *técnicas* (de *recogida de información* y de *análisis*) a utilizar.

Todo ello estará determinado por los *objetivos del estudio* y el *marco teórico de referencia*.

3.2.1. Pluralidad de estrategias de investigación

Como es constante en el ámbito de conocimiento de la *metodología*, no existe unanimidad entre los autores a la hora de diferenciar las diversas *estrategias de investigación*.

En la literatura consultada se observa la inexistencia de coincidencia plena, respecto a cuáles son las principales *estrategias de investigación social*. Si bien, no puede afirmarse que la discrepancia entre los autores sea abismal. Para ilustrar este extremo, el Cuadro 3.2 resume algunas *tipologías de estrategias de investigación*, seleccionadas y ordenadas por fecha de edición.

Bulmer (1992: 13) reconoce que su clasificación no es exhaustiva, aunque sostiene que los cuatro tipos diferenciados (a los que se refiere indistintamente como “*estilos*” y “*estrategias*” de investigación –al igual que hiciera posteriormente Yin (1989)– “cubre, probablemente las principales aproximaciones empleadas en nueve de cada diez monografías sobre investigación sociológica”.

Bulmer (1984) destaca, como *estrategia de investigación*, la que él denomina “*métodos y fuentes del historiador*”; en cambio, desestima otras estrategias importantes, como el “*experimento*”, al contrario de Yin (1989).

CUADRO 3.2. Principales estrategias de investigación social, según autores.

BULMER (1984)	YIN (1989)	MARSHALL Y ROSSMAN (1989)	BREWER Y HUNTER (1989)
Encuesta.	Experimento.	Experimento y cuasiexperimento.	Trabajo de campo.
Medidas no reactivas (datos preexistentes).	Encuesta.	Encuesta.	Encuesta.
Métodos y fuentes del historiador.	Análisis de datos de archivos.	Análisis de datos de archivos.	Experimentación.
Procedimientos interpretativos (trabajo de campo).	Historia.	Historia.	Investigación no reactiva.
	Estudio de casos.	Estudio de casos.	Aproximación multimétodo.
		Estudio de campo.	
		Etnografía	

En su monografía sobre la investigación mediante *estudio de casos*, Yin (1989) diferencia cinco *estrategias de investigación social*, sumando el “*experimento*” a la clasificación anterior. Su formación en historia y en psicología experimental le lleva a enfatizar la *experimentación* y los *métodos del historiador*. No obstante, la *estrategia* que este autor destaca más es la denominada “*estudio de casos*”.

A las *estrategias* apuntadas por Yin (1989, e. o. 1984), Marshall y Rossman (1989) añaden el “*estudio de campo*” y la “*etnografía*”. Si bien, estos autores consideran a las *etnografías* como casos especiales de *estudios de campo*.

Por último, Brewer y Hunter (1989) resumen a cuatro las *estrategias de investigación social* posibles: “*trabajo de campo*”, “*encuesta*”, “*experimentación*” e “*investigación no reactiva*” (haciéndose eco de la publicación de Webb *et al.*, de 1966, *Unobtrusive Measures: Nonreactive Research in the Social Sciences*). A ellas suman una quinta *estrategia*: la “*aproximación multimétodo*”. Ésta constituye la combinación de dos o más *estrategias de investigación*.

Los demás autores también partieron de una concepción de las *estrategias de investigación* como complementarias, abogando por una visión pluralista de las mismas. Sin embargo, no destacaron la *triangulación* o “*aproximación multimétodo*” como una *estrategia* específica y alternativa de investigación.

A partir de estas clasificaciones, propongo la siguiente *tipología de estrategias de investigación social*:

- a) *Uso de fuentes documentales y estadísticas.*
- b) *Estudio de casos* (etnográfico, biográfico, otros cualitativos y los cuantitativos).
- c) *Encuesta.*
- d) *Experimentación.*
- e) *La triangulación o aproximación multimétodo.*

La *estrategia* que Bulmer (1984) llama “*métodos y fuentes del historiador*”, y que tanto Yin (1989), como Marshall y Rossman (1989), resumen en la denominación de “*historia*”, optó por subsumirla en una *estrategia de investigación* más netamente sociológica que denominó: *uso de fuentes documentales y estadísticas*.

Ésta constituye una *estrategia básica* en cualquier *diseño de investigación* e incluye a las otras designaciones: *fuentes de datos preexistentes, análisis de datos de archivos, investigación no reactiva*.

Es una *estrategia básica* porque –como ya se expuso en la sección primera de este capítulo–, antes de proceder al diseño de cualquier investigación, resulta necesario efectuar una *revisión bibliográfica* exhaustiva sobre el tema concreto de estudio y el área de conocimiento específica donde éste se ubique. Ello favorece la familiarización con el *tema de investigación*, además de contribuir a la consecución de un buen *diseño de investigación*.

Las *estrategias* diferenciadas dentro de la *metodología cualitativa* (“*trabajo de campo*”, “*etnografía*”, “*estudio de casos*”) resuelvo enmarcarlas en la denominación tradicional de *estudio de casos*.

El *estudio de casos* constituye una *estrategia de investigación* de orientación diferente a la *encuesta*, la *experimentación*, y al *uso de documentos y estadísticas*; si bien, puede hacer uso de las mismas *técnicas de recogida y de análisis* de la información.

Tradicionalmente, el *estudio de casos* se ha ubicado en la *metodología cualitativa*. Pero, como apunta Yin (1989: 24-25), los *estudios de casos* “pueden basarse enteramente en evidencia cuantitativa y no necesitan incluir siempre observaciones directas y detalladas como fuentes de evidencia”.

Años más tarde, Bryman (1995: 170) insiste también en que “no todos los estudios de casos pueden ser adecuadamente descritos como ejemplos de investigación cualitativa, ya que algunas veces realizan un uso sustancial de métodos de investigación cuantitativa”.

Una persona, una familia, una comunidad, acontecimiento o actividad, puede constituir el *caso*, que será analizado mediante una variedad de *técnicas de recogida de información*.

El *estudio de casos* puede ser *único* o *múltiple*, depende del número de *casos* que se analicen (desde uno hasta cientos). Por otra parte, el análisis que se haga de los *casos* puede también variar. Ello determina el carácter que adquirirá el *estudio de casos*: *etnográfico, etnometodológico, biográfico* e, incluso, *cuantitativo*.

Cada una de las *estrategias* de investigación social referidas posee unas ventajas y unos inconvenientes. La elección entre ellas estará sobre todo determinada por:

- a) El *objetivo* del estudio.
- b) El *control* que el investigador desee ejercer en el desarrollo de la investigación.
- c) La *orientación* de la indagación, por fenómenos actuales (en un momento concreto), o hechos del pasado (históricos).

A estos aspectos se sumarían los relacionados con la *factibilidad de la investigación* y las *preferencias del investigador* (ya expuestos en la sección 3.1).

Sin embargo, como el clásico *dictum* de Trow (1957) recomienda, y del que se hacen eco numerosos autores –entre los cuales se encuentra Bulmer (1992: 15)–, el *problema de estudio* es lo que dicta “los métodos de investigación”, y no sólo las preferencias por un *estilo o estrategia de investigación*.

Así, por ejemplo, al *estudio de casos* suele describirse como una *estrategia de investigación* que:

- a) Se adecua a cuestiones del “cómo” y el “por qué”.
- b) En situaciones donde el investigador ejerce poco control sobre los acontecimientos.
- c) El foco de atención se encuentra en un fenómeno contemporáneo, dentro de algún contexto de la vida real. Aunque, también, puede obtenerse información de hechos del pasado en los *estudios de casos biográficos*.

Pero la aplicación del *estudio de casos* no sólo se limita a estas situaciones. De acuerdo con Hakim (1994: 61):

“Los estudios de casos son, probablemente, los más flexibles de todos los diseños de investigación. En el nivel más sencillo, proporcionan explicaciones descriptivas de uno o más casos. Cuando se utilizan de una forma intelectualmente rigurosa para lograr aislamiento experimental de factores sociales seleccionados, ofrecen la fuerza de la investigación experimental en entornos naturales.”

Los *estudios de casos* más famosos han sido tanto *descriptivos* (como es la famosa investigación de Whyte, *Street corner society*, en la que se hace uso de la *técnica de la observación participante*) como *explicativos* (sirva de ejemplo la obra publicada por Allison en 1971, *Essence of decision making. Explaining in the Cuban Missile Crisis*).

Pero, el *estudio de casos* también se adecua a los otros *objetivos de investigación*: *exploratorios*, *predictivos* y *evaluativos*. De él, se ha llegado incluso a afirmar que ocupa un lugar distintivo en la *investigación evaluativa* (Yin, 1989; Stake, 1994).

Las otras *estrategias de investigación* se detallarán en los capítulos siguientes: el *uso de fuentes documentales y estadísticas*, en el Capítulo 6; la *encuesta*, en el Capítulo 7; y la *experimentación*, en el Capítulo 8. Para una profundización en el *estudio de casos* remito al manual anteriormente citado de Miguel S. Valles, en esta misma colección.

A modo de conclusión de este subapartado, repárese en la conveniencia, a veces incluso exigencia, de una concepción pluralista de las diversas *estrategias de investigación social*.

Una misma *estrategia* puede permitir la consecución de distintos objetivos de investigación; y, a la inversa, un mismo objetivo puede suponer la aplicación de dos o más *estrategias de investigación* (como ya se expuso en el subapartado 2.1.2, dedicado a la *triangulación* o articulación de *estrategias* en una misma investigación).

Asimismo, para cualquier *objetivo de investigación* existe un amplio abanico de *estrategias* y técnicas a las que recurrir: “No importa lo que quieras averiguar, probablemente existan muchas formas de hacerlo” (Babbie, 1992: 89).”

3.2.2. Tipologías de diseños de investigación

Como en las *estrategias*, también existen diferentes clasificaciones de *diseños de investigación*. Depende de cuáles sean los *objetivos del estudio*, el *marco temporal* y el *marco contextual* de la observación (entorno natural o entorno artificial), primordialmente.

Tal vez una de las tipologías más conocidas e influyentes en la investigación social sea la que Campbell y Stanley (1970) publicaran en 1966 (posteriormente revisada por Cook y Campbell en 1977 y 1979). Pero, no es ésta la única tipología diferenciada.

El abanico de *diseños* no se limita a la clasificación originaria que Campbell y Stanley hicieron desde la lógica experimental. Puede ampliarse más, en función de cómo se trate la variable *tiempo* (*diseños seccionales* o *transversales* y *diseños longitudinales*), y de cuál sea el *objetivo* fundamental en la investigación. De ahí que en este manual se distingan tres *tipologías de diseños*. Éstos se resumen en el Cuadro 3.3.

- *Diseños preexperimentales, cuasiexperimentales y experimentales*

En su famoso libro sobre el *diseño experimental* (*Experimental and Quasi-experimental Designs for Research*), editado por vez primera en 1963, Campbell y Stanley diferenciaron tres tipos principales de *diseños de investigación*: los *diseños preexperimentales*, los *cuasiexperimentales*, y los *experimentales*. Esta tipología de *diseños* responde a los siguientes criterios de clasificación:

- a) Cómo se seleccionan las *unidades de observación*.
- b) El número de *observaciones* realizadas.
- c) El grado de *intervención del investigador* en el proceso de investigación.
- d) El *control* de posibles variables explicativas alternativas a las variables analizadas (*validez interna*).
- e) Posibilidad de *generalización de los resultados* de la investigación a otros contextos espaciales y temporales (*validez externa*).

CUADRO 3.3. Clasificación de los diseños de investigación.

<ul style="list-style-type: none"> • Según el <i>grado de cumplimiento de los supuestos de la experimentación</i>: <ol style="list-style-type: none"> 1. Diseños preexperimentales o correlacionales. 2. Diseños cuasiexperimentales. 3. Diseños experimentales. • Según el <i>tratamiento de la variable tiempo</i>: <ol style="list-style-type: none"> 1. Diseños seccionales o transversales. 2. Diseños longitudinales. <ul style="list-style-type: none"> — De tendencias. — De cohorte. — Panel. • En función de <i>los objetivos de la investigación</i>: <ol style="list-style-type: none"> 1. Diseños exploratorios. 2. Diseños descriptivos. 3. Diseños explicativos. 4. Diseños predictivos. 5. Diseños evaluativos. <ul style="list-style-type: none"> — Evaluación del impacto. — Evaluación del proceso. — Valoración de necesidades. — Evaluación mediante análisis de sistemas. — Análisis coste-beneficio. — Evaluación de conjunto.
--

1) *Diseños preexperimentales o correlacionales*

Entre sus *características* distintivas destacan:

- a) La *ausencia de manipulación* de las variables intervinientes en la investigación. El investigador se limita a la observación del fenómeno que analiza, sin introducir ninguna modificación o alteración en el mismo.
- b) Se efectúa una *única medición* del fenómeno, aunque se incluyan diversos aspectos del mismo.
- c) *Falta de control* de posibles fuentes de invalidación de la investigación, lo que resta poder explicativo a estos diseños.

Ahora bien, la aplicación de técnicas de *análisis multivariable* permite que los *diseños preexperimentales* (o *correlacionales*) también lleguen al establecimiento de relaciones causales, por medio del *control a posteriori*; no cuando se diseña la investigación (como acontece en los *diseños experimentales*), sino después de haberse recogido la información, en la fase de análisis.

Los *diseños preexperimentales* incluyen distintas modalidades, cuya comprensión precisa de la lectura del Capítulo 8, dedicado a la *experimentación* como *estrategia de investigación*.

- a) El *diseño de un único grupo con una sola medición* (posterior al *tratamiento*).
- b) *Diseño pretest-postest de un único grupo*.
- c) Comparación entre un *grupo experimental* y otro de *control*, pero sin ninguna *medición previa*.

Ninguno de estos tres *diseños* cumplen todos los requisitos de la *experimentación*: la existencia de al menos un *grupo de control* (no expuesto al *tratamiento* cuyos efectos quiere medirse); la *aleatorización* en la formación de los grupos; y la *medición antes y después del tratamiento experimental*. Por esta razón los autores los denominan “*diseños preexperimentales*”.

El primer diseño carece de *medición* (en la *variable dependiente*) previa al *tratamiento* (o manipulación de la *variable independiente*). Ello, unido a la carencia de un *grupo de control* que ayude a la eliminación de explicaciones alternativas al *tratamiento*, impide diferenciar los *efectos* debidos al *tratamiento* de los provocados por factores ajenos al mismo.

Los otros dos *diseños* resuelven una de esta dos salvedades, pero no las dos. En el segundo diseño (el *diseño pretest-postest de un único grupo*) se introduce la medición de la *variable dependiente* antes de la aplicación del *tratamiento*; ello permite conocer la evolución de los sujetos analizados. En cambio, en el tercer diseño se incluye el *grupo de control*, pero se carece de *medición* previa al *tratamiento*. Hecho que dificulta, igualmente, el establecimiento de *relaciones causales*.

Como ejemplo de *diseño preexperimental* o *correlacional* destaca la *encuesta* usual (no la *encuesta panel*), en la que se produce una única medición de la realidad social.

2) *Diseños experimentales*

Cabe definirlos como “observación controlada” (Caplow, 1977), gracias a una serie de actuaciones del investigador, dirigidas al control de posibles fuentes de invalidación de la investigación. Entre estas actuaciones destacan:

- a) La *manipulación experimental*. En los *diseños experimentales* el investigador puede manipular *a priori* (es decir, antes de la recogida de información) las va-

riables cuya influencia en la ocurrencia de determinados fenómenos trate de medir.

- b) La *formación de grupos de control*, que sean totalmente equivalentes al *grupo experimental*, salvo en la variable o variables independientes cuyos efectos se trata de medir.

Los individuos se asignan, de forma *aleatoria*, a cada uno de los grupos: el *experimental* y el de *control*. De esta forma se garantiza la equivalencia inicial en la constitución de ambos tipos de grupos.

Estas actuaciones favorecen el análisis de *relaciones causales*, al cumplirse las exigencias de *validez interna*: el control de posibles explicaciones alternativas a las analizadas.

No obstante, los *diseños experimentales* presentan una importante desventaja: la mayoría de ellos carecen de *validez externa*, debido, precisamente, a la *manipulación experimental* (a la alteración que el investigador introduce en la realidad que analiza). Ello imposibilita la generalización de los resultados de la investigación.

La falta de *validez externa* también suele deberse a las características de la *muestra*. Esto acontece cuando los sujetos que participan en el *experimento* no se seleccionan al azar, entre los que constituyen el *universo* o *población* del estudio, sino que se reclutan entre los voluntarios al *experimento*. Ello, sumado al usual escaso número de integrantes en la *muestra* (no superior a los 200 casos, generalmente), limita la posibilidad de generalización de los resultados de la investigación a otros contextos distintos al experimental (*validez externa*).

Bajo la rúbrica de *diseños experimentales* se incluye una variedad de *diseños* de investigación. Estos se detallan en el Capítulo 8, dedicado a la *experimentación*. La lectura de este capítulo es necesaria para la comprensión de la lógica experimental.

3) *Diseños cuasiexperimentales*

Se hallan a caballo entre los dos polos extremos de *diseños* anteriormente referidos, participando de las características definitorias de ambos.

En los *diseños cuasiexperimentales* puede haber (o no) *manipulación experimental* de la *variable independiente* o predictora, con el propósito de comprobar su efecto en la *variable dependiente* (cuyos valores –como su nombre indica– “dependerán” de los que tome la *variable independiente*). Pero, se distancian de los *diseños experimentales* en dos aspectos fundamentales:

- a) Los *diseños cuasiexperimentales* rara vez acontecen en el marco de un laboratorio, sino en el contexto de la vida real.

- b) La distribución de las unidades de observación (la *muestra* del estudio) al *grupo experimental* y al de *control* no se realiza de forma *aleatoria*. Por lo que no puede garantizarse la equivalencia inicial de los grupos de comparación.

Esta última limitación puede obviarse mediante la aplicación de *técnicas estadísticas bivariantes y multivariantes de control a posteriori*. Ello favorece la consecución de resultados similares a los alcanzados mediante los *diseños experimentales*.

De los *diseños preexperimentales* (o *correlacionales*) se distinguen por un hecho esencial: la actuación del investigador no se limita (en los *diseños cuasiexperimentales*) a la observación; sino que, por el contrario, puede estructurar la situación de forma que facilite su análisis.

En función de estas características definitorias de los *diseños cuasiexperimentales*, puede agruparse la amplia variedad de *diseños* existentes en dos categorías generales:

- a) Diseños en los cuales el investigador sí *manipula* la situación experimental, pero no existe un *grupo de control*.
 b) Diseños en los que el investigador no *manipula* la situación experimental, pero sí existen *grupos de control y experimental* equiparables.

Campbell y colaboradores distinguen tres grandes grupos de *diseños cuasiexperimentales*, que resultan de la combinación de las distintas opciones posibles:

- a) Diseños que no permiten la *inferencia causal*, bien por la ausencia de algún *grupo de control* equiparable, o bien por la inexistencia de medición de las variables antes de comprobarse la influencia de un *tratamiento* o *intervención*.
 b) Diseños con *grupo de control* no equivalente.
 c) Diseños de *series temporales* (a partir de la información extraída de fuentes estadísticas y documentales, de encuestas periódicas, o mediante otra estrategia de investigación). Lo que singulariza a esta variedad de *diseño* es la medición de la *variable dependiente* en distintos períodos de tiempo, con la finalidad de analizar su evolución.

Como esta clasificación de *diseños* se hace desde la vertiente experimental, su comprensión exige el conocimiento de los rasgos distintivos de la *experimentación* como *estrategia de investigación*. Razón por la cual se reitera en la necesidad de completar esta tipología de *diseños* con la lectura del Capítulo 8.

• *Diseños seccionales (o transversales) y diseños longitudinales*

Una clasificación distinta de los *diseños de investigación* responde a cómo se haya planificado la recogida de información: en un único momento (*diseños seccionales* o

transversales) o, por el contrario, en dos o más veces (*diseños longitudinales*). En este último caso, se pretende analizar la evolución de los fenómenos que se investigan a lo largo del “tiempo”. De esta forma, el “tiempo” pasa a ser variable esencial en este tipo de *diseños*.

1) Diseños seccionales o transversales

Se caracterizan por circunscribir la recogida de información a un único momento en el tiempo. El objetivo de la investigación puede ser *descriptivo* (describir las características de una población en una fecha concreta, como acontece en el Censo de Población, por ejemplo), *explicativo* (analizar, mediante una *encuesta*, las variables que inciden en la ocurrencia de un hecho o acción determinada, por ejemplo), o de otro orden. Lo que distingue a esta variedad de *diseño* es que la recogida de información se lleva a cabo de una sola vez, aunque se incluyan circunstancias temporales o contextos ambientales diferentes.

EJEMPLO DE DISEÑO SECCIONAL

Si un investigador quisiera comprobar cómo incide la situación económica de un país en los hábitos de consumo de sus habitantes, el *diseño de la investigación* sería *seccional*, si la información se recogiese en un único momento: mediante una *encuesta*, por ejemplo, aplicada una sola vez, en varios países seleccionados conforme a su nivel de desarrollo, y de forma simultánea.

El *diseño* sería, por el contrario, *longitudinal*, si se hiciera la misma *encuesta* en fechas diferentes, con la periodicidad que el investigador determinase en el *diseño* de la investigación. En este caso, el objetivo principal del estudio sería analizar la *evolución* de los hábitos de consumo, en tiempos y países distintos.

2) Diseños longitudinales

Esta modalidad de *diseño* se caracteriza por plantear el análisis del *problema de estudio* a lo largo del tiempo, con el propósito de observar su dinámica. La recogida de información se planifica, entonces, para su realización en varias fechas, que el investigador especifica en el *proyecto de investigación*.

La amplitud del período de observación y su cronología se halla relacionada con el objetivo de la investigación.

EJEMPLO DE DISEÑO LONGITUDINAL

Si el objetivo principal de la investigación fuese describir cómo la experiencia de desempleo influye en la dinámica familiar de las personas en paro, la concreción del tema y su limitación en el tiempo marcará la duración del estudio: por lo general, inferior a tres años. En todo caso, el período de investigación guardará relación con la duración media estimada de la situación de desempleo en la sociedad y momento histórico que se analice.

Por el contrario, describir la experiencia de la vejez, cómo se vive en distintas sociedades y a diferentes edades, implicaría un lapso de tiempo bastante superior. El estudio podría comenzar en el momento en que una *cohorte* de personas (o generación) cumple los 65 años de edad, y finalizar cuando sus supervivientes (de la *cohorte* escogida) alcancen los 90 años. Obviamente, conforme se avance en el tiempo, disminuirá el número de personas en la *muestra del estudio*; sobre todo, cuanto mayor sea el período de tiempo considerado. Asimismo, se podría optar por seguir:

- a) Un *diseño longitudinal de cohorte*, si se parte de la cohorte general de personas que en la fecha de inicio de la investigación cumple los 65 años de edad. De ellas se extraería, preferiblemente mediante un procedimiento aleatorio, una *muestra*. Para posteriores observaciones (por ejemplo, cada cinco años), se seleccionarían nuevas *muestras* del total de supervivientes de la *cohorte* inicial.
- b) Un *diseño longitudinal de panel*. Este tipo de *diseño* difiere del anterior en que la investigación se fundamenta en una única *muestra*. Son las mismas personas que conformaron la *muestra* inicial del estudio las que se analizan durante toda la investigación. El inicio sería el mismo: se extrae una muestra de la *cohorte* de personas que cumple 65 años al comienzo del estudio. Pero, para posteriores observaciones, no se elegiría una nueva *muestra*, sino que volvería a recogerse información de las mismas personas que constituyeron la *muestra* inicial de la investigación.

Habría, por tanto, varias modalidades de *diseños longitudinales*. Éstos tradicionalmente se han agrupado en los tres tipos siguientes: *diseños longitudinales de tendencias*, de *cohorte* y de *panel*.

- *Diseño longitudinal de tendencias*

El estudio se enfoca a la descripción de la población total, no de una parte de ella. Se trata de analizar su evolución, prestando atención a las *tendencias* y cambios de *tendencias* (de ahí su nombre) en las características investigadas.

EJEMPLO DE DISEÑOS LONGITUDINALES DE TENDENCIAS

Éste es el caso cuando se comparan los datos registrados en diferentes censos de población (como los de 1960, 1970, 1981 y 1991) para analizar su evolución; o cuando se contrastan diferentes sondeos de opinión, efectuados durante el desarrollo de una campaña electoral, con la finalidad de analizar la evolución del voto; o cuando se cotejan las respuestas emitidas a unas mismas preguntas en distintas encuestas, realizadas en fechas diferentes.

En cualquiera de estos casos, se confrontan las respuestas dadas a unas mismas cuestiones. El instrumento de medición no varía; pero sí la *muestra* del estudio. En cada observación puede analizarse una *muestra* diferente (aunque extraída de la misma población que constituye el foco de la investigación). Otras veces no se extrae una *muestra*, sino que se analiza a toda la población.

La existencia de bancos de datos, donde se almacenan distintas encuestas (atendiendo a su temática), facilita el análisis de *tendencias*. Además, en muchos centros de opinión suelen realizarse sondeos periódicos, donde se formulan las mismas preguntas que en sondeos precedentes, pero a distintas personas (aunque de características semejantes a las *muestras* anteriores). Con esta práctica se pretende determinar la influencia del *tiempo* en el cambio de *tendencias* en las respuestas a distintas *encuestas*.

• Diseño longitudinal de cohorte

Esta segunda modalidad de *diseño longitudinal* difiere del anterior en que el interés no se halla en la población total, sino en una subpoblación o *cohorte*.

La *cohorte* estará constituida por individuos que comparten una misma característica; la edad, habitualmente. La *cohorte de edad* puede definirse de forma amplia: personas nacidas en los años sesenta en España, por ejemplo. Pero, también de una manera más específica: personas que nacieron en 1961 o, aún más concretamente, en el mes de julio de 1961.

Aunque la *cohorte de edad* constituya la *cohorte* más típica, no es la única posible. Una *cohorte* puede constituirse a partir de un acontecimiento determinado: personas que se doctoraron en las universidades españolas en 1985; parejas que se casaron en Sevilla, en 1992, durante la Expo. Igualmente, la *cohorte* puede fijarse a partir de una experiencia concreta: mujeres que estuvieron en prisión 10 años; personas mayores de 40 años que llevan dos años en situación de paro ininterrumpido, por ejemplo.

De la *cohorte* escogida se analiza su evolución. Para ello se selecciona una *muestra* distinta de entre los individuos que conforman la *cohorte* seleccionada. Es decir, se observan distintos individuos (pertenecientes a la misma *cohorte*) en momentos diferentes.

EJEMPLO DE DISEÑO LONGITUDINAL DE COHORTE

Un grupo de investigadores desea analizar cómo varían las relaciones entre padres e hijos, a medida que estos últimos van independizándose del núcleo familiar. Para ello eligen una *cohorte* compuesta por personas que nacieron en el año 1955. Elaboran un cuestionario, formado por preguntas referidas al tema en cuestión. La encuesta deciden que se pase cada cinco años, a personas que integren la *cohorte* seleccionada.

En 1975 se extrae una *muestra* de entre aquellos jóvenes de 20 años de edad; en 1980 se extrae otra *muestra* de entre las personas de 25 años; en 1985, de entre aquéllos con 30 años; y así sucesivamente, hasta la fecha de conclusión de la investigación.

Como la *muestra* se escoge de los supervivientes de la *cohorte* inicial, su volumen se verá mermado conforme se avance en el tiempo

Pero la investigación no tiene por qué circunscribirse a una única *cohorte*. Puede compararse la evolución de *cohortes* diferentes.

EJEMPLO DE DISEÑO LONGITUDINAL DE VARIAS COHORTES

En una investigación que tenga como objetivo principal determinar cómo incide la experiencia carcelaria en la reincidencia en conducta delictiva (medida por el número de detenciones policiales o de la Guardia Civil), podría compararse la evolución de personas que estuvieron en prisión durante cinco años en dos períodos de tiempo diferentes: uno, en los años sesenta (durante el franquismo); dos, a principios de los años ochenta (durante la democracia).

• *Diseño longitudinal de panel*

En esta última variedad de *diseño longitudinal*, la atención del investigador se dirige, a diferencia de los *diseños* anteriores, a analizar la evolución de unos mismos individuos, que se eligieron al inicio de la investigación. Por lo que, no se procede a nuevas selecciones muestrales en cada fase posterior de la investigación.

EJEMPLO DE DISEÑO LONGITUDINAL DE PANEL

Un ejemplo clásico de este tipo de *diseño* es la investigación que Lazarsfeld, Berelson y Gaudet llevaron a cabo en EEUU, durante las elecciones presidenciales de 1940 (la campaña electoral Wilkie-Roosevelt): *The people's choice* (1944).

El estudio tenía como objetivo comprobar la influencia de la campaña electoral en las intenciones de voto de los electores del condado de Erie (Ohio). Para dicho propósito diseñaron un cuestionario, que pasaron varias veces (entre mayo y noviembre de aquel año) a la misma muestra de electores que extrajeron (siguiendo procedimientos *aleatorios*) a comienzos de la investigación (600 personas).

El *diseño panel* es de gran utilidad para indagar en las causas del “cambio”. Las personas que reiteradamente se observan son las mismas; lo que varía son las circunstancias en que se hallan tras el paso del tiempo. Esto permite conocer los factores que pueden haber contribuido al cambio que en ellos se observe.

En su contra, el *diseño panel* se enfrenta a dos problemas metodológicos fundamentales:

- a) El *desgaste de la muestra* con el consiguiente aumento de la *no respuesta*.

Éste representa un problema similar al denominado “mortalidad experimental”. Hace referencia a la dificultad de localizar, en fechas distintas, a las personas que forman la *muestra* del estudio. Ello puede deberse a un cúmulo de razones: la persona se niega a seguir participando en la investigación; ha cambiado de domicilio y no lo ha comunicado al equipo investigador; se encuentra enferma o ha fallecimiento, entre otras posibles razones.

Este problema se halla más presente en *estudios longitudinales* que abarcan un amplio período de tiempo.

El paso del tiempo actúa de forma negativa, propiciando el desinterés de los sujetos por participar en la investigación. Esto revierte en la reducción del *tamaño muestral*. Lo que introduce un sesgo importante en la investigación: las personas que abandonan el estudio pueden diferir de aquellas que permanecen en la investigación. Ello tendrá repercusión en la *validez externa* de la investigación (su significatividad y posibilidad de generalización).

- b) El proceso de medición puede suscitar *sesgos en mediciones posteriores*, por efecto del aprendizaje.

Las respuestas o actitudes que una persona manifieste en una fase de la investigación puede deberse a su actuación en fases anteriores, especialmente cuanto mayor es la proximidad entre las fechas de observación. Es decir, el hábito o el ejercicio de la investigación puede llevar a la no veracidad de las respuestas o conductas manifestadas.

A estos dos problemas fundamentales del *diseño panel* se suma uno común a todo *diseño longitudinal*: el mayor coste económico de la investigación, no sólo presente en la fase de recogida de información, sino también en la de análisis (por el amplio volumen de información que se maneja). Este problema adquiere mayor relieve, conforme aumenta el tamaño de la muestra y su diversidad espacial.

El *diseño transversal o seccional* también puede hacer viable el análisis del cambio (y a un coste económico considerablemente inferior al *diseño longitudinal*), si en el instrumento de medición (cuestionario, guión de entrevista abierta, ...) se incluyen preguntas concernientes al “pasado” de los sujetos: hechos, opiniones, actitudes. Por ejemplo, a qué partido votó en las tres últimas elecciones; cuáles eran sus actividades de ocio antes de casarse; qué opina sobre el divorcio ahora y cuál era su opinión hace cinco años.

Este tipo de estudio retrospectivo, al que Hakim (1994) se refiere como “diseño cuasi-longitudinal”, presenta, igualmente, un problema metodológico importante: los fallos en la memoria. La persona entrevistada puede no recordar acontecimientos de su vida pasada, sobre todo cuanto mayor es el lapso de tiempo transcurrido y menor repercusión tuvo el evento en su vida. Asimismo, puede sí recordarlos, aunque vagamente, y ser inexactas sus respuestas.

En suma, cualquier *diseño* de investigación presenta unas ventajas y, a su vez, unos inconvenientes. La elección entre un tipo de *diseño* u otro vendrá marcada por la conjunción de los tres factores señalados en el apartado 3.1. Concretamente, los *objetivos* de la investigación, los *recursos* (materiales, económicos y humanos) y el *tiempo* que se disponga para la realización del estudio.

- *Diseños exploratorios, descriptivos, explicativos, predictivos y evaluativos*

Una última tipología de *diseños* de investigación responde, precisamente, a cual sea el objetivo principal de la investigación. En 1989, dos autoras, Marshall y Rossman, publican, en su obra *Designing Qualitative Research*, una clasificación de *diseños* fundamentada en los *objetivos de la investigación*; si bien, los limitan a cuatro: “*exploratorios*”, “*descriptivos*”, “*explicativos*” y “*predictivos*”.

A estos objetivos genéricos habría, en mi opinión, que añadir otro fundamental: el *evaluativo*. De este modo resultaría la configuración de *diseños* de investigación que se detalla a continuación.

Sin embargo, téngase presente que esta clasificación de *diseños* no ha de entenderse como si se tratase de modalidades excluyentes. Una misma investigación puede incluir *objetivos* de diverso rango, en consonancia con las distintas fases en su desarrollo. En estos casos se estaría ante *diseños de investigación* complejos.

El estudio puede comenzar siendo *exploratorio* (si el equipo investigador dispone de escasa información sobre el objeto de conocimiento) para, posteriormente, proceder a la *descripción*, *explicación*, *predicción* y/o *evaluación*. Generalmente, los *diseños* que incluyen objetivos de rango superior suelen comprender, a su vez, objetivos inferiores como fases previas en su materialización. Tal es el caso de la *investigación evaluativa*, en cuyo *diseño* confluyen varias etapas: desde la familiarización con el *programa* o la *intervención* que se evalúa (mediante un *diseño exploratorio* y *descriptivo*) hasta, por ejemplo, la valoración de los *resultados del programa* (mediante un *diseño explicativo* que anteceda al *evaluativo*).

La interpretación de los *resultados* de cualquier *programa*, intervención, o conjunto de actividades, requiere que antes se haya descrito al *programa* y su funcionamiento real. A partir de esta descripción se procede a descubrir qué se debe al *programa* y qué cabe atribuir a la intervención de otros factores. Como afirma Alvira (1991a: 7, 10-11):

“Evaluar es más que investigar [...]. Evaluar es emitir juicios de valor, adjudicar valor o mérito de un programa/intervención, basándose en la información empírica recogida sistemática y rigurosamente.”

Por último, adviértase que no se trata de una taxonomía, ni tampoco de una escala. Podría, también, haberse resaltado otros objetivos (como el comparativo, por citar alguno). Pero se desestiman por considerar que quedan subsumidos en los objetivos anteriores.

1) *Diseño exploratorio*

Un diseño de investigación exploratorio se lleva a cabo para cubrir alguno o varios de los *propósitos* siguientes:

- a) Familiarización con el *problema de investigación* para deducir (a partir de la información reunida) qué aspectos requieren un análisis pormenorizado en indagaciones posteriores.
- b) Verificar la factibilidad de la investigación y documentar los medios que se precisan para hacerla viable.
- c) Comprobar qué *estrategia* (o estrategias) de investigación se adecúa más a su análisis.

En cada *estrategia*, seleccionar la *técnica* (o técnicas) de obtención de datos y de análisis más pertinentes para futuras indagaciones (más formalizadas).

El *estudio exploratorio* rara vez constituye un fin en sí mismo. “Rara vez proporcionan respuestas satisfactorias a las preguntas de investigación. Pueden sugerir los métodos de investigación que podrían proporcionar respuestas definitivas” (Babbie, 1992: 91). A ello contribuye la escasa representatividad de los casos que se analizan (ya por su volumen, ya por el procedimiento de selección muestral empleado).

2) *Diseño descriptivo*

Al igual que la exploración, la *descripción* constituye un paso previo en cualquier proceso de investigación. Antes de indagar en la explicación de cualquier evento hay que

proceder a su descripción mediante alguna o varias *estrategias de investigación* (*encuesta, uso de documentos y estadísticas o el estudio de casos*). De ellas, el investigador obtendrá información que le servirá en la caracterización del fenómeno que analiza.

EJEMPLO DE DISEÑO DESCRIPTIVO

El *censo de población* es un ejemplo tradicional de indagación *descriptiva*. En él se describen características sociodemográficas básicas de una población determinada, en un momento concreto.

3) Diseño explicativo

Después de la *descripción* procede la *explicación*: buscar posibles causas o razones de los hechos, acciones, opiniones o cualquier fenómeno que se analice.

EJEMPLO DE DISEÑO EXPLICATIVO

Describir el perfil (o perfiles) de la población delincuente juvenil constituye un *estudio descriptivo*. Pero, analizar las *causas* que convergen en la delincuencia juvenil implica un *diseño explicativo*, en el que se midan las interrelaciones e influencias de distintas variables.

4) Diseño predictivo

La *predicción* forma un objetivo específico en sí mismo, aunque lleva consigo, a su vez, alguno o varios de los objetivos anteriormente referidos. Para poder predecir cuál será la evolución futura de un determinado fenómeno habrá, previamente, que proceder a su análisis en el momento presente, tanto en la vertiente *descriptiva* como en la *explicativa*.

5) Diseño evaluativo

De acuerdo con Babbie (1992: 346), “la *investigación evaluativa* –algunas veces llamada *evaluación de programas*– se refiere más a un propósito de investigación que a un método de investigación específico”. Se configura como una forma de *investigación*

aplicada definida por la “aplicación sistemática de procedimientos de investigación social para asegurar la conceptualización, diseño, realización y utilidad de programas de intervención social” (Rossi y Freeman, 1991: 5).

En el *diseño evaluativo* se busca la aplicación de procedimientos de investigación (*estrategias, técnicas*), de manera sistemática y rigurosa. Con ello se pretende alcanzar conclusiones (*válidas y fiables*) sobre la efectividad del *programa* (o conjunto de actividades específicas) cuya actuación se evalúa.

Existen varias *modalidades de evaluación*, que pueden resumirse en las siguientes:

- *Evaluación del impacto*

Incluye las denominadas *evaluación de resultados, de efectividad y sumativa*. Representa una de las variedades de evaluación más populares. Persigue –como su nombre indica– el análisis del *impacto* o *efectos de un programa* para, a partir de la información obtenida, buscar su mejora (si se precisa). Ello requiere que previamente se:

- a) *Definan* los criterios de éxito: los *objetivos* o *metas* que el *programa* debería alcanzar.
- b) *Diferencien* los *resultados* debidos al *programa* de los causados por otros factores.
- c) *Especifiquen* las condiciones bajo las cuales el *programa* resultaría más eficaz.

- *Evaluación del proceso, formativa o del desarrollo*

Tiene como objetivo fundamental la *descripción del programa*: en qué consiste y cómo funciona, en la teoría y en la práctica. De ahí su complementariedad con la modalidad de evaluación anterior. La *evaluación del proceso* contribuye a:

- a) La comprobación de si el *programa* se está realizando en conformidad con los planes originales.
- b) Conocer las causas de su éxito o fracaso.

Para ello, es necesario que la *evaluación del proceso* incluya –siguiendo a Krisberg (1980)– los siguientes aspectos:

- a) Las condiciones y los supuestos que definan, operativa y conceptualmente, los rasgos del *programa* (supuestos teóricos, características organizativas, históricas).

- b) Los criterios y los procedimientos seguidos en la selección de los clientes o participantes en cada una de las alternativas del programa.
- c) Las actividades y los servicios creados para alcanzar los objetivos del *programa*.
- d) Los criterios a seguir para poder determinar la efectividad del *programa*.

- *Valoración de necesidades*

En este caso, el objetivo específico de la investigación lo constituye la identificación de las *necesidades prioritarias* que el *programa* debe atender. Este tipo de valoración suele llevarse a cabo cuando el programa a evaluar resulta amplio y complejo, con un gran número de componentes que precisan revisarse (Herman *et al.*, 1987).

- *Evaluación mediante análisis de sistemas*

Se parte de la consideración del programa como un conjunto de subsistemas (personal subalterno, dirección, gerencia) relacionados y vinculados, a su vez, a otros sistemas más amplios (familia, comunidad). Se busca el análisis de sus relaciones e interrelaciones.

- *Análisis coste-beneficio*

Se estudia la relación entre los *costes del programa* y sus *resultados*, expresados (generalmente) en términos monetarios.

Difiere del *análisis efectividad-coste* en que este último no traduce los beneficios del programa a términos monetarios, sino de *impacto real*.

- *Evaluación de conjunto*

Comprende tanto las *metas del programa* como su *repercusión*, las alternativas disponibles y/o los costes de su actuación. Por lo que conforma una variedad de *evaluación de síntesis*.

La *investigación evaluativa* puede –como bien indica Alvira (1985)– desarrollar cuatro fases principales: *evaluación de necesidades*, *planificación del programa*, *evaluación formativa* y *evaluación sumativa*; solamente una, o varias de ellas, o incluso todas conjuntamente.

Asimismo, el *diseño de una investigación evaluativa* puede llevar consigo la aplicación de una única *estrategia de investigación* (el *experimento* se muestra como una de las más idóneas) o varias a la vez. Igualmente, la *validez* de sus hallazgos puede requerir

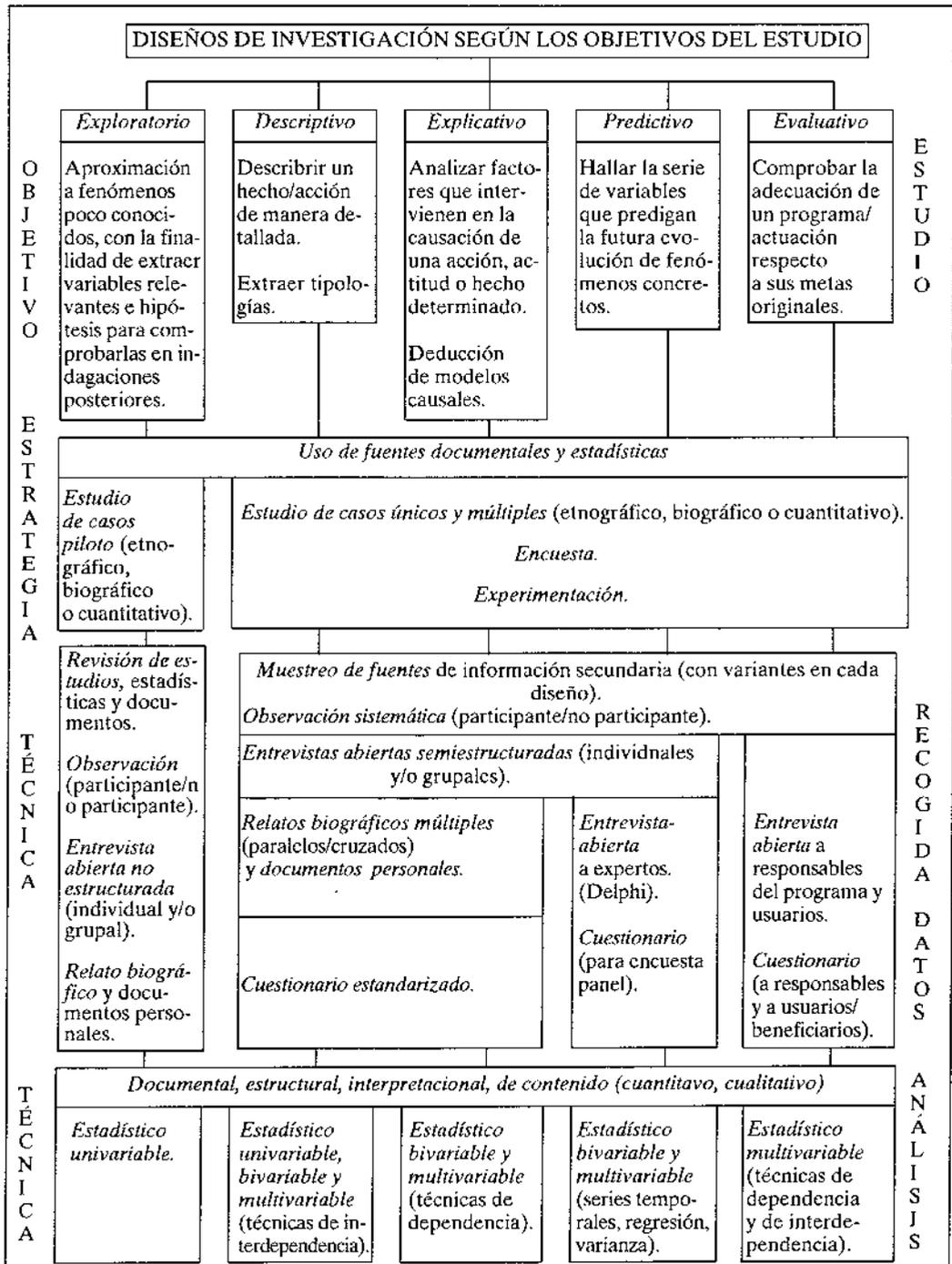


Figura 3.2. Tipología de diseños de investigación según los objetivos del estudio.

(como en los otros tipos de investigación) la utilización de varias *técnicas de recogida y de análisis* de información (tanto *cuantitativas* como *cualitativas*).

La Figura 3.2 sintetiza la variedad de *diseños de investigación*, dependiendo del *objetivo principal del estudio*. Para cada tipo de *diseño* se indican distintas *estrategias y técnicas* (de *recogida* y de *análisis*) alternativas para cumplir los *objetivos* marcados al inicio de la indagación.

La presentación esquemática no debe mal interpretarse. Se trata de categorías genéricas de *diseños*, cuyo valor clasificatorio se consigue a costa de sacrificar el reflejo fiel y directo de la práctica investigadora.

Como ya se ha hecho referencia, cualquier investigación puede perseguir más de un *objetivo*. Puede comenzar siendo descriptiva, y terminar siendo explicativa, por ejemplo. Depende, primordialmente, del estado de conocimiento que el investigador tenga, además del enfoque que pretenda dar a la indagación.

3.2.3. Ejemplo detallado tomado de una investigación real

Como ilustración de lo expuesto, a continuación se describe el desarrollo de una investigación real. Se trata de la *investigación evaluativa* a la que se ha hecho referencia en apartados anteriores: Cea D'Ancona, M.^a A. (1992): *La Justicia de Menores en España*, Madrid, CIS, *Monografía*, n.º 127.

La investigación tenía como *objetivo principal* evaluar la actuación de los (ya extintos) Tribunales Tutelares de Menores en España. La delimitación del *problema de estudio* supuso la incardinación de la investigación en el marco de la *metodología de la evaluación de programas*. Se quería descubrir si el sistema de justicia de menores alcanzaba sus propósitos originarios (rehabilitar al transgresor, reeducándolo y adaptándolo a la vida social); o, por el contrario, producía *resultados* no deseados contribuyendo, de este modo, al desarrollo de la “*carrera delictiva*”, como afirman los teóricos del *etiquetamiento*.

A tal fin se opta, primero, por efectuar una *evaluación del proceso* (o *formativa*), que describa el sistema de justicia de menores en España (en la teoría y en la práctica), a nivel normativo y organizacional. Una vez conocido su funcionamiento real (primer objetivo), se comprueba su efectividad o *impacto* (segundo y principal objetivo de la investigación).

Ambos objetivos se ajustan a dos modalidades de *evaluación*: la *evaluación del proceso* (o *formativa*), y la *evaluación del impacto* (o *sumativa*). A partir de ambas evaluaciones, se deducen los aspectos más necesitados de reforma en la justicia de menores (tercer objetivo específico de la investigación).

Para el cumplimiento de los objetivos marcados, se siguieron distintas fases. El primer paso fue el habitual acopio de *referencias bibliográficas*, no sólo relativas a la jurisdicción de menores y a la delincuencia juvenil, sino también de aspectos metodológicos concernientes a la *evaluación de programas* y las *técnicas de análisis de datos* (*bivariados* y *multivariados*) para, con posterioridad, poder abordar la parte empírica de la investigación.

A partir de la *revisión bibliográfica*, se concretan los *objetivos* y el *marco teórico* de la investigación. Para alcanzar cada uno de los *objetivos específicos* se procedió de la forma siguiente:

- *Evaluación del proceso*

Lo primero fue documentarse sobre el fundamento y el origen de la justicia de menores, instituciones y agentes colaboradores. La lectura comienza con el texto refundido de la legislación de los tribunales tutelares de menores aún vigente en España —en el momento de la investigación— (la LTTM de 1948), y prosigue con varios estudios de derecho comparado de menores en España, Alemania Federal, Francia, Inglaterra, Italia y Estados Unidos.

De estos estudios se extrajeron los puntos que diferencian y asemejan a los diversos sistemas legislativos en materia de creación, organización, competencia y funcionamiento de los tribunales e instituciones auxiliares; además de su concepción de la conducta delictiva y de la minoría de edad penal, como fundamento y razón de ser de la jurisdicción de menores.

Después se procede a analizar la práctica real del Tribunal Tutelar de Menores de Madrid (cómo opera en realidad). Los criterios seguidos para determinar la *muestra* de la investigación fueron los siguientes:

- Como el objetivo de la investigación era evaluar la intervención del Tribunal, la población cuya actividad delictiva “registrada” resultaba episódica y, por tanto, no relevante (al menos hasta la mayoría de edad penal) no es de interés. En cambio, sí interesan aquellos jóvenes que, bien por sus características personales, bien por la entidad de los hechos cometidos (o denunciados), quedaban bajo la tutela permanente del Tribunal (ya como primer *tratamiento*, ya después de medidas leves, como la *amonestación*).
- La *evaluación del impacto* de un *programa* requiere que la intervención del *programa* haya concluido, y no en fecha reciente. En estudios de delincuencia, el plazo de tiempo normal para medir la reincidencia es dejar transcurrir al menos dos años después de haber finalizado el *tratamiento* aplicado. Por esta razón, se excluyen del estudio los expedientes de menores abiertos o cerrados con posterioridad a 1986 (fecha de comienzo de la investigación, que duró cuatro años).
- La *muestra* debía de ser amplia y lo más representativa posible de la población de la que se extrae, con objeto de que los resultados del estudio puedan hacerse extensivos a la población. Asimismo, debía abarcar los diversos momentos históricos vividos en fechas próximas y que, probablemente, habrían repercutido en la política de tratamiento y prevención de la delincuencia juvenil. Todo ello contribuye a la *validez externa* del *diseño de investigación*.

Con estos requisitos en mente, se consultaron las estadísticas propias del Tribunal de Menores de Madrid, referentes a su actuación desde 1975 hasta 1983 (de fechas posteriores se carecía de información estadística). En las estadísticas se observaron cambios de tendencia en la política de tratamiento de la delincuencia desde 1975. Por lo que se incluyeron todos los casos con expedientes abiertos en 1975, 1977, 1979, 1981 y 1983. Éstos debían también reunir la condición de haber estado bajo tutela del Tribunal con anterioridad a 1986. La *muestra* final de la investigación la integraron 715 expedientes de menores que habían estado en libertad vigilada y/o internamiento antes de 1986.

De los distintos informes adjuntados al expediente (de la Policía, del Servicio de Orientación Psicopedagógico del Tribunal, del delegado de Libertad Vigilada y del Centro de Reforma o de educación, además de las diligencias previas y el *acuerdo* del Tribunal) se obtuvo información para describir:

- a) El *cliente del programa*: las características sociodemográficas y delictivas del menor de reforma. La aplicación de *técnicas de análisis multivariable* (de conglomerados, discriminante y factorial confirmatorio) contribuyó a la diferenciación de tres tipos de menores, rompiéndose el estereotipo del menor delincuente común.
- b) El *proceso* de aplicación de la norma legal, desde la instrucción del procedimiento penal hasta el *acuerdo* de medida. Se contrasta la información obtenida del expediente (lo que “es”), con lo dicho al respecto en la normativa legal vigente (lo que “debería ser”).

En concreto, se analizaron los distintos criterios seguidos por el Tribunal en la adjudicación de *tratamientos* a diferentes menores (a partir de la instrucción del procedimiento y del *acuerdo* tomado finalmente por el juez).

La aplicación del análisis *logit* permitió obtener las variables más relacionadas con la probabilidad de imponer al menor los distintos *tratamientos* (medidas leves, libertad vigilada e internamiento).

- c) *Aplicación del tratamiento*. La *evaluación del proceso* concluyó con la descripción de cómo se ejecuta realmente el *programa*: si los diferentes *tratamientos* acordados por el Tribunal se aplicaron en conformidad con las normas establecidas. Para ello se contrastó, nuevamente, la información extraída del expediente con las disposiciones expresas en el texto legal (LTTM de 1948).

• *Evaluación del impacto*

Toda *evaluación del impacto* debe, primero, definir claramente cuáles son los criterios de éxito y de fracaso del *programa*; y, segundo, diferenciar aquellos que se deben al programa de los causados por otros factores (McQuay, 1979).

Definir los criterios de éxito de la intervención judicial no resulta difícil. Habrá éxito siempre y cuando el *tratamiento* proporcionado consiga prevenir la *reincidencia* en

conducta delictiva, logrando la reinserción del menor en la sociedad. El problema surge en su operacionalización.

La *no-reincidencia* hay que medirla tanto durante el tiempo de aplicación del *tratamiento* como al menos dos años después de su conclusión. Del expediente del menor se extrajo el número de detenciones policiales registradas antes y durante el período de *tratamiento*. Para obtener información de detenciones posteriores al tiempo de *tratamiento*, se acudió a otras instancias: la Dirección General de la Policía y la Dirección General de Instituciones Penitenciarias.

Uno de los problemas más difíciles de resolver cuando se evalúa el *impacto* de un *programa* consiste en determinar las causas a las que se deben los *resultados* observados (no siempre atribuibles al *programa*). A veces será factible seguir un *diseño experimental*, siempre que puedan formarse grupos aleatoriamente y cuyos miembros difieran sólo en el *tratamiento* que reciban. Pero esto no siempre resulta viable. Entonces, habrá que recurrir a técnicas estadísticas de *control a posteriori* –de gran utilidad cuando no se dispone de ningún grupo de control (equivalente o no)– si quieren obviarse posibles hipótesis alternativas al *tratamiento*.

El *diseño* utilizado con mayor frecuencia en *investigaciones evaluativas* del sistema de justicia es el *diseño de un grupo de control no equivalente*: igualación de los miembros de cada uno de los grupos sólo respecto a algunas de las variables. Este *diseño*, complementado con *técnicas de análisis multivariable*, consigue –siguiendo a Crizzle y White (1980: 270)– resultados equivalentes a los que se obtendrían con el *diseño experimental*.

Para comprobar la eficacia de dos de los *tratamientos* (libertad vigilada e internamiento), primero se dividió a la población en tres grupos, de acuerdo con el *tratamiento* que hubiesen recibido: sólo libertad vigilada, sólo internamiento, y ambos *tratamientos*. Éstos no constituían grupos homogéneos (aunque pudieran asemejarse en algunos rasgos), ni habían sido formados *a priori*, de forma aleatoria. Los sujetos pasaron a uno u otro *tratamiento* en función de sus características sociodemográficas, delictivas u otras.

En esta fase de la investigación, interesaba conocer cómo incidía cada una de las circunstancias personales, sociofamiliares y legales del joven en la *reincidencia*; además de aspectos relacionados con la aplicación del *tratamiento* (tiempo de aplicación, número de diligencias policiales y de denuncias previas, durante y posteriores al *tratamiento*, adaptación del menor y motivo del cese). Sus efectos se controlaron por el tipo de *tratamiento*.

Primero, se calculó la proporción de individuos que reincidieron en conducta delictiva (para cada una de las variables consideradas). A continuación, se aplicó, de nuevo, el *análisis logit* para obtener las variables más relacionadas con la probabilidad de reincidir.

A partir de la información recabada en cada una de las fases que forman la investigación, se procedió a la *interpretación de los resultados del programa*. Los *resultados* resultaron ser bastante negativos. Dos de cada tres menores proseguían su *carrera delictiva*, después de haber abandonado la jurisdicción de menores. La información obtenida durante la *evaluación del proceso* ayudó a la búsqueda de:

- a) *Causas* posibles del elevado porcentaje de reincidencia obtenido.
- b) *Soluciones*: la modificación plena del *programa* (o conjunto de actuaciones de los Tribunales Tutelares de Menores en España).

3.3. La validez del diseño de investigación

A la hora de *evaluar* un *diseño de investigación* existen varios *criterios* a seguir. Tal vez el más fundamental es que el *diseño* se adecue a los *objetivos* principales de la investigación. Si éstos no se alcanzan, la investigación quedaría desaprobada, ya que el fin para el que se diseñó no logra satisfacerse. En la medida en que este *criterio de evaluación* no se cumpla, no procede pasar a considerar otros criterios.

Obviamente, en la consideración de si se han alcanzado los *objetivos* del estudio, también intervienen los otros dos aspectos presentes en todas las fases del *proyecto* de investigación: los *recursos* (materiales, económicos y humanos), y el *tiempo* disponible para formalizar la investigación. Estos dos aspectos son igualmente esenciales en cuanto que inciden en el buen término de la investigación y la consiguiente consecución de los *objetivos* propuestos.

Dificultades presupuestarias o problemas de tiempo para llevar a cabo el *diseño de investigación* original, pueden llevar a su modificación e, incluso, poner en peligro el logro de los *objetivos* fijados al inicio de la investigación.

Una vez que se cumple el criterio de los objetivos, se pasaría a analizar otros *criterios de evaluación*. De nuevo, la aportación de Campbell y Stanley (1970), Cook y Campbell (1977) y Reichardt y Cook (1979), merecía destacarse.

Estos autores propusieron cuatro *criterios de validez* en la *evaluación de los diseños de investigación cuantitativa*:

- a) Validez interna.
- b) Validez externa.
- c) Validez de constructo.
- d) Validez de conclusión estadística.

3.3.1. Validez interna

Hace referencia a la posibilidad de establecer relaciones de *causalidad* entre *variables* (*dependientes e independientes*), al haber eliminado (o *controlado*) otras *explicaciones alternativas*. De ahí que la comprobación de este tipo de *validez* en un *diseño de investigación* sea prioritaria, sobre todo en los *diseños explicativos*, más que en los *descriptivos*.

De acuerdo con Campbell y Stanley (1970), la *validez interna* constituye el “*sine qua non*” de cualquier indagación empírica. Si este criterio de *validez* no se satisface,

los resultados de la investigación serían cuestionables. Siempre surgirían posibles *explicaciones alternativas* a las relaciones observadas.

EJEMPLO DE VALIDEZ INTERNA

En una investigación sobre el *rendimiento académico* (medido por la calificación obtenida en un examen), quiere conocerse qué variables inciden más en la consecución de una buena calificación. De las distintas variables analizadas se observa la existencia de *relación positiva* entre las variables *horas de estudio*, a la semana, dedicadas a la asignatura (*variable independiente*) y *calificación* obtenida en el examen (*variable dependiente*): la calificación resulta más elevada, conforme aumenta el número de horas de estudio.

Para que esta relación sea *válida*, habría que *controlar* el efecto de *terceras y cuartas variables* que pudiesen mediar en la relación observada. Cuanto mayor sea el número de *variables perturbadoras* que el investigador controle, mayor grado de *validez* adquirirá su investigación.

Ello lleva a *controlar* el efecto de otras variables en la relación observada entre las variables *horas de estudio* y *calificación*. Entre esas variables se encuentran las tres siguientes: el cociente de inteligencia, la asiduidad en la asistencia a clase, y el grado de nerviosismo durante la realización del examen.

Para que se pudiese concluir que sí existe relación entre las variables *horas de estudio* y *calificación*, tendría previamente que haberse comprobado que, indistintamente del cociente de inteligencia, o de la asiduidad en la asistencia a clase, conforme el alumno aumenta sus *horas de estudio*, alcanza una mejor *calificación académica*. En caso contrario, la relación no se sostendría. A esto se le llama *control de terceras variables* (o de *explicaciones alternativas*).

El control de *explicaciones alternativas* puede efectuarse *a priori* o *a posteriori*.

a) *A priori*, en el diseño de la investigación. Esto es posible sobre todo cuando la *estrategia de investigación* es la *experimentación*.

El *diseño experimental* conlleva la formación de *grupos de control equivalentes al experimental*, salvo en el *tratamiento* o variable cuya influencia quiere comprobarse. La selección de los sujetos para constituir cada uno de los grupos sería totalmente aleatoria. Ello contribuye a neutralizar la influencia de *variables perturbadoras* que pudieran incidir en la relación.

b) *A posteriori*, en el proceso de análisis de la información.

Este tipo de control resulta más habitual en la práctica de la investigación social (en el análisis de datos de encuesta, por ejemplo). Su consecución se logra mediante la aplicación de *técnicas de análisis bivariable y multivariable*, que permiten la formación de grupos de sujetos iguales en función de los valores de la variable que se controle.

En todos los grupos formados debe observarse las mismas variaciones, dependiendo de los valores de la *variable independiente* que se analice, para poderse afirmar que esta variable explica la ocurrencia de la *variable dependiente*.

En suma, el mayor o menor grado de *validez interna* de un *diseño de investigación* depende del *control de explicaciones alternativas* a las relaciones observadas; es decir, del número de *variables perturbadoras* cuya influencia se haya neutralizado o controlado en el proceso de investigación.

3.3.2. Validez externa

Representa la posibilidad de *generalización de los resultados* de una investigación, tanto a la *población* concreta de la que se ha extraído la *muestra*, como a otros *tiempos* y *contextos*. Pero ello exige que éstos participen de las características presentes en el contexto espacial y temporal observado.

En la práctica de la investigación social, aunque la *población* a analizar sea muy pequeña, por razones de economía (presupuestaria) y de tiempo, fundamentalmente, suele observarse sólo a una parte de dicha *población*. Esta parte constituye la *muestra* de la investigación, que deberá seleccionarse de manera que constituya una representación, a pequeña escala, de la *población* de la que se ha extraído. De ello dependerá la posibilidad de que los *resultados* de la investigación puedan generalizarse y hacerse extensibles a la *población* de referencia.

Además del número de casos observados, la *representatividad de la muestra* (y la consiguiente *validez externa* de la investigación) también estará subordinada al procedimiento seguido en la selección de los elementos de la *muestra*.

Se recomienda seguir, preferiblemente, *procedimientos de selección aleatorios o probabilísticos* (que se exponen en el Capítulo 5). Cuando ello no sea factible (por carecer de un *marco de muestreo* adecuado, o por otra razón), procurar que en la *muestra* escogida se hallen incluidas las diferentes variedades de casos que componen el *universo o población de estudio*.

3.3.3. Validez de constructo

Hace referencia al grado de adecuación conseguido en la medición de los *conceptos* centrales de la investigación.

Como ya se ha hecho mención, cualquier *concepto* permite distintas posibilidades de medición. Asimismo, cualquier operacionalización de un *concepto* es difícil que cubra todas las *dimensiones del concepto*. En consecuencia, habrá que procurar operacionalizar los *conceptos teóricos* lo más rigurosamente posible, al menos los *conceptos* que sean fundamentales en la investigación. Ello contribuirá a reducir la duda de si se

habrían alcanzado los mismos resultados con una *operacionalización* distinta del *concepto*.

Para dicho propósito, se recomienda efectuar una *operacionalización múltiple*, porque permite una mejor aproximación al significado real del *concepto*. De lo que se trata es de buscar una serie de *medidas* (dos o más) para cada *concepto*. Pero, antes de afirmar que los hallazgos del estudio son válidos o inválidos, dependiendo de si existe convergencia (o divergencia) entre los resultados alcanzados en las distintas *mediciones*, ha de comprobarse si, realmente, se ha medido el mismo *concepto*.

EJEMPLO DE VALIDEZ DE CONSTRUCTO

Si la *actividad delictiva* se mide combinando *datos policiales* con la información obtenida mediante *encuesta* (tanto de *encuestas de victimización* como de las denominadas de *delincuencia autodenunciada*), lo más probable es que los resultados alcanzados, mediante estos tres procedimientos, no coincidan. La razón se halla en que en cada procedimiento mide un *concepto* distinto de *delincuencia*:

- a) Las *encuestas de autodenuncia* miden la actividad delictiva de personas que no han sido todavía etiquetadas oficialmente como "delincuentes". Generalmente, hacen referencia a infracciones triviales.
- b) Los *informes oficiales* (de la policía, tribunales) se refieren a infracciones graves. Suelen utilizarse para analizar la actuación judicial y/o policial, más que para medir la comisión de actividades delictivas.
- c) Las *encuestas de victimización* aluden a actos delictivos cometidos por otros.

En consecuencia, los datos no pueden coincidir porque las tres medidas empleadas en realidad no miden el mismo *concepto*. De ahí la insistencia en que cuando se diseñe una investigación, se procure asegurar la equivalencia de las distintas mediciones de un mismo *concepto*. Cuando esta exigencia no se cumpla, los resultados logrados de cada una de ellas no serán comparables.

3.3.4. Validez de conclusión estadística

Esta última variedad de *validez* se halla relacionada con el poder, adecuación y fiabilidad de la *técnica de análisis de datos* aplicada.

Como la práctica de la investigación social se fundamenta en la información extraída de una *muestra*, la adecuación del *tamaño muestral* (normalmente superior a 100 casos) con la *técnica analítica* utilizada (máxime en *técnicas de análisis multivariables*, que exigen una determinada proporción de casos por variables introducidas en el análisis) incide en la *significatividad estadística* de los *resultados de la investigación*; en la

posibilidad de generalización de los hallazgos obtenidos en la *muestra* a la *población* de la que ésta se ha extraído.

Además, antes de utilizar una *técnica analítica* concreta, ha de comprobarse si la información recabada satisface los supuestos exigidos para la práctica de la técnica de análisis elegida. En las *técnicas de análisis multivariable*, los supuestos más habituales son los de *normalidad, homocedasticidad e independencia* de los *términos de error*. Éstos se hallan relacionados con el *tamaño de la muestra* y el *nivel de medición de las variables*.

En el Capítulo 9 se contemplan distintos aspectos a considerar en el *análisis estadístico* para que la investigación cumpla requisitos de *validez*.

• *Síntesis de los criterios de validez*

La optimización de cada uno de estos *criterios de validez* en un mismo diseño de investigación no resulta una tarea fácil, en la práctica. Primar la *validez interna*, por ejemplo, puede conllevar la disminución de la *validez externa*, ya que ambos *criterios de validez* se contraponen. Esto es, la primacía del *control experimental* lleva consigo como contrapartida, la disminución de la capacidad de generalización de los resultados de la investigación. Mientras que en los *diseños explicativos* se prima la *validez interna*, en los *descriptivos* la *externa*, máxime cuando se trabaja con *muestras*.

No obstante, puede lograrse un *diseño de investigación* que consiga un grado aceptable en los cuatro *criterios de validez* señalados. Para dicha consecución, conviene tener presente las recomendaciones que figuran resumidas en el Cuadro 3.4.

CUADRO 3.4. Recomendaciones para aumentar la validez de un diseño de investigación.

<i>Validez interna</i>	<i>Validez externa</i>	<i>Validez de constructo</i>	<i>Validez estadística</i>
Creación de varios grupos de comparación equivalentes al de observación.	Selección de las unidades de la muestra mediante procedimientos aleatorios.	Delimitación clara y precisa de los conceptos teóricos.	Aumentar el tamaño de la muestra.
Efectuar varias mediciones.	Formar grupos heterogéneos de unidades de observación que incluyan varios contextos temporales y espaciales.	Operacionalización múltiple de los conceptos.	Formar grupos internamente poco heterogéneos.
Controlar todo suceso externo e interno a la investigación que puedan afectar a sus resultados.		Empleo de varias técnicas de obtención de información.	

Lecturas complementarias

- Alvira, F. (1994): "Diseños de investigación social: criterios operativos", en García Ferrando, M. et al. (comps.), *El análisis de la realidad social*, Madrid, Alianza Universidad, pp. 87-112.
- Babbie, E. (1992): *The practice of social research*. California, Wadsworth Publishing Company.
- Campbell, D. y Stanley, J. (1970): *Diseños experimentales y cuasiexperimentales en la investigación social*. Buenos Aires, Amorrortu.
- Hakim, C. (1994): *Research design: strategies and choices in the design of social research*. London, Routledge.
- Hernández Sampieri, R. et al. (1991): *Metodología de la investigación*. México, McGraw Hill.
- Pons, I. (1993): *Programación de la investigación social*. Madrid, CIS, Cuaderno Metodológico n.º 8.

Ejercicios Propuestos

1. Escoja una investigación y describa su diseño. En la descripción destaque si el diseño cumple criterios de validez. En caso negativo, indique qué fuentes invalidan el estudio.
2. A partir de una *idea*, formule el problema de investigación. Concrete los objetivos y el marco teórico del estudio. En su exposición, cite las referencias bibliográficas consultadas.
3. ¿Qué formas existen para controlar explicaciones alternativas? ¿De qué diseños son más características?
4. ¿En qué consiste la operacionalización múltiple? Cite un ejemplo.
5. A partir de un problema de investigación, diseñe un estudio seccional y transversal (en sus tres modalidades).
6. Durante varias semanas se observa el absentismo laboral de un grupo de personas que trabajan de noche en un hospital. Con el fin de reducir dicho absentismo, se incrementa el salario del personal del hospital que trabaja en el turno de noche. A la semana siguiente, y durante un mes, se observa el absentismo. ¿A qué tipo de diseño corresponde este ejemplo? Especifique las "amenazas" a la validez del diseño efectuado.



LA OPERACIONALIZACIÓN DE CONCEPTOS

La *operacionalización de conceptos teóricos* constituye una fase intermedia en el *proceso de investigación*. Antecede al diseño de la indagación empírica siendo, a su vez, consecutivo a la *formulación del problema* de estudio.

De la *teoría* (o *marco teórico* que conforma la investigación) se extraen unos conceptos y proposiciones. Los *conceptos* se traducen a *términos operacionales*. De ellos se deducen unas *variables empíricas* o *indicadores* que posibiliten la contrastación empírica del *concepto* que se analice.

El término “*operacionalización*” –tomado de las ciencias naturales– es el que habitualmente se emplea para denotar los estadios implicados en el proceso de asignación de *mediciones* a *conceptos*. El presente capítulo trata, precisamente, de la *práctica* de la *medición* y de las dificultades que acaecen en su desarrollo.

4.1. Fundamentos y principios de la operacionalización

En el proceso global de *operacionalización* hay que diferenciar –siguiendo a Blacklock (1982)– dos nociones fundamentales: la *conceptualización* y la *medición*.

- a) La *conceptualización* hace referencia al proceso teórico mediante el que se clarifican las *ideas* o *constructos teóricos*. Esta clarificación ha de hacerse de manera que la definición del *constructo teórico* comprenda el significado que se le suele asignar.
- b) La *medición* connota, en cambio, el proceso general que vincula las operaciones físicas de *medición* con las operaciones matemáticas de asignar números a objetos.

El proceso completo implicaría, en consecuencia, un triple nexo que relaciona los *conceptos teóricos* con las operaciones físicas de *medición*, y de éstas con los *símbolos matemáticos*.

“Si queremos que nuestras teorías sean generalizables a través de una variedad de entornos, o con respecto a una variedad de fenómenos, obviamente tenemos que *conceptualizar* nuestras variables de forma que las proposiciones que contengan estas variables puedan aplicarse en tales entornos y fenómenos diversos” (Blalock, 1982: 29).

Ello es necesario si se pretende la comparabilidad de las *operaciones de medición*. Pero, ¿qué se entiende por *medición*?

Tal vez la definición más popular de *medición* sea la proporcionada en 1951 por Stevens. De acuerdo con este autor, la *medición* consiste en la “asignación de números a objetos o acontecimientos de acuerdo con determinadas reglas” (Stevens, 1951: 22). Esta definición clásica de *medición* no es, sin embargo, compartida por la generalidad de los autores.

Duncan (1984: 126) matiza que la asignación de números debe hacerse de modo que los números “se correspondan con diferentes grados de *cualidad* –o propiedad– de algún objeto o evento”.

Otros autores, como Carmines y Zeller (1979), van más allá. Rechazan la definición de *medición* de Stevens, por considerar que se adecua a las ciencias físicas, pero no a las ciencias sociales. En su opinión, muchos fenómenos sociales son “típicamente demasiado abstractos para ser adecuadamente caracterizados o como objetos o como acontecimientos”.

Fenómenos como la *eficacia política*, la *alienación*, o la *disonancia cognitiva*, por ejemplo, son “demasiado abstractos para ser considerados cosas que pueden verse o tocarse [la definición de un objeto] o, meramente, como un resultado o consecuencia [la definición de un acontecimiento]” (Carmines y Zeller, 1979: 10).

La *medición* debería, en cambio, comprenderse –siguiendo a Carmines y Zeller (1979)– como el proceso de vincular *conceptos abstractos* a *indicadores empíricos*. Comprende, por tanto, consideraciones teóricas y empíricas.

- a) Desde el *punto de vista teórico*, el interés se halla en el *concepto*. Éste no es observable (y, por tanto, no es directamente medible), sino *latente*. Es decir, se halla representado por una respuesta que sí es *observable*.
- b) Desde el *punto de vista empírico*, la respuesta observable constituye el centro del proceso, tanto si ésta adquiere la forma de respuesta a una pregunta incluida en un *questionario* o *entrevista*, o la forma de una conducta grabada en un *estudio observacional*, por ejemplo.

En suma, la *medición* se centra “en la relación crucial entre el indicador(es) fundamentado teóricamente –eso es, la respuesta observable– y el concepto(s) no ob-

servable, latente” (Carmines y Zeller, 1979: 10). El problema se encuentra, precisamente, en la concreción de esa correspondencia (que debe haber entre el sistema conceptual y el empírico), para que se alcance una *medición* válida y fiable del fenómeno que se analiza.

Los *conceptos* pueden definirse como “símbolos lingüísticos que categorizan a los fenómenos” (Phillips, 1985: 77). Con frecuencia derivan de *reflexiones teóricas* (a las que se accede mediante una *revisión bibliográfica*); otras veces, proceden de *reflexiones propias* sobre la variedad de experiencias de la vida social.

Sea como fuere, se caracterizan por representar *constructos abstractos* y, en consecuencia, no directamente *observables* (como la “*anomia*”, la “*cohesión social*”, o la “*amistad*”, por ejemplo).

Los *conceptos* también sintetizan distintas variedades en que pueden clasificarse los objetos de conocimiento. De ahí que pueda afirmarse que los *conceptos* cumplen una función fundamental de síntesis, de denominación común, que englobe a una serie de observaciones, proporcionándolas un sentido. El *concepto* de “*tristeza*”, por ejemplo, proporciona una denominación común a una serie de manifestaciones distintas: llanto, desánimo, bajo tono de voz, negativa a hablar, comer, problemas de sueño.

Los *conceptos* difieren, esencialmente, por su mayor o menor grado de abstracción (el concepto de “*poder*”, por ejemplo, connota un grado de abstracción superior al *concepto* de “*educación*”).

Si bien, la generalidad de los *conceptos* constituyen variables “latentes”, “hipotéticas”, no directamente “observables”. Por lo que su concreción precisa de la traducción del *concepto teórico* a *indicadores*, a *variables empíricas* (*observables* o *manifiestas*), que midan las *propiedades latentes* enmarcadas en el *concepto*.

En toda *operacionalización* de *conceptos teóricos* se ha de partir de las siguientes consideraciones:

- a) Entre los *indicadores* y el *concepto* a medir ha de haber una plena correspondencia. Los *indicadores* han de seleccionarse y combinarse de manera que logren representar la *propiedad latente* que el *concepto* representa (su existencia e intensidad), en condiciones de *validez* y de *fiabilidad*.
- b) Los *indicadores* pueden materializarse en formas diversas (preguntas en un *cuestionario* o en una *entrevista abierta*, en el *registro* de una conducta observada, en *datos estadísticos* contabilizados en un censo, por ejemplo). Depende de cuál sea la *técnica de obtención de información* que el investigador haya seleccionado en el *diseño de la investigación*.
- c) En la *operacionalización*, como en todo proceso analítico, se asumen unos márgenes de incertidumbre. La relación entre los *indicadores* y la *variable latente* (el *concepto*), que tratan de medir, siempre será supuesta, nunca plenamente “cierta”; se consideran aproximaciones en términos de “probabilidad”.

Esto –como bien afirma González Blasco (1989: 236)– “evidentemente limita también el valor de la medida, pero es una limitación que hemos de aceptar si queremos medir”.

En conclusión, los *indicadores* se emplean para cuantificar, e inferir, la existencia o inexistencia de una *variable latente* (un *concepto*); aunque, siempre en términos de “probabilidad”. Se tratará, por tanto, de reducir el *error de medición* al mínimo posible.

4.2. La medición de variables: tipologías

“Los procesos de conceptualización y de operacionalización pueden verse como la especificación de variables y los atributos que las componen” (Babbie, 1992: 140).

Por *variable* generalmente se entiende cualquier cualidad o característica de un objeto (o evento) que contenga, al menos, dos atributos (categorías o valores), en los que pueda clasificarse un objeto o evento determinado.

Los *atributos* son las distintas *categorías* o *valores* que componen la variable. En función de ellos se clasifica a los objetos (o eventos) en un grupo u otro. *Variables* como la “*edad*” (años cumplidos), la “*altura*” (centímetros), o el “*nivel de ingresos*” (en pesetas), toman *valores* (numéricos). Por el contrario, *variables* como “*sexo*” (varón, mujer), “*estado civil*” (soltero, casado, viudo, separado, divorciado), o “*satisfacción conyugal*” (bastante satisfecho, satisfecho, ni satisfecho ni insatisfecho, insatisfecho, bastante insatisfecho) adoptan *categorías*.

La *medición de una variable* consiste, precisamente, en el proceso de asignar *valores* o *categorías* a las distintas características que conforman el objeto de estudio. Para que la *medición* se realice adecuadamente se recomienda, al menos, cumplir tres requisitos básicos:

a) *Exhaustividad*

La medición de la variable ha de efectuarse de forma que ésta comprenda el mayor número de *atributos* (*categorías* o *valores*) posible. El propósito es que ninguna observación quede sin poder clasificarse. De ahí la sugerencia cuando se diseña un *cuestionario*, por ejemplo, de incluir la opción de respuesta “*otros*” (especialmente en aquellas preguntas en las que caben otras respuestas diferentes a las dadas en el cuestionario), y la categoría común “*no sabe/no contesta*” (dirigida a aquellos que decidan no emitir ninguna respuesta).

b) *Exclusividad*

Los distintos *atributos* que componen la *variable* deben ser mutuamente excluyentes. Por lo que deberán definirse de manera que cualquier observación sólo pueda clasificarse en términos de un único *atributo*.

EJEMPLO DE EXCLUSIVIDAD

Para ilustrar la necesidad de cumplir el requisito de *exclusividad* en la *medición de una variable*, a continuación se comparan dos mediciones alternativas de la variable "edad".

Si esta variable se midiese como en (A), las personas de 18, 25, 50 y 65 años de edad no sabrían en qué *atributo* clasificarse, al estar estos valores incluidos en más de un intervalo.

En cambio, si la *medición de la variable* se efectuase como en (B), este problema desaparecería, cumpliéndose así el requisito de *exclusividad*.

A) Edad		B) Edad	
18 y menos	1	Menos de 18	1
18-25	2	18-25	2
25-50	3	26-50	3
50-65	4	51-65	4
65 y más	5	Más de 65	5
NS/NC	0	NS/NC	0

c) *Precisión*

Realizar el mayor número de distinciones posibles. Ello contribuye a la consecución de una información más *precisa*. Tiempo habrá para agrupar las distintas *categorías* o *valores* de las *variables*; generalmente, después de haberse recabado la información (en la fase de análisis), a la vista de la frecuencia que presente cada *atributo* de la *variable*. Pero, por el contrario, nunca será factible desglosar los *atributos* después de la obtención de los datos.

EJEMPLO DE PRECISIÓN

Si la variable "nivel de instrucción" se midiese como en (B), se obtendría una información más precisa y detallada que si se midiese como en (A).

A) Nivel de instrucción		B) Nivel de instrucción	
Sin estudios	1	No sabe leer ni escribir	1
Primarios	2	Sólo sabe leer y/o escribir	2
Profesionales	3	Primarios de 1º grado	3
Bachillerato	4	Primarios de 2º grado	4
Medios superiores	5	Formación profesional	5
Superiores	6	Bachillerato	6
		Medios superiores (Esc. Univ.)	7
		Superiores (facultades, ETS)	8
		Otros	9

Existen distintas modalidades de variables. En el Cuadro 4.1 se resumen los criterios principales de clasificación de las variables.

CUADRO 4.1. Tipologías de variables según criterios de clasificación.

<i>Nivel de medición</i>	<ul style="list-style-type: none"> • Variables <i>cualitativas</i> o <i>no métricas</i>. <ol style="list-style-type: none"> 1) Nominales. 2) Ordinales. • Variables <i>cuantitativas</i> o <i>métricas</i>. <ol style="list-style-type: none"> 3) De intervalo. 4) De razón o proporción.
<i>Escala de medición</i>	<ol style="list-style-type: none"> 1) Continuas. 2) Discretas.
<i>Función en la investigación</i>	<ol style="list-style-type: none"> 1) Independientes. 2) Dependientes. 3) Perturbadoras. <ul style="list-style-type: none"> — De control. — Aleatorias.
<i>Nivel de abstracción</i>	<ol style="list-style-type: none"> 1) Generales. 2) Intermedias. 3) Empíricas o indicadores.

• *Tipos de variables según el nivel de medición*

1) *Variables nominales*

Aquellas variables cuyos atributos sólo cumplen las condiciones esenciales de *exhaustividad* y *exclusividad*. Únicamente se hallan compuestas por distintas denominaciones, entre las que no puede establecerse ningún tipo de relación (de orden o de otra clase). Ello limita las posibilidades de análisis estadísticos en este tipo de variables.

A esta modalidad pertenecen las variables sexo, estado civil, nacionalidad, partido político, color del pelo, grupo sanguíneo, situación laboral; cualquier variable que indique una cualidad del objeto o evento que se analice, sin establecer ninguna graduación entre las distintas categorías que conforman la variable.

EJEMPLOS DE VARIABLE NOMINAL

Partido político		Facultades universitarias	
PSOE	1	CC. Políticas y Sociología	1
PP	2	CC. Económicas y Empresariales	2
IU	3	Psicología	3
CiU	4	Derecho	4
PNV	5	Ciencias de la Información	5
Otros	6	Medicina	6
		Farmacía	7
		Otras	8

Los números que se asignan a cada una de las *categorías* que forman la variable sirven para indicar la pertenencia a la *categoría*. Se asignan arbitrariamente y no denotan cantidades. Por lo que no se toman como *mediciones cuantitativas*.

2) *Variables ordinales*

Variables cuyos *atributos* participan de las características anteriormente referidas, a las que se suma la posibilidad de poderse “ordenar”, en el sentido de “mayor que” o “menor que”. No obstante, continúa sin poderse conocer la magnitud exacta que diferencia a un *atributo* de otro.

Las *variables ordinales* son, igualmente, *variables no métricas o cualitativas*. Expresan una “cualidad” del objeto o acontecimiento, no una “cantidad”.

Como ejemplos pueden citarse las variables clase social, nivel de estudios, ideología política, satisfacción laboral, calificación académica, curso académico o cualquier otra que comprenda categorías “ordenables” (en un sentido u otro).

EJEMPLOS DE VARIABLE ORDINAL

Partido político		Facultades universitarias	
Extrema izquierda	1	Católico muy practicante	1
Izquierda	2	Católico practicante	2
Centro izquierda	3	Católico no practicante	3
Centro	4	No católico pero creyente	
Centro derecha	5	de otra religión	4
Derecha	6	No creyente	5
Extrema derecha	7		

3) *Variables de intervalo*

Constituyen *variables cuantitativas* o *métricas*. Puede “cuantificarse” la distancia exacta que separa cada *valor* de la variable. Ello es posible gracias al establecimiento de alguna unidad física de medición estándar (años, pesetas, horas, minutos, centímetros, grados). Lo que posibilita que pueda afirmarse, por ejemplo, que la distancia que separa a aquellas personas de 15 y 16 años es la misma que la habida entre aquellos de 72 y 73 años. Esta capacidad permite la realización de la mayoría de las operaciones aritméticas, como se verá en el capítulo 9.

EJEMPLOS DE VARIABLE DE INTERVALO

Peso (gramos)		Nº de habitantes		Ingresos (pesetas)	
40 – 55	1	Menos de 2.000	1	Menos de 70.000	1
56 – 60	2	2.000 – 10.000	2	70.000 – 100.000	2
61 – 72	3	10.001 – 50.000	3	100.001 – 200.000	3
73 – 84	4	50.001 – 100.000	4	200.001 – 300.000	4
Más de 84	5	100.001 – 500.000	5	300.001 – 500.000	5
		Más de 500.000	6	500.001 y más	6

- Años cumplidos.....
- Tiempo que Vd. tarda en llegar a la facultad.....
- Puntuación en un test de inteligencia.....

4) *Variables de proporción o razón*

A las características del *nivel de intervalo* se suma la posibilidad de establecer un *cero* absoluto. Lo que permite el cálculo de “proporciones” y la realización de cualquier operación aritmética.

La mayoría de las *variables de intervalo* son, a su vez, *de razón* (ingresos, n.º de habitantes, n.º de veces que se asiste a un concierto, edad –los bebés tienen menos de 1 año–, por ejemplo); aunque no todas. Ello lleva a algunos autores, como Blalock (1978), a afirmar que la distinción entre *variables de intervalo* y *variables de razón* es puramente académica más que real. Una vez que se ha determinado la magnitud de la *unidad*, resulta difícil concebir la posibilidad de fijar cero unidades.

• Estos cuatro *niveles de medición* de las *variables (nominal, ordinal, de intervalo y de razón)* conforman una *escala acumulativa*. Cada *nivel* comparte las propiedades de los *niveles de medición* que le anteceden. De ahí que se recomiende:

- a) Escoger el *nivel de medición* más elevado posible, con el propósito de poder así acceder a un mayor abanico de técnicas analíticas (a aplicar en los datos que finalmente se obtengan).
- b) Tener siempre presente los objetivos de la investigación. Estos marcarán el rango de variación en la *medición*: si se requiere una información detallada o, por el contrario, genérica.

En suma, cuando el investigador tenga que medir *variables* que pueden ser a su vez *cuantitativas* o *cualitativas*, tendrá que decidir, por ejemplo, si proceder a una *medición de intervalo* u *ordinal*. Dependerá de cómo haya diseñado la investigación. Concretamente, de qué técnicas empleará para la recogida y el análisis de la información, en conformidad con los objetivos del estudio.

En general, se aconseja optar por la precisión y el detalle antes que por la generalidad. Tiempo habrá para resumir la información. Los *atributos* de la *variable* siempre podrán agruparse en *categorías* genéricas (durante la fase de análisis). En cambio, nunca podrá procederse a la inversa. Una vez recogida la información, el investigador no puede desmembrar *categorías* genéricas en *atributos* específicos.

EJEMPLOS DE VARIABLES EN DISTINTO NIVEL DE MEDICIÓN

Variables como “calificación académica” o “edad” pueden medirse comprendiendo *categorías* o *valores*. El investigador tendrá que elegir entre uno u otro según los objetivos de la investigación.

Calificación académica		Edad	
Nivel ordinal	De razón	Nivel ordinal	De razón
No presentado		Niño	0 – 13
Suspense	0 – 4	Adolescente	14 – 18
Aprobado	5 – 6	Joven	19 – 30
Notable	7 – 8	Adulto	31 – 50
Sobresaliente	9 – 10	Anciano	51 – 65
Matrícula Honor			Más de 65

Asimismo, tendrá que decidir si es suficiente conocer, por ejemplo, si la persona es de ideología de izquierdas o de derechas (*nivel ordinal*) o, por el contrario, necesita saber a qué partido votó en las últimas elecciones (*nivel nominal*). En caso de duda, se aconseja –como ya se ha indicado– anteponer el detalle a la generalidad.

- *Tipos de variable según la escala de medición*

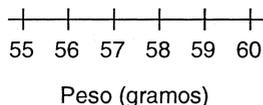
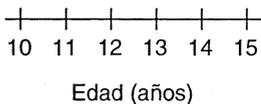
Un segundo criterio de clasificación de las variables responde a si en la *medición de la variable* se ha aplicado una *escala continua* o una *discreta*. Así se diferencia entre *variables continuas* y *discretas*.

1) *Variables continuas*

Aquellas variables en las que pueden hallarse *valores* intermedios entre dos *valores* dados, al conformar una escala ininterrumpida de *valores*.

EJEMPLOS DE VARIABLE CONTINUA

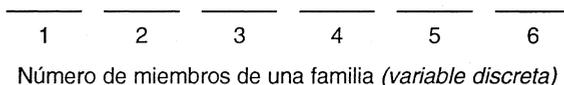
La variable "edad" es *continua* porque entre un año y otro caben valores intermedios. Así, entre los 12 y los 13 años hay infinitos valores: como 12 años, 8 meses y 15 días, por ejemplo. Lo mismo cabe decir de la variable "peso" o cualquier otra variable que sea *cuantitativa (métrica)*. Si bien, hay que matizar que no todas las variables de *intervalo* o de *razón* son, a su vez, *continuas*. Pueden ser *discretas*, como después se verá.

2) *Variables discretas*

Cuando en la escala de medición de la variable no cabe la posibilidad de hallar valores intermedios, comprendidos entre dos *atributos* de la *variable*.

La generalidad de las *variables* denominadas *cualitativas (nominales y ordinales)* son *discretas*. También lo son algunas *cuantitativas*, como el n.º de miembros de una familia (no puede haber 2 hijos y medio), o el n.º de coches vendidos, por ejemplo.

EJEMPLOS DE VARIABLE DISCRETA Y CONTINUA



<i>Variables discretas</i>	<i>Variables continuas</i>
Nº libros comprados	Nº libros leídos
Nº barómetros	Temperatura atmosférica registrada
Nº mesas en un aula	Longitud de las mesas
Entradas de cine pagadas	Películas vistas

• *Tipos de variables según su función en la investigación*

Las variables también difieren según el papel que cumplen en una investigación. Atendiendo a este tercer criterio de clasificación, se distingue entre *variables independientes, dependientes y perturbadoras (de control y aleatorias)*.

1) *Variables independientes, explicativas o predictoras (X)*

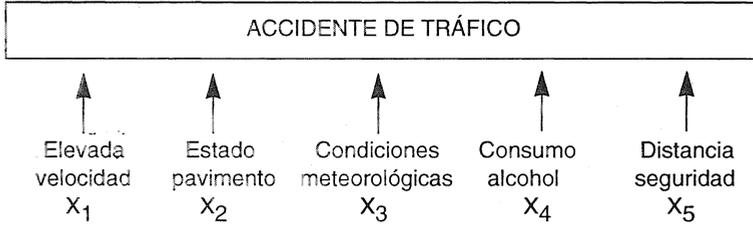
Aquellas variables cuyos *atributos* se supone que influyen en los que adopta una segunda variable (la *dependiente*). Figuran en las hipótesis de la investigación e indican posible “causas” de la variación de la variable que centra el interés de la indagación (la *dependiente* o *variable efecto*).

2) *Variables dependientes o criterio (Y)*

Variables cuyos atributos “dependen” –como su nombre indica– de los que adoptan las *variables independientes*. Ambos tipos de variables corresponden a los objetivos de la investigación.

EJEMPLO DE VARIABLES INDEPENDIENTES Y DEPENDIENTES

Un equipo de investigadores desea determinar las variables que inciden en la ocurrencia de accidentes de tráfico. En esta investigación, el sufrir un “accidente de tráfico” actuará como la *variable dependiente* (la que constituye el objeto de análisis); mientras que las posibles “causas” de la siniestralidad serán las *independientes* (conducir a elevada velocidad, el estado del pavimento, las condiciones meteorológicas, el consumo de alcohol, la experiencia en la conducción, la edad del conductor, el guardar la distancia de seguridad).



3) *Variables perturbadoras*

En la relación entre una *variable independiente* y la *dependiente* siempre cabe la posibilidad de que existan otras variables mediando en la relación. Ello contribuye a la existencia de explicaciones alternativas que hagan espúrea la relación observada entre la *variable dependiente* y la *independiente*.

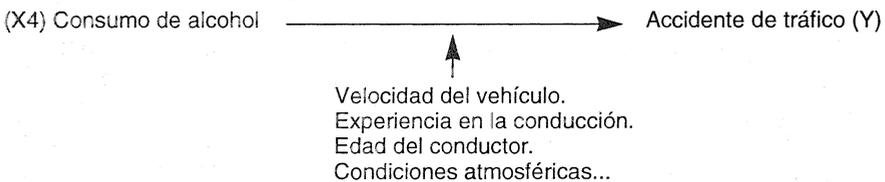
Si el efecto de esas terceras o cuartas variables se controla (bien antes *-a priori-*, o bien después de haberse recogido los datos *-a posteriori-*), dejan de ser *perturbadoras* y se convierten en *variables de control*. Como ya se expuso en el Capítulo 3, en toda investigación hay que procurar controlar el mayor número de *variables perturbadoras* posible. Ello favorece la *validez interna* de la investigación.

EJEMPLO DE VARIABLES DE CONTROL

Retomando el ejemplo anterior, escójase una de las *variables independientes* barajadas: el “consumo de alcohol”, por ejemplo. De la información reunida en la investigación se concluye que una alta proporción de las personas que sufrieron un accidente de tráfico presentaban un elevado índice de alcoholemia en la sangre.

Para que el investigador pueda afirmar que el consumo de alcohol (X₄) incide en la siniestralidad (Y), previamente deberá haber “controlado” el efecto de otras variables que pueden estar mediando en dicha relación. Por lo que tendrá que indagar en variables que incidan diferencialmente en personas con similares índices de alcoholemia.

Algunas de las variables que actuaban como *independientes* pueden convertirse, a su vez, en *variables de control* (cuando se analiza la influencia de otra *variable independiente*); y, a la inversa: una *variable de control* que haya mostrado incidencia pasa a ser (en la misma investigación) *variable independiente*, cuya influencia específica en la *variable dependiente* deberá cuantificarse.



Las *variables aleatorias* o *estocásticas* son, también, *variables perturbadoras*, pero de menor incidencia en el conjunto de la investigación. Su efecto en la relación entre la *variable independiente* y la *dependiente* puede reducirse siguiendo un procedimiento *aleatorio* de selección de la *muestra* (los casos a observar). La *aleatorización* favorece la heterogeneidad en la composición de la *muestra*, como se verá en el Capítulo 5.

EJEMPLO DE VARIABLES ALEATORIAS

En el ejemplo anterior, quizás algunas de las personas con un índice elevado de alcoholemia sufriera el accidente al maniobrar para salvar un obstáculo (para no atropellar a un perro, por ejemplo). La proporción de casos (respecto del total analizado), en los que dicha circunstancia concurre, será muy baja (apenas unos casos). Por esta razón, la no consideración de la incidencia de esta variable en la investigación no introduciría sesgos importantes en los resultados de la indagación. De ahí su calificación como *variable aleatoria* en esta investigación.

• Tipos de variables según su nivel de abstracción

Un último criterio de clasificación hace referencia al nivel de abstracción de la variable. De acuerdo con él, existen *variables generales*, *intermedias* e *indicadores*.

1) Variables generales

Aquellas variables que son tan genéricas y abstractas que no pueden ser directamente observadas. Su medición exige que se traduzcan a *variables intermedias* e *indicadores*.

Un ejemplo típico de *variable genérica* lo representa la variable “status social” porque necesita de *indicadores* concretos que ayuden a su medición.

2) Variables intermedias

Expresan alguna *dimensión* o aspecto parcial de los comprendidos en la *variable genérica*.

Por ejemplo, el “nivel educativo” para la medición de la variable “status social”.

3) Indicadores o variables empíricas

Representan aspectos específicos de las *dimensiones* que comprende un *concepto abstracto* o *variable genérica*. Se distinguen por ser directamente medibles.

Por ejemplo, los “cursos académicos cumplidos” como *indicador* para la *dimensión* “nivel educativo”.

En la siguiente sección se desarrollará el *proceso de operacionalización de conceptos teóricos*: cómo se pasa de *variables generales* a *indicadores* e *índices*. El propósito es transformar el *concepto teórico* en observable.

4.3. De los conceptos teóricos a los indicadores e índices

Para medir la ocurrencia de *conceptos teóricos*, se procede a su *operacionalización*. Primero, proporcionando una *definición operativa*, que comprenda el significado determinado que se da al *concepto*. Segundo, especificando los *indicadores empíricos* que representarán a los *conceptos teóricos*, y se utilizarán para medir el grado de existencia del *concepto* en determinados contextos.

En la indagación científica, la delimitación de los *conceptos* depende de dos tipos de definiciones: la *definición nominal* y la *operacional*.

- a) Una *definición nominal* o *teórica* es aquella que se asigna a un *concepto* (para definirlo y perfilarlo), pero que carece de las precisiones necesarias para medir los fenómenos a los que hace referencia el *concepto*.

Por *concepto* puede entenderse –siguiendo a Bollen (1989)– cualquier idea que vincule diferentes fenómenos (actitudes, comportamientos) bajo una misma etiqueta. Por ejemplo, el concepto de “alegría” reúne las distintas manifestaciones (verbales y físicas) de los individuos: reír, cantar, gritar, dar saltos, besos, abrazos.

- b) La *definición operacional*, en cambio, especifica cómo se medirá la ocurrencia de un *concepto* determinado en una situación concreta. En dicha definición se detallará el contenido del *concepto* que va a medirse, vinculando los *constructos* definidos teóricamente con los *procedimientos operacionales*.

Teóricamente, la *definición operacional* consiste en “la identificación de operaciones de investigación con ayuda de las cuales se puede decidir acerca de la presencia y la intensidad de aquellos hechos que permitirán la deducción de la presencia de los fenómenos conceptualmente caracterizados” (Mayntz *et al.*, 1983: 28).

EJEMPLOS DE DEFINICIÓN TEÓRICA Y OPERACIONAL

- *Concepto*: “El absentismo (escolar)”.
- *Definición teórica*: “Ausencia de la escuela en horario escolar”.
- *Definición operacional*: “Número de días al mes que el alumno falta a la escuela sin justificación”.

La *operacionalización de conceptos teóricos* fue pioneramente desarrollada por Lazarsfeld (1973a). Este destacado metodólogo distinguió las siguientes *fases* en el desarrollo de la *operacionalización*:

- 1) *Representación teórica del concepto* de forma que queden reflejados sus rasgos definitorios.
- 2) *Especificación del concepto*, descomponiéndolo en las distintas *dimensiones* o aspectos relevantes que engloba.
- 3) Para cada una de las *dimensiones* elegidas se selecciona una serie de *indicadores* (o *variables empíricas*), que “indiquen” la extensión que alcanza la *dimensión* en los casos analizados.
- 4) *Síntesis de los indicadores* mediante la elaboración de *índices*. A cada *indicador* se le asigna un *peso* o *valor*, de acuerdo con su importancia. A partir de estos *valores* se confecciona un *índice*, una medida común que agrupe a varios *indicadores* de una misma *dimensión conceptual* operacionalizada numéricamente.

A este proceso de *operacionalización* diseñado por Lazarsfeld, hay que hacer –de acuerdo con González Blasco (1989)– las precisiones siguientes:

- a) Por muchas *dimensiones* que se consideren, nunca puede abarcarse la totalidad de un *concepto*. Sobre todo, cuando el *concepto* incluye una gran variedad de aspectos.
- b) Operar con demasiadas *dimensiones* de un mismo *concepto* puede dificultar los análisis. Por esta razón resulta conveniente limitar el número de *dimensiones* a aquellas que sean más relevantes en el *concepto* considerado. Pero, como no existe ninguna regla teórica que contribuya a la delimitación del número de *dimensiones* a considerar, ésta se deja a la intuición del investigador.
- c) Después de haber escogido los *indicadores*, ha de volverse a considerar el fenómeno que se observa. Con ello se pretende comprobar si las medidas obtenidas mediante los *indicadores* reflejan los hechos observados. En caso afirmativo, se estudia la necesidad de ajustar el *concepto* que los hechos reflejen.

Los *indicadores* constituyen propiedades esencialmente *manifiestas* que, supuestamente, se hallan empíricamente relacionadas con una *propiedad latente* o no observable (*dimensión*). De ellos se exige que sean “expresión numérica, cuantitativa de la *dimensión* que reflejan” (González Blasco, 1989: 235). Por lo que, en cualquier *operacionalización de un concepto* habrá que encontrar, al menos, un *indicador* para cada *dimensión* del *concepto*.

No obstante, ha de tenerse presente que los *indicadores* representan “aproximaciones” (en términos de “probabilidad”) al *concepto* que miden. Ello se debe, pre-

cisamente, a las características de generalidad y abstracción que definen a todo *concepto*.

En consecuencia, tomar las “detenciones policiales” como único o principal *indicador* de “delincuencia”, por *ejemplo*, sería inexacto. Puede haberse cometido un acto delictivo y no haberse registrado ninguna detención policial o, a la inversa, una persona puede ser detenida sin haber cometido ningún acto delictivo. De ahí la recomendación –ya reiterada en capítulos precedentes– de acudir a la *operacionalización múltiple* como recurso para aumentar la *validez* de una investigación (*validez de constructo*).

Las razones a favor de la *operacionalización múltiple* (o *medición* de un mismo *concepto* por medios diferentes) pueden resumirse en dos fundamentales:

- a) Utilizar varias *medidas* para un mismo *concepto* contribuye a abarcar las distintas *dimensiones* que éste incluye.
- b) Proporciona una mayor precisión y validez de la *medición*, cuando coinciden los resultados de mediciones diferentes.

La elección de *indicadores* depende de los *objetivos de la investigación*, principalmente. Además se recomienda:

- a) Confeccionar una lista con el mayor número de *indicadores* posible. Después, a la vista de la información reunida, se procederá a la eliminación de aquellos *indicadores* que resulten ser no significativos para la medición de esa *dimensión* concreta.
- b) Acudir a *indicadores* ya validados en investigaciones previas. Ello favorece la comparación de los resultados alcanzados en estudios diferentes, además de suponer una garantía mayor en su aplicación.

La adecuación de un *indicador* puede variar en distintos contextos y momentos temporales. Si se recurre a aquéllos ya validados en indagaciones previas, podrá analizarse dicha variación.

El *análisis factorial* resulta de gran utilidad en la elucidación de las distintas *dimensiones* que comprende un *concepto*, pero con posterioridad a la recogida de información.

Esta *técnica analítica multivariable* se ocupa del análisis de las interrelaciones entre un gran número de variables (tomadas de respuestas a preguntas de un *cuestionario*, por ejemplo). El objetivo es explicar dichas variables en términos de sus *dimensiones latentes comunes*, denominadas “*factores*”. La condición exigida es que el número de *factores* (o de *dimensiones latentes*) sea inferior al número de variables originarias en el análisis.

Cada *factor* queda definido por las variables que muestren una mayor correlación respecto a él que hacia los otros *factores*. El estadístico más empleado para medir el

5

LA SELECCIÓN DE LAS UNIDADES DE OBSERVACIÓN: EL DISEÑO DE LA MUESTRA

La selección de las *unidades de observación* constituye un estadio primordial en todo proceso investigador. De cómo se materialice dicha elección dependerá, en buena medida, la calidad de la información que se recoja. Razón por la cual se recomienda al investigador que no escatime, ni tiempo ni esfuerzo, en la planificación y ejecución del *diseño de la muestra*.

Pero, ¿en qué consiste el diseño muestral?; ¿cómo hay que proceder para la consecución de una *muestra* apropiada?; ¿qué es una *muestra*?; ¿qué variedades de *diseños muestrales* existen? Los apartados siguientes tratarán de dar respuesta a éstas y otras cuestiones relacionadas con las *unidades de observación* en el proceso de investigación.

5.1. Fundamentos y clarificación terminológica

Una de las primeras decisiones a tomar en cualquier investigación es la *especificación y acotación de la población* a analizar. La concreción de ésta vendrá determinada por cuál sea el *problema* y los *objetivos* principales de la investigación.

Por *población* (o *universo* de estudio) comúnmente se entiende “un conjunto de unidades, para las que se desea obtener cierta información” (Sánchez Crespo, 1971: 11). Las *unidades* pueden ser personas, familias, viviendas, escuelas, organizaciones, artículos de prensa...; y la *información* vendrá dictada por los *objetivos* de la investigación (condiciones de habitabilidad de la vivienda, reacción familiar ante el fracaso escolar, la práctica de la religión en las escuelas).

En la *definición y acotación de la población* han de mencionarse características esenciales que la ubiquen en un espacio y tiempo concreto.

EJEMPLOS DE DEFINICIÓN DE LA POBLACIÓN (O UNIVERSO) DE ESTUDIO

- a) En una investigación sobre la ocupación del tiempo tras la jubilación, una posible definición del *universo* de estudio sería la siguiente:

“Población de 65 y más años que residen en el municipio de Madrid”.

Además, habría que precisar si se analizará a toda la *población* (indistintamente del lugar donde reside) o si, por el contrario, se excluirá a aquellos que, en el momento de la investigación, se hallen viviendo en instituciones (como residencias para la tercera edad). Depende que cuáles sean los objetivos del estudio.

- b) Si el objetivo fundamental de la investigación fuese analizar la corrupción política a través de la prensa escrita, una definición de la *población* sería:

“Artículos de prensa publicados en los periódicos *El País*, *ABC* y *El Mundo* entre 1990 y 1995”.

Una vez definida la *población*, se procede al *diseño de la muestra*: la selección de unas *unidades* concretas de dicha *población*.

Aunque el *universo* fuese de pequeña dimensión, por razones de economía (en tiempo y dinero), rara vez se observa a cada una de las *unidades* que lo forman. Por el contrario, se decide la extracción de una *muestra* de entre los integrantes del *universo*. Si bien, en este hacer también existen divergencias. Depende, fundamentalmente, de la *estrategia de investigación* que se haya escogido para la consecución de los objetivos del estudio.

Un *estudio de casos* o un *experimento*, por ejemplo, imponen menos exigencias en la *muestra* que una *encuesta* (más motivada a la consecución de una *muestra* “representativa”). Dicha *representatividad* estará subordinada al *tamaño de la muestra* y al procedimiento seguido para la selección de las *unidades muestrales* (si éste garantiza la *inclusión de toda la variedad de componentes del universo*).

Si, a partir de los datos obtenidos en una *muestra*, quieren inferirse las características correspondientes de la *población* (*parámetros poblacionales*), es imperativo diseñar una *muestra* que constituya una representación a pequeña escala de la *población* a la que pertenece. Para ello resultan esenciales las aportaciones llevadas a cabo en el campo de la estadística. Concretamente, en la Estadística Inferencial, configurada a partir de la teoría de las probabilidades. En ella se fundamentan los *diseños muestrales probabilísticos*, como después se verá.

Cualquier *diseño muestral* (máxime los *probabilísticos*) comienza con la *búsqueda de documentación* que ayude a la identificación de la *población* de estudio. Con el término *marco* se hace referencia al “listado” que comprende las unidades de la *población*. Puede ser un censo general de la población, un registro de individuos, o cualquier otro procedimiento que lleve a la identificación de los miembros de una *población*. De

él se espera que sea un descriptor válido de dicha *población*. Por lo que ha de cumplir una serie de *requisitos* mínimos:

- a) El *marco* ha de ser lo más completo posible, en orden a facilitar la representatividad de la *muestra*.

Téngase presente que la *muestra* escogida sólo podrá considerarse “representativa” de la *población* comprendida en el *marco de muestreo* elegido; es decir, de aquellos que han tenido probabilidad de participar en la *muestra*. Por esta razón, la *comprehensividad* se convierte en una exigencia básica de todo *marco muestral*.

EJEMPLOS DE COMPREHENSIVIDAD DEL MARCO DE MUESTREO

- a) La guía telefónica constituye el *marco de muestreo* habitual en una *encuesta telefónica*. Pero, su uso impone una serie de restricciones en la *muestra* que el investigador ha de conocer para la acotación de la *población* a analizar.

En primer lugar, la guía telefónica (aunque esté actualizada y completa) limita la *población* a las personas con un número de teléfono “registrado”. Por tanto, excluye a aquellos que carecen de teléfono. A éstos se suman los que, aun teniéndolo, no figuran en el listín telefónico (ya por expreso deseo del abonador de no figurar, ya por haberse suscrito en fechas posteriores a la última actualización de la guía, o por otra razón).

En segundo lugar, las guías no registran a todos los integrantes de la vivienda con teléfono. Sólo aparece uno de ellos; generalmente, el cabeza de familia. Este hecho puede llevar a una mayor presencia de varones en la *muestra*, si ésta se limita a las personas cuyo nombre figura en la guía telefónica.

También, se ha de tener presente que las personas con más de un número de teléfono tienen una probabilidad mayor de ser elegidas para formar parte de la *muestra*.

Todas estas circunstancias condicionan las posibilidades de generalización de la información obtenida en una *encuesta telefónica*. Esta únicamente podrá generalizarse a la *población* registrada en el *marco de muestreo* empleado; no al conjunto de la *población* que resida en el espacio geográfico que abarque el estudio.

- b) Si la finalidad de la investigación fuese conocer la opinión de los sociólogos españoles sobre su actividad profesional, un *marco de muestreo* idóneo sería el directorio de doctores y licenciados en sociología colegiados. Pero su uso también impone restricciones importantes en la *población* de estudio. La opinión que se recabe sólo podría ser “representativa” de los sociólogos que estén colegiados (y no de la totalidad de los sociólogos españoles). Por lo que, habría que comprobar qué proporción representan los colegiados en el conjunto de la *población*.

La “representatividad” de la *muestra* también se verá afectada (aun limitándose a la *población* colegiada) por la actualización del *marco de muestreo*. Si éste no incluye a los sociólogos que se hayan colegiado en fechas recientes, y, en contra, no

excluye a aquellos que han dejado de estar colegiados (por fallecimiento u otra causa), la *muestra* podrá incluso no ser “representativa” del colectivo de sociólogos colegiados. Ello dependerá de la magnitud de los excluidos y de los incluidos erróneamente.

Si en esta segunda investigación el investigador desease una mayor “generalización” de los resultados de su indagación, tendría que recurrir a otros *marcos muestrales*: un listado de profesores de sociología, un registro de sociólogos que hayan concursado en oposiciones para las administraciones públicas, por *ejemplo*. Estos *marcos muestrales* complementarían la información proporcionada por el *marco de muestreo* principal.

En suma, cualquier *marco de muestreo* se halla limitado a un conjunto de la *población*, siendo además vulnerable a la presencia de omisiones. Del volumen de éstas, junto a la capacidad del investigador de detectarlas y, en su caso, eliminarlas, dependerá la incidencia del *marco de muestreo* en la aparición de *sesgos* en la *selección de la muestra*.

- b) La *comprehensividad* del *marco muestral* conlleva, necesariamente, la exigencia de su *actualización* –como ya se ha hecho mención–.

En la medida en que el *marco muestral* se halle actualizado (en cortos períodos de tiempo, preferentemente) las posibilidades de omisiones se restringen. Por el contrario, aumenta la probabilidad de que éste contenga a los miembros reales de la población que “representa”.

EJEMPLO DE ACTUALIZACIÓN DE UN MARCO DE MUESTREO

Si en 1989 se pasase una *encuesta* a la población residente en Madrid, el *marco muestral* preferible sería el *Padrón de habitantes de 1986*. Este reflejará mejor el perfil de la población madrileña –a fecha de 1989– que el último *censo de población* (que a dicha fecha sería el de 1981).

- c) Cuando la investigación persigue la *generalización de los datos muestrales* (a la *población* que conforma el *marco muestral*), es preciso que cada componente de la población esté igualmente representado en el *marco de muestreo*.

Han de evitarse las *duplicidades*. Éste es un problema habitual cuando, para una misma *muestra*, se combinan dos o más listados diferentes.

Aquéllos que (por cualquier motivo) aparezcan más de una vez, tendrán una oportunidad superior de ser elegidos. Lo que favorece su sobrerrepresentación en la *muestra*. Por esta razón deberá eliminarse su duplicidad.

- d) El *marco muestral* tampoco ha de incluir *unidades* que no correspondan a la *población* que se analiza. La inclusión de estas *unidades* reduce la probabilidad de elección de las *unidades* que sí pertenecen a la *población*.

EJEMPLO DE SUPRESIÓN EN UN MARCO DE MUESTREO

Si se hiciera una encuesta a la población mayor de 40 años, habría que circunscribir la *población* a esta cuota de edad. Las personas de 40 años y menos deberían eliminarse del *marco muestral*. Su supresión podría hacerse *a priori* (antes de proceder a la extracción de la *muestra*), o *a posteriori* (una vez que la *muestra* ha sido seleccionada). Esta segunda opción suele ser más practicada. Consiste en sustraer (de la *muestra* obtenida) aquellas *unidades* que no pertenezcan a la *población* de interés.

- e) El *marco muestral* ha de contener información suplementaria que ayude a la localización de las *unidades* seleccionadas: teléfono, dirección.
- f) Ante todo, el *marco* debería ser fácil de utilizar. Aquellos que están en soporte magnético simplifican y automatizan bastante el proceso de extracción muestral. No sólo reducen los costes del diseño de la *muestra* (en tiempo y en dinero), sino que también contribuyen a la reducción de errores (o *sesgos*) en la *selección de la muestra*.

Por todos los aspectos señalados, es importante que el investigador evalúe la calidad del *marco muestral* antes de comenzar la extracción de la *muestra*. El conocimiento de cómo se ha procedido en su elaboración y la periodicidad de su actualización, reducirá los *sesgos* motivados por su desconocimiento.

Por último, insistir en que la “representatividad” de la *muestra* sólo podrá ser de la *población* contenida en el *marco muestral* empleado. De ahí que se recomiende:

“Cuando se informen resultados, el investigador debe decir a los lectores a quién se dio y a quién no se dio una oportunidad de ser seleccionado, cómo aquellos omitidos diferían de los otros y si había o no algunas personas muestreadas para quienes las oportunidades de selección no eran conocidas” (Fowler, 1988: 22).

5.2. El tamaño de la muestra

El número de unidades a incluir en la *muestra* constituye una de las decisiones preliminares en cualquier *diseño muestral*. En esta decisión participan diferentes factores. Éstos pueden resumirse en los seis siguientes:

- a) El *tiempo* y los *recursos* disponibles para llevar a cabo la investigación.
- b) La *modalidad de muestreo* seleccionada en orden a alcanzar los objetivos esenciales de la investigación.
- c) La diversidad de los *análisis de datos* prevista.
- d) La *varianza* o heterogeneidad poblacional.
- e) El *margen de error* máximo admisible para la estimación de los *parámetros poblacionales*.
- f) El *nivel de confianza* de la estimación muestral.

- *El tiempo y los recursos disponibles*

Sin duda, dos componentes esenciales en cualquier investigación son el *tiempo* y los *recursos* (económicos, materiales y humanos) concedidos al equipo investigador para la materialización del estudio propuesto.

En función de cuánta sea la dotación económica y los plazos temporales otorgados para cada fase de la investigación, el tamaño final de la *muestra* se verá aumentado o mermado.

- *La modalidad de muestreo seleccionada*

La elección de una (o varias) de las modalidades de *muestreo* (*probabilístico* o *no probabilístico*) se halla determinada –como ya se expuso en el Capítulo 3– por la confluencia de tres componentes claves en la práctica de cualquier investigación: los *objetivos*, los *recursos* y el *tiempo* dado para su realización. La incidencia de estos tres elementos no sólo influye en la selección de una *estrategia* concreta de investigación (y las consiguientes *técnicas de obtención* y *de análisis de datos*), sino también en la configuración del *diseño de la muestra*.

Como posteriormente se verá, las exigencias de *tamaño de la muestra* varían en consonancia con la modalidad elegida para la extracción de la *muestra*.

En general, los *diseños muestrales no probabilísticos* demandan un *tamaño muestral* inferior a los *diseños probabilísticos*. Estos últimos están más orientados al logro de *muestras* “representativas”. Lo que exige aumentar el tamaño de la muestra para reducir el *error muestral*.

En ambas agrupaciones de diseños (*probabilísticos* y *no probabilísticos*) se observan también variaciones entre las diferentes modalidades de muestreo. Así, por ejemplo, el *tamaño muestral* que se requiere para la realización de un *muestreo aleatorio sistemático* es inferior al necesario para efectuar un *muestreo aleatorio estratificado* (que garantice la presencia en la *muestra* de determinados subgrupos de población). Lo mismo acontece si se decide aplicar un *muestreo por cuotas*, con preferencia a otros *muestreos no probabilísticos* (como el *estratégico* o el de *bola de nieve*).

- *La diversidad de los análisis de datos prevista*

Cuando se proyecta cualquier investigación, antes de *diseñar la muestra* y la *estrategia* a seguir en la obtención de información, hay que anticipar la variedad de *análisis* que se estimen necesarios para la consecución de los *objetivos* de la investigación.

Si el equipo investigador cree de interés la aplicación de alguna (o varias) *técnicas estadísticas multivariantes*, deberá procurar que la *muestra* analizada incluya un número elevado de casos. Ello contribuye a la reducción de los errores de predicción y al aumento, consiguiente, del poder predictivo del modelo obtenido.

Para la realización de *análisis multivariantes* se precisa una cierta proporcionalidad entre el número de observaciones (*tamaño de la muestra*) y el número de variables incluidas en el estudio. En *regresión múltiple*, por ejemplo, la proporción es, al menos, de 20 veces más casos que *variables independientes* (si se pretende que el modelo resultante sea *estadísticamente significativo*). Esta proporción se duplica (40 veces más casos que *variables independientes*), cuando se realiza un análisis de regresión “paso a paso” (Tabachnick y Fidell, 1989). En caso contrario, habrá que eliminar *variables independientes* o proceder a su agrupación. Para *análisis bivariados* (de cruces de variables) también se precisa un *tamaño muestral* elevado. Especialmente, si las *variables* comprenden un número amplio de *categorías*.

- *La varianza o heterogeneidad poblacional*

Si la *muestra* intenta ser una representación a pequeña escala de una *población* (o *universo*), su tamaño se verá afectado por la *varianza poblacional*. Cuanto más heterogénea sea la *población*, mayor será su *varianza poblacional*. En consecuencia, necesitará un mayor *tamaño muestral* para que la variedad de sus componentes se halle representada en la *muestra*.

Las exigencias de *tamaño muestral* se reducen, en cambio, en universos homogéneos, caracterizados por una baja *varianza poblacional*.

El conocimiento de la homogeneidad o heterogeneidad poblacional resulta, por tanto, primordial en la decisión del *tamaño de la muestra*. Rodríguez Osuna (1991) sintetiza en dos las formas esenciales de acceder a dicho conocimiento en la práctica de la investigación empírica:

- a) La *experiencia* adquirida en estudios que se repiten con periodicidad (como los *Eurobarómetros* realizados mensualmente por la Comunidad Económica Europea (CEE); o la *Encuesta de Población Activa* del INE). Estos estudios proporcionan los valores de las *varianzas* para distintas variables. Lo que es de ayuda para el diseño muestral de futuras indagaciones empíricas.

Sin embargo, hay que matizar que ello es posible siempre y cuando ambas *poblaciones* (la del estudio previo y la de la investigación que se proyecta) coincidan.

- b) La realización de *estudios pilotos* previos a la investigación principal, que ayuden al cálculo de las *varianzas* de las variables de interés.

Cuando se desconoce el valor de la *varianza poblacional* (situación muy habitual en la práctica de la investigación social), se recurre al supuesto más desfavorable: se toma el producto de las probabilidades “P” (de aparición de un suceso) y “Q” (que indica la no ocurrencia del suceso o evento; siendo su valor igual a “ $1 - P$ ”) como equivalente a la *varianza poblacional*; ambas probabilidades presentarían el valor de 0,50.

Este es el supuesto más desfavorable porque conlleva un *tamaño muestral* superior (con el consiguiente incremento de los costes de la investigación) al correspondiente cuando se conocen los valores reales de dichas proporciones en la población, y éstos se alejan del valor 0,50.

A modo de ejemplo, véase el Cuadro 5.1. En él se resumen algunos *tamaños muestrales* correspondientes a *poblaciones infinitas* (mayores de 100.000 unidades), al *nivel de confianza* más empleado (2 sigma = 95,5% de probabilidad de acertar en la estimación del *parámetro* poblacional), para distintos valores de “P” y “Q” (expresados en porcentajes), y *límites de error* establecidos.

CUADRO 5.1. Tamaño muestral para poblaciones infinitas a un nivel de confianza del 95,5% (2 sigma).

Límites de error (%) para $\pm 2\sigma$	Valores presupuestos de P y Q (%)					
	1/99	10/90	20/80	30/70	40/60	50/50
0,1	39.600	360.000	640.000	840.000	960.000	1.000.000
0,5	1.584	14.400	25.600	33.600	38.400	40.000
1,0	396	3.600	6.400	8.400	9.600	10.000
1,5	176	1.600	2.844	3.733	4.267	4.444
2,0	99	900	1.600	2.100	2.400	2.500
2,5	63	576	1.024	1.344	1.536	1.600
3,0	44	400	711	933	1.067	1.111
3,5	32	294	522	686	784	816
4,0	25	225	400	525	600	625
5,0	16	144	256	336	384	400

En el cuadro puede observarse que el tamaño muestral aumenta, conforme los valores de “P” y “Q” se aproximan al valor 0,50 (indistintamente de cuál sea el *margen de error* fijado). Así, por *ejemplo*, para un *margen de error* de $\pm 2\%$, el *tamaño muestral* para una *población infinita* será de 900 unidades, si $P = 10$ y $Q = 90$. En cambio, el *tamaño de la muestra* será más del doble (2.500 unidades), cuando $P = 50$ y $Q = 50$.

Este incremento considerable en el *tamaño de la muestra* reduce, no obstante, las posibilidades de equivocación en la estimación de los *parámetros poblacionales*

correspondientes a las variables del estudio. Razón por la cual se adopta este supuesto, siempre que se desconozca el valor de la *varianza poblacional* respecto a las variables que se analizan.

Los *tamaños muestrales* que figuran en el Cuadro 5.1 se han obtenido a partir de la fórmula común para el cálculo del *tamaño muestral* para *universos infinitos*, a un *nivel de confianza* de 2 sigma (que supone un 95,5% de probabilidad de acertar en la estimación):

$$N = \frac{4\hat{P}\hat{Q}}{E^2}$$

donde “E” representa el *error muestral*.

Las fórmulas del *tamaño muestral* se comentan al final de este apartado.

- *El margen de error máximo admisible*

En el Cuadro 5.1 también puede observarse la correspondencia entre las variables *error* y *tamaño muestral*: los incrementos en el *tamaño de la muestra* repercuten en una mayor precisión en la *estimación* de los *parámetros poblacionales*, con la consiguiente reducción del *error muestral*. En cambio, en *muestras* pequeñas, el *error de muestreo* aumenta, manteniendo constante la *varianza poblacional*.

La Figura 5.1 representa, de forma gráfica, la relación negativa existente entre las variables *error* y *tamaño muestral*. Para su elaboración se han tomado los valores que aparecen en el Cuadro 5.1 correspondientes a la situación habitual de desconocimiento del valor de la *varianza poblacional* ($P = Q = 0,50$).

Como puede observarse en la Figura 5.1, a medida que aumenta el volumen del *tamaño de la muestra*, se produce un decrecimiento en el valor del *error muestral*. Si el *tamaño de la muestra* se amplía, por *ejemplo*, de 816 a 2.500 unidades, el *error muestral* disminuye en un 1,5% (pasando de representar un 3,5% a sólo un 2%).

También se advierte que a partir del 2% de *error*, se disparan los crecimientos en el *tamaño de la muestra* para alcanzar una mínima ganancia en la reducción del *error muestral*. Pasar del 2% al 1% de *error* supone un aumento de 2.500 a 10.000 unidades, con el incremento correspondiente en los costes (económicos y temporales) de la investigación; a cambio, sólo se logra una leve mejora en la precisión de la *estimación del parámetro poblacional* (un 1%). Por lo que, el investigador deberá encontrar un punto intermedio entre el *tamaño* y el *error muestral*, ante la tendencia observada en *muestras* grandes a proporcionar mínimos incrementos en adecuación en la *estimación de parámetros*.

El *error muestral* interviene en la decisión sobre el *tamaño de la muestra* siempre que el *diseño muestral* sea *probabilístico*. En este caso, el investigador fija el *error a priori*, sopesando la precisión que desea para sus *estimaciones*, con los costes que supondría la reducción del *error muestral*. Los *errores* comprendidos entre el 2,5% y el 2% son los más frecuentes en la investigación social.

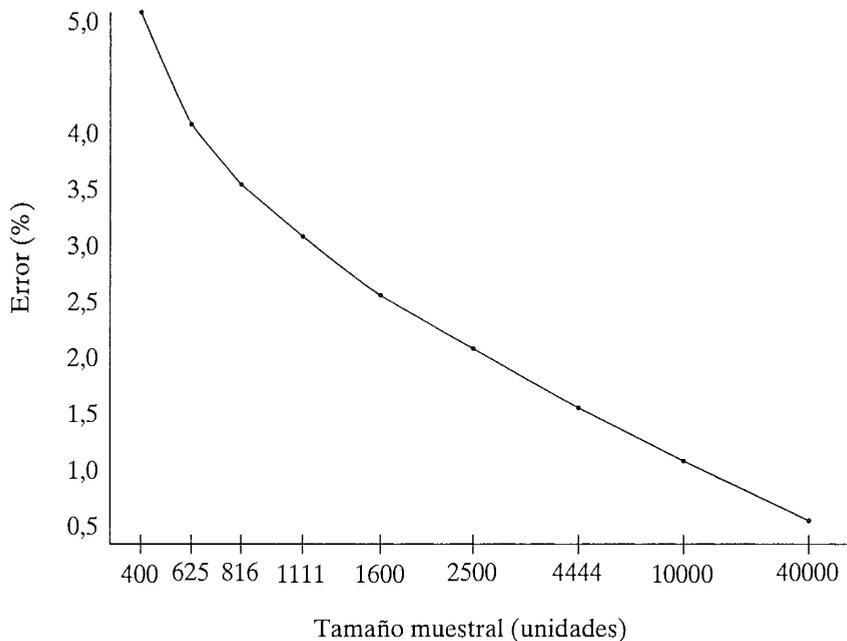


Figura 5.1. Representación gráfica de la relación entre las variables error y tamaño muestral para $P = Q = .50$ y nivel de confianza del 95,5%.

- *El nivel de confianza de la estimación*

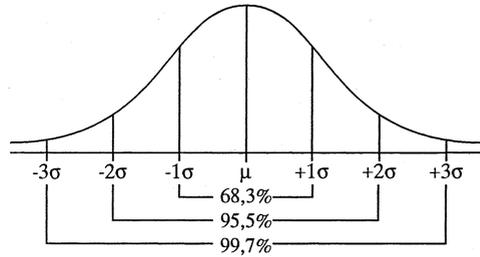
Un último elemento a considerar en la decisión sobre el *tamaño de la muestra* –aunque también sólo en los *diseños probabilísticos*– es el *nivel de confianza* que el investigador concede a sus *estimaciones*.

El *nivel de confianza* expresa –como su nombre indica– el grado de “confianza” (o de “probabilidad”) que el investigador tiene en que su estimación se ajuste a la realidad.

Tres son los *niveles de confianza* comunes en la investigación social. Corresponden a áreas bajo la curva normal acotadas por distintos valores de *desviación típica* (denominada *sigma* (σ) en referencia a la *desviación poblacional*). De ellos el más habitual es 2σ , que supone un 95,5% de probabilidad de acertar en la estimación a partir de los datos muestrales.

La *distribución normal* se aplica en la *estadística inferencial* para la estimación de la probabilidad de que un determinado evento acontezca. Representa una curva perfectamente simétrica, en forma de campana, y que admite infinitos valores (unidades “Z”: *unidades de desviación típica*). El área total bajo la *curva normal* es 1 (dado que la *probabilidad* siempre es un valor comprendido entre 0 y 1). En función de cuál sea el valor de “Z”, variará la *probabilidad* concedida al evento en cuestión (véase la tabla del área bajo la curva normal en el Anexo).

Valores de la distribución normal		Probabilidad comprendida en el intervalo
Entre	$\pm 1\sigma$	0,683
	$\pm 2\sigma$	0,955
	$\pm 3\sigma$	0,997



Si se toma una unidad de *desviación típica*, la *probabilidad* de acertar en la estimación sería del 68,3%. Si en vez de una se toman dos unidades de *desviación*, la *probabilidad* de acierto se incrementa hasta el 95,5% (habiendo sólo un 4,5% de *probabilidad* de equivocación en la *estimación del parámetro poblacional*).

La plena seguridad en la estimación (exactamente el 99,7%) se adquiere, si se fijan en tres las unidades de *desviación típica* (3 sigma). Pero, no es éste el *nivel de confianza* usual en la práctica investigadora, sino el 95,5% (equivalente a 2). Ante ello, la pregunta inmediata es ¿por qué?

Para encontrar la respuesta, véase el Cuadro 5.2. En él se ejemplifica cómo (para una misma *varianza poblacional*) varía el *tamaño muestral*, dependiendo de cuál sea el *nivel de confianza* adoptado. Incrementos mínimos en *nivel de confianza* (del 95,5% de probabilidad de acierto al 99,7%) supone añadir más del doble de unidades a la *muestra*.

Si para un *error* del 2% y *varianza* $P = Q = 50$, el *tamaño muestral* sería de 2.500 unidades (para un *nivel de confianza* de 95,5%), al aumentar el nivel de confianza al

CUADRO 5.2. Tamaño muestral para poblaciones infinitas a un nivel de confianza del 95,5% y del 99,7%.

Límites de error (%)	Nivel de confianza (%)	Valores presupuestos P Y Q (%)				
		10/90	20/80	30/70	40/60	50/50
1,0	95,5	3.600	6.400	8.400	9.600	10.000
	99,7	8.100	14.400	18.900	21.600	22.500
2,0	95,5	900	1.600	2.100	2.400	2.500
	99,7	2.025	3.600	4.725	5.400	5.627
2,5	95,5	576	1.024	1.344	1.536	1.600
	99,7	1.296	2.304	3.024	3.456	3.600
3,0	95,5	400	711	933	1.067	1.111
	99,7	900	1.600	2.100	2.400	2.500
4,0	95,5	225	400	525	600	625
	99,7	506	900	1.181	1.350	1.406

99,7%, la muestra pasaría a estar integrada por 5.627 elementos. La ganancia en precisión es ínfima, mientras que los costes de la investigación (en tiempo y dinero) supondrían más del doble de los presupuestados para un nivel de confianza del 95,5%.

Todos los aspectos referidos participan en el cálculo del *tamaño de una muestra probabilística*. La fórmula genérica para una *muestra aleatoria (simple o sistemática)* sería la siguiente, cuando el universo o población estuviere compuesto por más de 100.000 unidades:

$$n = \frac{Z^2 \hat{S}^2}{E^2} \quad \text{o} \quad n = \frac{Z^2 \hat{P}^2 (1 - \hat{P})}{E^2}$$

donde: “Z” representa las *unidades de desviación típica* correspondientes al *nivel de confianza* elegido (2σ o 3σ , fundamentalmente).

“ \hat{S}^2 ” constituye el valor de la *varianza poblacional*. Este equivale al producto de las proporciones \hat{P} y \hat{Q} siendo $\hat{Q} = 1 - \hat{P}$.

“E” denota el *error* máximo permitido que el investigador establece *a priori*.

Si el universo estuviere integrado por 100.000 unidades o menos, se trataría de una *población finita*. En este caso, habría que introducir un factor de corrección, quedando las fórmulas transformadas de la siguiente manera:

$$n = \frac{Z^2 \hat{S}^2 N}{E^2 (N - 1) + Z^2 \hat{S}^2} \quad \text{o} \quad n = \frac{Z^2 \hat{P} \hat{Q} N}{E^2 (N - 1) + Z^2 \hat{P} \hat{Q}}$$

donde: “N” es el tamaño de la población.

En *muestras aleatorias estratificadas y por conglomerados* se introducen otras variaciones que se comentarán cuando se describan ambas modalidades de muestreo (apartado 5.4).

EJEMPLOS DE CÁLCULO DEL TAMAÑO MUESTRAL

Se desea conocer el tamaño de la muestra necesario para una encuesta a la población mayor de 18 años, con objeto de medir el voto en las próximas elecciones municipales. El error máximo permitido será $\pm 2,5\%$.

- a) Si la encuesta se efectúa en Madrid capital, sin información previa sobre el porcentaje de voto. Para un nivel de confianza de 95,5% (2σ), el tamaño muestral necesario sería:

$$n = \frac{Z^2 \hat{P}\hat{Q}}{E^2} = \frac{2^2 \cdot 50 \cdot 50}{2,5^2} = 1.600 \text{ unidades}$$

En cambio, para un nivel de confianza del 99,7% (3σ), el tamaño muestral sería más del doble:

$$n = \frac{Z^2 \hat{P}\hat{Q}}{E^2} = \frac{3^2 \cdot 50 \cdot 50}{2,5^2} = 3.600 \text{ unidades}$$

- b) Si se parte de la predicción (a partir de sondeos realizados con anterioridad) de que el PSOE puede alcanzar el 24% del total de votos:

$$n = \frac{Z^2 \hat{P}\hat{Q}}{E^2} = \frac{2^2 \cdot 24 \cdot 76}{2,5^2} = 1.167 \text{ unidades para } 2\sigma$$

$$n = \frac{Z^2 \hat{P}\hat{Q}}{E^2} = \frac{3^2 \cdot 24 \cdot 76}{2,5^2} = 2.627 \text{ unidades para } 3\sigma$$

- c) Si la encuesta se llevase a cabo en el municipio de Torrelodones, con una población de hecho (censada en 1991) de 7.113 individuos.

$$n = \frac{Z^2 \hat{P}\hat{Q}N}{E^2(N-1) + Z^2 \hat{P}\hat{Q}} = \frac{4 \cdot 50 \cdot 50 \cdot 7.113}{(2,5^2 \cdot 7.112) + (4 \cdot 50 \cdot 50)} = 1.306 \text{ para } 2\sigma$$

$$n = \frac{Z^2 \hat{P}\hat{Q}N}{E^2(N-1) + Z^2 \hat{P}\hat{Q}} = \frac{9 \cdot 50 \cdot 50 \cdot 7.113}{(2,5^2 \cdot 7.112) + (9 \cdot 50 \cdot 50)} = 2.390 \text{ para } 3\sigma$$

5.3. El error muestral

Cuando se diseña una *muestra*, el objetivo primordial es conseguir un elevado nivel de adecuación en la selección de la *muestra*, respecto de la *población* a la que pertenece. Ello contribuye a que la investigación adquiera *validez externa* (como ya se vio en el Capítulo 3).

Pero, por muy perfecta que sea la *muestra*, como únicamente se analiza una parte de la *población* (y esa *muestra* sólo representa una de todas las posibles *muestras* que pueden extraerse de una misma población), siempre habrá alguna divergencia entre los valores obtenidos de la *muestra* (*estimaciones*) y los valores correspondientes en la *población* (*parámetros*). Esa disparidad se denomina *error muestral*: el grado de inadecuación existente entre las *estimaciones muestrales* y los *parámetros poblacionales*.

Como todos los resultados muestrales se hallan sujetos a *errores de muestreo*, cuando se analizan datos muestrales, debería comprobarse la *representatividad* de la *mues-*

tra e identificar (si los hubiese) los *sesgos* existentes. Ello proporciona un mayor rigor a las conclusiones que se deduzcan de la investigación efectuada.

Si la extracción de la *muestra* se ha llevado a cabo de forma aleatoria (*muestreo probabilístico*), podrá calcularse el *error muestral*, con la ayuda de la *teoría de la probabilidad*; en caso contrario (*muestreo no probabilístico*), no. Esta constituye una de las ventajas importantes de los *diseños muestrales probabilísticos*: la capacidad de determinar la precisión de las *estimaciones muestrales*.

Para el cálculo del *error muestral* (en *muestras probabilísticas*, exclusivamente), se acude al estadístico llamado "*error típico*". En la *teoría de la probabilidad*, este estadístico mide la extensión a la que las *estimaciones muestrales* se distribuyen alrededor del *parámetro poblacional*. Concretamente, se especifica que aproximadamente el 68% de las *estimaciones muestrales* se hallarán comprendidas entre ± 1 vez el *error típico del parámetro poblacional*; el 95,5%, entre ± 2 veces el *error típico*; y, finalmente, el 99,7%, entre ± 3 veces el *error típico*.

En suma, el *nivel de confianza* en la *estimación* aumenta conforme se amplía el *margen de error*. El *nivel de confianza* más utilizado es –como ya se indicó en el apartado anterior– 2σ (sigma). Este nivel de confianza expresa que hay un 95,5% de probabilidad de que la *estimación muestral* se halle comprendida en el intervalo definido por dos veces el *error típico del parámetro poblacional*.

En el cálculo del *error típico* intervienen los elementos siguientes:

- a) El *tamaño muestral*.
- b) La *varianza poblacional*.
- c) El *nivel de confianza* adoptado.
- d) El *tipo de muestreo* realizado.

• *El tamaño muestral*

No es la proporción de la población que constituye la muestra (la *fracción de muestreo* = n/N) lo que determina el *error muestral* (salvo si la *fracción* resulta elevada), sino el *tamaño de la muestra*. A medida que aumenta el *tamaño de la muestra*, decrece el *error muestral* (*cualquiera que sea el valor de la varianza poblacional*).

Dicha afirmación puede comprobarse en el Cuadro 5.3. En él figuran los *errores típicos* correspondientes a determinados *tamaños muestrales*. Estos *errores típicos* disminuyen conforme se incrementa el *tamaño muestral*, indistintamente de los valores presupuestos de P y Q. Incluso, véase que para reducir el *error típico* a la mitad, habría que cuadruplicar el *tamaño de la muestra*.

Por ejemplo, una *muestra* compuesta de 500 unidades en la situación más desfavorable ($P = Q = 50$) presenta un *error típico* de $\pm 4,5\%$ (a un n. c. del 95,5%). Para reducir dicho *error* a $\pm 2,2\%$, habría que multiplicar el *tamaño muestral* por 4, obteniéndose una muestra de 2.000 unidades. Si el *tamaño muestral* se amplía de 1.000 a 4.000 unidades, el *error* también se reduce a la mitad (de $\pm 3,2\%$ a $\pm 1,6\%$).

CUADRO 5.3. Errores típicos correspondientes a distintos tamaños muestrales a un nivel de confianza del 95,5% (2σ).

Número de elementos muestra	Valores presupuestos de P y Q (%)					
	1/99	10/90	20/80	30/70	40/60	50/50
50	2,8	8,5	11,4	13,0	13,9	14,2
100	2,0	6,0	8,0	9,2	9,8	10,0
200	1,4	4,3	5,7	6,5	7,0	7,1
500	0,9	2,7	3,6	4,1	4,4	4,5
1.000	0,6	1,9	2,6	2,9	3,1	3,2
1.500	0,5	1,6	2,1	2,4	2,5	2,6
2.000	0,4	1,3	1,8	2,0	2,2	2,2
3.000	0,4	1,1	1,5	1,7	1,8	1,8
4.000	0,3	1,0	1,3	1,4	1,5	1,6
5.000	0,3	0,9	1,1	1,3	1,4	1,4
10.000	0,2	0,6	1,8	1,0	1,0	1,0
50.000	0,1	0,2	0,3	0,4	0,4	0,4

- *La varianza poblacional*

El nivel de *heterogeneidad* de una *población* (respecto a una variable considerada), favorece el *error muestral*, excepto si se aumenta el *tamaño muestral* para incluir todas las distintas variedades que componen el *universo*.

El *error muestral* se halla más presente en *poblaciones* heterogéneas (de elevada *varianza poblacional*) que en *universos homogéneos*.

- *El nivel de confianza adoptado*

Como aumentar el *nivel de confianza* supone ampliar el *tamaño de la muestra*, obviamente un mayor *nivel de confianza* llevará consigo la reducción del *error muestral*. Pese a ello, hay que insistir en que, cuando se diseña una *muestra*, se valoran conjuntamente distintos aspectos, buscando el equilibrio entre precisión y coste.

Incrementos en el *tamaño de la muestra* conllevan una ampliación del *nivel de confianza* en la *estimación muestral* (descendiendo el *error*). Pero esta mejora en precisión repercute en unos costes superiores de la investigación. De ahí que el *nivel de confianza* más empleado sea el ya referido de 2σ (95,5% de *probabilidad* de acierto en la *estimación*).

- *El tipo de muestreo realizado*

El *error muestral* también se halla afectado por el procedimiento de selección de las unidades muestrales. En general, el *muestreo aleatorio estratificado* es (de los *di-*

seños muestrales probabilísticos) el que genera un menor *error muestral*. En cambio, el *muestreo aleatorio por conglomerados* es (de los *probabilísticos*) el que ocasiona un mayor *error muestral*. Aunque la agrupación de la *muestra en conglomerados* presenta la gran ventaja de reducir los costes (en tiempo y dinero) del trabajo de campo, éste a su vez repercute en una desventaja importante: incrementa el *error típico de la muestra*, con la consiguiente disminución en precisión de los estimaciones de la población.

A estas afirmaciones genéricas cabe añadir –de acuerdo con Fowler (1988)– las precisiones siguientes:

- a) El *muestreo aleatorio sistemático* ocasiona *errores de muestreo* equivalentes a las *muestras aleatorias simples*, siempre y cuando no exista ninguna estratificación en la muestra.
- b) Las *muestras aleatorias estratificadas* presentan menores *errores de muestreo* que las *muestras aleatorias simples* del mismo tamaño; pero, sólo si existe heterogeneidad entre los estratos.
- c) El *muestreo aleatorio por conglomerados* genera un mayor *error típico*, sobre todo cuanto mayor es el tamaño de los conglomerados en la última fase de muestreo.

Dadas estas variaciones en el *error muestral* debidas al diseño realizado, los cálculos del *error típico* han de adecuarse a las fórmulas apropiadas a cada *diseño muestral*.

Para una *muestra aleatoria simple* o *sistemática*, las fórmulas correspondientes al error típico (E) serían las siguientes:

	Universo infinito	Universo finito (≤100.000 unidades)
Error típico de la media	$E_{(\bar{x})} = \sqrt{\frac{\hat{S}^2}{n}}$	$E_{(\bar{x})} = \sqrt{\frac{\hat{S}^2}{n} \cdot \frac{N-n}{N-1}}$
Error típico de una proporción	$E_{(p)} \sqrt{\frac{\hat{P}\hat{Q}}{n}}$	$E_{(p)} \sqrt{\frac{\hat{P}\hat{Q}}{n} \cdot \frac{N-n}{N-1}}$

Las proporciones “P” y “Q” pueden expresarse tanto en porcentajes como en tantos por uno, obteniéndose los mismos resultados.

En todos los casos, los resultados obtenidos se multiplicarían por el *nivel de confianza* adoptado. Si es 95,5%, por 2σ ; y si fuera 99,7%, por 3σ .

A partir de los valores obtenidos, se calculan los *intervalos de confianza* para el *nivel de probabilidad* fijado. Ello posibilita conocer cuánto se aproxima la estimación muestral al *parámetro poblacional*.

EJEMPLOS DE ERROR MUESTRAL

- a) Un equipo de investigadores desea conocer los hábitos de consumo de la población juvenil española. Para ello entrevistan (mediante *cuestionario*) a 2.000 jóvenes distribuidos por toda España. Del estudio se concluye (entre otros aspectos) que la media de gastos mensuales es de 30.000 pts al mes, con una desviación típica de 5.600. Calcular el error de la estimación muestral. A partir de él, inferir cuál será el parámetro poblacional correspondiente a un nivel de probabilidad de 2σ .

$$E_{(\bar{x})} = \sqrt{\frac{\hat{S}^2}{n}} = \sqrt{\frac{5.600^2}{2.000}} = 125,21$$

Si el *error* se multiplica por 2σ , se realiza una estimación del *intervalo de confianza* que comprende el *parámetro poblacional*, sumando y restando a la estimación muestral el producto del *error* por el *nivel de confianza* elegido: $30.000 \pm (2)(125,21)$.

En consecuencia, hay un 95,5% de probabilidad de que la media de gastos mensuales de la población juvenil española se halle comprendida entre 29.750 y 30.250 pesetas al mes.

Si el *nivel de confianza* elegido hubiera sido 3σ , el *intervalo de confianza* sería ligeramente más amplio; por lo que sería menor la probabilidad de equivocación en la estimación. Los *límites de confianza* serían, respectivamente: $30.000 - (3)(125,2)$ y $30.000 + (3)(125,2)$. Es decir, habría un 99,7% de probabilidad de que la media de gastos mensuales de la población juvenil española se halle comprendida entre 29.624 y 30.376 pesetas al mes.

- b) De los 2.950 menores internados en un centro de reforma se ha extraído una muestra de 780 menores de distintas edades. De ellos, el 49% fueron acusados de hurto. ¿Cuál será la proporción de detenidos por hurto en la población total recluida en dicho centro a un nivel de confianza de 2σ ?

$$E_{(p)} = \sqrt{\frac{\hat{P}\hat{Q}}{n} \frac{N-n}{N-1}} = \sqrt{\frac{(49)(51)}{780} \frac{2950-780}{2950-1}} = 1,54\%$$

Los *límites de confianza* serían, respectivamente: $49 - (2)(1,5)$ y $49 + (2)(1,5)$. Por tanto habría un 95,5% de probabilidad de que el porcentaje de menores detenidos por hurto (en todo el centro) estuviese comprendido entre el 46 y el 52% de la población total.

Si la muestra fuese *aleatoria estratificada proporcional*, se introducirían las modificaciones siguientes en las fórmulas del *error típico*:

Error típico de la media:
$$E_{(\bar{x}_{prop})} = \sqrt{\frac{\sum_{i=1}^n n_i \hat{S}_i^2}{n^2}}$$

Error típico de una proporción:

$$E_{(Pprop)} = \sqrt{\frac{\sum_{i=1}^n n_i \hat{P}_i \hat{Q}_i}{n^2}}$$

donde " \hat{P}_i " es la proporción de la muestra en el estrato i que posee el atributo en cuestión.

" $\hat{Q}_i = 1 - \hat{P}_i$ ": la proporción que no posee dicho atributo.

" $\hat{S}_i^2 = \hat{P}_i * \hat{Q}_i$ ": estimación de la *varianza* de la variable de interés para la población en el estrato i .

" Σ ": sumatorio de todos los estratos, desde 1 hasta n .

" n_i ": tamaño de la muestra total.

EJEMPLO DE ESTIMACIÓN DEL ERROR MUESTRAL EN UNA MUESTRA ALEATORIA ESTRATIFICADA

Se quiere conocer el grado de aceptación de los nuevos planes de estudio por la población universitaria española. A tal fin se escoge una *muestra* integrada por 2.439 alumnos de la Universidad Complutense. La *muestra* se *afija* (distribuye) entre las distintas facultades de manera *proporcional* al número de alumnos registrados en dichas facultades. Para ello se multiplica la proporción que representa cada uno de los *estratos* (facultades diferenciadas) –columna (2)– por el *tamaño muestral* (2.439), obteniéndose las cifras que aparecen en la columna (3). En la columna (4) figuran los porcentajes de alumnos que en cada una de las facultades afirman estar satisfechos con los nuevos planes de estudio.

Facultades	(1) Población	(2) Pobl. prop.	(3) n_i	(4) P_i	(5) $n_i P_i (1 - P_i)$	(6) $r_i = n_i P_i$
Medicina	4.535	0,046	112	0,49	27,99	55
Farmacia	4.983	0,050	122	0,35	27,76	43
Veterinaria	2.682	0,027	66	0,49	16,49	32
Derecho	18.600	0,188	459	0,71	94,51	326
Económicas/Empresa.	13.755	0,139	339	0,83	47,83	281
Políticas/Sociología	6.888	0,070	171	0,42	41,66	72
Psicología	4.567	0,046	113	0,43	27,70	49
Filol./Filosofía/Pedag.	10.613	0,107	261	0,39	62,09	102
Geografía/Historia	6.063	0,061	149	0,51	37,24	76
CC. Información	13.900	0,141	344	0,59	83,21	203
Química/Física	5.568	0,056	137	0,69	29,30	95
CC. Exactas	2.411	0,024	59	0,47	14,70	28
Biología/Geología	4.388	0,044	107	0,44	26,36	47
	98.953		2.439		536,84	1.409

Para estimar el *error típico* de esta *muestra aleatoria estratificada proporcional*, se procede a multiplicar el número de alumnos que componen la muestra en cada estrato (columna 3) por la proporción de alumnos a favor de los nuevos planes de estudio (columna 4) y por la proporción de alumnos en contra ($\hat{Q} = 1 - \hat{P}$). Como resultado se obtienen las cifras que aparecen en la columna (5). Después, se suman dichos productos y se sustituyen en la fórmula correspondiente:

$$E_{(pprop)} = \sqrt{\frac{\sum_{i=1}^n n_i \hat{P}_i \hat{Q}_i}{n^2}} = \sqrt{\frac{536,84}{2.439^2}} = 0,0095$$

Para demostrar que el *muestreo estratificado* genera un menor *error típico* que el *muestreo aleatorio simple*, se calcula el *error típico* que se tendría si la *muestra* hubiera sido *aleatoria simple*. Con este propósito se estima la proporción de los alumnos encuestados que afirman estar a favor de los nuevos planes de estudio. En la columna (6) figura el número de alumnos que contestaron estar a favor en cada uno de los estratos considerados. Si se suman dichas cantidades (r_i) y se divide por el total de sujetos analizados (2.439), se obtiene la proporción global de alumnos a favor: el 57,8% de los encuestados.

$$P = \frac{\sum r_i}{n} = \frac{1.409}{2.439} = 0,578$$

Por tanto

$$E_{(P)} \sqrt{\frac{\hat{P}\hat{Q}}{n} \frac{N-n}{N-1}} = \sqrt{\frac{(0,58)(0,42)}{2.439} \frac{(98.953 - 2.439)}{(98.953 - 1)}} = 0,0099$$

En este caso, la diferencia entre ambos tipos de errores parece mínima. Pero, si se eleva el *error típico* al cuadrado (*varianza*) obtenido del *muestreo aleatorio estratificado*, y se divide por el cuadrado del *error típico* que se obtendría mediante un *muestreo aleatorio simple*, se comprueba el *efecto del diseño* (*deff*), enunciado por Kish (1965):

$$deff = \frac{(0,0095)^2}{(0,0099)^2} = 0,921$$

De esta forma se concluye que la *varianza* de la *muestra estratificada proporcionalmente* es un 8% más pequeña que la *varianza* de la *muestra aleatoria simple*.

Por tanto, si mediante un *muestreo aleatorio simple*, se quisiera alcanzar la misma precisión en la estimación de los parámetros poblacionales, habría que ampliar el *tamaño muestral* a 2.648 (esta cifra se obtiene dividiendo 2.439/0,921). Con ello se demuestra la mayor precisión alcanzada mediante un *muestreo aleatorio estratificado*. La diferencia entre ambos tipos de errores crece a medida que aumentan las variaciones observadas en los estratos respecto a la variable considerada.

Por último, si la *muestra* fuese por *conglomerados*, la fórmula correspondiente al *error típico* sería la siguiente:

$$E_{(\bar{x}_{congl})} = \sqrt{\left(1 - \frac{m}{M}\right) \frac{Sb^2}{m}}$$

donde: “M”: número de *conglomerados* en la *población*.

“m”: número de *conglomerados* seleccionados en la *muestra*.

“Sb²”: la *varianza* de los valores del *conglomerado* x_j. Se obtiene:

$$Sb^2 = \frac{1}{m-1} \sum_{i=1}^n (\bar{x}_i - \bar{x})^2$$

EJEMPLO DE ESTIMACIÓN DEL ERROR EN UNA MUESTRA ALEATORIA POR CONGLOMERADO

Adaptemos el ejemplo anterior a una *muestra por conglomerados*. Con este propósito se divide la población universitaria por curso académico y grupo. Imaginemos que fuesen 780 los grupos existentes en total. De ellos se decide escoger aleatoriamente 20 grupos (*conglomerados*) de 122 alumnos cada uno. De esta forma se obtendría una *muestra* total de 2.440 alumnos.

El *error típico* para esta *muestra por conglomerados* sería:

$$E_{(\bar{x}_{congl})} = \sqrt{\left(1 - \frac{m}{M}\right) \frac{Sb^2}{m}}$$

donde: “M”: número de *conglomerados* en la *población* (780).

“m”: número de *conglomerados* seleccionados en la *muestra* (20).

“Sb²”: la *varianza* de los valores del *conglomerado* a favor de los nuevos planes de estudio (0,099).

Entonces

$$E_{(\bar{x}_{congl})} = \sqrt{\left(1 - \frac{20}{780}\right) \frac{0,099}{20}} = 0,0698$$

Por tanto, se habría obtenido un *error típico* superior al alcanzado mediante la *estratificación* (0,0095) y el *muestreo aleatorio simple* (0,0099), también.

En general, las estimaciones de los *errores típicos* para *muestras aleatorias por conglomerados* (especialmente si el *diseño muestral* es *polietápico*) resultan bastante complejas; si bien, en la actualidad, algunos paquetes informáticos –como el SAS (*Statistical Analysis System*)– calculan los *errores típicos* para cualquier variedad de *diseño muestral*.

5.4. Tipos de muestreo: diseños muestrales probabilísticos y no probabilísticos

La *representatividad* de una *muestra* no sólo depende de la magnitud de su *tamaño*, sino también del procedimiento seguido para la selección de las unidades muestrales. Las modalidades de muestreo son variadas, aunque cabe agruparlas en dos amplias categorías (muestreo *probabilístico* y *no probabilístico*), dependiendo de si el azar interviene en todas las fases de la selección.

- *Muestreo probabilístico o aleatorio*

Se fundamenta en la *aleatorización* como criterio esencial de selección muestral. Ello favorece que:

- a) Cada unidad de la población tenga una *probabilidad* igual (y conocida *a priori*) de participar en la muestra.
- b) La elección de cada unidad muestral sea independiente de las demás.
- c) El cálculo de la adecuación de la estimación muestral (*error muestral*) a los parámetros poblacionales pueda hacerse dentro de unos márgenes de probabilidad específicos.

- *Muestreo no probabilístico*

A diferencia del anterior, la extracción de la *muestra* se efectúa siguiendo criterios diferentes de la *aleatorización* (como la conveniencia u otros criterios subjetivos). Ello da cabida a cualquier discreción por parte del equipo investigador. Además, repercute en:

- a) La desigual *probabilidad* de las unidades de la población para formar parte de la muestra.
- b) La dificultad de calcular el *error muestral*.
- c) La introducción de *sesgos* en el proceso de elección muestral. Esto redundará en riesgos superiores de invalidez de los hallazgos de la investigación.

No obstante, el *muestreo no probabilístico* presenta dos *ventajas* notorias que le hacen atractivo en la práctica investigadora:

- a) No precisa de la existencia de un *marco de muestreo*.
- b) Su materialización resulta más sencilla y económica que los *muestreos probabilísticos*.

La elección entre un tipo de muestreo u otro vendrá condicionada por la confluencia de cuatro aspectos básicos:

- a) La *dotación económica* de la investigación.
- b) El período de *tiempo* programado para su ejecución.
- c) La existencia de un *marco muestral* válido, que haga factible la *selección muestral aleatoria*.
- d) El *grado de precisión* que el investigador quiera dar a la indagación.

El *muestreo probabilístico* se adecúa más a *propósitos* de:

- a) Estimación de *parámetros*.
- b) Comprobación de *hipótesis (test de significatividad)*.

En cambio, el *muestreo no probabilístico* se muestra más apropiado para:

- a) La indagación *exploratoria (estudios piloto)*.
- b) Estudios *cualitativos*, más interesados en profundizar en la información aportada que en su representatividad estadística.
- c) Investigaciones sobre población “marginal” (como prostitutas, delincuentes, drogadictos, homosexuales, inmigrantes ilegales, e incluso miembros menos “marginales” como parejas cohabitantes, por ejemplo), de difícil registro y localización. Lo que complica la aplicación de *diseños muestrales probabilísticos*.

El Cuadro 5.4 resume la variedad de *procedimientos muestrales* clasificados en “probabilísticos” y “no probabilísticos”. Si bien, hay que precisar que éstos no forman compartimentos estancos. Un mismo *diseño muestral* puede compaginar diferentes variedades de muestreo en fases distintas de su desarrollo.

CUADRO 5.4. Modalidades de muestreo.

<i>Probabilísticos</i>	<i>No probabilísticos</i>
Simple Sistemático Estratificado Por conglomerados De rutas aleatorias	Por cuotas Estratégico “Bola de nieve”

5.4.1. Muestreo aleatorio simple

Constituye el prototipo de muestreo, en referencia al cual se estiman las fórmulas básicas para el cálculo del *tamaño* y del *error muestral*.

Como *muestreo probabilístico*, su realización exige la existencia de un *marco muestral* que cumpla las condiciones expuestas en el apartado 5.1. Una vez localizado, se asigna a cada unidad de la población (en el listado) un *número de identificación* (si éste no figura anexo al *marco muestral*, siguiendo un orden consecutivo) para, posteriormente, proceder a la extracción aleatoria de los integrantes de la *muestra*.

La selección muestral puede hacerse siguiendo una *tabla de números aleatorios*, mediante un programa de ordenador, u otro procedimiento que garantice que:

- a) Cada unidad de la población tenga una probabilidad igual de participar en la muestra. Esta probabilidad viene determinada por la *fracción de muestreo* = n/N (el cociente entre el *tamaño de la muestra* y el total de la *población*).
- b) La selección muestral sea totalmente *aleatoria* hasta alcanzar el *tamaño muestral* fijado. No obstante, se aconseja extraer un número mayor de unidades que las prefijadas en el *tamaño muestral* para las sustituciones (por imposibilidad de localizar a algunos de los integrantes originales de la *muestra*, o porque estos se nieguen a colaborar en la investigación, entre otras razones).

La elección de las unidades muestrales puede hacerse *sin reemplazamiento* (cuando cada unidad de la población puede ser elegida sólo una vez) o *con reemplazamiento* (si participa en elecciones sucesivas). Lo habitual es que el proceso de selección se efectúe *sin reemplazamiento*.

Cuando el *marco muestral* se halla en soporte magnético (como el último censo de población –de 1991–, por ejemplo), se facilita la selección muestral mediante ordenador. En este caso, el programa de ordenador configurado ejecuta todas las tareas correspondientes:

- a) Numera a cada uno de los integrantes del *marco muestral* (o listado de las unidades de la población).
- b) Genera su propia serie de *números aleatorios*.
- c) Selecciona e imprime la *lista* de las unidades muestrales finalmente elegidas.

Si el uso del ordenador no resulta viable, se recurre al *procedimiento tradicional*: utilizar una *tabla de números aleatorios*. Estas *tablas* comprenden múltiples combinaciones de números extraídos al azar (o por sorteo), de forma que cada dígito presente una probabilidad igual de figurar en cualquier punto de la tabla. La actuación, en este caso, sería:

- a) Elegir aleatoriamente un *punto de partida*: una columna o una fila cualquiera de la *tabla*. Puede comenzarse por cualquier lugar y moverse en cualquier dirección (hacia arriba, hacia abajo, a la derecha o a la izquierda). La única condición que se exige es que la elección no responda a un conocimiento previo de la distribución de los números en la *tabla*. Ello introduciría *sesgos* en la *muestra*.
- b) El número de dígitos extraídos de la *tabla* ha de corresponder con el número de dígitos de la población contenida en el *marco muestral*.
 Por *ejemplo*, si el *marco* estuviese compuesto por 10.000 unidades (5 dígitos), habría que numerar a cada unidad de la población con 5 dígitos: desde el primero (00001) hasta el último (10.000), correlativamente. A continuación, se escogerían números con 5 dígitos, desconsiderándose aquellas cifras que superen la última unidad poblacional numerada en el listado (10.000).
- c) El individuo al que pertenezca el número extraído pasará a formar parte de la *muestra*, salvo que en el *marco* no se adjunte un medio para su localización (dirección postal, teléfono).

EJEMPLO DE MUESTREO ALEATORIO SIMPLE

Supóngase que va a realizarse una encuesta a licenciados en Ciencias Políticas y Sociología. El equipo investigador elige tomar como *marco de muestreo* el directorio de “colegiados” (en el Ilustre Colegio Nacional de Doctores y Licenciados en Ciencias Políticas y Sociología de Madrid). Imaginemos que, a fecha de la investigación, hubieran 4.500 personas colegiadas. La *muestra* (si se fija un margen de error de $\pm 3\%$ y un nivel de confianza de 2σ) estaría compuesta por 891 colegiados.

Si se tomase el número de colegiado como *número de identificación*, no habría que asignar a cada unidad de la población un número.

Para la extracción de los números se comenzaría por una fila o columna de una *tabla de números aleatorios*. En el Cuadro 5.5 figura un extracto de la *tabla de números aleatorios* adjunta en el Anexo. Del cuadro se deduce que la decisión tomada ha sido comenzar en la primera columna, escogiendo los cuatro últimos dígitos, y continuar hacia abajo hasta el final de la columna. Cuando ésta concluye, seguir en la siguiente columna, pero ahora tomando los cuatro primeros dígitos e, igualmente, hacia abajo. El procedimiento continúa hasta que, al menos, se alcanza el *tamaño de la muestra* fijado. Se decide extraer 60 colegiados más para posibles sustituciones de la *muestra* inicial.

Descartando los número superiores a la cifra de 4.500 ($n.$ º total de colegiados), los sujetos seleccionados serían aquellos cuyo *número de identificación* ($n.$ º de colegiado) se correspondiese con las cifras siguientes: 4357, 2111, 0801, 1767, 3740, 0924, 0364, 1247, 2147, 3364, 4460, 0621, 2815, 3666, 4410, 0054, 0935, 3602.

La *probabilidad de selección* dada a cada unidad de la población sería 0,198, al ser la *fracción de muestreo*:

$$\frac{n}{N} = \frac{891}{4.500} = 0,198$$

CUADRO 5.A. Extracto de una tabla de números aleatorios.

9	6754	1767	6	5	5659	4410	5	4	7361	3483	3
3	4357	8604	0	5	3364	7172	6	4	5690	6633	4
0	6318	3740	3	4	9927	5771	5	5	0423	6737	2
6	2111	5282	0	0	7243	7993	1	8	9292	8476	7
4	7534	0924	3	6	7879	0054	4	2	3410	1274	0
9	8614	7599	3	8	4460	6284	6	5	9844	1492	2
2	4856	0364	8	4	4898	0935	1	9	8795	1864	4
9	6887	1247	9	8	0621	6622	3	8	6085	7828	5
9	0801	2147	2	4	2815	7740	8	3	7390	7676	6
5	5165	7731	2	8	3666	3602	8	2	8420	7021	9

La exigencia de disponer de un listado de la población de interés, y su numeración, hace desaconsejable este procedimiento de selección muestral en poblaciones de elevada magnitud. En estos casos, habría que acudir a otro tipo de *diseño muestral* que suponga un menor coste de realización.

El Cuadro 5.5 resume las ventajas y los inconvenientes principales del *muestreo aleatorio simple*.

CUADRO 5.5. Ventajas e inconvenientes del muestreo aleatorio simple.

<i>Ventajas</i>	<i>Inconvenientes</i>
Facilidad de los cálculos estadísticos.	Requiere listar y enumerar a las unidades de la población.
Elevada probabilidad de lograr la equiparabilidad entre las características de la muestra y las correspondientes a la población.	Resulta monótono y arduo en muestras y poblaciones elevadas (sobre todo cuando se emplean procedimientos manuales).
	La dispersión alcanzada en la muestra repercute negativamente en los costes de la investigación.

5.4.2. Muestreo aleatorio sistemático

Esta variedad de muestreo probabilístico es más practicada que la anterior. Exige igualmente la existencia de un listado de la población, pero difiere del *muestreo aleatorio simple* en dos aspectos fundamentales:

- a) Sólo la primera unidad de la *muestra* se elige al azar (mediante una *tabla de números aleatorios*, por sorteo u otro procedimiento), con la condición de que el número seleccionado sea inferior al *coeficiente de elevación*.

El *coeficiente de elevación* se calcula dividiendo el *tamaño del universo* por el *tamaño de la muestra* (N/n). Expresa el número de veces que la *muestra* se halla contenida en el *universo*.

- b) Los restantes elementos de la *muestra* se obtienen sumando, sucesivamente, el *coeficiente de elevación* (a partir del primer número elegido aleatoriamente), hasta completar, al menos, el *tamaño muestral*.

Si, previamente, no se ha extraído un excedente de unidades muestrales a considerar para las “sustituciones” (de los elementos de la *muestra*), en el momento de la *selección muestral* (o posteriormente), ha de calcularse un nuevo *coeficiente de elevación* que permita una nueva *selección sistemática* de las unidades muestrales no cubiertas en el *trabajo de campo* (por no estar localizables, por negación a cooperar u otra razón).

EJEMPLO DE MUESTREO ALEATORIO SISTEMÁTICO

Si la muestra del ejemplo anterior se eligiese de forma sistemática, se actuaría de la manera siguiente:

- a) Se calcula el *intervalo de selección*, determinado por el *coeficiente de elevación*

$$\frac{N}{n} = \frac{4.500}{891} = 5,1$$

- b) Se extrae al azar un número inferior a 5; por ejemplo, el 3. La unidad de la población a la que corresponda dicho número (en el marco muestral utilizado), constituye la primera unidad de la *muestra*.
- c) Las demás unidades muestrales se eligen sumando, sistemáticamente, el *coeficiente de elevación* a los números que se seleccionen a partir del primer número extraído al azar. En este ejemplo, el 3 sería el primer número. Le seguirían el 8, el 13, el 18, 23, 28, 33..., hasta reunir, como mínimo, las 891 unidades que integran la *muestra* del estudio.

La *muestra* así extraída se asemejaría a la lograda mediante un *muestreo aleatorio simple*

En el Cuadro 5.6 figuran algunas de las ventajas y de los inconvenientes del *muestreo aleatorio sistemático*.

CUADRO 5.6. Ventajas e inconvenientes del muestreo aleatorio sistemático.

<i>Ventajas</i>	<i>Inconvenientes</i>
<p>No precisa del uso continuo de una tabla de números aleatorios (u otro procedimiento de asignación aleatoria).</p> <p>No exige que el marco muestral sea un listado. Éste, en cambio, puede adoptar varias formas (fichas, papeletas,... hasta la propia presencia física de las unidades de la población).</p>	<p>Necesita del recuento constante de las unidades de la población.</p> <p>Antes de la selección muestral, hay que desordenar el marco muestral (si éste se encuentra ordenado de acuerdo con algún criterio que favorezca la mayor representación de determinados segmentos de la población en la muestra).</p>

5.4.3. Muestreo aleatorio estratificado

El *muestreo aleatorio estratificado* constituye una de las modalidades de muestreo más practicadas en la investigación social, cuando se dispone de información sobre características de la población de interés. Supone la clasificación de las unidades de población (contenidas en el *marco de muestreo* escogido), en un número reducido de grupos (*estratos*), en razón de su similitud, dictada por las características observadas en el *marco muestral*. Con ello se persigue que cada *estrato* tenga representación en la *muestra* final.

Al igual que en el *muestreo por cuotas*, en el *estratificado* la *muestra* se distribuye en diferentes grupos de población, en función de los valores que presente en las variables elegidas para la *estratificación*. No obstante, ambas modalidades de muestreo difieren en la forma de elección de las unidades muestrales. Mientras que en el *muestreo por cuotas* intervienen criterios subjetivos, en el *estratificado* sólo el *azar*. En él, la extracción de las unidades finales de la *muestra* (en cada *estrato*) se hace siguiendo exclusivamente procedimientos aleatorios de selección muestral (*muestreo simple, sistemático, de rutas aleatorias*).

Lynn y Lievesley (1991) destacan cuatro principios básicos en el diseño de un esquema de *estratificación*:

- a) *Elección de las variables de estratificación*. En general, la elección de las variables está condicionada a aquellas comprendidas en el *marco muestral* de referencia. De ellas se escogerán las más relacionadas con los objetivos de la investigación, con la finalidad de aumentar la efectividad de la *estratificación*.

De las variables también se exige que sean *categorías*. En caso contrario, habría que cambiar su nivel de medición; es decir, categorizarlas.

- b) *Orden de las variables de estratificación*. Como la primera variable de estratificación es la que más discrimina, en el primer estadio de *estratificación* debería

elegirse la variable de mayor relevancia para la investigación; en el segundo estadio, la segunda variable más importante; y, así, consecutivamente, debido al poder decreciente de la *estratificación*.

- c) *Número de variables de estratificación*. Para la mayoría de los objetivos de investigación, no se precisa diferenciar tres o cuatro estadios de *estratificación*. Tampoco se requiere la adopción de las mismas variables de estratificación para todas las muestras. Puede alcanzarse una mayor eficacia siguiendo un esquema de *estratificación* distinto para las variables incluidas en los diversos estadios de la *estratificación*.
- d) *Tamaño de los estratos*. La *estratificación* resulta más adecuada, cuando se divide la población en grupos de igual tamaño. En cambio, se alcanzan escasos beneficios, cuando se utilizan variables bastante correlacionadas entre sí. En este caso, los estratos diferenciados variarían en su tamaño, aunque no en su naturaleza, en contra del propósito primordial de la *estratificación*: la consecución de grupos de población muy homogéneos entre sí y diferentes del resto.

En suma, si con la *estratificación* se persigue el logro de una mayor precisión en la estimación muestral, ésta se alcanzará cuando se cumplan dos condiciones esenciales:

- a) Sean máximas las diferencias entre los estratos y mínimas dentro de cada estrato.
- b) Las variables de estratificación se hallen relacionadas con los objetivos de la investigación (con las variables independientes y/o dependientes). De no ser así, se obtendría una precisión similar a la alcanzada sin la *estratificación*.

EJEMPLO DE ESTRATIFICACIÓN

En un estudio sobre la incidencia del fracaso escolar en la conducta delictiva, se elegirían como *variables de estratificación* cualquiera de las habitualmente asociadas a la causalidad de ambos fenómenos sociales: el fracaso escolar y la delincuencia.

Es fácil que el investigador disponga de información sobre las variables sexo y clase social. En caso afirmativo, podría estratificarse en función de ambas variables. Atendiendo a estas variables, la población quedaría, por ejemplo, dividida en seis *estratos*: varones de clase alta, varones de clase media, varones de clase baja, mujeres de clase alta, mujeres de clase media, mujeres de clase baja.

De esta forma, el investigador asegura la representación en la *muestra* de los estratos diferenciados. Lo que será de utilidad para la posterior realización de comparaciones entre los datos obtenidos de cada uno de los grupos señalados.

Las *variables de estratificación* más empleadas son las variables sexo y edad, debido a que ambas se hallan recogidas en la mayoría de los *marcos muestrales*. A ellas pueden añadirse otras variables, como la clase social, la ocupación, el nivel de instrucción; depende del objetivo fundamental de la investigación.

En estudios a nivel nacional (e internacional), suele estratificarse por ubicación geográfica: ámbito territorial (país, comunidad autónoma, provincia, municipio), tipo de hábitat (urbano, semiurbano, rural), o por tamaño de hábitat (número de habitantes).

Tras la clasificación de la *población en estratos*, se procede a *afijar la muestra* en cada *estrato*. Por *afijación* se entiende la distribución del *tamaño muestral* global entre los *estratos* diferenciados. Esta distribución puede cumplirse de tres maneras distintas: de forma *simple*, *proporcional* u *óptima*.

- *Afijación simple*

Se asigna el mismo *tamaño de la muestra* a cada *estrato*. Con ello se busca la igual representación de los estratos en la muestra global. Lo que favorecerá el logro de estimaciones igualmente significativas en todos los estratos.

Esta equidistribución del *tamaño muestral* conlleva, no obstante, un *inconveniente* importante: favorece a los estratos de menor volumen de población (al concedérseles un tamaño muestral superior al correspondiente a su volumen poblacional), respecto de los estratos de mayor magnitud.

Para la *afijación simple* se divide el *tamaño muestral* global entre el número de *estratos*. La cantidad así obtenida expresará las unidades a observar en cada uno de los estratos.

- *Afijación proporcional*

Como su nombre indica, en la *afijación proporcional* la distribución de la *muestra* se hace “proporcional” al peso relativo del *estrato* en el conjunto de la población. Por lo que, a los estratos que reúnan un mayor número de unidades de población les corresponderá un tamaño muestral superior al de aquellos que representen un porcentaje inferior en la población.

Este criterio de afijación es el más seguido. Para su cálculo, se multiplica la *proporción* que representa el *estrato* en la población por el *tamaño muestral* a afijar.

- *Afijación óptima*

Al peso relativo del *estrato* en la población, en la *afijación óptima* se añade la *variabilidad* (o *heterogeneidad*) del *estrato* respecto a la variable considerada en la *estratificación*. En conformidad con este último criterio de *afijación*, les corresponderá un tamaño muestral superior a los *estratos* de mayor heterogeneidad y peso poblacional.

La realización de esta variedad de *afijación* precisa del conocimiento previo de la *varianza poblacional* en cada *estrato*, respecto de las variables que se hayan empleado para la *estratificación*. Esta información es difícil de conocer (los *marcos muestrales* usuales no la facilitan), a menos que se haya analizado con anterioridad a la misma población (en un *estudio piloto ex profeso* o en investigaciones periódicas). Esto explica la escasa práctica de este tipo de *afijación*, pese a su calificativo de “óptima”.

Para su cálculo se multiplica, en cada *estrato*, el *porcentaje* que representa el *estrato* en el conjunto de la población por la *varianza* (o *desviación típica*) correspondiente. Luego se suman todos los productos y se calcula la proporción que representa cada producto en el total. Por último, se multiplican estas proporciones por el tamaño de la muestra. De esta forma se obtiene el número de unidades de la población a observar en cada *estrato*.

EJEMPLO DE AFIJACIÓN DEL TAMAÑO MUESTRAL EN ESTRATOS

En una investigación sobre las actitudes delictivas de la población universitaria española, se decide la *estratificación* de la población universitaria por nivel académico, con la finalidad de garantizar la presencia en la *muestra* de los distintos niveles académicos. La *muestra* global está integrada por 2.500 unidades (con un *error* máximo de $\pm 2\%$ y *nivel de confianza* de 2σ). Esta *muestra* se *afija* en los estratos siguiendo alguno de los criterios siguientes:

Nivel de estudios universitarios	Porcentaje población	Varianza	Afijación		
			Simple	Proporcional	Óptima
Primer ciclo	45	1.900	833	1.125	970
Segundo ciclo	39	2.600	833	975	1.150
Tercer ciclo	16	2.100	833	400	380
			2.499	2.500	2.500

Afijación simple:
$$\frac{2500}{3} = 833$$

De cada estrato se escoge el mismo número de unidades muestrales

Afijación proporcional:

$$0,45 \cdot 2.500 = 1.125$$

$$0,39 \cdot 2.500 = 975$$

$$0,16 \cdot 2.500 = 400$$

Al primer estrato le corresponderá un mayor tamaño muestral porque reúne un volumen de población más amplio. En cambio, al tercer estrato sólo se le asignan 400 unidades muestrales.

Afijación óptima:

(A) Porcentaje	(B) Varianza	(C) (A) · (B)	(D) Proporción (C)	(E) Afijación óptima*
45	1.900	85.500	0,388	970
39	2.600	101.400	0,460	1.150
16	2.100	33.600	0,152	380
100		220.500	1.000	2.500

(*) La columna (E) se obtiene multiplicando cada cifra que figura en la columna (D) por el tamaño muestral a fijar (2.500 unidades).

Como el segundo estrato presenta la *varianza poblacional* más elevada y, al mismo tiempo, es el segundo con mayor representación poblacional, éste resulta ser el más favorecido en la *afijación óptima*.

Por último, comparéense las cifras obtenidas en las tres modalidades de *afijación*. Así podrán apreciarse las variaciones en la distribución de la *muestra* que producen los tres criterios de *afijación* existentes.

Las tres variedades de *afijación* pueden englobarse en dos amplias modalidades de *estratificación*: la *estratificación proporcional* y la *no proporcional*.

- *Estratificación proporcional*

La *estratificación* se hace de manera que garantice una probabilidad igual de selección para todos los estratos. Como difiere el conjunto de unidades que forman el estrato, los tamaños muestrales correspondientes a cada estrato también variarán en proporción similar.

- *Estratificación no proporcional*

La representación de los estratos en la *muestra* final no es proporcional a su peso en el conjunto de la población, al haberse dado una probabilidad desigual de selección en cada estrato (mediante la *afijación simple* o la *óptima*).

Esta “no proporcionalidad” puede deberse al deseo de analizar, con mayor detalle, unos estratos concretos, a los cuales les correspondería un tamaño muestral inferior, si se hubiese optado por la *estratificación proporcional*; o, simplemente, para propiciar la representatividad de las estimaciones muestrales en todos los estratos.

Uno de los *inconvenientes* fundamentales de la *estratificación no proporcional* es la necesidad de *ponderar la muestra*, si se desea la obtención de estimaciones muestrales para el conjunto de la población. Por el contrario, no se precisa de la *ponderación*, cuando sólo se realizan análisis individuales y/o comparativos entre los estratos.

Por *ponderación* se entiende el proceso de asignación de “pesos” a cada *estrato*, de manera que logre compensarse la desigual probabilidad de selección dada a cada unidad de población que compone el *estrato*. A tal fin, se comparan los datos muestrales con características de la población de interés publicadas en el último *censo de población*, *padrón de habitantes* u otro sondeo a cuyos datos se conceda una significativa *validez*. Por lo que, antes de *ponderar*, el investigador deberá comprobar la adecuación de los datos que toma como referente de las características poblacionales.

La *ponderación* puede efectuarse de varias formas. La más usual consiste en dividir el porcentaje que representa el estrato en el *universo* entre el porcentaje que representa en la *muestra*. De esta manera se obtienen los “pesos” o *coeficientes de ponderación*. Éstos se multiplicarán por cada una de las *estimaciones muestrales* de los estratos correspondientes para, así, poderse inferir el valor de los respectivos *parámetros poblacionales*.

En los ficheros de datos de la mayoría de los paquetes informático se fijan los *pesos* a asignar a cada subconjunto (*estrato*) de población. El programa automáticamente cumplimenta la *ponderación*, antes de proceder a la *tabulación* conjunta y posterior análisis de la información obtenida.

EJEMPLO DE PONDERACIÓN

Una institución alavesa encarga una encuesta para conocer los hábitos de consumo de droga entre la población (de ambos sexos) de 16 y más años de edad, por franjas de edad y zonas geográficas. El *tamaño de la muestra* se ha fijado en 1.600 entrevistas. Lo que supone un *error genérico de muestreo* del $\pm 2,5\%$, para un *nivel de confianza* del 95,5% (2σ) y un supuesto de *varianza máxima* ($P = Q = 0,50$).

Dada la distribución de la población alavesa en el territorio (en la que destaca la concentración en la capital de tres de cada cuatro habitantes), el equipo investigador decide realizar una *afijación inicial no proporcional* de la *muestra* total. Se asignan 900 entrevistas a la capital Vitoria (con 172.546 habitantes de 16 y más años de edad) y las 700 restantes distribuidas en el conjunto de los otros municipios (con una población global de 56.224 habitantes de 16 y más años).

El *objetivo* es obtener resultados con *errores de muestreo* similares en estas dos grandes zonas geográficas. Para devolver a los datos su proporcionalidad (y poder así analizar

conjuntamente la información recabada en cada estrato), se procede a la *ponderación de la muestra*:

<i>Cuadrillas que dividen administrativamente Álava</i>	<i>Población 16 y más años (*)</i>	<i>(A) Porcent. poblac.</i>	<i>Afijación propor. (**)</i>	<i>Afijación no propor. (***)</i>	<i>(B) Porcentaje muestra</i>	<i>Coefficiente ponderación (A/B)</i>
Añana	4.868	2,1	34	61	3,8	.553
Ayala	28.263	12,4	198	352	22,0	.564
Campezo	2.880	1,3	21	36	2,3	.565
La Guardia	8.066	3,5	56	100	6,2	.565
Salvatierra	7.434	3,2	51	92	5,7	.561
Vitoria	172.546	75,4	1.206	900	56,3	1.339
Zuya	4.713	2,1	34	59	3,7	.568
	228.770	100	1.600	1.600	100	

(*) Fuente: Censo de Población 1991, INE.

(**) Número de entrevistas que se habrían realizado, si se hubiese efectuado una *afijación proporcional*. Estas se obtienen multiplicando la proporción que representa cada *cuadrilla* (zona) en el conjunto de la población por el tamaño muestral (1.600).

(***) Número de entrevistas que en realidad se han efectuado: 900 en Vitoria y las 700 restantes distribuidas proporcionalmente entre el resto de las cuadrillas. Para ello se calcula nuevamente la proporción que representa cada cuadrilla, pero en el conjunto de las seis cuadrillas (eliminando a Vitoria), con una población total de 56.224 habitantes. Las proporciones así obtenidas se multiplican por el número de entrevistas a afijar en las seis cuadrillas (700). Por ejemplo, la primera cuadrilla representa ahora el 8,7% de la población total (exceptuando Vitoria). Multiplicando esa proporción por 700, se tendrían las 61 entrevistas a realizar; 27 más que si se hubiese hecho una *estratificación proporcional*.

En suma, si se hubiese *afijado proporcionalmente*, el *tamaño muestral* correspondiente a las cuadrillas (a excepción de Vitoria) habría sido bastante inferior al conseguido mediante la *estratificación no proporcional*. Mediante la *ponderación* (o equilibrio de la *muestra*) se devuelve a los datos su proporcionalidad, antes de proceder a la tabulación conjunta. Para ello se indican los *coeficientes de ponderación* (o pesos) en el fichero de datos y, automáticamente, se multiplica cada valor numérico obtenido por su correspondiente *coeficiente de ponderación*. De esta forma se obtienen las *estimaciones muestrales*, a partir del instrumento de medición aplicado en la investigación.

A modo de resumen, en el Cuadro 5.7 se indican algunas de las ventajas e inconvenientes que lleva consigo la aplicación de un *muestreo aleatorio estratificado*.

CUADRO 5.7. Ventajas e inconvenientes del muestreo aleatorio estratificado.

<i>Ventajas</i>	<i>Inconvenientes</i>
<p>Supone un menor error muestral y, por tanto, una mayor precisión de la estimación muestral.</p> <p>Asegura la representación de las variables de estratificación (y de las variables relacionadas con ellas).</p> <p>Pueden emplearse procedimientos muestrales variados en los distintos estratos.</p> <p>Facilita la organización del trabajo de campo.</p>	<p>Precisa más información del marco muestral que el muestreo aleatorio simple (para identificar a la población de cada estrato). Ello puede resultar costoso.</p> <p>Lleva consigo cálculos estadísticos complejos.</p>

5.4.4. Muestreo aleatorio por conglomerados

El *muestreo por conglomerados* también representa un procedimiento de selección aleatoria de un conjunto de individuos (referidos ahora como *conglomerados*).

Con el *muestreo aleatorio estratificado* comparte la característica básica de seccionar la población total en grupos, como fase previa a la extracción muestral. Si bien difiere de él en varios aspectos importantes:

- a) En el *muestreo estratificado* se busca la homogeneidad dentro del estrato y la heterogeneidad entre los estratos. En el *muestreo por conglomerados* es a la inversa: el *error muestral* disminuye conforme aumenta la heterogeneidad dentro del grupo (*conglomerado*). Ello se debe a la necesidad de que cada *conglomerado* constituya una representación, lo más ajustada posible, de la variedad de componentes del universo.
- b) En el *muestreo estratificado* se selecciona aleatoriamente una *muestra* para cada *estrato*. En el *muestreo por conglomerados* lo que se extrae es una muestra aleatoria de *conglomerados*. Sus integrantes formarán la *muestra*.
- c) En el *muestreo estratificado*, la unidad de muestreo es el individuo. En cambio, en el *muestreo por conglomerados* es el *conglomerado* (o conjunto de individuos).

Los *conglomerados* pueden ser las áreas geográficas que dividen a la población que se analiza (país, comunidad autónoma, municipio, distrito, áreas censales, vivien-

das); pero, también, organizaciones u instituciones (colegios, hospitales, tribunales, centros penitenciarios).

Sudman (1976) señala los siguientes aspectos a considerar en la elección de los *conglomerados*:

- a) Los *conglomerados* han de estar bien definidos y delimitados. Cada unidad de la población sólo puede pertenecer a un único *conglomerado*.
- b) El número de elementos que componen el *conglomerado* ha de ser conocido previamente (aunque sea de manera aproximada).
- c) Los *conglomerados* elegidos han de ser pocos, si realmente quieren reducirse los costes de la investigación.
- d) Los *conglomerados* deberían escogerse de manera que se consiguiera disminuir el aumento en el *error muestral*, generado por la agrupación (o *aconglomeración*).
- e) Los *conglomerados* no tienen por qué hallarse idénticamente definidos en todos los lugares.

Cuando se muestrean individuos u hogares en áreas urbanas, los *conglomerados* suelen ser bloques o conjuntos de bloques. En cambio, en las áreas rurales, los *conglomerados* serán segmentos geográficos limitados por carreteras y fronteras naturales (como ríos y lagos).

Si, a partir de una *muestra por conglomerados*, se extrae una nueva muestra, con referencia a cada uno de los *conglomerados* previamente elegidos, y así, sucesivamente, se está ante un *diseño muestral* muy habitual en la investigación social: el *muestreo polietápico por conglomerados*.

El *muestreo polietápico (o polifásico) por conglomerados* representa una extensión del *muestreo por conglomerados*. En él la unidad de muestreo final no son los *conglomerados*, sino subdivisiones de estos. Por lo que no se toman cada uno de los integrantes de los *conglomerados* elegidos aleatoriamente, sino sólo a una parte de ellos, escogidos también de forma aleatoria. Ello supone muestrear –como su nombre indica– en distintos niveles, implicando varios procedimientos de selección muestral (como la *estratificación*).

La modalidad de *muestreo polietápico por conglomerados* más sencilla implica la extracción muestral en dos fases (Henry, 1990):

- a) En la *primera fase*, se seleccionan las agrupaciones de los miembros de la población de estudio (conocidas como las *unidades de muestreo primarias*), que son análogas a los *conglomerados*.
- b) En la *segunda fase*, se eligen aleatoriamente los miembros de la población a observar, de las *unidades de muestreo primarias* previamente seleccionadas.

Pero, lo usual es que el número de estadios comprendidos en el *muestreo polietápico* sean tres o cuatro (Lynn y Lievesley, 1991). En el caso más general (en muestras

nacionales), se recurre a *muestreos polietápicos, estratificados, por conglomerados*: primero, se *afija* la *muestra* por *estratos*; después, se extraen (de forma aleatoria proporcional) los municipios, las secciones estadísticas; y, por último, los hogares en los que se realizarán las entrevistas (Osuna, 1989).

EJEMPLO DE MUESTREO ALEATORIO POR CONGLOMERADOS

Se quiere encuestar (mediante entrevista personal) a 1.300 profesores universitarios españoles, con el propósito de conocer su opinión sobre los nuevos planes de estudio.

Como la encuesta es de ámbito nacional, la selección de la *muestra* siguiendo el procedimiento *aleatorio simple* elevaría considerablemente los costes de la investigación. A la dificultad de encontrar un *marco muestral* que incluyera a todos los profesores universitarios españoles (tanto de universidades públicas como privadas), se sumaría el aumento de los costes por desplazamiento de los entrevistadores a puntos dispersos del país.

Lo más fácil sería acceder a un listado que incluyese al conjunto de las universidades españolas (públicas y privadas) para, a continuación proceder a la extracción *muestral polietápica por conglomerados*. Ésta podría consistir en:

- a) Listar todas las universidades españolas, tanto públicas como privadas. Este listado constituiría el *marco de muestreo* inicial de la investigación.

En la selección de la muestra nacional de universidades podría *estratificarse* por *tipo de universidad* (pública y privada; y dentro de esta última, diferenciándose por *confesionales* y *no confesionales*), y por *ubicación geográfica* (Comunidad Autónoma). Ello aseguraría la representación de las *unidades de muestreo primarias* (las universidades) en conformidad con las *variables de estratificación* elegidas: dispersión geográfica y tipo de universidad.

También conviene –para alcanzar una mayor *precisión*– que las *unidades primarias de muestreo* elegidas sean bastante heterogéneas respecto a las características que se analizan. De esta forma se asegura que la variedad del universo de estudio quede reflejada en la *muestra*.

Asimismo, se recomienda que –como regla– no proceda más del 7% de la *muestra* total de una sola *unidad de muestreo primaria* (Lynn y Lievesley, 1991). Lo que interesa es que haya un número elevado de *conglomerados* en la primera fase del *diseño muestral*. Esto favorece la heterogeneidad de la *muestra*, con la consiguiente reducción del *error muestral*.

- b) Elegir (de las universidades seleccionadas) una muestra de facultades mediante alguno de los procedimientos de selección aleatoria (*simple, sistemática, estratificada*).
- c) Para cada una de las facultades escogidas, extraer una muestra aleatoria de profesores. A tal fin se acude a un nuevo *marco muestral*: el listado de los profesores pertenecientes a las facultades elegidas. Éstos podrían, a su vez, muestrearse conforme a los criterios de categoría profesional (catedrático, titular, profesor asociado, ayudante), dedicación (a tiempo completo o parcial), y antigüedad docente, siguiendo un *muestreo aleatorio estratificado o por cuotas*.

Si se opta por el *estratificado*, la elección de los profesores cuyas características se ajusten a los *estratos* diferenciados será totalmente aleatoria (*muestreo simple, sistemático*). En cambio, si se decide el de *cuotas*, la elección de los profesores será arbitraria, con la única condición de que compartan las características fijadas en las *cuotas*.

El *diseño muestral* que resulte de la conjunción de estas fases se corresponderá con la precisión que el investigador quiera dar a las *estimaciones muestrales*. Esto tiene una repercusión directa en los costes de la investigación.

Así, por ejemplo, en la elección del número de facultades y de profesores a entrevistar en cada una de ellas, las opciones posibles son diversas. Dependerá del presupuesto destinado a la realización del *trabajo de campo* y de la heterogeneidad de los *conglomerados*, principalmente. El investigador tendrá que decidir si aumentar el número de facultades, disminuyendo el número de profesores a entrevistar en cada una de ellas, o proceder a la inversa (reducir el número de facultades, incrementando el número de profesores a encuestar). La primera opción supone aumentar –para un mismo *tamaño muestral* (aproximadamente, 1.300 unidades)– los costes (en tiempo y dinero) de la investigación, respecto de la segunda opción que resulte más económica.

A continuación se exponen algunas de las opciones posibles:

<i>Nº facultades</i>	<i>Nº profesores</i>
5	260
10	130
15	87
20	65
30	43
40	33

Si fuesen 5 las facultades elegidas (entrevistándose a 260 profesores, en cada una de ellas), los costes del trabajo de campo serían inferiores; pero también sería mayor el *error muestral* que si se eligiesen 40 facultades, y de ellas se entrevistase a 33 profesores.

En general, se recomienda aumentar el número de *conglomerados* (en este caso de facultades) con preferencia a elevar el número de individuos a observar en cada uno de ellos (los profesores). La razón está en la probable homogeneidad de los *conglomerados* elegidos.

Cuando éstos son bastante homogéneos, no se precisa añadir más elementos del *conglomerado* a la *muestra*, ya que se obtendría una información redundante, al ser similares las características de las unidades que forman el *conglomerado*. En este caso (cuando los *conglomerados* son homogéneos), se aconseja ampliar el número de *conglomerados* para, de esta forma, abarcar una mayor variedad de la población de interés. Por el contrario, si los *conglomerados* fuesen heterogéneos, la mejor opción (entendida como reducción del *error muestral*) sería la opuesta: reducir el número de *conglomerados*, aumentando las unidades a observar en cada uno de ellos.

La probabilidad de selección de cada uno de los integrantes de los *conglomerados* será igual al número de unidades a elegir de cada *conglomerado* dividido por el número de unidades contenidas en el *conglomerado*.

En suma, el *muestreo aleatorio por conglomerados* se muestra de especial interés cuando:

- a) Resulte difícil compilar una lista exhaustiva de todos los componentes de la población. Lo que imposibilitaría la práctica de otra variedad de *muestreo probabilístico*.
- b) Se quiera reducir la duración y los costes económicos del trabajo de campo en la investigación.
- c) Se realicen estudios de ámbito nacional o internacional, que supongan una considerable dispersión de la *muestra*.

El Cuadro 5.8 resume ventajas e inconvenientes principales del *muestreo aleatorio por conglomerados*.

CUADRO 5.8. Ventajas e inconvenientes del muestreo aleatorio por conglomerados.

<i>Ventajas</i>	<i>Inconvenientes</i>
<p>No exige un listado de toda la población de interés; sólo de las unidades del conglomerado.</p> <p>Al concentrar el trabajo de campo en un número limitado de puntos de muestreo, disminuyen los costes de la investigación.</p>	<p>Mayor error muestral (y, por tanto menor precisión de las estimaciones muestrales) porque los conglomerados suelen ser muy homogéneos*.</p> <p>Requiere cálculos estadísticos complejos en la estimación del error muestral, principalmente.</p>

* El error muestral puede reducirse aumentando el número de *conglomerados*.

5.4.5. *Muestreo por cuotas*

Una de las modalidades de muestreo más populares es el *muestreo por cuotas*. Esta constituye una variedad de *muestreo no probabilístico* que parte, igualmente, de la segmentación de la población de interés en grupos, a partir de variables sociodemográficas relacionadas con los objetivos de la investigación. Por lo que, su puesta en práctica conlleva, también, la elaboración de una *matriz* con las características básicas de la población que se analiza (proporciones de población diferenciadas por sexo y edad, nivel de instrucción, clase social). Esta información suele obtenerse del último *censo de población, padrón de habitantes* u otra fuente estadística similar.

El propósito es seleccionar una *muestra* que se ajuste a la distribución de las características fundamentales de la población. Si el objetivo del estudio fuese, por *ejemplo*, comprobar la influencia del nivel educativo de la personas en sus actividades de ocio, la población debería, al menos, *estratificarse* por nivel educativo. Ello ga-

rantizaría que en la *muestra* se encuentren representados los distintos grupos de población (diferenciados por nivel de instrucción). Lo que ayudaría a la realización de comparaciones entre ellos.

Además de los objetivos del estudio, en la elección de las variables intervienen otros factores: la precisión que el investigador desee para su indagación, junto a la accesibilidad de las variables elegidas. Ésta dependerá de la facilidad de su obtención en el *marco de muestreo* elegido pero, también, de su practicabilidad (si el entrevistador puede acceder fácilmente) a los grupos de población definidos por las variables escogidas.

Las *cuotas* más habituales son las determinadas por la conjunción de las variables sexo y edad, en consonancia con su mayor accesibilidad. La mayoría de los *marcos muestrales* contienen ambas variables. A esto se suma la relativa facilidad (para el entrevistador) de localizar visualmente a personas que pertenezcan a los distintos grupos de sexo y edad.

Lo mismo no cabe decir de otras variables, como la clase social, por ejemplo. Esta variable se enfrenta –de acuerdo con Moser y Kalton (1989)– a dos importantes trabas que dificultan su uso en el establecimiento de las *cuotas*:

- a) La carencia de fuentes estadísticas fiables para la delimitación de las *cuotas*. A ello contribuye la propia definición de la variable clase social, que supone la combinación de factores objetivos (ocupación, renta) y subjetivos (apariencia física, forma de vestir, de hablar).
- b) La vaguedad de la definición de la variable clase social, que provoca discrecionalidad en el juicio subjetivo del entrevistador, con la consiguiente introducción de sesgos importantes en la investigación.

Una vez confeccionada la matriz, se calculan –siguiendo un proceso similar al *muestreo estratificado proporcional*– las proporciones relativas (las *cuotas*) para cada celdilla de la *matriz*, a partir de la proporción que representa cada categoría de las variables seleccionadas en la población total. Con ello se busca la equiparación de la *muestra* con la *población* de interés.

Después de la delimitación de las *cuotas*, se proporciona a cada entrevistador su asignación correspondiente: el perfil y el número de personas a entrevistar en cada *cuota*.

EJEMPLO DE ASIGNACIÓN DE CUOTAS

En un *diseño muestral* en el que se seccione la población conforme a las variables sexo, edad y clase social, primero se delimitan las *cuotas*, en consonancia con el peso de las distintas categorías de las variables en el conjunto de la población. Después, se asigna a ca-

da entrevistador un número total de entrevistas a realizar, acordes con las *cuotas* fijadas. La asignación podría ser la siguiente:

		Clase social			Total
		Alta	Media	Baja	
Varones	18 – 30	1	2	1	4
	31 – 45	1	5	2	8
	46 – 65	2	6	1	9
	Más de 65	–	1	–	1
Mujeres	18 – 30	–	2	1	3
	31 – 45	1	6	2	9
	46 – 65	2	5	2	9
	Más de 65	–	1	1	2
Total		7	28	10	45

Lo que distingue al *muestreo por cuotas*, respecto del *estratificado*, es la libertad que se da al entrevistador para la elección de las unidades finales de la población a entrevistar. Aunque el azar intervenga en las fases iniciales del *diseño muestral* (en la elección de áreas o zonas geográficas, por ejemplo), la selección de los elementos concretos de la población es totalmente arbitraria. Es el propio entrevistador quien elige al entrevistado, en cualquier momento y lugar (en una calle comercial, a la salida del metro, en la parada del autobús, en un mercado, en un parque). La única condición que se le impone es que la persona se ajuste a las *cuotas* fijadas por el equipo investigador.

Este margen de libertad que se concede al entrevistador representa la principal debilidad de esta modalidad de muestreo:

- a) El entrevistador es libre de entrevistar a quien quiera o pueda. Lo que puede introducir *sesgos* importantes en el proceso de selección muestral, no siempre detectables por el equipo investigador. Por *ejemplo*, puede haber escogido a personas de su propio entorno sociofamiliar, a aquellas que haya estimado más propensas a participar en la investigación, o simplemente a las más accesibles.

También es fácil que, dentro de una misma *cuota*, se escoja a unos individuos con preferencia a otros. Por *ejemplo*, si se controla por edad, es probable que se entreviste más a personas de edades intermedias en los intervalos diferenciados, y no a aquellas situadas en los límites del intervalo.

Asimismo, el entrevistador puede ubicar a los sujetos en *cuotas* diferentes a las que realmente pertenecen, en aquellas donde se precisen casos (para completarlas). Esta última picaresca también resulta difícil de controlar.

- b) El desconocimiento de las probabilidades de selección no permite evitar los errores generados por ponderaciones incorrectas en el proceso de estimación. Tampoco permite la estimación de los *errores de muestreo*.

Todo ello repercute en el principal *inconveniente* de este tipo de muestreo: la *muestra* finalmente obtenida puede no ser representativa de la población que se analiza, aunque la *muestra* diseñada coincida con la distribución de la *población* en los controles de *cuotas* fijados. Como Azorín y Sánchez Crespo (1986: 21) afirman, “puede ocurrir que la muestra represente muy bien a la población para ciertas variables, y no la represente en absoluto en cuanto opiniones”.

Para solventar los *sesgos* inherentes en el *muestreo por cuotas*, éste suele complementarse con el *muestreo de rutas aleatorias*: para cada entrevistador se fija un *itinerario aleatorio* (definido en todos sus detalles), indicándole en qué puntos concretos ha de realizar cada entrevista. Estas indicaciones limitan la arbitrariedad del entrevistador.

En el Cuadro 5.9 se exponen ventajas e inconvenientes generalmente atribuidos al *muestreo por cuotas*.

CUADRO 5.9. Ventajas e inconvenientes del muestreo por cuotas.

<i>Ventajas</i>	<i>Inconvenientes</i>
<p>Resulta más económico (en tiempo y dinero) que los muestreos probabilísticos.</p> <p>Fácil de administrar.</p> <p>No precisa de un listado de la población.</p>	<p>Supone un mayor error muestral que los diseños probabilísticos.</p> <p>Inexistencia de algún método válido para calcular el error típico (al no ser un muestreo probabilístico).</p> <p>Limites en la representatividad de la muestra para las características no especificadas en los controles de cuotas.</p> <p>Dificultad para controlar el trabajo de campo.</p>

5.4.6. Muestreo de rutas aleatorias

El *muestreo de rutas aleatorias* también suele estar presente en el estadio final de un *diseño muestral* (en la selección de las unidades últimas de muestreo), comple-

mentando tanto a *muestreos no probabilísticos* (como el de *cuotas*) como a *probabilísticos* (el *muestreo de conglomerados* y/o el *estratificado*).

Se denomina *muestreo de rutas* porque se establece la “ruta” o *itinerario* que el entrevistador ha de seguir en la selección de las unidades muestrales. Las *rutas* se eligen de forma “aleatoria”, sobre un mapa del municipio concreto donde han de realizarse las entrevistas. En el mapa se indican múltiples puntos de partida o de inicio de rutas posibles. Normalmente, el *punto de partida* corresponde a un edificio determinado por una dirección (por ejemplo, la calle Serrano n.º 22). Pero, en entidades de población pequeñas (núcleos rurales, por ejemplo), el *punto de inicio de la ruta* suele corresponder a edificios centrales, como el ayuntamiento, la iglesia o la plaza del pueblo.

Una vez que se ha elegido aleatoriamente el comienzo de la *ruta*, el entrevistador deberá tomar una dirección u otra, siguiendo las normas fijadas por el equipo investigador. En estas *normas* se indican:

- a) Los *giros* que han de realizarse: a la izquierda o a la derecha.
- b) Los *edificios* en los que debe entrarse: como los terminados en una cifra específica (por ejemplo, los acabados en 2).
- c) Si el *edificio* comprende más de una *vivienda*, se especifica la elección de la *escalera* (si hay más de una); de la *planta*; de la *vivienda* en la planta (siguiendo una *tabla aleatoria*); y, por último, de la *persona* a entrevistar en cada *vivienda* (de acuerdo con las *cuotas* fijadas).

En las *hojas de ruta* el entrevistador registra las actuaciones seguidas hasta la localización de los sujetos de la *muestra*, así como cualquier incidencia de utilidad para la supervisión del *trabajo de campo*.

Este procedimiento de selección muestral por *rutas aleatorias* presenta la gran *desventaja* de no garantizar que todas las unidades de la población tengan la misma probabilidad de ser elegidas, aunque la designación de las *rutas* sea *aleatoria*. Además, existe el peligro de sobrerepresentación de las personas que se encuentran, con mayor frecuencia, en las viviendas (amas de casa, jubilados, personas en paro).

Para obviar dicha desventaja se aconseja complementar el *muestreo por rutas* con el *muestreo por cuotas*. También se reitera –véase Arber (1993)– la necesidad de insistir, al menos tres veces, en aquellas viviendas donde previamente no se haya obtenido respuesta.

5.4.7. Muestreo estratégico

El *muestreo estratégico* o *de conveniencia* responde a una modalidad de *muestreo no probabilístico*, en el que la selección de las unidades muestrales responde a criterios subjetivos, acordes con los objetivos de la investigación. Por lo que comparte las ventajas y los inconvenientes básicos de cualquier *muestreo no probabilístico* (Cuadro 5.10):

CUADRO 5.10. Ventajas e inconvenientes del muestreo estratégico.

<i>Ventajas</i>	<i>Inconvenientes</i>
Simplicidad y economía del diseño muestral.	Imposibilidad de estimar el error típico.
Fácil de administrar.	Dificultad en la generalización de los resultados de la investigación más allá de los casos analizados, por la introducción de sesgos en la elección de la muestra.
No precisa de un listado de la población.	

Esta variedad de *muestreo no probabilístico* es habitual en *estudios cualitativos*, no interesados en la generalización estadística. Pero, también es frecuente en los *experimentos* realizados con personas que se ofrecen voluntarias; e, igualmente, en *estudios piloto*, en los que se comprueba la validez del instrumento de medición elaborado (un cuestionario, un guión de entrevista). Estos estudios no siempre precisan de la selección rigurosa de las unidades muestrales. Basta con escoger (y no necesariamente de forma aleatoria) a individuos que compartan las características de la *muestra* diseñada para la investigación definitiva.

EJEMPLO DE MUESTREO ESTRATÉGICO

Un *diseño muestral* viable para el estudio sobre las trayectorias laborales de los jóvenes madrileños, mediante *entrevistas en profundidad*, sería *estratégico*, si la selección de los jóvenes madrileños se hiciera de forma no aleatoria, sino consecuyente con criterios marcados por el investigador.

El investigador podría decidir que, para cumplir los objetivos de su estudio, basta con seleccionar un número reducido de jóvenes que constituyan casos “representativos” de la variedad juvenil actual (en sus comportamientos y características sociodemográficas). Para ello opta por delimitar la selección muestral a dos distritos polares de Madrid, como son los distritos de Salamanca y de Carabanchel. De ellos escoge, de forma deliberada, casos que juzga que “representan” al conjunto de la población cuyas características analiza.

El procedimiento de selección muestral concluye cuando se llega a la “*saturación teórica*”: cuando la información comienza a ser redundante, no aportando ningún nuevo punto de vista analítico. Por lo que se decide la conclusión de la recogida de información.

5.4.8. Muestreo de “bola de nieve”

Esta última variedad de *muestreo no probabilístico* difiere del anterior en que las unidades muestrales van escogiéndose, sucesivamente, a partir de las referencias aportadas por los sujetos a los que ya se ha accedido. Como, a su vez, los nuevos casos identifican a otros individuos en su misma situación, la *muestra* va aumentando, como una “bola de nieve”, hasta que el investigador decida cortar.

Este tipo de muestreo es de gran utilidad, cuando se carece de un *marco de muestreo* que recoja a la población de interés. Especialmente, en poblaciones que, por sus propias características –generalmente población “marginal”–, son difíciles de identificar y de localizar. En estos casos, la localización de uno de los integrantes de la *población* se convierte en doblemente valiosa, al suponer la identificación de otras personas pertenecientes a la misma *población*. Para ello se pregunta a los sujetos ya “descubiertos”, si conocen a otros en su misma situación.

Este procedimiento muestral comparte las ventajas y los inconvenientes comunes a los *diseños muestrales no probabilísticos*.

5.5. Ejemplos de diseños muestrales tomados de investigaciones reales

Como ilustración de lo expuesto en este capítulo, a continuación se detallan dos *diseños muestrales* correspondientes a dos encuestas nacionales: la Encuesta Nacional de Juventud 1988, dirigida por Zárrega, y el 2.º Informe de la Universidad Complutense, dirigido por A. de Miguel. Ambos *diseños muestrales* resultan de interés por la conjugación que hacen de varios procedimientos de selección muestral.

- Informe Juventud en España, 1988, *dirigido por José Luis de Zárrega Moreno* (Ministerio de Asuntos Sociales, Instituto de la Juventud, 1989).

La descripción del diseño de la muestra figura en el anexo de la investigación. En él se detallan los aspectos siguientes:

Universo de estudio

- a) *Definición*: población de 15 a 29 años, residente en el territorio español (de la península Ibérica e islas Baleares), en hogares familiares situados en núcleos de población de distinto tamaño. Por lo que, se exceptúa a los jóvenes que residen en Canarias, Ceuta y Melilla, junto con aquellos que viven fuera del hogar familiar.
- b) *Universo de referencia*. Se toma como universo de referencia el expuesto en el *Padrón municipal de 1986*, al ser ésta la fuente estadística más próxima al momento de la investigación (1988). El último *censo de población* era de fecha de 1981.

A efectos de distribución de la población en entidades dentro de cada municipio, se toman los datos del *Nomenclator del Censo de población de 1981*.

Para la fijación de las *cuotas* de sexo-edad-estado civil, el referente es *La encuesta de población activa*, publicada por el INE.

c) *Estratificación del universo*. La *población* se estratifica por región y tipo de hábitat. Para la *estratificación* (por región) se aplica la actual división político-administrativa del país en comunidades autónomas. Para la formación de los *estratos* según el tipo de hábitat, se aplica el criterio del tamaño del municipio, distinguiéndose las ocho categorías siguientes:

- 1) Menos de 2.000 habitantes.
- 2) De 2.000 a 5.000 habitantes.
- 3) De 5.001 a 10.000 habitantes.
- 4) De 10.001 a 50.000 habitantes.
- 5) De 50.001 a 100.000 habitantes.
- 6) De 100.001 a 250.000 habitantes.
- 7) De 250.001 a 1.000.000 habitantes.
- 8) Más de 1.000.000 habitantes.

El cruce de ambas variables (comunidad autónoma y tamaño del municipio) genera 128 estratos. De ellos, 26 quedan vacíos, por no existir en esa comunidad autónoma algún municipio del tamaño correspondiente.

Tamaño de la muestra

Considerados los objetivos del estudio, se estima aceptable un *error de muestreo* máximo de $\pm 1,5\%$ (con un *nivel de confianza* del 95,5%) en los datos a nivel nacional, para lo cual se precisa una *muestra* de 4.300 individuos, al menos.

Asimismo, se establece un *tamaño mínimo* de 625 entrevistas –que proporciona un *error de muestreo* máximo de $\pm 4\%$ – para las comunidades mayores y aquellas que presentan una diferenciación cultural superior respecto al conjunto nacional, con la finalidad de facilitar análisis específicos en cada una de ellas. Este criterio se aplica a las submuestras de las CC.AA. de Andalucía, Madrid, Cataluña, Galicia y País Vasco.

Por otra parte, también se adopta el criterio de ampliar las submuestras de solteros y de casados para tener datos significativos en ambos subconjuntos de población (de muy diferentes modos de vida).

Con estos criterios se diseña una *muestra básica* –con ampliaciones regionales– formada por 4.320 jóvenes, distribuidos por comunidades autónomas y tipos de hábitat como figura en el Cuadro 5.11. La asignación de entrevistas se hace de forma proporcional (*afijación proporcional*) al tamaño del *estrato* en la comunidad autónoma, respecto del total de España (según la fuente estadística consultada).

CUADRO 5.11. Distribución de la muestra base.

	N.º entrevistas	Distribución de las entrevistas según tamaño del municipio							
		Menos 2.000	2.000 5.000	5.001 10.000	10.001 50.000	50.001 100.000	100.001 250.000	250.001 1 millón	Más de 1 millón
Galicia	262	11	103	138	137	78	45	48	-
	+ 298								
Asturias	109	2	7	9	30	20	17	24	-
Cantabria	52	6	9	4	9	6	18	-	-
País Vasco	223	29	25	42	149	72	127	116	-
	+ 337								
Navarra	52	11	12	4	7	-	18	-	-
Aragón	112	30	9	6	16	-	-	51	-
Baleares	61	1	7	8	18	-	-	27	-
Cataluña	574	42	40	31	114	55	86	28	178
Castilla-León	260	99	29	14	23	23	42	30	-
Rioja	25	6	3	3	3	-	10	-	-
Extremadura	110	26	25	20	22	6	11	-	-
Madrid	474	10	8	8	29	34	76	-	395
	+ 85								
Castilla-La Mancha	163	42	28	30	31	21	11	-	-
C.Valencia	359	27	29	32	123	22	27	99	-
Andalucía	667	37	80	77	186	53	61	173	-
Murcia	96	1	1	10	32	6	17	29	-
Total	3.600	380	415	436	929	396	566	625	573
	+ 720								
	4.320								

A esta muestra básica se agrega una *ampliación de muestra*: por cada sujeto caso seleccionado en la *muestra básica* se elige (adicionalmente) un sujeto soltero del mismo sexo y grupo de edad. De este modo, se agregan 922 entrevistas adicionales a la *muestra básica*. La *muestra total* está, por tanto, formada por 5.242 jóvenes, distribución por sexo, edad y estado civil como se expone en el Cuadro 5.12.

CUADRO 5.12. Distribución de la muestra total por sexo, edad y estado civil.

	Varones			Mujeres		
	Total	Solteros	Casados	Total	Solteras	Casadas
15 - 19	747	746	1	746	757	7
20 - 24	795	733	62	859	715	144
25 - 29	1.001	705	296	1.076	664	412
Total	2.543	2.184	359	2.699	2.136	563

(Los desequilibrios introducidos en la muestra se corrigen durante el proceso de análisis, mediante *coeficientes de ponderación* que restituyen a los subconjuntos muestrales su peso correspondiente en la distribución del universo).

Selección de los puntos de muestreo

La primera etapa de la selección de la muestra consiste en la elección de los municipios dentro de cada extracto. En total se seleccionan 262 municipios, distribuidos por región y tipo de hábitat del modo siguiente (Cuadro 5.13):

CUADRO 5.13. Distribución geográfica de los puntos de muestreo.

	Distribución de las entrevistas según tamaño del municipio								Total
	Menos 2.000	2.000 5.000	5.001 10.000	10.001 50.000	50.001 100.000	100.001 250.000	250.001 1 millón	Más de 1 millón	
Galicia	1	12	9	9	4	1	1	–	37
Asturias	–	1	1	2	1	1	1	–	7
Cantabria	1	2	–	1	1	1	–	–	6
País Vasco	4	3	3	10	4	3	1	–	28
Navarra	1	2	–	1	–	1	–	–	5
Aragón	4	1	1	1	–	–	1	–	8
Baleares	–	1	1	1	–	–	1	–	4
Cataluña	6	5	2	8	3	4	1	–	30
Castilla y León	12	4	1	2	1	2	1	–	23
Rioja	1	–	1	–	–	1	–	–	3
Extremadura	3	3	2	2	1	1	–	–	12
Madrid	1	1	1	2	2	4	–	–	12
Castilla-La Mancha	6	4	2	2	1	1	–	–	16
Com. Valenciana	4	4	2	8	1	2	2	–	23
Andalucía	5	10	5	12	3	3	4	–	42
Murcia	–	–	1	2	1	1	1	–	6
Total	49	53	32	63	23	26	14	2	262

(La selección de los municipios dentro del estrato se hace aleatoriamente, asignando a cada municipio una probabilidad proporcional a su población.)

La segunda etapa consiste en la selección de núcleos de población dentro de cada municipio. Para esta selección se siguen los criterios siguientes:

- a) En cada municipio menor de 50.000 habitantes se elige al azar una entidad de población, con probabilidad proporcional al número de habitantes registrados en el *Nomenclator del Censo de 1981*.
- b) En los municipios mayores de 50.000 habitantes se calcula la proporción de la población que reside en la entidad principal y en el conjunto de las restantes entidades. Si, repartiendo el número de entrevistas asignadas al municipio (en proporción directa a esas poblaciones), corresponde al conjunto de las entidades menores de seis o más entrevistas de la *muestra básica*, entonces se selecciona aleatoriamente una de dichas entidades menores, y a ésta se le asignan las entrevistas. En caso contrario, todas las entrevistas asignadas al municipio se realizan en la entidad principal.

Proceso de selección de individuos

En una primera fase, la elección de los entrevistados se hace aplicando un *muestreo polietápico*, con selección de hogares por el procedimiento de *itinerarios aleatorios* y de individuos (dentro del hogar) mediante *números aleatorios*. En cada estrato muestral se aplica además un sistema de *cuotas* de edad y sexo para corregir las desviaciones que pudieran producirse por la incidencia de las negativas a participar en la investigación y las ausencias en los domicilios.

En una segunda fase del *trabajo de campo*, la correspondiente a la ampliación de individuos casados, se emplea un sistema de selección por *cuotas* de edad, sexo y estado, dentro de zonas elegidas aleatoriamente en las entidades de población de la *muestra básica*.

Estimación de los errores de muestreo

Con el procedimiento de selección muestral utilizado no es posible un cálculo exacto de los *errores muestrales*. En estos casos, suele tomarse como referencia los *errores* que corresponderían a un *muestreo aleatorio simple*, ya que, si bien la selección “en racimos” –como la del procedimiento de “*itinerarios aleatorios*”– incrementa esos errores, la *estratificación de la muestra* los disminuye.

- *Diseño muestral realizado por Valles Martínez, M. y Cea D’Ancona, M.ª A., para el 2.º Informe de la Universidad Complutense: De Miguel, A. (1994): La Sociedad Española 1993-1994, Madrid, Alianza.*

Para la realización de este *informe* se llevaron a cabo dos operaciones de encuesta: una dirigida a la población española en su conjunto, y otra a las personas de más de 64 años (mayores) que viven en residencias para la tercera edad. La primera, a personas de distintos grupos de edad y situación económica (jóvenes, adultos activos, amas de casa, y mayores). La segunda constituye, en cambio, una *muestra estratégica*, hecha con el propósito de contrastar la opinión de los ancianos institucionalizados con la de aquellos que siguen viviendo en su domicilio o en su entorno familiar.

A) *La encuesta general*

Ambito y universo

Aunque la *muestra* es prácticamente nacional –pues sólo se excluye a Ceuta y a Melilla–, la población de los municipios más rurales (en cada comunidad autónoma) no forma parte del *universo*, por razones de coste, fundamentalmente. Ahora bien, para

delimitar el margen de mayor ruralidad, teniendo en cuenta la diferente configuración de lo rural y de lo urbano en el territorio español, se considera la densidad y el tamaño medio de los municipios en cada comunidad autónoma, así como la división municipal en entidades singulares de población. Por lo que se decide la siguiente inclusión de municipios por comunidad autónoma (Cuadro 5.14):

CUADRO 5.14. Municipios incluidos en la muestra por comunidad autónoma.

<i>A</i> <i>Municipios de 1.000 y más habitantes</i>	<i>B</i> <i>Municipios de 3.000 y más habitantes</i>	<i>C</i> <i>Municipios de 5.000 y más habitantes</i>
Aragón Castilla La Mancha Castilla y León Extremadura Navarra La Rioja	Cantabria Cataluña Madrid País Vasco C. Valenciana	Andalucía Asturias Balears Canarias Galicia Murcia

Estratificación y afijación

Antes de proceder a la selección de los *puntos de muestreo*, se agrupan los municipios mediante el cruce de las 17 comunidades autónomas y el tamaño de la población, dividido conforme a las 9 categorías siguientes:

- 1) De 1.000 a 3.000 habitantes.
- 2) De 3.001 a 5.000 habitantes.
- 3) De 5.001 a 10.000 habitantes.
- 4) De 10.001 a 20.000 habitantes.
- 5) De 20.001 a 50.000 habitantes.
- 6) De 50.001 a 100.000 habitantes.
- 7) Municipios de más de 100.000 habitantes, no capitales de provincia.
- 8) Capitales de provincia.
- 9) Capitales autonómicas.

La *afijación* de las entrevistas a los distintos *estratos* se hace de forma proporcional al número de habitantes (*Censo de población, 1991*) del conjunto de municipios que componen el *estrato*.

Selección de los puntos de muestreo

En cada estrato se eligen los municipios, en función de su tamaño y el número de entrevistas afijadas. El procedimiento es directo, aleatorio y con probabilidad pro-

porcional a la población de cada municipio. Después se extraen, dentro de cada municipio, las entidades singulares de población. Por ello se barajan los criterios siguientes:

- a) *Municipios de menos de 10.000 habitantes.* Selección aleatoria de una entidad singular de población, con probabilidad proporcional al número de *habitantes de hecho* registrados en el *Nomenclator de 1986 ó 1981*. Éste es el caso de las provincias de Huelva, Zaragoza, Sta. Cruz de Tenerife, Toledo, Salamanca, Valladolid, Barcelona, Valencia, Orense y La Rioja. La entidad singular capital también entra en el sorteo.
- b) *Municipios de más de 10.000 habitantes.* *Afijación proporcional* de las entrevistas del municipio a la entidad singular capital y al conjunto de entidades singulares no capitales.

Cuando al conjunto de entidades singulares capitales corresponden seis o más entrevistas, se elige aleatoriamente (y con *probabilidad proporcional*) una de ellas, donde se hacen las entrevistas. En la entidad singular capital se realizan las correspondientes o todas, si en el supuesto anterior no se alcanzan las seis entrevistas.

c) *Criterios complementarios:*

- 1) Considerar entidades singulares elegibles sólo aquellas que cuentan con una población al menos 10 veces superior al número de entrevistas afijadas al municipio.
- 2) Aunque corresponda aplicar el criterio *b*, por el tamaño del municipio, si éste se compone de “barrios” y “lugares” de peso demográfico parecido al de la entidad singular capital, se utiliza el criterio *a*.

En resumen, han sido 111 los municipios seleccionados (pero 131 las entidades singulares de población visitadas, y por tanto los *puntos de muestreo efectivos*); lo que supone una dispersión aceptable. Más aún, si se tiene en cuenta que la *muestra* toca 42 de las 50 provincias españolas. (Ver cuadro 5. 15.)

Selección de los entrevistados

En el procedimiento de selección de los entrevistados se tienen en cuenta *cuotas* de sexo, edad y condición socioeconómica, tomando como referencia las estimaciones de la *Encuesta de población activa del INE*, correspondientes al tercer trimestre de 1991. Estos son los datos para los grandes subconjuntos de población que han compuesto la *muestra* general (Cuadro 5.16).

CUADRO 5.15. Puntos de muestreo por comunidades autónomas, provincias y municipios, con especificación del estrato al que pertenecen y el número de entrevistas teóricas a realizar.

<i>Comunidad autónoma</i>	<i>Provincia</i>	<i>Municipio</i>	<i>Estrato</i>	<i>Entrevistas</i>
Andalucía	Huelva	Aljaraque	3	23
	Cádiz	Algodonales Ubrique San Fernando Algeciras Cádiz	3	22
			4	38
			6	39
			7	37
			8	73
	Córdoba	Pozoblanco Córdoba	4	37
			8	73
	Granada	Granada	8	73
	Jaén	Martos	5	31
	Málaga	Coín Ronda Estepona Marbella	4	38
			5	40
			5	40
6			38	
Sevilla	Gínés Montellano Morón Sevilla	3	23	
		3	23	
		5	40	
		9	90	
Aragón	Huesca	Biescas	1	11
		Fraga	4	23
Zaragoza	Ateca Zaragoza	1	12	
		9	90	
Asturias	Oviedo	Gozón	4	33
		Avilés	6	36
		Oviedo	8	60
Balears	Mallorca	Muro	3	25
		Mahón	5	20
		Palma Mallorca	9	40
Canarias	Tenerife	Santiago Teide	3	20
		Candelaria	4	25
		Realejos	5	49
		Laguna	7	33
		Sta. Cruz Ten.	8	72
Cantabria	Santander	Santoña	4	25
		Santander	9	30

CUADRO 5.15. (Continuación)

<i>Comunidad autónoma</i>	<i>Provincia</i>	<i>Municipio</i>	<i>Estrato</i>	<i>Entrevistas</i>
Castilla-La Mancha	Albacete	Balazote	1	14
	Ciudad Real	Retuerta	1	13
		Porzuna	2	16
		Puertollano	5	34
Guadalajara	Guadalajara	8	48	
Toledo	Casarrubios Puebla Almorá Sonseca	1	14	
		3	24	
		3	25	
Castilla y León	Palencia	Palencia	8	51
	León	Boñar	1	14
		Villabalter	5	35
	Salamanca	Sta. Marta Tor. Salamanca	3	19
8			52	
Valladolid	Fresno Viejo Mojados Tudela Duero Valladolid	1	14	
		1	14	
		2	16	
		9	44	
Cataluña	Barcelona	Sta. Coloma C.	2	17
		Argentona	3	26
		Franquèses V.	4	35
		Montcada i R.	5	33
		Rubí	6	34
		Cerdanyola V.	6	35
		Badalona	7	69
		Hospitalet Ll.	7	69
		Barcelona	9	216
	Girona	Cassa Selva Salt	3	25
5			34	
Lleida	Lleida	8	39	
Tarragona	Ametlla Mar Deltebre Valls	2	17	
		4	35	
		5	34	
Comunidad Valenciana	Alicante	Bigastro	2	21
		Muro Alcoy	3	24
		Villena	5	35
		Orihuela	5	35
		Alcoy	6	46
	Castellón	Castellón Pl.	8	77

CUADRO 5.15. (Continuación)

<i>Comunidad autónoma</i>	<i>Provincia</i>	<i>Municipio</i>	<i>Estrato</i>	<i>Entrevistas</i>
Comunidad Valenciana	Valencia	Puig	3	24
		Alginet	4	33
		Utiel	4	32
		Untinyent	5	34
		Valencia	9	101
Extremadura	Badajoz	Puebla Maes.	1	14
		Fuente Cantos	3	17
		Don Benito	5	25
		Badajoz	8	35
Cáceres	Piornal	1	15	
	Talayuela	3	17	
Galicia	Coruña	Carreira	3	20
		Arzúa	3	21
		Ribeira	5	29
		Santiago	6	36
Lugo	Lugo	8	67	
Orense	Celanova	3	20	
	Barco Valdeo.	4	33	
Pontevedra	Tui	4	34	
Madrid	Madrid	Morata Tajuña	3	27
		S. Fernando H.	5	38
		S. Sebastián R.	6	46
		Fuenlabrada	7	60
		Móstoles	7	62
		Madrid	9	397
Murcia	Murcia	Cehegín	4	34
		Águilas	5	34
		Murcia	9	66
Navarra	Navarra	Doneztebe	1	18
		Barañain	4	17
		Pamplona	9	25
País Vasco	Vizcaya	Muskiz	3	17
		Ermua	4	39
		Galdakao	5	33
		Barakaldo	6	51
	Guipúzcoa	Lazkao	3	17
	Álava	Vitoria	8	97
La Rioja	Logroño	Logroño	9	30

CUADRO 5.16. Estructura sociodemográfica de la muestra.

<i>Características</i>	<i>Universo * Hab. (miles)</i>	<i>Muestra N°entrev.</i>	<i>% Universo y Muestra</i>
Jóvenes 18-29			
Varones	3.807	623	13,9
Mujeres	3.572	585	13,0
Total	7.379	1.208	26,9
Amas de casa 30-64	4.493	735	16,3
Adultos activos 30-64			
Varones	6.680	1.094	24,3
Mujeres	3.126	512	11,4
Total	9.814	1.605	35,7
Mayores 65 y más			
Varones	2.440	400	8,9
Mujeres	3.363	551	12,2
Total	5.803	951	21,1
Total	27.489	4.500	100,0

* Fuente: EPA. Año 1991 (Tercer trimestre).

Todas las entrevistas se realizan, cara a cara, por entrevistadores en los hogares familiares de los núcleos de población elegidos como *puntos de muestreo*. La selección de los hogares se hace conforme a los patrones de zonificación urbanística (centro, ensanche, extrarradio).

Estimación de los errores de muestreo

Debido a la combinación de varios procedimientos muestrales (*estratificación, conglomerados, cuotas*), no es posible, en rigor, el cálculo de los *errores* derivados del manejo de muestras y del *diseño muestral*. En estas circunstancias, habituales en los sondeos de opinión, se toman como referencia los *errores* que corresponderían a un *muestreo aleatorio simple*. Se trata de una solución sencilla, pero operativa y fundamentada, puesto que respecto al *muestreo aleatorio simple* la *estratificación* mejora el *error* y el *muestreo por conglomerados* lo aumenta.

Por otro lado, conviene recordar que los *errores de muestreo* también dependen de la altura del porcentaje que se analice y del tamaño de las subdivisiones muestrales.

B) Encuesta a mayores en residencias

La muestra estratégica a mayores en residencias se circunscribe al territorio de la Comunidad de Madrid. Por ésta y otras razones, los resultados de la investigación se toman como meramente indicativos, incluso para el ámbito madrileño.

La selección de residencias y de entrevistados está guiada por criterios de heterogeneidad (a modo de *muestreo por cuotas*) en las siguientes variables:

- a) *Titularidad de las residencias de la tercera edad.* Se tiene en cuenta la distinción entre residencias públicas y privadas, y la distribución porcentual tanto de residencias como de plazas en cada categoría.

Las cifras de referencia para el año 1988, según la Consejería de Integración Social de la Comunidad de Madrid, cuantifican las residencias privadas en un 89% (con un 58% de las plazas), mientras que las residencias públicas representan los porcentajes complementarios: 11% las residencias, y 42% las plazas.

En esta *muestra estratégica* son 24 las residencias visitadas. De ellas, 4 son públicas (en las que se realizan el 46% del total de entrevistas) y 20 privadas (donde se hacen el 54% de las 507 entrevistas).

- b) *Localización espacial de las residencias.* La diferenciación a este respecto es entre las residencias ubicadas en Madrid capital y las emplazadas en la periferia restante.

Aunque la fuente citada indica la concentración (65% de las residencias) en Madrid capital, se opta por sobrerrepresentar el ámbito periférico a costa del capitalino, en aras de una mayor dispersión espacial de la *muestra*. Además, en términos de plazas, el municipio de Madrid cuenta con menos plazas públicas (32%) que el resto del territorio de la Comunidad de Madrid, según un estudio del Ayuntamiento de Madrid (1988: 98 y ss). Éste ha sido el resultado (Cuadro 5.17):

CUADRO 5.17. Muestra a mayores en residencias, según hábitat.

	Madrid cap.	Periferia	Total
Residencias	8	16	24
Entrevistas	90	387	507

- c) *Sexo y edad de los mayores institucionalizados.* La no disponibilidad de datos publicados (con carácter oficial y suficientemente desagregados), hace que se empleen estas dos variables básicas de forma aproximada. Para ser precisos, no se considera como población residente para entrevistar a los mayores deno-

minados “no válidos” o “asistidos” (aquéllos necesitados de mayores cuidados), debido a la dificultad mayor para entrevistarles. Sin embargo, se decide dar *cuotas* de sexo y edad a los entrevistadores, con la finalidad de evitar el *sesgo muestral* hacia los residentes más jóvenes, con mejor salud o más dispuestos.

En el Cuadro 5.18 se describe la *muestra a mayores en residencias*.

CUADRO 5.18. Mayores entrevistados en residencias, según edad y sexo.

<i>Edad</i>	<i>Ambos sexos</i>		<i>Varones</i>		<i>Mujeres</i>	
-78 años	158	31%	64	36%	94	29%
78 y +	349	69%	114	64%	235	71%
Total	507	100%	178	100%	32	100%

Lecturas complementarias

Azorín, F. y Sánchez Crespo, J. L. (1986): *Métodos y aplicaciones del muestreo*. Madrid, Alianza.

Henry, G. T. (1990): *Practical sampling*, Newbury Park, Sage.

Levy, P. y Lemeshow, P. (1991): *Sampling of population: methods and applications*. New York, Wiley.

Moser, C. A. y Kalton, G. (1989): *Survey methods in social investigation*. Aldershot, Gower Publishing Company.

Rodríguez Osuna, J. (1991): *Métodos de muestreo*, Cuaderno Metodológico n.º 1. Madrid, CIS.

Rodríguez Osuna, J. (1993): *Métodos de muestreo. Casos prácticos*, Cuaderno Metodológico n.º 6. Madrid, CIS.

Ejercicios Propuestos

1. Compare los dos diseños muestrales desarrollados en el apartado 5.5, destacando las diferencias y semejanzas entre ellos.
2. Se desea conocer la opinión de los jueces sobre el nuevo Código Penal. Concrete las fases del diseño de la muestra.
3. Un equipo de investigadores desea entrevistar a enfermos de sida para analizar la repercusión de la enfermedad en su vida diaria. ¿Qué diseño muestral se adecuaría más a este objetivo de investigación? Especifique cómo seleccionaría las unidades muestrales.
4. En un municipio de 185.000 habitantes se extrae una muestra de 10.000 individuos para conocer la renta mensual por habitante en dicho municipio. Ésta resulta ser de 95.000 ptas al mes, con una desviación típica de 3.600. Calcular el error que representa la muestra y los intervalos de confianza. Afijar la muestra e indicar qué tamaño se precisaría para un error máximo del $\pm 3\%$.

<i>Clase social</i>	<i>Porcentaje</i>	<i>Desviación</i>
Alta	37,7	2.500
Media	39,9	2.100
Baja	22,4	1.600

5. A continuación se exponen los datos correspondientes a una encuesta de 2.000 entrevistas a jóvenes entre 18 y 30 años. ¿Qué tipo de afijación se ha efectuado? Pondere la muestra, si lo estima necesario.

<i>Hábitat</i>	<i>Universo % (por mil)</i>	<i>Muestra</i>	<i>A favor de ampliar el aborto</i>
Urbano	404	1000	75
Intermedio	364	600	67
Rural	232	400	58

9

EL ANÁLISIS DE LOS DATOS

Una vez que ha concluido la recogida de información, comienza la fase de *análisis e interpretación de los datos*. Ésta es una fase compleja e igualmente crucial en el proceso de investigación. Si bien, en los últimos años se ha visto beneficiada por los grandes avances producidos en el campo de la informática. En concreto, la expansión de paquetes estadísticos informáticos (adaptados a ordenadores personales), que cubren el amplio espectro analítico existente (tanto *cuantitativo* como *cualitativo*). La implantación de estos programas ha adquirido tal relieve que, en la actualidad, sería inconcebible un *análisis de datos* ajeno al procesamiento informático.

En las páginas siguientes se ofrece una panorámica de las *técnicas cuantitativas de análisis de datos*. La notable pluralidad de éstas, unido a condicionantes de espacio, limitan su exposición a una mera descripción de rasgos distintivos de cada una de ellas. Para un mayor conocimiento y profundización, se remite a la consulta de bibliografía especializada en *análisis de datos* (véase la selección de textos que figuran al final del capítulo, y en la bibliografía general al término del libro).

9.1. Los preliminares del análisis de datos

“El paso más importante antes de comenzar a diseñar un proceso de entrada de datos es determinar qué programas se utilizarán para analizar los datos y convenciones concretas respecto a los formatos de ficheros y de *missing data*, que pueden manejarse para esos programas” (Fowler, 1988: 128). De ahí que la exposición de las *técnicas estadísticas de análisis* comience por sus *preliminares* esenciales: la creación de los *ficheros de datos*, junto a la depuración de la información (con especial referencia a la “no-respuesta”).

9.1.1. La creación de los ficheros de datos

Cada programa informático comprende instrucciones concretas para la creación de *ficheros de datos* propios y *ficheros de definición de los datos*. No obstante, la mayoría de estos programas permiten el acceso a ficheros elaborados por otros programas.

Primero, se confecciona un *fichero de datos* para la introducción de la información obtenida durante la investigación. Los *ficheros de datos* presentan formas diversas, dependiendo del programa informático que se maneje. La más común es el fichero de datos en formato *ASCII* (*American Standard Code for Information Interchange*). Este se compone de filas y columnas de números. Cada *fila* representa los datos recabados de cada sujeto o caso observado. En las *columnas*, en cambio, figuran los valores de las variables correspondientes a cada caso (véase el Cuadro 9.1). Este tipo de formato se precisa para ejecutar la mayoría de los *paquetes estadísticos genéricos* (como el SPSS, el BMDP o el SAS).

CUADRO 9.1. Extracto de un fichero de datos.

```
001211611450021402032540210011011934212416421102111005312123211200502013
002312911340010292016332221012101321122216221210211120143223612520202041
00381161243002433201634111061111430031111120000221130271124413211311122
004911291150021402032530200143221400222122111003132122232040207110002110
005211911450021212011141110923322113002221121121101001104422342614001103
006311811460731242014542210923411230012111211202002202432232211300310012
007611711440021212017541111042311230012211222200001002432236211040053212
```

Los *paquetes estadísticos genéricos* también pueden, sin embargo, leer *ficheros de datos* creados por *bases de datos* (como el DBASE) u *hojas de cálculo* (como LOTUS). Estos ficheros incluyen nombres para las variables, además de información relativa a la amplitud de la variable y la *codificación* de los valores de “no-respuesta” (*missing data*).

Cuando los datos se han registrado en *formato ASCII*, hay que elaborar también un *fichero de definición de datos*. Este comprende instrucciones precisas para la lectura de los datos que componen el *fichero de datos*. Estas instrucciones incluyen información primordial sobre las variables analizadas en la investigación: *columnas* donde se hallan ubicadas en el *fichero de datos*; las *etiquetas* dadas a las variables y a sus atributos; así como la especificación de cómo se *codifican* los valores “*missing*” (o sin respuesta), para su tratamiento en el análisis de los datos.

En el Cuadro 9.2 figura un extracto de un *fichero de definición de datos* en SPSS. Comienza con el nombre del fichero donde se han grabado los datos. Prosigue con la relación de variables y el número de las columnas donde éstas se encuentran registradas. Después, se asignan *etiquetas* a las variables que lo precisen, al igual que a sus valores. En el SPSS, como en los demás programas informáticos, se dan al usuario instrucciones

para la correcta elaboración de estos ficheros. También se expone cómo transportar y convertir ficheros de otros programas, su lectura y demás especificaciones necesarias para el análisis de los datos.

CUADRO 9.2. Extracto de un fichero de definición de datos en SPSS.

```
Data list file= 'menores.dat'.
/id 1-7 anio 8 sexo 9 edad 10 nacional 11 etnia 12 anorma 13 alcohol 14
heroína 15 cannabis 16 estudio1 to estudio2 17-20 desescol 21 profepa 22
situacpa 23 apreiviv 24 ambiente 25 fuga 26 compania 27.
Variable labels desescol 'desescolaridad'
/profepa 'profesion del padre'
/situacpa 'situación empleo padre'
/apreiviv 'apreciación vivienda'.
Value labels sexo 1 'varon' 2 'mujcr'
/nacional 1 'cspanol' 2 'extranjero'
/estudio1 to estudio2 01 'primero EGB' 02 'segundo EGB' 03 'tercero EGB'
04 'cuarto EGB' 05 'quinto EGB' 06 'sexto EGB' 07 'septimo EGB'
08 'octavo EGB' 09 'FP' 10 'BUP' 11 'compensatoria' 12 'otros'
13 'analfabeto'.
Missing values anio to compania (0).
Frecuencias anio to compania /statistics=all /hbar.
Save outfile='menores.sys'.
```

Aparte de estos ficheros, pueden crearse *subficheros* específicos con objeto de facilitar los análisis. Especialmente, cuando el tamaño de la muestra es elevado y el *fichero de datos* originario adquiere un gran volumen.

Estos *subficheros* incluyen una *muestra de variables*, restringida a aquellas que sean de interés para la ejecución de análisis concretos. También, pueden representar una submuestra aleatoria de los casos observados en la investigación. En este último caso, la finalidad no sería tanto el ahorro de tiempo en el análisis de la información, sino la *validación* de los resultados estadísticos obtenidos de la otra submuestra de la muestra global. Así, por *ejemplo*, para la *validación de análisis multivariantes* (como el de *regresión*), se recomienda la división de la *muestra* en dos *submuestras*: una, para la consecución del modelo; y la otra, para su *validación*.

9.1.2. La depuración de la información

A la creación de ficheros sigue la *depuración de los datos*, como antesala del análisis. El investigador ha de identificar posibles *errores* cometidos en la *grabación de los datos*. A tal fin resulta conveniente solicitar al programa la relación de *frecuencias* de todas las *variables* introducidas en el *fichero de datos* ("*frecuencias*" en SPSS, o "*univariate*"

en SAS, por ejemplo). Esta relación incluye un listado con todos los *valores* de cada *variable*, la *frecuencia (absoluta y relativa)* de cada uno de ellos, los casos sin respuesta (*missing data*), *estadísticos univariados* y las *representaciones gráficas* que se soliciten.

De esta relación, se observará si existen anomalías en los *valores* de las *variables* codificadas. Más concretamente, si alguna de ellas incluye valores ajenos al recorrido o *rango* definido de la *variable*.

Por *ejemplo*, si en la variable sexo, que se ha codificado con sólo dos opciones de respuesta (1 ‘varon’ 2 ‘mujer’), aparecen casos con valores superiores a dos (3, 4, 5 u otro), éstos corresponderían a sujetos que han sido erróneamente codificados o grabados en el ordenador. Por lo que, habría que proceder a su comprobación y corrección posterior.

Algunos programas informáticos (como el SAS o el SPSS) proporcionan, además, especificaciones dirigidas a la *depuración de los datos* (tanto durante su introducción, como una vez concluida ésta).

Si se observan incongruencias en los *valores* de las *variables* (como la anteriormente ejemplificada), ha de procederse a su localización y corrección. Los *errores* pueden deberse a fallos en la introducción de los datos en el ordenador, pero no siempre. También pueden ser ocasionados por deficiencias en la recogida de información.

En el primer caso, la corrección resulta más viable e inmediata: se revisan los *cuestionarios* (u otro instrumento de recogida de datos que se haya empleado), hasta localizar los casos en que se han grabado mal los *valores* de las *variables*; posteriormente, se introducirían los *códigos* correctos de las *variables* correspondientes.

Pero, si los *errores* se deben a un mal registro de la información en el instrumento de medición, las posibilidades de corrección se restringen. La dificultad de contactar de nuevo con la fuente de información (las *unidades muestrales*) lleva, no a la transformación de *códigos numéricos*, sino a la eliminación de aquellos casos con datos incorrectos o inconsistentes. Éstos se sumarían a aquellos que originariamente no proporcionaron información alguna. De esta forma, se incrementaría el volumen de los llamados “*missing values*” (o *valores* con los que se codifican las respuestas en blanco o incorrectas, y que se dan por *perdidos*).

Cuando unos casos específicos presentan muchos “*missing values*” en la mayoría de las *variables*, suele decidirse su exclusión del *fichero de datos* (salvo que al investigador le interese el análisis y descripción de estos casos). Igualmente, si de alguna variable se obtiene escasa información (teniendo un elevado porcentaje de valores *missing*), también suele optarse por su exclusión para el resto de los análisis.

En general, antes de proceder al *análisis de los datos*, el investigador evalúa los *porcentajes de respuesta* (para cada *variable*), y los “*outliers*” registrados en la *matriz de datos*.

Por “*outliers*” se entiende cualquier observación o caso que muestre inconsistencia con la serie global de datos. Su identificación requiere la realización de *análisis univariados* (para cada una de las *variables*), tanto numéricos como gráficos. Además, la mayoría de los programas informáticos incluyen instrucciones específicas para la detección de “*outliers*”.

En cuanto a la “*no-respuesta*”, su evaluación resulta igualmente exigida. Diversos autores, como Bourque y Clark (1994), recomiendan la *comparación* de las caracte-

rísticas demográficas de la *muestra* con las correspondientes a la *población* de la que ésta procede. Para ello se emplea el *Censo de Población*, u otra fuente de datos estadísticos o estudio, que describa al conjunto de la *población*.

Si de este análisis se dedujese la no *representatividad* de la *muestra*, el investigador deberá establecer la magnitud de las diferencias entre la *población* y la *muestra*. Esto es importante para la delimitación de las posibilidades de *inferencia* de las *estimaciones muestrales*. En palabras de Arber (1993: 71):

“La capacidad para realizar inferencias de una muestra a una población se basa en el supuesto de que la muestra lograda no esté sesgada por la *no-respuesta*. En la medida en que aquellos que no responden difieran de forma significativa de aquellos que sí responden, el investigador tiene una muestra sesgada.”

9.2. El análisis estadístico univariable

En el diseño de la investigación ya se prevén los análisis a realizar con la información reunida en el desarrollo de la investigación. Aunque ha de matizarse que el *proyecto de análisis* no es inmutable. Depende, en gran parte, de la *cantidad* y *calidad* de los datos que se recaben. De ahí la importancia que adquieren, en cualquier indagación, los *análisis exploratorios*, como paso exigido y previo a la decisión de qué técnica analítica (*bivariable* y *multivariable*) se va a aplicar.

9.2.1. La distribución de frecuencias

En la *exploración de los datos*, primero se procede a un análisis exhaustivo de cada *variable* incluida en la *matriz de datos* (*análisis univariable*). Para cada una de las *variables* se calcula su *distribución* o *tabla de frecuencias*. La *tabla de frecuencias* —como puede verse en el Cuadro 9.3— incluye los distintos *valores* que presenta la *variable* (distribuidos en *clases* o *categorías*), acompañados por su *frecuencia* (es decir, el número de veces en que aparecen).

En la primera columna (encabezada por el rótulo “*value label*”) figuran los distintos atributos que componen la *variable*. La siguiente columna (“*value*”) muestra el *valor* dado a cada *atributo*. En la tercera columna (la denominada “*frequency*”) se hallan las *frecuencias absolutas*; o sea, el número de casos (de la *muestra*) que comparten cada uno de los *valores* de la *variable*.

Para conocer la importancia de cada *valor*, y a efectos comparativos, se obtienen las *frecuencias relativas* o *porcentuales*, que representan cada *valor* en el conjunto de la *muestra* (columnas 4.^a y 5.^a). Primero, se calculan los *porcentajes* para toda la *muestra*; segundo, exclusivamente para aquellos casos que han proporcionado información al respecto (columna de “*valid percent*” o porcentaje válido). En esta columna no se consideran, por tanto, los “*missing values*”.

Este desglose de casos, en función de si aportan o no información, permite conocer la proporción de “*no-respuesta*” de cada *variable*. Este conocimiento adquiere especial relevancia para posteriores análisis.

Por último, se calculan las *frecuencias relativas acumuladas* (“cum percent”), a partir de las *frecuencias* contenidas en la columna 5.^a. Estas *frecuencias acumuladas* denotan la proporción de casos (válidos) que se encuentran por debajo, o por encima, de un determinado *valor* de la *variable*.

Cuando la *variable* está medida a nivel de *intervalo*, se aconseja la previa *agrupación* de los *valores*. Ello facilitará su presentación en una *tabla de frecuencias* de menores dimensiones. A este respecto, algunos autores –como Bryman y Cramer (1995)– sugieren que el número de *categorías* diferenciadas esté comprendido entre 6 y 20. Argumentan que menos de 6 y más de 20 categorías pueden distorsionar la *forma* de la distribución de la *variable*.

Los *estadísticos univariados* que figuran en el Cuadro 9.3 se comentan en la sección 9.2.3, dedicada a su exposición.

CUADRO 9.3. Ejemplo de una tabla de frecuencias para la variable “estudio” en SPSS.

Value label	Value	Frecuency	Percent	Valid Percent	Cum Percent
Primero EGB	1	8	1.1	1.1	1.1
Segundo EGB	2	8	1.1	1.1	2.2
Tercero EGB	3	31	4.3	4.3	6.6
Cuarto EGB	4	59	8.3	8.3	14.8
Quinto EGB	5	119	16.6	16.7	31.5
Sexto EGB	6	152	21.3	21.3	52.8
Septimo EGB	7	177	24.8	24.8	77.6
Octavo EGB	8	72	10.1	10.1	87.7
FP	9	27	3.8	3.8	91.5
BUP	10	7	1.0	1.0	92.4
Compensatoria	11	1	.1	.1	92.6
Analfabeto	13	53	7.4	7.4	100.0
	0	1	.1	MISSING	
	TOTAL	715	100.0	100.0	
Primero EGB	8				
Segundo EGB	8				
Tercero EGB	31				
Cuarto EGB	59				
Quinto EGB	119				
Sexto EGB	152				
Septimo EGB	177				
Octavo EGB	72				
FP	27				
BUP	7				
Compensatoria 1					
Analfabeto	53				
Mean	6.566	Std Err	.091	Median	6.000
Mode	7.000	Std Dev	2.420	Variance	5.856
Kurtosis	1.644	S E Kurt	.183	Skewness	1.000
S E Skew	.091	Range	12.000	Minimum	1.000
Maximum	13.000	Sum	4688.000		
Valid Cases	714	Missing Cases	1		

9.2.2. Representaciones gráficas

La información contenida en una *tabla de frecuencias* también puede representarse de forma gráfica. Ello ayuda a la visualización global de la concentración, o dispersión, de los datos en la variable considerada.

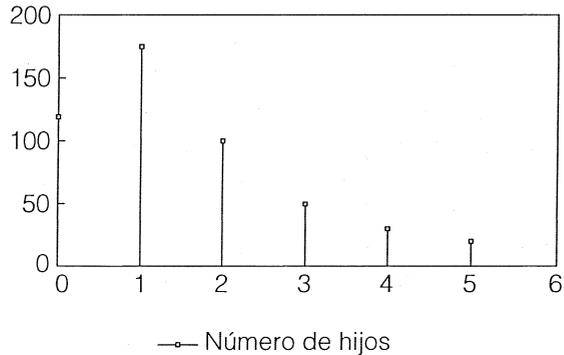
Dos de los *gráficos* habituales en la representación de *frecuencias* son el *diagrama de barras* y el *histograma*. A ellos se suman otros también usuales en el *análisis exploratorio*, como el diagrama de “*tronco y hoja*” y la “*caja*”; o el *polígono*, las *ojivas*, y el *gráfico de sectores*, entre la amplia variedad gráfica existente.

• *Diagrama de barras*

Consiste en una serie de “*barras*” (una para cada categoría de la variable), cuyas longitudes expresan las *frecuencias* de cada *atributo* de la *variable*.

EJEMPLO DE DIAGRAMA DE BARRAS

Número de hijos	Frecuencia absoluta
0	120
1	170
2	90
3	50
4	25
5	15
	470



• *Histograma*

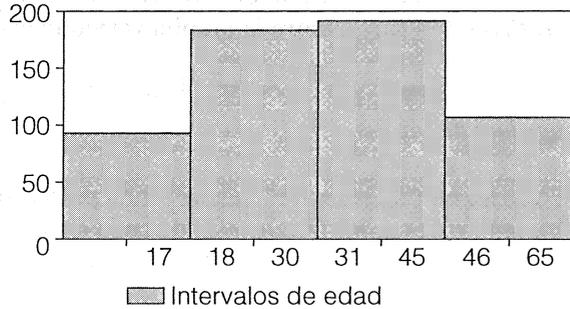
Un gráfico de contenido similar al anterior, aunque más adecuado a la representación de variables *métricas* (de *intervalo* o de *razón*).

En lugar de *barras*, el gráfico se compone de *rectángulos* unidos. Razón por la que se exige que la variable esté agrupada en *intervalos*. Estos forman las bases de cada uno de los *rectángulos*, que están delimitados por los *límites (reales)* de los respectivos *intervalos*. En cambio, la altura de los *rectángulos* es proporcional a la *frecuencia (absoluta o relativa)* de cada *intervalo*.

La suma total de las áreas de los rectángulos será igual a 1 (dado que la suma de todas las proporciones es la unidad).

EJEMPLO DE HISTOGRAMA

Edad	Frecuencia absoluta
Menos de 18	90
18 – 30	170
31 – 45	185
46 – 65	115
	560



• *El “Tronco y las Hojas” (“Stem-and-Leaf”)*

Constituye un gráfico parecido al *histograma*, pero integrado por los *dígitos* de los *valores* de las *variables*. Los *dígitos* se dividen entre dos. Los situados a la izquierda del punto (el *tronco*) figuran ordenados verticalmente, en orden creciente (de arriba a abajo). Por el contrario, los *dígitos* a la derecha del punto (las *hojas*), se disponen horizontalmente, aunque también en sentido creciente (de menor a mayor).

El *dígito* a la izquierda (la *columna*) que comprenda más *valores* a la derecha será aquél en el que se agrupen un mayor número de casos en la distribución. Por esta y otras razones, este gráfico suele tomarse como referente de las medidas de *tendencia central* de una distribución de frecuencias (a las que se hará referencia en el apartado 9.2.3).

EJEMPLO DEL GRÁFICO “EL TRONCO Y LAS HOJAS”

En el siguiente gráfico, puede observarse que la *fila* tercera representa los valores de mayor frecuencia en la distribución. En concreto, los *valores* 36, 36, 37, 37, 38, 38, 39, 39, 39.

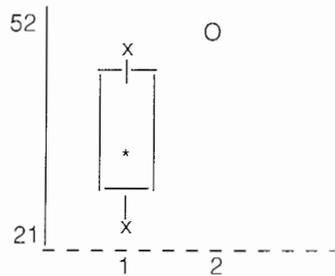
2	•	1 2 3 4
3	•	0 1 2 2 3 3 5
3	•	6 6 7 7 7 8 8 9 9 9
4	•	1 1 1 5 6 7 7
5	•	2 2

• La “Caja” (“Box-plot o Box-and-Whisker”)

Ofrece una visión global de la distribución, más sintética que la anterior. En ella la *variable* se representa de forma que el 50% de los casos queden comprendidos en el interior de la “caja”. En los extremos se sitúan, respectivamente, el 25% superior e inferior.

Con un asterisco se señala la *mediana*; con una “X” los valores máximos y mínimos; y con una “O”, los “*outliers*”. De esta manera se proporciona (gráficamente) información referente a la *mediana*, el *primer cuartil* (el 25% de los casos iniciales), el *tercer cuartil* (el 25% finales), y el *recorrido intercuartílico* (el 50% de los casos centrales) de la distribución de frecuencias. Ello exige que el nivel de medición mínimo de la variable sea el *ordinal*. En caso contrario, no podría estimarse el valor de la *mediana*, ni de ningún estadístico que precise de la ordenación de los *valores* de la variable, en un sentido creciente o decreciente (los *cuantiles*).

EJEMPLO DEL GRÁFICO LA “CAJA”



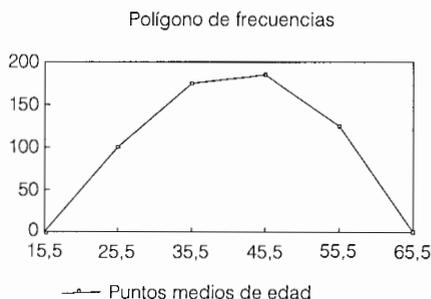
• Polígono

Un gráfico lineal que se traza sobre los *puntos medios* de cada *intervalo*, a una altura proporcional a su frecuencia (*absoluta* o *relativa*). Por lo que, de su visualización también se extraerán los *valores* con mayor (y menor) *frecuencia* en una distribución.

Los *puntos medios* se calculan sumando los *límites* de los *intervalos*, y dividiéndolo entre 2. De esta forma se obtiene el *valor* que representará al *intervalo* a todos los efectos. Por ejemplo, $(21 + 30)/2 = 25,5$.

EJEMPLO DE POLÍGONO DE FRECUENCIAS

Edad	Frecuencia absoluta	Punto medio
21 – 30	90	25,5
31 – 40	170	35,5
41 – 50	185	45,5
51 – 60	115	55,5
	560	



• Ojivas

Polígonos de frecuencias acumuladas que muestran la *frecuencia* de casos por encima, o por debajo, de un determinado *valor* de la distribución.

La *ojiva* será “*menor que*”, si se consideran los casos que hay por debajo de un *valor*. Por el contrario, será “*mayor que*”, cuando se representan los casos que comparten un *valor* superior de la distribución.

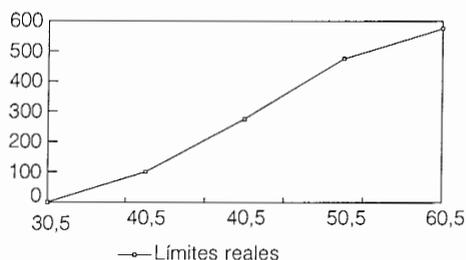
En el eje horizontal (de *abscisa*) se sitúan los *límites reales* de los *intervalos* de la *variable*, mientras que en el eje vertical (de *ordenada*) se disponen las *frecuencias acumuladas* (*absolutas* o *relativas*) de cada *intervalo*.

Para el cálculo de las *frecuencias acumuladas* se tiene en cuenta si interesa conocer el número (o la proporción) de casos que hay por debajo (*ojiva “menor que”*), o por encima (*ojiva “mayor que”*) de un *valor* específico de la distribución.

EJEMPLO DE OJIVAS “MENOR QUE” Y “MAYOR QUE”

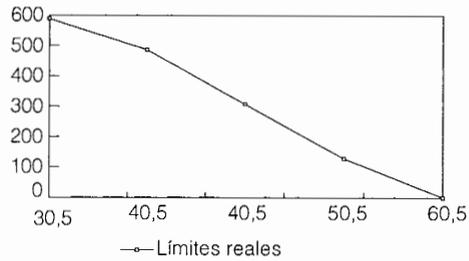
Ojiva “menor que”

Edad	Frecuencia absoluta	Frecuencia acumulada
21 – 30	90	90
31 – 40	170	260
41 – 50	185	445
51 – 60	115	560
	560	



Ojiva "mayor que"

Edad	Frecuencia absoluta	Frecuencia acumulada "más de"
21 - 30	90	560
31 - 40	170	470
41 - 50	185	300
51 - 60	115	115
	560	



• Gráfico de sectores (o en forma de "tarta")

A diferencia de los gráficos anteriores, éste se representa mediante un círculo, dividido en "sectores", cuyos ángulos indican el porcentaje de casos que comparten cada atributo de la variable.

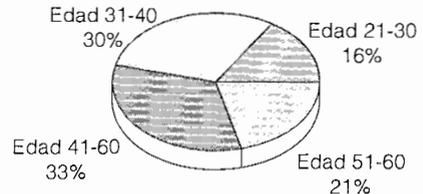
Para su obtención manual, se multiplican las *frecuencias relativas* (porcentual) de los *intervalos* (o valores) de la variable por 3,6 (que son los *grados* que corresponden a cada unidad porcentual: 360/100). De esta forma se transforman los *porcentajes* en *grados*, lo que posibilita su representación en una circunferencia.

No obstante, el empleo de programas gráficos (como el Harvard Graphics) hace innecesaria la realización de estos cálculos. Basta con conocer la *frecuencia* de cada *valor* de la distribución.

Los "sectores" del gráfico suelen representarse con colores o trazas diferentes, que proporcionen un mayor contraste entre ellos.

EJEMPLO DE GRÁFICO DE SECTORES

Edad	Frecuencia absoluta	Frecuencia relativa (%)	Grados
21 - 30	90	16	58
31 - 40	170	30	108
41 - 50	185	33	119
51 - 60	115	21	75
	560		360



Para un conocimiento más amplio de éstas y demás representaciones gráficas comunes en la investigación social, se remite a textos específicos sobre *Gráficos* (como el de Alaminos (1993), y otros citados en la bibliografía).

9.2.3. Estadísticos univariados

Aparte de los *gráficos*, en la *exploración univariada* también se emplean estadísticos para medir, de manera precisa, la distribución de los valores de una variable (véase Cuadro 9.3). Su uso dependerá, sobre todo, del nivel de *medición* de la *variable*.

Los *estadísticos univariados* se agrupan en tres grandes apartados:

- a) Medidas de tendencia central.
- b) Medidas de dispersión.
- c) Medidas de forma de la distribución.

• *Medidas de tendencia central*

Describen cómo se agrupan los *atributos* de una *variable* alrededor de un “valor típico” de la distribución. Por lo que, proporcionan una síntesis de la información contenida en la distribución.

Las medidas de *tendencia central* más empleadas en la investigación social son las siguientes: la *media*, la *mediana* y la *moda*.

- 1) La *media* es la medida más representativa, siempre y cuando la variable sea cuantitativa (de *intervalo* o de *razón*). Su cálculo precisa de la participación de todos los valores de la distribución. Cada uno de ellos se multiplica por sus respectivas *frecuencias absolutas*. Después, se suman todos los productos, y el resultado se divide por el número total de casos. De esta forma se obtiene el *promedio* de los valores de la distribución, que es como se define la *media aritmética*. El *inconveniente* fundamental de este estadístico es que se ve afectado por *valores* muy extremos en una distribución. Cuando la variable incluye *valores* muy dispares (o extremos), la *media* pierde representatividad a favor de la *mediana*, como *medida de tendencia central*.
- 2) La *mediana* es el valor que divide a la distribución en dos partes iguales. Se sitúa en el medio de la distribución. Lo que permite el conocimiento de los *valores* de mayor representación en la *muestra*. Su cálculo precisa que la variable sea, al menos, ordinal.
- 3) La *moda* denota el valor de mayor frecuencia en una distribución; aquel que más casos comparten. La distribución puede ser *unimodal* (una sola *moda*), pero también *bimodal* (dos *modas*), o *multimodal* (más de dos *modas*). Esto dificulta su interpretación, en caso de coexistir varias *modas*. A este *inconveniente* primordial se suma otro importante: en su cálculo no intervienen todos los *valores* de la distribución. Si bien, su principal *ventaja* es su universalidad. Pue-

de estimarse para cualquier tipo de variable, ya que el *nivel de medición* mínimo exigido es el *nominal*.

Además de estos estadísticos, existen los *cuantiles* como *medidas de tendencia no central*. Éstos representan *valores* que dividen a la distribución en partes iguales. Los *cuartiles* la dividen en cuatro partes iguales (cada una de ellas incluye al 25% de los *valores*); los *deciles* en diez partes; y los *percentiles* en cien partes. Su cálculo se asemeja al de la *mediana*. De hecho, el *cuartil* dos, el *decil* cinco, y el *percentil* cincuenta, expresan el *valor mediano*.

- *Medidas de dispersión*

Al conocimiento de los *valores centrales* de la distribución le sigue la medición de su representatividad: la mayor o menor *variabilidad* existente en torno a la *media* o la *mediana* de la distribución.

Las medidas de *dispersión absolutas* más comunes son el *rango* (o recorrido), la *desviación típica*, y la *varianza*.

- 1) El *rango* o *recorrido* expresa el número de valores incluidos en la distribución. Estos se obtienen de la diferencia entre el valor superior y el inferior. Su comprensión es sencilla, aunque presenta, en su contra, un *inconveniente* importante: es susceptible a la distorsión proporcionada por valores muy extremos en una distribución. Ello se debe a la única consideración del primer y último valor de la distribución.

Para solventar esta deficiencia, suelen aplicarse otras medidas de *rango*, que emplean un mayor volumen de información. Si bien, éstas exigen que la variable sea, al menos, *ordinal*. Se trata de los *recorridos*: *intercuartílico* (la diferencia entre el tercer *cuartil* y el primero; por lo que incluye al 50% de los valores centrales de la distribución), *semiintercuartílico* (el 25% de los valores centrales), *interpercentílico* (el 80% de los valores centrales, al ser la diferencia entre el *percentil* 90 y el 10), y el *semiinterpercentílico* (el 40%).

- 2) La *desviación típica* es el promedio de la desviación de los casos con respecto a la *media*. Como indicador de heterogeneidad (o de dispersión de los valores de una distribución), su estimación se exige siempre que se calcule la *media*, porque ayudará a la interpretación de su representatividad en la distribución.

Al igual que la *media*, su valor viene expresado en la unidad de medición de la *variable*, y únicamente puede calcularse cuando la variable es *cuantitativa*.

- 3) La *varianza* constituye otra medida de heterogeneidad de una distribución. Se define como el cuadrado de la *desviación típica*. Su valor expresa el grado de heterogeneidad de una población respecto a la *variable* medida, siendo sus características similares a la *desviación*.

A partir de la *desviación*, puede calcularse una medida de *dispersión relativa* que ayudará en su interpretación. Se trata del *coeficiente de variabilidad de Pearson*. Una medida estandarizada que se obtiene del cociente entre el valor de la *desviación* y la *media aritmética*. Su valor se expresa en porcentajes, siendo de utilidad en la comparación de la homogeneidad de dos o más grupos respecto a una o más variables.

Por último, cuando se calcula la *mediana*, también suelen estimarse estadísticos que midan su representatividad en la distribución. Concretamente, la *desviación media de la mediana* y el *coeficiente de variabilidad de la mediana de Pearson* (obtenido del cociente entre la *desviación media de la mediana* y la *mediana*). Su interpretación es análoga a los estadísticos anteriores.

Para un conocimiento más detallado, remito a la consulta de cualquier manual de estadística.

- *Medidas de la forma de la distribución*

Este término comprende la disposición espacial de los *valores* en una distribución: su *asimetría* y *curtosis*.

- 1) La *asimetría* constituye un indicador de la *agrupación* de las *frecuencias* en la *curva* de una distribución. Cuando su valor es cero, expresa que la *curva* es *simétrica*; es decir, que coinciden los valores de la *media*, la *moda* y la *mediana*. Si es mayor de cero, significa que la distribución es *asimétrica a la derecha* (*o de sesgo positivo*). Los casos se agrupan a la izquierda de la *curva*. Esto significa la mayor presencia en la distribución de *valores* inferiores a la *media*.

En cambio, si el valor del coeficiente de sesgo es inferior a cero, denota que la agrupación de los *valores* se produce a la derecha de la *curva*. Por lo que habrá una mayor representación de los *valores* superiores a la *media*.

En general, un valor superior a 0.8 (positivo o negativo) indica que la *asimetría* de la variable es importante.

- 2) La *curtosis* hace referencia a la mayor o menor *concentración* de *valores* en torno a la *media* de la distribución. Si existe una elevada concentración, la distribución será *leptocúrtica*. Esto significa que la *media* es muy representativa, al haber muy poca dispersión respecto a ella.

La distribución será *platicúrtica*, en el caso contrario: cuando existe una elevada *dispersión* de *valores* respecto a la *media*. Lo que expresa su escasa representatividad.

Con estos estadísticos lo que se pretende es comprobar si existen valores muy extremos en una distribución. En caso de existir, puede decidirse su transformación, con la finalidad de alcanzar una distribución que se aproxime a la *normal* (*simétrica* y *mesocúrtica*).

Tras los primeros *análisis univariados*, el investigador adquiere una descripción inicial de cada una de las *variables* que componen la investigación. Este conocimiento le

puede llevar a introducir *modificaciones en la configuración de las variables*, de forma que se faciliten los análisis posteriores. Estas *modificaciones* consisten, esencialmente, en agrupaciones de *variables* o de sus *atributos*.

Cuando en una distribución se observa que una o varias *categorías* presentan una *frecuencia* muy baja, habría que proceder a su *agrupación* con otras *categorías* de similar significado (si se pretende, con ellas, efectuar *análisis bivariantes* o *multivariantes*). Si la variable es *nominal*, la agrupación se produce con *categorías* que presentan un nexo común. Pero, si la variable es *ordinal*, de *intervalo*, o de *razón*, han de agruparse sólo los *valores* que se hallen más próximos en la escala (por *ejemplo*, las categorías “extrema derecha” con “derecha”).

A la nueva *categoría* (formada de la agrupación de dos o más categorías originales) habría que darle una nueva denominación, que resuma la variedad de *atributos* que comprende. De este modo las posibilidades de análisis, de variables con *atributos* de escasa representación en la *muestra*, se amplían; aunque la agrupación lleve consigo pérdida de información.

Igualmente, puede *agruparse variables* similares, con la finalidad de componer una medida única que sintetice la información contenida en *variables* análogas.

9.3. El análisis bivariante

Después de la realización de los *análisis exploratorios*, procede la realización de *análisis bivariantes*, tanto con fines *descriptivos* (describir al conjunto de la población observada), como *explicativos* (analizar posibles relaciones causales entre dos variables: la independiente y la dependiente).

9.3.1. Las tablas de contingencia

En la investigación social, la práctica habitual es la confección de *tablas de contingencia*, formadas del cruce, al menos, de dos *variables*. Estas *tablas* generalmente se obtienen mediante los comandos CROSSTABS y TABLES, en la mayoría de los programas estadísticos.

En estos comandos se especifican las *variables* a cruzar. Si puede establecerse una *relación causal* entre ellas, el orden convencional de exponer las *variables* es: *dependiente* “by” *independiente*. De esta forma, la *variable dependiente* figurará en las *filas*, mientras que la *independiente* en las *columnas*.

Si se desean introducir *variables de control* (para la eliminación de *explicaciones alternativas*), estas terceras y, a veces, cuartas variables, se añaden a las anteriores, siendo igualmente precedidas por la preposición “by” (por *ejemplo*, CROSSTABS = EDAD BY HERMANOS; CROSSTABS = MEDIDA BY DELITO BY SEXO). No obstante, en el programa informático que se emplee, se especifican las instrucciones propias para la ejecución del comando.

Además de las *variables*, ha de delimitarse la información que se precise: *frecuencias absolutas, porcentajes* (horizontales, verticales, totales), y *estadísticos* que midan el grado y la significatividad de la relación entre las *variables*.

El Cuadro 9.4 muestra una *tabla de contingencia simple* obtenida mediante el comando CROSSTABS de SPSS. En él se solicitaron las *frecuencias absolutas* ("count"), y las *relativas*: porcentajes horizontales ("row") y verticales ("column"), para cada *casilla* de la *tabla*. A ello se sumaron los *estadísticos de contingencia*.

CUADRO 9.4. Una tabla de contingencia simple mediante SPSS.

Crosstabulation:		Delito By medida			
Medida	Count	Amones- tación 1	Libertad vigilada 2	Interna- miento 3	Row total
	Row Pct Col Pct				
Delito	1	40 51.3 19.3	26 33.3 11.7	12 15.4 12.2	78 14.8
Hurto	2	12 19.7 5.8	36 59.0 16.1	13 21.3 13.3	61 11.6
Intimidación con armas	3	116 45.8 56.0	93 36.8 41.7	44 17.4 44.9	253 47.9
Robo sin intimidación	4	10 13.7 4.8	43 58.9 19.3	20 27.4 20.4	73 13.8
Insumisión paterna	5	29 46.0 14.0	25 39.7 11.2	9 14.3 9.2	63 11.9
Otras infracciones					
	Column total	207 39.2	223 42.2	98 18.6	528
<u>Chi-Square</u>	<u>D.F.</u>	<u>Significance</u>	<u>Min E.F.</u>	<u>Cells with E.F. < 5</u>	
41.32788	8	.0000	11.322	None	
<u>Statistics</u>	<u>Symmetric</u>	<u>With DELITO Dependent</u>	<u>With MEDIDA Dependent</u>		
Lambda	.07069	.00000	.13443		
Uncertainty Coefficient	.03488	.03033	.04102		
Somers'D	.04951	.05219	.04710		
Eta		.07227	.17561		
<u>Statistics</u>	<u>Value</u>	<u>Significance</u>			
Cramer's V	.19783				
Contingency Coefficient	.26943				
Kendall's Tau B	.04958	.0952			
Kendall's Tau C	.04959	.0952			
Pearson's R	.04051	.1764			
Gamma	.07346				
Number of Missing Observation = 0					

Cada *casilla* es el resultado del cruce de una *fila* con una *columna* (es decir, del cruce de un *atributo* de una *variable* con el *atributo* de la otra *variable*).

La lectura de las *tablas* con frecuencia se limita a comentarios porcentuales. Se contrastan los *porcentajes* de cada *casilla* para comprobar la existencia de variaciones entre los distintos *atributos* de las *variables*. A tal fin se calculan los *porcentajes* a partir de los *marginales* de la *tabla*.

Si se toma como base el total de *filas*, el *porcentaje* será *horizontal*, y las comparaciones porcentuales (entre los subgrupos) se efectuarán verticalmente. Por el contrario, cuando la base la constituye el total de *columnas*, se procede a la inversa: el *porcentaje* será *vertical*, y las comparaciones de porcentajes en sentido horizontal.

El investigador deberá escoger entre uno u otro tipo de *porcentaje* (horizontal o vertical), en conformidad con los *objetivos* del estudio y las *hipótesis* que compruebe. Si bien, existen mayores restricciones en *estudios explicativos* que en los *descriptivos*.

Si la finalidad de la investigación es la búsqueda de *relaciones causales*, los *porcentajes* se estiman sólo en el sentido de la variable *independiente*. Esta variable suele situarse en las *columnas*, y la *dependiente* en las *filas*; salvo que el elevado número de *atributos* de la variable *independiente* desaconseje su ubicación en las *columnas*. Esta disposición de las variables responde a la mayor facilidad de lectura (en la cultura occidental) en sentido horizontal frente al vertical. Los *porcentajes* serían, por tanto, *verticales* y las comparaciones horizontales.

Las diferencias porcentuales deben superar un determinado valor (al menos superior al 5%) para que puedan considerarse importantes. Depende del *error muestral* que derive de los tamaños de las *bases* sobre las que se calculan los *porcentajes*. Si estos *tamaños muestrales* son bajos, la diferencia porcentual ha de ser superior, si de ella quiere deducirse la existencia de *asociación* entre las *variables*.

En la exposición de la *tabla* (en el informe de la investigación) ha de indicarse, explícitamente, la dirección en la que se han calculado los *porcentajes*. Como sólo se aportan datos porcentuales, se recomienda poner entre paréntesis las *bases* de los *porcentajes*. Ello ayudará a la interpretación de las diferencias que en ellos se observen.

También se aconseja encabezar la *tabla* con un *título* que describa, sucintamente, el contenido de la *tabla*. En el *título* han de especificarse las *variables* comprendidas en la *tabla* y su relación.

A modo de ejemplo, la *tabla de contingencia* expuesta en el Cuadro 9.4 puede transformarse, en el *Informe*, como se expone en el Cuadro 9.5. Si bien, ha de matizarse que ésta constituye una de las posibles alternativas.

Pese a la existencia de convencionalismos en el formato de las tablas, el investigador es libre de diseñar el formato que más se ajuste a su estilo particular, y al contenido de la *tabla*.

La *tabla* expuesta en el Cuadro 9.5 constituye una *tabla simple* porque en ella figuran sólo dos *variables*. Pero también suelen componerse *tablas complejas*, a partir de

CUADRO 9.5. Menores clasificados por medida del tribunal, según el tipo de delito (porcentaje vertical).

Medida	Tipo de delito					
	Hurto	Robo con intimidación	Robo sin intimidación	Insumisión paterna	Otros delitos	Total
Amonestación	51	20	46	14	46	39
Libertad vigilada	33	59	37	59	40	42
Internamiento	16	21	17	27	14	19
Total	100 (78)	100 (61)	100 (253)	100 (73)	100 (63)	100 (528)

la conjunción de varias *variables* (independientes y/o dependientes). En estos casos, los análisis se centran en las variaciones en los *valores* extremos de las *variables*.

La *lectura porcentual*, aunque ilustrativa, resulta, no obstante, insuficiente. Precisa del complemento de *estadísticos* que gradúen la *asociación* entre las *variables* y su *significatividad*. Este complemento adquiere un mayor protagonismo cuando de las *tablas* quiera deducirse una *relación causal*.

Dos *variables* se hallan relacionadas si sus *atributos* varían conjuntamente. Para la graduación de esta relación, se acude a alguno de los *estadísticos de contingencia*, dependiendo del nivel de medición de la variable. Aquí sólo se enumeran. Consúltese en un manual de estadística su formulación.

- a) Si la variable es *nominal*, los estadísticos que miden el grado de *asociación* entre dos *variables* son: Phi cuadrado, “C” de Pearson, “V” de Cramer, “Q” de Yule, Lambda, Tau-Y de Goodman y Kruskal...; además del coeficiente “d” o de diferencia de proporciones.
- b) *Variables ordinales*: Rho de Spearman, Tau-A, Tau-B y Tau-C de Kendall, Gamma de Goodman y Kruskal, “D” de Sommer, entre otros.
- c) *Variables de intervalo*: a los estadísticos anteriores se suma el coeficiente de correlación producto-momento de Pearson.

Cada uno de estos estadísticos indican la fuerza y la dirección de la *asociación* entre dos variables. Su *signo* expresa la dirección de la *correlación* (positiva o negativa); mientras que el *valor* numérico (que oscila entre 0 y 1), la magnitud de la relación. De este modo:

- a) El valor “.00” denota *inexistencia* de *asociación*.
- b) “-1.00”, *correlación perfecta negativa* (conforme aumenta el valor de la variable independiente, disminuye el valor correspondiente a la dependiente).

- c) “1.00”, *correlación perfecta positiva* (al incremento de la variable independiente le sigue el aumento, también, de la dependiente).

Una vez medida la relación entre las variables, se comprueba su *significatividad*. Los datos analizados siempre pertenecen a una *muestra*, de las múltiples posibles, que pueden extraerse de una misma *población*. Razón por la cual, se exige la comprobación de la *significatividad* de los estadísticos y sus posibilidades de *inferencia* a la *población*.

En las *tablas de contingencia*, se aplica el *test de la Chi-Cuadrado* (X^2). Este estadístico se obtiene de la comparación entre las *frecuencias observadas* (en la *muestra*) y aquellas que cabría esperar en caso de inexistencia de relación entre las variables. Su valor se compara con el *teórico* (aquel que figura en una tabla de la X^2), para unos grados de libertad determinados y un nivel de significación escogido por el investigador. El *nivel de significación* habitual es .05. Este supone una posibilidad de error en la estimación del 5%. Los *grados de libertad*, en cambio, vienen marcados por las dimensiones de la *tabla*: número de *filas* (i) y de *columnas* (j). Concretamente, g. l. = (i-1)(j-1).

Cuando el valor de la X^2 empírico (el obtenido en la *muestra*) supera al teórico (el marcado en una tabla de la X^2), se deduce la *significatividad* estadística de la relación observada entre las *variables*. En caso contrario (X^2 empírico < X^2 teórico), se desestima la relación bivariable, por su no *significatividad*. La relación se consideraría, entonces, meramente casual, debida a *errores muestrales*.

En la salida de ordenador (véase Cuadro 9.4, por ejemplo) el valor de la X^2 aparece acompañado con su *significatividad* (“*significance*”). El valor que figura bajo este rótulo ha de ser inferior a .05 para que el valor de la X^2 sea significativo a un nivel de .05.

En cualquier manual de estadística en las ciencias sociales (como el de Blalock, 1978; García Ferrando, 1985; o Siegel, 1985) puede encontrarse una exposición detallada de estos y demás *estadísticos de contingencia*.

9.3.2. Otros análisis bivariantes

Aparte de las *tablas de contingencia*, existen otras *técnicas de análisis bivariantes*, como el de *regresión* y *varianza simple*. Ambas técnicas analíticas miden la relación de *dependencia* entre dos variables, si bien imponen mayores restricciones que las *tablas de contingencia*. Su cumplimiento exige que la *variable dependiente* sea *métrica* o cuantitativa. Ello determina su menor aplicabilidad en la investigación social, en la que predominan las *variables cualitativas* (*no métricas*).

- *Varianza simple*

Esta técnica analítica es muy aplicada en los *diseños experimentales*, en la comprobación de los efectos de los tratamientos experimentales. A ello contribuye la es-

pecificidad del análisis. Su finalidad es comprobar la existencia de diferencias grupales respecto a una única *variable dependiente* (*métrica*). Para ello se manipula una *variable independiente*, en función de cuyos valores se forman distintos grupos de tratamiento. Constituidos los grupos, se comprueba la media de cada uno de ellos respecto a la *variable dependiente*. Si se observan diferencias entre las *medias grupales*, se procede después a la comparación de las *varianzas grupales*, y a la medición de su *significatividad*.

Interesa que la *varianza* entre los grupos supere a la *varianza intragrupal*. Ello expresaría una mayor heterogeneidad entre los grupos, frente a una escasa variabilidad dentro de ellos. Por lo que podría afirmarse la existencia de diferencias entre los grupos.

La *significatividad* de las diferencias grupales se comprueba mediante los estadísticos “t” (si únicamente se han formado dos grupos de tratamiento), y “F” (si son más de dos los grupos creados).

En ésta, como en cualquier prueba de *significatividad*, se comparan los *valores empíricos* (“t” y “F”) con los *teóricos* (mostrados en las tablas de la “t” de Student y de la “F” de Fisher, correspondientes). El proceso es similar al descrito en el *test de la Chi-Cuadrado*. Se fija el *nivel de significatividad*, en función de la precisión que el investigador desee para su estimación (.05, generalmente); y los *grados de libertad* (ahora determinados por el *tamaño muestral* y el número de *variables independientes* consideradas). Siempre que el *valor empírico* supere al *teórico*, las diferencias grupales observadas en la *muestra* adquirirán *significatividad estadística*. Podrán, por tanto, hacerse extensibles al *universo* del que se extrajo la *muestra* (en los niveles de probabilidad fijados).

• *Regresión simple*

Constituye otra técnica de *dependencia* en la que se analiza la relación entre una única *variable independiente* (*métrica o no métrica*) y una *dependiente* (*métrica*). Pero, a diferencia de la técnica analítica anterior, la finalidad del análisis es la *predicción* del valor de la *variable dependiente* a partir del conocimiento de la *independiente*. Se cuantifica la relación existente entre ambas variables; y, se establece el grado de confianza o *significatividad* de la estimación efectuada.

La *correlación* entre las dos variables (*dependiente e independiente*) se mide mediante el *coeficiente R de Pearson*. Éste expresa el grado de *covariación* entre las variables, según se aproxime a “0” (inexistencia de asociación) o a “1” (asociación perfecta). También informa de la dirección de la asociación: creciente (si el signo es positivo) o decreciente (si es negativo).

En el *análisis de regresión*, la idea que subyace es la consecución de una *recta de regresión* que presente el mejor “ajuste” de los casos respecto a las variables analizadas. Esta *recta* tiene su expresión matemática en la siguiente *ecuación de regresión*:

$$y = a + bx + e$$

donde: “y” denota el valor de la variable *dependiente*.

“a” es el *intercepto* o punto de la *recta* que corta al eje de las Y.

“b” es la *pendiente* de la *recta* (también referido como el *coeficiente de regresión*). Su *valor* expresa la cantidad de variación de la variable *dependiente* por cada unidad de variación de la *independiente*. Su *signo* denota si se produce aumento (pendiente creciente; signo positivo) o disminución (pendiente decreciente; signo negativo).

“e” representa el *error de la estimación*: la inadecuación de la *ecuación de regresión* en la predicción del valor de la variable *dependiente*.

Esta *ecuación* permite la predicción del valor de la variable *dependiente* a partir de valores conocidos de la *independiente*. Los *coeficientes* se obtienen, generalmente, siguiendo el *criterio de mínimos cuadrados* (hacer mínima la distancia que separa los *puntos* –obtenidos de la confluencia de ambas variables en cada uno de los casos– y la *recta de regresión*).

La *significatividad* de los *coeficientes* se comprueba mediante el estadístico “t”, con $n-1$ *grados de libertad* (siendo “n” el número de observaciones). En cambio, la *significatividad* de la *correlación* se comprueba mediante el estadístico “F”. Como en cualquier prueba de *significatividad*, los valores de “t” y de “F” *empíricos* han de superar los *teóricos* (determinados en las tablas respectivas) para que el modelo de *regresión* sea significativo estadísticamente.

9.4. El análisis multivariable

Los análisis *univariantes* y *bivariantes* con frecuencia se muestran insuficientes para cubrir los objetivos de la investigación. El proporcionar una visión conjunta e integrada, que describa y/o explique la realidad que se analiza, demanda la realización de *análisis multivariantes* (de más de dos variables al mismo tiempo). De otra forma no podrían medirse las influencias e interrelaciones existentes entre grupos de variables. Como hace tiempo reconociera García Ferrando (1979: 198):

“Las distribuciones bivariantes en sociología aparecen demasiado simplistas para lograr adecuadas explicaciones científicas.”

La peculiaridad del *análisis multivariable* reside en operar con un número elevado de variables, y de manera simultánea, basándose en el cálculo matricial. Kendall (1975) lo define como el conjunto de técnicas estadísticas que permite el análisis simultáneo de más de dos variables en una muestra de observaciones. A esta definición, Dillon y Goldstein (1984) añaden la posibilidad de análisis sincrónicos de mediciones en más de una *muestra*.

Para el *análisis multivariable* existe un amplio abanico de *técnicas*. En conformidad con un extenso grupo de autores (véase Kendall, 1975; Dillon y Goldstein, 1984; o Hair *et al.* 1992, por ejemplo), estas técnicas pueden agruparse en dos grandes categorías (*técnicas de dependencia* y *técnicas de interdependencia*), en función de si se diferencia, o no, entre variables *dependientes* e *independientes*.

En la elección de la *técnica* concreta a aplicar intervienen, básicamente, el objetivo de la investigación, y las características de las variables que se analicen (su número y nivel de medición).

A continuación se ofrece una breve descripción de cada una de las *técnicas multivariadas* comúnmente referidas. Para un mayor conocimiento de cada una de ellas remito a la bibliografía especializada. Lo que sigue es una mera visión panorámica de las *técnicas multivariadas*.

9.4.1. Técnicas multivariadas de dependencia

Un conjunto de técnicas analíticas unidas por un mismo propósito: medir la existencia de *relaciones causales* entre un conjunto de variables, el grado y *significatividad* de la misma. Sin embargo, difieren en el número de variables *dependientes* que incluyen, y en el nivel de *medición* exigido (*métrico* o *no métrico*). En la Figura 9.1 se esquematizan las principales alternativas, comúnmente señaladas, en el análisis de la *dependencia*.

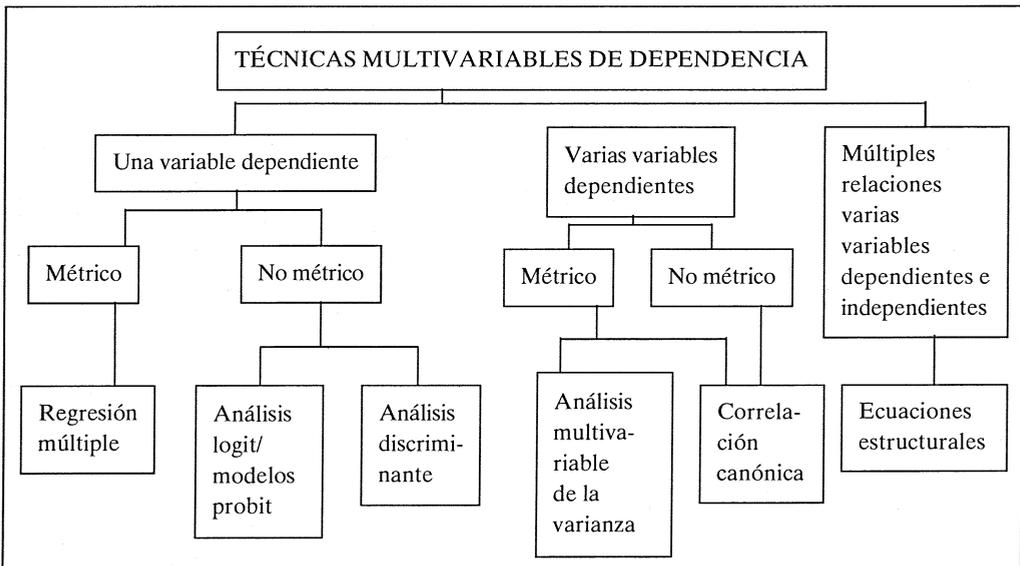


Figura 9.1. Técnicas de análisis multivariable de dependencia.

- *Regresión múltiple*

Una técnica de *dependencia* muy versátil, y utilizada, cuando se busca la *predicción* del valor de una variable *dependiente* (o *criterio*), a partir de valores conocidos en una serie de variables *independientes* (o *predictoras*). Para su realización se exige el cumplimiento de una serie de *supuestos* básicos:

- Linealidad*: la interdependencia entre las variables ha de responder a un modelo lineal. Esto quiere decir, que a cada cambio en una unidad en una variable corresponda un cambio, también en una unidad, en la otra variable.
- Normalidad*: la correspondencia de la distribución de los datos (para la variable *dependiente* y las *independientes*) con la *curva normal*. Esto permite la utilización de los estadísticos “*F*” y “*t*” para la comprobación de la *significatividad*.
- Homocedasticidad* o igualdad de las *varianzas* de los términos de error en la serie de variables *independientes*.

Para que pueda medirse la relación de las variables *independientes* con la *dependiente* se precisa que la *varianza* de los valores de la variable *dependiente* sea igual en cada valor de las variables *predictoras*.

- Aditividad*: los efectos de las variables *independientes* han de poderse sumar entre sí, para poder predecir la variable *dependiente*.
- Ausencia de *colinealidad* (de correlación) entre las variables *independientes*, con objeto de que puedan medirse sus efectos concretos en la variable *dependiente*.
- Ha de haber un número elevado de observaciones. Al menos, debería haber 20 veces más casos que variables *independientes*. Máxime si el modelo de regresión se obtiene “paso a paso”: incorporando una a una las variables *independientes* hasta que no exista ninguna más con poder predictivo significativo.

Cuando no se obtiene esta proporción, habría que optar por alguno de los siguientes remedios: eliminar alguna variable *independiente*, o agrupar varias variables creando una nueva variable (que sea una combinación de dos o más variables *independientes*).

Como en *regresión simple*, en *regresión múltiple* se obtiene una *ecuación de regresión*, con tantos *coeficientes* como variables *independientes*. A partir de ella, podrá predecirse el valor medio de la variable *dependiente*. Asimismo, se mide el grado de *correlación* existente entre las variables (mediante el estadístico *R múltiple*) y su *significatividad* (gracias al estadístico “*F*”).

El valor del R^2 *múltiple* (el *coeficiente de determinación*) expresará el porcentaje de *varianza* de la variable *dependiente* que es explicado por las *independientes*. Cuanto más elevado sea su valor, mejor para la predicción de la variable *dependiente*.

- *Análisis logit*

Los análisis *logit* y *probit* se definen como modelos analíticos alternativos (al modelo de *probabilidad lineal*) para variables *dependientes cualitativas*. Ambos modelos (*logit* y *probit*) se asemejan bastante. Si bien, el *análisis logit* ha alcanzado un mayor desarrollo, en parte debido a su mayor practicabilidad. De ahí que se le conceda un mayor protagonismo en esta breve reseña analítica.

El *análisis logit* constituye un modelo de respuesta cualitativa, que mide la relación entre una serie de variables *independientes (métricas o no métricas)* y una única variable *dependiente categórica (no métrica)*. Para ello se basa en el análisis de la *razón de probabilidad* de una variable *dependiente*, en función de las *independientes*. De lo que se trata es de graduar la *probabilidad* de que determinadas variables *independientes* provoquen la ocurrencia de un evento concreto (la variable *dependiente*).

Del análisis se obtiene una *ecuación* similar a la de *regresión*. Los *coeficientes* (calculados siguiendo el criterio de *máxima verosimilitud*, o el de *mínimos cuadrados*) se interpretan como en *regresión*. La *significatividad* de cada uno de ellos también se comprueba mediante el estadístico “*t*”.

La peculiaridad del análisis reside, no obstante, en el cálculo de los *incrementos de probabilidad*. Éstos se calculan para cada variable con *coeficiente* significativo. Se considera el valor del *coeficiente* (β) y el *valor medio* de la variable en la *muestra* (X_j), siguiendo la expresión de McFadden (1974):

$$\text{Prob}(Y = 1) = F(X', \beta) = \frac{e^{X'_j \beta}}{1 + e^{X'_j \beta}}$$

La *significatividad* del modelo global se comprueba mediante el estadístico X^2 . Este estadístico se complementa con el *porcentaje de aciertos* (el porcentaje de observaciones que logran ser correctamente estimadas mediante los *coeficientes logit*), como un indicador del éxito del modelo en la predicción de la variable *dependiente*.

- *Análisis discriminante*

Una técnica de clasificación y de asignación de individuos a grupos, a partir del conocimiento previo de sus características.

Su objetivo básico es la estimación de la relación existente entre una serie de variables *independientes* y una única variable *dependiente no métrica* (o *categórica*).

El número de *categorías* que incluya la variable *dependiente* determina los grupos formados a partir de la *muestra*. Estos *grupos* deben, previamente, haberse configurado mediante la aplicación de una *técnica multivariable de interdependencia* (como el análisis de *conglomerados* o el *factorial*).

Con el *análisis discriminante* lo que se pretende es la comprobación de si los *grupos* creados por otra técnica analítica son adecuadamente caracterizados por las variables que les definen (de acuerdo con los análisis previos).

Asimismo, se trata de conocer cuál es la combinación de variables (*funciones discriminantes*) que hace máxima la diferenciación entre los grupos. El conocimiento de estas variables ayudará a la predicción de la probabilidad de pertenencia de un individuo concreto a uno de los grupos diferenciados. Ello dependerá de los valores que presente en las variables *independientes* analizadas.

La ejecución del *análisis discriminante* exige, igualmente, el cumplimiento de unos *supuestos* claves, como son:

- a) La *normalidad* de las variables *independientes*. Se recomienda que éstas sean *métricas*. En caso contrario, habría que introducir modificaciones en el análisis.
- b) *Matrices de covarianzas* iguales en cada grupo.
- c) Inexistencia de *colinealidad* entre las variables *independientes*, para evitar que su información sea redundante en la explicación de la variable *dependiente*.
- d) Ha de haber, al menos, 20 casos por cada variable *independiente* introducida en el análisis. Esto contribuye a la *significatividad* estadística del modelo obtenido.

La distancia entre los grupos se mide mediante el estadístico D^2 de Mahalanobis. A éste se añaden los estadísticos *lambda de Wilks*, F y X^2 , en la comprobación de la *significatividad* de las variables *independientes* en la diferenciación entre los grupos.

La contribución de estas variables se cuantifica mediante los *coeficientes de función discriminante estandarizados* (similares a los *coeficientes beta* en *regresión*) y los *factores de carga discriminantes* (los “*discriminant loadings*”). Sobre todo, por estos últimos, debido a su mayor precisión.

Como en el *análisis logit*, en el *discriminante* también se calcula la *razón de aciertos*, como medida del éxito del modelo en la predicción de la variable *dependiente* (la clasificación de los grupos).

- *Análisis multivariable de la varianza*

También conocido como MANOVA. Representa una extensión del análisis univariado de la varianza (ANOVA), mediante el cual se exploran, simultáneamente, las posibles relaciones existentes entre: varias variables *independientes no métricas* (normalmente referidas como *tratamientos*) y dos o más variables *dependientes métricas*. Lo que permite la medición de las *correlaciones* entre las variables *dependientes* y entre las *independientes*.

El uso de este análisis se adecúa, igualmente, a los *diseños experimentales*, en la comprobación de los efectos de distintos *tratamientos*. El fin que se pretende es la determinación de la existencia de diferencias, en los *valores medios* de las variables *dependientes*, en cada *grupo de tratamiento*. Ello exige que:

- a) Las variables *dependientes* se hallen distribuidas *normalmente*.
- b) Se parta de *matrices de varianza-covarianza* iguales en todos los *grupos*. En caso contrario, no podría medirse el efecto específico del *tratamiento* dado (debido a que se partiría de grupos inicialmente diferentes).
- c) El *tamaño muestral* ha de superar al necesario para la *varianza simple*, si se pretende que el modelo adquiera *significatividad* estadística.

La *significatividad* de las diferencias multivariadas entre los *grupos* se comprueba mediante cuatro criterios primordiales: la *raíz máxima de Roy*, la *Lambda de Wilks*, la *traza de Hotelling* y la de *Pillai*. Si estos tests multivariados resultan significativos, se procede a la comprobación de las diferencias grupales en cada una de las variables *dependientes*. Para ello se acude al estadístico “*F*”.

• *Correlación canónica*

Una *técnica de dependencia* que permite la comprobación de la existencia de interrelación entre una serie de variables *dependientes* y otra serie de variables *independientes*. Ambas pueden ser tanto *métricas* como *no métricas*.

La finalidad principal del análisis es la obtención de *combinaciones lineales* de cada serie de variables (*dependientes* e *independientes*). Estas combinaciones han de maximizar las *correlaciones* entre las variables.

El grado de relación entre la serie de *variables canónicas* se mide mediante el *coeficiente de correlación canónica*. Éste, elevado al cuadrado (R^2), representa el porcentaje de *varianza* de una combinación de variables *dependientes canónicas* que es explicada por una combinación de las variables *independientes*. Su *significatividad* se comprueba, de nuevo, mediante el estadístico “*F*”.

Las contribuciones de cada variable (*dependiente* e *independiente*) a la combinación o serie de *variables canónicas* respectiva se comprueba, al igual que en el *análisis factorial* y *discriminante*, mediante los *factores de carga* (“*canonical loadings*”). Estos han de ser $\geq .30$ para que se consideren relevantes. La contribución de las variables también puede comprobarse mediante los *pesos canónicos* (“*canonical weights*”), si bien estos últimos presentan una mayor inestabilidad que los anteriores (los *factores de carga*).

• *Ecuaciones estructurales*

También conocidas como modelos LISREL, en referencia a uno de los programas estadísticos más populares para su resolución.

Esta última *técnica de dependencia* puede catalogarse como una extensión del *análisis factorial* y de *regresión múltiple*, por dos razones fundamentales:

- a) En las relaciones de *dependencia* se representan variables *latentes* o constructos (no observadas), a partir de valores conocidos en las variables *manifiestas* (u observadas) –a semejanza con el *análisis factorial*–.
- b) Las diversas *relaciones causales* tienen su expresión matemática en varias *ecuaciones de regresión* (ahora denominadas *ecuaciones lineales estructurales*).

Pero, a diferencia de las *técnicas de dependencia* anteriormente expuestas, en los *modelos de ecuaciones estructurales* se comprueba (simultáneamente) diversas *relaciones causales*. No sólo entre varias variables *independientes* (*métricas o no métricas*) y *dependientes* (*métricas*), sino entre ellas mismas, también (dependientes con dependientes e independientes con independientes). De ello resulta un *modelo causal* de mayor complejidad, aunque más ajustado a la pluridimensionalidad de las *relaciones causales*.

Las *relaciones causales* también se representan gráficamente, mediante un diagrama de “sendero” (“*path*”). En él se diferencian las variables *latentes* de las *observadas*, además de los *errores de medición* (de cada modalidad de variable), y el tipo de relación entre ellas (por ejemplo, una flecha bidireccional indica la existencia de correlación entre dos variables).

Para cada variable que recibe una flecha se formula una *ecuación*. En esa *ecuación*, dicha variable actúa como variable *dependiente*.

La realización de esta técnica analítica exige el cumplimiento de los mismos supuestos que en las anteriores (como la *independencia* de las observaciones; la *linealidad* de las relaciones; la *normalidad* multivariable; la selección *aleatoria* de las unidades muestrales; y un elevado *tamaño muestral*).

Los *parámetros* se obtienen mediante el método de *mínimos cuadrados generalizados*, o por el de *máxima verosimilitud*. Si bien, el primer método muestra imprecisión, conforme desciende el tamaño de la *muestra* y aumenta la complejidad del modelo. La *significatividad* de los *parámetros* se comprueba mediante la “*t*”, además de considerarse el *error* de la estimación.

Como se diferencia entre variables *latentes* y *observadas*, se comprueba la *fiabilidad* ($\geq .70$) y la *validez* ($\geq .50$) de la *medición* de los *constructos teóricos* (*variables latentes*). A ello se suma el *ajuste* global del modelo alcanzado respecto al inicial.

Existen distintos *índices de ajuste*. Entre ellos destacan los siguientes: el estadístico X^2 (aunque ahora interesan los valores bajos, porque expresan un mejor *ajuste* entre el modelo obtenido y el propuesto inicialmente); el *índice GFI*, que oscila entre .00 y 1.00 (cuanto más se aproxime a 1, mejor); los *índices TLI* ($\geq .90$), *NFI* ($\geq .90$), *AGFI* ($\geq .90$), y *AIC* ($\geq .70$), entre otros.

9.4.2. Técnicas multivariadas de interdependencia

A diferencia de las técnicas analíticas anteriores, las de *interdependencia* presentan un menor poder predictivo. Mediante ellas se analiza la existencia de *asociación* o

relación mutua entre varias variables, sin diferenciar entre *dependientes* e *independientes*.

La Figura 9.2 representa las principales técnicas de interdependencia. La elección entre una u otra responde a los mismos criterios fundamentales mencionados en las *técnicas de dependencia*. Concretamente, el objetivo de la investigación y el nivel de medición de las variables. Ahora la diferencia básica entre las técnicas se establece en función del nivel de *medición* mínimo exigido en las variables para su cumplimentación: *métrico* o *no métrico*.

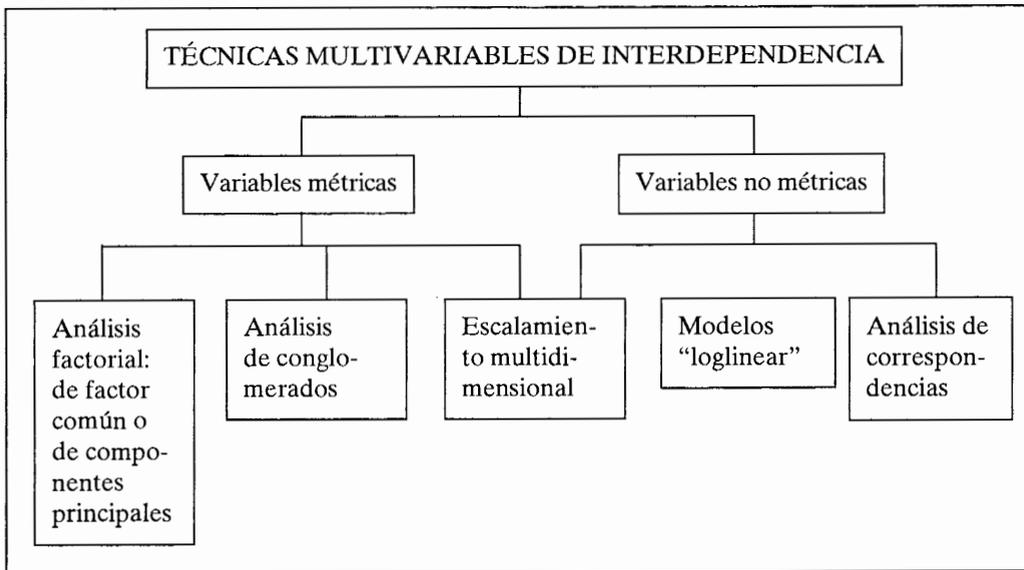


Figura 9.2. Técnicas de análisis multivariable de interdependencia.

- *Análisis factorial*

Una denominación genérica que engloba distintos procedimientos centrados en el análisis de la variación total que una variable comparte con otras variables. Su objetivo fundamental es resumir, la información contenida en un conjunto de variables interrelacionadas, en un número reducido de *dimensiones latentes* comunes (o *factores*).

El *análisis factorial* es *exploratorio*, cuando no se conoce, previamente, cuáles son los "factores". Éstos se determinarán, precisamente, tras el *análisis factorial*. En cambio, el análisis será *confirmatorio*, si se parte de unos "factores" especificados *a priori*, mediante un *análisis factorial exploratorio* u otra técnica multivariable (como el análisis de *conglomerados*). Estos *factores* representan a un conjunto de variables *empíricas* u observadas. Entonces, se tratará de "corroborar" la adecuación de estas variables (empíricas) en la medición de las *dimensiones* de los *conceptos teóricos*.

Además de esta clasificación básica, se diferencia entre análisis factorial de *componentes principales* y de *factor común*. En el *análisis de componentes* se persigue la obtención de *combinaciones lineales* de variables que logren explicar la mayor proporción de *varianza* conjunta. Para ello se tiene en cuenta tanto la *varianza específica* de cada variable, como la *varianza conjunta* (o compartida con otras variables). Por el contrario, en el *análisis de factor común*, la identificación de los *factores latentes* responde sólo a la *varianza común* de una serie de variables.

En cualquiera de las modalidades de *análisis factorial*, el investigador deberá concretar el método para la extracción de los *factores*: *ortogonal* (quartimax, varimax o equimax) u *oblicuo*. En el *ortogonal*, los *factores* se extraen de manera que sean totalmente independientes unos de otros. La extracción *oblicua* permite, en cambio, la existencia de correlación entre los *factores*.

En la decisión de cuántos *factores* escoger han de compaginarse distintos criterios como, por ejemplo, el criterio de *raíz latente* (*autovalor* superior a 1), o el de *porcentaje acumulado de varianza* (al menos superior al 60%).

Las *correlaciones* de las variables *empíricas* con los *factores* se comprueban mediante los *factores de carga* ("*factor loadings*"). Éstos han de ser $\geq .30$ para considerarse significativos.

Una vez definidos los *factores*, se les asigna un nombre o *etiqueta*, que refleje el contenido de las variables *empíricas* que representa.

Las investigaciones comentadas en los Capítulos 4 y 10 ilustran la aplicación de esta técnica analítica. Razón por la que se remite a su lectura.

- *Análisis de conglomerados*

Una variedad de análisis cuya finalidad principal es la clasificación de un grupo de individuos u objetos en un número reducido de grupos. Estos *grupos* han de ser mutuamente excluyentes; han de estar compuestos por individuos lo más similares posible entre sí y diferentes de los integrantes de otros grupos.

A diferencia del *análisis discriminante*, los *grupos* no se hallan predefinidos. Precisamente se aplica esta técnica analítica para la definición de los grupos. Posteriormente, éstos pueden ser corroborados mediante otras técnicas multivariantes (como el *análisis discriminante* o el *factorial confirmatorio*).

El *grupo* se define en función del algoritmo de clasificación que se emplee en la agrupación de los sujetos. En general, se busca la agrupación que haga máxima la distancia entre las *medias grupales* y mínima la *desviación intragrupal* (de los individuos integrantes del grupo respecto a su *centroide* o *media grupal*).

Para la obtención de estos grupos, o *conglomerados*, puede elegirse entre distintos procedimientos, agrupados en dos amplias categorías: *jerárquicos* y *no jerárquicos*. Los primeros se dividen, a su vez, en *aglomerativos* y *disociativos* (o *divisorios*).

- a) El procedimiento *jerárquico aglomerativo* comienza con tantos grupos como individuos en la *muestra*. Paulatinamente van reduciéndose el número de grupos

y aumentando, en contra, el número de individuos en cada grupo. El programa finaliza cuando un único *conglomerado* agrupa a todos los individuos.

- b) El procedimiento *jerárquico disociativo* (o *divisorio*) procede a la inversa. Parte de un único grupo, que integra a todos los individuos. Poco a poco, este grupo va seccionándose en diferentes subgrupos hasta que, al final, existen tantos grupos como individuos.

En ambos procedimientos *jerárquicos* de formación grupal, la elección del número de *conglomerados* se realiza conforme a una variedad de criterios. Entre ellos destacan: el criterio de la *distancia mínima*; de la *distancia máxima*; el *promedio de las distancias*; el *método Ward*; o el de los *centroides*. Mientras que los primeros criterios consideran la distancia de los individuos, el último tiene en cuenta la distancia entre las *medias (centroides)* de las variables. En cualquiera de estos procedimientos *jerárquicos* los datos pueden visualizarse mediante un *dendograma*: una representación gráfica en forma de árbol.

En los procedimientos de agrupación *no jerárquica* se parte, a diferencia de los anteriores, de una especificación previa de los grupos que desean formarse. De lo que se trata es de encontrar representantes para cada uno de los grupos. Un individuo pertenecerá al grupo cuya distancia a su *centroide* sea menor.

Una vez que los grupos se han constituido (por cualquiera de los procedimientos referidos), se procede a la comparación de las *varianzas* de las variables en relación a los grupos (*intergrupo e intragrupo*). Ello permite comprobar si los grupos presentan diferencias en los valores de las variables consideradas en el análisis.

En función del valor del estadístico *F* (en cada una de las variables), se rechazan aquellas variables que no diferencian a los grupos. A partir del valor de la *media* de las variables se extrae, por último, los *rangos* que más separan a los integrantes de cada grupo.

• Escalamiento multidimensional

Una variedad analítica análoga al *análisis factorial*. Con él comparte un mismo objetivo: la obtención de un número reducido de *dimensiones*, que permitan caracterizar a determinados objetos o sujetos. Si bien difiere (entre otros aspectos) en el número de *dimensiones* a obtener. Mientras que el *análisis factorial* no impone restricciones al respecto (de hecho el análisis puede efectuarse con un número elevado de *dimensiones* o *factores*), el *escalamiento multidimensional* aconseja su reducción al menor número posible. Ello responde a condicionamientos impuestos para la representación gráfica de los resultados de la investigación.

En el *escalamiento multidimensional* los datos se representan como puntos en un espacio “multidimensional”. La distancia habida entre ellos se considera una expresión gráfica de su semejanza o disimilaridad. Esta representación se hace más fac-

tible e interpretable cuando la información puede reducirse a dos o tres *dimensiones*, como máximo.

Para la concreción de estas *dimensiones* se acude a uno (o varios) *criterios de bondad de ajuste*: el “stress” de Kruskal, el “s-stress” de Young, la *correlación múltiple al cuadrado*, o el *diagrama de Shepard*.

Para que la configuración de las *dimensiones* presente un *ajuste* adecuado, los valores correspondientes a “stress” y “s-stress” han de ser bajos. En cambio, la *correlación múltiple al cuadrado* ha de ser elevada. Su valor se interpreta como proporción de *varianza* explicada por las distancias respectivas. Por su parte, el *diagrama de Shepard* ha de reflejar una tendencia ascendente (o creciente).

Dependiendo del nivel de medición de las variables, se diferencia entre *escalamiento multidimensional “métrico”* y “no métrico”.

a) *Métrico*, cuando las variables son de *intervalo* o de *razón*. Lo que favorece su configuración en una escala *continua*.

A partir de una *matriz de correlaciones*, o de *distancias* entre objetos, se trata de situar a éstos en un espacio *multidimensional*. Para ello se transforma la similaridad o disimilaridad percibida en ellos, en distancias (*euclidianas*).

b) *No métrico*, si las variables son *cualitativas (nominales u ordinales)*. En este caso, se parte de una *matriz de rangos*. De ésta se obtiene información de la similaridad de los objetos. La configuración final será aquella que mejor represente a los *rangos* de la *matriz* inicial.

Como en el *análisis factorial*, el *escalamiento multidimensional* puede ser *exploratorio y/o confirmatorio*. Depende de la finalidad del mismo.

• Modelos “log-linear”

También conocidos como modelos “lineales logarítmicos” o modelos “log-lineales”. Su denominación deriva de la transformación logarítmica operada en los datos (las frecuencias observadas en las variables) para facilitar su aditividad: la posibilidad de sumar los distintos efectos de un conjunto de variables *no métricas* (a semejanza de las variables *métricas* en el *análisis de regresión*).

Las *tablas de contingencia* resultan muy útiles en el análisis de variables *cualitativas*. Pero, cuando se añaden terceras o cuartas variables, a modo de variables de *control*, la interpretación de la *tabla* se complica. En estos casos, sobre todo conforme aumenta el número de variables, se aconseja la aplicación de *modelos log-linear*. Éstos se muestran adecuados para el análisis de las interrelaciones entre una serie de variables *no métricas*, que conforman una *tabla de contingencia multidimensional*.

Las casillas de la *tabla* se traducen a componentes, denominados *parámetros lambda*. Estos miden el efecto de los valores de las variables que conforman cada casilla. Si

el valor Z correspondiente a cada λ es ≥ 1.96 , el *parámetro lambda* se considera relevante en la interpretación del modelo. De él se pretende obtener la probabilidad de que un individuo concreto comparta una combinación específica de atributos de un conjunto de variables.

Para la comprobación del *ajuste del modelo* se acude a los estadísticos X^2 y la *razón de verosimilitud*. Los valores de ambos estadísticos tienden a coincidir conforme aumenta el tamaño de la muestra.

- *Análisis de correspondencias*

Una de las *técnicas de interdependencia* de más reciente desarrollo, también adecuada al análisis de variables *cualitativas*.

Parte de la configuración de las variables en una *tabla de contingencia*. Su objetivo fundamental es la representación de las distancias de las *filas* y las *columnas*, que integran la *tabla*, en unos *ejes cartesianos*. Para ello se transforma el valor de la X^2 en una medida métrica de distancia.

Como en el *escalamiento multidimensional*, la proximidad de los puntos mide la similitud existente entre ellos. En el *análisis de correspondencias*, la proximidad muestra, concretamente, la asociación entre las categorías de las variables.

Primero se procede (como en el *análisis factorial* y el *escalamiento multidimensional*) a la identificación del número adecuado de *dimensiones* que categoricen al objeto de estudio. A tal fin, se examina, igualmente, el porcentaje acumulado de *varianza* explicada.

Para cada *dimensión* se deriva, a continuación, unos *autovalores* (“*eigenvalues*”). Éstos expresan la contribución relativa de cada *dimensión* en la explicación de las *varianzas* de las variables.

Una vez establecida la *dimensionalidad*, se identifica la *asociación* o relación existente entre las categorías de las variables mediante su *proximidad*. Ésta ha de comprobarse o en las *filas* o en las *columnas*. Depende de su ubicación. Un valor elevado de X^2 indica un fuerte grado de “correspondencia” entre los atributos de las variables.

De esta forma se obtiene la reducción dimensional de las proporciones de objetos en una serie de atributos. Al mismo tiempo, se extrae la representación de los objetos, relacionados con esos atributos.

9.5. Paquetes estadísticos disponibles

Sin la mediación del ordenador no sería viable ninguno de los análisis estadísticos reseñados. Especialmente, los análisis *multivariantes*.

La gran revolución experimentada en los últimos años en el campo de la informática hace que la información que se dé quede obsoleta en un breve período de tiempo.

po. Por esta razón, únicamente van a nombrarse algunos de los principales *paquetes estadísticos*, para conocimiento del lector.

La distinción entre los *paquetes estadísticos* responde a su especificidad. Se diferencia entre programas “genéricos” (que ejecutan la generalidad de los análisis estadísticos), y los “específicos” (especializados en técnicas analíticas concretas).

Entre los *paquetes estadísticos genéricos* destacan el SPSS, BMDP, SAS, ESP, y OSIRIS. En sus manuales respectivos se detallan tanto aspectos técnicos (de funcionamiento del programa), como analíticos (para la interpretación de los análisis estadísticos). Dicha interpretación será la misma, independientemente del programa utilizado.

Los *paquetes estadísticos específicos* son numerosos y variados. A modo de ejemplo se señalan los siguientes: LISREL y EQS (para el análisis de *ecuaciones estructurales*), LIMDEP (análisis *logit*), INDSCAL (*escalamiento multidimensional*), ECTA (*tablas de contingencia*), SPAD (análisis de *correspondencias*), o BROCOLI (*series temporales*).

A estos paquetes estadísticos se suman otros específicos para la realización de *gráficos* (como el SYSTAT, STATGRAPHICS, o el HARVARD GRAPHICS). También hay que mencionar las *bases de datos numéricas* (como DBASE o SIR-DB). Estas bases están diseñadas para la exportación automática de datos en la mayoría de los paquetes estadísticos (en especial, los *genéricos*).

Finalmente, cabe mencionar que en los últimos años comienzan a comercializarse *sistemas expertos en análisis de datos* (como el GLIM). Éstos funcionan a modo de consultor experto en *técnicas cuantitativas de análisis*.

Lecturas complementarias

- Bisquerra, R. (1989): *Introducción conceptual al análisis multivariante*. Barcelona, PPU.
- Garrido Luque, A.; J. L. Álvaro Estramiana (1995): *Técnicas de análisis estadístico en ciencias sociales*. Madrid, Universidad Complutense.
- Hair, J. et al. (1992): *Multivariate data analysis*. New York, McMillan.
- Sánchez Carrión, J. J. (comp.) (1984): *Introducción a las técnicas de análisis multivariable aplicadas a las ciencias sociales*. Madrid, CIS.
- Sánchez Carrión, J. J. (1989): *Análisis de tablas de contingencia: el uso de los porcentajes en las ciencias sociales*. Madrid, CIS, *Monografía* n.º 105.
- Sánchez Carrión, J. J. (1995): *Manual de análisis de datos*. Madrid, Alianza Universidad.
- Sánchez Carrión, J. J. y Torcal, M. (1992): *Utilidades del SPSS/PC+. Presentación de informes, grabación de datos y creación de gráficos y mapas*. Madrid, Alianza.
- Spiegel, M. (1991): *Estadística. Teoría y problemas resueltos*. Madrid, McGraw Hill.



EL ANÁLISIS DE CONTENIDO CUANTITATIVO

En el análisis de *documentos* (y, en general, de cualquiera de las modalidades de comunicación verbal, e incluso visual) adquiere un destacado protagonismo el *análisis de contenido*. Aunque algunos autores (Krippendorf, 1980, 1990; y algunos de sus seguidores: López Aranguren (1989), O'Connell y Layder (1994), entre otros) lo cataloguen como “técnica de investigación”, considero que las características del *análisis de contenido* corresponden más bien a una “técnica de análisis”, de utilidad en cualquiera de las *estrategias de investigación* anteriormente referidas. Desde el *uso de fuentes documentales* y el *estudio de casos*, hasta la propia *encuesta* (en la *codificación de preguntas abiertas*).

La posibilidad de utilización de esta técnica surge siempre que el análisis se centre en el “contenido” *manifiesto* y *latente* de la información “verbal” (extraída de distintas fuentes: documentos, prensa escrita, transcripciones de entrevistas, grabaciones radiofónicas, de programas de televisión) o “visual” (de la observación de imágenes estáticas –fotografías, cuadros– y/o en movimiento –una grabación en vídeo, por ejemplo).

Pero, ¿en qué consiste el *análisis de contenido*?, ¿qué peculiaridades distinguen a esta variedad de análisis?, ¿cuáles son sus límites y posibilidades? A estas y otras preguntas tratará de darse respuesta en las páginas siguientes.

10.1. El análisis de contenido como técnica de análisis

El *análisis de contenido* se perfiló, en sus orígenes, como una técnica analítica eminentemente *cuantitativa*. Berelson (1952) la definió como la descripción *objetiva, sistemática* y *cuantitativa* del *contenido manifiesto* de la comunicación. No obstante, en la actualidad se reconoce la pluralidad analítica existente, de hecho, bajo la rúbrica

común del “análisis de contenido”. Éste no se limita a la “cuantificación” del *contenido manifiesto* de la comunicación; sino que también aborda la interpretación del *contenido latente*.

Tesch (1992) diferencia dos tipos genéricos de *análisis de contenido*: el *clásico* y el *etnográfico*. El primero se ajusta a la definición de Berelson (1952), mientras que el segundo (el *etnográfico*), responde a una redefinición desde posturas *cualitativas* (Altheide, 1987), que pone el acento en el *análisis reflexivo de documentos*. Este segundo tipo se dirige, no tanto a la *descripción*, como a la *comprensión* de los *significados latentes*, y a la *verificación* de relaciones teóricas.

Dada la especificidad del presente manual (sobre la *metodología cuantitativa*), la exposición que aquí se hace del *análisis de contenido* se ciñe a la modalidad *clásica* o *cuantitativa*. Para su complementación con el *análisis de contenido etnográfico*, remito al manual de metodología cualitativa de esta misma colección, y a cualquier texto especializado en el *análisis cualitativo*.

10.1.1. Características esenciales del análisis de contenido cuantitativo

Los tres rasgos definitorios del *análisis de contenido cuantitativo* están ya presentes en la definición clásica de Berelson (1952): *objetividad*, *sistematicidad* y *cuantificación* de los *contenidos manifiestos* de la comunicación. A éstos se suma otro fundamental: la *inferencia* de los “datos” al “contexto” de referencia. De hecho, éste es –siguiendo a Krippendorff (1980, 1990)– el propósito primordial del análisis: realizar, a partir de ciertos datos, *inferencias* válidas y replicables, que puedan aplicarse a su *contexto*. Ello exige que el análisis reúna condiciones mínimas de *validez* y de *fiabilidad*, en el proceso de medición.

En el *análisis de contenido cuantitativo* el énfasis no recae en los aspectos semánticos o sintácticos de los textos, sino en la “cuantificación” de sus integrantes (palabras, expresiones, frases, temas); es decir, en la medición de su “frecuencia” de aparición en el texto.

Esta “cuantificación” permite, precisamente, la comparación de los “contenidos” de diferentes documentos y, sobre todo, la *inferencia*. Ésta sería –de acuerdo con Weber (1994)– difícilmente alcanzable por otros medios. Pero, para que pueda, a partir de los *contenidos “manifiestos”*, inferirse los *contenidos “latentes”* (no explícitos en el documento, y relativos tanto al mensaje, como a sus agentes emisores y receptores), la “cuantificación” ha de hacerse de forma *sistemática* y *objetiva*.

La *objetividad* se halla relacionada con el uso de procedimientos “rigurosos y replicables” (Scott, 1990: 130). No sólo se han de elegir *indicadores* que constituyan una representación “válida” de los *conceptos*. También ha de lograrse la coincidencia en la *codificación* de un mismo documento, por personas diferentes, si se pretende que la *clasificación* efectuada se considere “fiable”.

En suma, el *análisis de contenido cuantitativo* puede caracterizarse como un análisis no limitado a la *descripción*, sino orientado a la *inferencia*. La *descripción* (o enu-

meración de las características del texto) constituye –en conformidad con Bardin (1986)– la primera etapa del análisis; la *interpretación* (la significación acordada de estas características), la última. Entre ambas se encuentra la *inferencia*, como un procedimiento intermedio, que permite el paso de la *descripción* a la *interpretación*.

10.1.2. Fases de un análisis de contenido

Bardin (1986) diferencia tres momentos cronológicos fundamentales en la organización del *análisis de contenido*:

- a) El *preanálisis* o fase de organización, de *operacionalización* de las ideas originarias de la investigación. Incluye:
 - 1) La especificación de los *objetivos* de la investigación y la formulación de las *hipótesis*.
 - 2) Elaboración de los *indicadores* en que se apoyará la interpretación de los datos. Éstos se obtendrán a partir de las *definiciones operativas* de los *conceptos teóricos* básicos en la investigación.
 - 3) Elección de los *documentos* a analizar: la *muestra* del estudio.
- b) *Explotación del material*. Incluye:
 - 1) Elección de las *unidades de codificación* (palabras, expresiones, frases, temas, imágenes...; depende de qué se analice). De ellos, se cuantificará la *frecuencia*, e *intensidad*, de aparición en el documento. También pueden efectuarse enumeraciones de carácter temporal (cuándo) y espacial (dónde aparecen).
Además de las *unidades de codificación*, se identifican las *unidades de contexto*, siguiendo la recomendación de Holsti (1969). Una misma palabra, por ejemplo, puede adoptar significados diferentes, dependiendo del *contexto* donde se ubique. Razón por la que se aconseja su contextualización. Ésta ayudará en la *interpretación* del significado de las *unidades de codificación*.
 - 2) Determinación de las *categorías* a emplear en la *codificación* del “contenido” de los documentos, a partir de las *hipótesis* de la investigación.
Esta fase es exigida, cuando se realiza un *análisis de contenido categorial* o temático. Si bien, ésta no es la única variedad de *análisis de contenido*, aunque sea la más habitual.
 - 3) El *registro* y *cuantificación* de los *datos*, de acuerdo con los *indicadores* elegidos.
- c) *Tratamiento e interpretación de los resultados*. A los datos reunidos se les da tratamiento estadístico, a semejanza de lo descrito en el Capítulo 9. Se comienza con el *análisis univariable*. Si las características de los datos lo permiten, puede concluirse con la aplicación de alguna técnica de análisis multivariable.
Los resultados se someten a *pruebas estadísticas* y a *tests de validez*. En caso afirmativo, el investigador podrá hacer *inferencias* e *interpretaciones* de los

resultados de la investigación. Estas *interpretaciones* pueden actuar como punto de partida en una nueva investigación.

En la Figura 10.1 se resumen las fases principales en un *análisis de contenido cuantitativo*.

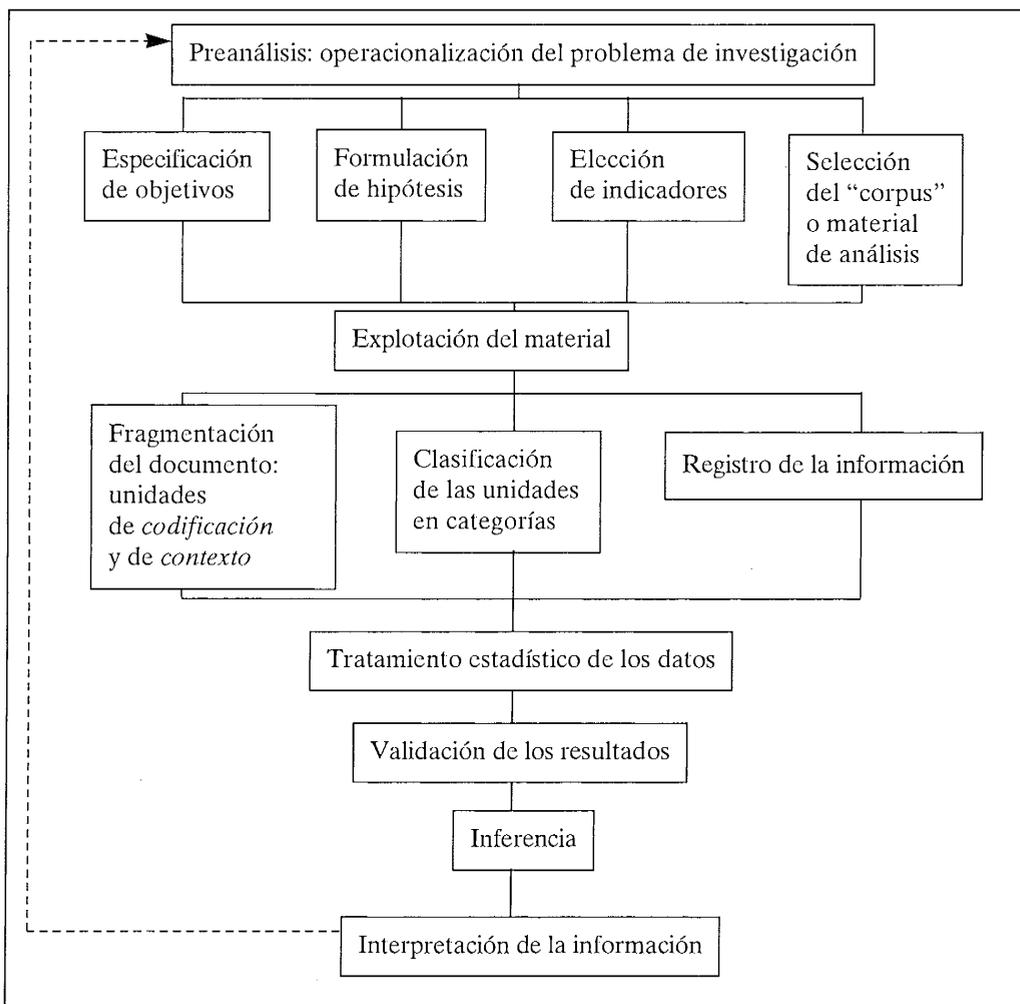


Figura 10.1. Fases principales en la realización de un análisis de contenido cuantitativo.

De las fases representadas en el gráfico, únicamente se comentarán, en continuación, los aspectos más específicos del *análisis de contenido*. Sobre los demás, remito a los capítulos iniciales del libro.

10.2. La selección del “corpus” de documentos o material de análisis

En función de los objetivos de la investigación, se eligen las *unidades de análisis* y la *estrategia* a seguir en la recogida de información.

Para efectuar un *análisis de contenido*, previamente ha de delimitarse el “*corpus*” o material a analizar. Esta delimitación incluye tanto a la *población* de documentos, como la concreción del procedimiento de *muestreo* a seguir en su selección.

Salvo que el *universo* sea de escasa amplitud, lo habitual es proceder a la extracción de una *muestra* representativa de dicho *universo*. En esta selección se siguen los mismos procedimientos muestrales descritos en el Capítulo 5. Concretamente, se consideran las características de la población de documentos (escritos, visuales, etc.), y el tipo de *inferencia* que desea realizarse. En general, se recomienda seguir procedimientos de selección *aleatorios*. Con estos se logra una mayor representatividad de la *muestra*.

La *selección muestral* comienza con la elección de la *fente de comunicación* (periódicos, revistas, libros, canciones, imágenes). Prosigue con la extracción de *documentos* y, en caso de que éstos sean muy extensos, de *partes* de estos documentos.

EJEMPLO DE UNA SELECCIÓN MUESTRAL PARA UN ANÁLISIS DE CONTENIDO

Imaginemos que se desea analizar el tratamiento dado a los temas de corrupción en la prensa española. Uno de los posibles *diseños muestrales* incluiría las *fases* siguientes:

- a) *Elección de fuentes*. Habría que decidir qué periódicos representarán el “*corpus*” de la investigación. El investigador puede optar entre elegirlos “intencionalmente”, o seleccionarlos de forma *aleatoria*, siguiendo cualquiera de los procedimientos expuestos en el Capítulo 5 (desde el *muestreo aleatorio simple*, hasta el de *conglomerados*).

Si se busca una *muestra* representativa de la pluralidad de periódicos nacionales (de ámbito estatal, autonómico, o provincial), puede hacerse un *muestreo aleatorio estratificado* por Comunidades autónomas e ideología de las editoriales de los periódicos.

- b) Seleccionados los periódicos, se procede a la *elección*, en cada uno de ellos, de unos *números* concretos. En esta segunda fase del *diseño muestral*, la selección atiende, habitualmente, a “*fechas*” de publicación. Como indican Sánchez Carrión (1985) y Weber (1985, 1994), los “*contenidos*” de los periódicos difieren por día de la semana (laboral o festivo) y período del año (diferenciando, sobre todo, el período vacacional del resto del año). Estos y otros aspectos deberán tenerse presentes en la selección (intencional o aleatoria) de las *fechas* de edición de los periódicos que forman la *muestra*.
- c) *Extracción de textos en los documentos de la muestra*, si las dimensiones de éstos lo exigen. En este último estadio del *diseño muestral*, la selección suele ser, a diferencia de las anteriores, *intencional*. Se buscan, de forma intencional, “aquellas unidades que tratan sobre el tema objeto de estudio” (Sánchez Carrión, 1985: 96).

Weber (1994) recomienda el análisis del texto completo, siempre que sea posible. Ello ayudará a preservar la coherencia semántica del texto como unidad. Cuando no sea factible el análisis del documento completo, el investigador deberá considerar la *estructura del texto* en la *selección muestral*. En concreto:

- 1) Podrían excluirse las secciones de *introducción* y de *cierre*.
- 2) Cada *muestra* debería constar de uno o más *párrafos* enteros para no alterar la coherencia semántica.
- 3) Las *frases* no deberían muestrearse. Analizar frases separadas (incluso aquellas que se han extraído de un mismo tema), destruye la coherencia semántica. Esto dificulta la *validación* posterior e *interpretación* de los resultados del *análisis de contenido*.

Por último, respecto al *tamaño de la muestra*, éste dependerá, principalmente, de la variabilidad de los textos dentro de un mismo periódico (Sánchez Carrión, 1985). Cuánto más homogéneo sea su *contenido*, menor *tamaño muestral* se precisará para representar a la *población*.

10.3. La explotación del material

A la selección de documentos (*población, muestra*) sigue la determinación de criterios para la extracción de la información en ellos “contenida”. Los *criterios* varían dependiendo, fundamentalmente, de los objetivos de la investigación y de las características de los documentos. No obstante, pueden señalarse unas *pautas comunes* en los *análisis de contenido cuantitativo*. Éstas conciernen a la creación del *esquema de codificación* y la preparación del documento para su posterior tratamiento estadístico. En concreto, se diferencian tres estadios básicos en esta fase esencial al *análisis de contenido*:

- a) La fragmentación del documento en *unidades de codificación* y de *contexto*.
- b) La clasificación de las *unidades* en *categorías*.
- c) El *registro de la información*.

10.3.1. La fragmentación del documento en unidades de codificación y de contexto

Indistintamente de cómo se vaya a analizar la información, una vez elegidos los documentos, éstos han de fragmentarse en *unidades* que ayuden a su organización y análisis.

Las *unidades* se dividen en *unidades de codificación* (o *registro*) y las *unidades de contexto*. Las *unidades de codificación* (o *registro*) se definen como “la mínima porción de contenido que el investigador aísla y separa por aparecer allí uno de los símbolos,

palabras, *slogans* o temas que el investigador considera significativos” (López Aranguren, 1989: 394). Corresponden, por tanto, “al segmento de contenido que será necesario considerar como unidad de base con miras a la categorización y al recuento frecuencial” (Bardin, 1986: 79).

Las *unidades de codificación* que se emplean con mayor frecuencia son las siguientes:

- a) La *palabra*. Ésta constituye la *unidad de codificación* más básica. A ello contribuye su fácil identificación. Si bien, como indica Weber (1994), ha de tenerse presente el problema de los distintos sentidos que puede tener una misma palabra, dependiendo del *contexto* en el que se emplea. No todos los paquetes informáticos disponen de rutinas de desambiguación o de contextualización de palabras específicas.

La *codificación de palabras* puede incluir a todas las comprendidas en el texto, o restringirse, por el contrario, a aquéllas consideradas “claves”, en conformidad con los objetivos de la investigación.

- b) La *frase*, cuando se está interesado en la cuantificación de frases que contienen un mismo significado (por ejemplo, “a favor” o “en contra” de un determinado tema).

Si la *frase* es extensa o compleja, puede decidirse su división en unidades más pequeñas.

- c) El *tema* en torno al cual gira el discurso del documento. Holsti (1969) lo define como la unidad textual que hace referencia a uno de los siguientes aspectos: quién percibe; el agente de la acción; la misma acción; y el objeto de la acción.

Por ejemplo, el *tema* representa una *unidad de codificación* habitual en estudios de motivaciones, opiniones, actitudes, valores, o creencias (Bardin, 1986).

- d) El *párrafo*. La *codificación de párrafos* enteros es practicada, sobre todo, en el *análisis de contenido manual*, cuando se carece de la ayuda de programas informáticos al efecto.

Weber (1994: 265) descarta esta cuarta *unidad de codificación* por su baja fiabilidad. La experiencia de este autor le lleva a afirmar que resulta “más difícil alcanzar una fiabilidad elevada cuando se codifican unidades grandes (como párrafos), que cuando se codifican unidades más pequeñas como palabras”.

- e) El *texto entero*, si éste puede codificarse en su globalidad. Ello exige que el *texto* sea *breve* (titulares de periódicos, editoriales, o las respuestas a una *pregunta abierta*, por ejemplo). En caso contrario, será “difícil alcanzar una elevada fiabilidad cuando se codifican textos completos” (Weber, 1994: 265). A ello se suma otra condición importante: la idea dominante o principal ha de ser suficiente para el objetivo buscado (Bardin, 1986: 81).
- f) El *personaje* o actor al que se le atribuyen las acciones que figuran en los documentos analizados.

Ésta constituye una *unidad de codificación* adecuada “siempre que el texto narre una acción” (Sánchez Carrión, 1985: 97). Bardin (1986) cita como ejemplos, las obras de ficción (films, emisiones, novelas...), que pueden analizarse siguiendo a sus personajes.

De estas *unidades de codificación*, el investigador escogerá una o varias. La elección dependerá de la conjunción de tres *aspectos* básicos:

- a) Los objetivos de la investigación.
- b) Las características del documento.
- c) Los medios disponibles para la realización del *análisis de contenido* (manual o mediante ordenador).

Además de las *unidades de codificación*, en los documentos también se diferencian las *unidades de contexto*. Éstas son de gran utilidad para la comprensión de las *unidades de codificación*.

Bardin (1986: 82) define las *unidades de contexto* como “el segmento del mensaje cuyo tamaño (superior a la unidad de registro) es óptimo para captar la significación exacta de la unidad de registro”. Si la *unidad de registro* fuera, por ejemplo, la *palabra*, la *unidad de contexto* sería la *frase*; del *tema*, el *párrafo*; del *párrafo*, el *documento entero*.

La *unidad de contexto* hace, por tanto, referencia al “contexto” donde se ubica la *unidad de codificación*. Su identificación resulta, en consecuencia, básica para la comprensión del *significado* de la *palabra*, la *frase*, o cualquiera que sea la *unidad de codificación* elegida.

10.3.2. La clasificación de las unidades en categorías

En el *análisis de contenido* “categorial”, las *unidades de registro* se codifican en *categorías*, que singularicen su “contenido”. Estas *categorías* proporcionan una denominación genérica a unidades de contenido similar, contribuyendo, de esta forma, a su clasificación. Para ello, las *categorías* han de cumplir los tres *requisitos* –referidos en el Capítulo 4, apartado 4.2– comunes en la medición de variables: *exhaustividad*, *exclusividad* y *precisión*.

No sólo han de cubrir toda la variedad del objeto de estudio. También han de evitar la “ambigüedad” en el proceso de medición, si se quiere que el *análisis de contenido* reúna condiciones de *validez* y de *fiabilidad*. Quiere esto decir que cada *unidad de registro* sólo puede clasificarse en una única *categoría*.

A esos tres *requisitos* básicos se añaden otros dos fundamentales:

- a) La *pertinencia* o adecuación de las *categorías* a los *objetivos* de la investigación.
- b) La *homogeneidad*: la creación de *categorías* ha de responder a un único *principio de clasificación*.

López Aranguren (1989: 398-401) diferencia los siguientes *principios de clasificación*. De ellos derivan la mayoría de las *categorías* empleadas en el análisis de material escrito:

- a) *Materia, tema o asunto*: ¿de qué trata la comunicación?
- b) *Dirección o punto de vista*: el tratamiento o presentación (favorable-desfavorable-neutral; positiva-negativa) que la comunicación hace de la materia o asunto del que trata.
- c) *Criterio o valor*: sobre qué se basa la clasificación en términos de “dirección” o punto de vista. Por ejemplo, “moralidad-inmoralidad”, “fuerza-debilidad”.
- d) *Autoridad, fuente o prueba*: personas, grupos, instituciones u objetos que en la comunicación se citan en apoyo de una afirmación, o en cuyo nombre se hacen declaraciones o afirmaciones.
- e) *Objetivo o meta* que se persigue.
- f) *Medio o método* para alcanzar los objetivos.
- g) *Actores*: personas, grupos o instituciones que en el texto se representan como realizadores de determinados actos. Por ejemplo, países, grupos religiosos, económicos, etc.
- h) *Características personales y rasgos de carácter*: edad, sexo, optimismo, depresión, entre otras.
- i) *Origen*: lugares de origen de las comunicaciones.
- j) *Destino*, cuando la comunicación no se dirige al público en general, sino a grupos específicos (como estudiantes, jubilados, mujeres, por ejemplo).
- k) *Forma de la afirmación*: hechos, expectativas, preferencias, demandas, ...
- l) *Intensidad*, cuando se clasifica el componente emocional o sentimental de las clasificaciones, el entusiasmo, la agitación o la excitación.
- m) *Mecanismo o recurso estilístico o retórico*, que se emplea con fines propagandísticos, persuasivos o de tendencias.

EJEMPLO DE CLASIFICACIÓN DE UNIDADES EN CATEGORÍAS

Como ilustración de lo expuesto, a continuación se va a exponer el *sistema de categorías* desarrollado en una investigación real, en la que se aplica el *análisis de contenido*. Se trata del estudio llevado a cabo por M. Clemente y M. A. Vidal, publicado en 1994 con el título “La violencia simbólica: la televisión como medio generador de la delincuencia” (en la revista *Apuntes de Psicología*, n.º 41-42, pp. 47-60).

En esta investigación, los autores organizaron la información extraída de series de dibujos animados –emitidos en cadenas públicas (TVE1, TVE2, Telemadrid) y privadas (Tele 5, Antena 3 y Canal Plus)– en un *sistema de categorías* diseñado al efecto. Este *sistema* se confecciona conforme a tres *criterios de clasificación*:

- a) Quién realiza el acto agresivo.
- b) Sobre quién se comete el acto agresivo.
- c) De qué acto agresivo se trata.

En los dos criterios de clasificación primeros, los autores diferencian las mismas siete agrupaciones de categorías:

<p>a. Sexo</p> <ol style="list-style-type: none"> 1. Varón 2. Mujer 3. Sin determinar 	<p>e. Individual o grupal</p> <ol style="list-style-type: none"> 1. Uno sólo 2. Grupo 3. No está
<p>b. Edad</p> <ol style="list-style-type: none"> 1. Niños 2. Adolescentes 3. Adultos 4. Sin especificar 	<p>f. Rol</p> <ol style="list-style-type: none"> 1. Protagonista 2. Habitual 3. Otros
<p>c. Apariencia</p> <ol style="list-style-type: none"> 1. Humano 2. Animal 3. Planta 4. Objeto 	<p>g. Calificación del rol</p> <ol style="list-style-type: none"> 1. Bueno 2. Malo 3. Sin especificar
<p>d. Capacidad de decisión individual</p> <ol style="list-style-type: none"> 1. La tiene 2. No la tiene 	

En cambio, en la especificación de “qué acto agresivo se realiza”, los autores distinguen tres agrupaciones de categorías:

<p>a. Tipos de agresiones</p> <ol style="list-style-type: none"> 1. No verbales 2. Verbales leves o actos irónicos 3. Verbales graves o fuertes 4. Amenazas 5. Agresiones físicas leves (una sola) 6. Agresiones físicas fuertes (más de una) 7. Agresiones con objetos no destinados a agredir 8. Armas blancas 9. Armas 	<p>b. Grado de proactividad del acto</p> <ol style="list-style-type: none"> 1. Proactivo 2. Reactivo
	<p>c. Grado de daño manifiesto</p> <ol style="list-style-type: none"> 1. No aparece 2. Leve 3. Fuerte 4. Muerte

La información analizada corresponde a series de dibujos animados emitidos desde el 29 de noviembre al 5 de diciembre de 1993, ambos días inclusive (tanto en días de diario, como en fines de semana). En conformidad con el sistema de categorías señalado, la información se cuantificó en forma de frecuencia (absoluta y porcentual) de aparición. A modo de ejemplo, en el Cuadro 10.A se recoge la tabulación efectuada para los primeros cinco sistemas de categorías.

CUADRO 10.A. Distribuciones de frecuencias para cinco sistemas de categorías.

Categoría		Características del agresor		Características del agredido	
		Frecuencia	Porcentaje	Frecuencia	Porcentaje
Sexo	1	382	64	257	43
	2	86	14	168	28
	3	131	22	174	29
Edad	1	30	4	83	14
	2	76	11	82	14
	3	271	51	230	38
	4	291	44	206	34
Apariencia	1	287	49	226	38
	2	192	33	235	39
	3	28	5	23	4
	4	79	13	111	19
Capacidad de decisión	1	461	78	387	79
	2	132	22	106	21
Grupo	1	494	83	410	68
	2	93	16	190	31
	3	4	1	8	1

Una vez diseñado el *esquema de codificación*, se recomienda (véase Bardin, 1986; o Weber, 1994) su comprobación en una pequeña *muestra* (de la misma población de documentos), previa a su utilización definitiva para el registro de la información. Esta pre-comprobación ayudará a precisar la adecuación del *esquema de codificación* elaborado para la consecución de los objetivos de la investigación (la clasificación de las *unidades de registro*, de *contexto*, y el *sistema de categorías* creado).

10.3.3. El uso de programas informáticos para el análisis textual

Lo hasta ahora expuesto constituye el procedimiento usual, cuando el *análisis de contenido* se realiza manualmente. Sin embargo, como ya señalaba Mochmann en 1985, cada vez van adquiriendo mayor protagonismo los “procedimientos de *clasificación empírica*”. En ellos, el investigador no define, *a priori*, los *conceptos teóricos*. Por el contrario, éstos resultan de la realización de procedimientos estadísticos (como el *análisis factorial* o el *análisis de conglomerados*), a partir de las *matrices de correlaciones* de palabras incluidas en los textos.

Como ilustración de este proceder, se hace mención del programa WORDS, diseñado con la finalidad de excluir cualquier influencia del investigador en la *categorización* de los textos. A tal fin, se siguen las *fases* siguientes:

- a) El documento se divide en *segmentos* (páginas, párrafos, entrevistas, por ejemplo).
- b) Se suprimen las *palabras funcionales* (artículos, conjunciones).
- c) Se determina la *frecuencia* de aparición de cada *palabra* en cada *segmento*.
- d) Se selecciona un *subconjunto* de esas *palabras* para su análisis.
- e) Se calcula una *matriz de correlaciones* de las palabras incluidas en ese *subconjunto*.
- f) A partir de la *matriz* se realiza un *análisis factorial*. De él se obtendrá un número de *factores*, que corresponderán a los principales *temas* de *contenido* del documento analizado (Mochmann, 1985: 37).

Asimismo, la mayoría de los programas informáticos para *análisis de contenido* no sólo proporcionan *listados de palabras*, acompañados con su *frecuencia* de aparición en el texto. También pueden proporcionar el *contexto* que arroja a la *unidad de codificación*. Esto adquiere especial relevancia en casos de ambigüedad, cuando la *palabra* presenta distintas connotaciones en *contextos* diferentes. Para su correcta clasificación en *categorías*, se crean, entonces, *índices* o *diccionarios generales*. Éstos actúan a modo de *thesaurus*, de diccionarios, que recogen los posibles sentidos de las *palabras*, en función del *contexto* en que se hallen.

El *General Inquirer* (el primero de los programas para *análisis de contenido*, desarrollado inicialmente por Stone y colaboradores, en 1966), por ejemplo, incluye varios *diccionarios generales*, especializados en áreas de conocimiento. Entre ellos se encuentran los siguientes: “Harvard Psychological Dictionary”, “Stanford Political Dictionary”, “Need Achievement Dictionary”, “Santa Fe Anthropological Dictionary”, o “Laswell Value Dictionary” (LVD).

Estos *diccionarios* proporcionan al investigador numerosas *categorías* (generalmente de 60 a 150, e incluso más), así como definiciones o reglas para la correcta clasificación de *palabras* en *categorías*. A tal fin, se considera el sentido que adopte en el *contexto* en que aparece.

Las *categorías* se representan mediante abreviaciones o *etiquetas*, que figuran en tablas. El número al final de la *palabra* indica el número de sentidos particulares de las palabras. De esta forma, la creación de *diccionarios generales* ayuda a la sistematización y estandarización del procedimiento de clasificación de las *palabras* en *categorías*.

EJEMPLO DE DICCIONARIO GENERAL

La presente ilustración se ha tomado de la obra de Bardin (1986: 100-101). Corresponde al *Stanford Political Dictionary* (de Holsti), para el análisis de documentos políticos. En él pue-

den registrarse hasta un total de 4.000 palabras, aproximadamente, según seis o cuatro conceptos clave. Por ejemplo:

<i>Conceptos clave</i>	<i>Número de palabras</i>	<i>Ejemplos de palabras</i>
Afecto positivo	977	Mutuo, natural, normal, puro.
Afecto negativo	1.513	Escándalo, profano, rehusar, repugnante.
Fuerza	1.391	Aceros, piedra, espada, duro.
Debilidad	579	Delgado, ignorante, debajo.
Actividad	1.218	Reacción, reino, contacto, viaje.
Pasividad	722	Inmutable, espera.
Sobreestimación	128	Absolutamente, exactamente, siempre.
Subestimación	50	Aunque, aparentemente, aproximadamente
Negación	6	Diferente, ni uno ni otro.

Además de los *diccionarios*, la mayoría de los programas incluyen el procedimiento KWIC (*Key Words In Context*), que relaciona las *palabras* con el *contexto*. Por ejemplo, el programa TEXTPACK (de Mohler y Zuell, 1990, ZUMA, reléase 4.0) incluye seis operaciones básicas, entre las cuales se incluye el procedimiento KWIC y la elaboración de *índices* (o *diccionarios*). Las operaciones que incluye este paquete son:

- a) *Frecuencia de palabras* diferentes para uno o más textos (procedimiento *FREQ*).
- b) *Índice*: un listado de palabras y sus referencias en el texto (procedimiento *XREFF*).
- c) *Palabras claves en su contexto* (*concordancia*), mediante el procedimiento *KWIC*.
- d) *Comparación de vocabulario*, mediante el procedimiento *WORDCOMP* (*Word Comparison*). Este procedimiento está diseñado para la comparación de palabras con una denominación común o diferente.
- e) *Selección de textos utilizando información externa* al texto, en forma de fichero numérico (un fichero *SPSS*, por ejemplo).

En estos casos puede aplicarse el programa *SUBSEL*. Éste permite la selección de un texto, que corresponde a una combinación específica o valores en un fichero numérico.

- f) *Codificar/categorizar palabras* que figuran en secciones concretas del texto, y su posterior inclusión en un fichero numérico (*TAGCODER*).

A partir de esta información podrán después efectuarse análisis estadísticos (mediante paquetes estadísticos como el *SPSS*, *BMDP* o *SAS*, por ejem-

plo). Para ello se utilizarán no *palabras*, sino *grupos de palabras* clasificadas en *categorías*, en función de su significado común.

Otros programas que tan sólo se mencionan aquí son: EVA (creado para el análisis de los titulares de los diarios), ANACONDA (para la codificación de las *preguntas abiertas* de los *cuestionarios*), SPENCE (para el análisis de los *protocolos* de las *entrevistas* psiquiátricas), y el QUESTER (dirigido al análisis del *contexto*) (Mochmann, 1985).

A estos programas hay que sumar los desarrollados en fechas más recientes. Como son el LEXICLOUD (por el Laboratorio de Lexicometría y Textos Políticos de la Escuela Normal Superior de Fontenay-St.Cloud, para el establecimiento de relaciones estadísticas entre las unidades léxicas), y el sistema ALCESTE (creado por Reinert en 1990, para el análisis léxico por contexto de un conjunto de segmentos de texto). El detalle de estos programas puede encontrarse en el artículo de Julia Behar (1993).

La ayuda de estos u otros programas facilita, considerablemente, la práctica del *análisis de contenido*. Entre sus *ventajas* destacan las siguientes:

- a) Pueden analizarse, simultáneamente, textos complejos, y de considerable magnitud, en un breve período de tiempo.
- b) Se introduce una mayor estandarización y sistematización en el análisis, que mediante procedimientos manuales. Ello facilita la comparabilidad de los resultados analíticos.
- c) Los errores de *codificación* se reducen considerablemente, al no depender ésta de la actuación del equipo investigador.

A estas *ventajas* hay que añadir la reducción en costes (económicos y temporales) que supone la aplicación de *scanners* (con reconocimiento de caracteres) para la grabación óptica de textos en el ordenador. Atrás ha quedado el tedioso procedimiento primigenio de transformación del texto en *tarjetas perforadas*, para su posterior lectura por el programa informático. Además, el uso de *scanners* presenta otra ventaja importante: la reducción de los *errores de grabación*. Los textos se introducen en el ordenador tal como aparecen en el original.

10.4. El tratamiento estadístico y la interpretación de los resultados

En el *análisis de contenido cuantitativo* se aplican las mismas *técnicas estadísticas* descritas en el Capítulo 9. A la *segmentación* de los textos sigue el *listado de frecuencias* de aparición de *palabras*, *frases*, o cualquiera que sea la *unidad de codificación* escogida en la investigación.

En este estadio inicial (a veces único), las *listas KWIC de palabras claves* son de gran utilidad. Estas listas muestran el *contexto* más inmediato de la *palabra*, lo que ayudará en la interpretación de su significado.

Las palabras que únicamente sirven para articular el texto (artículos, pronombres, conjunciones, preposiciones) se excluyen de los análisis. El interés, en cambio, se dirige a la cuantificación de palabras “claves”; aquellas que conforman el léxico del discurso (sustantivos, adjetivos, verbos, adverbios).

Pero, como reconoce Weber (1994: 294):

“Aunque las listas de frecuencias de palabras revelan cambios o diferencias en el énfasis entre los documentos, deben utilizarse con precaución. Las frecuencias de palabras no revelan mucho sobre la asociación entre las palabras.”

Cuando el interés del investigador va más allá de la mera descripción frecuencial de las unidades de codificación por separado, puede aplicarse cualquiera de las *técnicas multivariadas* descritas en el Capítulo 9. Lo que normalmente exige la previa codificación numérica del *contenido manifiesto* del material verbal.

Una de las *técnicas multivariadas* más aplicadas en el *análisis de contenido* es el *análisis factorial*. En éste, las *variables observadas* suelen ser ahora las *categorías* diferenciadas. Las *variables latentes*, por el contrario, hacen referencia a *temas* genéricos, “latentes” en el texto e, igualmente, deducidos de las *variables observadas* que conforman cada *factor* obtenido. De este modo, como en cualquier *análisis factorial*, las *categorías* que correlacionan con un mismo *factor* (es decir, aquellas que presenten un “*factor loading*” superior a .30) se interpretarán como representando un mismo *tema* en el texto analizado.

Los resultados numéricos se *validan e interpretan* en relación con el *tema* al que hacen referencia. Los resultados pueden corroborar las *hipótesis* de la investigación o, por el contrario, contradecirlas. Lo que puede provocar la realización de una nueva investigación.

Como ilustración de lo aquí expuesto, se han seleccionado dos investigaciones reales, en las cuales el *análisis de contenido cuantitativo* se ha materializado de forma distinta. La primera investigación ejemplifica la aplicación del *análisis factorial*; mientras que en la segunda, el análisis se restringe al recuento de *frecuencias*.

10.4.1. Ejemplos desarrollados de análisis de contenido en dos investigaciones reales

- Cabrera, J. (1992): *La nación como discurso. El caso gallego*, Madrid, CIS, *Monografía*, n.º 126.

En esta investigación se analizan distintas manifestaciones del sistema ideológico nacionalista:

- a) El nivel *léxico* y *semántico*, que constituyen las estructuras superficiales del campo de la manifestación.
- b) El nivel *semiológico*, que trata de desentrañar la significación profunda de dicho sistema.

Para cada nivel se ha optado por una *técnica de análisis* diferente, combinándose técnicas *cuantitativas* (*lexicometría*) y *cualitativas* (análisis *semántico-categorial* y análisis *semiológico*).

El *corpus* que se somete al *análisis lexicométrico* (del vocabulario que conforma el discurso) está integrado por las respuestas obtenidas mediante *entrevista abierta*. El *guión* de la *entrevista* comprende tres aspectos concretos:

- a) La biografía de cada entrevistado.
- b) El proceso de socialización nacionalista.
- c) La posición teórico-nacionalista de los entrevistados.

La *palabra* se toma como *unidad de medida*, procediéndose a su “*lematización*”, que comprende:

- a) La *reagrupación de las formas heterogéneas de un mismo vocablo*. Consistió en reducir las formas plurales a singulares, los femeninos a masculinos, y las diversas formas de un verbo a su expresión en “*infinitivo*” (a excepción de los “*participios*”. Éstos se analizaron contextualizadamente, con el propósito de determinar si constituían un adjetivo).
- b) La *separación de las formas homógrafas que muestran vocablos diferentes*. Se analizaron formas con igual grafía que hacían referencia a contenidos diferentes. Estos *homógrafos* fueron identificados con un dígito, o un *dígito* y una letra, a modo de *sufijos*.

Por *ejemplo*, la forma adjetiva de “gallego” se distinguió de sus formas sustantivas, quedándose como sigue:

- “Gallego” = adjetivo
- “Gallego2” = sustantivo
- “Gallego2a” = sustantivo referido al idioma

En general, se adoptaron las claves siguientes: “2” para sustantivos; “2a” para sustantivos referidos a idiomas; “3” para adjetivos; “4” para adverbios; y, “5” para aquellos nombres propios de personas o ciudades que podían inducir a confusión.

Finalizada la *lematización*, se efectúa el recuento mecánico de los textos. Para ello se diferencia, a nivel *léxico* (de vocabulario), dos tipos diferentes de *palabras*:

- a) *Palabras de relación o gramaticales*, aquellas que por sí mismas carecen de significado: artículos, pronombres, preposiciones y conjunciones.
- b) *Palabras lexicales* o formas plenas de significado. Éstas conforman el *léxico* del discurso: sustantivos, adjetivos, adverbios y verbos.

La investigación se centra en el “discurso lexical”, en el análisis estadístico de los rasgos de contenido. En el Cuadro 10.2 figura la distribución general de frecuencias para cada uno de los 12 sujetos cuyos discursos forman la *muestra* del estudio. Los diversos “subcorpus” se distribuyen en función de la extensión del:

- a) *Número de ocurrencias (N)*: el número de ítems. Refleja, por tanto, su extensión.
- b) *Vocabulario (V)*: número de formas diferentes empleadas.
- c) *Hapax*: aquellas formas cuya frecuencia de aparición es igual a 1.
- d) *Número de ocurrencias funcionales (FF_(n))*: ítems que se corresponden con el vocabulario funcional (FF_(v)).
- e) *Número de ocurrencias lexicales (FL_(n))*, con la extensión del vocabulario correspondiente.

Cada uno de estos valores figura acompañado de su correspondiente *valor de rango (r)* para cada columna. El *corpus* total (suma de los diversos *subcorpus* correspondientes a cada emisor) tiene una extensión de 133.249 palabras. Éstas se distribuyeron porcentualmente en los diversos *subcorpus* como se refleja en el Cuadro 10.1.

CUADRO 10.1. Distribución general de frecuencias.

	N	r	V	r	Hapax	r	FF _(n)	r	FF _(v)	r	FL _(n)	r	FL _(v)	r
GM	5524	1	1113	1	613	2	3066	1	162	2	2458	1	951	1
VA	6521	2	1151	2	598	1	3766	2	160	1	2755	2	991	2
FE	8378	3	1601	7	875	8	4624	3	175	5	3754	3	1426	7
GS	9593	4	1479	3	762	4	5543	5	183	8	4050	5	1296	3
JB	9605	5	1580	5	821	7	5314	4	177	6	4291	6	1403	5
CH	9763	6	1503	4	771	5	5722	6	171	3	4041	4	1332	4
PA	11802	7	1581	6	698	3	6618	8	174	4	5184	7	1407	6
CC	12098	8	1869	10	959	10	6556	7	189	9	5542	10	1680	10
LS	12655	9	1751	9	924	9	7218	9	190	10	5437	8	1561	9
RA	13359	10	1695	8	792	6	7845	10	178	7	5514	9	1481	8
LF	16794	11	2206	11	1136	11	9166	11	198	12	7628	11	2008	11
BA	17157	12	2284	12	1158	12	9492	12	192	11	7665	12	2092	12
Total	133249				1010		74930				58319			
							%N: 56,23				%N 43,77			

En el Cuadro 10.2 destacan los discursos de los sujetos BA y LF, cuya extensión supera a la del resto de los entrevistados. Ambos acumulan algo más de la cuarta parte de todo el *corpus*. Los discursos de RA, LS, CC y PA forman un segundo bloque. Sus discursos superan el millar de *ítems*, pero no alcanzan a superar el millar quinientos mil. Por lo que, acumulan el 37,5% del total.

Respecto a la *lexicalidad*, en esta investigación se confirma una norma observada en la generalidad de los *análisis de contenido cuantitativo*: las *formas funcionales* (FF) superan el 50% de la extensión general. En este caso, el 56%. En la investigación se analizará, en consecuencia, el 44% restante que conforma el “discurso lexical”: las formas que caracterizan el discurso sobre la nación.

A nivel léxico se diferencia entre:

- a) *Vocabulario original*: toda forma que sólo se encuentra en un emisor de la muestra.
- b) *Vocabulario común*: formas que tienen al menos una ocurrencia en todos los emisores.
- c) *Vocabulario de base*: formas que tienen una distribución regular en todos los emisores. Por lo que no introducen diferencias.
- d) *Vocabulario específico*: el que caracteriza a cada uno de los emisores.

El subconjunto de formas así obtenido constituye el *corpus* sobre el que se realizan los análisis. El propósito es determinar las especificidades de cada emisor.

Para el cumplimiento de dicho objetivo, se aplica un *análisis factorial de componentes principales* mediante el paquete estadístico BMDP. La *matriz de correlaciones* forman cada uno de los 12 discursos que componen la *muestra* del estudio. Estos *discursos* actúan como *variables* y las 347 *palabras* como *casos*. De lo que se trata es de analizar las interrelaciones existentes entre ellos, en busca de semejanzas y especificidades.

La *matriz de correlaciones* entre las 12 variables se muestra en el Cuadro 10.2.

CUADRO 10.2. Matriz de correlaciones.

	1	2	3	4	5	6	7	8	9	10	11	12
1	1.00											
2	-.04	1.00										
3	-.10	-.06	1.00									
4	-.04	-.07	-.04	1.00								
5	-.12	-.05	-.03	-.03	1.00							
6	-.00	.05	-.07	-.08	-.10	1.00						
7	-.05	-.11	-.06	-.13	-.07	-.05	1.00					
8	-.05	-.10	-.09	-.16	-.12	-.05	.11	1.00				
9	-.14	.00	-.00	-.08	-.13	-.06	-.13	-.11	1.00			
10	-.15	-.10	-.03	.07	.06	-.19	-.25	-.30	.05	1.00		
11	-.04	-.09	-.15	-.10	-.10	-.13	-.20	-.20	-.12	-.08	1.00	
12	-.12	-.06	-.08	-.17	-.10	-.12	-.17	-.25	-.13	-.06	-.08	1.00

La *matriz factorial* de los 4 factores que resultan de la agrupación de las 12 variables, en función de la semejanza de los discursos, se rota por el procedimiento más habitual: *rotación ortogonal varimax*. Se busca la disposición de los *componentes* o *factores* de manera que cada uno de ellos tuviera saturaciones elevadas en pocas variables (en este caso, emisores de los discursos), y muy bajas en el resto.

El Cuadro 10.3 muestra las *cargas factoriales* de cada variable en los 4 factores. Todas las variables presentan, al menos, un “*factor loading*” superior a .25, en alguno de los factores. Como es usual en este tipo de análisis, los factores extraídos aportan de más a menos información: el primer factor explica más *varianza* (14%) que el segundo (11%), y éste más que el tercero (10%). La *varianza* explicada por el tercer factor es ligeramente superior a la explicada por el cuarto. Véase el Cuadro 10.3.

CUADRO 10.3. Matriz factorial rotada oblicuamente mediante el procedimiento varimax (*factor loadings* superiores a .25).

Emisor	Factor 1	Factor 2	Factor 3	Factor 4
GM	–	.517	–	–.394
VA	–	–	.570	–
FE	–	–.432	–	–
GS	–.268	–	–	–.681
JB	–	–	–.476	–
CH	–	–	.629	–
PA	.690	–	–	–
CC	.724	–	–	–
LS	–	–.562	–	–
RA	–.563	–.271	–.395	–
LF	–.253	.604	–	–
BA	–	–	–	–.726
Varianza	13,78%	10,82%	10,33%	10,06%

Para cada *factor* se presentan las formas (o *palabras*) positivas y negativas que lo componen, y sus emisores. En el Cuadro 10.4 se extraen algunas de las *palabras* correspondientes al factor 1. Este factor está caracterizado por los emisores CC, PA y RA, con unos pesos de .724, .690 y .563, respectivamente.

En los *valores positivos* se encuentran la mayoría de las *formas* que caracterizan tanto a CC como a PA. Destacan, especialmente, las correspondientes a los *ámbitos lingüísticos* (“lingua”, “catalán2a”, “catalán”, “portugués2a”, “ortografía”, etc.), con resonancias de elementos discursivos del galleguismo de la época republicana.

El lado *negativo* del primer factor está fuertemente dominado por RA. En él se encuentra un léxico de corte más político, de claras referencias partidistas y electo-

rales (“PSOE”, “CiU”, “voto”, “OTAN”, “votar”, “AP”); junto a un tímido eco del léxico sociológico y de estratificación social (“social”, “clase”, “sector”, “sociedades”). En estos ámbitos el entrevistado RA mostraba una clara especialización.

CUADRO 10.4. Extracto de algunas de las formas positivas y negativas que componen el factor 1.

Factor 1				Emisores
Formas positivas				
lingua	3.490	hispanico	1.540	CC .724
catalán2a	3.100	protagonismo	1.540	PA .690
catalán	2.750	mesmo	1.530	
portugués2a	2.710	galeguista	1.490	
ortografía	2.620	universal	1.470	
mentalidade	2.580	forza	1.460	
Formas negativas				
pensar	-3.928	problema	-1.200	RA -.563
cousa	-2.078	caer	-1.183	
socia	-1.900	formulación	-1.122	
PSOE	-1.397	votar	-1.027	
diferente	-1.315	AP	-1.000	
OTAN	-1.209	traballar	-1.000	

A partir de las *formas positivas y negativas* que configuran cada *factor*, se da a cada uno de ellos las “etiquetas” o denominaciones siguientes:

- Factor 1: cultural-liberal / político-partidista
- Factor 2: sociológico-administración / político– estratificación social
- Factor 3: cultural-identitario / histórico-social
- Factor 4: estratégico-economía política / dispersión político-cultural

Los 4 *factores* se cruzan entre sí (el 1.º con el 2.º; el 1.º y el 3.º; el 1.º y el 4.º, y así), con la finalidad de comprobar la diferenciación léxica entre las distintas generaciones de nacionalistas entrevistados. Cada uno de los cruces se representan gráficamente mediante *planos de proyecciones* de las *palabras* que forman cada *factor*. De esta manera se obtienen las *estructuras léxicas* de las manifestaciones ideológicas.

El primer estadio de la investigación (el correspondiente al *análisis de contenido cuantitativo*) se da por concluido. Para un mayor conocimiento de esta primera etapa de investigación, así como el desarrollo del *análisis cualitativo*, remito a la publicación de esta excelente investigación de Julio Cabrera.

- Llovet, J. J. (1992): “El control de la prensa sobre la profesión médica: el caso de EL PAÍS”, *Revista Española de Investigaciones Sociológicas*, n.º 59, pp. 261-285.

En esta segunda investigación se estudia el control social sobre la medicina a través de la prensa. Para ello se realiza un *análisis de contenido* centrado en el diario *El País*.

La *muestra* abarca las noticias, cartas de lectores, editoriales, etc., publicados en el lustro 1985-1989, referidos a negligencia y conductas médicas incorrectas o lesivas. Se descartaron, por el contrario, las informaciones que aludían a hechos o procesos de este carácter ocurridos fuera de España; al igual que los derivados de la atención psiquiátrica (salud mental), por ser ésta un área muy específica.

En primer lugar se lleva a cabo un rápido peinado visual de los titulares (antetítulos y títulos), aparecidos en las secciones del diario en las que corrientemente pueden leerse comentarios o información sobre el tema. Pocos meses del año quedan en blanco, sin esta clase de información.

En el Cuadro 10.5 se exponen las *distribuciones porcentuales* de las unidades de redacción de *El País*. Éstas se refieren a negligencias y actuaciones incorrectas, o lesivas, de médicos, según la subsección del diario y género periodístico (del 1 de enero de 1985 al 31 de diciembre de 1989).

CUADRO 10.5. Distribuciones porcentuales de unidades de redacción de *El País*.

<i>Frecuencia porcentual</i>	
Subsección del diario	
“Opinión”/“Cartas al director”	22,9
“Sociedad”/“Tribunales”	22,2
“Sociedad”/“Sanidad”	21,5
“Sociedad” (sin especificación)	15,6
“España”	7,3
“Madrid”	5,2
“Opinión”/“Editorial”	2,8
“Domingo”	1,0
“Sociedad”/“Ciencia”	0,3
“Contra-tapa”	0,3
“Opinión”/“El Ombudsman”	0,3
“Opinión”/“Fe de errores”	0,3
Total	100 (288)
Género periodístico	
Noticia	70,1
Carta de lector	22,9
Editorial	2,8
Reportaje, informe, reseña	2,8
Artículo o columna firmada	1,0
Errata	0,3
Total	100 (288)

También se calculan las *frecuencias porcentuales* de las noticias, según el papel social o institucional de los enunciadores: jueces o magistrados, médicos denunciados, familiares o amigos de los pacientes perjudicados, pacientes perjudicados, abogados de los pacientes, colegios de médicos, etc.

De toda la información extraída del periódico *El País*, se deduce la contribución de la prensa como factor preponderante en la apertura, esclarecimiento y sensibilización de los ciudadanos, y pacientes, hacia la problemática sanitaria.

Como puede verse, el uso que del *análisis de contenido* se ha hecho en esta segunda investigación es bastante más simple que el efectuado en la primera investigación. Se limita a la descripción de frecuencias y comentario de extractos de textos que aparecieron en prensa durante el período analizado. No obstante, el principal interés en la exposición de esta segunda investigación es ilustrar un modo de hacer *análisis de contenido*, también usual en la investigación social.

10.5. La aplicabilidad del análisis de contenido y los límites a la inferencia

La aplicación del *análisis de contenido* se ha vinculado, tradicionalmente, al estudio de la comunicación. En 1969 Holsti sintetiza su *aplicabilidad* en la resolución de los siguientes interrogantes de la comunicación:

a) *Qué se dice:*

- Describir tendencias y diferencias en el “contenido” de la comunicación.
- Relacionar las características de las fuentes con los mensajes que éstas generan.
- Evaluación del “contenido”.
- Extraer los patrones culturales presentes en el mensaje.

b) *Cómo se dice:*

- Comprobar si en el mensaje se hace uso de técnicas persuasivas.
- Analizar el estilo del texto.

c) *A quién:*

- Qué peculiaridades presentan aquéllos a quienes está destinado el mensaje.

d) *Por qué:*

- Identificar las intenciones y otras características de los emisores del mensaje.

e) *Quién lo dice:*

- Determinar la autoría del emisor.

f) *Con qué efectos:*

- Analizar el flujo de la comunicación.

Sin embargo, la aplicación que del *análisis de contenido* se hace desde la vertiente metodológica cuantitativa se halla condicionada por *limitaciones* importantes:

- a) Aunque la *cuantificación* ayude a la sistematización y estandarización del análisis textual, no siempre se profundiza en los *contenidos no manifiestos* de los mensajes.
- b) La interpretación de las *unidades de codificación* no ha de restringirse al mero recuento de *frecuencias* de aparición. Por el contrario, exige su previa contextualización. Precisamente, porque unas mismas palabras, en contextos diferentes, pueden implicar significados distintos.

Estas limitaciones pueden solventarse complementando el *análisis de contenido* “*categorial*” con el “*contextual*”, el “*cuantitativo*” con el “*cualitativo*” (este último más centrado en la indagación de las *estructuras semánticas* y *semiológicas* de los documentos). Como bien reconocen Ruiz Olabuenaga e Ispizua (1989: 185), “los mejores análisis de contenido son aquellos que utilizan la técnica de la *triangulación*”. Ello permite una mejor captación de su pleno significado, la definición de la situación y del punto de vista del emisor.

10.5.1. Cuestiones de validez y de fiabilidad

La *inferencia* en el *análisis de contenido* está condicionada, como en cualquier procedimiento analítico, a la *validez* y *fiabilidad* de la *medición* aplicada. En este caso, la síntesis hecha de los “contenidos” de los documentos que se analizan.

La *validez* indica la *adecuación* del sistema de clasificación seguido en la medición de los *conceptos* teóricos; mientras que la *fiabilidad* expresa la *consistencia* del procedimiento de medición.

- Respecto a la *fiabilidad*, han de arbitrarse unas reglas de *codificación* que permitan que distintos codificadores alcancen, independientemente unos de otros, los mismos resultados. Esto es más fácilmente alcanzable cuando el *análisis de contenido* se realiza mediante ordenador. Una vez que las *reglas de codificación* se han definido, éstas se aplican automáticamente, y de igual manera en cada uno de los casos analizados.

Krippendorff (1980, 1990) diferencia tres tipos de *fiabilidad* pertinentes en el *análisis de contenido*:

- a) *Estabilidad*: grado en el que varían los resultados de la clasificación a lo largo del tiempo.

Deberían alcanzarse los mismos resultados en la clasificación de las *unidades de codificación*, siempre que se emplee la misma agrupación de *categorías*. Ello exige la eliminación de cualquier ambigüedad en la clasificación de las *unidades* en categorías.

- b) *Reproducibilidad*: grado al que se alcanzan los mismos resultados, cuando un mismo texto se codifica por varias personas.

En este caso, la *infiabilidad* se atribuye, sobre todo, a arbitrariedades en la *codificación*. Estas pueden deberse a ambigüedades en las instrucciones de codificación. Pero, también a errores de grabación o a diferencias cognitivas entre las personas que realizan la *codificación*.

- c) *Adecuación*: grado en el que la clasificación del documento se corresponde con un estándar o norma.

Como rara vez se establecen *codificaciones estándar*, este último criterio de *fiabilidad* apenas se aplica.

Para la cuantificación de la *fiabilidad* se han propuesto algunos *coeficientes*. Sánchez Carrión (1985) destaca los dos siguientes:

- a) La *razón de codificaciones de acuerdo*, entre el total de codificaciones:

$$C.R. = \frac{2M}{N_1 + N_2}$$

donde: "M" corresponde al número de codificaciones en las que coinciden los 2 codificadores.

"N₁" y "N₂", el número de codificaciones efectuadas.

- b) El *índice de fiabilidad* (p_i) de Scott:

$$p_i = \frac{\text{proporción de acuerdo observado} - \text{proporción de acuerdo esperado}}{1 - \text{proporción de acuerdo esperado}}$$

Este índice oscila entre .00 y 1.00. El valor 1.00 representa el acuerdo total entre los codificadores; mientras que .00, el desacuerdo.

La *validez*, en cambio, denota correspondencia entre los *constructos teóricos* y las *variables empíricas* empleadas. Esto exige, igualmente, la eliminación de cualquier ambigüedad en la clasificación de las *unidades de codificación en categorías*.

En concreto, se diferencian varios tipos de *validez* en el *análisis de contenido* (López Aranguren, 1989; Krippendorff, 1990; Weber, 1994):

- a) *Validez pragmática* o de utilidad de los resultados del análisis en el estudio del tema que se investiga.
- b) *Validez de constructo*: correspondencia entre la medida aplicada con alguna otra utilizada del mismo *constructo*.
- c) *Validez de contenido*: grado en el que los *indicadores* cubren la variedad de significados incluidos en el *concepto*.

- d) *Validez predictiva*: grado de acuerdo/correspondencia entre las predicciones obtenidas y los hechos observados.

Las *inferencias* hechas a partir del análisis de la información han de “mostrar un alto grado de acuerdo con los estados, atributos, sucesos o propiedades del contexto de los datos al que esas inferencias se refieren, y un alto grado de desacuerdo con las características contextuales que esas diferencias excluyen lógicamente” (Krippendorff, 1990: 232).

- e) *Validez semántica*: grado en el que las *unidades de codificación* clasificadas juntas poseen connotaciones similares.

En el *análisis de contenido* “se logra una alta validez semántica cuando la semántica del lenguaje de los datos se corresponde con la fuente, el receptor o cualquier otro contexto respecto al que se examinan dichos datos” (Krippendorff, 1990: 231).

Los límites a la *inferencia* se hallan, en conclusión, en la ambigüedad de los significados de las *palabras* y de las *reglas de codificación*. Por lo que se recomienda su explícita delimitación.

Lecturas complementarias

- Bardin, L. (1986): *Análisis de contenido*. Madrid, Akal.
- Behar, J. (1993): “Aproximación al análisis textual informatizado”. *Anuario de Psicología*, n.º 59, pp. 61-78.
- Clemente Díaz, M. (1992): “El análisis de contenido”. En Clemente, M. (comp.): *Psicología Social: métodos y técnicas de investigación*. Madrid, Eudema, pp. 169-207.
- Krippendorff, K. (1990): *Metodología del análisis de contenido: teoría y práctica*. Barcelona, Paidós.
- López-Aranguren, E. (1994): “El análisis de contenido”. En García Ferrando, M. *et al.* (comp.): *El Análisis de la Realidad Social*. Madrid, Alianza, pp. 461-492.
- Mochmann, E. (1985): “Análisis de contenido mediante ordenador aplicado a las ciencias sociales”. *Revista Internacional de Sociología*, vol. 43 (1), pp. 11-44.
- Sánchez Carrión, J. J. (1985): “Técnicas de análisis de los textos mediante codificación manual”, *Revista Internacional de Sociología*, vol. 43 (1), pp. 89-118.
- Weber, R. P. (1994): “Basic content analysis”. En Lewis-Beck, M. S. (comp.) *Research Practice*. California, Sage, pp. 251-338.

EJERCICIOS PROPUESTOS

1. Describa una investigación en la que se haya hecho un análisis de contenido cuantitativo.
2. Analice el contenido de los editoriales del próximo fin de semana que aparezcan en tres periódicos.
3. ¿Qué peculiaridades presentan los criterios de validez y de fiabilidad en el análisis de contenido? Señale semejanzas y diferencias con los enunciados en el Capítulo 4.
4. Un grupo de investigadores desea analizar la influencia de la prensa escrita en el desarrollo de una campaña electoral. Para ello deciden analizar los contenidos de periódicos publicados durante la última campaña electoral. Diseñe la investigación: especifique cómo se seleccionaría la muestra del estudio, cuáles serían las unidades de codificación y de contexto, y trace el plan de análisis.