

individuos, vamos a extraer en la muestra 45 varones del barrio A, es decir el 11,2% de los individuos de la muestra. Esto es exactamente el porcentaje de varones que viven en el barrio A sobre todos los adultos de la ciudad. Se puede fácilmente comprobar que el mismo mecanismo funciona para todos los segmentos en que hemos dividido la población. El resultado es que tenemos una muestra extraída de forma aleatoria que es representativa de la población de la ciudad con respecto al sexo y al barrio de residencia.

Naturalmente, no se puede decir nada de su representatividad con respecto a todas las otras propiedades —es decir, infinitas menos dos. Lo que podemos hacer, si queremos, es segmentar la población teniendo en cuenta también una tercera propiedad (por ejemplo, el nivel de instrucción), e incluso una cuarta (por ejemplo, el grupo de edad). Pero cada vez que se considera una nueva propiedad el número de segmentos a ordenar en secuencia crece en proporción geométrica: si con el sexo y cinco barrios de residencia los segmentos eran 10, considerando cuatro niveles de instrucción ellos serán 40, y considerando también seis grupos de edad, serán 240. Con una población de 80.000 individuos, el tamaño promedio de cada subgrupo (por ejemplo, varones del barrio A con instrucción baja y menores de 25 años) va a ser de 333 individuos. Reduciéndose la diferencia entre el número de individuos que hay en cada segmento (en el ejemplo, 200) y el número de individuos en cada subgrupo, se acrecienta la probabilidad de que la extracción aleatoria sistemática atribuya a un subgrupo un porcentaje de extraídos no exactamente correspondiente a su porcentaje en la población (aun si las diferencias no pueden matemáticamente exceder una unidad en cifras absolutas).

El principal inconveniente de este método es, por lo tanto, la complicación procedimental vinculada con el manejo de subgrupos formados por el producto lógico de categorías en diferentes propiedades. Pero en el caso de que se atribuya gran importancia a la representatividad de la muestra con respecto a pocas propiedades, el método que se ilustró es la manera más simple¹⁰ de conseguirla sin violar la naturaleza aleatoria de la muestra.

¹⁰ Una forma alternativa es el así llamado "muestreo estratificado". Se definen substratos con un criterio cualquiera y luego se sortean submuestras dentro de cada substrato. Por otro lado, a menudo se usa un muestreo estratificado "no proporcional" para sobrerrepresentar o subrepresentar un segmento particular de la población estudiada.

CAPÍTULO 7

CONCEPTOS DE PROPIEDADES. VARIABLES, FIDELIDAD Y FIABILIDAD

7.1. La definición operativa

En las columnas de una matriz de datos se encuentran las variables. Una variable es un vector de signos que representan los estados de los casos en las propiedades que interesan. Habitualmente, pero no necesariamente, estos signos son números.¹

A veces la relación entre un estado en una propiedad y el signo que lo representa es directa e intuitiva: si la entrevistada es Morena Ruiz y tiene 18 años, tal estado en la propiedad "edad" será representado por el mismo 18 en la celda ubicada en el cruce entre el vector-fila relativo a Morena Ruiz y el vector-columna representando la edad. Para saber que este 18 es el número de años de Morena Ruiz, tenemos que saber cuál es el titular del vector-fila y qué propiedad se representa en el vector-columna. Sólo si tenemos estas informaciones, este 18 se convierte de un mero número en un dato —y lo mismo para todos los otros números en la matriz. Pues sólo se puede hablar de una matriz de datos si todos los números (y, más generalmente, los signos) que se ven son interpretados, es decir, son DATOS.

Continuando con Morena Ruiz, es probable que estemos interesados no sólo en su edad, sino también en su título de estudio (en el caso: licenciada). Pero no hay cifras que representen directamente este título (ni los otros títulos).

Por lo tanto, para poner esta información en las celdas relativas necesitamos una convención que conecte el título de estudio "licenciada" a un número particular, y lo mismo para cada otro título que deseamos registrar. Una convención del género se llama PLAN DE CODIFICACIÓN.

Un plan de codificación de la propiedad "título de estudio" podría ser:

¹ Se ponen números para facilitar el análisis estadístico de las relaciones entre variables. Pero este hábito conlleva el riesgo de que se hagan operaciones estadísticas sobre números que sólo son tales en apariencia.

- 0 = ningún título
- 1 = sólo educación primaria
- 2 = sólo educación secundaria
- 3 = licenciatura
- 4 = doctorado
- 9 = información faltante.

Cabe subrayar el hecho de que en cada plan de codificación se necesita prever un número particular para informaciones faltantes. Para esta función de señalar la falta de informaciones se elige habitualmente un número al margen del abanico de los otros: un 0 o un 9 si se usan números de una cifra, un 0 o un 99 si se usan números de dos cifras, y así sucesivamente.

Se puede comentar que el plan de codificación que hemos visto es bastante grosero —en el lenguaje metodológico, poco sensible. La SENSIBILIDAD es un cociente entre el número de estados en una propiedad que consideramos como categorías en nuestro plan de codificación y el número de estados diferentes que tiene la misma propiedad.²

Para acrecentar la sensibilidad podemos considerar una gama más amplia de estados en nuestro plan de codificación; por ejemplo:

- 0 = ningún título
- 1 = sólo educación primaria
- 2 = sólo educación secundaria
- 3 = sólo educación terciaria
- 4 = licenciatura
- 5 = especialización
- 6 = maestría
- 7 = doctorado
- 8 = posdoctorado
- 9 = información faltante.

El plan de codificación es parte de un conjunto de reglas y convenciones³ que permiten convertir una propiedad de los objetos o sujetos en la realidad en una variable en la matriz de datos. Este conjunto se llama DEFINICIÓN OPERATIVA y las partes que lo integran varían según la naturaleza de la propiedad. La variación más considerable está vinculada al tipo de unidad de análisis, y por consiguiente al tipo de técnica de recolección.

² Cuando desconocemos el número de estados diferentes, obviamente el cociente queda indeterminado. Pero el concepto de sensibilidad mantiene su utilidad: sea cual sea este número, agregando una categoría se acrecienta la sensibilidad, y eliminando una categoría la sensibilidad disminuye.

³ Se entiende que estas reglas y convenciones son específicas no sólo de cada propiedad, sino también de cada investigación, en el sentido de que dependen de la elección del o de los investigadores. Como observa Corbetta (2003: 90), en las ciencias físicas "se forman dentro de la comunidad científica convergencias que permiten llegar a definiciones de carácter intersubjetivo". Pero en las ciencias sociales estamos muy alejados de esta situación.

Si queremos recoger informaciones acerca del título de estudio de los entrevistados, la definición operativa incluye:

- 1) el texto de la pregunta relativa (ejemplo: "¿Cuál es su título de estudio?");
- 2) un plan de codificación como el que hemos visto arriba;
- 3) el hecho de que este plan sea directamente planteado al entrevistado junto con la pregunta pidiéndole que elija una de las respuestas alternativas que se le ofrecen, o sólo sea empleado como una lista de voces a las cuales atribuir las respuestas espontáneas de cada entrevistado a la pregunta relativa;
- 4) la indicación del vector-columna en que se deben poner las informaciones acerca del título de cada entrevistado —informaciones que se convirtieron en datos representando cada una de ellas con el correspondiente número del plan de codificación;
- 5) si el investigador sospecha que un número bastante alto de entrevistados puede mentir acerca de su título, y si él dispone de los recursos necesarios (tiempo, fondos, colaboradores confiables), puede también insertar en la definición operativa un procedimiento con el que se controlará la sinceridad de las declaraciones de los entrevistados —por ejemplo, recurriendo a documentos oficiales.⁴

Veamos ahora como se presenta la definición operativa cuando la unidad de análisis es una provincia u otro agregado territorial: imaginemos que la propiedad que interesa sea la proporción de población licenciada en cada provincia. En este caso —como para la mayoría de las propiedades relativas a provincias, Estados u otros agregados territoriales— se recurre al censo.

La definición operativa indicará:

- 1) las tablas de la publicación del censo donde se encuentran informaciones sobre el número de licenciados y el número de habitantes con más de 20 años de edad;
- 2) la operación matemática que se debe realizar (una división del número de licenciados por el número de habitantes de la edad requerida);
- 3) la ulterior operación matemática que se realizará para hacer más legible el resultado (por ejemplo, multiplicar los cocientes por 1.000 y redondear el resultado para eliminar los decimales);
- 4) la indicación del vector-columna en que se debe poner el número resultante;
- 5) un procedimiento de control mecánico, por ejemplo, la ejecución de los cálculos al revés para ver si se obtienen los mismos números que aparecen en las tablas de las que se iniciaron los cálculos;
- 6) una valoración de la plausibilidad de los resultados. En efecto, puede pasar que la proporción de población licenciada en una provincia rural

⁴ Acerca del problema del control de la fidelidad de los datos, véase el apartado 7.3.

y marginal aparezca más alta que la misma proporción en una provincia urbanizada y terciarizada. No se puede excluir que ésta sea la situación efectiva; pero su escasa plausibilidad sugiere revisar los cálculos e incluso poner en duda los mismos datos en las tablas del censo.

En los ejemplos hemos presentado las partes de dos definiciones operativas que eran específicas de las propiedades en cuestión (título de estudio individual y proporción de población licenciada en la provincia). Pero si consideramos que una definición operativa es el conjunto de todas las operaciones, intelectuales y materiales, que permitan pasar de estados reales a números (en un vector-columna) que representen esos estados en la matriz, debemos concluir que, junto con las reglas y convenciones específicas de la propiedad en cuestión, una definición operativa incluye reglas y convenciones más generales y comunes a todas las propiedades que se indagan en esa investigación específica y en todas aquellas análogas. Por ejemplo, si la técnica de recolección de los datos es una entrevista con cuestionario, la definición operativa de cada propiedad investigada incluye el arte de redactar el texto de las preguntas de manera simple, el arte de ordenarlas en el cuestionario de forma correcta y eficaz, el arte de contactar al entrevistado y plantearle preguntas obteniendo respuestas sinceras, y, por lo tanto, el nivel de adiestramiento y motivación de los entrevistadores —en pocas palabras, toda la sabiduría que se ha acumulado en décadas de desarrollo de los sondeos.

La tesis de que estos aspectos propios de una técnica de recolección en general pertenecen a la definición operativa de cada propiedad investigada con esa técnica puede parecer extraña si no se reflexiona sobre el hecho de que una pregunta mal formulada o mal planteada puede obtener respuestas radicalmente diferentes de las que obtendría si fuera formulada de la mejor manera, y que la misma pregunta sobre una cuestión embarazosa obtiene a menudo respuestas diferentes según se ponga al inicio, a medio camino o al final de un cuestionario.⁵

7.2. La fidelidad de los datos

Un dato tiene escasa fidelidad si no representa correctamente el correspondiente estado en la propiedad según las convenciones de la definición operativa.⁶ Como hemos visto antes, una causa de datos de escasa fidelidad son las declaraciones mentirosas de los entrevistados: se tiende a reducir la

⁵ Décadas de investigaciones específicas han averiguado de forma conclusiva los efectos determinantes de variaciones en el texto de una pregunta, en su posición en un cuestionario, en el número y sucesión de las alternativas de respuesta, en el modo en que se presentan, etcétera.

⁶ Para aclarar lo que significa la oración en cursiva, piénsese en una definición operativa que prescribe registrar la propiedad edad en décadas; en este caso el número que representa una edad de 10 años y una de 19 será el mismo (supongamos, el número 1). En eso, se tratará de una definición operativa con un plan de codificación poco sensible; pero los datos 1 que corresponden a edades entre 10 y 19 años representarán la edad en la forma prevista por la definición operativa.

edad y el ingreso, a elevar el título de estudio y la frecuencia con la que se leen libros, se va al teatro y se siguen programas culturales por televisión. Raramente se admiten hábitos sexuales heterodoxos, simpatías por partidos y grupos extremos, conductas ilegales (evasión impositiva, ausentismo en el trabajo, etcétera).

Los investigadores que estudian habitualmente las opiniones de la gente saben por experiencia que los entrevistados —si no tienen un interés particular y opiniones firmes en el tema de la pregunta— tienden a declarar como propias las opiniones que creen ortodoxas y compartidas por la mayoría. Este fenómeno, que no se puede considerar una mentira en sentido estricto, es conocido como “deseabilidad social”, y es tan difundido que hace poco fiables los resultados de encuestas sobre opiniones realizadas con preguntas sencillas y directas, ya que es fácil para el entrevistado imaginar cuál sea la respuesta convencional y por lo tanto “socialmente deseable”.⁷

Además, los entrevistados pueden dar involuntariamente informaciones que generan datos no fidedignos. Puede ser que no entiendan la pregunta o no le atribuyan el mismo significado que le da el investigador. Puede ser que la entiendan correctamente pero se equivoquen acerca de su propio estado en la propiedad investigada. Puede ser que entiendan mal el mecanismo de algunas técnicas sofisticadas que imaginaron psicólogos y metodólogos para registrar opiniones y actitudes⁸ y alteren sus estados sin quererlo.

Si la pregunta implica el recuerdo de acciones y situaciones pasadas, la experiencia ha mostrado que la memoria presenta algunos sesgos sistemáticos, en el sentido de que se tiende a acercarse en el tiempo algunos tipos de eventos y a alejar otros (véase Montesperelli 2005: cap. 2).

Naturalmente, el entrevistado no es la única fuente de datos no fidedignos en un sondeo. El entrevistador, a su vez, puede leer mal algunas preguntas o algunas respuestas del plan de codificación; puede influir al entrevistado manifestando sus propias preferencias o sus expectativas acerca de las respuestas; puede equivocarse asignando una categoría errada a una respuesta del entrevistado.

Faltas materiales de transcripción se pueden producir en cada eslabón de la cadena de pasajes que une la respuesta vocal del entrevistado a una cifra en el soporte electrónico que representa la matriz de datos.

Se podría pensar que estas faltas materiales de transcripción son la única causa posible de escasa fidelidad cuando las informaciones se refieren a unidades de análisis como provincias o Estados, y son recogidas por censos o encuestas de departamentos estadísticos públicos. Pero no es así: la gran mayoría de estos datos que convencionalmente se llaman “duros” —para distin-

⁷ El problema de la “deseabilidad social” fue identificado por primera vez por Allen Edwards (1957b) analizando las respuestas al afamado test Minnesota Multiphasic Personality Inventory. Otros importantes estudios del problema son Wiggins (1968), Orvik (1972), DeMaio (1984), Presser y Traugott (1992), Corbetta (2003).

⁸ Véase el capítulo 8.

guirlos de los datos de sondeo, que serían “blandos”—proviene de entrevistas, y por lo tanto sufren de las mismas causas de infidelidad que hemos examinado antes.⁹

7.3. Para controlar la fidelidad

Dado que la escasa fidelidad de los datos empeora la calidad y reduce la credibilidad de una investigación, se debería dedicar mucha más atención al control de la fidelidad de la que se dedica habitualmente.¹⁰

En el apartado 7.1, hablando de la definición operativa, ya se hizo referencia a algunas formas de ese control. Ahora se reseñará la gran variedad de estas formas.

La más rápida—que no cuesta nada de dinero e insume poco tiempo, y por lo tanto no hay alguna razón para evitar—se desarrolla explotando la matriz misma,¹¹ y puede seguir dos caminos. El más simple es una inspección de las distribuciones de frecuencia de cada variable, es decir, de cuántos datos tiene cada uno de los códigos numéricos posibles en esa variable.¹² De todas formas, esta inspección es el primer paso del análisis, porque nos permite tomar contacto con los datos que hemos recogido. Además, nos permite descubrir si algunos datos tienen códigos que no estaban previstos en el plan de codificación (por ejemplo, un dato con código 3 en la variable “sexo” cuando el plan prevea asignar 1 a las mujeres, 2 a los varones y 9 a las informaciones faltantes).

El análisis de las distribuciones de frecuencia de cada variable de una matriz para encontrar códigos no previstos se llama *wild code check* (control de códigos “locos”). Si la unidad de análisis es una provincia y la variable en cuestión es el número de habitantes, o de licenciados, o de hospitales, ese control se realiza de forma diferente, buscando a qué provincias corresponden los valores más altos, o más bajos, en cada variable cuantitativa, para ver si las cifras que encontramos conciben más o menos con lo que sabemos de esa provincia: por ejemplo, si resulta que La Rioja tiene más licenciados que la provincia de Buenos Aires, esto es una señal de que hubo una falta material en la codificación de los datos (por ejemplo, se agregó un cero al final de la cifra, con el efecto indeseado de multiplicarla por diez).

El otro camino con el que podemos controlar la fidelidad de (algunos) datos

⁹ Muchos de los que han trabajado con estos datos “duros” denuncian que son escasamente confiables. Véanse por ejemplo Taylor y Hudson (1972) y los ensayos recogidos en Merritt y Rokkan (1966).

¹⁰ En efecto, como se observa en un texto clásico de epistemología, “no hay manipulación matemática que pueda remediar una escasa precisión de los datos que se manipulan” (Cohen y Nagel 1934: 292).

¹¹ Sin embargo, si bien estos controles rápidos en la matriz permiten detectar datos no fidedignos, no permiten descubrir cuáles deberían ser los datos fidedignos. Para lograr esto, tenemos que reexaminar el proceso de recolección, e incluso repetirlo.

¹² Acerca de las distribuciones de frecuencia, véase el apartado 8.1.

explotando la matriz misma se basa en el hecho de que algunas combinaciones de categorías son social o jurídicamente imposibles en una sociedad dada: no puede haber curas católicos de sexo femenino, ni jueces que no tengan el título de abogado, ni electores menores de dieciocho años, etcétera.

Como se puede entender, esta forma de control, que se llama *consistency check* (control de coherencia), sólo se puede aplicar a un número limitado de situaciones—lo que de todos modos no es una buena razón para no aplicarlo.

Todas las otras formas de control se aplican confrontando los datos con informaciones que se encuentran fuera de la matriz. Ya hemos visto un caso hablando en el apartado 7.1 de la valoración de la plausibilidad de los resultados de una definición operativa cuando la unidad de análisis es una provincia u otro agregado territorial. En el mismo apartado se vio que el recurso a documentos oficiales permite controlar algunas declaraciones de los entrevistados (por ejemplo, declaraciones sobre el hecho de que votaron en las últimas elecciones, o que tienen o no tienen teléfono, licencia de conducir, etcétera).¹³

Como el precedente, ese tipo de control es conceptualmente simple y tiene un resultado en principio fiable. Pero necesita una labor paciente y tediosa y una considerable inversión de tiempo y dinero, porque los documentos oficiales no siempre son tan fáciles de consultar como se proclama oficialmente. Además, con la difusión del culto de la privacidad en los países occidentales, el número de documentos oficiales que se pueden consultar se ha ido reduciendo dramáticamente. De todas formas, este tipo de control sólo puede aplicarse a estados en propiedades que sean oficialmente registradas—seguramente una minoría entre las que interesan al investigador social.

Un tercer tipo de control no sufre de esta limitación, en el sentido de que puede aplicarse a cualquier propiedad individual. Pero necesita grandes inversiones de tiempo y dinero y—a diferencia de las otras formas—una gran destreza y sabiduría de parte de los que ejercen los controles.

Según nuestro conocimiento, esta forma de control—imaginada, por ejemplo, por McNemar (1946)—fue realizada sistemáticamente sólo en un centro de investigaciones metodológicas de la Universidad de Lodz, en Polonia, liderado por Jan Lutinski, que la llamó “entrevista sobre la entrevista”. La situación en la que se desarrolló esta técnica de control es muy particular y merece ser analizada en detalle.

Es sabido que los regímenes autoritarios o totalitarios son hostiles a las ciencias sociales, porque tienden a dar una conciencia crítica a la población dificultando su adoctrinamiento. Además, como el marxismo sostenía ser la única ciencia legítima de la sociedad, después de la toma del poder en Rusia, los comunistas procuraron erradicar a las nascentes ciencias sociales.¹⁴

¹³ En los Estados Unidos y en algunos otros países se puede controlar el ingreso declarado por un entrevistado con su declaración fiscal (Katona 1951).

¹⁴ Suerte radicalmente diferente tocó a la psicología, que fue impulsada para crear el nuevo “hombre soviético”. Con Vygotskij, Lurija, Leontiev y otros, la psicología rusa estuvo en efecto a la vanguardia en el período de entreguerras.

Lo mismo hicieron al final de la Segunda Guerra Mundial en Europa oriental; con la excepción de Polonia, donde la sociología tenía gran tradición y alto prestigio.¹⁵ Los gerentes comunistas polacos no se arriesgaron a callar y desterrar a los sociólogos y en cambio trataron de explotar sus conocimientos y su sabiduría para conocer mejor lo que la población pensaba realmente del régimen.

En aquel tiempo se decía comúnmente que cada polaco tenía tres verdades: una para el partido, una para la Iglesia y una para sus amigos de confianza. Obviamente, los sondeos que el partido organizaba periódicamente para conocer el estado de la opinión pública sobre tal o cual cuestión sólo podían brindarle respuestas que reproducían exactamente las verdades oficiales, y por lo tanto tenían un valor informativo nulo. Los políticos se daban cuenta de que para tener un cuadro más realista debían permitir a los sociólogos investigar lo que realmente pensaba la gente.

En este contexto se desarrolla la actividad del Centro Metodológico de Lodz (la tercera ciudad de Polonia), una institución pública cuya única tarea era refinar las herramientas del sondeo para mejorar la fidelidad de los datos. Además, el centro no sufría ninguno de los condicionamientos que caracterizan la actividad del metodólogo (a menudo presupuestos insuficientes, tiempo siempre limitado, presiones para publicar rápidamente, tal vez sutiles presiones en dirección a los resultados deseados por los que encargan o por el teórico).

Después de organizar un sondeo sobre un problema determinado, el centro seleccionaba un pequeño subconjunto de los entrevistados y encomendaba cada uno de ellos a uno de sus investigadores, que debía dedicarle todo el tiempo necesario para ganar su confianza, invitándolo entre otras cosas a comentar libremente los temas del cuestionario para llevarlo a dar respuestas más sinceras que las dadas al cuestionario estandarizado (Pilichowski y Rostocki 1978).¹⁶ Lo más difícil era obviamente convencer a los entrevistados de que no corrían ningún riesgo como consecuencia de lo que decían. Esto estaba garantizado por los líderes políticos, que no tenían interés en cortar esos valiosos canales de informaciones confiables, y por lo tanto respetaban el estricto anonimato que los investigadores a su vez garantizaban a los entrevistados.

Hemos dedicado gran atención a la actividad de este centro polaco para mostrar qué particulares y extrañas condiciones deben darse para posibilitar un trabajo en profundidad acerca de la fidelidad de los datos de encuestas.¹⁷

¹⁵ Piénsese en Ossowski, Chalasiniski, Znaniecki.

¹⁶ El metodólogo italiano Gilli propuso efectuar esta profundización durante la misma entrevista inicial (1971: 117). Pero de esta forma ésa se convertiría en una entrevista en profundidad —una herramienta típica de la aproximación no estándar— (véase el capítulo 12).

¹⁷ En verdad, algunos controles de fidelidad de las respuestas estandarizadas a través de sucesivas preguntas abiertas se realizan en algunos centros de investigaciones en Inglaterra y Estados Unidos (véanse por ejemplo Schuman 1966; Belson 1981). Son controles llevados a cabo a menudo durante la misma entrevista y siempre por entrevistadores —no por investigadores como en Polonia. Nada que ver con el empeño y la profundidad del trabajo del Centro de Lodz.

No es nuestra intención proponer este camino para controlar la fidelidad de centenares de respuestas a los millares de cuestionarios estandarizados de un sondeo de masas. El tiempo, el dinero y la habilidad requeridos para plantear preguntas en una entrevista estandarizada son inconmensurables con la inversión de tiempo, dinero y habilidad necesaria para controlar la fidelidad de las respuestas en el modo que hemos descrito.

Tampoco se quiere afirmar que una “entrevista sobre la entrevista” a la manera del Centro de Lodz va a garantizar la fidelidad de las respuestas a preguntas sobre opiniones y actitudes. Incluso los investigadores más hábiles y persuasivos no pueden garantizar esto.

Lo que se pretende afirmar es que éste es el único camino para controlar la fidelidad de respuestas a preguntas sobre opiniones y actitudes, cuando por fidelidad se entienda la correspondencia entre el dato que se anota en la matriz y lo que el entrevistado opina realmente. Las herramientas de control de la fiabilidad que los psicólogos han desarrollado durante el siglo pasado, y que se usan en todas las ciencias sociales, no logran este resultado, a pesar de sus pretensiones. Trataremos de mostrarlo en el apartado siguiente.

7.4. Los coeficientes de fiabilidad: una crítica

El largo camino que conduce a la familia actual de coeficientes que tratan de medir la fiabilidad (en inglés *reliability*) empieza en la astronomía del siglo XVIII.¹⁸ En dicho siglo la consolidación del paradigma kepleriano-galileano-newtoniano, con su robusto *corpus* teórico, permitió calcular en abstracto y con gran seguridad la posición de los planetas conocidos entonces. Confrontando estas posiciones teóricas con las posiciones observadas mediante telescopios, los astrónomos se dieron cuenta de que hay siempre diferencias entre unas y otras. Dado que tenían plena confianza en sus teorías, no dudaron de que esas diferencias se debieran a fallas de las herramientas y/o de los hombres que las manejaban.

Nace así el problema de la precisión de las herramientas y de los observadores —que se convertirá en un problema central en la filosofía de la ciencia. El primer ensayo sobre “fallas de observación” es un artículo de 1757 del matemático inglés Thomas Simpson. Poco después, al problema se dedicaron dos grandes matemáticos y astrónomos franceses, el conde de Lagrange y el marqués de Laplace. A principios del siglo XIX el matemático alemán Karl Friedrich Gauss recolectó unas series de observaciones de la posición de los planetas registradas durante años por varios observatorios astronómicos europeos. Confrontando las posiciones de un planeta cualquiera observadas por varios telesco-

¹⁸ En realidad, parece que los astrónomos sumerios usaban el promedio de las posiciones observadas de una estrella como estimación de su posición efectiva. Eso muestra la gran sofisticación gnoscológica de esos sabios, plenamente conscientes de la posibilidad de un desajuste entre la realidad y lo que se observa de ella.

pios con su posición teórica, calculada según la teoría astronómica, Gauss encontró que las posiciones observadas se disponían alrededor de la posición calculada, y que había muchas diferencias pequeñas y pocas diferencias grandes en ambos lados de dicha posición. En otras palabras, las fallas de observación tenían una distribución campanular, que tomará el nombre de ese matemático (distribución gaussiana).¹⁹

Si se suman los cuadrados²⁰ de estas diferencias entre las posiciones observadas y la posición calculada, se puede cuantificar el nivel de (im)precisión del complejo de las herramientas y operadores. La medida de esta imprecisión es exactamente la varianza de las observaciones, es decir, la media de sus desviaciones cuadradas de la posición calculada.²¹ Obviamente, el inverso de esta varianza mide la precisión.

Nótese que ésta es una medida global, que se refiere —como se decía— al complejo de herramientas y observadores. Para cuantificar la precisión de cualquier pareja telescopio/observador se debería sólo considerar sus observaciones, y para atribuir la falla a uno u otro se deberían alternar varios observadores en el mismo telescopio y varios telescopios con el mismo observador. Pero los astrónomos no han seguido este camino.

En el siglo XIX, la idea de confrontar los resultados de observaciones repetidas del mismo fenómeno para cuantificar la precisión de las herramientas de medición (por ejemplo, la balanza) se difunde en las ciencias físicas. Al final del siglo, el psicólogo alemán Hermann Ebbinghaus trata de aplicar la idea en sus investigaciones sobre la memoria (1885). Poco después, el gran psicólogo inglés Charles Spearman, en el mismo ensayo (1904) en que esboza los fundamentos del análisis factorial, propone medir la fiabilidad de un test calculando el coeficiente de correlación²² entre dos vectores de puntajes obtenidos en una muestra de sujetos a los que el test fue aplicado en dos momentos diferentes. Se trata de la primera codificación de la técnica conocida como control de fiabilidad *test-retest*.

Puesto que esta técnica es todavía regularmente usada en psicología, se describirá en detalle su aplicación:

- a) para “medir” la fiabilidad de un test (que comprende habitualmente un cierto número de distintas preguntas o pruebas) se le aplica en un cier-

¹⁹ La particular forma de distribución campanular que se llama distribución normal había sido definida en abstracto por el matemático francés Abraham De Moivre (1733). Sin embargo, Gauss fue el primero en encontrar esta importante aplicación práctica.

²⁰ Gauss introduce el criterio de sumar los cuadrados de las desviaciones porque, si no se eliminase el signo con la elevación al cuadrado, las desviaciones negativas tenderían a anular las desviaciones positivas. Desde entonces este criterio ha sido la regla en estadística.

²¹ Sobre el concepto de varianza, véase el apartado 8.3.

²² Esta fundamental herramienta del análisis estadístico (véase el apartado 15.6) había sido recientemente propuesta por Francis Galton (1888) y perfeccionada por Karl Pearson (1896). Sobre este período de la psicología británica, fundamental para el desarrollo de las actuales técnicas de análisis de datos en el enfoque estándar, véase Piovani (2006).

to día a una muestra de sujetos (habitualmente, los estudiantes de la clase universitaria de psicología dictada por el psicólogo que propone el test);²³

- b) se registran las respuestas a las preguntas (o las soluciones de las pruebas) y se le atribuye un puntaje a cada respuesta;
- c) a cada sujeto se le atribuye la suma o el promedio de sus puntajes, y esa cifra, siendo el resultado global de cada individuo en el test, se pone en un vector-columna (un puntaje para cada sujeto);
- d) después de algún tiempo (habitualmente, una o más semanas) se repite la fase (a) con los mismos sujetos, las mismas preguntas o pruebas y en una situación en cuanto posible análoga;
- e) se repite la atribución de puntajes a los sujetos, como sub (b) y (c), y se ponen las cifras en otro vector-columna;
- f) se calcula un coeficiente de correlación entre los dos vectores;
- g) esta cifra, llamada “coeficiente de correlación test-retest”, se considera una medida de la fiabilidad del test en consideración.

Muchos psicólogos y psicómetras han expresado varias críticas a esta forma de control de la fiabilidad:

- a) en el intervalo de tiempo entre el test y el retest, el estado de algunos sujetos en la propiedad investigada puede efectivamente cambiar, como pueden cambiar algunos aspectos de la conducta del investigador que aplica el test, o de la situación en general.²⁴ Todo eso va a reducir el coeficiente de correlación y por consiguiente nuestra estimación de la fiabilidad del test; para limitar este sesgo sería por lo tanto aconsejable un intervalo muy corto entre test y retest;
- b) por otro lado, la primera administración puede cambiar artificialmente el estado del sujeto en la propiedad, y esto en varias formas:
 - 1) el sujeto puede acordarse las respuestas que dio en la primera ocasión y repetir las en la segunda para mostrarse coherente:²⁵ esto va a acrecentar la aparente fiabilidad del test, y para limitar el sesgo sería aconsejable un intervalo muy largo entre test y retest;
 - 2) si el test concierne a cualquier forma de habilidad, la primera administración puede actuar como entrenamiento, y la prestación de algunos sujetos puede mejorar en el retest:²⁶ esto va a reducir la aparente fiabilidad del test, y para limitar el sesgo sería aconsejable un intervalo largo entre test y retest;
 - 3) la primera administración del test puede acrecentar el interés de algunos sujetos hacia los problemas investigados, y ellos pueden re-

²³ La misma observación en Corbetta (2003: 153).

²⁴ Véanse, por ejemplo, Webb *et al.* (1966), Converse (1970), Berka (1983).

²⁵ Véanse, por ejemplo, Kuder y Richardson (1937), Ingram (1977).

²⁶ Véanse, por ejemplo, Cronbach (1949), Anastasi (1953), Corbetta (2003).

flexionar, informarse, etcétera²⁷ —lo que puede cambiar sus respuestas a las preguntas y por consiguiente reducir la aparente fiabilidad del test;

- 4) por lo demás, la primera administración del test puede tener un abanico de otros efectos²⁸ (acostumbramiento y desencanto, menor miedo y por lo tanto mayor sinceridad, mayor control de la situación y por lo tanto mejor capacidad de adivinar las respuestas “socialmente deseables”, etcétera).

Por este complejo de motivos Cronbach (1947), Anastasi (1954) y otros aconsejan señalar siempre el intervalo entre la primera y la segunda administración del test —lo que sería determinante si los estados sólo cambiasen espontáneamente, como función lineal del tiempo. Pero hemos visto que no es así: hay muchos otros factores de distorsión, que varían de situación a situación y que no se pueden determinar fácilmente —y todavía menos controlar.

Como se vio, el problema medular que —de una forma u otra— llamaba la atención crítica de algunos psicólogos era la distancia temporal entre las dos administraciones del test. En las primeras décadas del siglo XX, los esfuerzos de los psicómetras apuntaban a eliminar esta distancia, y con ella todos los problemas que traía aparejados.²⁹ ¿Pero cómo lograr igualmente dos vectores de cifras para correlacionarlos y calcular un coeficiente, con su aspecto tan sólido y científico?

Una solución hallada fue realizar dos tests sobre el mismo tema, con el mismo número de preguntas o pruebas, formuladas de manera análoga, que se bautizaban como “paralelos”. Los dos vectores deseados se producían obviamente sumando los puntajes logrados por cada sujeto al contestar las preguntas (o solucionando las pruebas) de cada test. Si la correlación resultaba alta, eso confirmaba el supuesto de equivalencia de los dos tests y simultáneamente la fiabilidad de cada uno y del conjunto de los dos.³⁰

A principios de la década de 1930 los psicómetras propusieron una técnica aún más sencilla: administrar un solo test y luego crear los dos vectores deseados simplemente sumando por separado los puntajes conseguidos por un sujeto en las preguntas que tienen un número par en la secuencia del test y en las que tienen un número impar.³¹ El coeficiente de correlación así producido fue bautizado *split-half* (subdivisión a mitad) *reliability coefficient*.

Algunos críticos³² han notado que la división entre preguntas con número par y preguntas con número impar es arbitraria, y que cambiando el criterio de

²⁷ Véanse, por ejemplo, Campbell y Stanley (1963), Bohrnstedt (1970), Carmines y Zeller (1979).

²⁸ Véase, por ejemplo, Scott (1968).

²⁹ Incluso el problema práctico de convencer a los sujetos testeados la primera vez de dejarse testear una segunda vez con una herramienta que habitualmente es larga y aburrida, por motivos que veremos pronto.

³⁰ Sobre esta técnica, bautizada *parallel forms*, véanse Thurstone (1928), Anastasi (1954).

³¹ Véanse Brownell (1933), Zubin (1934), Kuder y Richardson (1937).

³² Kuder y Richardson (1937), McKennell (1977), Giampaglia (1985).

división interna el coeficiente *split half* podría cambiar de modo no desatendible. Pero la mayoría de los psicómetras tienen una preocupación diferente: dividiendo por dos el número de preguntas (o pruebas) se reduce automáticamente el nivel del coeficiente.³³ Se adopta con entusiasmo una fórmula que permite “corregir” (es decir: aumentar) el valor del coeficiente neutralizando el efecto de la reducción del número de preguntas/pruebas.

Quien tiene experiencia acerca de las reacciones de los sujetos investigados a frases planteadas en forma de batería sabe que muchos tienden a acordar con ellas sin reflexionar, independientemente de su sentido, por motivos que van de la prisa a un malentendido respecto de la institución que encarga el estudio, al deseo de quitarse un peso de encima. Este fenómeno —conocido como *response set*— es un mal endémico de las baterías, y acerca de sus causas existe una literatura vastísima.³⁴ Obviamente no hay ninguna garantía de que una respuesta que un sujeto da para quitarse un peso de encima corresponda a sus opiniones efectivas. Pero cada sujeto que acuerda con todas las oraciones de una batería va a aumentar inevitablemente el nivel de cualquier coeficiente asentado en la correlación entre vectores de cifras.³⁵

Este patente absurdo no parece preocupar a ninguno: si el único problema es acrecentar el nivel de un coeficiente de correlación entre dos vectores de cifras en la matriz de datos,³⁶ todo lo que coadyuve a este objetivo es bienvenido. Nadie se plantea el problema de la correspondencia de dichas cifras con los estados efectivos de los sujetos investigados.³⁷ El mundo real ha desaparecido. Un metodólogo sueco lo ha declarado simple y abiertamente: “La acumulación de estos estudios nos ha enseñado a confiar en estas técnicas [...]. Usándolas, los investigadores ya no tienen que encarar el largo y aburrido proceso de controlar la fiabilidad de sus datos” (Zetterberg 1954: 124).

Un progreso meramente técnico se produjo a fines de la década de 1930, cuando dos integrantes del más importante laboratorio psicométrico nortea-

³³ Muchas fórmulas de coeficientes que “midan” la fiabilidad de un test (por ejemplo el alfa de Cronbach) son construidas de tal forma que el nivel del coeficiente depende también del número de preguntas o pruebas: por este motivo muchos tests psicométricos son largos y aburridos. Para aumentar artificialmente el nivel del coeficiente se somete a los estudiantes “cautivos” a una faena que a muchos puede parecer sin sentido. Lo han observado, entre otros, McKennell (1970), Davies (1977), Ferrand y Martel (1986).

³⁴ Mencionamos por ejemplo Cronbach (1946), Block (1965), Hui y Triandis (1985), Greenleaf (1992), Johnson *et al.* (2005), y los ensayos en Berg (1967) y en Dijkstra y Van der Zouwen (1982).

³⁵ “Un alto valor del coeficiente puede depender totalmente de factores de sesgo” (Bartlett *et al.* 1960: 703).

³⁶ La idea de producir dos vectores para correlacionarlos y obtener un coeficiente se difunde de la psicología a otras disciplinas. Como en el *test-retest* se variaba el tiempo y en el *split half* “formas equivalentes” se variaba la herramienta, se pueden variar los observadores que valoran un texto cualquiera (por ejemplo, en el análisis de contenido) y obtener un coeficiente de acuerdo entre codificadores (*inter-individual* o *inter-coder reliability coefficient*: Zetterberg 1954; Singer 1982).

³⁷ Por lo menos algunos psicólogos admiten que en estos casos no se puede hablar de fiabilidad, y en cambio bautizan “coeficiente de estabilidad” al *test-retest* y “coeficiente de coherencia interna” al *split-half* (Cronbach 1947; Anastasi 1954; McKennell 1970).

americano (Kuder y Richardson 1937) propusieron una técnica que superaba el límite que compartían todos los coeficientes formulados hasta entonces, es decir basarse en una correlación entre dos vectores solos. Con este nuevo coeficiente se calculaba la correlación entre todas las respuestas a las preguntas o pruebas de un test, con tal de que ellas fuesen dicotómicas. Esta limitación fue superada en la segunda posguerra, cuando Cronbach (1951) propuso el coeficiente *alpha*, aplicable a todo tipo de batería. En las décadas siguientes la evolución técnica siguió siendo rápida: entre otros, se han formulado coeficientes basados en el análisis factorial (Heise y Bohrnstedt 1970) y en el análisis de componentes principales (Armor 1974).³⁸

Un punto a destacar es que desde su inicial concepción por Spearman, esta línea de herramientas para controlar la fiabilidad queda vinculada a los tests que emplean los psicómetros en sus laboratorios, es decir baterías de preguntas o pruebas de formato igual. No es aplicable a ninguna de las otras herramientas que se usan en un sondeo, y menos aún —obviamente— a todas las herramientas para recoger informaciones sobre unidades no individuales (agregados territoriales, etcétera).³⁹

Además, la idea misma de que se pueda asignar a un test una etiqueta numérica que atestigüe para siempre su nivel de fiabilidad, como si fuera un atributo permanente y definitivo, sufre de una falla gnoseológica aún más grave que los varios límites técnicos que hemos visto arriba. Más que de una falla simple, se trata de un complejo de fallas que conciernen el problema de la generalizabilidad.

La más obvia es que no es legítimo generalizar a todas las poblaciones en todos los tiempos y situaciones un resultado obtenido con una población específica (habitualmente estudiantes de psicología) en un tiempo y situación específicos (habitualmente la clase del autor del test, y apenas éste ha sido imaginado). Esto parece obvio, pero “los coeficientes de fiabilidad encontrados en un grupo son transferidos a todos los otros grupos sin cuestionamientos” (Tukey 1954: 40). Se encuentran frecuentemente compilaciones de tests sobre varios temas donde cada test va acompañado por su coeficiente de fiabilidad;⁴⁰ muy raramente se reportan los resultados de controles sucesivos a la primera determinación del coeficiente.⁴¹

Una observación más sutil es que la generalización es ilegítima incluso entre el mismo grupo de sujetos a los que se aplica el test para calcular su coeficiente de fiabilidad, en cuanto:

- a) cada sujeto tiene una manera particular de reaccionar a cada pregunta o prueba de un test;

³⁸ Una exhaustiva crítica técnica de todos los coeficientes de fiabilidad de la segunda generación está en Giampaglia (1986).

³⁹ Una observación análoga en Corbetta (2003: 108).

⁴⁰ Véanse, por ejemplo, Fiske (1971), Cronbach *et al.* (1972), Robinson *et al.* (1991).

⁴¹ Una excepción es el manual de Shaw y Wright (1967), que reporta variaciones de un coeficiente .98 inicial a un .71 en el control sucesivo.

- b) cada pregunta, o prueba, de un test tiene su particular nivel de inteligibilidad, de aceptabilidad, de agrado;
- c) la reacción de cada sujeto particular a cada pregunta, o prueba, particular no es estable, sino que puede variar en cada situación específica.

Reduciendo toda esta variabilidad a un simple coeficiente que se pretende universal y definitivo se echa por la ventana la especificidad y no fungibilidad de los seres humanos, que había sido de alguna forma (aunque mínimamente) reconocida aplicando los tests a una pluralidad de individuos en lugar de sentar todas las teorías sobre un individuo aislado.

Como ha observado el psicólogo Frank:

“Las técnicas universalmente usadas en psicología sólo dan informaciones acerca de un complejo de individuos, y nada sobre cada sujeto particular [...]. Estos tests, imaginados cuando no había otra manera de controlar la fiabilidad, ahora se han convertido en fetiches y bloquean la búsqueda de otras formas de control” (1939: 400).

Una línea crítica diferente ha sido expresada por el antropólogo Naroll:

“El término ‘fiabilidad’ en su significado técnico de congruencia entre dos o más vectores no es atinado. El significado no técnico de ‘fiable’ es ‘digno de confianza’. No creo que los científicos sociales vayan a tener más confianza en una herramienta por efecto de un test de fiabilidad. Se deberían llamar ‘tests de congruencia’ —lo que efectivamente son— más que ‘tests de fiabilidad’ —lo que en el sentido ordinario del término no son para nada” (1968: 265-6).

A varias décadas de distancia de estos escritos, la situación permanece problemática. Los controles basados en los coeficientes descritos en este apartado se han vuelto todavía más técnicos y alejados de todo contacto con situaciones reales. Para controlar realmente la fidelidad, sólo tenemos los procedimientos que se han descrito (apartados 7.1 y 7.3). Algunos (*wild code check*, *consistency check*, inspección de documentos oficiales) se pueden aplicar sólo raramente; otros (la entrevista sobre la entrevista, y formas parecidas) demandan una enorme inversión de tiempo, recursos y habilidad. Aparte de esto, no hay otro camino para mejorar la fidelidad de los datos que una gran atención al formular las definiciones operativas, soportada por una gran experiencia de todas las situaciones de peligro (para la fiabilidad) que se producen a lo largo del proceso de recolección de datos.

Algunas sugerencias serán dadas en los capítulos siguientes.

7.5. ¿Las definiciones operativas registran o manipulan opiniones y actitudes?

En el apartado precedente hemos visto que, limitándose al análisis de datos en la matriz y olvidando mirar afuera en el mundo, se pueden lograr coeficien-

tes de fiabilidad tranquilizadores que, sin embargo, no tienen ninguna relación con las situaciones reales.

Naturalmente, la responsabilidad de esta manera de proceder no es propia de la matriz, sino de la actitud de quienes la usan. La matriz en sí es un instrumento neutro, que se puede usar también para descubrir indicaciones significativas sobre dichas situaciones reales. Lo hizo hace medio siglo Donald T. Campbell, el metodólogo más dotado de espíritu crítico y de sensibilidad gnoseológica.⁴² Con la ayuda del psicólogo Donald W. Fiske pasó revista a los miles de investigaciones depositadas en los archivos de la Universidad de Michigan, con el fin de identificar todas las que habían registrado propiedades psíquicas (opiniones, actitudes, valores) con tres o más definiciones operativas diferentes. Para cada una de las investigaciones que cumplían con estos requisitos Campbell y Fiske construyeron una matriz del tipo que se ve en la tabla 7.1, que bautizaron *multitrait multimethod matrix* (matriz con muchas propiedades y muchas técnicas).

Tabla 7.1: *Multitrait multimethod matrix*

	Al	Aa	At	El	Ea	Et	Pl	Pa	Pt
Al	1	p	p	t			t		
Aa	p	1	p		t			t	
At	p	p	1			t			t
El	t			1	p	p	t		
Ea		t		p	1	p		t	
Et			t	p	p	1			t
Pl	t			t			1	p	p
Pa		t			t		p	1	p
Pt			t			t	p	p	1

En el ejemplo de la tabla, se consideran tres propiedades psíquicas (autoritarismo, extraversión, particularismo) cada una registrada con tres diferentes técnicas (escala de Likert, escala autoanclante, termómetro: véase apartado

⁴² Se trata del mismo que en obras clásicas (Campbell y Stanley 1963; Cook y Campbell 1979) criticó la simpleza de muchos "experimentos" en las ciencias sociales, estableciendo una serie de condiciones y de procedimientos que permiten reducir la gran distancia epistemológica entre los experimentos de las ciencias físicas y lo que se puede lograr en las ciencias sociales. Véase apartado 1.2.

8.5). La tabla representa una matriz de correlación (véase apartado 15.6) Cada fila representa una propiedad particular registrada con una técnica particular: así Al significa autoritarismo registrado con una escala de Likert, Aa significa autoritarismo registrado con una escala autoanclante, etcétera. Como en todas las matrices de correlación, la secuencia de elementos en las filas se repite idéntica en las columnas, y en cada celda está el coeficiente de correlación entre la variable en la fila correspondiente y la variable en la columna correspondiente. Naturalmente, en las celdas de la diagonal descendente están las correlaciones de una variable consigo misma, que tienen el valor 1. En la tabla, las celdas señaladas con una p contienen la correlación entre variables que registran la misma propiedad con técnicas diferentes; las celdas señaladas con una t contienen la correlación entre variables que registran propiedades diferentes con la misma técnica; las celdas blancas contienen la correlación entre variables que registran propiedades diferentes con técnicas diferentes.

Si la técnica de registro no tuviese ningún impacto, es decir no influyese mínimamente en la forma en que los estados en las propiedades son convertidos en datos, no habría razón para esperar que las correlaciones entre variables que registran propiedades diferentes con la misma técnica fuesen más altas que las correlaciones entre variables que registran propiedades diferentes con técnicas diferentes; las únicas correlaciones fuertes se deberían encontrar entre las variables que registran la misma propiedad con técnicas diferentes. Usando los símbolos de las celdas de la tabla 7.1, la expectativa sería:

$$p > t = \square$$

Por el contrario, Campbell y Fiske encontraron que en la gran mayoría de las matrices *multitrait multimethod* construidas con los datos de las investigaciones depositadas en los archivos, la situación efectiva era:

$$t \gg p > \square$$

Es decir, los coeficientes de correlación entre las variables eran casi siempre más o mucho más altos cuando sólo la técnica era la misma que cuando sólo la propiedad era la misma. En otras palabras, las técnicas determinaban los estados que sólo deberían haber registrado.

Cuando Campbell y Fiske divulgaron sus resultados (1959), el artículo tuvo el impacto de un sismo en la comunidad de investigadores que adoptaban la aproximación estándar. Durante años no se publicó un solo informe de investigación sobre opiniones y actitudes en el que no se apelase ritualmente a él (y luego se siguiese empleando las mismas técnicas que el artículo ponía en tela de juicio). Lentamente, incluso los homenajes formales se hicieron raros, y hoy en día el artículo se menciona sólo como un destacado antecedente de la tendencia actual hacia la triangulación metodológica (véase apartado 2.4).

CAPÍTULO 8

CLASIFICACIÓN, CONTEO, MEDICIÓN, CONSTRUCCIÓN DE ESCALAS

Hasta ahora hemos considerado las propiedades (y por lo tanto las variables que las representan en la matriz de datos) globalmente. En este capítulo consideramos los diversos estados en cada propiedad, las consecuencias de la naturaleza de esos estados sobre la manera en que se transforman en datos y el tipo de análisis al cual se pueden someter las variables correspondientes.

Algunas propiedades, tanto de objetos físicos como de individuos o de unidades de otro tipo, tienen un número finito de estados, claramente distintos (o distinguibles) el uno del otro: en una palabra, discretos. Y DISCRETAS se llaman las propiedades que tienen estados discretos.

Las otras propiedades tienen un número infinito de estados imperceptiblemente diferentes el uno del otro: en este caso se habla de propiedades CONTINUAS.

Veremos pronto ejemplos. Por el momento, cabe destacar que la distinción entre propiedades discretas y propiedades continuas es la más importante cuando se trata de recoger informaciones y transformarlas en datos por medio de una definición operativa.

Las propiedades discretas se dividen en tres clases según el tipo de relación que existe entre sus estados.

Si los estados son sólo diferentes el uno del otro, pero no hay ningún tipo de relación cuantitativa entre ellos, se habla de propiedades CATEGORIALES. Ejemplos: la nacionalidad de un individuo, el tipo de régimen político de un estado, el tipo de producción agrícola prevaleciente en un territorio, etcétera (véase apartado 8.1).

Si los estados tienen relaciones ordinales (de mayor y menor) entre ellos, es decir, se pueden ordenar a lo largo de una escala, se habla de propiedades ORDINALES. Ejemplos: el nivel de estudio alcanzado por un individuo, su grado jerárquico en una organización civil o militar; el grado de centralidad de una comuna en su provincia, el grado de dureza de un metal (escala de Mohs), etcétera (véase apartado 8.2).

Si los estados tienen relaciones cardinales, en el sentido de que se puede legítimamente establecer un cociente entre ellos, se habla de propiedades DISCRETAS CARDINALES. Ejemplos: el número de componentes de una familia, el nú-