

# MODELOS ESTRUCTURALES – MATRICES DE COVARIANZAS

## 1. Regularidades y estructuras

La sociología existe en la medida en que existe regularidad y estructura. Esto no implica o excluye la existencia del cambio. Toda sociedad es histórica, y en ese sentido, cambiante. Sin embargo, la posibilidad de conocer y más aún, de explicar, exige de una pauta o regularidad. Así, el cambio podrá ser objeto de estudio siempre que no sea de tipo aleatorio. Es importante advertir, asimismo, que esta regularidad no es de tipo “normativo”, impuesta por unas “normas sociales”. Debe diferenciarse entre diferentes fuentes de regularidad social, señalando aquella que puede ser objeto de la investigación sociológica. Como señala E. Lamo “Ahora bien, es sabido que sobre el concepto científico-newtoniano de ley se superpuso otro más antiguo: el que establece una conexión de tipo *normativo* entre fenómeno antecedente o fenómeno consecuente, especialmente en la obra de Talcott Parsons, ha tratado de encontrar en dicha normatividad la causa de las regularidades sociales. Ello supondría reducir el concepto de la ley social al de norma. Creo, pues, conveniente, antes de nada, rechazar esta posibilidad teórica para devolverle al concepto de ley su autonomía científica en sociología”<sup>1</sup>.

El problema radica en que el sistema normativo puede hacer *comprensibles* las regularidades del comportamiento individual, pero no *explicar* tales comportamientos ni tampoco las *regularidades* del agregado y ambos fenómenos son sustancialmente diferentes. Así, Von Hayek ha propuesto distinguir tajantemente entre “el sistema de reglas de conducta que gobierna el comportamiento de los miembros individuales del grupo [...] por un lado, y por otro, el orden o pauta de acciones que resulta de aquel para el grupo como totalidad”<sup>2</sup>. De este modo, la norma interpretada individualmente no responde de los comportamientos agregados. Es decir, la regularidad normativa (o ética) permite *comprender* acontecimientos *singulares*, por ejemplo, *acciones*, pero

---

<sup>1</sup> Emilio Lamo de Espinosa, *La sociedad reflexiva*, CIS, Madrid, 1990, p. 87.

<sup>2</sup> Von Hayek, “Notes on the Evolution of the Systems of Rules of Conduct”, *Studies in Philosophy, Politics and Economics*, Citado en E. Lamo de Espinosa, *op. cit.*, p. 8.

no *explicar* regularidades típicas del agregado<sup>3</sup>, pues éstas no dependen sólo de las acciones. Del análisis, concluye E. Lamo<sup>4</sup>, se debe distinguir radicalmente el orden normativo del social. No cabe en todo caso señalar una situación determinística, en la medida que existen posibilidades de cambio en los dos niveles individual y agregado. Es importante la distinción entre la acción normativa (basada en el debe hacer/ debe ser) y la de tipo probabilístico que se postula para la regularidad no normativa. Las regularidades sociales, en términos probabilísticos es lo que se denominaría “ley sociológica”. Así, “(la sociología) necesita que haya hechos sociales; pero si además quiere ser ciencia *nomotética* necesita que haya leyes sociales”<sup>5</sup>. No debe en todo caso, confundirse la noción de ley con la de teoría. El rasgo diferencial entre leyes y teorías se encuentra en que las leyes hacen referencia a las características “empíricas” de los fenómenos, o sea, a aspectos observables y no a “conceptos teóricos” o abstractos.

Tal y como señala M. Navarro, la aproximación desde el constructivismo afirma la naturaleza no determinista en la constitución de cualquier regularidad social al mismo tiempo que destaca el papel contingente e histórico de las construcciones sociales. “El constructivismo social que inicia la sociología weberiana parte, pues, de acciones sociales individuales entrelazadas, más o menos permanentes o transitorias. Cualquier relación social puede quebrarse en todo momento y por ello, sólo podemos afirmar la regularidad en términos de probabilidad de conducta y no de normas: “la relación social *consiste* sólo y exclusivamente –aunque se trate de formaciones sociales como “Estado, “iglesia”, “corporación”, “matrimonio”, etc.- en la probabilidad de que una forma determinada de conducta social, de carácter recíproco por su sentido, haya existido, existe o pueda existir”<sup>6</sup>”.

---

<sup>3</sup> K. Popper, *Rationality and the Status of the Rationality Principle* (1967), citado por J.C. Zapatero en “K. Popper y la metodología de las ciencias sociales”, en *Cuadernos Económicos de ICE*, 3-4 (1977), p. 103.

<sup>4</sup> E. Lamo de Espinosa, *La sociedad reflexiva*, *op. cit.*, pp. 88-91.

<sup>5</sup> *Ibíd.*, p. 95.

<sup>6</sup> Max Weber, *Economía y sociedad*. México, Fondo de Cultura Económica, 1969, p.22, citado por M. Navarro, "Apuntes para una teoría de la cultura económica", en VV. AA. *Escritos de Teoría Sociológica en homenaje a Luis Rodríguez Zúñiga*, Madrid, CIS, 1992, p.792

Las regularidades que estamos considerando serían en todo caso, agregadas y probabilísticas. Wilfredo Pareto, ilustre antecedente de la sociología matemática destacaba los problemas en la detección de regularidades: Las distorsiones que se aprecian en las diferentes regularidades son producto de los errores y limitaciones propias del ser humano. En ese sentido, afirmaba que “(7) Hablando propiamente no puede haber excepciones a las leyes económicas y sociológicas, en la misma forma que las otras leyes científicas. Una uniformidad no uniforme no tiene sentido. Pero las leyes científicas no tiene una existencia objetiva. La imperfección de nuestro espíritu no nos permite considerar los fenómenos en su conjunto y estamos obligados a estudiarlos separadamente. En consecuencia, en lugar de uniformidades generales que están y que quedarán siempre ignoradas, estamos obligados a considerar un número infinito de uniformidades parciales, que crecen, se superponen y se oponen de mil maneras. Cuando consideramos una de esas uniformidades, y que sus efectos son modificados u ocultos por los efectos de otras uniformidades, que no tenemos la intención de considerar, decimos de ordinario, pero la expresión es impropia, que la uniformidad o la ley considerada sufre de excepciones. Si es admitida esta forma de hablar, las leyes físicas, y aún las matemáticas, comportan excepciones, lo mismo que las leyes económicas. (...) (8) Una ley o una uniformidad no es verdadera sino bajo ciertas condiciones, que nos sirven precisamente para indicar cuáles son los fenómenos que queremos destacar del conjunto. Por ejemplo, las leyes químicas que dependen de la afinidad son diferentes, según que la temperatura se mantenga en ciertos límites, o los sobrepase. Hasta una cierta temperatura los cuerpos no se combinan; más allá de esa temperatura se combinan, pero si aumenta todavía más allá de cierto límite se disocian. (9) Esas condiciones son unas implícitas y otras explícitas. No se debe hacer entrar entre las primeras más que las que son entendidas fácilmente por todos y sin el menor equívoco; si no eso sería un jeroglífico y no un teorema científico. No hay proposición que no se pueda certificar como verdadera bajo ciertas condiciones a determinar. Las condiciones de un fenómeno son parte integrante de ese fenómeno, y no pueden ser separadas. (10) No conocemos, ni podremos jamás conocer, un fenómeno concreto en todos sus detalles; siempre hay un residuo. (...). (11) Puesto que no conocemos enteramente ningún fenómeno concreto, nuestras teorías de esos fenómenos no son más que aproximadas. (...)”<sup>7</sup>.

Destaca en la noción de Pareto que el fenómeno es indisociable de sus condiciones de realización pero también que un mismo proceso puede generar regularidades distintas, incluso contradictorias. En ese sentido, M. Navarro destaca como la progresión de los esquemas de decisión instrumental adquieren una mayor complejidad conformen se difunden como culturalmente adecuados para la toma de decisiones. “El hecho sociológico básico que hay

---

<sup>7</sup> W. Pareto, *Manual de economía política*, Genève, Librairie Droz, 1964.

que resaltar y que guía toda la concatenación de fenómenos que se han expuesto hasta este momento, es la aparición en nuestras sociedades de conductas económicas definidas, que han tenido, a su vez, influencia en la aparición de un complejo de conductas no económicas, todas ellas con un carácter masivo y que se hacen más centrales, en la vida de los hombres, que tienen un carácter abstracto y complejo, que requieren un mayor número de conocimientos y que ofrecen una previsible perspectiva de aumentar el número de complejidad y de generalizarse socialmente”<sup>8</sup>.

Para M. Navarro, la difusión de un esquema de actuación como es el desarrollado dentro de la cultura económica adquiere una generalización y complejidad cada vez mayor, lo que puede conducir a fenómenos sociales contradictorios. “La vida social es tan compleja que las mismas fuerzas que llevan al mercado y a la democracia desarrollan otras, en alguna medida contrarias, como la burocracia. El desarrollo del mercado y la competencia, genera, monopolio y burocracia, regulaciones, controles y salvaguardias. La aparición y la evolución del capitalismo han propiciado estos fenómenos. Pueden no ser necesarios, pero se han manifestado de ese modo”<sup>9</sup>. Esta complejidad por la que una dinámica social puede asociar y provocar de algún modo el desarrollo de otras contrarias es lo que contribuye a hacer borrosas determinadas regularidades sociales.

La noción misma de “regularidad” puede aparecer denominada como “ley social” en un sentido “blando”. Otros autores como J. Elster prefieren matizar sustituyendo la palabra “ley” por “mecanismo” “(...) debilidad de la teoría más famosa de la explicación científica, la propuesta por Carl Hempel. El sostiene que la explicación equivale a la deducción lógica del acontecimiento a explicar, con leyes generales y declaraciones de las condiciones iniciales como las premisas. Una objeción es que las leyes generales pueden expresar correlación pero no causa. Otra es que las leyes, aunque sean genuinamente causales pueden ser anticipadas por otros mecanismos. Es por eso que aquí he puesto el acento en los mecanismos y no en las leyes. Esto no es un

---

<sup>8</sup> M. Navarro, "Apuntes para una teoría de la cultura económica", en VV. AA. *Escritos de Teoría Sociológica en homenaje a Luis Rodríguez Zúñiga*, Madrid, CIS, 1992, p. 786.

<sup>9</sup> *Ibíd.* p.793.

profundo desacuerdo filosófico. Un mecanismo causal tiene un número finito de eslabones. Cada eslabón se debe describir mediante una ley general y en ese sentido por una 'caja negra' acerca de cuyos engranajes internos permanecemos en la ignorancia. Pero para los fines prácticos (los fines del científico social en acción) es importante el lugar del acento dinámico de la explicación científica: el impulso a producir explicaciones cada vez más finas"<sup>10</sup>. Esta matización se corresponde en un sentido global con los planteamientos de Pareto y responde a la dinámica entre generalización y detalle en la explicación para alcanzar una mejor determinación de las regularidades.

Otro aspecto interesante procede desde el ámbito de la generalización y lo local. En este contexto reaparece el debate sobre la utilidad local de las generalizaciones. Nuevamente la dicotomía entre sociología como ciencia o tecnología. P. Von Morpurgo, reflexiona como<sup>11</sup> "El carácter de las ciencias sociales está íntimamente vinculado a su reivindicación de universalidad. Si se puede determinar la regularidad de algunos fenómenos, y por consiguiente establecer leyes que tengan un alcance casi universal, esa reivindicación se puede aceptar. No obstante, con una vasta aplicación de este principio se corre el riesgo de ignorar la diversidad local. En efecto, en estos últimos cincuenta años, las ciencias sociales y sus repercusiones han cobrado mayor fuerza gracias a la mejora de sus métodos e instrumentos de investigación aplicables a escala reducida. Es evidente que, en materia de elaboración de políticas, los resultados y recomendaciones adecuadamente detallados y relacionados con un entorno específico son más pertinentes que las vastas generalizaciones". No parece, sin embargo, que ambas tareas sean excluyentes, sino más bien complementarias.

La característica distintiva en la sociología aplicada es el empleo de métodos formales para la construcción de modelos en el análisis de los fenómenos sociales. Así, T. Fararo en la introducción de su manual afirma como<sup>12</sup> "este libro pretende ser una contribución a acelerar la formalización de las teorías así

---

<sup>10</sup> J. Elster, *Tuercas y tornillos* Barcelona, Gedisa. 1996 p. 16

<sup>11</sup> P. von Morpurgo, *op. cit.*,

<sup>12</sup> T. Fararo, *Mathematical Sociology: an introduction to fundamentals*. New York, Wiley, 1973. p.15

como a restituir la capacidad explicativa de la teoría mediante el empleo de modelos”. Los dos conceptos clave son formalización y representación mediante modelos. La actividad de la sociología matemática no se dirige a una labor de formalización *per se* de la teoría sociológica, sino que contribuye a ello mediante el desarrollo de modelos formalizados, generalmente de carácter matemático. La observación es importante, dado que la función del análisis lógico se circunscribe a la elaboración de los modelos. Es evidente que el empleo y desarrollo de modelos dentro de una teoría ayuda a su formalización y potencia la consistencia interna. Esta tarea constituye una aproximación metodológica bastante definida a la realidad social.

En ese sentido, cabe recordar como una de las características del método científico es, según Hanson, la búsqueda de un modelo en el que insertar los datos: “En una disciplina de búsqueda en crecimiento, la investigación se dirige, no a reordenar viejos hechos y explicaciones en modelos formales más elegantes, sino más bien al descubrimiento de nuevos esquemas de explicación”<sup>13</sup>. A este respecto, la misma opinión sostiene Allais “Cuando se analizan los fenómenos sociales sobre todo los económicos, se revela la existencia de regularidades tan sorprendentes como las que encontramos en las ciencias físicas (...). Toda ciencia se basa en modelos, ya sean descriptivos o explicativos o estén destinados a hacer pronóstico o a tomar decisiones”<sup>14</sup>. En ese sentido, la noción de representación es uno de los núcleos importantes en la actividad de la sociología aplicada.

El estudio de modelos estructurales lleva asociado habitualmente el penetrar en el futuro, ya sea a efectos de control metodológico (como parte de un diagnóstico del ajuste), o como finalidad del modelado. Por ello, el modelar implica el descubrir pautas regulares y sistemáticas en los datos, de modo que nos permitan construir un modelo matemático que explique las relaciones conocidas entre los fenómenos sociales. A partir de ello, se intentará predecir el comportamiento futuro más probable en el fenómeno social.

---

<sup>13</sup> Citado en T.A. Sebeok “One, Two, Three...Uberty” en U. Eco y T.A. Sebeok (eds.) *El signo de los tres*. Dupin, Holmes, Peirce. Barcelona, Lumen. 1989. p.74

<sup>14</sup> M. Allais, “La pasión por la investigación” en M. Szenberg (ed) *Grandes economistas de Hoy*, Madrid, Debate, 1994.

Sin embargo, en este escenario de modelado, para que el futuro sea predecible debe de ser semejante al pasado en lo que se refiere a los factores que generan el comportamiento de la serie. Es decir, las reglas que rigen el cambio deben mantenerse estables en el futuro, dado que el "mecanismo" modelado se pone a funcionar para generar los valores futuros de la variable. Por lo tanto, cualquier cambio brusco en el medio donde se desenvuelve el "mecanismo", o en los elementos que componen el "mecanismo" podrán inducir a error. Podemos pensar que la eficacia del modelo dependerá, por lo tanto, de la "profundidad" en que estén operando sus reglas en el sistema social, por un lado, y del contexto "evolucionario" o "revolucionario" en el que se desarrolla el fenómeno social en estudio.

Un modelado elaborado sobre una realidad social epidérmica (epifenoménica), determinara un "mecanismo" con predicciones excesivamente contingentes. La eficacia del modelado de una estructura de covarianzas dependerá más de la reflexión sociológica que le concede un ámbito de existencia, determinando los "mecanismos" a revelar, que de las técnicas instrumentadas para acceder a ellos. En otro sentido, el proceso o fenómeno social en estudio puede desenvolverse en un entorno social en equilibrio dinámico o en profunda y rápida transformación. La posible existencia de discontinuidades (una guerra civil, una revolución, un golpe de estado, etc.) matiza las probabilidades futuras de realización del fenómenos social modelado. De este modo, un conjunto relacionado de variables o indicadores se modela según sus propias características, pero su futuro se evalúa por sus circunstancias o contexto.

La investigación sociológica recorrió un camino importante desde los diseños correlacionales, el análisis de regresión y el "análisis de senderos". A principios de la década de los 70 quedaba claro que desde diferentes disciplinas se estaba considerando un problema parecido. Así, el análisis de senderos en sociología, los modelos de ecuaciones estructurales en econometría o los modelos de análisis factorial en psicología respondían a un mismo intento de modelado. Esta aproximación se consolidaría en una estrategia general de modelado denominado modelos estructurales de covarianzas. Los modelos estructurales se emplean tanto en la medición como en la explicación. Ambas

dimensiones pueden combinarse, de modo que simultáneamente a la realización de una medición (por ejemplo, de variables latentes) se puede estructurar un diseño explicativo. Los denominados modelos causales son un área de actividad dentro de este campo. Es importante contextualizar el concepto de explicación (o causalidad) a un doble nivel, epistemológico y operativo.

Esta dinámica se vería muy influenciada por el impacto “operativo” de los nuevos métodos de computación, que ayudarían a generar modelos matemáticos de elevada sofisticación, como los desarrollados en la Escuela de Wisconsin. El motor del proceso no fue un salto cualitativo en los desarrollos de la sociología que exigiese de expresión matemática, sino un impacto tecnológico que operando sobre los procedimientos (que no sobre la teoría) intento arrastrar a ésta tras sí, enganchada en el atractivo de los modelos matemáticos.

Como podremos evaluar más adelante, la investigación primero nombra y después numera. Cuando existen números sin nombre (como en el análisis factorial exploratorio o modelos estructurales de medición) la desorientación está servida y el resultado queda disponible a las veleidades argumentales del investigador. De este modo, los desarrollos de la computación (numerología) sólo podía potenciar los procedimientos e indirectamente, el método “positivista” que se articula sobre dichos procedimientos. Indirectamente, en la medida que los nuevos procedimientos no se vertebran claramente sobre él; de hecho, en términos funcionalistas, los procedimientos son denominados métodos, pero muy raramente se encontrará una articulación explícita entre investigación y teoría.

La tendencia iniciada en los años 70 iniciaba una autonomización del método (en los términos anteriores), que conduciría no ya al contraste de teorías (realmente difícil) sino de “resultados”. Es el refugio en el metodologismo que Coser (1975) denunciará y que Hope (1981<sup>a</sup>), recuperando a Lakatos, afirmaría como proceso ritual “El gran error de la construcción de modelos reside en concentrarse en cuestiones de ajuste y el examen de las desviaciones del modelo, con la exclusión de la investigación del significado del modelo en sí mismo, de manera que éste no constituye ya un instrumento de progreso, sino más bien el centro



ritual de un proceso degenerativo”. No eran, en general, ideas sociológicas servidas con técnicas estadísticas, sino desarrollos estadísticos ilustrados con una sociología simplificada. La sociología matemática entendida en ese modo entró en crisis, evidentemente. Ya desde principios de los años 70 (Golthorpe, Hope, Bertaux y Kreckel en el coloquio de Constanza) y especialmente en su segunda mitad, surgen fuertes críticas que cuestionan el ateoricismo en que se mueven estos metodólogos y la no significatividad de los resultados obtenidos mediante modelos matemáticos muy elaborados sobre ideas sociológicas rudimentarias.

Destaca especialmente el discurso de Coser (1975) como presidente de la Asociación Americana de Sociología, donde crítica duramente la sustitución de ideas por números. En esa línea crítica aparecen varias aportaciones, como son los artículos de Crowder (1974), Bertaux (1967<sup>a</sup>), Burawoy (1977) y Horan (1978). Como ya comentábamos, la eclosión de modelos producidos por la primera “sociología matemática” era esencialmente una explosión inducida desde el desarrollo de otros ámbitos (estadística, computación, etc.) y que en esas condiciones, tendía a carecer de intereses propios y contenido.

## **2. Describir, explicar, comprender**

Previamente a desarrollar en profundidad el concepto operativo de causalidad en el entorno del modelado estructural es importante destacar la lógica subyacente a este tipo de modelos. La noción de explicación causal se encuentra directamente articulada sobre la noción de necesidad o necesario en el contexto del discurso científico. La noción de necesario se interpreta no en el sentido de que algo debe necesariamente acontecer, sino más bien en la idea de que algo que acontece debe de tener necesariamente una explicación. El conocimiento científico sólo reconoce como campo de actuación el de aquellos acontecimientos susceptibles de una explicación racional, es decir, que tienen una razón para suceder. Las cosas suceden por algún motivo, son causadas por algo. Esta idea que es central en el pensamiento de tipo científico, se encuentra en la base de la argumentación de los modelos de covarianzas. En definitiva la pregunta es ¿qué variables pueden dar cuenta de los cambios que

estamos apreciando en otras variables? En ese sentido es destacable como la idea de necesidad no esta referida a la obligatoriedad de que algo suceda, sino a que exista una explicación para ello. En ese sentido, los modelos estructurales de covarianzas pretenden explicar y no sólo describir.

Émile Durkheim (1858-1917) presenta, una estrecha relación con la sociología matemática, y en especial con los modelos estructurales, tanto en el método como en la identificación del objeto. Un ejemplo notable de ello es su análisis del suicidio. Se han hecho críticas muy válidas de los errores concretos de interpretación y las inferencias erróneas (en particular su recurso constante a la falacia ecológica<sup>15</sup>; por otra parte, fue uno de los primeros en usar el análisis multivariable, al igual que las correspondencias externas e internas, y muchos de sus resultados se han visto ampliamente confirmados por la investigación posterior. H. C. Selvin, sociólogo matemático, afirma como, especialmente en el sistema de argumentación y empleo de una aproximación multivariante anticipadora del análisis causal: “Son pocos, si los hay, los trabajos posteriores que pueden igualar la claridad y potencia con que Durkheim ordenó sus datos para comprobar y afinar su teoría. La grandeza de su trabajo es aún más impresionante si se tiene en cuenta que Durkheim careció incluso de un instrumento tan rudimentario como el coeficiente de correlación. Con todo, la metodología de *El suicidio* sigue siendo de gran importancia para todos aquellos que actualmente se dedican a la investigación empírica, y no sólo para los historiadores de la sociología. Durkheim señaló y resolvió muchos de los problemas que se plantea la investigación hoy en día<sup>16</sup>”.

En ese sentido, se encuentran en Durkheim tres aspectos centrales: la búsqueda de regularidades, el empleo de métodos formales y su aproximación empírica. Tal y como afirmaba en su crítica a Jankelevitch, “He aquí otro libro más de generalidades filosóficas sobre la naturaleza de la sociedad, y de generalidades a través de las cuales resulta difícil percibir un conocimiento práctico, íntimo y familiar de la realidad social. En parte alguna da el autor impresión de haber entrado en contacto directo con los hechos de que habla,

---

<sup>15</sup> H.C. Selvin, “Durkheim’s Suicide and problems of empirical search”. *AJS*, 62, 1958, pp. 607-19. En S. Lukes, *Émile Durkheim, su vida y su obra*, CIS, Madrid, 1984, pp. 204-205.

<sup>16</sup> *Ibid.*, pp. 204-205.

ya que las ideas generales que desarrolla no aparecen ilustradas por ningún ejemplo concreto, ni aplicadas a ningún problema sociológico concreto. Por grande que pueda ser el talento dialéctico y literario de los autores, nunca se protestará lo bastante del escándalo de un método que ofende tanto, como lo hace éste a nuestras prácticas científicas, y que, no obstante, es todavía empleado con harta frecuencia. Hoy ya no está permitido especular sobre la naturaleza de la vida, sin antes haberse familiarizado con las técnicas de la biología. ¿En virtud de qué privilegio se ha de permitir que el filósofo especule acerca de la sociedad, sin entrar en contacto con los hechos sociales concretos?”<sup>17</sup>.

Por otra parte, Durkheim aplicaba en su trabajo como sociólogo los métodos que predicaba. Como aprecia el testimonio de un asistente a sus conferencias “El señor Durkheim habla un lenguaje claro, preciso y, dentro de lo posible, técnico. Usa símbolos concretos y especificaciones numéricas. Si por ejemplo está disertando sobre la familia, emplea un diagrama que muestra las fases sucesivas por las que ha atravesado dicho grupo social, indicándolas con el dedo; si lo que estudia es el suicidio, escribe en la pizarra las cifras proporcionadas por las estadísticas oficiales. En resumen, sea cual fuere su tema, salta a la vista su esfuerzo por expresarse con claridad. Es inútil esperar de sus discursos exhibiciones literarias o parrafadas retóricas, llamadas al sentimiento o extrapolaciones metafísicas, actitudes a las que tan bien parecen prestarse los temas sociológicos. Es obvio que el señor Durkheim huye de las “frases”, de las generalidades y de las grandes síntesis; que no quiere generar “problemas” ni construir sistemas. Observa los hechos, los analiza y los explica mediante leyes. En una palabra, lo que hace no es tanto llegar a conclusiones como registrar resultados y expresarlos en proposiciones que califica de “fórmulas” en virtud de su precisión y de su rigor. Por supuesto que el señor Durkheim es siempre elocuente y con frecuencia hasta extremos insospechados. Cualquiera que le haya oído no olvidará fácilmente el irresistible entusiasmo que es capaz de levantar en su auditoria cuando quiere. Pero incluso en esos momentos lo que mueve a estos oyentes es el enorme

---

<sup>17</sup> *L'Année Sociologique*, vol. X. Reseñas: Jankelevitch Nature et société, pp. 171-174. Citado por S. Lukes, *op., cit.*, p. 401.

interés de las ideas expresadas; lo que les excita es la intensidad del análisis, conciso, apremiante y tendente siempre hacia la constatación; lo que les vence y abruma es la velocidad de la conferencia que, concentrada y concisa, aunque prolija, está como impaciente por llegar a su fin [...] Escuchando al señor Durkheim se tiene la clara sensación de que no pretende persuadir sino convencer, de que se dirige menos a los sentimientos que a la razón humana”<sup>18</sup>.

Max Weber (1864-1920), como ha sido frecuentemente citado, no siempre respetó sus propuestas metodológicas en la práctica de la investigación. Así, aparecen referencias directas al empleo de leyes en el análisis de los fenómenos sociológicos en *El político y el científico*. Amparo Almarcha destaca “En la formación de una metodología específica de las Ciencias Sociales, M. Weber da un gran paso al definir los conceptos en términos de ‘probabilidad’ y de asentar, en consecuencia, las leyes sociales sobre la base de la entonces recién creada teoría de la probabilidad”<sup>19</sup>. No es sin embargo, una contribución específica a la sociología matemática europea, que habitualmente no considera la estadística como un elemento central en su historia. Esta es una diferencia interesante con la economía matemática, donde los métodos estadísticos multivariantes, incluyendo las ecuaciones estructurales ocupan una posición privilegiada.

Como destacará Paul Lazarsfeld (1964) “Resulta curiosa la historia de los métodos empíricos que acabo de mencionar. Todos ellos fueron elaborados al principio en Europa Occidental, pero las universidades de este continente nunca los reconocieron verdaderamente. A principios de siglo llegaron a ser preponderantes en los Estados Unidos, y de aquí se volvieron a importar a Europa, sobre todo después de la Segunda Guerra Mundial”<sup>20</sup>.

---

<sup>18</sup> Delprat, G. “L’enseignement sociologique a la L’Université de Bordeaux”, *Revue Philomatique de Bordeaux et du Sud-Ouest*, 3er año, agosto de 1900, p. 357. Citado por S. Lukes *Émile Durkheim. Su vida y su obra*. Madrid: CIS, 1984, pp. 104-105.

<sup>19</sup> A. Almarcha, A. de Miguel, J. de Miguel y J. L. Romero, *La documentación y organización de los datos en la investigación sociológica*, Madrid, Confederación Española de Cajas de Ahorros, 1969, p. 16.

<sup>20</sup> P. F. Lazarsfeld, “La investigación social empírica y las relaciones interdisciplinarias”, *Revista Internacional de Ciencias Sociales*, vol. XVI, (1964), p. 620.

Theodor Adorno (1973) planteaba la necesidad, desde el punto de vista de la investigación de comprender. La explicación solamente es útil cuando se orienta a la comprensión, no puede ser un fin en sí misma. La explicación es una vía para la comprensión. Explicar para comprender. En definitiva, saber lo que los números significan evitan la ceguera de la explicación “per se”. Adorno (1973) planteaba “Así es como deben ser entendida las consideraciones de Habermas en torno a las leyes históricas de la evolución, es decir, en el contexto de la inminente determinación objetiva de lo particular mismo. La teoría dialéctica se niega a contrastar el conocimiento histórico y social como conocimiento de lo individual, ajeno al conocimiento de las leyes, dado que lo que se considera exclusivamente individual (la individualización es una categoría social) lleva cruzados dentro de sí algo singular y universal; incluso la necesaria distinción entre ambos tienen carácter de una falsa abstracción. Las tendencias evolutivas de la sociedad, como la tendencia a la concentración, a la superacumulación y a la crisis, son modelos del proceso de lo general y singular. Hace ya mucho tiempo que la sociología empírica se ha dado cuenta de lo que pierde en contenido específico por culpa de la generalización estadística. A menudo viene contenido en un detalle algo decisivo desde el punto de vista de lo general y se escapa a la mera generalización. De ahí la fundamental necesidad de completar los sondeos estadísticos por medio de “estudios de casos”. El objetivo de los métodos sociológicos cuantitativos debería ser, asimismo, la comprensión cualitativa; la cuantificación no es un fin en sí misma, sino un medio para dicho fin. Los propios técnicos estadísticos se inclinan de mejor grado a reconocerlo que la lógica usual de las ciencias sociales.”

En palabras de Coleman<sup>21</sup> “Con demasiada frecuencia se considera la sociología matemática como un fin en sí misma, en vez de insertarse en el contexto de los problemas que tradicionalmente han preocupado a la sociología. Las matemáticas son una herramienta, y si quiere ser una herramienta útil para la sociología, debe ser usada reconociendo el estado en que la sociología se encuentra en la actualidad”. Las llamadas a la sensatez efectuadas por muchos de los sociólogos matemáticos del momento no

---

<sup>21</sup> J. S. Coleman, *Introduction...op.*, cit., p. 4.

siempre se tenían en cuenta, produciéndose algunos excesos evidentes. Estos excesos, encontraron eco de forma totalmente explícita en Coser. Especialmente destacan los comentarios del discurso presidencial efectuado por el Lewis Coser en la Asociación Americana de Sociología en 1975<sup>22</sup>.

No obstante, la sociología no debería de avanzar exclusivamente en las áreas donde son factibles los datos y las mediciones precisas. De hecho observaciones cualitativas pueden generar ideas que conduzcan a desarrollos teóricos que terminen en mediciones y tratamientos estadísticos más refinados.

Limitar las investigaciones que se apoyan sobre datos descriptivos porque se expresaran simplemente mediante tabulaciones restringe seriamente la potencia teórica de los sociólogos. La amenaza es el educar a los nuevos sociólogos a no preocuparse de aquellos problemas donde es difícil conseguir datos y concentrarse en aquellos donde sí lo es. La consecuencia para la sociología sería producir una especie de visión de túnel por el cual algunos problemas sociales serían estudiados exhaustivamente mientras que otros prácticamente no serían percibidos. Por otro lado, Coser advertía como la fascinación por las nuevas herramientas de investigación, como era el caso de los ordenadores, así como los nuevos procedimientos y técnicas que permiten una mayor precisión podían conducir a olvidar que las mediciones son simplemente un medio para alcanzar una mejor explicación y análisis de los fenómenos sociales.

En resumen, si los conceptos y las nociones teóricas son débiles ninguna medición por precisa que sea incrementará la capacidad explicativa de la ciencia. La falacia de la precisión, según Coser, consiste en creer que es factible compensar las debilidades teóricas gracias a la fortaleza metodológica. Así preocuparse con la precisión en las mediciones antes de clarificar teóricamente lo qué vale la pena medir y lo qué no, así como tener claro que lo que uno está midiendo es un camino cerrado en el progreso de cualquier análisis sociológico. En ese sentido Coser indica estar de acuerdo con el

---

<sup>22</sup> L. A. Coser, "Presidential Address: Two methods in search of a Substance" *American Sociological Review*. December, vol.40 n°6, (1975), pp. 691-700.

empleo de las nuevas tecnologías en la investigación social, pero que su mayor preocupación procede de los abusos de sus instrumentos.

Especialmente cuando se aplican indiscriminadamente y la elección de los problemas para investigar dependen de las técnicas y procedimientos disponibles, siguiendo la ley del instrumento formulada por Kaplan<sup>23</sup> cuando afirma “Dadle a un niño un martillo y encontrará que todo lo que está a su alcance necesita ser golpeado”. El hecho que destaca Coser, es que si bien en principio los nuevos métodos y tecnologías pueden ayudar a alcanzar una mayor sofisticación teórica, se les tiende a utilizar como “ayudantes mágicos”, o como “atajos” para producir teoría.

El planteamiento que mantendremos aquí es la subordinación del lenguaje matemático a la investigación sociológica, empleando sus potencialidades para arrojar luz sobre aquellos aspectos que la sociología desde dentro no acaba de definir. Al mismo tiempo, hay que considerar que surgen nuevas matemáticas que tienden a seducir a los sociólogos, siempre en busca de una solución externa a sus propios problemas. No parece un camino sensato, en la medida que toda matemática y estadística contiene una definición del mundo oculta en sus presupuestos. Deberá ser la sociología la que invente sus propios bagajes matemáticos. Al igual que en las ciencias naturales los cambios de explicación del mundo generan nuevas matemáticas para hablar sobre ellos (así como se entretienen inventando matemáticas para especular mundos) la sociología debe definir sus mundos e imponer sus presunciones sobre el lenguaje que se emplee para hablar de ellos.

### **3. Elementos de los modelos estructurales**

Los modelos causales se apoyan tanto sobre nociones epistemológicas, como metodológicas y de procedimiento matemático. Expondremos, pues, aquellos aspectos básicos que deben considerarse en el desarrollo de modelos matemáticos.

---

<sup>23</sup> A. Kaplan, *The Conduct of Inquiry* San Francisco, Chandler 1964, p. 28.

### **3.1. Los conceptos teóricos**

Rara vez los conceptos teóricos representan un fenómeno social unidimensional. La dimensionalidad se refiere al número de cualidades distintas que son inherentes a un concepto teórico. Por ejemplo, “status social”, “alienación”, o “prejuicio” son conceptos teóricos que engloban aspectos bastantes diferentes. Así, en el concepto “prejuicio” se incorporan aspectos tales como “tendencia a discriminar”, “presencia de sentimientos negativos”, o “el contenido de estereotipo” que pueda existir en el prejuicio. En general, la caracterización multidimensional de los conceptos teóricos contribuyen al desarrollo de la ciencia. Esto, porque al presentar diferentes lecturas de un concepto teórico se favorece la formulación de hipótesis a la vez que se indican posibles vías de acceso a su medición. En el caso de los prejuicios contra los gitanos, se podría medir este mediante las creencias despectivas que pudiesen existir, o por el rehuir a relacionarse con personas de esta raza, lo que ofrece un mayor repertorio de estrategias de medición.

Un aspecto reconocido por la mayoría de los científicos sociales es el hecho de que es necesaria una formulación más estricta de los conceptos teóricos. En palabras de Blalock: “la cuidadosa reelaboración de las teorías verbales constituye una de las tareas más desafiantes con que nos enfrentamos. La principal tarea de esta empresa consistirá en clasificar conceptos, eliminar o consolidar variables, traducir las teorías verbales existentes a un lenguaje común, investigar la bibliografía a la búsqueda de proposiciones, y detectar las presunciones que conectan las principales proposiciones con el trabajo teórico serio”. Se precisa, pues, de una tarea previa de formalización que permita a la encuesta muestral el tener una mayor eficacia teórica de la que posee actualmente. Este esfuerzo desborda, con mucho, las tareas estrictamente metodológicas. Tal y como lo expresa Lazarsfeld (1951) “¿Cómo nos ponemos a formar tales categorías, en primer lugar? ¿Por qué escoger ciertos elementos de la situación, y no otros? ¿Por qué combinarlos precisamente en estas categorías? Se puede argüir acertadamente que no podemos redactar un conjunto de instrucciones manuales para categorizar los fenómenos sociales: tales instrucciones no serían más que un programa general para desarrollar la



teoría social. No podemos escribir un manual sobre cómo formar fecundos conceptos teóricos de la misma manera que escribimos manuales sobre cómo seleccionar muestras o redactar cuestionarios”. Este es un reconocimiento a la evidencia de que el desarrollo de la investigación social sólo puede ser una consecuencia de un desarrollo previo de la teoría sociológica.

La investigación social es siempre investigación de algo, y ese algo debe de estar definido, desde una matriz teórica previa que identifique los límites lógicos y consecuencias teóricas de aquello que se evidencie empíricamente. Sin embargo, existe un desarrollo técnico de la investigación social que no se relaciona directamente con una eficacia teórica, dado que la teoría sociológica ha padecido siempre de una escasa formalización de conceptos, planteando más imágenes globales que detalles particulares, lo que constituye un obstáculo a su cuantificación. La investigación social es posible actualmente porque el investigador participa de la sociedad que estudia, de modo que posee un conocimiento, en gran parte intuitivo, de las características que son relevantes. Cicourel (1982) lo expresa claramente: “la falta de una teoría social desarrollada obliga a todos los investigadores en sociología a emplear conceptos vulgares que reflejan los conocimientos comunes a los sociólogos y a los miembros “medios” de la comunidad o sociedad. Suponiendo desde el principio que el sociólogo y sus sujetos constituyen una cultura común que cada uno entiende más o menos de la misma manera, los sentidos “obvios” de las preguntas operativizadas del cuestionario en que se basan los indicadores incorporarán propiedades sólo vagamente definidas en la teoría social, pero cuya importancia para el proyecto de investigación se da por supuesta”. Hay que aceptar, en parte, la crítica de Cicourel en el sentido de que se debe de mejorar la definición teórica de los conceptos sociológicos, mostrando aquellos aspectos que puedan orientar una cuantificación más precisa. No obstante, el compartir una cultura común es una ventaja indudable y un punto de partida obvio para la cuantificación de conceptos teóricos.

Es decir, está claro que se debe desarrollar la medición de los conceptos teóricos desde la teoría que los genera, pero también es cierto que no nos encontramos en un camino equivocado, sino sólo al comienzo de la vía

correcta. Utilizar como punto de partida un conocimiento vulgar, en tanto que socialmente compartido, es lo que permite que la investigación social recoja actualmente una parte, mayor o menor, de la realidad de los fenómenos sociales. Ciertamente este conocimiento no está poseyendo una gran trascendencia teórica, pero sí la está teniendo práctica, tal y como demuestran los gastos cada vez más crecientes en investigación de mercados. Recordemos que el origen de la encuesta no se dio en el ámbito universitario, ni en el de las instituciones científicas, sino en la vida pública moderna, en la necesidad urgente de conocimientos prácticos, demandados por un sistema económico competitivo y un régimen político democrático.

Se hace necesario un nuevo estilo de teoría sociológica capaz de crear nuevas estrategias de medición. Para que esto sea posible hay que hacer operativos los conceptos; es decir expresarlos de modo que sean mensurables.

Una definición operativa consiste en la descripción detallada de los procedimientos utilizados en la investigación para asignar las unidades de análisis a las diferentes categorías de una variable. Esto incluye, en una encuesta, las preguntas efectuadas, que categorías se ofrecen como respuestas (o en todo caso las reglas empleadas para codificar si la pregunta es abierta), las instrucciones completas para recoger los datos, etc. Formalmente debe de recurrirse a definiciones preexistentes del mismo concepto generadas por otros científicos, de modo que las conclusiones sean equiparables. No obstante, está abierta la vía a generar nuevas definiciones operativas siempre y cuando tengan en consideración las ya existentes.

En ese sentido, la interpretación de un concepto teórico debe considerar como una de las características más destacables de la investigación científica es el estar basada sobre el empirismo. Este es un modo de conocer o comprender el mundo que descansa directa o indirectamente en aquello que experimentamos a través de nuestros sentidos: vista, oído, sabor, olor y tacto. La información o los datos son aceptados como tales en la investigación científica en la medida en que puedan ser observados o "sentidos" en algún modo, bajo condiciones específicas por individuos que poseen un aparato sensorial, inteligencia y habilidad normales. Así, un neutrón no es visible directamente, pero puede

observarse indirectamente mediante las fotografías que muestran la trayectoria dejada por los neutrones. Hay que notar que los conceptos “observable” o “empírico” poseen un significado más amplio en la actividad científica que en la filosofía. Para los filósofos, lo observable se aplica a propiedades que puedan ser percibidas directamente por los sentidos, como el color, forma, etc. Para el científico, observable es cualquier cosa que pueda ser asociada a los resultados de una medición percibible. Por ejemplo, una temperatura de 40 grados es considerada observable por los científicos porque aparece asociada a la altura de una columna de mercurio en el termómetro. Para un filósofo no sería un observable porque no existe una percepción sensorial directa de esa magnitud de calor. Por ello, en la actividad científica el empirismo aparece frecuentemente en la forma de observación indirecta, donde los instrumentos se emplean para ayudar y extender la habilidad que posean los científicos para observar. De algún modo, se trata de un empirismo sofisticado, dado que extiende la habilidad humana de observación mediante una “ortopedia” instrumental.

Decir que la actividad científica se basa en el empirismo es afirmar que la única evidencia admisible a favor o en contra de una teoría o hipótesis debe de ser observable, ya sea directa o indirectamente, mediante alguna manifestación tangible. Las implicaciones de este principio son muy amplias. Por ejemplo, que el criterio de autoridad, tradición, revelación, intuición o cualquier otra forma no empírica de conocimiento no puede ser aceptada como evidencia científica. Tal como afirman Katzer, Cook y Crouch (1978) “La evidencia es lo fundamental. Si los resultados de estudios bien efectuados están en desacuerdo con la autoridad, entonces la autoridad puede muy bien haberse equivocado. El procedimiento correcto para que un experto pueda recurrir este veredicto es efectuar otro estudio. O, si dos informes de investigación están en desacuerdo, entonces deben de efectuarse más estudios adicionales hasta resolver la cuestión. La idea es que el conocimiento acerca del mundo se obtiene mejor mediante una observación cuidadosa del mundo, no mediante la contemplación de lo que uno piensa sobre el mundo”.

Esto que actualmente se considera un principio rector del conocimiento científico actual, es enunciado como una regla por Isaac Newton, en sus reglas para la investigación de la naturaleza. “Regla cuarta. En la Física experimental, los teoremas derivados por inducción de los fenómenos, si no se dan presuposiciones contrarias, deben ser tenidos por precisamente o muy aproximadamente ciertos, hasta que aparecen otros fenómenos gracias a los cuales aquellos teoremas alcanzan mayor precisión o son sometidos a excepciones. Así debe hacerse, para que el argumento de la inducción no sea abolido a fuerza de hipótesis”.

El principio de empirismo en la ciencia también significa que la actividad científica debe de limitarse a problemas o cuestiones que puedan ser solventables mediante observación. Así, la determinación de si un problema es tratable científicamente se decide por la pregunta ¿Es un tema que se pueda resolver mediante evidencias empíricas?. De ser la respuesta “no”, su tratamiento está más allá de la ciencia.

Sin embargo, tal y como destacara Adorno (1973), esto no siempre es tan evidente. En gran medida, la imposición de restricciones con respecto a lo que es preguntable en función a la posibilidad o no de responder de forma empírica supone una restricción innecesaria. Especialmente cuando la respuesta empírica pasa por la desagregación en partes del problema, lo que diluye las preguntas posibles “La Ciencia ha de ocuparse exclusivamente de problema solubles. Rara vez, sin embargo, los plantea el material de manera tan concluyente. En este mismo espíritu define Popper el método de las ciencias sociales "como el de las ciencias de la naturaleza". Consistiría en "ensayar intentos de solución para sus problemas (los problemas de los que parte). Se proponen y se critican soluciones. Cuando el intento de solución no resulta asequible a la crítica objetiva, es descartado como algo no científico, aunque quizá, sólo provisionalmente". El concepto de problema que es aquí utilizado no es menos atomístico que el criterio wittgensteiniano de verdad; postula que cuanto caiga legítimamente en el ámbito de la sociología puede ser descompuesto en problemas particulares. A pesar del “sentido común” del que a simple vista parece estar impregnada, esta tesis de Popper se convierte,

rigurosamente interpretada, en una censura inhibitoria del pensamiento científico. Marx no propuso la “solución” de ningún “problema” –ya en la idea misma de “proponer” viene implícita la ficción de consensus como fiador de la verdad en sí-; y, sin embargo, ¿deja de ser por ello “El Capital” ciencia social? En el Contexto de la sociedad, la llamada solución de un problema presupone la existencia de dicho contexto. La panacea del “prueba y error” procede a costa de momentos, de tal modo que una vez eliminados éstos, los problemas son arreglados “ad asum scientiae” y convertidos posiblemente, en pseudoproblemas. La teoría tiene que operar mentalmente a través de las interrelaciones –que la descomposición cartesiana en problemas particulares tiende a suprimir- mediando hacia los hechos. Incluso en determinados casos en los que un intento de solución no resulta sin más asequible a la “crítica objetiva”, como Popper la califica, es decir, a su refutación, puede ocurrir que el problema tenga una importancia central desde el punto de vista de la cosa. Interrogarse entorno a sí por la fuerza de su propia dinámica la sociedad capitalista camina, como Karl Marx enseñaba, hace su desmoronamiento o no, es enunciar una pregunta que únicamente tiene sentido en la medida en la que no se manipulan el propio preguntar: es una de las más importantes de entre todas las que podría plantearse la ciencia social”.

En ese sentido, las preguntas que nos hacemos sobre la sociedad pueden abrir “ventanas” que muestran realidades nuevas. Una pregunta acertada expondrá a la luz una realidad que podría estar oculta hasta ese momento. Por el contrario, una pregunta mal formulada puede ocultar la realidad del campo de visión del investigador. Existe en esos términos una contradicción evidente entre los planteamientos positivistas puros y la reflexión crítica. En ese sentido, tal y como señalara Adorno, la actividad investigadora no se rige exactamente por criterios “formalizadores” y en la práctica la flexibilidad es muy importante, en la medida que el investigador “crea” y define las condiciones que generan “datos”.

Los conceptos teóricos, con los que pensamos el mundo, se plasman desde la perspectiva empírica en un conjunto de variables o indicadores, no necesariamente percibibles directamente. Un modelo de covarianzas plantea,

esencialmente, que existe un sistema de relaciones de interdependencia entre un conjunto de variables. En ese sentido los conceptos importantes para el análisis causal están articulados entre sí. La noción de explicación conlleva la distinción de tipos de variables, así como de las relaciones entre ellas. La complejidad que adopten las relaciones nos lleva a la idea de sistemas y tipos de sistemas. Los modos que empleemos para hablar de dichos sistemas de relaciones nos conducen a las notaciones alternativas para poder expresarlos.

Una de las primeras consecuencias del análisis causal es el refinamiento conceptual de las variables. Estas poseerán una cualidad explicativa especial dependiendo de su posición estructural. El análisis no se limita a las variables manifiestas sino que incluye las variables latentes. De este modo, un modelo puede combinar simultáneamente variables manifiestas y variables latentes, que vendrán definidas por un conjunto de variables identificadoras de la variable latente. Por una parte, en función a su *visibilidad* encontramos las variables manifiestas, las variables latentes y las variables indicadoras que actúan de forma instrumental para determinar los efectos de las variables latentes. Si atendemos al *criterio explicativo*, la principal distinción atañe a las variables que se intenta explicar y aquellas otras que no. Tendremos por lo tanto variables que explican y variables que son explicadas. Estos dos conjuntos de variables no son disjuntos. Esto implica que tendremos variables que explican y a su vez son explicadas.

Las variables exógenas son variables que sólo aportan explicación, pero que ellas mismas no son explicadas dentro del modelo que se propone. En el orden causal implican factores de explicación no explicada. Las variables endógenas intervinientes son aquellas que explican y son explicadas dentro del modelo explicativo que estamos desarrollando. Por último, las variables endógenas no intervinientes son aquellas que vienen explicadas dentro del modelo pero no contribuyen a explicar ninguna otra. Realmente, un modelo estructural consiste en narrar una historia donde se explica algo. Nuestro repertorio de personajes está limitado a los que consideramos dentro del modelo, sus acciones se explicarán entre sí. Habrá personajes (variables) que actúan dentro la trama, pero no sabemos que les dirige o explica. Habrá personajes cuyos actos están

explicados por las acciones de otros y que a su vez influyen en unos terceros. Por último, estarán aquellos que, dentro del ámbito de la historia que intentamos contar, son explicados al máximo en sus actuaciones pero cuyo impacto en los demás no es tenido en cuenta, por esta vez, en la historia que se está contando.

En ese sentido, un modelo causal no deja de ser una narración donde se intenta explicar el por qué de algo, con un número limitado de personajes (variables exógenas, endógenas intervinientes o endógenas). La explicación nunca es completa, ni del conjunto ni de cada personaje endógeno, por lo que restan residuales entre la explicación y los comportamientos que se detectan (varianza de las variables). La lógica explicativa indica que cuanto más residuales peor explicación, cuanto menos residuales, es decir, cuanto mejor se puedan predecir los cambios en el comportamiento de los actores-variables mejor modelo. Podemos pues diferenciar entre descripción, desarrollada mediante la introducción de variables intervinientes, que refinan el modelo y lo hace más detallado, y explicación, que dependerá del grado de ajuste con la realidad de las relaciones que encontramos.

Una noción importante es la de variable común. Esta variable común (A) puede explicar dos variables relacionadas entre sí (BC) sin disminuir su relación. En este caso la variable común valida indirectamente la relación entre B y C. Pero también una variable común a otras puede explicarlas simultáneamente. Consideremos un ejemplo intuitivo: dos personas que se conocen y se relacionan (B y C), pero que ambas son inicialmente amigas de otra tercera persona (A) que los presenta y que casi siempre está cuando los otros dos se encuentran. Puede plantearse una relación de orden entre los dos primeros pero lo evidente es que el conocido común explica porque se reúnen. El número de reuniones entre ambos puede ser elevada (B,C), pero cuando se tiene en cuenta cuantas reuniones se realizan sin presencia del amigo común (A), este número deviene insignificante. Es decir, la reunión (covariación) entre B y C es dependiente de A, y si se controla por ella, la relación entre B y C puede devenir "espuria". Una relación espuria se revela cuando la asociación entre dos variables viene explicada por otra común a las dos. Así, la noción de

causa común es aquella variable explicativa que siendo previa a otras dos que covarían, explica o no la covariación entre estas. Como veremos, la noción de causa común es muy importante para falsar las hipótesis causales.

En resumen, introduciendo el máximo de causas comunes (sean exógenas o endógenas) incrementamos la capacidad explicativa del modelo y su robustez a la falsación, mientras que la introducción de variables endógenas intervinientes no comunes refinan el análisis, haciéndolo más descriptivo.

Una vez que poseemos un lenguaje para expresar ideas y explicaciones, el siguiente paso es contar una historia e intentar compararla con los datos que aportan la realidad (tal y como ha sido seleccionada y definida). Esta fase se denomina de especificación (contar una historia) y ajuste (comparar lo que decimos con los datos de que disponemos). Evidentemente, ninguna explicación es completa ni definitiva. No obstante, la existencia de una realidad (un proceso, una historia, acontecimientos) a explicar se impone inexcusable. En ese sentido, la tarea del modelado, que pudiera parecer absurda, adquiere su verdadero sentido, deviene reveladora en la medida que exige una explicación al mundo, tanto de lo que creemos saber como de lo que no sabemos.

Pasemos pues a introducirnos en el vocabulario y notaciones que nos forzarán a reconstruir lo que incluso damos por evidente.

### **3.2. *La argumentación***

El planteamiento del problema a investigar debería sugerir qué observaciones contribuirían a ofrecer soluciones. Así, en primer lugar, el problema de investigación debería expresar o plantear una pregunta acerca de la relación entre dos o más variables (Kerlinger, 1973). Recordemos como el principio de empirismo exigía una posibilidad de contrastación empírica de las afirmaciones. Esto descarta en el establecimiento de preguntas (hipótesis) asuntos religiosos, morales o filosóficos, así como preguntas generales del tipo ¿qué causa A? ó ¿cuales son los efectos de B? En segundo lugar, las variables deben de ser



observables o al menos potencialmente observables. Consideremos, en ese sentido, las siguientes preguntas:

- ¿Producen las noticias de suicidios de personajes famosos un incremento en los suicidios nacionales?
- ¿Se incrementa la represión política en un país cuando se incrementa la inversión extranjera?

Las respuestas a dichas preguntas son las denominadas hipótesis, si bien las hipótesis pueden plantearse en términos distintos a sistemas de relaciones como es en el caso de distribuciones de variables. Una hipótesis viene definida como una relación esperada, pero sin confirmar, entre dos o más variables, o una distribución esperada. Es factible generar hipótesis desde diferentes fuentes, que irían desde la teoría, la observación directa, intuiciones, etc. Algunas veces, la formulación de hipótesis es el resultado principal de una investigación previa, donde se determina y define el problema. En otros casos, una hipótesis o un conjunto de ellas puede, de forma no explícita, conducir toda una actividad investigadora. En principio, aún cuando las hipótesis pueden plantearse de modos diferentes, todas ellas especulan acerca de la naturaleza o la forma de una relación entre variables. Así, una hipótesis planteada adecuadamente indicara la naturaleza de la relación, es decir que variable predice a la otra, así como su forma, en que modo los cambios en una variable están relacionados con los cambios en la otra.

La noción de relación (o causa) es central en la tarea inductiva (cuyo eje central es la hipótesis). No obstante, la discusión sobre la idea de causalidad ha sido crucial durante mucho tiempo. No obstante, tal y como destacara Stuart Mill en su *Sistemas de lógica*, "Siendo la noción de causa la raíz de toda la teoría de la inducción, es indispensable que esta idea sea fijada y determinada con la mayor precisión posible al principio de nuestra investigación. A la verdad, si la lógica inductiva exigiese para constituirse que las disputas tan largas y encarnizadas de las diferentes escuelas filosóficas sobre el origen y el análisis de la idea de causalidad fuesen decididamente terminadas, la promulgación o, por lo menos, la adopción de una buena teoría de la inducción podría ser considerada por mucho tiempo como desesperada. Pero, felizmente, la ciencia

de la investigación de la verdad por la vía de la prueba es independiente de las controversias que perturban la ciencia del espíritu humano, y no está necesariamente obligada a proseguir el análisis de fenómenos intelectuales hasta este último límite, que sólo podría contentar a un metafísico.” En ese sentido, las hipótesis que especifican la forma de una relación se dicen que son testables, por que permiten, suponiendo que las variables sean medidas adecuadamente, determinar si la hipótesis es verdadera o falsa, o al menos si es probablemente verdadera o probablemente falsa. En la mayoría de los casos, el optar por formular una hipótesis testable en un tipo u otro no conduce a mundos incompatibles. Existen muchos modos de plantear hipótesis testables, si bien las más comunes en la actividad científica son las siguientes.

1. Afirmaciones planteadas “si, entonces”.
2. Afirmaciones planteadas matemáticamente.
3. Afirmaciones planteadas sobre continuidad.
4. Afirmaciones planteadas sobre diferencias.

1. “Si, entonces”. Este tipo de hipótesis plantea que si un fenómeno o condición sucede, entonces otro sucederá a su vez. Por ejemplo, “si una persona tiene un alto nivel de educación formal, entonces tendrá pocos prejuicios raciales”. Como veremos, en lógica este tipo de argumentos se denominan implicaciones. En resumen, una implicación consiste en la conexión entre dos afirmaciones simples, cada una de ellas indicando la presencia de una categoría o valor en una variable. En investigación social este tipo de implicaciones es bastante escasa.

2. Formulaciones matemáticas. En algunas ocasiones, la hipótesis adopta una formulación matemática explícita, indicando algún tipo de relación funcional. Por ejemplo,

$$y = f(x)$$

En cierto modo, una afirmación matemática es equivalente a una implicación lógica. Así,  $y = f(x)$  afirma “si y sólo si  $x$  adopta este valor, entonces  $y$  adoptara este otro valor”.

3. Hipótesis de continuidad. Adoptara la forma “cuanto más grande (o menor) es X, más grande (o menor) es Y”. En definitiva, afirma que los incrementos o decrementos en una variable (X) están asociados con incrementos o decrementos en otra variable (Y). Por ejemplo, “cuando los ingresos medios familiares crecen, el consumo familiar crece”.

4. Hipótesis de diferencias. Las afirmaciones de este tipo dicen que una variable difiere (adopta un valor diferente) según las categorías de otra variable. Eje. “Los hombres tienen más prejuicios que las mujeres”.

Las hipótesis planteadas en términos de continuidad o de diferencias dependen de la presencia de variables con niveles de medición discretos o continuos. Así, variables continuas pueden ser agregadas para crear categorías, (ingresos altos, medios y bajos) de tal forma que la hipótesis podrá ser de diferencias o de continuidad, según decida el investigador. Sin embargo, cuando alguna variable en la hipótesis es discreta dará origen a una hipótesis basada en diferencias.

Tanto las hipótesis expresadas como continuo o en base a diferencias especifican claramente la forma de la relación. Sin embargo, ambos tipos son ambiguos acerca de la naturaleza de la relación. Es decir, no explicitan la posible dirección en la relación causal postulada entre las variables. Por ello, debe cuidarse el explicitar claramente que variable se considera causa y cual es efecto, o si la relación es recíproca, etc.

En términos de investigación científica las relaciones entre variables se plantean realmente en términos probabilísticos. Esto se refleja en el uso de las palabras “tiende” o “en general”. Parecería más correcto explicitar las hipótesis con estos correctores, en la medida que advierte sobre el carácter probabilístico de la afirmación. En esencia, las hipótesis se refieren a conceptos teóricos y sus relaciones, si bien su operativización exige, en la mayoría de los casos la sustitución de conceptos teóricos por variables.

### **3.3. Explicación y causalidad**

Existe una dificultad evidente en establecer la noción de causalidad. En general, se consideran necesarias tres condiciones, Selltitz et al., (1959) para que se pueda plantear la relación de causalidad entre dos variables. La primera es la necesaria covariación entre las variables que se suponen ligadas causalmente. La segunda de las condiciones impone una asimetría u orden causal (temporal) entre la variable causa y la variable efecto. La tercera de las condiciones es algo más problemática, en la medida que exige que no existan otros factores o variables comunes que puedan explicar la covariación de las variables que se suponen relacionadas causalmente. De suceder esto, la desaparición de esa causa común supondría la desaparición de la supuesta relación entre causa y efecto, Es evidente que la explicación causal es fuertemente dependiente de la teoría. Existe un número potencialmente infinito de variables (reconocidas como tal o no) que pueden estar influenciando sobre las dos simultáneamente. No existe ningún test estadístico que pueda revelar la existencia de estas variables. En ese sentido, la construcción de un modelo causal supone el cierre formal (abierto en la realidad) de un sistema. Para ello se selecciona un conjunto finito de variables y se proponen y evalúan relaciones entre ellas. Es muy importante tener presente la contingencia del modelo.

Es especialmente importante recordar esta provisionalidad porque, en evitación de la parálisis investigadora, la recomendación tal y como señala Asher (1983) consista en operar sobre la base de que todas las variables están consideradas. En palabras de Blalock (1964) “no importa cuan elaborada sea una teoría, siempre es necesario establecer algunas presunciones. Particularmente, debemos asumir hasta cierto punto que los factores externos que afectan al modelo son estadísticamente despreciables. La aleatorización puede ayudar a resolver este problema, pero la plausibilidad de este tipo de presunción simplificadora es siempre una cuestión de grado.” Como tendremos ocasión de comprobar, las presunciones a las que alude Blalock se refieren preferentemente al error.

En términos generales, hablaremos de causalidad cuando los cambios en una variable (llamémosle X) producen (tarde o temprano) cambios en otras variables (llamémosle Y). No obstante, es preciso establecer una serie de aclaraciones de las condiciones en que se plantea ese "producir" cambios en otra variable.

Estamos hablando de *cambios medios*. Estos cambios, al estar referidos a grupos de individuos (o diferentes unidades de análisis, pero siempre conjuntos de ellas) nos lleva a hablar básicamente de promedios o tendencias. Es decir, que la relación observada lo es para segmentos mayores o menores de población; esto implica que al ser promedios puede que dicho cambio detectado no exista realmente (es decir, ninguna unidad de análisis se ve afectada realmente de esa forma); debe entenderse entonces que, no sólo no es obligatorio o necesario el cambio a todas las unidades de análisis (es bastante probable que existan excepciones individuales o de grupos específicos de individuos), sino que además podría suceder perfectamente que los coeficientes de cambio que se determinan empíricamente no existan realmente en la magnitud estimada. Una relación causal, en este sentido, es probabilística y no obliga a todos (a diferencia de las leyes en una democracia, que sí obliga a todos). Cuestión diferente es el tratamiento o etiqueta con que se identifiquen las excepciones al promedio. Las excepciones en términos de distribución pueden ser casos atípicos o casos extremos.

En el diagnóstico previo a cualquier análisis causal es imprescindible el estudiar estos casos que se encuentran fuera de la norma de variación o de covariación. Es decir, que se comportan de modo distinto al patrón detectado en el conjunto de la población para la que poseemos datos. Existen razones para ello. Primera, validar los datos descartando errores de medición o grabación. Segundo, explicarnos porque esos casos o unidades de análisis son distintos. Podría suceder que dichos casos con un comportamiento extraño sean simplemente un caso de visibilidad. El riesgo evidente es que la producción de los datos o de las mediciones haya excluido (o se hayan excluido) segmentos de sociedad con un comportamiento equivalente. En ese sentido, el patrón detectado es simplemente consecuencia de la homogeneidad producida por la tarea de recoger información, lo que ha hecho más visible a una parte de la sociedad e invisible a otras. Es urgente explicar los casos

atípicos y extremos respondiendo a la pregunta ¿son unidades de análisis en condiciones individuales particulares o representan un caso de visibilidad de los segmentos sociales invisibilizados en el proceso de traducción de la realidad social en datos e información?

En cualquiera de los tres casos, error de grabación, casos individuales o segmentos invisibles, debe de tomarse una decisión antes de proceder al análisis. El motivo es evidente; los casos extremos ejercen una influencia notable sobre los ajustes. En el caso de errores de grabación la solución es quizás más evidente, bastando con corregirlos o suprimirlos. En las otras situaciones, una vez explicados los casos atípicos o extremos, retirarlos del análisis sólo parece legítimo si son casos particulares y además hemos aprendido lo que significan con relación al resto de los datos. La decisión en el caso que se determine mediante análisis que dichos casos representan situaciones de visibilidad excepcional de segmentos sociales (y condiciones sociales) ocultos sólo podría ser redefinir de que se esta hablando y de que no se esta hablando. Sería inaceptable la ocultación de los casos atípicos o extremos y mantener la ficción de investigar la sociedad.

No debe olvidarse que la transformación de la sociedad en información/datos implica una traducción de esta que aporta, a su vez, una fuente de variabilidad. En ese sentido, la variación, la covariación o cambios medios determinados entre dos variables puede verse condicionada por el proceso de obtención de los datos. De este modo, la traducción de la realidad con unas escalas y no otras, con unas variables y no otras, con un método y no otro, puede producir variabilidad en los coeficientes sintéticos y por lo tanto en los mismos cambios medios.

En esta revisión del significado de los cambios medios y sus condiciones de existencia e interpretación (validación, visibilidad/invisibilidad, casos particulares, traducción) nos falta considerar aquellos que siendo reales y habiendo superado todas las consideraciones anteriores dependen de condiciones teóricas para ser definidos como parte causal en un proceso explicativo. La relación empírica debe ir acompañada de una relación lógica que le haga verosímil. Se trata de un caso especial de relaciones espurias, donde una estrecha relación estadística no responde a ninguna explicación lógica (no existe contigüidad en el espacio teórico entre las variables o

conceptos relacionados). En ese sentido, la noción de causalidad es evidentemente teórica, en la medida que la relación no es por sí sola suficiente para definir causalidad.

La noción de causa es una noción asimétrica. Implica una ordinalidad temporal, una secuencia ordenada de covariaciones. La covariación entre dos variables no prueba por sí sola la presencia de causación. Como ya hemos advertido, un coeficiente de covariación expresa la fuerza con que está asociada la variación de dos variables, pero no afirma nada acerca de una hipotética dirección causal entre ambas. Un ejemplo de ello es cuando se presentan correlaciones expresas, es decir, asociaciones entre variables que no están relacionadas teóricamente. Cuando no existe una relación teórica, la asimetría tiende a ser arbitraria. En ese sentido, las covariaciones espurias sin contigüidad explicativa acostumbran a estar, a su vez, desordenadas lógicamente. La causalidad u orden de una asociación es un constructo teórico, que primero considera la asociación entre un conjunto de variables y no otras y segundo establece un orden entre ellas. En resumen, el concepto de correlación habla de coordinación en la variabilidad de dos variables, mientras que el de causa lo hace de subordinación entre ellas (es decir, varían conjuntamente, pero una de ellas dependiendo de la otra, de algún modo a sus órdenes).

Esto es importante en la medida que la noción de causalidad implica que el ajuste empírico (determinación de una relación mediante estimación) es condición necesaria pero no suficiente para definir causalidad. Por el contrario, la contigüidad lógica en un contexto cultural puede llegar, erróneamente, ha aceptarse teóricamente como condición suficiente de causalidad, incluso en ausencia de evidencia empírica. Precisamente, esa es parte de la crítica de la sociología matemática con respecto a la sociología exclusivamente teórica: su tendencia a satisfacerse con la contigüidad lógica como condición suficiente para explicar el mundo. Recordemos que el análisis causal intenta explicar la sociedad, pero en todo momento tiene el freno de lo que puede saber y en que condiciones. Este freno a reconstruir el mundo sobre condiciones suficientes (propio de las ideologías) y atender a las urgencias de las condiciones necesarias es uno de los elementos centrales de la sociedad matemática y, evidentemente, del análisis causal.

Cuando en un análisis causal se afirma que "X" es causa de "Y" no se afirma que "X" sea única causa de "Y". Existirán, probablemente, fenómenos sociales que dependerán de una sola y exclusiva causa. Sin embargo, es difícil que en un sistema tan interconectado como es el sistema social este caso sea el más frecuente. Ello es cierto incluso para coeficientes de correlación iguales a 1, donde la relación sea perfecta, en la medida que pueden existir otras variables que contribuyan en explicar la variabilidad de la dependiente, coordinada con la otra variable explicativa.

Como casi siempre, será desde la interpretación que ofrezca la teoría como se resolverán las cuestiones de dependencia. Recordando nuevamente el aspecto central, en un modelo causal se presume implícitamente que las causas no son únicas, y precisamente la importancia del residual (perturbación) en una relación indicará la importancia, en cantidad y calidad, de las variables que no están presentes en el modelo.

Los modelos sociológicos son porosos a la realidad que los rodea. Esto nos conduce al concepto de "Control". La noción de control es posiblemente la más alejada de la práctica de la investigación sociológica. La idea central es la necesidad de controlar el máximo de factores que puedan estar influenciando el proceso o fenómeno en observación. De este modo, es factible atribuir de forma única la explicación de la variabilidad observada. En ciencias naturales nos conduce directamente a los diseños experimentales. En Ciencias Sociales la psicología ha sido, probablemente, la disciplina que más se plantea el empleo de este tipo de diseños. En sociología, estos están, evidentemente vedados. La investigación sociológica se centra sobre todo en los diseños correlacionales y en todo caso, en los diseños cuasi-experimentales. En ese sentido, hablar de control es hablar de errores en la investigación. No obstante, eliminar todos los sesgos es prácticamente imposible, dado que el lenguaje mismo estructura nuestra percepción del mundo. Los sesgos alcanzan incluso a la selección de los problemas para investigar o las mismas estrategias utilizadas.

En el caso de la investigación experimental, existen diversos procedimientos empleados en el transcurso de la investigación destinados a minimizar los errores. Así, la asignación aleatoria a grupos experimentales. Es importante recordar que el tipo de diseño de investigación empleado afecta a la



aplicación de técnicas de análisis. Por ejemplo, los problemas del análisis de varianza con dos factores o más, empleando datos de encuestas, y por ello con casos desiguales en cada celdilla. Este ideal de control del proceso de investigación es una referencia importante en la actividad denominada ciencia. El empleo de los procedimientos de control elimina las explicaciones confusas de los sucesos en estudio, y es uno de los rasgos principales por el que difiere el conocimiento científico del casual. Mientras que el investigador puede diseñar estudios para recoger información y testar determinadas explicaciones o respuestas, los hallazgos están frecuentemente abiertos a diferentes interpretaciones. La idea de control pretende emplear procedimientos que efectivamente descarte aquellas explicaciones que realmente no son coherentes con la información. En el caso de la investigación sociológica, el control es muy limitado y se reduce, en el caso de la investigación cuantitativa, a control estadístico.

Los datos que provienen de diseños no experimentales aportan evidencias menos concluyentes que aquellos otros que se obtienen de diseños experimentales. Uno de los motivos por los que la evidencia es más sólida en términos experimentales que en no experimentales viene dado por el hecho de que en diseños experimentales existe un control sobre las fuentes de variabilidad. Es decir, al mutilar o aislar el sistema de relaciones, la atribución a las variables de control de la variabilidad apreciada en las variables dependientes es neta, dado que se intenta que no existan otras fuentes de variabilidad que influyan. Sin embargo, en investigación sociológica, los sistemas causales se ajustan sobre datos que provienen de contextos permeables a la realidad que le rodea (variables no incluidas en el análisis, variables desconocidas). Esto implica que la permeabilidad de la información y su sensibilidad a factores no controlados obliga a un escrutinio minucioso de la variación no explicada. Dentro del método científico, el método correlacional es una alternativa al método experimental siempre y cuando se acepten y reconozcan sus limitaciones.

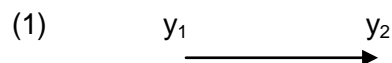
Como resumen, para que se pueda afirmar la noción de causalidad, es decir, que las variaciones en una variable (X) causan las variaciones en otra (Y), deben de estar presentes tres condiciones:

- (1) La variable que se considera causa (X) debe preceder temporalmente/lógicamente a la variable efecto (Y). (criterio de asimetría)
- (2) Las variables que se consideran causa (X) y efecto (Y) deben de covariar. (coordinación estadística)
- (3) No deben existir explicaciones alternativas que, de un modo satisfactorio, den cuenta teórica de la covariación observada entre X e Y. (criterio de control desde la perspectiva correlacional).

#### **4. La determinación teórica del orden explicativo**

Como hemos mencionado anteriormente, consideradas dos variables, éstas pueden mostrar una coordinación estadística de tal forma que covarían. Sobre dicha coordinación observada puede articularse una relación teórica de subordinación entre ellas, donde una característica correspondiente a un caso (sea entrevistado, país, en definitiva, unidad de análisis) cambia de valor en función a los valores que adopte otra característica diferente. Por ejemplo, horas de ver televisión en función de ingresos. Probablemente, existe un umbral de ingresos donde otros ocios reemplazan las horas de televisión. Por lo tanto, de existir correlación estadística, las horas de televisión está subordinada a los ingresos o renta disponible del entrevistado.

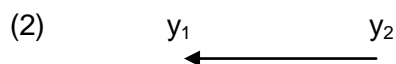
Parece evidente que el orden de subordinación que se establezca entre dos variables se "construye" sobre la existencia de covariación entre los valores de dos variables. De no existir covariación estadística, la subordinación (relación causal) entre variables pierde fundamento empírico. Analíticamente, podemos diferenciar cuatro posibilidades en ese orden causal para el caso de dos variables<sup>24</sup>:



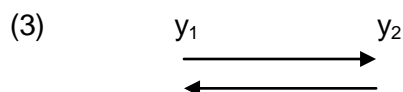

---

<sup>24</sup>Hemos notado la relación mediante una flecha, es decir, mediante un grafo orientado. Cuando la relación de subordinación es unidireccional implicará un sólo grafo orientado, cuando es una relación de interdependencia serán dos grafos, cada uno de ellos indicando una dirección opuesta. En el caso de coordinación estadística, sin relación de dependencia, se expresa mediante un grafo con dos cabezas de flecha.

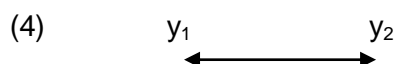
En este caso, "y<sub>1</sub>" puede influir en "y<sub>2</sub>" pero no lo contrario.



En este segundo caso, "y<sub>2</sub>" es la que influye en la variabilidad de "y<sub>1</sub>", siendo imposible, teóricamente lo contrario.



Aquí las variables "y<sub>1</sub>" e "y<sub>2</sub>" se influyen mutuamente, donde se retroalimentan en su variación. Por ejemplo, los presupuestos de defensa de USA e URSS durante la guerra fría.



"y<sub>1</sub>" e "y<sub>2</sub>" presentan una coordinación estadística, sobre la que no presumimos ningún orden teórico. Básicamente, existe una covariación no explicada.

En el momento de construir la secuencia de relación entre variables sobre la base del criterio de subordinación lógica pueden surgir problemas importantes. Algunas de ellas sólo se resolverán desde la aceptación de una teoría previa, especialmente en el caso de variables que acostumbran a ser de naturaleza exógena (como entre Religión e Ideología política). No obstante, si es posible establecer unas orientaciones metodológicas para la determinación teórica del orden causal.

La noción de orden causal se establece habitualmente sobre la de "tiempo". Esencialmente, se resumen en la afirmación "lo que sucede después no puede causar lo que sucedió antes". Si algo cambia, lo hace en función de un cambio previo en la variable de la que depende. Si instrumentamos la determinación del orden causal en función a qué variable cambia primero, es útil operativizar esta secuencia temporal.

Es útil pensar que las variables tienen una fecha de "inicio" (o de fabricación) y otra de "término". Así, la fecha de inicio de una variable es el momento antes del cual no puede existir valor para esa variable. La fecha de término de una variable es el momento tras el que no existen cambios de valor posible. Por ejemplo, la variable "participar en la última Guerra Civil española", la variable tiene una fecha de inicio en 1936, y de terminación en 1939. Antes de 1936 no existía valor ni variable, después de 1939 el valor que posea cada caso es fijo (sí o no). Siguiendo este criterio de inicio y término de una variable (cuando esto es posible) es factible el establecer una serie de reglas para determinar el orden causal entre variables (allí donde la teoría o la incertidumbre no prescriban un orden de subordinación en la covariación).

**a)** Consideremos que  $y_2$  varía en función de  $y_1$  cuando la variable  $y_2$  tiene una fecha de inicio posterior a la fecha de terminación de  $y_1$ . Por ejemplo, la variable "actitud hacia la entrada de España en la OTAN" y "participación en la última Guerra Civil española". La opinión de los españoles acerca de la entrada de España en la OTAN, variará en función a su participación en la Guerra Civil, dado que la fecha de inicio (entrada de España en la OTAN) es posterior a la de terminación de "participación en la última Guerra Civil española"<sup>25</sup>.

**b)** Podemos considerar que  $y_2$  depende de  $y_1$  cuando la variable  $y_1$  se vincula a un paso, etapa o momento anterior de  $y_2$ , dentro de una secuencia temporal bien conocida. Un ejemplo de secuencia en sociología es el ciclo vital, donde se describe una trayectoria vital que es seguida por la inmensa mayoría de los que viven en una sociedad. Así, la secuencia vendría a ser, esquemáticamente,

- (1) características de la familia donde se crece
- (2) educación escolar
- (3) primer trabajo a tiempo completo
- (4) primer matrimonio
- (5) nacimiento de hijos

---

<sup>25</sup>La anticipación de una variable (es decir, pensar sobre su comportamiento futuro, puede hacer variar otra que le antecede temporalmente. Sin embargo, hay que considerar que la variable "anticipación de  $y_2$ " es una variable distinta y anterior a la variable " $y_2$ ".

(6) disolución de matrimonio por divorcio o muerte.

Y así continuadamente, concretándose tanto como el investigador desee.

De hecho, el ciclo vital de un individuo se trenza, por así decirlo, con el de sus padres y sus propios hijos o su entorno social. De este modo prácticamente todo aquello que acontece en el ciclo vital de un individuo, se podría hacer depender, teóricamente, del entorno familiar en que este desarrolla su personalidad y sus costumbres, así sucesivamente.

**c)** Otro criterio es cuando una de las variables posee valores o atributos que se pueden considerar definitivos para cada individuo, mientras que la otra registra la posibilidad de que un individuo cambie de valor o atributo. Por ejemplo, una variable en la que difícilmente se cambia de valor es la variable género, y de hecho esta variable se emplea habitualmente para explicar la variación de otras variables sociológicas.

**d)** Otra regla, que más bien es una orientación, afirma que si los valores o atributos que cada individuo o caso tiene registrado en una variable  $y_1$  son relativamente estables, lentos de cambiar, mientras que los valores que registra otra variable  $y_2$  son volátiles o poco estables,  $y_2$  dependerá normalmente de  $y_1$ . Así por ejemplo, si las variables son "preferencias religiosas" y "opinión sobre el gobierno", probablemente la segunda dependerá de la primera.

Es difícil infravalorar la importancia que posee una correcta especificación del modelo, tanto por sus consecuencias teóricas como por la prima de ajuste que se obtendrá al postular relaciones con probabilidad de ocurrir. Estas orientaciones anteriores ayudan a establecer con mayor consistencia el orden causal, en el sentido de postular una dirección causal, bidireccionalidad o simple coordinación estadística sin determinación de orden.

#### **4.1. La lógica como pensamiento**

El proceso de razonamiento en la ciencia consiste en intentar obtener conclusiones acerca del orden que exista en el mundo empírico en base a evidencias observacionales. Para ello, la herramienta es la lógica. La lógica es un lenguaje formalizado con vida propia, tal y como sucede con las matemáticas. En ese sentido, contiene sus propios demonios y limitaciones cuando se trata de "hablar" sobre la sociedad. Tal y como destaca Adorno "Sin

proponérselo, niega precisamente la complejidad de las situaciones sociales tal y como éstas vienen indicadas por término hoy tan excesivamente solicitados como los de alienación, cosificación, funcionalidad y estructura. El método lógico de reducción a elementos, a partir de los cuales se edifica lo social, elimina virtualmente las contradicciones objetivas.” En ese sentido, la lógica estudia la relación entre evidencias y conclusiones (Salmon, 1973), pero la forma de empleo de las evidencias (términos relacionados) es anterior a la evaluación de la argumentación. Una vez efectuada una inferencia, la lógica puede decirnos si ésta es correcta o no. En otras palabras, si las evidencias justifican las conclusiones. Es una operación mecánica, que analiza la estructura del argumento, no los genera.

Esta presentación de la lógica es bastante selectiva y supone básicamente una introducción de la cuestión. Siguiendo a Wheelwright (1962) entendemos que los elementos fundamentales del análisis lógico consisten en términos, proposiciones y argumentos.

*Términos:*- Los “términos” son los elementos más simples del análisis lógico. Pueden venir dados por una palabra o una frase. Un término no es por sí mismo falso o verdadero. Podemos entender o comprender un término pero no afirmarlo o negarlo. En ese sentido, término es análogo a concepto en investigación y, al igual que esta, son evaluados por su significado.

*Proposición:*- Es un juicio expresado acerca de un término o términos. Desde un punto de vista gramatical es un aserto. A diferencia de los términos, las proposiciones son por definición verdaderas o falsas. Existen diferentes tipos de proposiciones en análisis lógico. La lógica tradicional se ocupaba principalmente de las denominadas proposiciones categóricas. Un tipo de proposición con un interés especial para la investigación social son las denominadas proposiciones hipotéticas o condicionales. Básicamente, una proposición condicional consiste en dos afirmaciones unidas por las preposiciones “si” y “entonces”. Por ejemplo, “Si estamos en 1995, entonces la bomba atómica existe”. Ejemplos de estos son "todos los hombres son mortales" o "ningún efecto existe sin causa". Otros dos tipos de proposiciones bastante comunes son "disyuntivas" o "conjuntivas". Tal y como destacaba Wittgenstein en el Tratado lógico-filosófico, “Las proposiciones lógicas

describen la armazón del mundo, o, mejor, la presentan. No “tratan” de nada, presuponen que los nombres tienen significado, y las proposiciones elementales, sentido; y ésta es su conexión con el mundo. Es claro que debe manifestar algo sobre el mundo el hecho de que resulten tautologías de uniones de símbolos que tienen esencialmente un carácter determinado. Este es el punto decisivo. Decimos que, en los símbolos que utilizamos, algunas cosas son arbitrarias y otras no. La lógica expresa solamente esto último; pero esto significa que en lógica nosotros no expresamos por medio de signos lo que queremos, sino que en lógica habla la naturaleza misma de los signos esencialmente necesarios. Si nosotros conocemos la sintaxis lógica de un lenguaje de signos cualquiera, entonces todas las proposiciones de la lógica están ya dadas.”

En una proposición condicional la parte introducida por el “si” es denominada antecedente y la expresión que sigue a “entonces” es llamada consecuente. Así, en nuestro ejemplo, “Si estamos en 1995, entonces la bomba atómica existe”, “Si estamos en 1995” es el antecedente y el consecuente será “entonces la bomba atómica existe”. Una proposición condicional afirma que el antecedente implica el consecuente. Si el antecedente es verdadero, el consecuente también lo es. No obstante, dentro de esta situación pueden plantearse diferentes tipos de relaciones entre antecedente y consecuente. Una relación es, por ejemplo, de tipo definicional, donde la verdad del consecuente se deriva en tanto que es parte (es decir, está contenida) en la definición del antecedente. Así, “Si esta sociedad está económicamente desarrollada entonces su producto interior bruto es muy alto”. En este tipo de proposición condicional el consecuente es parte de la definición operativa del antecedente.

Otra interpretación importante para la investigación científica es la lectura de la relación “Si, entonces” como una conexión causal. “Si se rompen los vínculos emocionales de pertenencia al grupo, entonces se incrementaran las tasas de suicidio”. Este tipo de proposición no valida el consecuente desde un punto de vista conceptual. La verdad de esta proposición condicional debe ser establecida empíricamente.

*Argumentos:-* Desde el punto de vista de la lógica, un argumento es un conjunto de proposiciones, una de las cuales se presume que se infiere

necesariamente o probablemente de las demás. En principio, una vez que se efectúa un razonamiento, este puede ser explicitado en la forma del argumento, donde una afirmación sigue a otra hasta concluir en una última afirmación. La afirmación última que concluimos se denomina conclusión, mientras que las diferentes afirmaciones que se ofrecen para aceptar la conclusión se denominan premisas.

El principal tipo de argumento estudiado por los lógicos es el silogismo. Los silogismos son argumentos compuestos por tres proposiciones: dos premisas y una conclusión, que se infiere lógicamente desde la primeras. Así,

"Todos los hombres viven en una cultura,  
los sociólogos son hombres,  
luego, los sociólogos viven en una cultura."

Evidentemente, el argumento que subyace a un razonamiento puede contener más de tres proposiciones, pero argumentos más largos pueden ser descompuestos en series de silogismos, por lo que el silogismo puede considerarse la unidad básica del análisis lógico. Así, para analizar la lógica de un razonamiento científico debemos identificar en primer lugar la estructura silogística del razonamiento y, posteriormente, examinar los silogismos con respecto a su validez o invalidez. La validez de un silogismo depende sólo de la relación entre premisas y conclusión. En un silogismo válido se presenta la siguiente relación entre premisas y conclusiones. Si las premisas son verdaderas, entonces la conclusión debe también ser verdad. No se plantea si las tres afirmaciones son verdaderas o no realmente, sino que si las premisas lo fuesen, la conclusión lo sería a su vez. Por ello, un silogismo válido en su estructura puede contener proposiciones falsas en la realidad, e inversamente, un silogismo inválido en su estructura puede estar compuesto de afirmaciones ciertas.

Debe quedar claro que la validez de un silogismo no depende de la verdad de sus premisas. En ese sentido, la lógica como disciplina posee una aplicación más amplia que las definidas por la actividad científica, en la medida que se puede analizar argumentos independientemente de si dicen algo acerca



de la realidad. En resumen, los argumentos, proposiciones y términos son considerados desde planteamientos diferentes.

- Argumento: validez según la estructura del razonamiento
- Proposiciones: verdad, en base a evidencias empíricas
- Términos: evaluación de su significado; operativización del concepto.

Es decir, los términos son evaluados por su significado, las proposiciones por su verdad y los argumentos (silogismos) lo son en su validez (Wheelwright, 1962). La actividad científica, en todos sus campos, intenta establecer un conocimiento acerca del mundo empírico. Ello implica que para afirmar sus conclusiones acerca de la realidad debe evaluarse la verdad que contienen sus afirmaciones (correspondencia entre ellas y las evidencias empíricas observadas) así como la validez de sus argumentos (Salmon, 1973). Podríamos afirmar que no sólo debe preocupar la validez del argumento, o la verdad de las proposiciones, sino que la preocupación de la actividad científica debe ampliarse hasta los mismos términos, en base a definiciones operativizables de los conceptos utilizados.

Los argumentos, o en definitiva, los razonamientos se acostumbran a clasificar en dos tipos principales: deducción e inducción. No obstante, con bastante frecuencia desde Pierce aparece un tercer tipo denominado abducción. Desde el punto de vista clasificatorio de la codificación o fortaleza del análisis lógico, la mayor contundencia corresponde al análisis lógico deductivo, siendo más débil el inductivo y especialmente débil el abductivo, más ligado a la teoría del descubrimiento. Así, una de las diferencias esenciales viene dada por el grado de certeza con que puede afirmarse la conclusión en base a las premisas. En el caso del razonamiento deductivo, si el argumento es lógicamente válido la conclusión será necesariamente verdadera si las premisas lo son. Cuando se razona desde una lógica inductiva, o más aún desde una abductiva, podemos afirmar que la conclusión es probablemente cierta (pero no necesariamente) en el caso que las premisas sean ciertas a su vez.

En el caso deductivo sería inconsistente negar la conclusión si las premisas son ciertas, dado que en ésta no existe nada que no se afirmara

previamente. En el caso de la inducción, es factible que la conclusión sea falsa aún cuando las premisas sean ciertas, dado que la información contenida en la conclusión va más allá de la afirmada en las premisas. Así, en términos de argumentación, una deducción es válida o inválida, pero no existe punto intermedio.

*Comprobación de validez en argumentos deductivos.* Existen métodos más sofisticados para testar la validez de una forma argumental, como son diagramas de Venn, reglas silogísticas o tablas de verdad. Podrá profundizarse en el tema a través de cualquier manual de lógica formal o simbólica.

Un argumento inductivo puede aceptarse o no como convincente en función de la probabilidad que posea la conclusión de ser verdadera dadas las evidencias que aportan las premisas. Nos encontramos ante dos procedimientos argumentales que no son excluyentes, sino que son aplicables dinámicamente en la actividad científica, generando inducciones desde datos empíricos y planteando desde ellas deducciones confirmatorias que busquen más evidencias empíricas.

Como hemos advertido, la validez de un argumento deductivo depende de la relación entre premisas y conclusiones. Veamos seguidamente algunas estructuras argumentales frecuentes en el razonamiento científico. Una primera tarea es determinar un sistema de notación. Usualmente se substituyen las afirmaciones que constituyen el argumento mediante letras. Siendo “p”, “q” y “r” afirmaciones cualquiera, las siguientes son tres formas argumentales válidas referidas a proposiciones condicionales (si, entonces).

1. Afirmación de antecedente

Si p, entonces q.

p.

Luego q.

2. Negación de consecuente

Si p, entonces q.

No q.

Luego no p.

### 3. Cadena argumental

Si  $p$ , entonces  $q$ .

Si  $q$ , entonces  $r$ .

Luego si  $p$ , entonces  $r$ .

Obsérvese de dónde procede el nombre de las formas argumentales. En la forma “afirmar el antecedente”, la primera premisa es una proposición condicional, la segunda premisa afirma el antecedente de este condicional. La forma “negar el consecuente”, siendo la primera premisa una afirmación condicional, la segunda premisa niega el consecuente. En la forma cadena argumental, dos premisas se vinculan (como en una cadena) mediante una afirmación común que es consecuente de la primera y antecedente de la segunda. Es factible generar cadenas argumentales tan largas como se quiera, siguiendo la mecánica explicitada, dónde el consecuente de una premisa es a su vez el antecedente de la siguiente, teniendo la conclusión el antecedente de la primera premisa, y como consecuente el de la última premisa.

Las formas argumentales inválidas son denominadas falacias. Las siguientes falacias son bastantes frecuentes.

#### 1. Falacia de afirmar el consecuente

Si  $p$ , entonces  $q$ .

$q$ .

Luego  $p$ .

#### 2. Falacia de negar el antecedente

Si  $p$ , entonces  $q$ .

No  $p$ .

Luego no  $q$ .

Ambos argumentos son inválidos porque es factible que las premisas sean ciertas y la conclusión falsa. Generalmente las falacias vienen dadas por combinaciones peculiares de falsedad y verdad en las premisas y las conclusiones. Manheim (1977) ofrece un resumen de varias combinaciones de

premisas falsas y verdaderas, argumentos validos e inválidos, así como de conclusiones falsas y verdaderas.

1. Si todas las premisas son verdaderas y el argumento es valido, la conclusión debe ser verdad.

2. Si todas las premisas son verdad y la conclusión falsa, el argumento debe ser inválido.

3. Si el argumento es valido y la conclusión es falsa, al menos una premisa debe ser falsa.

En todo caso, el mejor método para evitar posibles confusiones es determinar la validez de un argumento deductivo evaluando su forma. La utilidad del análisis lógico se revela en la evaluación de teorías. Para ello, es necesario descomponerlas en series de proposiciones y, posteriormente organizar las proposiciones en argumentos, determinando posteriormente su validez. Este procedimiento facilita no sólo su validación lógica, sino también la operativización de forma explícita de aquello que se propone (aquello que es comprobable y aquello que no lo es); asimismo, ayuda a detectar posibles contradicciones o presunciones implícitas.

## **4.2. Los argumentos inductivos**

Como afirmamos anteriormente, mediante la inducción pretendemos llegar a conclusiones que exceden la información contenida en las premisas. Dado que un presupuesto de la ciencia es su intención abstraer y generalizar, es decir determinar patrones regulares que subsumen la variabilidad de las observaciones particulares, la inducción es una herramienta necesaria. En la inducción nos interesamos, como hiciéramos con la deducción, por la relación existente entre las premisas y las conclusiones. Sin embargo, analizar un razonamiento inductivo es mucho más complicado dado que debe evaluarse el grado en que las premisas soportan las conclusiones; aún cuando las premisas de un argumento inductivo correcto sean verdaderas, la conclusión solamente puede probablemente verdadera.

Vamos a aproximarnos al problema de la lógica inductiva en dos enfoques sucesivos. En primer lugar, el razonamiento inductivo en tanto que parte del proceso de generalización inductiva que forma parte de la propuesta de explicaciones científicas, es decir desde la óptica exploratoria, y en segundo

lugar el papel de la inducción en el testado de hipótesis ya preexistentes, de acuerdo al método hipotético deductivo.

1) La generalización inductiva afirma algo acerca de una clase entera de objetos o sucesos, en base a información sobre una parte de la clase. Por ejemplo, sobre la base de diez observaciones, por ejemplo diez sindicalistas afiliados a un partido político, podemos inducir que todos los sindicalistas están afiliados a ese partido político. Expresado más formalmente:

Todos los miembros observados de p son q  
luego, todos los p son q.

Esta parece ser una conclusión aceptable cuando todos los casos observados presentan las mismas características. Ello recordando que no afirmamos algo acerca de lo que conocemos, de los casos observados, sino también de los casos que desconocemos. No obstante, un caso más complicado se presenta cuando afirmamos que un porcentaje de los casos observados presentan una característica, e inferimos que un porcentaje semejante de la clase total poseerá dicha característica. Por ejemplo, si el 80% de los sindicalistas observados están afiliados a un partido político, inferimos que un porcentaje equivalente (el 80%) de todos los sindicalistas (incluidos lo no observados, evidentemente) están afiliados a ese partido político; la diferencia con el primer argumento es que en él un 100% de los casos observados poseían una determinada característica y por lo tanto ese 100% se transformaba en “todos” para el resto de la clase. En este segundo caso contamos con una fracción de ese todo, con lo que se abre la puerta a una incertidumbre. Por ello, la inducción nos permite afirmar:

El 80% de los miembros observados de p son q  
luego, el 80% de p son q.

Como en todas las inferencias inductivas, las generalizaciones inductivas van más allá de los hechos observados, considerando que no se observa toda la población. Por ello, lo mejor que se puede decir de ellos es que

son probablemente ciertos. En cierto modo, nos estamos refiriendo, si bien desde el ámbito de la inducción lógica, a los mismos problemas que se consideran en el ámbito de la teoría del muestreo, donde desde un subconjunto de casos particulares establecemos un marco de probabilidades a partir del cual se les confieren las atribuciones del conjunto más amplio y desconocido de casos al que pertenece.

Precisamente esta proximidad a la teoría del muestreo hace que algunos argumentos inductivos sean más sólidos que otros. En principio, la solidez de un argumento inductivo depende de lo razonable que sea suponer que los elementos observados (pertenecientes a una determinada clase) sean representativos de la clase completa. Toda aquella estrategia que garantice la similitud entre clase total y observaciones tenderá a apoyar la solidez de un argumento inductivo. Barker (1974), sugiere las siguientes reflexiones.

En general, cuanto más parecidas sean las observaciones, en características diferentes a las evaluadas, más débil será la inferencia. Es decir, si los casos son homogéneos en muchas características, tenderán a parecer casos particulares dentro de una clase total más diversa. De ello se puede afirmar lo contrario, cuanto más diversos son los casos, es decir, heterogéneos en características diferentes a la considerada, más sólido será el argumento inductivo.

Asimismo, en principio, cuanto mayor sea el número de observaciones que apoyan la influencia, más sólido es el argumento, siempre que se suponga una asociación entre el incremento del número de casos y el incremento de la heterogeneidad.

Otro criterio importante es que cuanto más vinculada este la hipótesis con otras de importancia para el área de conocimiento, más sólido será el argumento. Así, si postulamos una hipótesis entre sindicalismo y partidos políticos, podría estar relacionada con otras sobre ideología política y posicionamiento social. En todo caso, cuanto más amplia o extensa es la generalización menor será la probabilidad de ser aceptada, y su argumentación será más débil. Así, para el caso de “todos los sindicalistas son ....” su probabilidad de ser aceptada es baja. Si se restringe la generalización a sólo un segmento, por ejemplo, “todos los sindicalistas del metal en la ciudad de ...” el argumento tendrá más probabilidades de ser aceptado. En relación a ello, lo

habitual es que cuanto más precisión se invoque en el argumento, este se entenderá como más improbable; así en “el 80,5% de los sindicalistas están afiliados...” en tanto que afirmación basada sobre una muestra es mucho más débil que “la mayoría de los sindicalistas....”.

En definitiva, para evaluar la fuerza de un argumento inductivo, deben de considerarse los cinco factores mencionados de modo conjunto. Así, hay que apreciar el número de observaciones que da pie a la inducción, en conjunción con el grado de igualdad o desigualdad de dichos casos, el alcance y la precisión de la generalización inductiva, y todo ello en relación al conocimiento establecido con anterioridad.

La actividad científica se apoya en gran medida en la producción de generalizaciones empíricas en base a este tipo de inducción. Constituye la zona abierta del conocimiento basado en la observación de la realidad. Hume en su “Tratado de la naturaleza humana” planteaba la relación entre hipótesis y probabilidad en los siguientes términos; “la probabilidad, como no descubre las relaciones de las ideas consideradas como tales, sino únicamente las de los objetos, tiene, en algunos aspectos, que fundarse en las impresiones de la memoria y de los sentidos, y en otros, en nuestras ideas. Si no hubiese mezcla de impresión alguna en nuestros razonamientos probables, la conclusión sería completamente quimérica; y si no hubiese mezcla de ideas, la acción de la mente, al observar la relación, sería, propiamente hablando, sensación, no razonamiento. Es, por lo tanto, necesario que en todo razonamiento probable haya algo presente a la mente, o visto o recordado; y que de eso nosotros infiramos algo que está en conexión con ello, que no ha sido visto ni recordado.

La única conexión o relación de objetos que puede conducirnos más allá de las impresiones de la memoria y de los sentidos es la de causa y efecto, y ello porque ésa es la única en la cual podemos fundar una inferencia cierta de un objeto a otro. Las ideas de causa y efecto se derivan de la experiencia, la cual nos informa que tales objetos particulares han estado unidos, en todos los casos pasados, el uno al otro; y como se supone que un objeto semejante a uno de ellos está inmediatamente presente en su impresión, de ahí nosotros presumimos la existencia de uno semejante a su acompañante habitual. Según esta explicación, que es, creo, incuestionable en cada uno de sus puntos, la probabilidad se funda en la presunción de una semejanza entre aquellos

objetos de los cuales hemos tenido experiencia y aquellos de los que no hemos tenido ninguna, y, por lo tanto, es imposible que tal presunción se origine de la probabilidad. El mismo principio no puede ser a la vez causa y efecto de otra cosa, y ésta es, tal vez, la única proposición sobre esta relación que es cierta o intuitivamente o por demostración.” No debemos olvidar que las hipótesis constituyen, en definitiva, intentos de explicaciones de fenómenos sociales.

### **4.3. El testado de hipótesis**

El método hipotético-deductivo es la clave principal del testado de hipótesis dentro del marco de una teoría científica. Su nombre podría provocar desconcierto, en la medida que se está considerando el método inductivo. No hay tal. Hipotético implica que las explicaciones científicas testadas son “hipótesis” en el sentido que se postula su verdad, si bien no existe certeza de ello. Una vez que una explicación es entendida como una hipótesis, se deduce sus posibles consecuencias en términos observables, a efectos de testado. De ese modo, formar hipótesis y deducir las consecuencias que se derivan de dichas hipótesis, constituyen los dos primeros pasos de los cuatro que constituyen el núcleo del método hipotético-deductivo. El tercer paso consiste en comprobar las observaciones para ver si las consecuencias deducidas se cumplen, siendo el cuarto y último el efectuar inferencias acerca de las hipótesis en base a las observaciones. Estas inferencias tendrán un tratamiento lógico diferente según verifiquen o falseen, como tendremos ocasión de comprobar. Desde el punto de vista del análisis lógico, nos interesan especialmente los pasos uno (formulación de hipótesis), dos (deducción lógica de consecuencias) y cuatro (inferencias desde los datos), si bien en el ámbito inductivo importan el uno y dos.

El primer paso consiste en generar explicaciones (hipótesis), que puedan explicar un conjunto de fenómenos sociales. Si seguimos a E. Durkheim, podemos considerar una serie de asociaciones:

- Los viudos y divorciados cometen más suicidios que los casados
- Las parejas sin hijos producen más suicidios que las parejas con hijos
- Los habitantes de las ciudades cometen más suicidios que los que viven en el campo.



De las asociaciones anteriores es factible generar una explicación que de cuenta de todas esas variaciones, y por lo tanto, postular una hipótesis, según la cual la solidaridad social que exista en una sociedad explica las tasas de suicidio. Es evidente que esta generalización, que esta explicación va más allá de los datos o evidencias que le respaldan. El proceso siguiente consiste en postular la explicación como premisa dentro de un argumento deductivo y deducir desde ella los hechos que actualmente le avalan en tanto que evidencias empíricas.

- Si la solidaridad social de un grupo es más elevada que la de otro grupo distinto, entonces su tasa de suicidio será menor.
- La solidaridad social es más elevada entre los individuos casados que entre los viudos o divorciados.
- Luego la tasa de suicidios es menor entre los casados que entre los viudos y divorciados.

De un modo equivalente actuaríamos para el resto de las evidencias, de tal modo que se establezca la probabilidad de que la premisa (hipótesis) sea acertada en su intento de explicar la realidad. No obstante, explicar los hechos conocidos de este modo implica que disponemos de una hipótesis operativa para interpretar un conjunto de fenómenos sociales, pero no que dicha hipótesis sea cierta para explicar la realidad social más allá de las evidencias que le han dado forma. Es evidente la dependencia que posee la hipótesis o explicación, en esta primera etapa, de las evidencias locales y parciales que, en definitiva, han substanciado la formulación que adoptó. Esta situación es propia de la actividad inductiva tal y como señalara J. Stuart Mill en “Sistemas de lógica”. “La inducción propiamente dicha, en cuanto distinta de las operaciones mentales, impropriamente designadas algunas veces con este nombre, que hemos tratado de caracterizar en el capítulo anterior, puede ser brevemente definida como una generalización de la experiencia. Consiste en inferir de algunos casos particulares en que un fenómeno es observado, que se encontrará en todos los casos de una determinada clase, es decir, en todos los casos que se parezcan a los primeros en lo que ofrecen de esencial. Por qué

medio las circunstancias esenciales pueden ser distinguidas de las que no lo son y por qué algunas circunstancias son esenciales y otras no, todavía no estamos en disposición de explicarlo. Es preciso ante todo observar que hay un principio implicado en el enunciado mismo de lo que es la inducción, un postulado relativo al curso de la naturaleza y al orden del universo, a saber: que hay en la naturaleza casos paralelos; que lo que sucede una vez sucederá siempre que las dichas circunstancias se presenten. Esto es, digo, un postulado implicado en cada inducción. Y si consultamos el curso actual de la naturaleza, encontraremos en él la garantía. El universo, en cuanto lo conocemos, está constituido de modo que lo que es verdad en un caso cualquiera, es verdad también en todos los casos de una cierta naturaleza. La única dificultad consiste en saber cuál es esta naturaleza.”

Además de la posible “localidad” de la hipótesis, explicando hechos específicos, también es cierto que pueden existir otras explicaciones alternativas que expliquen esos mismos hechos. Así, en el caso de la hipótesis de Durkheim sobre el suicidio, podrían existir características comunes que explicasen ese comportamiento diferenciado entre grupos. Por ejemplo, que la enfermedad mental sea más frecuente entre aquellos que viven en las ciudades, viudos y divorciados o familias sin hijos, siendo ésta responsable de las tasas diferenciadas de suicidio que se encuentran en este grupo.

En definitiva, la hipótesis de E. Durkheim podrá ser pensada como más o menos probable en función a los hechos que explica (diversidad y abundancia) y en relación a la importancia de las explicaciones alternativas a ese mismo fenómeno, con las que compite. Una vez definida la situación, donde una hipótesis se apoya sobre unas evidencias empíricas determinadas y además compite con otras hipótesis o explicaciones alternativas, la consolidación de una hipótesis se desarrollara en función a su capacidad para explicar hechos observables adicionales. Estos hechos observables adicionales deben de ser deducidos desde la propia hipótesis, que es la que tiene que definir su existencia. Este proceso de redescubrimiento de la realidad observable, deductivamente, desde una premisa/hipótesis preestablecida empíricamente, constituye la clave central del conocimiento científico occidental. Este proceso no implica exclusivamente la comprobación o establecimiento de que una determinada realidad (asociación o distribución)

existe, sino que también la interpreta y por lo tanto “comprende” desde una lectura determinada porque ocurre un determinado fenómeno social.

Así, las hipótesis se generan desde un conjunto conocido y limitado de observaciones, y se comprueba su posible veracidad en función a su capacidad para predecir el futuro. Este futuro debe entenderse en la cronología del sujeto de conocimiento y no en el del suceso a comprobar. Para el que investiga intentando establecer una hipótesis, la comprobación de aquello que deduce se encuentra en su futuro particular, pero la observación correspondiente puede haber sucedido en el pasado, ocurrirá en el presente o, también, a su vez, en el futuro.

El segundo paso, en el proceso hipotético-deductivo, implica como hemos señalado, deducir consecuencias desde la explicación que hemos propuesto. Continuando con el ejemplo de Durkheim, podríamos afirmar que

- Si la solidaridad social de un grupo es más elevada que la de otro grupo distinto, entonces su tasa de suicidio será menor.
- Los que profesan la religión católica poseen una mayor solidaridad social que los protestantes
- Luego la tasa de suicidios será menor entre los que profesan la religión católica que entre los protestantes

En esta tarea deductiva, es fundamental controlar la pertinencia de la consecuencia que se ofrece. Por ejemplo, que la solidaridad social está asociada a las creencias religiosas y además de una forma no espuria.

El tercer paso en la tarea hipotética-deductiva viene definido por la necesidad de recoger observaciones que confirmen o nieguen la presencia de esa asociación entre solidaridad y religión, así como entre religión y tasa de suicidios. Esta etapa presenta todas las complejidades de la operativización de los conceptos, determinando las variables o factores, así como la medición y recolección de datos. Esta fase introduce un margen de incertidumbre en el establecimiento de las evidencias que avalen o nieguen la existencia de la relación que se postula, en la medida que debe evaluarse en detalle toda la construcción de variables y recolección de datos. La validez y fiabilidad de toda

esta etapa afecta a las conclusiones que se puedan establecer desde el punto de vista de la lógica tras la investigación empírica.

La cuarta y última fase en el método hipotético-deductivo encontramos que, desde la información que aporta la investigación empírica de las relaciones y distribuciones, la hipótesis (explicación) puede verse confirmada desde los hechos que predecía o por el contrario puede verse contradicha por la realidad. En ambos casos, en el caso de verificación o de falsación de la hipótesis, nos encontramos con dos lógicas diferentes: lógica de confirmar o verificar hipótesis y la lógica de contradecir o falsar hipótesis.

#### **4.4. Lógica de confirmar o verificar hipótesis**

Supongamos que los datos confirman la predicción que afirmaba el que la tasa de suicidio era menor entre los católicos que entre los protestantes. Entonces podemos considerar que la hipótesis que explica la relación entre solidaridad social y suicidio se “confirma”. No obstante, no podemos decir que la veracidad de la hipótesis este probada, sólo que la probabilidad de que sea cierta se incrementa.

Concluir desde la confirmación de las observaciones que la hipótesis está probada implicaría efectuar el siguiente argumento:

- Si la hipótesis es verdadera, entonces los hechos predecidos son verdaderos.
  - los hechos predecidos son verdaderos.
  - luego la hipótesis es verdadera.

Es decir, notado mediante los ya conocidos “p” y “q”

Si p, entonces q.

q.

Luego p.

Lo que constituye en lógica deductiva la falacia de afirmar el consecuente; por lo tanto desde el punto de vista deductivo define un argumento no válido que no prueba la conclusión. En definitiva, los hechos que se confirman (relación o

distribución) son el consecuente de la hipótesis (q). Al igual que anteriormente afirmábamos con respecto a los hechos explicados por la hipótesis (solidaridad social, por ejemplo) que estos también podrían ser explicados por hipótesis alternativas (las demencias), de la misma forma una predicción derivada de una hipótesis podría tener origen en otras hipótesis diferentes. Así, deductivamente no podemos afirmar que se pruebe la verdad de la hipótesis por el hecho de que la realidad confirme sus predicciones.

Desde una lógica inductiva si podemos, sin embargo, afirmar que si una hipótesis predice un hecho observable y esto se ve corroborado por las evidencias, su probabilidad de ser cierta se incrementa, (aún cuando no este probada definitivamente). Siguiendo el esquema que se describía en el primer paso, durante la generación contemporánea de la hipótesis en relación a los hechos, cada predicción que se confirma supone una evidencia más en apoyo de la hipótesis. Así, podemos afirmar que, tras corroborar la relación existente entre religión y suicidio, entenderíamos que son cuatro las columnas empíricas que sustentan la hipótesis.

- Los viudos y divorciados cometen más suicidios que los casados
- Las parejas sin hijos producen más suicidios que las parejas con hijos
- Los habitantes de las ciudades cometen más suicidios que los que viven en el campo.
- La tasa de suicidios es menor entre los que profesan la religión católica que entre los protestantes.

De este modo, en lógica confirmatoria, cada predicción viene a convertirse en una premisa adicional del argumento inductivo sobre el que se apoya la hipótesis o explicación.

#### **4.5. Lógica de contradecir o falsar hipótesis**

A efectos de concluir algo acerca de una hipótesis en base a su contrastación empírica, la lógica de falsar es diferente en sus resultados a la lógica de verificar. Esto viene dado porque el razonamiento que se argumenta desde una

evidencia empírica contradictoria que rechaza la hipótesis es deductivamente valido, adoptando la forma siguiente.

- Si la hipótesis es verdadera, entonces los hechos predecidos son verdaderos.
  - los hechos predecidos no son verdaderos
  - luego la hipótesis no es verdadera (es falsa)

Argumentada en notación adopta la siguiente forma silogística

Si p, entonces q.

No p.

Luego no q.

Denominada negación de consecuente y que es deductivamente valida. Por ello, existe una corriente que mantiene que las hipótesis no pueden ser verificadas (probadas), pero si pueden ser falsadas (negadas). No obstante, como ya sabemos, el criterio lógico analiza exclusivamente el argumento, pero no examina en detalle la validez de la inferencia, es decir, que la predicción está claramente implicada en la hipótesis, o que las mediciones efectuadas poseen una correlación epistémica importante o no (en definitiva que la predicción sea verdaderamente falsa o no). Dado que los testados de hipótesis emplean implícitamente dichas presunciones, dándolas por correctas, falsar una hipótesis no constituye una base determinante para rechazar categóricamente una hipótesis, es decir descartarla, suprimirla definitivamente del repertorio explicativo. Consideramos que en principio, si el confirmar una hipótesis incrementa su probabilidad de ser cierta, el ser falsada debe hacer decrecer esa probabilidad. Sin embargo, verificar o falsar no tienen el mismo peso en la práctica de la actividad científica. El falsar una hipótesis tiene mucho más peso para rechazar una hipótesis, que el correspondiente a verificar para aceptarla, en gran parte debido a la propia estructura del argumento desde un punto de vista deductivo.

En general existen dos estrategias para aportar evidencias empíricas a favor de una hipótesis y con ello consolidarla como una explicación valida de una realidad social. Una de ellas consiste en ofrecer múltiples confirmaciones de la

hipótesis en base a predicciones. La otra consiste en eliminar hipótesis alternativas mediante falsaciones.

En la estrategia confirmatoria es aceptado que cuanto mayor es el número de test predictivos que respaldan una hipótesis, y así como diversos los ámbitos en que se prueba, mayor es la probabilidad de que sea cierta. En el mismo sentido, cuanto mayor número de hipótesis alternativas son falsadas con relación a los datos empíricos así como la acumulación de falsaciones sobre una misma hipótesis, mayor probabilidad de ser cierta posee la hipótesis no falsada. En ese sentido, donde la predicción funciona como testado, una parte importante de la actividad científica está orientada a buscar explicaciones alternativas a hechos explicados o por explicar, siendo en definitiva el mejor test aquel que confirma una hipótesis y falsa, simultáneamente, las alternativas. Así, en el caso del estudio de Durkheim sobre el suicidio, una hipótesis alternativa importante venía planteada por las enfermedades mentales. Durkheim falsea dicha hipótesis y valida la suya efectuando un test de asociación entre enfermedades mentales y cifras de suicidio.

*Falsación de hipótesis causales.* Como hemos podido apreciar, un aspecto esencial de la noción "causalidad" es que su naturaleza no es exclusivamente empírica. En ese sentido, su falsación no puede apoyarse exclusivamente en un test estadístico que compruebe su validez empírica. La cuestión con las hipótesis causales es que el concepto "fuerza causal" no es directamente observable, dado que es una noción teórica. Desde los datos sólo es factible determinar el grado de covariación. En ese sentido, la falsación se planteara en términos distintos a la existencia o no de covariación. "Una hipótesis causal es falsada si la fuerza de la asociación espuria entre las variables de interés es igual a la covariación observada entre esas variables". Básicamente, porque se ofrece una explicación alternativa a esa relación que se apreciaba. Es decir, la falsación de una hipótesis de causalidad se apoya en el modo en que la covariación (no si existe o no) entre dos variables pueda ser explicada teórica y estadísticamente por otra tercera variable. En todo caso, si la covariación y la relación espuria no son iguales, no significa que la hipótesis causal se haya

validado. En la medida que los datos proceden de diseños no experimentales aún cabrán dos explicaciones posibles:

- a) Existe un efecto causal entre las variables
- b) Una causa común importante ha sido omitida

En ese sentido, una de las características del análisis causal, como parte del método correlacional, es que siempre está abierto a la posibilidad de variables comunes no presentes (explicaciones alternativas) y que puedan falsar relaciones causales aceptadas anteriormente. Desde esta perspectiva, el análisis causal posee un gran potencial de autocrítica en la medida que su eficacia depende de la capacidad de los investigadores para pensar variables importantes que puedan haber sido omitidas de la explicación. Así, la covariación es importante, pero subordinada a la reflexión sobre la sociedad y sus componentes. Al mismo tiempo, la reflexión es importante para reconocer variables o dimensiones potencialmente útiles para explicar, pero siempre es una reflexión coordinada con las variables que ya se emplean en la actualidad.

Una conclusión importante a destacar es que las explicaciones, por mejor ajuste a los datos que puedan demostrar, estarán en un riesgo importante de falsación siempre que empleen una visión parcial en la selección de variables (es decir, de la sociedad que se hace visible). No es nada bueno diseñar un modelo causal sin pensar muy a fondo que podemos estar excluyendo u olvidando, dado que la completitud en el diagnóstico es una pieza central en la existencia del modelo. No será excesivo el espacio que dediquemos a recordar la necesidad de diagnosticar variables comunes. La calidad y potencia explicativa del modelo dependerá de ello en gran medida.

No debemos olvidar el significado real de la incorporación de causas comunes: la extensión de las teorías (explicaciones). Es fundamental introducir las causas comunes a las variables para tener una teoría causal lo más completa posible. Los problemas principales que plantea la introducción de causas comunes son de dos tipos: matemáticos al incrementar la posibilidad de generar problemas de identificación del sistema y teóricos, dado que la introducción de una causa común a dos variables implica la introducción de dos



nuevas hipótesis en la teoría. Con ello, las causas comunes actúan en el ámbito teórico simultáneamente al incrementar la potencia explicativa del modelo y hacerlo más descriptivo y refinado (completo).

Recordando nuevamente el aspecto central, en un modelo causal se presume implícitamente que las causas no son únicas, y precisamente la importancia del residual (perturbación) en una relación indicará la importancia, en cantidad y calidad, de las variables que no están presentes en el modelo.

En ese sentido, las ecuaciones mediante las que se representan los modelos intentan expresar matemáticamente la definición que el investigador hace de la realidad. Esto implica una simplificación importante que mostrara un margen de error mayor o menor en función tanto de la definición de la realidad, lo que existe y lo que no existe, como de la forma operativa que se le ha dado finalmente.

De este modo, el error en los modelos procede de dos fuentes de variabilidad diferentes: el error de las variables y el error de las ecuaciones. El error de las variables se deriva fundamentalmente de los defectos en las medidas de las variables y otras causas aleatorias. Este error se da en toda clase de datos manejados en los modelos, ya tengan éstos carácter observacional o experimental. Se les conoce, como errores de medición y son básicamente atribuibles a la forma operativa que se les ha dado o al procedimiento mismo de obtención de datos.

Otro tipo de error es el que aparece asociado a las ecuaciones. En ese sentido el error asociado a una ecuación es un dato importante a considerar en la medida que indica la “parte” no explicada de la argumentación. Expresado de otro modo, un error elevado indicara una explicación incorrecta (no están todas las variables importantes) o una argumentación incorrecta (especificación incorrecta). Este error esta compuesto, por lo tanto, de las variables no introducidas en el modelo así como de errores en la determinación de la fórmula matemática exacta que une las variables entre sí (esencialmente aditiva o interacción). La primera variabilidad del error procede del carácter no experimental de los datos, lo que supone el riesgo de explicaciones

incompletas. Este error puede reducirse mediante una evaluación más detallada de la realidad (pensando más y mejor). El segundo es un error de análisis que puede ser reducido mediante una mejor especificación del modelo, es decir, realizando una tarea previa de análisis exploratorio más detallada.