CIS

Centro de Investigaciones Sociológicas

Cuadernos Metodológicos

Métodos de análisis causal

Juan Díez Medrano

El presente manual instroduce al lector al estudio de modelos causales mediante la técnica denominada LISREL. En base a datos correspondientes a España, este libro construye poco a poco un modelo de complejidad cada vez mayor. Coincidiendo con este desarrollo, se van explicando sucesivamente modelos de regresión simple y múltiple, path-análisis, modelos de ecuaciones estructurales, análisis factorial confirmatorio y modelos de relaciones estructurales con variables latentes. Aunque no supone en absoluto un sustituto a manuales con carácter mas matemático, así como al manual del programa LISREL, trata de clarificar en lo posible el significado de los conceptos y la lógica que hace de estos modelos un instrumento indispensable a la hora de estudiar empíricamente la realidad social.

3

CIS 3 Métodos de Análisis Causal. Juan Díez Medrano

CIS

Centro de Investigaciones Sociológicas

Juan Díez Medrano nació en Madrid en 1961.

Estudió en el Liceo Francés de Madrid, donde obtuvo su Baccalauréat en 1979. Posteriormente cursó estudios de Sociología en la Facultad de Ciencias Políticas y Sociología de la Universidad Complutense de Madrid, obteniendo su licenciatura en 1984.

Cursó luego estudios en la Universidad de Michigan y obtuvo el Phd en Mayo de 1989. El tema de su tesis doctoral fue un estudio comparado de la evolución del nacionalismo vasco y catalán. En la actualidad es **Assistant Professor** en el Departamento de Sociología de la Universidad de California, en San Diego. Se ha especializado en el estudio del nacionalismo y en técnicas cuantitativas aplicadas a las Ciencias Sociales.

Entre sus diversas publicaciones destaca un trabajo conjunto con Juan Díez Nicolás y Blanca García Mon, titulado *El significado de ser de izquierdas en la España actual,* publicado en la Revista Española de Investigaciones Sociológicas (Reis) en 1989.

ISBN 84-7476-161-1
00001
9 788474 761610

CIS

Centro de Investigaciones Sociológicas

Cuadernos Metodológicos

Métodos de análisis causal

Juan Díez Medrano

3

COLECCIÓN «CUADERNOS METODOLÓGICOS», NÚM. 3

Primera edición, abril de 1992

Impreso y hecho en España
Printed and made in Spain

Diseño de la cubierta: Carlos Sendín

NIPO: 004-92-018-2
ISBN: 84-7476-161-1
Depósito legal: M. 14.794-1992

Fotocomposición: EFCA, S. A.
Avda. Doctor Federico Rubio y Galí, 16. 28039 Madrid

Impreso en Closas-Orcoyen, S. L. Polígono Igarsa
Paracuellos de Jarama (Madrid)

Índice

1. INTRODUCCIÓN: MODELOS LINEARES ESTRUCTURALES CON VARIABLES LATENTES 5

2. CAUSA Y MEDICIÓN 9

LA NOCIÓN DE CAUSA 9
EL PROBLEMA DE LA MEDICIÓN 11
CAUSALIDAD Y DISEÑOS DE INVESTIGACIÓN NO EXPERIMENTALES 15

3. DE LA REGRESIÓN SIMPLE AL PATH-ANÁLISIS 21

EL MODELO DE REGRESIÓN SIMPLE 21
ANÁLISIS DE REGRESIÓN MÚLTIPLE 25
PATH-ANÁLISIS 29

4. MODELOS DE ECUACIONES ESTRUCTURALES 35

FORMULACIÓN 35
IDENTIFICACIÓN DE UN MODELO 40
ESTIMACIÓN 41
BONDAD DE AJUSTE 44

5. ANÁLISIS FACTORIAL CONFIRMATORIO 51

ERROR DE MEDIDA Y PROBLEMAS 51
ANÁLISIS FACTORIAL CONFIRMATORIO 52

6. MODELOS DE RELACIONES ESTRUCTURALES CON VARIABLES LATENTES 61

FORMULACIÓN 61
COEFICIENTES ESTANDARIZADOS Y SIN ESTANDARIZAR 66
MEDIAS Y CONSTANTES 67
COMPARACIÓN DE GRUPOS 68
CONCLUSIÓN 69

Bibliografía comentada 71

1

Introducción: Modelos lineares estructurales con variables latentes

El propósito de este pequeño manual es el de presentar del modo más simple posible una técnica estadística cada vez más utilizada por sociólogos, psicólogos, politólogos, e incluso economistas, para el análisis de relaciones causales. Los modelos de ecuaciones estructurales con variables latentes, más conocidos como LISREL *(Linear Structural Relations)* debido al nombre del programa desarrollado por JÖRESKOG y SÖRBOM para el análisis de estos problemas, abarcan toda clase de relaciones lineares causales entre variables, desde el análisis de regresión simple hasta modelos más complicados en los cuales la red de relaciones causales es más compleja y donde las variables utilizadas en el análisis están medidas por varios indicadores. En este libro se pretende avanzar paso a paso, desde los modelos más simples hasta los modelos más complejos, evitando al máximo formulaciones matemáticas y poniendo el énfasis sobre las aplicaciones prácticas. Aquel que necesite una introducción más sistemática y compleja puede acudir a las fuentes bibliográficas citadas al final de esta obra.

LISREL es una técnica utilizada para el análisis de relaciones causales y no causales entre variables. Comparte con el path análisis el que permite analizar tanto relaciones directas como indirectas. Un ejemplo de relación causal implica el decir que la cantidad de educación recibida por la persona determina el nivel de sus ingresos. Esta relación puede representarse como sigue:

$$X \rightarrow Y$$

X representa la variable educación, mientras que Y representa la variable ingresos. Un ejemplo de relación causal en la cual se producen efectos directos e indirectos es el siguiente: Imaginemos que nos ponemos a reflexionar sobre la relación entre educación e ingresos y decidimos que la relación es más compleja. Por un lado, el grado de educación proporciona el conocimiento necesario para desempeñar actividades mejor remuneradas. Por otro, el grado de educación determina el grado de contacto con gente con mayores ingresos, y es este grado de contacto con gente con mayores ingresos el que determina el tipo de trabajo que uno consigue y el nivel de ingresos asociado con él. En este caso, el modelo propuesto sugiere que la educación tiene un efecto directo sobre el nivel de ingresos y un efecto indirecto, a través del grado de contacto con gente con mayores ingresos. Puede ser representado de la siguiente manera:

$$\begin{tikzcd} X \arrow[rr] \arrow[dr] & & Y \\ & Z \arrow[ur] & \end{tikzcd}$$

Aquí, X representa el grado de educación, Z representa el grado de contacto con gente con mayores ingresos, e Y representa el nivel de ingresos. Pues bien, LISREL permite cuantificar la magnitud de estos efectos directos e indirectos, es decir el cambio que se produce en Y por cada unidad de cambio en X o Z. Por ejemplo, imaginemos que la educación del individuo está medida en años de educación completados, y que el nivel de ingresos está medido en miles de pesetas. Si el coeficiente que representa el efecto causal de X sobre Y es igual a 10, ello quiere decir que por cada año de educación adicional recibido se produce un cambio en los ingresos individuales de diez mil pesetas.

LISREL también permite tomar en consideración una faceta largamente ignorada por los científicos sociales cuantitativos, como es el del error de medición. Cuando los científicos sociales estudian la sociedad suelen empezar por formular hipótesis sobre la relación entre conceptos abstractos. Por ejemplo, se señala que cuanto mayor es el conservadurismo político de la persona, mayor es su número de hijos. Tanto la variable conservadurismo político como la variable número de hijos son conceptos abstractos cuya relación causal no puede ser cuantificada hasta disponer de indicadores que los midan. Esta traducción de cada concepto abstracto en indicadores que los midan conlleva una serie de problemas no fáciles de resolver y que presuponen una definición precisa de cada concepto abstracto antes de la elección de indicadores que los midan. En primer lugar, la relación entre el indicador y el concepto puede ser más o menos vaga. Por ejemplo, la posición más o menos favorable de un individuo hacia la nacionalización de la banca es un indicador más o menos bueno del concepto 'Izquierdismo Político'. Es bueno en la medida en que la nacionalización de la banca ha sido una medida política tradicionalmente apoyada por partidos socialistas y comunistas. Sin embargo, no cubre todo el campo del izquierdismo político y, al mismo tiempo, hay constancia de movimientos conservadores que han favorecido la nacionalización de la banca. Finalmente las actitudes ante la nacionalización de la banca pueden traducir no sólo un mayor o menor izquierdismo político sino también una consideración del grado de eficacia de tal medida. Imaginemos por otro lado que, como indicador del número de hijos, utilizamos la respuesta de los individuos a la pregunta «¿Cuántos hijos ha tenido?». En este caso no hay duda de que la respuesta a esta pregunta mide perfectamente el concepto 'Número de Hijos'. Sin embargo, aquí nos encontramos con problemas adicionales, puesto que puede haber error de medida debido a que algunos individuos pueden tener mala memoria, quieren olvidar a hijos o hijas fallecidos/as, quieren engañar al entrevistador. En definitiva, incluso en un caso tan claro como éste es posible que el indicador escogido no represente fielmente el concepto analizado.

La mayor parte de las técnicas estadísticas omiten considerar este problema de medición que hace que los coeficientes obtenidos para la relación causal entre dos variables conceptuales sean altamente cuestionables. ¿Qué grado de confianza podemos otorgar a un determinado coeficiente de regresión relativo a la asociación causal entre izquierdismo político y número de hijos, medidos por los indicadores arriba citados? Incluso en el caso de que nos limitemos a decir que ese coeficiente simplemente refleja la relación entre actitudes ante la nacionalización de la banca y número de hijos, estaríamos asumiendo que los dos indicadores se han medido sin error. Técnicas más recientes han tratado de solucionar estos problemas a través de la construcción de escalas o índices basados en distintos indicadores de un mismo concepto. Los valores obtenidos para cada uno de estos indicadores son utilizados de manera más o menos arbitraria para la confección de estos índices o escalas. Por ejemplo, el investigador selecciona tres indicadores de izquierdismo político —actitud ante la nacionalización de la banca, actitud ante la nacionalización de empresas, actitud ante el intervencionismo gubernamental en la dirección de la economía— y suma los valores obtenidos para cada individuo, creando así una escala de izquierdismo político. O bien decide atribuir mayor importancia al primer indicador, de modo que el valor obtenido para éste se multiplica por dos, el valor para los otros dos indicadores se divide por dos, y finalmente se obtiene la suma de estos tres valores ponderados. Imaginemos que tenemos dos individuos: el primer individuo obtiene un valor 4 para el primer indicador, y un valor 1 para los otros dos. El segundo individuo obtiene un valor de 3 para el primer indicador y un valor de 2 para los otros dos. Si aplicamos el primer método de confección de la escala, el individuo uno tendría un valor de 6 (4+1+1) en la escala de izquierdismo político, mientras que el individuo dos tendría un valor 7 en esta escala. Por el contrario, aplicando el segundo método, el individuo uno tendría un valor de 9 {(2 · 4) + (1 · 0,5) + (1 · 0,5)} en esta escala, mientras que el individuo dos tendría un valor de 8. El grado relativo de conservadurismo político de los individuos depende, por lo tanto, de la decisión tomada para ponderar los valores de los tres indicadores que componen la escala. Por lo tanto, un primer problema consiste en la arbitrariedad existente al otorgar distintas ponderaciones a cada uno de los indicadores. Un segundo problema consiste en justificar la elección de estos tres indicadores para representar el concepto de izquierdismo político.

El análisis factorial trató de dar solución parcial a estos problemas de manera empírica, dejando que la relación entre distintos indicadores y los conceptos que representan, así como la ponderación de cada indicador respecto a esos conceptos, se basaran en el distinto grado de asociación entre dichos indicadores en una base de datos concreta. Por ejemplo, supongamos que en una encuesta se incluyen 20 preguntas, constituyendo 20 indicadores. Estos indicadores miden una serie de conceptos subyacentes que, utilizando el análisis factorial, el investigador trata de encontrar. El análisis factorial dirá, por ejemplo, que la autoidentificación ideológica de los entrevistados y sus actitudes ante la nacionalización de la banca, la nacionalización de empresas y el intervencionismo económico gubernamental, representan un mismo factor, correspondiéndole al investigador el dar un nombre a tal factor (*i.e.* 'Izquierdismo Político'). Al mismo tiempo, el análisis factorial determina el coeficiente de ponderación correspondiente a la relación de cada uno de estos indicadores con dicho factor. El problema esencial de este método consiste en que basa sus resultados exclusivamente en la información proporcionada por una muestra de datos, sin dar lugar en absoluto a las decisiones teóricas del investigador. El peligro fundamental consiste en que otra muestra nos dijese que en realidad los cuatro indicadores citados no forman un factor sino dos o más factores, y que además sus valores de ponderación son distintos a los obtenidos en la primera muestra. Además, el análisis factorial no nos dice nada sobre lo bien que esos indicadores miden el concepto o factor obtenido. LISREL trata de solventar este problema de dos modos: En primer lugar, requiere la toma de decisión previa por parte del investigador, guiado por consideraciones teóricas, respecto a la relación de cada indicador con distintos conceptos. En segundo lugar, estima el grado de acierto del investigador a la hora de relacionar cada indicador con distintos conceptos. Finalmente, concede flexibilidad al investigador a la hora de establecer las ponderaciones de cada indicador con respecto a cada concepto. El investigador puede dejar que éstos sean determinados empíricamente o puede determinarlos él mismo, teniendo en cuenta que LISREL, en cualquiera de las dos situaciones, proporcionará índices que permiten evaluar el acierto del investigador al establecer o dejar establecer dichas ponderaciones.

En definitiva, LISREL permite saltar cualitativamente desde la medición de relaciones estructurales entre indicadores a la medición de relaciones estructurales entre los conceptos medidos por dichos indicadores, al mismo tiempo que nos proporciona índices que nos permiten evaluar la bondad global de nuestros modelos teóricos.

2
Causa y medición

La noción de causa

Desde hace mucho tiempo humanistas y científicos sociales han tratado de desvelar las relaciones entre acontecimientos sociales, de manera a comprender, manipular y predecir. Una de las estrategias que han utilizado es la búsqueda de causas y efectos. Ahora bien, la definición de qué es lo que constituye una causa y la determinación del tipo de evidencia que necesita ser recogida para establecer causalidad ha sido fuente de innumerables debates que los párrafos siguientes tratan de resumir.

La definición clásica del concepto de causalidad nos la ofrece HUME, que en su *Tratado sobre la Naturaleza Humana*, señala tres criterios fundamentales: (1) Contigüidad entre Causa y Efecto, (2) la precedencia temporal de la Causa sobre el Efecto, y (3) la conjunción constante entre Causa y Efecto; es decir, que siempre que la Causa esté presente se observe el Efecto y que siempre que la Causa esté ausente no se observe el Efecto. Tanto para HUME como para la mayoría de los positivistas éstos son los tres criterios necesarios para poder decir que una determinada relación es causal. Así, RUSSELL (1913) define la causalidad de la manera siguiente: «Dado un acontecimiento e_1, existen un acontecimiento e_2 y un intervalo temporal T, de tal manera que cada vez que e_1 se da, e_2 le sigue tras un intervalo T.

Desde la perspectiva positivista, por tanto, la distinción entre meras regularidades o correlaciones y las relaciones causales, que implican la existencia de un argumento teórico previo que ligue Causa y Efecto, no es posible y no se plantea. Para aquellos, sin embargo, que piensan que dicha distinción es posible y necesaria, el criterio enunciado por John Stuart MILL, según el cual para que podamos definir una relación como causal debemos haber excluido todo mecanismo causal alternativo, se convierte en determinante a la hora de determinar la existencia de una relación de Causa y Efecto. En otras palabras, desde el momento en que aceptamos la posibilidad de distinguir entre relaciones causales y relaciones no causales, el criterio principal sobre el que se apoyará cualquier tesis sobre la existencia de una relación causal determinada es este principio introducido por MILL.

Ello nos conduce hacia la búsqueda de métodos adecuados que nos permitan eliminar, del análisis de una determinada relación causal, causas alternativas potenciales. La mayor o menor capacidad que nuestro diseño de investigación tenga a la hora de eliminar causas alternativas es la que determina la llamada Validez Interna de nuestro diseño de investigación. Esta validez interna es máxima en los experimentos, en los cuales los individuos u objetos que constituyen nuestra unidad de análisis son asignados de manera aleatoria a un grupo experimental y a un grupo de control. En muchas ocasiones, sin embargo, razones de diversa índole (éticas, financieras, etc.) impiden la utilización del método experimental, en cuyo caso el investigador trata de diseñar investigaciones que permitan eliminar un máximo de causas alternativas potenciales. Por ejemplo, la introducción de variables de control en análisis estadísticos, o la utilización del método comparado, en base al método de la concordancia y el método de las diferencias, propuestos por MILL. El método de la concordancia implica comparar dos casos que difieran en todo, menos en el efecto, para de esta manera constreñir al máximo el número de causas potenciales, que son aquellas pocas variables causales potenciales en las que los dos casos coincidan. El método de las diferencias, por el contrario, implica comparar dos casos que no difieran en casi nada, menos en el efecto estudiado, de manera a constreñir el número de causas potenciales, que son aquellas pocas variables causales potenciales en las que los dos casos difieren.

Si bien el criterio para determinar la existencia de una relación causal propuesto por MILL ha conllevado el desarrollo de diseños de investigación que permitan eliminar un máximo de causas alternativas potenciales, el criterio de la conjunción constante entre Causa y Efecto ha conllevado el desarrollo de métodos para determinar los límites contextuales de dicha conjunción constante. Se tiende a aceptar que toda conjunción constante puede ser probada únicamente dentro de determinados límites, impuestos por lo exhaustivo de nuestra observación empírica. Incluso aquellas relaciones que nos parecen más constantes podrían no darse en el futuro, al que no tenemos acceso. Para determinar el grado de regularidad con el que se producen determinadas relaciones, los metodólogos han propuesto diversos métodos de mayor o menor complejidad.

En general, se puede decir que cuanto mayor sea el número de ocasiones en que se observa una determinada relación causal, mayor es nuestra confianza en su generabilidad. Sin embargo, no es lo mismo observar dicha relación causal en varios contextos similares que observarla en contextos que difieran tanto como sea posible entre sí. Nuestra capacidad para decidir la magnitud del ámbito contextual en el que se da tal relación causal es mayor en este segundo caso. Sin embargo, el método más adecuado para determinar la generabilidad de una relación causal es el muestreo probabilístico. Los métodos de muestreo nos ayudan a determinar en qué contextos se produce una determinada relación causal y nos ayudan a determinar con un cierto grado de confianza si la relación causal observada en una muestra se da en la población de la que se extrajo dicha muestra.

Llegados a este punto, merece la pena volver a Bertrand RUSSELL para reconsiderar el concepto de causa. RUSSELL señala con acierto que encontrar procesos causales que cumplan el requisito de la conjunción constante es altamente improbable y que, incluso si ésta se da en el presente y en el pasado, no sabemos qué ocurrirá en el futuro. Además un análisis de la práctica científica demuestra que el investigador raramente se dedica exclusivamente a la búsqueda de tales procesos causales. RUSSELL propone abandonar de una vez por todas el concepto de causa y que, en lugar de ello, nos dediquemos a formular leyes científicas, es decir, relaciones funcionales del tipo $e = mc^2$. Ateniéndose, en sentido amplio a esta recomendación, los científicos se han dedicado a especificar el sentido y la magnitud de relaciones entre tipos de acontecimientos y a circunscribir el ámbito en el que se aplican determinadas relaciones. Por ejemplo, en lugar de intentar determinar si el descenso de los tipos de interés va siempre seguido de un aumento en la tasa de crecimiento económico, tratamos de determinar cuál es la probabilidad de que un descenso en los tipos de interés se traduzca en un mayor crecimiento económico, o tratamos de estimar cuál es el impacto sobre el crecimiento económico de cambios en los tipos de interés. Además, intentamos medir cómo varía esta relación a través del tiempo y el espacio.

Si siguiéramos utilizando la definición tradicional de causa, sólo podríamos utilizar este término cuando una relación cumple los cuatro criterios citados en esta sección. En la práctica, sin embargo, los investigadores hablan de causa cuando existe una relación entre dos variables y si se puede descartar la idea de que tal relación sea espúrea. En lugar de distinguir entre relaciones causales y relaciones no causales, distinguimos entre relaciones causales fuertes, relaciones causales débiles y relaciones no causales.

El problema de la medición

La mayoría de nuestras descripciones, explicaciones o predicciones sobre la sociedad se basan en la medición consciente o inconsciente. Ello no significa que todo conocimiento deba basarse en la medición de características sociales. Ciertamente algunas corrientes filosóficas creen en la existencia de otras fuentes de conocimiento: revelaciones, intuiciones, razonamientos deductivos. Hay que tener en cuenta, sin embargo, que incluso el razonamiento más deductivo se basa en ciertos supuestos previos que uno contrasta con la realidad. De este modo, incluso la lógica deductiva se basa en la medición previa de la realidad objetiva. La lógica inductiva se convirtió en la rama dominante dentro de la ciencia únicamente en tiempos recientes. Hoy, las ciencias sociales tienden a basar sus conocimientos en la medición sistemática de la realidad.

La medición consiste en dos etapas principales: la primera es la definición de lo que se va a medir. La segunda es la operacionalización de lo que va a medirse, en base a esa definición. Respecto al primer punto, hay que notar que uno de los problemas más serios con los que se enfrentan las ciencias sociales es la falta de consenso respecto a la definición de los conceptos utilizados, así como la incapacidad para darse cuenta de que la falta de concordancia entre los resultados de diversos tests para determinadas hipótesis se debe a que los investigadores están utilizando definiciones y operacionalizaciones distintas de un mismo concepto. Por tanto, el primer paso en toda investigación consiste en la definición de los conceptos utilizados. Para BOLLEN (1989):

> Un concepto es una idea que une distintos fenómenos (*i.e.* actitudes, comportamientos, rasgos) bajo una misma etiqueta. El enfado, por ejemplo, es el elemento común que une distintas manifestaciones individuales como el elevar la voz, el lanzamiento de objetos, el enrojecimiento o el comportamiento descontrolado. El concepto de enfado resume una serie de características exhibidas por los individuos. Otros conceptos juegan un papel similar [...] ¿Podemos acaso hablar de la existencia de conceptos? Los conceptos tienen la misma realidad o falta de realidad que otras ideas. Los crea la gente que cree que determinados fenómenos comparten algo en común. El concepto identifica aquellos rasgos compartidos por dichos fenómenos.

Una vez que un concepto ha sido seleccionado, los cuatro pasos siguientes son: (1) dar un significado al concepto, (2) identificar las dimensiones que lo representan, (3) crear indicadores, y (4) especificar la relación entre los indicadores y las variables latentes (o conceptos). El primer paso conlleva la formulación de una definición teórica. Una definición teórica explica en términos lo más simple posibles el significado de un concepto. Esta definición teórica cumple determinadas funciones. Una de ellas es la de unir una etiqueta al concepto. En segundo lugar, la definición teórica enumera las dimensiones del concepto. Cada concepto, efectivamente, consta de varias dimensiones. Las dimensiones de un concepto no pueden ser divididas fácilmente en componentes adicionales... En tercer lugar, la definición teórica provee información respecto al tipo de indicadores que uno debe seleccionar.

El concepto de terrorismo puede servir de ejemplo. La CIA lo define como sigue: «La amenaza o el uso de la violencia por motivos políticos por individuos o grupos, en favor o en contra de la autoridad gubernamental establecida, cuando dichas acciones tienen como fin el conmocionar o intimidar a un grupo mayor que el de sus víctimas inmediatas.» Esta definición tiene, como se puede apreciar, tres dimensiones fundamentales. La primera es cualquier amenaza o uso de violencia por motivos políticos; la segunda es la existencia de un objetivo que trasciende al individuo o grupo objeto de tal violencia; la tercera es el que la acción sea en favor o en contra de la autoridad gubernamental establecida.

El paso siguiente en el proceso de medición, la selección de indicadores, depende directamente de la definición teórica. Este paso es denominado fase de definición operacional. La definición operacional describe los procedimientos a seguir para seleccionar indicadores de la variable latente o variables latentes (si el concepto tiene varias dimensiones) que representan a un determinado concepto. En determinadas situaciones la variable o variables latentes son operacionalizadas en base a respuestas individuales a un cuestionario. Otras veces dichos indicadores provienen de estadísticas recogidas por la administración pública, ya sean datos del censo o de los registros civiles, etc... Un indicador determinado es apropiado en la medida en que provea al investigador con una variable empírica que corresponda al significado asignado al concepto. Muchas veces, sin embargo, la información proporcionada por un indicador no es completa, y una cantidad considerable de sentido común entra en juego a la hora de asignar indicadores a un determinado concepto. La diversidad de las fuentes utilizadas, la profesionalidad de los codificadores, son factores que influyen sobre la calidad de los indicadores. En el caso del terrorismo, la calidad de su medición dependerá de la diversidad de fuentes de información utilizadas, del cuidado tomado por los reporteros, e incluso de la capacidad de acceso a determinadas áreas geográficas. Por ejemplo, las fuentes de información occidentales tienden más a hablar de actos terroristas cuando estas acciones son tomadas por países o grupos hostiles al mundo capitalista occidental.

Prácticamente todas las medidas que utilizamos contienen error de medición. Por eso el cuarto paso en el proceso de medición consiste en formalizar ese tipo de errores. No hay criterios definitivos respecto a la definición de un concepto; tal como indiqué anteriormente, las ciencias sociales se caracterizan por su falta de consenso respecto a conceptos importantes. Sin embargo, significaría un importante avance el empezar a crear consenso tratando de (1) especificar claramente las definiciones utilizadas, (2) utilizar en las investigaciones las definiciones utilizadas más frecuentemente, y (3) siempre que sea necesario y posible, evaluar la calidad representativa de la realidad de las diferentes definiciones de los conceptos utilizados. En cuanto a la operacionalización de los conceptos, los científicos sociales han desarrollado técnicas para juzgar la bondad de distintos indicadores a la hora de medir los conceptos utilizados. Los dos criterios principales utilizados son la validez y la fiabilidad.

La validez de un concepto se refiere al grado en que una variable determinada mide lo que se supone debe medir. ¿Hasta qué punto, por ejemplo, podemos decir que la información sobre terrorismo incluida en los medios de comunicación mide realmente el terrorismo? ¿Hasta qué punto los tests de inteligencia miden la inteligencia? ¿Hasta qué punto el Producto Nacional Bruto mide el valor real de los bienes y servicios producidos en un país? Estas cuestiones se refieren a la validez de distintos indicadores, que nunca se puede determinar de modo absoluto. Sin embargo, aunque nunca podemos probar la validez de un concepto, sí que podemos obtener evidencia que determine el grado mayor o menor de validez de un indicador.

Existen medios teóricos y empíricos de determinar la validez de un indicador. Los primeros definen lo que se denomina validez de contenido. La validez de contenido se refiere al grado en que los indicadores de un concepto cubren todas sus dimensiones. En tanto en cuanto lo hagan podemos hablar de la validez de contenido de los indicadores. La pregunta crucial que nos podemos hacer entonces es ¿cómo sabemos cuáles son las dimensiones de un concepto? Para responder a esta pregunta debemos volver al primer paso dentro del proceso de medición del que hablamos en la primera parte de este capítulo. Es decir, que para poder tomar en cuenta todas las dimensiones de un concepto es necesario tener una definición teórica previa de ese concepto. En determinadas ocasiones, sin embargo, nuestros instrumentos de medida no nos permiten considerar todas las dimensiones de un determinado concepto. Lo mejor en tales ocasiones es reconocer el carácter parcial de nuestros resultados.

La limitación principal del criterio de la validez de contenido de un indicador es que depende de la definición teórica del concepto. Para la mayoría de los conceptos utilizados en las ciencias sociales no existe un consenso absoluto sobre su definición teórica. El investigador debe en estas situaciones no sólo proveer una definición teórica aceptada por los demás colegas sino también obtener indicadores que cubran completamente todas las dimensiones del concepto. Se han sugerido distintos métodos empíricos para determinar la validez de los indicadores para un concepto determinado. Ninguno de ellos es enteramente apropiado. Uno de estos métodos consiste en evaluar la validez de un indicador en base a su asociación con un indicador que supuestamente mide perfectamente el concepto en consideración. Otro de los métodos evalúa la validez de un indicador en base al grado en que sus asociaciones con otros indicadores de otros conceptos se adecúa a las predicciones sobre dichas asociaciones. BOLLEN, finalmente sugiere utilizar como criterio la asociación entre el concepto y su indicador, obtenida de modo empírico utilizando LISREL. Dado que ninguna de las estrategias empíricas es completamente acertada, la estrategia teórica es cuanto menos una condición necesaria a la hora de evaluar la validez de un indicador. La utilización de métodos empíricos debería completarla, siempre con plena conciencia de sus limitaciones.

La fiabilidad de un indicador es la consistencia con que mide un concepto. No es igual a su validez y, de este modo, podemos tener medidas que son a la vez fiables pero inválidas. Por ejemplo, el peso proporcionado por una balanza bien calibrada es un indicador fiable del peso real de un objeto determinado. Pero este mismo peso es un indicador fiable pero no válido de la temperatura de un cuarto.

La fiabilidad de un indicador sólo se puede medir de forma empírica y existen diversos métodos. BOLLEN cita unos cuantos de estos métodos: El primero de ellos es la técnica consistente en repetir una misma medición dos veces y en calcular el coeficiente de correlación entre las dos mediciones. El segundo, consiste en utilizar dos indicadores de un mismo concepto en dos ocasiones distintas y en calcular la correlación entre los valores obtenidos en las dos mediciones. Los coeficientes de correlación obtenidos con los dos métodos constituyen medidas de fiabilidad. El tercer método enumerado por BOLLEN es el consistente en tomar varios indicadores de un mismo concepto y dividirlos de forma arbitraria en dos grupos. Dentro de cada grupo se combinan los valores obtenidos para cada uno de los indicadores y, finalmente, se calcula el coeficiente de correlación entre la variable compuesta obtenida para cada grupo. Este coeficiente de correlación señala la fiabilidad de los indicadores incluidos en los dos grupos. BOLLEN, por su parte, sugiere tomar como medida de fiabilidad el coeficiente de correlación múltiple entre un indicador y el concepto o conceptos que éste mide supuestamente. En un capítulo ulterior exploraremos de modo más detallado las técnicas sugeridas por BOLLEN para determinar la validez y fiabilidad de un indicador determinado.

Es, por lo tanto, importante reconocer que las medidas que utilicemos para cada concepto utilizado en nuestra investigación contienen un determinado grado de error. Algunas veces este error es debido al azar; en otras ocasiones se trata de un error sistemático. Nuestra función como investigadores consiste en proporcionar definiciones claras de los conceptos que utilicemos, asegurarnos de que nuestros indicadores miden cada una de las dimensiones de un concepto determinado o, si no es así, ser explícitos respecto a las limitaciones de nuestros indicadores y, finalmente, utilizar varias medidas alternativas del mismo concepto, que, en determinados casos puedan ser introducidas dentro de un mismo modelo estadístico que tenga en cuenta la existencia de error de medida.

Causalidad y diseños de investigación no experimentales

Cualquier diseño de investigación interesado en la explicación de relaciones causales entre variables intentará demostrar que:

1. El indicador x sirve para medir el concepto X y lo hace con fiabilidad.
2. El indicador y sirve para medir el concepto Y y lo hace con fiabilidad.
3. La relación entre X e Y no sólo existe sino que además es causal, es decir, no se debe al efecto causal simultáneo de terceras variables T.

Ejemplo de relación causal entre X e Y:

$$\begin{tikzcd} X \arrow[rr] \arrow[dr] & & Y \\ & T \arrow[ur] & \end{tikzcd}$$

Por ejemplo, el nivel de estudios de una persona *(X)* determina sus ingresos *(Y)*, o el nivel de estudios de una persona determina sus ingresos debido a los conocimientos que proporciona *(T)*. Tanto uno como otro modelo implican la existencia de una relación causal entre el nivel de estudios e ingresos.

Ejemplo de relación no causal entre X e Y:

$$\begin{tikzcd} & X \arrow[d, bend left, leftrightarrow] \\ T \arrow[ur] \arrow[r] & Y \end{tikzcd}$$

Por ejemplo, alguien podría argumentar que la relación entre nivel de estudios e ingresos es espúrea, es decir que aunque existe una asociación entre las dos variables, ésta no se debe al efecto causal del nivel de estudios sobre los ingresos, sino a que tanto el nivel de estudios como los ingresos vienen determinados por el origen social de las personas.

Una vez descartada la posibilidad de que la relación entre las dos variables sea espúrea, el investigador diseña su investigación de manera que pueda explicar dichas relaciones causales mediante la intervención de terceras variables. Es decir que, por ejemplo, el investigador no se contenta con demostrar que existe una relación causal entre el nivel de estudios de una persona y sus ingresos, sino que también intenta explicar esta relación causal por la mediación de terceras variables (por ejemplo, el nivel de conocimientos proporcionado por un determinado nivel de estudios).

Pasamos de un modelo como el siguiente:

$$X \rightarrow Y$$

a otro modelo más complejo como el siguiente:

$$\begin{tikzcd} & Y \\ X \arrow[ur] \arrow[r] & T \arrow[u] \end{tikzcd}$$

Además, el investigador se interesa muchas veces en analizar la relación causal entre dos variables, dados varios contextos diferentes. Por ejemplo, intenta averiguar si la relación causal entre el nivel de estudios y los ingresos de las personas es igual en España que en Estados Unidos. En este ejemplo, el país considerado constituye la variable *T* utilizada para analizar la relación causal entre *X* e *Y* en distintos contextos.

Sólo la distribución aleatoria de los objetos de análisis (ya sean personas, agregados de personas, o lo que sea) en un grupo experimental y otro de control permite alcanzar estos objetivos. Por ejemplo la mejor manera de determinar si el nivel de estudios de las personas determina sus ingresos consistiría en asignar aleatoriamente, es decir mediante cualquier procedimiento que se base en la suerte o azar, quién va a la Universidad, quién abandona sus estudios al final de la secundaria, para después de algún tiempo medir los ingresos obtenidos por estas personas después de haber iniciado su ejercicio profesional. Por supuesto, tanto este experimento como muchos otros chocan con obstáculos éticos que hacen que la mayoría de las veces nos tengamos que contentar con aproximaciones al diseño experimental. En estas situaciones, la asignación aleatoria de los objetos de estudio al grupo experimental y al grupo de control no es posible y ello hace que nunca estemos seguros de que la asociación entre dos variables es causal y no espúrea. Lo más que podemos hacer es ir eliminando hipótesis alternativas.

Un último objetivo ligado a muchos diseños de investigación es el comparar el efecto causal de una variable con el de otras. Por ejemplo, podemos intentar determinar si el nivel de estudios es más importante a la hora de determinar los ingresos de las personas que los ingresos de los padres de dichas personas. En última instancia tratamos de hallar un número de variables determinado que nos permita predecir de la manera más precisa posible la variación de la variable dependiente *Y*. Por ejemplo, intentamos encontrar un número determinado de variables que nos permita predecir con máxima certeza los ingresos de las personas.

En resumen, el investigador diseña su trabajo de modo que pueda determinar si dos o más variables están relacionadas, de manera que pueda determinar si dicha relación es causal, de manera que determine qué variables median o preceden a esa relación causal, de manera que pueda determinar la variabilidad de la relación causal en varios contextos distintos, de manera que determine la existencia de otras variables explicativas y su efecto causal relativo sobre la variable dependiente, y de manera que maximice nuestra capacidad de predicción de la variable dependiente *Y*.

La mayoría de las veces el investigador analiza relaciones causales dentro de diseños no experimentales, es decir en situaciones en las cuales los objetos de estudio no han sido asignados aleatoriamente a los grupos experimental y de control. Por ejemplo, no podemos decidir de forma aleatoria sobre el nivel de estudios alcanzado por una persona. Independientemente de cuestiones éticas, es obvio que la sociedad se opondría a que un investigador decidiese aleatoriamente cuándo deben acabar sus estudios los individuos escogidos para un estudio de las relaciones entre nivel de estudios e ingresos. Los individuos escogidos para nuestro estudio vienen ya con un nivel de estudios dado que no podemos determinar libremente. En dichas situaciones, la única manera de establecer una relación causal entre nivel de estudios *(X)* e ingresos *(Y)* consiste en eliminar cuantas variables explicativas potenciales sea posible. Otras vías posibles para reforzar empíricamente la hipótesis sobre la existencia de una relación causal entre las dos variables consistiría en verificar el cumplimiento de predicciones respecto a las variables *(T)* que median aquella relación o en verificar el cumplimiento de predicciones respecto al tipo de relación existente entre las dos variables cuando variamos el contexto donde esta relación se desarrolla. Estas dos últimas alternativas en ningún modo tienen el mismo valor analítico derivado de descartar el máximo número de variables explicativas alternativas.

En resumen, cuando nos hallamos en condiciones no experimentales no nos queda más remedio que introducir variables de control *(T)* en el análisis, para asegurarnos que la relación entre X e Y no es espúrea. Las variables T incluidas dentro de modelos estadísticos cumplen varias funciones. En primer lugar pueden cumplir una función interpretativa. Es decir que introducimos T en el análisis para mostrar de qué modo se produce la relación entre X e Y o, en otras ocasiones, para mostrar qué variables antecedentes originan la relación entre X e Y:

a1.

$$\begin{tikzcd} & T \arrow[dr] & \\ X \arrow[ur] \arrow[rr] & & Y \end{tikzcd}$$

a2.

$$\begin{tikzcd} & X \arrow[dr] & \\ T \arrow[ur] \arrow[rr] & & Y \end{tikzcd}$$

En otras ocasiones introducimos variables T porque pensamos que la relación entre X e Y es espúrea:

a3.

$$\begin{tikzcd} & X \arrow[d, bend left, leftrightarrow] \\ T \arrow[ur] \arrow[r] & Y \end{tikzcd}$$

En otras ocasiones introducimos variables T porque queremos determinar la capacidad explicativa relativa de distintas variables:

a4.

$$\begin{tikzcd} X \arrow[dr] \arrow[dd, bend right, leftrightarrow] & \\ & Y \\ T \arrow[ur] & \end{tikzcd}$$

Somos nosotros los que a partir de argumentaciones teóricas decidimos el rol que juega la variable T dentro de nuestro análisis. Los datos nunca nos dirán cuál de los tres modelos presentados en las líneas precedentes es más adecuado. Lo único que nos dirá el análisis estadístico es la validez de nuestro modelo así como el valor de los distintos componentes de nuestro modelo. Hay ocasiones en las cuales nos es fácil decidir cuál de los cuatro modelos precedentes es más adecuado. Estas ocasiones se producen cuando existe clara precedencia temporal de unas variables sobre otras. Sin embargo, el diseño de investigación utilizado en algunas ocasiones facilita y en otras dificulta la determinación de un orden de precedencia. Los diseños longitudinales (por ejemplo, dos tandas de entrevistas al mismo grupo de personas) la facilitan, mientras que los diseños no longitudinales (por ejemplo, una sola entrevista a un grupo determinado de personas) la dificultan.

Imaginemos, por ejemplo, que queremos analizar la relación causal entre la ideología política y los ingresos. Se podría argumentar que la gente de izquierdas, al valorar menos el bienestar material, tenderá a comportarse de manera a obtener menos ingresos, mientras que la gente de derechas, al valorar más dicho bienestar material, tenderá a comportarse de manera a obtener más ingresos. Del mismo modo se podría argumentar que la gente que tiene menos ingresos tenderá a desarrollar una ideología más de izquierdas, mientras que la gente que tiene más ingresos tenderá a desarrollar una ideología más de derechas. Un diseño no longitudinal difícilmente podrá demostrarnos el orden de precedencia causal de una variable sobre la otra. Lo más que podremos observar en la mayoría de las ocasiones es el grado de asociación entre las dos variables. Los datos nunca nos dirán el orden de precedencia apropiado. Sin embargo, si tenemos datos para las mismas personas recogidos en dos momentos distintos, podemos al menos establecer la relación causal de una variable sobre la otra en un período de tiempo determinado.

En resumen, los diseños de investigación no experimentales, mayoritariamente utilizados por científicos sociales, impiden el que podamos estar seguros sobre la existencia de una relación causal entre dos variables. Nuestro esfuerzo va encaminado la mayoría de las veces a introducir variables T en nuestros modelos que nos permitan determinar la importancia relativa de nuestra variable X a la hora de explicar Y, así como la posibilidad de que la relación entre X e Y sea espúrea. Una vez descartadas una serie de variables explicativas alternativas para la asociación existente entre X e Y, podemos empezar a introducir otras variables T en nuestros modelos que nos ayuden a interpretar la relación entre X e Y. Ahora bien, los datos nunca nos dirán el rol que juega T dentro de estos modelos, del mismo modo que no nos dicen si X causa Y o viceversa. Solamente nuestra teoría puede indicarnos tanto lo uno como lo otro.

3

De la regresión simple al path-análisis

En los capítulos que siguen iremos elaborando un modelo cada vez más complejo para el análisis de relaciones causales basado en la utilización del programa LISREL. Ello conllevará la repetición de temas abordados más extensamente por libros especializados. El objeto de este manual es, no obstante, el mostrar cómo podemos estimar cualquier modelo linear utilizando LISREL. El análisis de regresión simple y múltiple, el path-análisis, el análisis factorial confirmatorio, y el path-análisis basado en las relaciones entre factores o variables latentes, son todas ellas técnicas específicas dentro de lo que globalmente podemos llamar modelos lineares, es decir modelos en los cuales asumimos que las relaciones entre variables son lineares.

El modelo de regresión simple

En los modelos de regresión simple intentamos analizar la relación causal entre dos variables *x* e *y*. Recordemos brevemente los criterios fundamentales enunciados para poder determinar que una relación es causal:

- Existe una asociación entre *x* e *y* más o menos fuerte.
- *x* precede a *y*, al menos teóricamente si no en el momento de la medición.
- Hemos rechazado a través del diseño o estadísticamente toda explicación *t* alternativa.

En todo modelo de regresión simple un supuesto básico, raramente hecho explícito, es que *x* e *y* son indicadores que miden con absoluta validez y fiabilidad los conceptos *X* e *Y* que pretenden representar. La variable *x* es llamada variable independiente y la variable *y* es llamada variable dependiente. La relación causal linear entre *X* e *Y* puede representarse matemáticamente de la siguiente manera:

$$Y = \text{Alpha} + \text{GA}(X)$$

(Utilizaremos la notación LISREL de ahora en adelante para familiarizar al lector. Los manuales de estadística generalmente utilizan el símbolo Beta (BE) a la hora de representar la pendiente de esta relación linear. LISREL, sin embargo, subdivide los coeficientes Beta en dos tipos, GAmma y BEta, dependiendo del carácter de las variables consideradas. Cuando la variable independiente no es, asimismo, causada por otra variable, se trata de una variable exógena, la denominamos *X* y el coeficiente que representa su efecto causal sobre *Y* es llamado GAmma. Cuando la variable independiente es, asimismo, causada por otra variable se trata de una variable endógena; la denominamos *Y* y el coeficiente que representa su efecto causal sobre la otra variable *Y* es llamado BEta.)

Naturalmente, es ilusorio pensar que *X* sea la única variable que afecte a *Y*, o que, en otras palabras, nuestro conocimiento de *X* nos permita predecir perfectamente el valor de *Y*. Generalmente, existen otras variables *X* que sirven para explicar *Y* y que no incluimos en el modelo. De este modo, un modelo más apropiado incluirá siempre un efecto residual que resume el efecto sobre *Y* de todas aquellas otras variables causales *X* no incluidas en el modelo:

$$y = \text{Alpha} + \text{GA}(X) + \text{ZE}$$

Este nuevo modelo indica que la variabilidad de *Y* es función de la variabilidad de *X* y de un elemento residual que incluye el efecto de todas aquellas variables que de algún modo influyen sobre *Y*. El coeficiente GAmma representa la magnitud del efecto de *X* sobre *Y*. Indica cuántas unidades de medida cambia *Y* por cada unidad de cambio en *X*. El coeficiente Alpha representa el valor de *Y* cuando *X* es igual a cero. Con la intención de simplificar la presentación de los modelos que vamos a desarrollar de ahora en adelante, asumiremos que tanto la variable *X* como la variable *Y* han sido centradas, es decir que hemos cambiado sus escalas de medida respectivas, sustrayendo a cada uno de sus valores la media de *X* o la media de *Y* según se trate de *X* o de *Y*. Al centrar las dos variables, el efecto de *X* sobre *Y* permanece inalterado, pero el coeficiente Alpha pasa a ser igual a cero. De este modo, nuestro modelo de regresión linear puede ser reformulado como sigue:

$$Y = \text{GA}(X) + \text{ZE}$$

Un ejemplo servirá para aclarar este tema. Supongamos que nos interesa estudiar la relación entre educación e ingresos. Nuestra hipótesis es que, cuanto mayor es el nivel de estudios de los individuos, mayores son sus ingresos, y queremos evaluar la elasticidad de la variable ingresos ante cambios en la variable educación. En este ejemplo la variable *X* es el nivel de estudios y la variable *Y* es el nivel de ingresos. Nuestro primer supuesto es que tanto una como otra variable han sido medidas perfectamente. Nuestro segundo supuesto es que la relación entre las dos variables es linear. Utilizaremos el programa LISREL para resolver este modelo. Las instrucciones proporcionadas son las siguientes:

```
Modelo 1 Regresión Simple
ni = 6 no = 376 ma = cm
la
*
ideol fcohab reldad nacban ingr educ
da fi = Reg1.lis
sd fi = Reg1.sd
se
5 6/
mo nx = 1 fi ny = 1 ne = 1 te = zero ly = id
ou se tv
```

(Explicar el lenguaje de programación de LISREL ocuparía más páginas de las que disponemos. Cualquier persona interesada puede acudir al manual del programa.)

El modelo resultante es el siguiente:

$$Y = 0{,}479X + \text{ZE}$$

El coeficiente 0,479 indica que por cada unidad de cambio en el nivel de estudios alcanzado por la persona se produce un cambio de 0,479 unidades en la variable ingresos. Dado que la variable nivel de estudios tiene en este ejemplo 9 valores y que la variable ingresos tiene 12 valores, este resultado sugiere que un cambio de nueve unidades en el nivel de estudios se traduce en un cambio de $9 \cdot 0{,}479 = 4{,}31$ unidades en la variable ingresos, es decir un 36 % del recorrido total de la variable ingresos. La magnitud de la asociación no es, pues, excesiva, pero tampoco desdeñable. Otra forma de evaluarla es a través del coeficiente de correlación simple, que mide la fuerza de la asociación entre las dos variables. Este coeficiente, que fluctúa entre -1 y 1 y cuyo valor se acerca a sus límites a medida que aumenta el grado de asociación, es en esta ocasión igual a 0,49. Se trata de una asociación moderada, tirando a fuerte.

Si elevamos al cuadrado dicho coeficiente, obtenemos R^2, que nos indica la proporción de varianza explicada en Y por X. Efectivamente, si no tuviésemos información para predecir los valores de Y y tuviésemos que predecir el nivel de estudios de una persona cualquiera, nuestra mejor apuesta sería la de sugerir el valor medio para dicha variable en nuestra muestra. En este caso, se trata del valor 5,01, equivalente a unos ingresos entre 75.000 y 100.000 pesetas. Al introducir en nuestro modelo información sobre el nivel de estudios de esas mismas personas, nuestra capacidad predictiva aumenta en relación a nuestra capacidad predictiva cuando sólo teníamos información respecto a y. El coeficiente de correlación múltiple al cuadrado nos indica el grado de mejora con respecto a la media obtenido en nuestra predicción, al introducir información sobre el nivel de estudios de las personas. En este caso, un coeficiente de correlación múltiple igual a $0{,}49 \cdot 0{,}49 = 0{,}236$ indica que la información sobre el nivel de estudios nos ayuda a explicar un 24 % de la varianza en Y, es decir de aquella variación existente alrededor de la media de Y. Un 76 % de esta variación queda todavía sin explicar, sin embargo $(1 - 0{,}236 = 0{,}764)$.

Por el momento, claro está, hemos procedido como si estuviéramos estudiando a la población en su conjunto, y como si los resultados obtenidos al estimar el modelo de regresión simple se aplicaran directamente a la población. Esto no es así, sin embargo, pues estamos operando con una muestra extraída de la población. Ello no obsta para que sigamos interesados en saber qué ocurre con esta asociación entre las dos variables en la población en general. Para ello acudimos a la teoría sobre distribuciones muestrales. Ésta nos dice que dado un coeficiente determinado para la relación entre dos variables en la población, si extraemos una infinidad de muestras de igual tamaño y por los mismos procedimientos, y en base a cada una de ellas estimamos el correspondiente coeficiente de regresión, los coeficientes obtenidos van a seguir una distribución particular, la t de STUDENT, con una media igual al coeficiente real en la población y con una desviación típica, o variación, específica. Sabiendo esto, podemos calcular tests de significación estadística que nos digan la probabilidad de obtener un coeficiente de regresión como el obtenido, dado un determinado valor de ese coeficiente en la población. Naturalmente, ignoramos tal valor y por eso utilizamos muestras. Sin embargo, si nuestro interés radica simplemente en saber si la relación entre las dos variables en la población es diferente de cero (ausencia de toda relación), no tenemos más que calcular la probabilidad de obtener un coeficiente de regresión muestral determinado o mayor en términos absolutos, dado un coeficiente de regresión igual a cero en la población. Si la probabilidad es muy pequeña podemos descartar la hipótesis de que no existe relación alguna entre las dos variables en la población, sujetos siempre a un pequeño margen de error determinado por la probabilidad de obtener un coeficiente de regresión muestral determinado si en la población dicho coeficiente fuera igual a cero.

En general, los científicos sociales rechazan la hipótesis de no asociación cuando dicha probabilidad es inferior al 5 %. Para calcular dicha probabilidad, no hay más que saber la distribución muestral del coeficiente de regresión dado, si éste fuese igual a cero en la población. Ésta, ya lo dijimos arriba, tendrá una media igual al coeficiente existente en la población, que aquí suponemos igual a cero, y una determinada dispersión que cualquier programa estadístico calcula y que se denomina error estándar del coeficiente de regresión *(Standard Error)*. En base a éste podemos estimar la probabilidad que andamos buscando. Para ello no tenemos más que dividir nuestro coeficiente muestral por su error estándar. Esta operación nos da el valor del coeficiente muestral dentro de la distribución de t. No se trata sino de una estandarización debida al hecho de que la distribución muestral de un determinado coeficiente puede tener muy diversos errores estándar. La estandarización se produce al dividir el coeficiente muestral por el error estándar. Es lo que denominamos valor t. Para saber la probabilidad de obtener un valor t determinado no tenemos sino que acudir a la tabla de valores de la distribución de t existente en cualquier manual de estadística. En el caso que nos interesa, el valor de t obtenido es de 10,751. Si estuviéramos interesados en rechazar la hipótesis de que en la población la relación entre nivel de estudios e ingresos es igual a cero cuando la probabilidad de obtener un coeficiente igual o mayor al obtenido en nuestra muestra, dado un coeficiente igual a cero en la población, sea igual o inferior al 5 %, un valor tan alto como este de 10,751 nos permitiría rechazar tal hipótesis cómodamente. Efectivamente, cualquier valor de t superior a 1,96 nos permitiría rechazar la hipótesis de no relación entre las dos variables en la población. Por tanto, este pequeño ejemplo nos sugiere que existe una relación moderada entre nivel de estudios e ingresos y que esta relación es significativamente diferente de cero desde un punto de vista estadístico.

En este pequeño subcapítulo hemos tratado de resumir los rudimentos básicos del análisis de regresión simple desde un punto de vista interpretativo, y hemos sugerido cómo estimar dicho modelo utilizando LISREL. Los índices básicos que hemos introducido han sido el coeficiente de regresión, el coeficiente de correlación, el coeficiente de correlación múltiple, y el test de significación estadística para el coeficiente de regresión en base a la utilización de la distribución muestral de la t de STUDENT. En los siguientes subcapítulos iremos complicando este modelo hasta llegar al path-análisis.

Análisis de regresión múltiple

El análisis de regresión múltiple consiste en aquellos modelos en los cuales la variable dependiente es explicada por más de una variable independiente. Estimamos este tipo de modelos cuando queremos cumplir tres objetivos principales: El primero de ellos es explicar tanto como nos sea posible la variación en la variable dependiente; el segundo de ellos consiste en evaluar el efecto causal relativo de distintas variables independientes; el tercero de ellos consiste en observar el comportamiento del efecto causal de una variable cuando introducimos una variable explicativa adicional en el modelo. La formulación del modelo es la siguiente:

$$Y = \text{GA}_1\ (X_1) + \text{GA}_2\ (X_2) + \ldots\text{GA}(X_n) + \text{ZE}$$

La significación de los coeficientes en este modelo cambia sin embargo. Cada uno de los coeficientes GAmma sigue significando el número de unidades de cambio en Y por cada unidad de cambio en X_n. Sin embargo, esta vez se trata del efecto causal de X_n sobre Y, controlando o manteniendo constantes las demás variables independientes en el modelo. Se trata del efecto linear medio ponderado de X_n sobre Y para cada una de las combinaciones de valores de las demás variables independientes en el modelo.

Imaginemos, por ejemplo, que nos interesa analizar la relación entre nivel de estudios y autoidentificación política. La variable nivel de estudios tiene nueve valores y la variable autoidentificación política tiene siete valores que oscilan desde la extrema izquierda hasta la extrema derecha. Hemos estimado este modelo utilizando LISREL:

```
Modelo 2: Regresión Simple
ni = 6 no = 376 ma = cm
la
*
ideol fcohab reldad nacban ingr educ
da fi = Reg1.lis
sd fi = Reg1.sd
se
1 6/
mo nx = 1 fi ny = 1 ne = 1 te = zero ly = id
ou se tv
```

El coeficiente de regresión cuyo valor representa el efecto del nivel de estudios sobre la autoidentificación política en nuestra muestra es igual a −0,017. Ello significa que por cada cambio en una unidad de la variable nivel de estudios se produce un cambio de −0,017 unidades en la variable de autoidentificación política. Para apreciar lo reducido de este efecto, no tenemos sino que ver cuál sería la elasticidad de la variable dependiente si pasáramos del valor 0 correspondiente a las personas que son analfabetas hacia el valor 8 que corresponde a las personas con estudios universitarios. Ésta sería igual a $-0{,}017 \cdot 9 = -0{,}153$. Es decir que no cambiaríamos ni siquiera una unidad en la variable dependiente. Ello implica que el efecto del nivel de estudios sobre la autoidentificación ideológica es mínimo. De hecho el coeficiente de correlación al cuadrado es igual a 0,001, lo cual indica que el nivel de estudios explica un 0,1 % de la varianza en la variable autoidentificación política. Resulta también que, dado que el error estándar para el efecto de la variable nivel de estudios es igual a 0,028 y que, consiguientemente, el valor *t* para este coeficiente es igual a −0,629, la relación no es estadísticamente significativa.

Visto esto podríamos plantearnos varias cuestiones: La primera de ellas es si podemos aumentar nuestra capacidad predictiva sobre la variable *Y* a base de introducir un número mayor de variables explicativas en nuestro modelo. La segunda de ellas es la de ver si el reducido efecto del nivel de estudios sobre la variable dependiente es debido a que existe una tercera variable asociada tanto a la variable dependiente como a la variable nivel de estudios que está atenuando la relación entre nivel de estudios y autoidentificación ideológica. Introduciéndola en el modelo obtendríamos un nuevo coeficiente para la variable nivel de estudios, controlando por esta tercera variable, es decir eliminando la distorsión ejercida por esta tercera variable sobre el análisis de la relación entre nivel de estudios y autoidentificación ideológica.

Imaginemos que nuestro interés principal radica en conocer el efecto que el nivel de conocimientos de las personas tiene sobre su autoidentificación ideológica. Dado que el nivel de estudios de una persona está asociado, como sabemos, a sus ingresos, y que es probable que sus ingresos afecten a su autoidentificación ideológica, es posible que el coeficiente anteriormente obtenido para la variable nivel de estudios represente no solamente el efecto del nivel de conocimientos sobre la autoidentificación ideológica, sino también el efecto del nivel de ingresos. Para conocer el efecto del nivel de estudios depurado del efecto del nivel de ingresos, debemos incluir la variable ingresos en el modelo predictivo. Una vez hecho esto, el coeficiente de regresión para la variable nivel de estudios representará el efecto de esta variable, independientemente del nivel de ingresos de las personas.

La tercera razón por la que introduciríamos una tercera variable en el modelo es la de saber qué variable tiene mayor impacto sobre la variable dependiente de nuestro modelo o, en otras palabras, qué variable tiene mayor capacidad predictiva sobre la variable dependiente de autoidentificación ideológica. Por ejemplo, nuestro propósito al introducir la variable ingresos en el modelo puede ser a la vez la de eliminar el efecto distorsionante que ésta tiene sobre el efecto de la variable nivel de estudios y el de medir el efecto particular de la variable ingresos una vez eliminado el efecto distorsionante de la variable nivel de estudios.

El programa LISREL correspondiente a este nuevo modelo, que incluye como variables independientes a la variable nivel de estudios y a la variable ingresos es el siguiente:

```
Modelo 3: Regresión Múltiple
ni = 6 no = 376 ma = cm
la
*
```

```
ideol fcohab reldad nacban ingr educ
da fi = Reg1.lis
sd fi = Reg1.sd
se
1 5 6/
mo nx = 2 fi ny = 1 ne = 1 te = zero ly = id
ou se tv
```

Corresponde al siguiente modelo:

$$\text{Ideología} = GA_1\text{Estudios} + GA_2\text{Ingresos} + ZE$$

El modelo obtenido es el siguiente:

$$\text{Ideología} = -0{,}034 \cdot \text{Estudios} + 0{,}035 \cdot \text{Ingresos} + ZE$$

Como vemos, el coeficiente para la variable nivel de estudios ha pasado de $-0{,}017$ a $-0{,}034$ cuando introducimos la variable ingresos en el modelo. Podemos decir que la variable ingresos estaba atenuando la relación existente entre nivel de estudios y autoidentificación ideológica. En cualquier caso, este efecto sigue siendo muy pequeño. Pasar del valor correspondiente a las personas analfabetas al valor correspondiente a las personas con estudios universitarios implica un cambio en la variable dependiente igual a $9 \cdot (-0{,}034) = -0{,}306$, es decir menos de una unidad. Su efecto sigue siendo insignificante. El efecto de la variable ingresos es también muy pequeño: El cambio que se produciría en la variable dependiente al pasar de ningún ingreso mensual a ganar más de 450.000 pesetas mensuales, sería de $12 \cdot 0{,}035 = 0{,}42$. (Hay doce categorías en la variable ingresos.)

En total, este modelo explica únicamente un 0,4 % de la varianza de la variable autoidentificación política, es decir prácticamente nada. Es necesario aclarar que en los modelos de regresión múltiple, el coeficiente de correlación múltiple, R al cuadrado, corresponde a la bondad del modelo en su conjunto, en lugar de representar el coeficiente de correlación de una variable en particular con la variable dependiente. Una de las características de este índice es el de que su valor aumenta simplemente con introducir más variables en el modelo. Parace lógico, por tanto, intentar comprobar si el aumento en la capacidad predictiva de nuestro modelo, que aunque ínfima, pasa de un 0,1 % a un 0,4 % de la varianza, es debido simplemente a la introducción de una variable adicional. Un test estadístico apropiado es el siguiente:

$(n - k - 1/p) \cdot (R^2 - R_0^2 / 1 - R^2)$, donde n es el número de casos, k es el número de variables en el modelo, p es el número de variables en el modelo reducido, R^2 es el coeficiente de correlación múltiple para nuestro modelo más complejo, y R_0^2 es el coeficiente de correlación múltiple de nuestro modelo más simple. En este ejemplo obtendríamos ((376 − 2 − 1)/1) · ((0,4 − 0,1/1) − 0,4) = 186,5. Resulta que los valores de este test tienen una distribución muestral particular, que denominamos *F*. Para saber si la diferencia entre el R^2 de los dos modelos es estadísticamente significativa no tenemos más que escoger un nivel de significación, generalmente igual a 0,05, y consultar la tabla correspondiente a la distribución de *F* para determinar el valor crítico correspondiente a ese nivel de significación. Un nivel de significación del 0,05 implica que la probabilidad de rechazar la hipótesis de que los dos modelos explican un porcentaje de varianza idéntico en la población, cuando en realidad tal hipótesis es cierta, es igual a un 5 %. El valor crítico de *F* para tal nivel de significación debe tener en cuenta también los grados de libertad del test, que en este caso son iguales a p y $n - k - 1$. El valor crítico en este caso es alrededor de 6,72, de modo que, dado que 186,5 es mucho mayor, podemos descartar la hipótesis nula y admitir que la introducción de la variable ingresos en el modelo mejora de forma estadísticamente significativa la capacidad explicativa del modelo.

Parece que el efecto de la variable ingresos es ligeramente superior al efecto de la variable nivel de estudios; sin embargo conviene primeramente determinar si tales efectos son estadísticamente significativos. Ésta es una norma que uno debería seguir a la hora de juzgar los modelos de regresión múltiple. En primer lugar, se determina si los efectos son estadísticamente significativos y luego, si lo son, se procede a determinar si su magnitud es suficientemente grande como para tomarlos en cuenta. En este caso, volvemos a utilizar la *t* de STUDENT para determinar si las dos variables tienen efectos significativos sobre la autoidentificacion ideológica de los entrevistados. El valor de *t* es igual a 1,076 para la variable nivel de estudios mientras que es igual a −1,073 para la variable ingresos. Por tanto, ninguno de los dos valores es estadísticamente significativo. No podemos rechazar la hipótesis de que ni el nivel de estudios ni los ingresos de los individuos tienen un efecto causal sobre el autoposicionamiento ideológico de los entrevistados. Por supuesto, estamos hablando de efectos lineares. Es posible que tengan efectos no lineares, y éstos podrían ser estimados. Sin embargo, a efectos de este manual, no continuaremos por esa vía.

Path-análisis

En el subcapítulo anterior hemos analizado un modelo en el que tanto el nivel de estudios como los ingresos determinan el autoposicionamiento ideológico de las personas. Ahora bien, al hacer esto excluimos de nuestro análisis información sobre la relación entre nivel de estudios e ingresos. Parece lógico pensar, sin embargo, que el nivel de estudios determina el nivel de ingresos de las personas. Ello nos llevaría a tratar de estimar un modelo más complicado para el análisis de las relaciones estructurales entre las tres variables que hemos introducido en nuestro análisis hasta el momento. Este nuevo modelo especificaría una relación causal directa entre nivel de estudios e ingresos, una relación causal directa entre nivel de estudios e ideología, una relación causal directa entre ingresos e ideología, y finalmente una relación causal indirecta entre nivel de estudios e ideología a través de la variable ingresos. Esta última relación existe en la medida en que cambios en el nivel de estudios se traducen en cambios en el nivel de ingresos, que a su vez se traducen en cambios en el autoposicionamiento ideológico de las personas. Este modelo puede representarse como sigue:

$$\text{Ingresos} = GA_1 \cdot \text{Estudios} + ZE_1$$

$$\text{Ideología} = GA_2 \cdot \text{Estudios} + BE_1\text{Ingresos} + ZE_2$$

(Nótese que en este modelo la relación entre ingresos y autoposicionamiento ideológico está representada por el coeficiente Beta en lugar de un coeficiente Gamma. La razón de esto radica en que ingresos es una variable endógena, es decir una variable a su vez causada por otra variable.)

El efecto indirecto del nivel de estudios sobre la ideología puede entenderse si transformamos la segunda ecuación:

$$\text{Ideología} = GA_2 \cdot \text{Estudios} + BE_1 \cdot \text{Ingresos} + ZE_2 \Leftrightarrow$$

$$\text{Ideología} = GA_2 \cdot \text{Estudios} + BE_1 \cdot (GA_1 \cdot \text{Estudios} + ZE_1) + ZE_2$$

$$\Leftrightarrow$$

$$\text{Ideología} = GA_2 \cdot \text{Estudios} + BE_1GA_1 \cdot \text{Estudios} + (BE_1ZE_1 + ZE_2)$$

$$\Leftrightarrow$$

$$\text{Ideología} = \text{Efecto Directo} + \text{Efecto Indirecto} + \text{Error}$$

El coeficiente GA_2 representa el efecto indirecto del nivel de estudios sobre la ideología, el producto BE_1GA_1 representa el efecto indirecto del nivel de estudios sobre la ideología a través de los ingresos, y $BE_1ZE_1 + ZE_2$ es un nuevo término residual.

Para estimar este modelo, el programa LISREL indicado sería el siguiente:

```
Modelo 4: Path-Análisis
ni = 6 no = 376 ma = cm
la
*
ideol fcohab reldad nacban ingr educ
da fi = Reg1.lis
sd fi = Reg1.sd
se
1 5 6/
mo nx = 1 fi ny = 2 ne = 2 te = zero ly = id be = fu, fi ps = di, fr
fr be 1 2
ou se tv
```

Para estimar este modelo hay que cumplir una serie de requisitos básicos: En primer lugar el valor esperado para las variables independientes tiene que ser igual a cero, dado que estamos utilizando variables centradas; en segundo lugar, el valor esperado para el error ZE tiene que ser también igual a cero; en tercer lugar, la correlación entre los errores ZE y las variables exógenas debe ser igual a cero; finalmente, la matriz (I-B) obtenida al restar la matriz de coeficientes de regresión entre las variables' endógenas de la matriz idéntica (es decir aquella que contiene unos en la diagonal y ceros fuera de ella) debe ser no singular, es decir, que debe tener una matriz inversa.

El output para este modelo es el siguiente:

$$\text{Ingresos} = 0{,}479 \cdot \text{Estudios} + ZE_1$$

$$\text{Ideología} = 0{,}035 \cdot \text{Ingresos} + -0{,}034 \cdot \text{Estudios} + ZE_2$$

La formulación en función de los efectos directos e indirectos del nivel de estudios sobre el autoposicionamiento ideológico es la siguiente:

$$\begin{aligned}\text{Ideología} &= -0{,}034 \cdot \text{Estudios} + (0{,}035 \cdot 0{,}479) \cdot \text{Estudios} + (0{,}035 \cdot ZE_1 + ZE_2)\\ &= -0{,}034 \cdot \text{Estudios} + (0{,}017) \cdot \text{Estudios} + (0{,}035 \cdot ZE_1 + ZE_2)\end{aligned}$$

En esta muestra se observa, por tanto, que ni lo estudios ni los ingresos parecen tener un efecto linear importante sobre el autoposicionamiento ideológico. Independientemente de sus ingresos, las personas con mayor nivel de estudios son un poco más izquierdistas que las personas con menor nivel de estudios e, independientemente de su nivel de estudios, las personas con menores ingresos tienden a ser más de izquierdas que las personas con mayores ingresos. Por otro lado se observa que el nivel de estudios de las personas tiene efectos opuestos sobre su ideología. Por un lado, al ganar más dinero, las personas con mayor nivel de estudios tienden a ser más conservadoras, mientras que, por otro lado, un mayor nivel de estudios implica mayor izquierdismo. Este último efecto tiende a ser mayor que el anterior, pero, en cualquier caso, al ser de signo diferente, el efecto total del nivel de estudios es muy cercano a cero ($-0{,}034 + 0{,}017 = -0{,}017$).

El efecto de cualquiera de estas dos variables, en cualquier caso, sobre el autoposicionamiento ideológico, es muy pequeño, y no significativo estadísticamente. Únicamente el efecto de la educación sobre los ingresos es sustancial y estadísticamente significativo.

Nótese que los coeficientes que hemos obtenido al estimar este modelo son los mismos que se obtuvieron al estimar los modelos de regresión simple y regresión múltiple arriba expuestos. La estimación del modelo de regresión simple para medir el efecto del nivel de estudios sobre el autoposicionamiento ideológico nos proporciona el efecto total de la variable independiente sobre la variable dependiente. Este efecto incluye tanto el efecto directo como aquellos efectos indirectos que pueda ejercer el nivel de estudios. Por otro lado, el análisis de regresión múltiple nos proporciona los coeficientes para el efecto directo del nivel de estudios, no mediado por la variable ingresos. La diferencia entre el efecto directo obtenido al estimar el modelo de regresión simple y el efecto indirecto obtenido al estimar el modelo de regresión múltiple, es el efecto indirecto del nivel de estudios, a través de la variable ingresos.

$$\text{Ideología} = GA_1 \cdot \text{Estudios} + GA_2BE_1 \cdot \text{Estudios} + \ldots GA_nBE_{n-1} \cdot \text{Estudios} + ZE$$

Los coeficientes GA_2BE_1 hasta GA_nBE_{n-1} son los efectos indirectos de la variable estudios a través de otras variables. Esta ecuación puede transformarse como sigue:

$$\text{Ideología} = (GA_1 + GA_2BE_1 + \ldots GA_nBE_{n-1}) \cdot \text{Estudios} + ZE$$

De este modo vemos cómo el coeficiente de estudios en un modelo de regresión simple incluye el efecto directo y todos los efectos indirectos de la variable independiente. Por otro lado el coeficiente de regresión múltiple para la variable estudios representa el efecto no mediado por la variable ingresos. Aunque se le llama efecto directo, hay que resaltar que éste incluye todos los efectos indirectos a través de variables intervinientes no incluidas en el modelo. En cualquier caso, lo que este ejemplo demuestra es que sin necesidad de estimar un path análisis, uno puede obtener una descomposición del efecto total de una variable independiente sobre una variable dependiente en sus componentes directos e indirectos, simplemente estimando un modelo de regresión simple y un modelo de regresión múltiple que además de la variable independiente incluya aquellas variables intervinientes entre ésta y la variable dependiente. El coeficiente de regresión múltiple para la variable de interés constituye el efecto directo, mientras que la diferencia entre el coeficiente de regresión simple y el coeficiente de regresión múltiple para esta variable constituye el efecto indirecto global.

Además de tomar en consideración los coeficientes de regresión que ligan a las distintas variables entre sí, existen otras medidas de bondad de ajuste para estos modelos. Una de ellas es el coeficiente de correlación múltiple al cuadrado, que nos indica la proporción de varianza en cada una de las variables endógenas del modelo explicada por éste. En nuestro ejemplo, estos coeficientes son iguales a 0,236 para la variable ingresos y 0,004 para la variable autoposicionamiento ideológico. Esto quiere decir que nuestro modelo predice un 24 % de la varianza en ingresos, y un 0,4 % de la varianza en el autoposicionamiento ideológico. Tal como otros estudios han puesto de relieve, por lo tanto, variables estructurales tan importantes como pueden ser los ingresos y el nivel de estudios son malos predictores del autoposicionamiento ideológico de los entrevistados. Ello podría deberse a que son variables complejas cuyos múltiples efectos sobre la ideología tienden a ser opuestos, de forma que su efecto total tiende a ser ínfimo.

La otra medida de bondad de ajuste, ésta referida al modelo en su conjunto, es el coeficiente de determinación. Esta medida, cuya magnitud varía también entre cero y uno, mide el efecto conjunto de las variables exógenas sobre las variables endógenas, es decir la proporción de la varianza de las variables endógenas explicada en conjunto por las variables exógenas. En nuestro ejemplo, el coeficiente de determinación es igual al 0,238. Nótese que este coeficiente, al ser global, puede, en ocasiones, representar de modo inadecuado la bondad de ajuste del modelo, al ocultar partes del modelo que no están tan bien especificadas. En nuestro caso, 0,238 es el resultado de ponderar los dos coeficientes de correlación múltiple al cuadrado de los que hablamos antes, 0,236 y 0,004. El coeficiente de determinación 0,238 por sí solo ocultaría el hecho de que la variación en la variable autoposicionamiento ideológico no está bien explicada por nuestro modelo.

Por tanto, a la hora de evaluar la bondad de un determinado modelo de ecuaciones estructurales conviene tomar en cuenta no una sino varias medidas de bondad de ajuste, empezando por los coeficientes de regresión múltiple y acabando por el coeficiente de determinación.

4
Modelos de ecuaciones estructurales

Formulación

Tanto la técnica de la regresión simple, como la de la regresión múltiple, como el path análisis son subcategorías de lo que se han denominado modelos de ecuaciones estructurales con variables observadas. Estos modelos analizan las relaciones causales y no causales entre variables que se supone miden perfectamente los conceptos bajo estudio. Es decir, la noción de error de medición queda excluida de este tipo de análisis. La hipótesis de partida de todos estos modelos, así como la de cualquier otro modelo basado en ecuaciones estructurales, es que nuestro modelo reproduce exactamente la matriz de varianzas y covarianzas que estamos estudiando. Ésta es una noción sobre la que merece la pena detenerse.

La sociedad puede concebirse como un conjunto interrelacionado de variables y la labor de los investigadores consiste en, guiados por la teoría, detectar aquellas interrelaciones realmente importantes. Supongamos que un investigador determinado recoge datos sobre una serie de variables para una población dada. Por ejemplo, respecto a las variables nivel de estudios, ingresos, asistencia a la iglesia y autoposicionamiento ideológico. Cada individuo varía en su nivel de estudios, sus ingresos, su frecuencia de asistencia a la iglesia y su autoposicionamiento ideológico. Podemos, por tanto, decir que cada variable tiene una varianza particular. Al mismo tiempo, todas estas variables están relacionadas entre sí con mayor o menor intensidad. Estas relaciones las denominamos covarianzas, para indicar que cada par de variables varía de forma conjunta con una intensidad determinada.

En definitiva, la estructura interna de una base de datos formada por un conjunto determinado de variables puede resumirse a través de una matriz de varianzas y covarianzas para este grupo de variables. Estos datos no nos dicen nada, sin embargo, sobre las relaciones de causalidad entre las distintas variables. Además, a medida que el número de variables aumenta, la imagen de la realidad presentada por la matriz de varianzas y covarianzas resulta en exceso complicada; en busca de parsimonía, nos interesa excluir de tal imagen aquella información que sea irrelevante, es decir aquellas asociaciones excesivamente débiles. Finalmente, la matriz de varianzas y covarianzas bivariadas no nos dice nada sobre cuál es la asociación entre dos variables determinadas una vez que eliminamos el efecto de terceras variables; es decir, no nos dice nada sobre el posible carácter espúreo de algunas de esas covarianzas. Por todas estas razones, el investigador, guiado por el conocimiento teórico, diseña modelos que intentan representar del modo más simple posible la realidad enfrascada en las variables recogidas, especificando relaciones causales y no causales.

Por ejemplo, el investigador puede, en base a las cuatro variables arriba enumeradas, especificar un modelo que represente su estructura relacional. Un modelo posible es aquél basado en las siguientes ecuaciones estructurales:

$$\text{Ingresos} = GA_1 \cdot \text{Estudios} + ZE_1$$

$$\text{AS.Iglesia} = GA_2 \cdot \text{Estudios} + ZE_2$$

$$\text{Ideología} = GA_3 \cdot \text{Estudios} + BE_1 \cdot \text{Ingresos} + BE_2 \cdot \text{AS.Iglesia} + ZE_3$$

Este modelo excluye, por lo tanto, una asociación causal cualquiera entre Ingresos y Asistencia a Misa, y asociaciones de signo contrario a las arriba citadas, como, por ejemplo, un efecto causal de la Ideología sobre el Nivel de Estudios. Si éstas fueran las únicas relaciones existentes en la población en cuestión entre estas cuatro variables y si pudiéramos cuantificar la magnitud de los distintos coeficientes, a este modelo le correspondería una determinada matriz de varianzas y covarianzas. (Mientras que a cada combinación particular de parámetros le corresponde una determinada matriz de varianzas y covarianzas, lo contrario no es cierto.) LISREL consiste precisamente en generar esa matriz de varianzas y covarianzas que corresponde a nuestro modelo y en compararla con la matriz real de varianzas y covarianzas que estamos estudiando, para determinar hasta qué punto se diferencian. Cuanto más parecidas sean, mejor es el modelo, puesto que ello quiere decir que el modelo reproduce el sistema de relaciones existente en la realidad.

Por supuesto, la situación se complica cuando utilizamos muestras. Nuestro objetivo sigue siendo el de comparar la matriz de varianzas y covarianzas que corresponde a nuestro modelo con la matriz de varianzas y covarianzas existente en la población en su totalidad. Sin embargo, nuestros datos, al ser muestrales, probablemente difieren en determinado grado de los datos referidos a la población en su conjunto. Por ejemplo, la covarianza entre el nivel de estudios de los entrevistados y sus ingresos será ligeramente diferente de la covarianza entre estas dos variables en la población total. Ello quiere decir que a la hora de comparar la matriz de varianzas y covarianzas resultante de nuestro modelo con la matriz de varianzas y covarianzas correspondientes a nuestros datos muestrales deberemos tener en cuenta que subyacente a esta última se encuentra la matriz de varianzas y covarianzas existente en la población.

Lo primero que debemos hacer, por lo tanto, para evaluar la bondad de ajuste de nuestro modelo es estimar la matriz de varianzas y covarianzas correspondientes a nuestro modelo. Ello requiere en primer lugar estimar los parámetros o coeficientes correspondientes a nuestro modelo, puesto que se puede demostrar que toda matriz de varianzas y covarianzas puede expresarse en función de las relaciones estructurales subyacentes entre las variables y expresadas a través de parámetros o coeficientes. Por ejemplo, si aplicamos las reglas para estimar varianzas y covarianzas, podemos observar que la varianza de la variable ingresos, expresada en desviaciones respecto a la media, puede expresarse en función de determinados parámetros derivados de nuestro modelo teórico:

$$\mathrm{Var(Ing)} = \mathrm{Cov(Ing,Ing)} \Leftrightarrow$$

$$\mathrm{Var(Ing)} = \mathrm{Cov}\,(GA_1 \cdot \mathrm{Est} + ZE_1, GA_1 \cdot \mathrm{Est} + ZE_1) \Leftrightarrow$$

$$\mathrm{Var(Ing)} = \mathrm{Cov}\,(GA_1 \cdot \mathrm{Est}, GA_1 \cdot \mathrm{Est}) + \mathrm{Cov}\,(GA_1 \cdot \mathrm{Est}, ZE) + \mathrm{Cov}\,(ZE_1, GA_1 \cdot \mathrm{Est}) + \mathrm{Cov}\,(ZE_1, ZE_1) \Leftrightarrow$$

$$\mathrm{Var(Ing)} = GA_1^{\,2}\,\mathrm{Cov}\,(\mathrm{Est},\mathrm{Est}) + \mathrm{Cov}\,(ZE_1, ZE_1)$$

[Debido a que uno de los supuestos básicos de todos estos modelos es que la asociación entre las variables independientes del modelo y las variables residuales no incluidas en el modelo es igual a cero.]

$$\Leftrightarrow \mathrm{Var(Ing)} = GA_1^{\,2}PH + PS$$

[GA es el coeficiente para la relación causal entre el nivel de estudios y los ingresos; PH es la varianza de la variable estudios estimada, no la real; PS es la varianza para la variable residual ZE_1 en el modelo.]

Por tanto, cualquier matriz de varianzas y covarianzas se puede expresar en función de los parámetros que la producen. Para estimar dichos parámetros, partimos de nuestra hipótesis de partida que es la de que la matriz de varianzas y covarianzas para las variables incluidas en nuestro estudio es idéntica a la que sería generada por nuestro modelo. Dada esta hipótesis, la matriz de varianzas y covarianzas a la que da lugar nuestro conjunto de datos se puede expresar en función de los parámetros de nuestro modelo. Para ilustrarlo, debemos partir de la idea de que la matriz de varianzas y covarianzas se compone de tres submatrices, la correspondiente a las varianzas y covarianzas entre las variables endógenas, la correspondiente a las varianzas y covarianzas entre las variables exógenas, y la correspondiente a las covarianzas entre las variables endógenas y las variables exógenas. La primera submatriz puede expresarse en función de los parámetros del modelo a través de la fórmula siguiente:

$$\mathbf{(I - BE)^{-1} (GA \cdot PH \cdot GA' + PS) (I - BE)^{-1}},$$

[Las letras en negrita denotan que las siglas corresponden a matrices y no a simples coeficientes; por ejemplo BE es la matriz de relaciones estructurales entre variables endógenas, GA es la matriz de relaciones estructurales entre variables exógenas y variables endógenas.]

La segunda submatriz puede expresarse en función de los parámetros del modelo a través de la fórmula siguiente:

$$\mathbf{PH}$$

[Donde PH es la matriz de varianzas y covarianzas entre variables exógenas implicada por el modelo.]

La tercera submatriz, finalmente, puede expresarse en función de los parámetros del modelo a través de la fórmula siguiente:

$$\mathbf{PH \cdot GA' (I - BE)^{-1}},$$

En nuestro ejemplo tenemos cuatro variables. La matriz de varianzas y covarianzas para estas cuatro variables tiene 0,5(4) (4 + 1) = 10 elementos [es decir 0,5($p + q$) ($p + q + 1$) elementos; p es el número de variables exógenas y q el número de variables endógenas]: una varianza para cada una de las variables (4) y seis covarianzas entre ellas. Si nuestro modelo reproduce perfectamente la matriz de varianzas y covarianzas de nuestros datos, cada una de estas varianzas o covarianzas puede expresarse en función de los parámetros de nuestro modelo de la manera siguiente:

1. La varianza del nivel de estudios:

Se trata de una variable exógena, x. Por tanto, su parámetro correspondiente es la matriz PHI de varianzas y covarianzas entre las variables exógenas de nuestro modelo. Como sólo hay un elemento —nivel de estudios—, el elemento correspondiente es la varianza de x subyacente a nuestro modelo —en teoría las dos deberían ser iguales.

2. Las varianzas para las variables Ingresos, Asistencia a Misa y Autoposicionamiento Ideológico:

Como son variables endógenas, la fórmula correspondiente es:

$$\mathbf{(I - BE)^{-1} (GA \cdot PH \cdot GA' + PS) (I - BE)^{-1}},$$

Para la variable Ingresos ésta es igual a:

$$(1 - 0)^{-1} (GA_1{}^2 \cdot PH + PS_1) (1 - 0)^{-1} = GA_1{}^2 \cdot PH + PS_1$$

(Como ya indicamos arriba.)

Para la variable Asistencia a Misa, ésta es igual a:

$$(1 - 0)^{-1} (GA_2{}^2 \cdot PH + PS_2) (1 - 0) = GA_2{}^2 \cdot PH + PS_2$$

Para la variable Ideología, ésta es bastante más complicada y requiere el uso del cálculo matricial. En cualquier caso, ésta es igual a:

$$BE_1{}^2GA_1{}^2PH + BE_1{}^2PS_1 + BE_2{}^2GA_2{}^2PH + BE_2{}^2PS_2 + GA_3{}^2PH + PS_3 + 2BE_1BE_2GA_1GA_2PH + 2BE_1GA_1GA_3PH + 2BE_2GA_2GA_3PH$$

La covarianza entre la variable Ingresos y la variable Asistencia a Misa puede ser expresada en función de los parámetros del modelo como sigue:

$$GA_1GA_2PH$$

La covarianza entre la variable Ingresos y la variable Autoposicionamiento Ideológico puede ser expresada en función de los parámetros del modelo como sigue:

$$BE_1 (GA_2{}^2PH + PS_1) + BE_2GA_2GA_1\ PH + GA_1GA_3PH$$

La covarianza entre la variable Asistencia a Misa y Autoposicionamiento Ideológico se puede expresar como sigue:

$$BE_1GA_2GA_1\ PH + BE_2 (GA_2{}^2PH + PS_2) + GA_2GA_3PH$$

Nos restan por determinar las covarianzas entre la variable exógena, nivel de estudios y las variables endógenas nivel de ingresos, asistencia a misa y autoposicionamiento ideológico. Siguiendo la fórmula para esta tercera submatriz resulta que la covarianza entre la variable nivel de estudios y nivel de ingresos puede representarse de la siguiente manera en función de los parámetros del modelo:

$$PH \cdot GA_1$$

La covarianza entre la variable nivel de estudios y asistencia a misa puede representarse de la manera siguiente:

$$PH \cdot GA_2$$

Finalmente, la covarianza entre la variable nivel de estudios y autoposicionamiento ideológico puede representarse de la manera siguiente:

$$PH \cdot GA_1BE_1 + PH \cdot GA_2 \cdot BE_2 + PH \cdot GA_3$$

Por tanto, toda matriz de varianzas y covarianzas puede sustituirse por una matriz equivalente en base a los parámetros del modelo causal que produce tal matriz de varianzas y covarianzas. Dicho de otro modo, dado un conjunto de relaciones estructurales entre distintas variables dentro de una población, se producirá una determinada matriz de varianzas y covarianzas. Si nuestra hipótesis de que nuestro modelo es perfecto es cierta, cada una de las varianzas y covarianzas de nuestros datos tienen que ser función de los parámetros de nuestro modelo teórico. Esto es lo que hemos hecho, obteniendo diez ecuaciones estructurales. El paso siguiente para verificar si nuestro modelo es verdaderamente el que daría lugar a tal matriz de varianzas y covarianzas es el de estimar el valor numérico de cada uno de los parámetros del modelo. Ello implica introducir dos conceptos nuevos, el de la identificación de un modelo y el de la estimación del valor de los parámetros.

Identificación de un modelo

Imaginemos, por ejemplo, que la covarianza entre nivel de estudios y asistencia a misa es igual a 0,236. Dado que hemos representado esta covarianza en función de los parámetros del modelo como $PH \cdot GA_2$ nuestro objetivo es el de encontrar valores para PH y GA_2 que una vez multiplicados den como resultado 0,236. Los valores 2 y 0,118 son una de las posibilidades, pero hay otras como, por ejemplo, 4 y 0,059. Sin embargo, no todas las combinaciones de valores son posibles puesto que nuestros datos han dado lugar a diez ecuaciones estructurales distintas en las que aparecen de manera repetida los coeficientes PH y GA_2. Por ejemplo, estos dos coeficientes aparecen en la fórmula para la covarianza entre nivel de estudios y autoposicionamiento ideológico. Los valores numéricos que otorguemos a PH y GA_2 deben permitirnos resolver todas las ecuaciones en las que aparecen estos dos coeficientes. Decimos que nuestro modelo está identificado si cabe una solución única para cada uno de los parámetros. Es decir si, por ejemplo, la información de que disponemos nos permite decir que PH sólo puede tener un valor determinado para que podamos resolver las diez ecuaciones estructurales.

Determinar si un modelo está identificado o no presenta problemas, especialmente en modelos complejos. Existen, sin embargo, una serie de reglas aplicadas a los modelos que estamos decribiendo, en los cuales cada concepto está medido por un solo indicador. A continuación se citarán las reglas más simples. El lector interesado puede acudir a libros más especializados para encontrar otras estrategias posibles.

Una primera regla es la regla de la *t*. Si el número *t* de parámetros a estimar es inferior al número de varianzas y covarianzas en el modelo, habremos cumplido una condición necesaria pero no suficiente para establecer la identificación de un modelo. En nuestro caso, tenemos nueve y no diez parámetros a estimar, puesto que sabemos que PHI es igual a la varianza del nivel de estudios. Como tenemos un número de varianzas y covarianzas igual a diez, hemos cumplido una de las condiciones necesarias para la identificación de un modelo.

La segunda regla es la «Null B rule», por la cual una condición suficiente pero no necesaria para la identificación de un modelo es que éste no postule ninguna asociación entre las variables endógenas del modelo. Como en nuestro ejemplo tanto el nivel de ingresos como la asistencia a misa, variables endógenas, afectan a la tercera variable endógena, el autoposicionamiento ideológico, podemos decir que la «Null B rule» no se cumple.

La tercera regla señala que si un modelo es recursivo entonces está identificado. Se trata ésta de una condición no necesaria pero sí suficiente para la identificación de un modelo. Un modelo es recursivo si no existen relaciones recíprocas, directas o indirectas, entre las variables. Dado que nuestro ejemplo plantea un modelo recursivo, podemos decir que nuestro modelo está identificado.

Estimación

Si nuestro modelo está identificado, cada uno de los parámetros obtenidos tendrá un valor único. Si nuestro modelo es el verdadero y si nuestros datos abarcan a toda la población, estos parámetros tendrían valores que nos permitirían reproducir la matriz de varianzas y covarianzas. Si nuestro modelo es el verdadero pero nuestros datos son muestrales, entonces es posible que no sea factible encontrar valores para nuestros coeficientes que reproduzcan completamente las varianzas y covarianzas muestrales. Sin embargo, podemos encontrar valores que minimicen las diferencias entre las varianzas y covarianzas muestrales y aquéllas obtenidas a partir de nuestro modelo. Si nuestro modelo es falso, entonces no podremos encontrar valores que reproduzcan la matriz de varianzas y covarianzas. El propósito de los tests de bondad de ajuste es el de determinar hasta qué punto las diferencias obtenidas entre la matriz de varianzas y covarianzas muestral y la matriz de varianzas y covarianzas estimada a partir de nuestro modelo son debidas al azar o a que nuestro modelo es falso.

La estimación de los coeficientes se hace siempre desde el supuesto de que nuestro modelo es el bueno y que el que no podamos encontrar coeficientes que puedan reproducir la matriz de varianzas y covarianzas muestrales se debe a la muestra. Por eso, siempre que el modelo está identificado, LISREL estimará aquellos coeficientes que nos permitan reproducir lo más fehacientemente posible la matriz de varianzas y covarianzas muestrales. Son los tests de bondad de ajuste los que nos permiten establecer si la falta de identidad entre las dos matrices de varianzas y covarianzas se debe al azar o a la inadecuación del modelo.

Existen muchas medidas de bondad de ajuste sin que se pueda decir que una de ellas es la mejor. Ello es fácil de explicar. Como se ha indicado anteriormente, la bondad de ajuste depende de comparar la matriz de varianzas y covarianzas muestral con la generada por nuestro modelo. Si la diferencia es suficientemente grande, podemos excluir el azar como posible explicación para esta diferencia y, por lo tanto, rechazar nuestro modelo. La primera cuestión es cómo determinar la magnitud de esa diferencia.

Tomemos como ejemplo nuestro modelo. Éste implica la toma en consideración de una matriz formada por un total de diez varianzas y covarianzas. A cada una de ellas le corresponde una varianza o covarianza estimada a partir de las características de nuestro modelo. Lo normal es que cada una de éstas difiera ligeramente de su varianza o covarianza muestral correspondiente. Por ejemplo, la covarianza muestral entre nivel de estudios y asistencia a misa será ligeramente diferente de aquella obtenida a partir del modelo, cuya fórmula correspondiente es $PH \cdot GA_2$. ¿Cómo resumimos en un índice el conjunto de las discrepancias entre los valores muestrales y los estimados? Una posible vía sería la de computar la media de los errores. Otra posibilidad sería la de computar la raíz cuadrada de la media de la suma de desviaciones cuadradas entre valores reales y valores estimados. Hay muchas estrategias posibles. Para entender por qué utilizamos unas y no otras tenemos que dar un paso atrás, sin embargo, y volver al proceso de estimación de la matriz de varianzas y covarianzas derivadas de nuestro modelo:

Como dijimos anteriormente, a no ser que nuestro modelo sea perfecto y que estemos estudiando la población en lugar de una muestra, será imposible encontrar coeficientes tales que se reproduzca idénticamente la matriz de varianzas y covarianzas muestral. Nuestra hipótesis es la de que nuestro modelo es perfecto y que, una vez tenido en cuenta el error muestral, deberíamos obtener coeficientes para nuestros parámetros a partir de los cuales se obtendrá una matriz de varianzas y covarianzas estimada lo más parecida posible a nuestra matriz de varianzas y covarianzas muestrales. Cuando decimos parecida volvemos al problema indicado en el párrafo anterior, es decir la búsqueda de un índice que resuma las discrepancias entre las varianzas y covarianzas obtenidas y las muestrales, y que es lo que se trata de minimizar al estimar los coeficientes, debido a que nuestra hipótesis es que deberían ser iguales.

Varias son las funciones que se utilizan a la hora de minimizar las diferencias entre la matriz muestral y la obtenida. Las más conocidas son la de Maximum Likelihood, Unweighted Least Squares y Generalized Least Squares. Estas tres funciones tienen tres características: Primero, dan lugar a una cifra que se intenta minimizar. Segundo, esta cifra es igual a cero si la matriz muestral y la obtenida son idénticas. Tercero, su valor, siempre positivo, se aleja de cero cuanto más discrepan las dos matrices. Finalmente, se trata de funciones continuas, es decir funciones que pueden dar lugar a cualquier valor de cero al infinito. Cada una de estas funciones presupone ciertas condiciones y supone ciertas ventajas.

La función de Maximum Likelihood presupone que las variables x e y tienen una distribución conjunta multinormal. ¿Qué quiere decir esto? Se dice que una variable x tiene una distribución normal si tiene forma de campana. Cuando hablamos de la distribución conjunta de dos variables nos referimos al hecho de que a cada valor x le corresponde un valor y; cada unidad de análisis está formada por dos dimensiones y, por lo tanto, la distribución conjunta va a tener dos dimensiones. Si esta distribución conjunta en dos dimensiones tiene forma de campana decimos que es multinormal. Un segundo supuesto es el de que nuestro modelo es correcto.

Dados estos supuestos, los coeficientes obtenidos por este procedimiento tienen las siguientes propiedades:

En primer lugar, cuando la muestra es suficientemente grande (alrededor de 100 casos y más) proporciona coeficientes estimados carentes de sesgo; ello quiere decir que si extrajésemos un número infinito de muestras de 100 o más casos y calculásemos cada vez el valor de estos coeficientes, el valor medio de estos coeficientes sería el correspondiente a la población total. En segundo lugar, en muestras grandes, esta función produce estimadores eficientes; es decir que si una vez obtenidas todas esas muestras, además de calcular el valor medio calculamos la desviación típica de esos valores, obtendremos un valor mínimo comparado con el que se obtendría con otros métodos. Ello nos permite determinar la significación estadística de los coeficientes, es decir que nos permite determinar si dichos coeficientes son diferentes de cero debido simplemente al azar. Otra propiedad importante es que da igual en cuanto al valor de la función a minimizar el que utilicemos la matriz de correlaciones o la matriz de varianzas y covarianzas como input en nuestro análisis. Finalmente, si transformamos la escala de las variables x e y es fácil convertir los coeficientes obtenidos en base a la nueva escala en aquellos coeficientes correspondientes a la escala original. Es decir que si, por ejemplo, en vez de medir la variable ingresos en pesetas, la medimos en duros, el coeficiente que obtengamos se puede transformar fácilmente en el valor que le correspondería si nuestra variable ingresos la hubiéramos medido en pesetas. Estas dos últimas propiedades son la propiedad de escala invariante y de escala libre, respectivamente.

La función Unweighted Least Squares difiere de la anterior en que no requiere una distribución conjunta multinormal de las variables x e y. Sin embargo, produce coeficientes no eficientes. Es decir que su distribución muestral no tiene la mínima desviación típica. Además, ni tiene la propiedad de escala invariante ni la de escala libre.

Dos supuestos son necesarios para utilizar la función Generalized Least Squares. La primera es la de que la distribución muestral de las covarianzas tenga como valor esperado el valor de la población; la segunda es la de que las variables x e y tengan una distribución conjunta multinormal o sin excesiva kurtosis (es decir, una campana ni muy plana ni muy estirada hacia arriba). Dado este supuesto, sus propiedades son muy similares a las de la función Maximum Likelihood.

En la práctica, todas estas funciones producen valores para los coeficientes y, por lo tanto, matrices estimadas de varianzas y covarianzas muy similares entre sí en condiciones normales. Tradicionalmente se suele utilizar la función de Maximum Likelihood cuando todas las variables son variables de intervalo o pueden ser conceptualizadas como tales al tener al menos seis valores. Generalized Least Squares, y una de sus versiones, Weighted Least Squares es utilizada sobre todo cuando las variables utilizadas son dicotómicas o ordinales. Métodos con los cuales analizar relaciones entre variables categóricas utilizando este método no se han desarrollado todavía.

Bondad de ajuste

Estamos en condiciones de volver ahora al tema de la bondad de ajuste. Todas estas funciones producen valores que en la medida en que difieren de cero indican que nuestro modelo no es perfecto. Sin embargo, ¿cómo determinar el grado de imperfección y cómo determinar si este grado de imperfección se debe simplemente al azar dado que estamos utilizando una muestra? Algunos de los índices que vamos a sugerir contestan a la primera pregunta y otros a la segunda. Deben utilizarse de manera complementaria con los índices correspondientes a parámetros particulares (por ejemplo, la magnitud y significación estadística de un determinado coeficiente) y con los índices que indican la proporción de varianza explicada por nuestro modelo (ya sean los coeficientes de correlación múltiple o el coeficiente de determinación).

Todos estos índices requieren una condición previa y es que el modelo esté sobreidentificado. Tal como se indicó anteriormente, un modelo está identificado si para cada parámetro cabe un coeficiente único, es decir si, por ejemplo, el número y características de nuestras ecuaciones estructurales son tales que el coeficiente para GA_1 puede tener un valor determinado único al resolver el sistema de ecuaciones estructurales. Si ello no es así el modelo en su conjunto está subidentificado. Cuando la información de que disponemos es tal que permite estimar un determinado coeficiente de varias maneras distintas decimos que el modelo está sobreidentificado. Por ejemplo, si el parámetro PHI pudiera ser estimado en base a la ecuación PHI = $\mathrm{Var}(x_1)$ y a la ecuación PHI = $\mathrm{Cov}\,(x_1,y_1)$. La solución a esta situación es única pero tenemos el constreñimiento de que PHI debe ser igual tanto a la varianza de x_1 como a la covarianza entre x_1 e y_1.

Dicho esto, un test ampliamente utilizado es el de Chi-cuadrado que puede ser utilizado tanto con la función Maximum Likelihood como con la función Generalized Least Squares. El test consiste en multiplicar el valor mínimo obtenido para cualquiera de las dos funciones por $N - 1$, donde N es el número de casos. Este test, basado en el supuesto de que nuestro modelo es perfecto, nos dice la probabilidad de que la divergencia entre la matriz de varianzas y covarianzas muestrales y la implicada por nuestro modelo sea debida al azar. En otras palabras si, como dijimos, el valor de la función minimizada sería igual a cero si el modelo fuese perfecto, cualquier valor de esta función que sea diferente de cero puede indicar o bien que el modelo no es perfecto o bien que esta diferencia entra dentro de lo esperado dado que estamos trabajando con una muestra. El factor $N - 1$, multiplicado por el valor de la función minimizada tiene una distribución muestral conocida como es la distribución de Chi-Cuadrado. Es decir que aunque en la población total el valor de esta función sea igual a cero, es decir si el modelo es perfecto, cabe esperar que, si extraemos una muestra de esta población, el valor muestral de esta función sea diferente de cero simplemente por mero azar. Si extraemos muestra tras muestra y calculamos siempre el valor de la función minimizadora, unos valores aparecerán con más frecuencia que otros, de manera que una vez obtenidas una infinidad de muestras podríamos, hipotéticamente, representar gráficamente la distribución de tales valores. Dicha distribución nos diría la probabilidad de que obtengamos un determinado valor muestral para nuestra función minimizadora si el valor en la población es igual a cero. La forma de tal distribución para el valor minimizado no sigue una pauta conocida o estándar; sin embargo, la forma de la distribución para el producto del valor minimizado por $N - 1$ sí que nos es conocida: se trata de la distribución de Chi-Cuadrado. Al sernos conocida podemos calcular inmediatamente la probabilidad de que obtengamos un valor muestral determinado para la función minimizada cuando su valor para la población total es igual a cero, es decir cuando el modelo es perfecto. Si la probabilidad es suficientemente pequeña (generalmente utilizamos 5 % como probabilidad límite) podemos concluir que es muy cuestionable el que nuestra hipótesis nula, el que nuestro modelo sea perfecto, sea válida. Podemos, por tanto, concluir que nuestro modelo no es perfecto. Para calcular dicha probabilidad hay que tener en cuenta que Chi-Cuadrado no es una sola distribución sino una familia de distribuciones cuya forma exacta viene determinada por lo que se denomina como grados de libertad. Los grados de libertad para este test se calculan sustrayendo el número de parámetros a estimar, GA_1 etc., del número total de parámetros conocidos, es decir el número de varianzas y covarianzas, o $1/2\,(p + q)\,(p + q + 1)$, donde q y p son el número de variables x e y en el modelo.

Volvamos a nuestro ejemplo para la explicación de la autoidentificación ideológica. En este modelo tenemos cuatro variables y, por tanto, el número de parámetros conocidos es $1/2\,(4 \cdot 5) = 10$. Los parámetros a estimar son nueve, puesto que uno de los constreñimientos del modelo es que PHI = Var (x1). Por tanto, tenemos $10 - 9 = 1$ grados de libertad. El valor de Chi-Cuadrado para nuestro modelo y que el programa LISREL nos proporciona es igual a 0,17. Este valor se diferencia de 0 y si estuviéramos tratando con la población concluiríamos que nuestro modelo no es perfecto. Sin embargo, dado que estamos utilizando una muestra, lo que debemos saber es la probabilidad de obtener un valor muestral igual a 0,17 si el valor de Chi-Cuadrado en la población fuese igual a 0. Esta probabilidad con un grado de libertad es de 0,677, es decir mucho mayor que un 5 %. Es decir que si el valor de Chi-Cuadrado en la población fuese igual a 0 la probabilidad de que una muestra cualquiera con 486 casos proporcionase un valor de Chi-Cuadrado igual a 0,17 es de un 67,7 %. Esta probabilidad es tan elevada que no podemos rechazar la hipótesis nula de que nuestro modelo es perfecto, es decir de que hemos representado perfectamente la estructura causal entre dichas variables.

El test de Chi-Cuadrado se basa en ciertos supuestos que conviene tener en cuenta: En primer lugar, la hipótesis de que la distribución conjunta de las variables exógenas no tiene kurtósis; en segundo lugar, que estamos analizando la matriz de varianzas y covarianzas y no la matriz de correlaciones; en tercer lugar, que la muestra es suficientemente grande ya que muestras pequeñas dan lugar a una tendencia al rechazo de la hipótesis nula; y en cuarto lugar, que la hipótesis de que nuestro modelo es perfecto es válida. Nótese, por otro lado, que no rechazar la hipótesis nula no implica poder probarla. Se podrían estimar otros modelos con los mismos datos que no permitiesen rechazar la hipótesis nula. En cambio, otro problema es que rechazar la hipótesis nula significa que nuestro modelo no es perfecto. Pero ¿qué modelo sociológico pretende ser perfecto? Este último punto sugiere que no lancemos todo por la borda si la hipótesis nula es rechazada.

A partir de que el test de Chi-Cuadrado plantea algunos problemas, diversos autores han sugerido la utilización de medidas adicionales de la bondad de ajuste de un modelo. El programa LISREL, además del Coeficiente de determinación y del test de Chi-Cuadrado, proporciona dos medidas adicionales, el GFI (Goodness of Fit Index) y el AGFI (Adjusted Goodness of Fit Index). Mientras que el Coeficiente de Determinación estima la proporción de la varianza de las variables endógenas explicada por el modelo, el GFI estima la cantidad relativa de varianzas y covarianzas entre todas las variables explicadas por el modelo. Alcanza un valor máximo de uno y, aunque suele ser superior a cero, puede, en ocasiones, ser negativo. El AGFI corrige por el número de grados de libertad. Dado el procedimiento para estimar los grados de libertad, sabemos que cuanto más complejo es un modelo, es decir cuantos más parámetros hay que estimar, menor es el número de grados de libertad. Tal como ocurre con el coeficiente de correlación múltiple, cuanto más complejo es un modelo, más probable es que aumente nuestra capacidad de predicción aun a costa de una pérdida de parsimonia. Corregir por el número de grados de libertad es un procedimiento por el cual se trata de premiar modelos que tienen mayor parsimonia. Esto es lo que hace el AGFI, de manera que podamos eliminar el efecto de la menor o mayor complejidad de un determinado modelo. El único problema de estas dos medidas es que algunas simulaciones han sugerido que el valor esperado muestral para estas medidas aumenta con el número de casos. En nuestro ejemplo el valor para GFI es igual a 1,000 y el valor para el AGFI es igual a 0,998. Tanto un valor como otro sugieren una bondad de ajuste excelente.

Existen, por supuesto, otros índices para medir la bondad de ajuste de un determinado modelo. A continuación presentaremos cuatro de ellos, Delta 1, Delta 2, Rho 1 y Rho 2. Todas estas medidas comparan nuestro modelo con un modelo base que, tradicionalmente, suele ser aquel que estipula una falta absoluta de asociación entre las variables del modelo; se trata, por tanto, de comparar nuestro modelo con el peor modelo posible.

El índice de Delta 1 consiste en la fórmula siguiente:

[Chi-Cuadrado (base) − Chi-Cuadrado (modelo)] / Chi-Cuadrado (base)

Si nuestro modelo es tan malo como el peor modelo posible, el valor de Delta 1 será igual a cero, y si nuestro modelo es perfecto, de modo que el Chi-Cuadrado para el modelo defendido es igual a cero, el valor de Delta 1 será igual a uno. Esta medida, sin embargo, presenta varios problemas. En primer lugar, su valor aumenta necesariamente con la complejidad del modelo, es decir a medida que disminuye el número de grados de libertad. Si queremos primar la parsimonía de nuestros modelos deberemos corregir, sobre todo al comparar modelos, por el número de grados de libertad. Otro problema con este índice es que la media de su distribución muestral tiende a aumentar a medida que aumenta el tamaño de la muestra. En vista de ello, BOLLEN (1988) ha propuesta Delta 2, que corrige por el número de grados de libertad y que reduce la dependencia de Delta 1 respecto al tamaño muestral. La fórmula para Delta 2 es la siguiente:

[Chi-Cuadrado (base) – Chi-Cuadrado (modelo)] / [Chi-Cuadrado (base) – gl (modelo)]

(donde «gl» significa «grados de libertad»).

Delta 2 tiende hacia uno, no varía obligatoriamente entre cero y uno, y su valor puede ser mayor que uno en determinadas ocasiones.

Rho 1 es casi idéntico a Delta 1 excepto que tanto el Chi-Cuadrado para el modelo defendido y el modelo base son divididos por sus grados de libertad respectivos. De esta manera se premia a los modelos más parsimoniosos o simples. Sin embargo, tal como ocurría con Delta 1, la media muestral de este índice tiende a aumentar con el tamaño de la muestra.

Finalmente, Rho 2 es un índice que parte del supuesto de que el mejor modelo no es aquel para el cual Chi-Cuadrado es igual a cero, sino aquel para el cual el valor esperado de Chi-Cuadrado dividido por sus grados de libertad es igual a uno. De ahí que se mantenga el numerador de Rho 1 pero que el denominador de Rho 2 sea igual a Chi-Cuadrado (base) – 1. Es decir que comparamos la diferencia entre nuestro modelo y el peor modelo posible con la diferencia entre el mejor modelo posible y el peor modelo posible. Si nuestro modelo es tan bueno como el mejor modelo posible, Rho 2 debería ser igual a uno. Rho 2 está menos influenciado que Rho 1 por el tamaño de la muestra y los valores de los dos índices convergen a medida que aumenta el tamaño de la muestra.

En líneas generales, parece indicado fiarse más de Delta 2 y de Rho 2 cuando la muestra es pequeña, mientras que cualquiera de ellos es bastante adecuado cuando la muestra es grande. La experiencia sugiere que un modelo adecuado debería proporcionar valores por encima de 0,9, dado que normalmente nuestro modelo base es el peor modelo posible.

En el ejemplo que estamos manejando, los valores para estos índices son los siguientes:

Delta 1: 0,999; Delta 2: 1,004; Rho 1: 0,995; Rho 2: 1,028

Está claro que, en este ejemplo, todas las medidas de bondad de ajuste indican un ajuste excelente, ya sea el coeficiente de determinación, el Chi-Cuadrado o cualquiera de los otros índices sugeridos. Sin embargo, medidas más puntuales, referidas a los distintos componentes del modelo, sugieren que el modelo es bastante pobre a la hora de explicar nuestra variable de mayor interés como es el autoposicionamiento ideológico. Nuestra decisión final depende, pues, de aquello que primemos a la hora de estimar un modelo. En el ejemplo presente, si el interés radica en explicar el autoposionamiento ideológico, uno tendería a modificar el modelo ya que está claro que ni los ingresos ni el nivel de estudios ayudan a explicarlo. Únicamente tiene interés la relación entre el grado de asistencia a misa y el autoposicionamiento ideológico.

Es importante, en cualquier caso, que toda modificación del modelo se haga en base a una muestra independiente si uno quiere evitar aprovecharse del azar, que haría que nuestra probabilidad de estimar un modelo que parezca adecuado desde el punto de vista de la significación estadística aumente con el número de modelos que estimemos. Una precaución en estos casos consiste en dividir la muestra en dos y realizar todo el análisis exploratorio con una mitad de la muestra hasta encontrar el modelo adecuado. Una vez encontrado, y satisfecho el investigador por la plausibilidad teórica del modelo, es cuando uno lo somete a verificación con la otra mitad de la muestra.

5
Análisis factorial confirmatorio

Error de medida y problemas

En este capítulo se profundiza en el tema del error de medida y sus consecuencias. Éste es un problema raramente considerado por aquellos que se dedican al análisis cuantitativo. Tendemos a considerar que nuestros indicadores son medidas sin error de los conceptos que en realidad nos interesan. Sin embargo, esto no es así y la consecuencia principal de no tener en cuenta el error de medición es que los resultados que obtenemos al estudiar relaciones entre variables conceptuales suelen carecer de significado.

Conviene por tanto decir algo sobre las consecuencias de este error de medida y proponer métodos que permitan estimar tanto la validez como la fiabilidad de nuestras medidas.

El error de medición es la diferencia entre el valor real de una variable determinada y el valor del indicador que estamos utilizando para medirla. Este error puede ser sistemático o aleatorio. Por ejemplo, imaginemos que nos preguntan por nuestro grado de felicidad y que para ello nos presentan una escala de felicidad que va del 0 al 10. Supongamos que objetivamente nuestro grado de felicidad se sitúa en el siete; es muy difícil, sin embargo, que cuando nos hagan la pregunta sepamos verbalizar nuestro grado de felicidad en esa escala, puesto que no estamos acostumbrados a contabilizarla de ese modo. Por tanto, y abstrayendo de factores coyunturales que afecten nuestro estado de humor en el momento que nos hacen la pregunta, la respuesta que demos va a fluctuar alrededor del valor real. Cuanto más desconexión exista entre nuestra manera de visualizar un determinado concepto y el instrumento de medición utilizado, mayor será la fluctuación de nuestras respuestas, es decir la varianza del error de medida. De este modo, si llamamos KSI a la variable exógena real y D al error de medición de una variable exógena determinada, si llamamos ETA a la variable endógena real y E al error de medición de una variable endógena determinada, nos encontramos con que:

x = KSI + D e y = ETA + E, siempre que KSI y x estén medidas en la misma escala y que ETA e y estén medidas en la misma escala también. Es decir que el valor para un indicador es igual al valor real más el error de medición.

Las consecuencias del error de medición sobre la estimación de relaciones entre variables KSI y ETA pueden ser considerables y no necesariamente predecibles. En el caso de la regresión simple, el error de medida en la variable independiente conduce a la subestimación del coeficiente de regresión. Cuanto mayor sea el error mayor será la subestimación. El error de medición tanto en la variable independiente como en la variable dependiente también conlleva la atenuación del coeficiente de correlación múltiple al cuadrado entre dos variables. Una vez que pasamos el umbral de la regresión simple, sin embargo, es difícil generalizar sobre las consecuencias del error de medición. Dependiendo del caso, los coeficientes obtenidos estarán infraestimados, sobreestimados o serán iguales a los reales. Por ello es muy importante tomarlo en cuenta a la hora de estimar modelos de ecuaciones estructurales y ésta es una de las razones principales por las que tanto el análisis factorial como la técnica llamada LISREL se desarrollaron.

Análisis factorial confirmatorio

El análisis factorial confirmatorio parte de la aceptación de que prácticamente todos los indicadores que utilizamos incluyen error de medición, y de que ningún indicador es una medida totalmente válida y fiable del concepto que se trata de medir. A través del análisis factorial confirmatorio tratamos de estimar la relación entre los conceptos en los que estamos interesados y los indicadores utilizados para medirlos. Este tipo de análisis se diferencia del análisis factorial tradicional o exploratorio en varios puntos. El principal es que en el análisis factorial exploratorio la relación entre distintos conceptos y los varios indicadores escogidos por el investigador no es determinada por el investigador antes del análisis. Éste incluye ciertos indicadores en el programa y deja que el paquete estadístico utilizado determine la relación entre éstos y distintos factores en número más o menos limitado a los que el investigador otorga un nombre una vez observadas las pautas de relación. En el análisis factorial confirmatorio el investigador determina con antelación qué indicadores están relacionados con cada concepto, también llamado variable latente o factor, y los nombres de cada factor. Lo único que el investigador no suele especificar es la magnitud de tal relación, aunque le cabe la posibilidad de hacerlo y dejar que LISREL le indique lo acertado de la elección de tales magnitudes. Un ejemplo de análisis factorial confirmatorio consistiría en elaborar una definición para el concepto de izquierdismo político, pensar en las dimensiones de este concepto y escoger una serie de indicadores que lo representen. Supongamos que definimos izquierdismo político como la defensa de actitudes políticas favorables a la eliminación de controles institucionales sobre la moral y al control estatal sobre la propiedad. Por supuesto, esta definición sería discutible y un trabajo definitorio más intenso debería preceder a la elección de indicadores. La definición propuesta contiene dos dimensiones:

— Actitud favorable a la eliminación de controles sociales sobre la moral.
— Actitud favorable al control estatal sobre la propiedad.

Imaginemos que disponemos de un cuestionario con los siguientes ítems:

— Autoposicionamiento ideológico (x_1).
— Actitud ante la nacionalización de la Banca (x_2).
— Actitud ante la cohabitación (x_3).
— Frecuencia de asistencia a Misa (x_4).

Podemos considerar a las dos últimas variables como indicadores de la primera dimensión, mientras que la segunda sería un indicador de la primera dimensión. La primera variable no mide ninguna de las dos dimensiones pero representa el izquierdismo subjetivo de cada individuo.

A través del análisis factorial confirmatorio intentaremos cumplir varios objetivos. En primer lugar, intentaremos determinar si dichos indicadores miden de verdad distintas dimensiones de un mismo concepto. Es decir, intentaremos determinar la validez de nuestros indicadores como medidas del concepto considerado. En segundo lugar, intentaremos determinar la fiabilidad de nuestros indicadores. Finalmente, intentaremos estimar la magnitud de la relación entre nuestros indicadores y el concepto de izquierdismo político. Hay varias cuestiones que no podremos resolver. En primer lugar, no podremos probar la validez de nuestros indicadores. Lo más que podremos deducir de nuestro análisis es si debemos rechazar la hipótesis de que estos indicadores están midiendo un mismo concepto. El que no podamos rechazarla no quiere decir que nuestra hipótesis quede probada. En segundo lugar, no podremos probar que, aun midiendo un mismo concepto, estas variables midan el concepto que nos interesa. Para esta tarea lo que tenemos que hacer es determinar si los indicadores miden nuestro concepto previamente definido en todas sus dimensiones. Ésta es una cuestión ciertamente subjetiva. Por otra parte, no podremos probar que nuestra definición es la correcta. No hay definiciones correctas; hay definiciones con más tradición y con menos tradición; hay definiciones más útiles y definiciones menos útiles.

En el análisis exploratorio, sin embargo, dejaríamos que el programa de turno determinase el número de factores subyacentes a nuestros indicadores y la relación de cada indicador con cada factor. No sólo pasamos a depender de la muestra particular que estamos analizando, sino que nos ahorramos un esfuerzo definitorio necesario y terminamos estimando modelos que están de hecho subidentificados porque tenemos menos información que parámetros a estimar. Estos comentarios, por supuesto, aluden a tipos ideales de análisis factorial confirmatorio y exploratorio. En la práctica las diferencias entre las dos estrategias pueden ser mucho menores. Por otro lado, el análisis factorial exploratorio puede ser útil a la hora de simplificar una matriz de datos excesivamente grande, de manera que el investigador pueda entonces pasar a la estimación de modelos mucho más informados teóricamente.

En el ejemplo que hemos descrito, el análisis factorial confirmatorio consiste en resolver el siguiente sistema de ecuaciones:

$$x_1 = \mathrm{L}x_1 \cdot \mathrm{KSI} + \mathrm{D}_1$$
$$x_2 = \mathrm{L}x_2 \cdot \mathrm{KSI} + \mathrm{D}_2$$
$$x_3 = \mathrm{L}x_3 \cdot \mathrm{KSI} + \mathrm{D}_3$$
$$x_4 = \mathrm{L}x_4 \cdot \mathrm{KSI} + \mathrm{D}_4$$

Tenemos cuatro indicadores, x_1 a x_4, que están relacionados con un mismo factor KSI. La relación estructural entre factores latentes y sus indicadores se expresa a través del símbolo Lx si se trata de una variable exógena y del símbolo Ly si se trata de una variable endógena. Lx expresa el número de unidades de medición que cambia x por cada cambio de una unidad en el factor latente. Sin embargo, cada una de estas medidas contiene un grado determinado de error, D para variables exógenas y E para variables endógenas, que hay que estimar.

Otros análisis factoriales confirmatorios son posibles. Por ejemplo, se podría formular la hipótesis de que el valor de x_1 es función a la vez del izquierdismo político y de otros conceptos o factores. También se podría pensar que algunos de los errores de medición, D, están correlacionados entre sí. Así, por ejemplo, si tanto x_1 y x_2 tienen valores que se han obtenido al formular un determinado tipo de pregunta —Díganos si está muy de acuerdo, de acuerdo, en desacuerdo, muy en desacuerdo...— mientras que x_1 y x_4 son respuestas que se han obtenido con otro tipo de preguntas, podemos formular como hipótesis que las respuestas a x_1 y x_2 son función a la vez del tema sobre el que se pregunta a los entrevistados como de la estructura de la pregunta misma. Ello podemos representarlo especificando una correlación entre los errores de medición para las variables x_1 y x_2.

Los únicos requerimientos adicionales en esta clase de modelos es que, si las variables están centradas, el valor esperado para los factores latentes es igual a cero, así como el de los errores. Por otra parte, es necesario que el error de medición de las variables endógenas no esté asociado en absoluto ni con los factores latentes ni con los errores de medición de las variables exógenas. Finalmente, el error de medición de las variables exógenas no debe estar asociado en absoluto ni con los factores latentes ni con los errores de medición de las variables endógenas.

La estimación de un análisis factorial confirmatorio sigue los mismos pasos enunciados para el análisis de relaciones estructurales entre variables perfectamente medidas. La hipótesis nula es que nuestro modelo reproduce perfectamente la información contenida en nuestra matriz muestral de varianzas y covarianzas. Si ello es así, cada una de las varianzas y covarianzas muestrales puede expresarse en función de los parámetros a estimar por el modelo. Como siempre, asumiremos, para simplificar la presentación analítica, que las variables x están centradas, es decir expresadas en desviaciones respecto a sus medias respectivas. Si esto es así, la matriz de varianzas y covarianzas entre las variables x puede expresarse como sigue, en función de los parámetros del modelo:

$$\mathbf{LX} \cdot \mathbf{PHI} \cdot \mathbf{LX}' + \mathbf{TD}$$

Recordemos que aquí los símbolos representan matrices. **LX** (en negrita) representa la matriz de coeficientes de asociación entre cada indicador x (líneas) y cada uno de los factores especificados por el modelo (columnas). En nuestro ejemplo, como sólo tenemos un factor latente, LX tiene cuatro líneas y una columna. PHI es la matriz de varianzas y covarianzas para los factores latentes exógenos del modelo. Como en nuestro ejemplo sólo hay un factor latente, esta matriz sólo incluye un elemento, correspondiente a la varianza de este factor latente. Finalmente, TD es la matriz de varianzas y covarianzas entre los errores de medición de los distintos indicadores x. Como en nuestro ejemplo hay cuatro indicadores, cada uno con su error de medición, esta matriz tiene cuatro líneas y cuatro columnas, que incluyen las varianzas para los errores de medición en la diagonal y las covarianzas entre los errores de medición fuera de la diagonal. Dada esta relación entre la matriz de varianzas y covarianzas del modelo y los distintos parámetros a estimar, la solución del modelo incluido en nuestro ejemplo requiere resolver las siguientes ecuaciones:

$$\begin{aligned}
\mathrm{Var}(x_1) &= (\mathrm{L}x_1)^2 \cdot \mathrm{PHI} + \mathrm{TD}_1 \\
\mathrm{Var}(x_2) &= (\mathrm{L}x_2)^2 \cdot \mathrm{PHI} + \mathrm{TD}_2 \\
\mathrm{Var}(x_3) &= (\mathrm{L}x_3)^2 \cdot \mathrm{PHI} + \mathrm{TD}_3 \\
\mathrm{Var}(x_4) &= (\mathrm{L}x_4)^2 \cdot \mathrm{PHI} + \mathrm{TD}_4 \\
\mathrm{Cov}(x_1, x_2) &= \mathrm{L}x_1 \cdot \mathrm{L}x_2 \cdot \mathrm{PHI} \\
\mathrm{Cov}(x_1, x_3) &= \mathrm{L}x_1 \cdot \mathrm{L}x_3 \cdot \mathrm{PHI} \\
\mathrm{Cov}(x_1, x_4) &= \mathrm{L}x_1 \cdot \mathrm{L}x_4 \cdot \mathrm{PHI} \\
\mathrm{Cov}(x_2, x_3) &= \mathrm{L}x_2 \cdot \mathrm{L}x_3 \cdot \mathrm{PHI} \\
\mathrm{Cov}(x_2, x_4) &= \mathrm{L}x_2 \cdot \mathrm{L}x_4 \cdot \mathrm{PHI} \\
\mathrm{Cov}(x_3, x_4) &= \mathrm{L}x_3 \cdot \mathrm{L}x_4 \cdot \mathrm{PHI}
\end{aligned}$$

Como se puede apreciar, tenemos diez piezas de información —las varianzas y covarianzas— y nueve parámetros que estimar. Para poder estimar el modelo, es necesario, sin embargo, otorgar una escala al factor latente. Efectivamente, ¿cómo podemos hablar del número de unidades de cambio en x por cambio de una unidad en KSI cuando no sabemos la escala de KSI, es decir cuando no sabemos qué quiere decir una unidad de cambio en KSI? Para resolver un análisis factorial confirmatorio, por lo tanto, tenemos que decidir previamente cuál es la escala de KSI. Generalmente se suelen adoptar dos estrategias: La primera consiste en otorgarle arbitrariamente la escala de cualquiera de los indicadores del modelo. Por ejemplo, podemos, en nuestro ejemplo, darle la escala de x_1. Si esto es así, un cambio de una unidad en KSI se traduce en un cambio de una unidad en x_1 más el error de medida:

$$x_1 = 1 \cdot \mathrm{KSI} + \mathrm{D}_1$$

La segunda estrategia que se suele adoptar es la de dar a KSI una media igual a cero y una varianza igual a uno, es decir, suponer que el factor latente está estandarizado. Si esto es así, Lx es el cambio que se produce en x por cada cambio de una desviación típica en KSI. Dado este necesario constreñimiento, el número de parámetros a estimar queda reducido a ocho cuando establecemos la escala de KSI de cualquiera de las maneras que hemos sugerido.

Para resolver este sistema de ecuaciones que nos proporcionará un valor para cada uno de los parámetros del modelo y que, por lo tanto, nos permitirá evaluar la bondad de ajuste de nuestro modelo, tenemos que volver al tema de las reglas de identificación. Se trata de presentar unas reglas que nos indiquen si existe una única solución posible para cada uno de los parámetros a estimar por nuestro modelo. Tal como ocurrió cuando nos referimos a los modelos basados en variables perfectamente medidas, hay reglas necesarias y reglas suficientes. Aquí presentaremos las reglas más simples. Una regla necesaria pero no suficiente es que el número de parámetros a estimar sea inferior o igual al número de parámetros conocidos. En nuestro ejemplo, el número de parámetros conocidos es de diez y el número de parámetros a estimar es de ocho. Por tanto, según la regla de t —que es como se llama— nuestro modelo cumple un requisito necesario para su identificación.

Una condición suficiente para que un modelo factorial confirmatorio esté identificado es que todos los indicadores se relacionen con un único factor, que cada factor esté medido por lo menos por tres indicadores, y que los errores de medición no estén correlacionados. Nuestro modelo cumple perfectamente esta regla puesto que nuestros indicadores se relacionan únicamente con un factor, puesto que este factor está medido por cuatro indicadores (más de tres), y puesto que no hay asociación especificada entre los errores de medición para cada uno de los indicadores. Una tercera regla de identificación se aplica a factores medidos por dos indicadores. Una condición suficiente es que todos los factores del modelo estén correlacionados entre sí, que cada indicador se relacione únicamente con un factor y que los errores de medición no estén correlacionados entre sí. Una segunda condición suficiente aplicable a modelos con factores medidos por únicamente dos factores es que cada factor del modelo se relacione al menos con otro factor, que cada indicador se relacione únicamente con un factor y que los errores de medición no estén correlacionados entre sí.

En modelos más complejos no hay reglas simples para determinar si el modelo está identificado. JÖRESKOG y SÖRBOM sugieren un test empírico consistente en estimar primeramente el modelo y guardar la matriz de varianzas y covarianzas predicha por el modelo. Después, se trata de ejecutar el programa nuevamente, utilizando como matriz de varianzas y covarianzas de partida la obtenida al estimar el modelo previamente. Si los parámetros obtenidos al ejecutar el programa las dos veces —1) con la matriz muestral de varianzas y covarianzas, y 2) con la matriz de varianzas y covarianzas estimada en el primer paso— son los mismos, entonces se puede decir que el modelo está identificado.

Para estimar los parámetros de este tipo de modelos, las funciones minimizadoras utilizadas son las mismas que mencionamos en el anterior capítulo, es decir Maximun Likelihood, Unweighted Least Squares y Generalized Least Squares. Las medidas de bondad de ajuste son las mismas que se mencionaron también en el capítulo anterior. Finalmente, es necesario señalar que los resultados obtenidos al estimar estos modelos nos permiten determinar la validez y fiabilidad de nuestros modelos. Para ello, BOLLEN (1989) propone las siguientes medidas. La validez de un determinado indicador es el coeficiente de correlación parcial al cuadrado para la relación entre ese indicador y cualquiera de los conceptos con los que está relacionado. Se trata de un índice lógico dado que la validez de un indicador consiste en el grado en que éste mide lo que se pretende que mida. Si el autoposicionamiento ideológico del entrevistado realmente mide su grado de izquierdismo político, entonces uno esperaría una correlación muy elevada entre el indicador de autoposicionamiento ideológico y el factor izquierdismo político. Utilizamos el coeficiente de correlación parcial al cuadrado por la simple razón de que nuestro modelo puede haber establecido que un determinado indicador mide varios factores a la vez. El coeficiente de correlación múltiple al cuadrado total indicaría el grado de asociación de nuestro indicador con el conjunto de factores del que depende.

BOLLEN también sugiere que el coeficiente de correlación múltiple al cuadrado es el índice más adecuado de fiabilidad. Recordemos que tradicionalmente la fiabilidad de un indicador consiste en el grado en que sucesivas mediciones, utilizando ese indicador, proporcionan el mismo resultado cuando el valor del concepto que se mide no ha cambiado. Para BOLLEN, dado que la consistencia o replicabilidad de un resultado utilizando un indicador cualquiera es difícil de operacionalizar y otros problemas cuya descripción alargaría el libro demasiado, es mejor utilizar una definición alternativa para el concepto de fiabilidad. Ésta sería igual a la magnitud de los efectos directos de todas las variables de un modelo, exceptuando el error de medida sobre un determinado indicador. Es, en definitiva, el coeficiente de correlación múltiple al cuadrado. Las medidas de validez y fiabilidad están por tanto relacionadas. Tanto la una como la otra tienen un valor mínimo de cero y un valor máximo de uno. La validez de una medida, por otra parte, jamás puede exceder su fiabilidad, y éstas son iguales cuando cada indicador mide un sólo concepto, como en nuestro ejemplo. Por otro lado, el lector puede observar que la validez y fiabilidad de un indicador no se pueden establecer de forma definitiva, sino que dependen del modelo estimado.

Para estimar el modelo factorial confirmatorio presentado en los párrafos anteriores, hemos utilizado el siguiente programa en LISREL:

```
Model 5: Análisis Factorial Confirmatorio.
da ni = 6 no = 376 ma = cm
la
*
Ideol Fcohab Reldad Nacban Ingr Educ
km file = spss.lis
sd file = spss2.lis
se
1 2 3 4/
mo nx = 4 nk = 1 lx = fu, fr
fi lx 1 1
va 1.00 lx 1 1
ou se tv rs ef mi ss
```

Los resultados de estimarlo son los siguientes:

$$\text{Autopos} = 1 \cdot \text{IzqPol} + D_1$$
$$\text{Nacban} = 0{,}412 \cdot \text{IzqPol} + D_2$$
$$\text{Fcohab} = 0{,}619 \cdot \text{IzqPol} + D_3$$
$$\text{Reldad} = 0{,}545 \cdot \text{IzqPol} + D_4$$

Si tenemos en cuenta la codificación de cada una de estas cuatro variables, resulta que cuanto menor es el izquierdismo político (recordemos que la escala de Izquierdismo Político la da la variable de autoidentificación política que va de izquierda a derecha) mayor es el desacuerdo con la nacionalización de la banca (esta variable está codificada de mayor a menor acuerdo), mayor es la oposición hacia la cohabitación (esta variable está codificada de actitud más favorable a actitud menos favorable), y mayor es la asistencia a misa (esta variable está codificada de mayor a menor asistencia). Por ejemplo, una unidad de cambio en el izquierdismo político se traduce en 0,619 unidades de cambio en la variable sobre la actitud ante la cohabitación.

Para analizar la bondad de ajuste de este modelo, es decir lo acertados que estamos al suponer que los cuatro indicadores miden un mismo concepto, el izquierdismo político, podemos utilizar índices parciales y totales. Un índice parcial (por referirse a elementos concretos del modelo) es la significación estadística de los tres coeficientes no fijados de antemano. Todos ellos lo son al nivel de confianza del 95 %, y por tanto no podemos rechazar la hipótesis de que son indicadores del concepto izquierdismo político. Dado que cada indicador corresponde a un sólo concepto o factor latente, la validez y la fiabilidad de cada uno de los indicadores es la misma y puede ser medida por el coeficiente de correlación múltiple. Este coeficiente es igual a los siguientes valores para cada uno de los indicadores del modelo:

R^2 autoposicionamiento político: 0,512
R^2 actitud ante la nacionalización de la banca: 0,155
R^2 actitud ante la cohabitación: 0,365
R^2 asistencia a misa: 0,300

Vemos, por tanto, que la validez y la fiabilidad de los distintos indicadores son elevadas pero no excesivamente; parece por otro lado que la actitud ante la nacionalización de la banca es la medida que más desentona con respecto a este factor. En definitiva, estos resultados sugieren que las respuestas a estas preguntas, si bien dependen del grado de izquierdismo político de los entrevistados, dependen también de otros factores no incluidos en el modelo.

Vayamos ahora a medidas de bondad de ajuste global. El coeficiente de determinación es igual a 0,678 que indica que nuestro modelo explica un 68 % de la varianza de los indicadores de Izquierdismo Político. Se trata de un índice elevado que sugiere que nos hallamos ante un buen modelo. El índice de Chi-Cuadrado para este modelo es igual a 14,66 que, con dos grados de libertad, es significativo al 5 %. Ello sugiere que nuestro modelo no reproduce de forma perfecta la estructura de asociaciones entre las variables que estamos manejando. El GFI y el AGFI producen, por otro lado, unos valores iguales a 0,981 y 0,903 respectivamente que, aunque elevados, no lo son tanto como para indicar que se trate de un modelo muy bueno. Finalmente, los valores para Delta 1, Delta 2, Rho 1 y Rho 2 son, respectivamente, 0,93, 0,939, 0,791 y 0,814. De nuevo, nos hallamos ante valores que se hallan en la frontera que separa un modelo realmente bueno de un modelo regular.

Resumiendo, si tomamos toda esta información en bloque, parece que nos hallamos ante un modelo adecuado pero no excepcional, en el cual lo más problemático es el papel jugado por la variable sobre la nacionalización de la banca. El investigador, a partir de aquí debe decidir si modificar o no su modelo, y para ello debe volver a su marco teórico para determinar su posible modificación. Posiblemente sería arriesgado modificar el modelo teórico y eliminar simplemente la variable problemática sin razones muy poderosas. Al fin y al cabo, la justificación para su utilización como medida del izquierdismo político es muy clara. No olvidemos que la estatalización de grandes sectores económicos jugó un papel fundamental en los programas de izquierdas hasta tiempos muy recientes. Quizás lo que ha ocurrido es que con la crisis del socialismo estas políticas ya no se ven como necesarias por gran parte de la gente que se considera de izquierdas. Puede que exista un corte generacional respecto a la significación del ser de izquierdas que convendría analizar, comparando la relación de las actitudes ante la nacionalización de la banca con el izquierdismo político entre jóvenes y personas más mayores.

6

Modelos de relaciones estructurales con variables latentes

Formulación

En los dos capítulos anteriores nos hemos ocupado de analizar: 1) relaciones estructurales entre variables perfectamente medidas, y 2) modelos de medición de factores o variables latentes. En este capítulo juntamos estos dos temas y consideramos modelos causales entre factores latentes medidos por distintos indicadores. De nuevo, el propósito del investigador consiste en evaluar dichos modelos, una vez estimadas las matrices de varianzas y covarianzas basadas en dichos modelos. La hipótesis de partida vuelve a ser que el modelo estimado es el correcto, es decir que la matriz muestral de varianzas y covarianzas es igual a la generada por nuestro modelo. Este modelo tiene dos componentes, uno estructural y otro de medición. El modelo estructural se representa de la siguiente manera:

$$\mathbf{ETA} = \mathbf{BE} \cdot \mathbf{ETA} + \mathbf{GA} \cdot \mathbf{KSI} + \mathbf{ZE}$$

Estas matrices representan a las variables latentes endógenas (ETA), las variables latentes exógenas (KSI), los coeficientes de regresión entre variables endógenas (BE), los coeficientes de regresión entre variables exógenas y variables endógenas (GA), y al coeficiente residual (ZE).

El modelo de medición se representa a su vez de la siguiente manera:

$$x = \mathbf{LX} \cdot \mathbf{KSI} + \mathbf{D} \; y = \mathbf{LY} \cdot \mathbf{ETA} + \mathbf{E}$$

Estas matrices representan a los indicadores exógenos (x), los indicadores endógenos (y), los factores latentes exógenos (KSI), los factores latentes endógenos (ETA), los coeficientes de regresión entre factores exógenos y sus indicadores (LX), los coeficientes de regresión entre factores endógenos y sus indicadores (LY), los errores de medición para los indicadores exógenos (D), y los errores de medición para los indicadores endógenos (E). Todo modelo LISREL, por tanto, incluye ocho matrices de parámetros: GA, BE, LX, LY, PHI, PSI, TD y TE. Como ya hemos indicado anteriormente, si nuestro modelo es correcto, la matriz muestral de varianzas y covarianzas muestrales puede ser expresada en función de los parámetros del modelo.

En primer lugar, la matriz muestral de varianzas y covarianzas entre indicadores endógenos puede expresarse de la siguiente manera:

$$\mathbf{LY\ (I - BE)^{-1}\ (GA \cdot PHI \cdot GA' + PSI)\ [(I - BE)^{-1}]'\ LY' + TE}$$

La matriz muestral de covarianzas entre indicadores endógenos y exógenos puede expresarse de la siguiente manera:

$$\mathbf{LY\ (I - BE)^{-1} \cdot GA \cdot PHI \cdot LX'}$$

Finalmente, la matriz muestral de varianzas y covarianzas entre variables exógenas puede expresarse de la siguiente forma:

$$\mathbf{LX \cdot PHI \cdot LX' + TD}$$

Para ilustrar esto, volvamos a nuestro ejemplo. Para ello complicaremos el modelo hasta ahora desarrollado, de manera que las ecuaciones correspondientes sean las siguientes:

$$\text{Ingresos} = GA_1 \cdot \text{Estudios} + ZE_1$$

$$\text{Ideología} = LY_1 \cdot \text{Izqpol} + E_1 = LY_1 \cdot (GA_2 \cdot \text{Estudios} + BE_1 \cdot \text{Ingresos} + ZE_2) + E_1$$

$$\text{As.Misa} = LY_2 \cdot \text{Izqpol} + E_2 = LY_2 \cdot (GA_2 \cdot \text{Estudios} + BE_1 \cdot \text{Ingresos} + ZE_2) + E_2$$

$$\text{Natban} = LY_3 \cdot \text{Izqpol} + E_3 = LY_3 \cdot (GA_2 \cdot \text{Estudios} + BE_1 \cdot \text{Ingresos} + ZE_2) + E_3$$

$$\text{Cohab} = LY_4 \cdot \text{Izqpol} + E_4 = LY_4 \cdot (GA_2 \cdot \text{Estudios} + BE_1 \cdot \text{Ingresos} + ZE_2) + E_4$$

Nuestro modelo consta de una variable exógena que asumimos que está perfectamente medida, Nivel de Estudios, y de dos variables endógenas, una de ellas perfectamente medida, Ingresos, y la otra que constituye un factor latente, Izquierdismo Político, medido por cuatro indicadores, Autoposicionamiento Ideológico, Asistencia a Misa, Actitud ante la Nacionalización de la Banca, y Actitud ante la Cohabitación. La variable exógena tiene un efecto causal sobre las dos variables endógenas, y la variable Ingresos influye sobre el factor latente Izquierdismo Político. Representar cada varianza o covarianza para las variables contenidas en este ejemplo en función de los parámetros del modelo sería sin embargo excesivamente complejo por la magnitud del modelo. Nos limitaremos a señalar algunas de estas ecuaciones. Por ejemplo, y basándonos en la fórmula para las varianzas y covarianzas entre los indicadores exógenos y endógenos, la covarianza entre la variable ingresos y la variable nivel de estudios es igual a:

$$GA_1 \cdot PHI$$

La covarianza entre la variable autoposicionamiento ideológico y la variable nivel de estudios es igual a:

$$(LY_1 \cdot GA_2 \cdot PHI) - (LY_1 \cdot BE_1 \cdot GA_1 \cdot PHI) = LY_1PHI\ (GA_2 - BE_1GA_1).$$

Como siempre, para poder estimar este modelo debemos asumir que nuestro modelo es perfecto. Además debemos determinar si el modelo está identificado. Una condición necesaria pero no suficiente para que el modelo esté identificado es que el número de parámetros a estimar sea inferior al número de parámetros conocidos. Nuestro modelo tiene un número igual a $1/2(p + q)\,(p + q + 1) = 1/2\,(1 + 5)\,(1 + 5 + 1) = 21$ varianzas y covarianzas conocidas. Tenemos que estimar dos coeficientes GAmma, un coeficiente BEta, dos elementos PS, cuatro elementos TE, y tres elementos LY puesto que al cuarto elemento será asignado el valor uno para fijar la escala del factor Izquierdismo Político. El coeficiente PHI, por otro lado, es igual a la varianza de la variable Nivel de Estudios. Son, por lo tanto, doce los elementos a estimar. Así pues, desde este punto de vista nuestro modelo cumple esta condición necesaria para la identificación del modelo. Una condición suficiente para la identificación, pero no necesaria, es la regla de los dos pasos. Ésta señala que, en primer lugar, el modelo de medición debe estar identificado. Éste es el modelo que relaciona a los factores con sus indicadores. Todas las relaciones estructurales del modelo son eliminadas. Siguiendo esta regla, nuestro modelo de medición está identificado, puesto que dos de nuestros factores, Nivel de Estudios e Ingresos, están medidos sin error (LX y $LY_1 = 1$ y TD y $TE_1 = 0$), y que el factor medido por cuatro indicadores cumple los requisitos establecidos en el anterior capítulo: cada uno de los indicadores está relacionado con un solo factor, hay tres o más indicadores para este indicador y los errores de medida no están correlacionados entre sí. El segundo paso consiste en determinar la identificación del modelo estructural:

$$\text{Ingresos} = GA_1 \cdot \text{Estudios} + ZE_1$$

$$\text{Izqpol} = GA_2 \cdot \text{Estudios} + BE_1 \cdot \text{Ingresos} + ZE_2$$

Este modelo está identificado porque cumple una condición suficiente para la identificación de un modelo, como es el hecho de que se trata de un modelo recursivo.

Puesto que tanto el modelo de medición como el modelo estructural están identificados, podemos asegurar que nuestro modelo está identificado.

Una vez establecida la identificación del modelo, debemos escoger una función minimizadora para asignar valores a los parámetros del modelo, de manera que la matriz de varianzas y covarianzas estimada sea lo más parecida posible a la matriz muestral de varianzas y covarianzas. Como ocurrió con anteriores modelos, las tres funciones minimizadoras utilizadas tradicionalmente son Maximum Likelihood, Unweighted Least Squares y Generalized Least Squares.

Finalmente, la bondad de ajuste del modelo se establece utilizando los distintos índices reseñados anteriormente.

Para estimar el modelo presentado en los párrafos previos, hemos utilizado el siguiente programa LISREL:

```
Model 6: Modelo Estructural con Factores Latentes
da ni = 6 no = 376 ma = cm
la
*
Ideol Fcohab Reldad Nacban Ingr Educ
km file = spss.lis
sd file = spss2.lis
se
1 2 3 4 5 6/
mo nx = 1 fi ne = 5 ne = 2 ga = fu,fr be = fu,fi ly = fu,fr te = di,fr ps = di,fi
ly 1 1 ly 5 2 ly 1 2 ly 2 2 ly 3 2 ly 4 2 ly 5 1 te 5 5 fr be 1 2
va 1.00 ly 1 1 ly 5 2
ou se tv rs ef mi ss
```

Los resultados que obtenemos son los siguientes:

$$\text{IzqPol} = -0{,}75 \cdot \text{Estudios} + 0{,}023 \cdot \text{Ingresos} + ZE_2$$
$$\text{Ingresos} = 0{,}479 \cdot \text{Ingresos} + ZE_1$$
$$\text{Autopos} = 1{,}000 \cdot \text{Izq.Pol} + E_1$$
$$\text{Nacban} = 0{,}405 \cdot \text{Izq.Pol} + E_2$$
$$\text{Fcohab} = 0{,}708 \cdot \text{Izq.Pol} + E_3$$
$$\text{AsMisa} = -0{,}604 \cdot \text{Izq.Pol} + E_4$$

Si nos fijamos, primeramente, en el modelo de medición para el izquierdismo político, observamos que los coeficientes que lo relacionan con cada uno de sus indicadores son muy parecidos a los obtenidos al estimar el modelo de medición de forma aislada. Esta estabilidad en los coeficientes, que siguen siendo estadísticamente significativos, sugiere que aunque no se trata de un modelo de medición perfecto, es todavía moderadamente bueno. La validez y fiabilidad de los indicadores, medida por los coeficientes de correlación múltiple para cada uno de ellos es de 0,452, 0,132, 0,346 y 0,395 respectivamente, es decir muy parecida a la que se obtuvo anteriormente para cada uno de ellos.

Al pasar al modelo estructural, observamos que los entrevistados con mayor nivel de estudios tienden a ser más de izquierdas una vez que controlamos por el nivel de ingresos, y que los entrevistados con mayores ingresos tienden a ser más de derechas una vez que controlamos por el nivel de estudios. Tanto uno como otro efecto son significativos estadísticamente. Lo más interesante desde el punto de vista de este manual consiste en comparar este modelo que incluye una variable dependiente medida por cuatro indicaciones con el path análisis que estimamos anteriormente, en el que el izquierdismo político estaba medido únicamente por el indicador de autoidentificación ideológica. La diferencia fundamental es que mientras que antes el efecto del nivel de estudios era muy pequeño y no significativo estadísticamente, en este modelo el coeficiente para esta variable se multiplica por dos y además se convierte en significativo. Las conclusiones teóricas de analizar uno u otro modelo serían totalmente diferentes. Si nos fiásemos del primero, concluiríamos que el nivel de estudios no afecta al grado de izquierdismo político, mientras que si nos fiamos del segundo, posiblemente una decisión más acertada, debemos concluir que sí que tiene un efecto significativo.

A continuación podemos evaluar la bondad de ajuste de nuestro modelo. En primer lugar, nuestro modelo explica sólo un 3,5 % de la varianza en el grado de izquierdismo político. Se trata de una capacidad explicativa muy pequeña pero notablemente superior a la que se encontró al estimar el path análisis (0,4 %). El coeficiente de determinación para este modelo es de un 26 %, fundamentalmente debido a que el nivel de estudios explica un 24 % de la varianza en el nivel de ingresos. No es éste un índice que debamos utilizar en este caso a la hora de evaluar el modelo, dado que su capacidad explicativa es muy distinta para las dos variables endógenas.

El coeficiente de Chi-Cuadrado para este modelo es igual a 45,8 con ocho grados de libertad, que es significativo al nivel de 5 %. Ello nos permite decir que nuestro modelo no reproduce perfectamente la matriz de varianzas y covarianzas de nuestros datos. El GFI y el AGFI son iguales a 0,959 y 0,891 respectivamente. Se trata por tanto de un modelo adecuado pero no especialmente brillante. Esta impresión nos la corrobora la inspección de los otros coeficientes de bondad de ajuste, Delta 1, Delta 2, Rho 1 y Rho 2, que son iguales respectivamente a 0,869, 0,890, 0,755 y 0,789. Indudablemente no se trata de un gran modelo.

En resumen, todos los índices analizados sugieren que nos hallamos ante un modelo pasable pero esencialmente incompleto. Por un lado, la variable dependiente no está perfectamente medida, puesto que la validez y fiabilidad de sus indicadores no es excesiva, especialmente la correspondiente a la actitud ante la nacionalización de la banca. Por otro lado, está claro que el izquierdismo político depende de algo más que la educación y los ingresos de los individuos. Es más, éstos explican muy poco. Sin embargo, debemos ser cautos, puesto que ni la variable dependiente ni las variables independientes están muy bien medidas. Esto lo sabemos para el caso de la variable dependiente y podría sospecharse de la medida del nivel de estudios y de la medida de ingresos para los entrevistados. Antes de reconsiderar la validez teórica del modelo convendría quizás medir mejor tanto unas como otras.

Coeficientes estandarizados y sin estandarizar

Existe cierta controversia sobre la conveniencia de utilizar coeficientes estandarizados o sin estandarizar en el análisis estadístico. El problema subyacente en estas discusiones es el de cómo juzgar las magnitudes absolutas y relativas de un coeficiente. Una primera equivocación es la de creer que la magnitud de un coeficiente está determinada por su significación estadística. Esto es absolutamente falso. La significación estadística de un coeficiente depende, es cierto, de su magnitud, pero también de otros factores como el tamaño de la muestra y la varianza de las variables independientes y dependientes. Cuanto mayor es la magnitud de la relación, mayor es la probabilidad de obtener una relación estadísticamente significativa. Cuanto mayor es el tamaño de una muestra, mayor es la probabilidad de obtener una relación estadísticamente significativa. Por último, cuanto menor es la varianza de las variables independientes y dependientes, mayor es la probabilidad de obtener una relación estadísticamente significativa.

Por tanto, cuando evaluamos la magnitud de un determinado coeficiente, debemos tener en cuenta que su significación estadística está parcialmente desligada de la magnitud de su efecto. Cuando comparamos el efecto de dos variables independientes sobre una variable dependiente para una muestra determinada, debemos tener en cuenta que su significación estadística dependerá tanto de la magnitud relativa de los coeficientes como de la varianza de las dos variables independientes. Por lo tanto, la significación estadística de un coeficiente no es el mejor indicador de su magnitud absoluta o relativa. Por otra parte, no podemos hablar de la magnitud de un determinado coeficiente a no ser que sea significativo estadísticamente.

Un primer paso, por lo tanto, a la hora de hablar de magnitudes, es el de centrarnos en aquellos efectos que son significativos estadísticamente. Una vez hecho esto, aunque se puede decir si un coeficiente es grande o pequeño, determinar si indica una relación fuerte o débil es arbitrario y depende del objeto de la investigación. Si se trata de utilizar los resultados de una investigación para establecer una determinada política, la magnitud de un coeficiente dependerá del coste de manipular la variable independiente contrastado con los efectos de tal manipulación. Cuando comparamos coeficientes entre distintas variables el problema se complica porque al estar medidas en distintas escalas no son comparables. Por ejemplo, nos basta cambiar de pesetas a duros en una variable independiente para que su coeficiente se divida en cinco. Este problema de comparabilidad ha llevado a buscar algún procedimiento para transformar dichos coeficientes y favorecer la comparación. Uno de estos procedimientos consiste en estandarizar los coeficientes para tratar de eliminar el efecto de escala. Esto se consigue multiplicando cada uno de ellos por la razón entre la desviación típica para la variable independiente y la desviación típica para la variable dependiente.

La utilidad de los coeficientes estandarizados, por otra parte, es limitada. No debemos, por ejemplo, comparar coeficientes estandarizados para una misma variable independiente en distintas muestras puesto que las diferencias que se observen dependerán en parte de la desviación típica de la variable en cuestión en las distintas muestras. Otra precaución consiste en no intentar solucionar el problema de la variación en la escala de las variables independientes a base de analizar la matriz de correlaciones en lugar de la matriz de varianzas y covarianzas. La matriz de correlaciones, efectivamente, elimina el problema de escala al indicar el cambio en desviaciones típicas de una variable asociado con cambios de una desviación típica en la otra. Aunque los coeficientes que se obtienen por este procedimiento son correctos, los tests de significación no lo son, puesto que no incorporan la distribución muestral de la razón entre la desviación típica de las variables independientes por las variables dependientes. Es mejor, por lo tanto, analizar la matriz de varianzas y covarianzas y, si acaso, computar los coeficientes estandarizados con posterioridad.

Medias y constantes

Toda la presentación anterior se ha basado en la utilización de variables centradas, es decir expresadas en función de sus desviaciones respecto a la media. Es posible, sin embargo, estimar tanto la media para los factores latentes como el elemento constante en cada uno de los modelos de regresión incluidos en los modelos. Esto conlleva una reespecificación de los modelos, de la manera siguiente:

$$\mathbf{ETA = Alpha + BE \cdot ETA + GA \cdot KSI + ZE}$$

$$y = vy + \mathbf{LY \cdot ETA + E}$$

$$x = vx + \mathbf{LX \cdot KSI + D}$$

La necesidad de estimar nuevos parámetros complica el problema de la identificación del modelo. Ésta se suele resolver igual que se resolvió el problema de escala para las variables latentes, fijando determinados parámetros arbitrariamente. La solución típica consiste en dar el valor cero a la constante de una de las ecuaciones que relacionan cada una de las variables latentes con sus indicadores. Es decir que uno de los indicadores para cada variable latente provee a la vez la escala y el origen una vez que se fija el LX o LY correspondiente en el valor uno, y el *vy* o *vx* correspondiente en cero. Otra alternativa consiste en fijar la varianza de KSI o ETA en un valor igual a uno, y la media de la variable latente en el valor cero. La introducción del cómputo de medias y constantes, sin embargo, en ningún modo afecta al cómputo de los demás elementos del modelo ni a sus índices de bondad de ajuste.

Comparación de grupos

En determinadas ocasiones, interesa determinar hasta qué punto un modelo determinado es aplicable a dos o más grupos diferentes. La comparabilidad de este modelo puede evaluarse en cuanto a su forma y en cuanto al valor de los coeficientes para los distintos parámetros estimados. En nuestro ejemplo, podríamos comparar a los jóvenes y a los viejos, para determinar si la forma del modelo representada por sus ecuaciones es igual y, posteriormente, para determinar si los coeficientes del modelo son iguales para viejos y jóvenes. En general, a la hora de comparar modelos se establece una jerarquía de comparación. Para los modelos estructurales con variables perfectamente medidas, la jerarquía de hipótesis que se establece suele ser la siguiente:

1. La forma del modelo es la misma para los grupos comparados.
2. Los coeficientes BEta y GAmma son los mismos para los grupos comparados.
3. Los coeficientes BEta, GAmma y PSi son iguales para los grupos comparados.
4. Los coeficientes BEta, GAmma, PSI y PHi son iguales para los grupos comparados.

Para los análisis factoriales confirmatorios la jerarquía que se establece es la siguiente:

1. La forma del modelo es la misma para los grupos comparados.
2. Los coeficientes LX y TD son iguales para los grupos comparados.
3. Los coeficientes LX, TD y PHI son iguales para los grupos comparados.

Finalmente, en modelos que combinan el estudio de relaciones estructurales con la utilización de variables latentes, la jerarquía de comparaciones que se suele establecer es la siguiente:

1. La forma es la misma.
2. Los coeficientes LX y LY son los mismos.
3. Los coeficientes LX, LY, BEta y GAmma son iguales.
4. Los coeficientes LX, LY, BEta, GAmma, TD y TE son iguales.
5. Los coeficientes LX, LY, BEta, GAmma, TD, TE y PSi son iguales.
6. Los coeficientes LX, LY, BEta, GAmma, TD, TE, PSi y PHi son iguales.

Por supuesto, según las características de la investigación, este orden puede ser modificado. Se suele ir verificando una a una cada una de las hipótesis, de manera que no se avanza al paso siguiente si no se ha verificado la hipótesis anterior.

LISREL permite someter a prueba cada una de estas hipótesis utilizando tests de Chi-Cuadrado. Este test establece si la hipótesis nula puede o no ser rechazada. Los grados de libertad en este test son $1/2(G)\ (p + q)\ (p + q + 1) - t$, donde G representa al número total de grupos. Por otra parte, cuando la hipótesis que se somete a prueba incluye las condiciones de una hipótesis menos restrictiva además de alguna otra, uno puede comparar el valor de Chi-Cuadrado correspondiente a las dos hipótesis puesto que la diferencia entre los dos valores se distribuye también como una función de Chi-Cuadrado con un número de grados de libertad igual a la diferencia de grados de libertad entre los dos modelos. Por ejemplo, si uno quisiera comprobar si los coeficientes PS para dos grupos son iguales, una vez comprobado que tanto los coeficientes LX, LY, BEta, GAmma, TD y TE, compararíamos los coeficientes de Chi-Cuadrado obtenidos al poner a prueba tanto la hipótesis más restrictiva como la menos restrictiva.

Conclusión

Este manual supone una introducción básica a los modelos de ecuaciones estructurales con variables latentes (LISREL), que constituyen el armazón metodológico sobre el que se funda el análisis de relaciones causales a nivel cuantitativo. Tal como dijimos al principio, esta técnica tiene como principal objetivo el de permitir el análisis combinado tanto de efectos directos como de efectos indirectos, y la inclusión dentro de los modelos de la noción de error de medición. Hemos intentado minimizar el uso de notación estadística, de manera que queden claros los conceptos; sin embargo, ésta es indispensable para entender verdaderamente una técnica tan compleja como LISREL. Por ello, referimos al lector interesado a la bibliografía incluida al final de este manual. Algunos de estos manuales pueden ser entendidos perfectamente si se invierte un poco de tiempo en el aprendizaje del cálculo matricial, y en la memorización de los símbolos utilizados. Tampoco hemos abordado temas más especializados relacionados con esta técnica. Sin embargo, conviene enumerarlos. En primer lugar, esta técnica permite no sólo la estimación de relaciones estructurales entre variables latentes sino que también permite estimar relaciones estructurales entre variables latentes de segundo orden, es decir variables latentes bajo las cuales subyacen otras variables latentes medidas por sus respectivos indicadores. También, hemos dejado de lado modelos en los cuales las variables observadas o medidas influyen sobre variables latentes en lugar de ser afectadas por ellas. En tercer lugar, hemos omitido hablar de las complejidades añadidas por la inclusión en estos modelos de variables dicotómicas u ordinales. Conviene decir que las versiones más recientes de programas relacionados con esta técnica permiten estimar modelos en los cuales uno utiliza variables dicotómicas y ordinales. La función minimizadora utilizada para estimar estos modelos es Weighted Least Squares, una versión especializada de la función Generalized Least Squares.

Los avances técnicos relacionados con esta técnica son muy rápidos en la actualidad así como los avances en el terreno de la programación. Aunque el programa LISREL de JÖRESKOG y SÖRBOM, ya en la versión número 7, fue pionero y es el más conocido, en tiempos recientes han sido puestos a la venta otros programas como EQS, desarrollado por BENTLER. En general, estos programas son más o menos iguales en sus aspectos técnicos, aunque siempre hay alguna diferencia en el tipo de índices de bondad de ajuste u otros aspectos relacionados con el output que justifican la puesta a la venta de estos programas. También existen diferencias notables en el lenguaje de programación utilizado, aunque unos como otros se caracterizan por venir acompañados de manuales muy fáciles de entender. Esperamos que el acceso que los investigadores tengan a estos manuales y programas favorezca el desarrollo de la sociología cuantitativa en España.

Bibliografía comentada

La bibliografía que se presenta a continuación es breve. Contiene los manuales básicos en los que se ha basado este cuaderno, y que cubren de manera a la vez exhaustiva y simple los temas principales tratados en él.

ALWIN, D. F., y HAUSER, R. M. (1975), «The Decomposition of effects in path-analysis», *American Sociological Review*, 40:37-47. Artículo espléndidamente escrito en que se describe cómo descomponer los efectos causales de una variable sobre otra, mediante demostraciones matemáticas extremadamente fáciles de seguir.

BENTLER, P. M. (1985), *Theory and Implementation of EQS: A Structural Equations Program*, Los Ángeles, BMDP Statistical Software. BENTLER es una de las personas que más han contribuido al desarrollo de las técnicas descritas en este libro, más desde el campo de la psicología. En este libro describe brevemente las técnicas y cómo aplicarlas utilizando el programa EQS que él desarrolló. EQS es muy parecido al programa LISREL, aunque algunos lo prefieren porque para programar no hay sino que escribir las ecuaciones que se quieren estimar, sin tener que entrar en las complejidades de programación en LISREL.

BLALOCK, H. M. (1967), *Causal Models in the Social Sciences*, Chicago, Aldine-Atherton. Libro clásico en que se describe la lógica del análisis causal en ciencias sociales, presentándose numerosos tipos de modelos y su solución matemática.

BOLLEN, K. A. (1989), *Structural Equations with Latent Variables*, Nueva York, Wiley and Sons. Éste es el mejor libro que se ha publicado hasta el momento sobre la técnica denominada LISREL y sobre el que se basa el manual aquí presentado. Es muy claro, exhaustivo y lleno de ejemplos. Aunque hay que dedicarle tiempo, contiene toda la información que se necesita para seguirlo, siempre que se quiera dedicar algo de tiempo al cálculo matricial. La mayoría de los interesados podrán aprender todo lo que necesiten sobre esta técnica en este libro que esperemos se traduzca pronto.

COOK, T. D., y CAMPBELL, D. T. (1979), *Quasi-Experiments*, Boston, Houghton Mifflin Company. Éste es un libro básico para entender la lógica experimental aplicada a las ciencias sociales. Describe todo tipo de diseños experimentales y quasi-experimentales, especificando sus ventajas y problemas.

DUNCAN, O. D. (1975), *Introduction to Structural Equation Models*, Nueva York, Academic Press. Este libro describe, utilizando el lenguaje de las correlaciones, todo aquello relacionado con el path análisis. Se trata de un libro pequeño y muy fácil de seguir.

FOX, J. (1984), *Linear Statistical Models and Related Models*, Nueva York, Wiley and Sons. Un libro imprescindible para entender todo aquello relacionado con los modelos de regresión y sus derivados, LISREL inclusive. Aunque es más complicado que otros manuales por su uso constante del cálculo matricial, con un poco de paciencia se puede entender y se convierte entonces en un libro de consulta indispensable.

JÖRESKOG, K. G., y SÖRBOM, D. (1988), *LISREL 7: A Guide to the Program and Applications*, Chicago, SPSS. Inc. Manual para la última edición de LISREL. Se trata de un manual escrito con mucha mayor claridad que los anteriores y que incluye numerosos ejemplos.

Números publicados

1. **Métodos de muestreo**
   Jacinto Rodríguez Osuna
2. **Metodología de la evaluación de programas**
   Francisco Alvira Martín

Próximos números

Análisis de regresión múltiple
Mauro Guillén

Historias de vida en las ciencias sociales
Juan José Pujadas

Métodos de muestreo. Casos prácticos
Jacinto Rodríguez Osuna