

TEMA 1

Antecedentes de los modelos de ecuaciones estructurales y uso actual

- 1.1. ANTECEDENTES PSICOMÉTRICOS
- 1.2. ANTECEDENTES BIOMÉTRICOS Y ECONOMÉTRICOS
- 1.3. MODELOS DE ECUACIONES SIMULTÁNEAS CON VARIABLES LATENTES

Para ilustrar los antecedentes de los modelos de ecuaciones estructurales, conviene tener en cuenta las preguntas o los problemas sustantivos a los que esta metodología ha intentado dar respuesta. En pocas palabras, los modelos de ecuaciones estructurales tratan de representar hipótesis teóricas sobre las medias, varianzas y covarianzas de los datos observados a partir de un pequeño número de parámetros definidos desde un modelo hipotético elaborado por el investigador (Kaplan, 2000). Esta visión realmente cubre un gran espectro de posibilidades. Sin embargo, una revisión de sus orígenes podría ayudar a delimitar las problemáticas que esta metodología ha tratado de resolver.

1.1. ANTECEDENTES PSICOMÉTRICOS

Los modelos de ecuaciones estructurales representan una combinación de dos tradiciones estadísticas. Por un lado, el Análisis Factorial desarrollado dentro de la Psicología, cuyo objetivo es explicar un conjunto de variables observadas (ítems de un test, por ejemplo) mediante un número menor de variables latentes no observables, denominadas factores (dimensiones teóricas) (Barbero et al., 2011). Y, por otro lado, los modelos de ecuaciones simultáneas desarrollados fundamentalmente dentro de la tradición econométrica y que, recientemente, han sufrido un gran avance gracias a la genética (Kline, 1998).

Se pueden considerar tres elementos decisivos en el desarrollo del Análisis Factorial: a) La apertura del laboratorio antropomórfico de Galton (1822-1911) para el estudio de las diferencias individuales; b) El desarrollo de la correlación de Pearson (Pearson y Lee, 1904) que permitió comenzar a estudiar las relaciones entre variables; y c) La interpretación de Spearman (1904) considerando que la correlación entre dos variables indica que ambas podrían tener un factor común.

Sir Francis Galton, primo de Charles Darwin, abrió un laboratorio antropomórfico en Londres donde medía una gran variedad de variables físicas y sensoriales de los sujetos. Consideraba que la inteligencia se podría medir a través de la actividad sensoriomotora. Una vez recogidos los datos, comenzó a elaborar distribuciones de frecuencias y comprobó que la forma era aproximadamente igual tanto para variables psicológicas como anatómicas, es decir similar a la distribución normal. Sin embargo, ante la can-

tividad ingente de datos, Galton se planteó el problema de cómo encontrar relaciones entre todas las variables medidas. Para responder a esta cuestión, recurrió a la correlación, siendo de esta forma el primero en aplicar la media, varianza, mediana y distribución normal a datos psicológicos. Posteriormente, esta tarea fue continuada por Karl Pearson (Barbero et al., 2015).

Por su parte, Charles Spearman entendía que la correlación existente entre distintos test de habilidad mental indicaba la existencia de un factor general común a todos ellos que podría explicar su asociación. En términos formales, ello se puede traducir de la siguiente manera:

$$\rho_{ij} = \lambda_i \lambda_j \quad (1.1)$$

donde:

- ρ_{ij} es la correlación poblacional entre el test i y el j .
- $\lambda_i \lambda_j$ son los pesos o saturaciones que relacionan el test i y j respectivamente con el factor general común.

Si retomamos la visión anterior sobre los modelos de ecuaciones estructurales, podríamos decir que la ecuación 1.1 expresa la correlación poblacional mediante un conjunto de parámetros estructurales. Es decir, la correlación poblacional entre el test i y el j es igual al producto de los coeficientes que relacionan el factor general común con cada uno de los test. Esto es lo que hoy en día llamamos factores o dimensiones de primer orden.

Louis Leon Thurstone es otra figura relevante en el desarrollo del Análisis Factorial. Este autor entendía que no existía un factor general común que explicara factores más específicos, sino que pensaba que existían grupos de factores mayores referidos como "habilidades mentales superiores" (Thurstone, 1938). De acuerdo con el principio de parsimonia, Thurstone pensaba que cada factor simple debería estar compuesto por la máxima covariación no compartida del conjunto de variables observadas como sea posible. Sin embargo, para lograr esta estructura tenía que permitir que dichos factores simples correlacionaran entre sí (Thurstone, 1947). Por tanto, en este contexto multifactorial, el factor de habilidad general tendría que existir en un nivel superior al de los factores simples de primer orden para poder explicar su variación común (Mulaik, 1972). Esto es lo que hoy en día llamamos factores o dimensiones de segundo orden.

Posteriormente, entre la década de los 50 y 60 el Análisis Factorial fue ganando una popularidad creciente, en parte gracias al desarrollo de procedimientos estadísticos cada vez más sofisticados que permitían poner a prueba el número de factores subyacentes a los datos empíricos de las variables. Ello se pudo realizar gracias al desarrollo de métodos de estimación como el de *máxima verosimilitud* (en inglés, *maximum likelihood*, o *ML*; Jöreskog, 1967; Lawley y Maxwell, 1971). Minimizando el logaritmo negativo de la función de máxima verosimilitud, es posible calcular el test de razón de verosimilitudes que, basado en la distribución de χ^2 , indica hasta qué punto el modelo (número de factores, en este caso) se ajusta a la forma en que se relacionan las variables empíricas. Posteriormente, Jöreskog y Goldberger (1972) desarrollan otro método de estimación denominado *mínimos cuadrados generalizados* (en inglés, *generalized least squares*, o *GLS*) que trata de reducir al mínimo los errores de estimación.

Sin embargo, aparte del número de factores latentes, también interesa conocer cómo se relacionan estos factores entre sí. Como respuesta a esta situación, Jöreskog (1969) desarrolla los fundamentos del Análisis Factorial Confirmatorio (AFC) que permite testar hipótesis tanto sobre el número de factores como sobre la relación que establecen entre sí, así como entre los factores y las variables observadas. En contraposición, a las técnicas de Análisis Factorial previas que se les adjudicó el apellido Exploratorio (Análisis Factorial Exploratorio o AFE). Desde una perspectiva histórica, el AFC permitió poner a prueba la idea de Thurstone, ya que ahora era posible especificar un modelo donde un conjunto de ítems (subtest) pesara exclusivamente en algunos factores y no en otros.

Hemos de reconocer que, en la investigación en psicología y otras ciencias sociales, el AFE y el AFC son técnicas de una enorme tradición. Sin embargo, el alcance de los modelos de ecuaciones estructurales es mucho más amplio, ya que permiten también poner a prueba la forma en que se relacionan los factores o variables latentes. Este sistema complejo de relaciones se puede representar mediante sistemas de ecuaciones simultáneas.

1.2. ANTECEDENTES BIOMÉTRICOS Y ECONOMÉTRICOS

Además del Análisis Factorial, el otro componente que define los modelos de ecuaciones estructurales es el "*path*" análisis, o análisis de senderos o vías. Básicamente, estos "*paths*" permiten especificar las relaciones

de dependencia existentes entre distintas variables a partir de pesos de regresión con una dirección concreta (normalmente, unidireccional, donde una variable x explicaría el cambio en otra variable y , por ejemplo). Este enfoque tiene su origen en la biometría a partir de los trabajos de Wright (1934, 1960), que mostró que la correlación entre variables podría relacionarse con los parámetros de un modelo representado mediante un diagrama "path" o de senderos. Además, demostró que las ecuaciones del modelo permitían estimar distintos tipos de efectos (directos, indirectos y totales) entre las variables.

Otro hito importante se debe a la econometría derivado de los trabajos de Haavelmo (1943), cuyo interés se centraba en encontrar un modelo matemático que relacionara la interdependencia entre las variables económicas mediante sistemas de ecuaciones simultáneas. De manera formal:

$$y = By + \Gamma x + \zeta \quad (1.2)$$

donde:

- y es el vector de variables endógenas o dependientes que han de ser explicadas por el modelo.
- x es el vector de variables exógenas o independientes que explican y .
- ζ es el vector correspondiente a los términos de error.
- B y Γ son las matrices de los coeficientes correspondientes a las matrices y y x , respectivamente.

En el desarrollo de los modelos de ecuaciones simultáneas fue clave la "Cowles Commission of Research in Economics" compuesta por un nutrido grupo de estadísticos teóricos y econométricos de la Universidad de Chicago (Berndt, 1991). Este grupo se encargó de desarrollar las ecuaciones simultáneas a partir de los métodos de estimación de máxima verosimilitud.

1.3. MODELOS DE ECUACIONES SIMULTÁNEAS CON VARIABLES LATENTES

Los modelos de ecuaciones estructurales surgen a partir de la combinación de las dos técnicas mencionadas anteriormente y, fundamentalmente, gracias a los trabajos de Jöreskog (1973). En palabras de Jöreskog y

Sörbom (1989, pág. i): *"La idea de combinar características de modelos psicométricos y econométricos nació en la mente del primer autor en la primavera de 1970, y fue inspirada por los trabajos del profesor Arthur S. Goldberger publicados en Psychometrika, 1971. La primera versión de LISREL fue Linear Structural Equation Modeling for Latent Variables, donde cada variable latente estaba representada por un único indicador observable, posiblemente poco fiable. Este modelo fue presentado por primera vez en la conferencia titulada: "Structural Equation Models in the Social Science" desarrollada en Madison, Wisconsin, en noviembre de 1970...El modelo LISREL se generalizó entre 1971 y 1972 incluyendo modelos previamente desarrollados para variables latentes con múltiples indicadores, para AFC (Jöreskog, 1969), para AFC simultáneo en distintas poblaciones (Jöreskog, 1971), y para modelos más generales de estructuras de covarianzas (Jöreskog, 1970, 1973, 1974)..."*

Grosso modo, según Jöreskog (1973), los modelos de ecuaciones estructurales se componen de dos partes:

1. El modelo de medida, que permite relacionar las variables observadas o indicadores con los factores latentes mediante el AFC.
2. El modelo estructural, que relaciona las variables latentes entre sí mediante las ecuaciones simultáneas.

A partir de aquí comienza el desarrollo de *softwares*, entre los que nos gustaría destacar tanto LISREL como el paquete *lavaan* de R como dos de las mejores opciones para implementar modelos de ecuaciones estructurales en la actualidad.

Los modelos de ecuaciones estructurales han experimentado un notable desarrollo en los últimos 40 años. Destaca el desarrollo de programas informáticos que hacen cada vez más flexible la aplicación de esta técnica a partir de datos cuya distribución se aleja de la normal. Entre los avances más notables cabe destacar el desarrollo de nuevos métodos de estimación que permiten estimar modelos cuyas variables son dicotómicas o categóricas, o son variables continuas con distribuciones no-normales (Kaplan, 2000).

Una de las principales características de esta técnica es que la toma de decisiones se basa en una continua interacción entre los sustratos conceptuales de la teoría y los índices de ajuste que se van obteniendo en cada modelo teórico puesto a prueba. De esta forma, en la medida que se examina un determinado modelo teórico, los índices de ajuste basados,

exclusivamente, en criterios estadísticos, nos orientan sobre posibles modificaciones a realizar. A su vez, estas modificaciones han de ser sometidas al juicio del investigador, cuyo criterio teórico es fundamental y lo más importante para valorar la inclusión de las modificaciones sugeridas desde las fuentes de desajuste. En la Figura 1.1 se presenta un esquema general sobre las distintas fases del modelo.

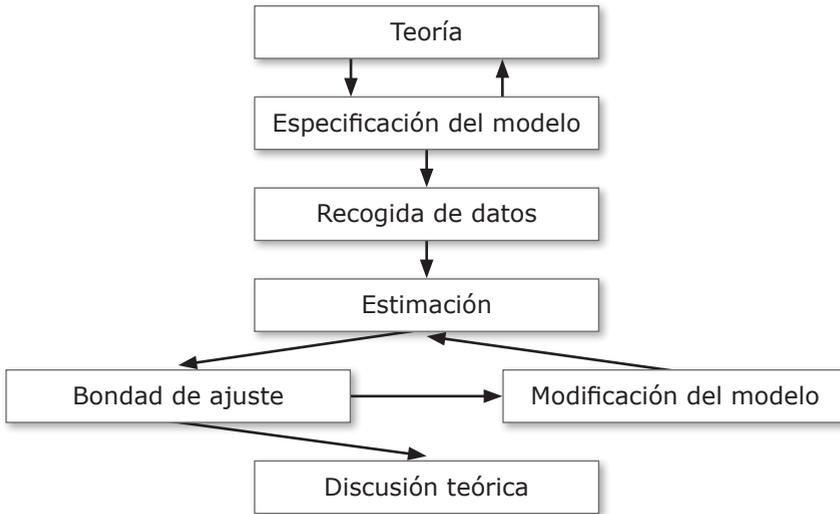


Figura 1.1. Esquema general fases modelado de ecuaciones estructurales (tomado de Kaplan, 2000).

En primer lugar, y siempre desde un enfoque deductivo, es necesario y obligatorio contar con un modelo teórico que ha de hacerse explícito mediante un diagrama que represente la combinación de las variables de interés a través de relaciones de dependencia. Una vez que el modelo se hace explícito, se recogen los datos necesarios seleccionando la muestra de sujetos más adecuada sobre la que extraer las inferencias pertinentes. Tras depurar y examinar previamente los datos, se estiman los parámetros especificados que tratan de reproducir, en la medida de lo posible, las correlaciones, varianzas, covarianzas y medias presentes en los estadísticos obtenidos en nuestra muestra de estudio. Precisamente, el objetivo de los índices de bondad de ajuste es cuantificar la discrepancia entre la matriz de varianzas-covarianzas reproducida por los parámetros estimados y la empírica (la que obtenemos en nuestra muestra). Estos índices de bondad de ajuste orientan al investigador sobre la adecuación del modelo

teórico puesto a prueba (especificado mediante los parámetros a estimar) a la hora de reproducir la matriz de varianzas-covarianzas empírica de la muestra. A partir de esta información, el investigador puede introducir las modificaciones que crea que pueden contribuir a la mejora del ajuste, si es que éstas tienen algún sentido teórico.

A título de ejemplo, mostramos el modelo comprensivo del sistema educativo propuesto por Oakes (1986). En la Figura 1.2, se observa que el modelo se compone por una serie de variables de "entrada", de "proceso" y de "resultados", y dichas variables interactúan entre sí. Así, por ejemplo, los recursos económicos que recibe una escuela afectan de manera directa a la calidad de la misma. A su vez, cuanto mejor sea la calidad de la escuela mejor serán las clases impartidas, y mejor será el rendimiento académico de los estudiantes. Algunos aspectos a destacar al observar la Figura 1.2. son:

1. Lo complejo que es describir un modelo teórico sobre el comportamiento de las variables. Hacer explícitas las relaciones entre las variables puede resultar harto complicado si el investigador no tiene un conocimiento profundo de los conceptos que pueden ser relevantes y cómo se pueden relacionar unas con otras.
2. A pesar de contar con un modelo teórico relativamente claro, a veces, es complicado determinar qué indicadores del constructo se han de utilizar. Así, por ejemplo, entre los constructos relevantes encontramos las "aspiraciones y actitudes" de los estudiantes, pero ¿cómo medimos estas aspiraciones?, y ¿qué tipo de actitudes hemos de investigar?
3. En tercer lugar, es necesario considerar que el modelo se desarrolla en un sistema jerárquico. Es decir, los resultados de los estudiantes se deben en gran medida a las variables que actúan en el proceso de enseñanza-aprendizaje que se desarrolla en la clase. A su vez, las variables que inciden en el "nivel clase", se verán afectadas por cuál sea la política de la región o país.

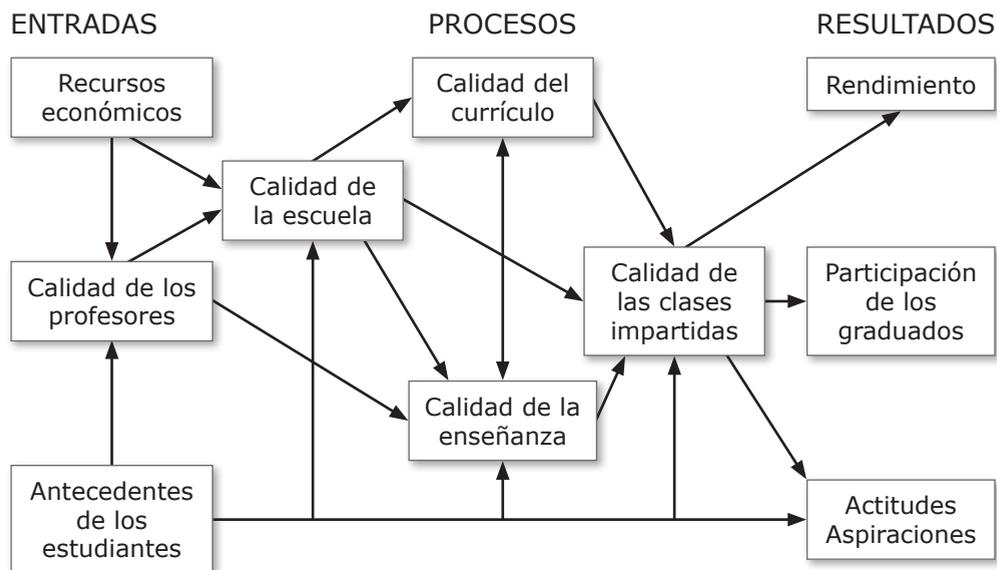


Figura 1.2. Modelo comprensivo del sistema educativo, adaptado de Oakes (1986).

Como se menciona anteriormente, es deseable tener modelos lo más detallados posible. Sin embargo, en modelos excesivamente desplegados o exhaustivos como el anterior, pueden aparecer dificultades en su interpretación. Esto conlleva a que la capacidad para derivar indicadores o variables que se puedan medir directamente se vea mermada, dado que no sabremos a ciencia cierta qué hemos de medir. Los modelos excesivamente complicados suelen ser interpretables y de utilidad, casi exclusivamente, por sujetos especializados, pero no para otras muchas audiencias implicadas, entre las que se suelen encontrar los responsables en las tomas de decisiones (Holgado, 2002).

Por ello, conviene matizar que la medición, al margen del contexto teórico, no consolida por sí misma las teorías (Barbero et al., 2015). En este sentido, Miller (1982, pág. 115) entiende que: *"...muchos psicólogos se han precipitado en buscar números antes de saber lo que esos números pueden significar...todavía se pueden encontrar psicólogos que llevan a cabo mediciones complejas y exageradamente precisas sólo para demostrar hasta qué punto el psicólogo puede ser científico. Hay gente que no admite que, si apenas merece la pena hacer una cosa, hacerla bien sigue sin merecer la pena"*.

Sin la ayuda de los programas informáticos sería muy laborioso realizar cualquier tipo de análisis mediante modelos de ecuaciones estructurales.

Durante casi 25 años, LISREL ha sido prácticamente el único programa disponible. Actualmente, la situación ha cambiado mucho, y en el mercado podemos encontrar una gran variedad de programas. Entre ellos, podemos citar AMOS (Analysis of Moment Structures; Arbuckle, 1997), CALIS (Covariance Analysis and Linear Structural Equations; Hartmann, 1992), EQS (Equations; Bentler, 1995); Mplus (Muthen y Muthen, 1998), RAMONA (Reticular Action Model or Near Approximation; Browne et al., 1997), o paquetes de R como *sem* (Fox, 2006; Fox et al., 2010) o *lavaan* (Rosseel, 2011), que tienen la ventaja crucial respecto a los anteriores de ser de libre acceso y disponer de innumerables tutoriales.

Muchos de estos programas permiten de una manera sencilla y directa poner a prueba los modelos teóricos de los investigadores. Sin embargo, esta aparente facilidad puede encerrar graves inconvenientes. Por un lado, estos programas, por defecto, a veces toman decisiones por el investigador sobre la especificación del modelo, poniéndose a prueba modelos que pueden estar alejados de lo que el investigador trata de examinar, incluso sobre modelos sin lógica teórica alguna se podría encontrar un ajuste perfecto simplemente por una cuestión instrumental. Otras veces, no tienen en cuenta la naturaleza métrica de las variables, sesgando tanto los parámetros que se estiman como la interpretación conceptual de los resultados. Así pues, la utilidad y fácil aplicación que nos brindan muchos programas debe estar sometida a una correcta interpretación de lo que se está haciendo. Por ello, creemos que es muy necesario entender bien la lógica subyacente a los modelos de ecuaciones estructurales y para ello, desde nuestro punto de vista, LISREL ofrece la posibilidad de entender bien dicha lógica, extrapolable a cualquier otro programa, y al mismo tiempo, permite utilizar uno de los mejores programas sobre la temática que nos ocupa. De manera complementaria, también hablaremos largo y tendido sobre los mismos modelos y de su implementación a través de la sintaxis del paquete *lavaan* de R.

Algunas revistas donde podemos encontrar aplicaciones y desarrollos teóricos sobre modelos de ecuaciones estructurales son:

- Applied Measurement in Education
- Applied Psychological Measurement
- Behavior Research Methods
- Behaviormetrika

- Biometrika
- British Journal of Mathematical and Statistical Psychology
- Educational and Psychological Measurement
- Educational Measurement: Issues and Practices
- European Journal of Psychological Assessment
- International Journal of Testing
- Journal of Multivariate Analysis
- Multivariate Behavioral Research
- Psychological Bulletin
- Psychological Methods
- Psychometric Monographs
- Psychometrika
- Quality and Quantity: International Journal of Methodology
- Structural Equation Modeling

De entre todas ellas, merece una mención especial la última (*Structural Equation Modeling*), cuyo primer volumen fue publicado en enero de 1994. En la revista se pueden encontrar artículos que abordan todas las disciplinas académicas cuyo interés se centra en los modelos de ecuaciones estructurales, incluyendo psicología, sociología, educación, política o economía.