

MODELOS DE ECUACIONES ESTRUCTURALES: OPERACIONES Y CONCEPTOS BÁSICOS

Lic. Juan Carlos Correa

RESUMEN

El uso de modelos de ecuaciones estructurales exige la comprensión de diversas operaciones y conceptos asociados con la teoría clásica de los tests, la teoría de distribución asintótica y algunas aplicaciones del álgebra matricial a la estadística inferencial, en el contexto de la contrastación de hipótesis teóricas. El propósito del presente artículo es destacar la relación entre estos conceptos y operaciones en el uso general de esta metodología. El apéndice incluido al final amplía algunas ideas expuestas a lo largo del artículo.

Palabras claves: modelos de ecuaciones estructurales, especificación, identificación, estimación, contrastación, variables observadas, variables latentes.

MODELOS DE ECUACIONES ESTRUCTURALES: ANTECEDENTES Y NOCIONES PRELIMINARES

Los conceptos y operaciones involucradas en el uso de los modelos de ecuaciones estructurales (SEM) encuentran sus antecedentes en la tradición del análisis factorial (Spearman, 1904), la teoría clásica de los tests (Lord y Novick, 1968) y el enfoque tradicional sobre el contraste de hipótesis estadísticas (Fisher, 1936). Desde estas perspectivas, se advirtió una doble contribución para la investigación en las ciencias del comportamiento. Primero, la necesidad de distinguir conceptual y metodológicamente la diferencia entre variables latentes, variables de disturbancia o error y variables observadas (ver nota 2 del apéndice); y en segundo lugar, se desarrolló la posibilidad de modelar, validar y estimar el ajuste de sus relaciones a la luz de un marco de referencia teórico de interés (Breckler, 1990; Bollen, 2002; MacCallum y Austin, 2000).

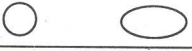

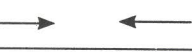
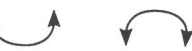
Clásicamente, suele distinguirse entre dos tipos generales de modelos, a saber: “modelos de medición” y “modelos estructurales” (ver nota 1 del apéndice). Un modelo

de medición es aquel en el que se sostiene que las relaciones entre distintas variables observadas (como los puntajes de un ítem en un test) son explicadas por, o explican la varianza de, una o más variables latentes (constructos teóricos) en conjunción con sus correspondientes términos de error. El segundo tipo de modelo general es el que se conoce como “modelo estructural”, en el cual se plantea la relación entre una o más variables latentes. Esta distinción, sin embargo, ya no suele plantearse en la práctica, pues en ambos casos siempre se define la existencia hipotética de una estructura de relaciones entre variables, cuya validez o ajuste debe estimarse por la vía de cuatro procedimientos concatenados, a saber: *la especificación, la identificación, la estimación de parámetros y la evaluación de modelos.*

ESPECIFICACIÓN DE MODELOS

La especificación consiste en establecer cuántas variables (observadas, latentes y de error) son incluidas en un modelo, cuáles son consideradas endógenas (variables dependientes) y cuáles exógenas (variables independientes), así como también el tipo de relación que las vincula (unidireccional o “causal” vs. bidireccional o “no causal”). La especificación se sustenta sobre la base de un criterio racional o marco de referencia teórico que permita sostener *a priori* las relaciones propuestas entre variables. Tales relaciones se asumen como hipótesis específicas a un problema de investigación y, entre otros aspectos, tienen especial vínculo con el número de parámetros a estimar dentro del modelo y los grados de libertad con los cuales se asumirá su contrastación y evaluación (Bentler y Dudgeon, 1996). Existen dos formas de realizar la especificación. La primera consiste en usar diagramas de rutas que sintetizan en forma gráfica las relaciones entre variables. La segunda emplea un sistema de ecuaciones lineales que las representa. La nomenclatura del sistema de ecuaciones se incluye en la nota 1 del apéndice, para utilizarlo en conjunción con varios ejemplos que ilustran dicho procedimiento. La especificación por la vía de diagramas de ruta sigue las convenciones estándares empleadas en análisis de rutas basados en modelos de regresión múltiple (e.g., Asher, 1981; Kerlinger y Pedhazur, 1973; Werts y Linn, 1970) y en la tabla 1 se sintetizan estas convenciones.

Tabla 1. Convenciones de representación gráfica para la especificación de modelos

Representación Gráfica	Significación
	Variable latente o no directamente observada (e.g., factores, errores).
	Variable observada o directamente medida.
	Ruta unidireccional que sugiere la influencia “causal” de una variable sobre otra, expresada en términos de pesos de regresión.
	Ruta bidireccional que sugiere la relación “no causal” entre variables, generalmente expresada en términos de covarianza o correlación.

Por conveniencia, los términos “indicadores”, “ítems” o “variables observadas” se usarán como sinónimos, al igual que los términos “variables latentes”, “factores” o “constructos”. Dentro de la perspectiva de los SEM, se plantea una distinción entre dos tipos de “modelos de medición”, los cuales están orientados a la validación de mediciones basadas en el uso de los tests. Al primero de ellos se les conoce como modelos de indicadores reflectivos; mientras que al segundo se les denomina modelos de indicadores formativos (Bollen, 1984; 1989; Bollen y Lennox, 1991). En ambos casos se plantean diversas relaciones entre variables observadas y latentes, tales como las que se presentan en los modelos reflejados en la figura 1.

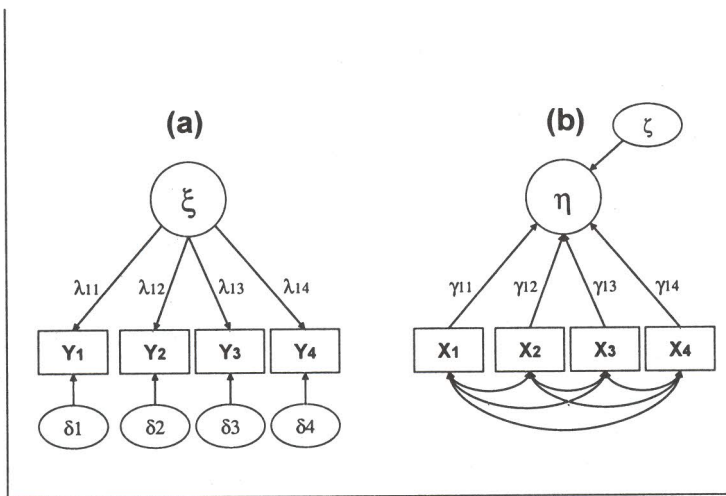


Figura 1. (a) Modelo de medición de indicadores reflectivos y (b) modelo de medición de indicadores formativos.

En el modelo de la figura 1(a) se presenta un modelo de medición de indicadores reflectivos, también conocido como modelo de análisis factorial confirmatorio (CFA). En ellos se considera que las distintas variables observadas (i.e., Y_1, Y_2, Y_3, Y_4) son dependientes de una latente (i.e., ξ) que las afecta en forma común, con una magnitud λ_{ij} que representa el coeficiente de determinación o carga factorial entre la j -ésima variable latente y la i -ésima variable observada. Adicionalmente, el modelo señala que cada variable observada posee un término de error de medición (i.e., $\delta_1, \delta_2, \delta_3, \delta_4$) que representa aquella proporción de su varianza no explicada por la variable latente. Un aspecto fundamental de este modelo es el hecho de que la varianza de la variable latente (ξ) determina a la varianza común de las variables observadas.

En la figura 1(b), en cambio, se presenta un modelo de medición con indicadores formativos. Dicho modelo es un caso especial de regresión múltiple, con la peculiaridad de que la variable dependiente resulta ser una variable latente (véase Bollen, 2002; pp. 619-624). En estos modelos, al considerar a la variable latente (η) como dependiente de distintas variables observadas (i.e., X_1, X_2, X_3, X_4), se asume que cada una de éstas la afectan de forma simultánea con una magnitud γ_{ij} que representa al coeficiente de determinación entre cada variable observada y la variable latente. Adicionalmente, el modelo señala que la variable latente posee un término de error o disturbancia (ζ_1) que representa aquella proporción de su varianza no explicada por la influencia simultánea de las variables observadas.

Aunque ambos modelos de la figura 1 introducen cuatro variables observadas y una variable latente, la diferencia fundamental entre ellos consiste en su especificación, pues mientras en el modelo 1(a) se asume que la variable latente es la que afecta a las variables observadas, en el modelo 1(b) se asume lo contrario; es decir, las variables observadas afectan a la variable latente. Desde un punto de vista gráfico puede observarse la diferencia por la dirección de las rutas entre variables observadas y latentes. Sin embargo, otras consideraciones de mayor importancia y trascendencia para la evaluación de modelos de medición pueden señalarse, convenientemente, más adelante.

En cualquier tipo de modelo especificado, cada ruta y también cada variable latente representa una hipótesis específica, la cual está en función de las relaciones entre las variables observadas. Además, cada hipótesis se define como un parámetro dentro del modelo, de modo que con la especificación se plantea la relación de cada hipótesis específica con uno o más parámetros del modelo propuesto; siendo éstos usualmente expresados como coeficientes de varianzas y covarianzas implicadas en matrices paramétricas. De allí que todas las hipótesis planteadas se contrasten de manera global mediante la siguiente hipótesis general:

$$\Sigma = \Sigma(\theta) \quad (\text{Ecuación 1})$$

Dicha hipótesis, plantea la igualdad entre las matrices S y $\Sigma(\theta)$. La matriz que se encuentra al lado izquierdo de la igualdad representa a una matriz de varianza-covarianza que expresa las relaciones entre "q" variables observadas de una población. La matriz $\Sigma(\theta)$ es una que contiene un conjunto θ de parámetros incluidos en la especificación de un modelo de medición propuesto. A dicho modelo se le especifica con el propósito de observar en qué medida puede juzgarse como una representación ajustada a las relaciones de las "q" variables implicadas en la población (Bentler y Bonett, 1980).

Dado que los parámetros incluidos en la especificación dependen del tipo de modelo de medición representado en términos de indicadores reflectivos o indicadores

formativos (Bollen, 1984), conviene hacer alusión al conjunto de parámetros que se incluyen en cada uno de estos modelos. En general, las matrices paramétricas incluidas en la especificación de un modelo de medición con indicadores reflectivos son:

$$\Sigma(\theta) = \Lambda \Phi \Lambda' + \Theta \delta \quad (\text{Ecuación 2})$$

Donde ΛX representa a la matriz de estructura factorial que expresa el grado de asociación entre las variables observadas y las variables latentes; $\Lambda X'$ representa la matriz transpuesta de la estructura factorial, que resulta de invertir las filas por columnas que contienen a los coeficientes de dicha matriz. Por otro lado, Φ representa a la matriz de varianza-covarianza de las distintas variables latentes que explican la varianza común de las variables observadas. Finalmente, $\Theta \delta$ representa a la matriz de varianza-covarianza de los términos de error de medición para las variables observadas.

En contraste, las matrices paramétricas incluidas en la especificación de un modelo de medición con indicadores formativos son:

$$\Sigma(\theta) = \Gamma X + \zeta \quad (\text{Ecuación 3})$$

Donde ΓX representa a la estructura factorial que expresa el grado de asociación entre la variable latente y sus múltiples indicadores o variables observadas, mientras que ζ representa al término de error o disturbancia de la variable latente η que aparece en el modelo de la figura 1(b). Una vez especificados los parámetros del modelo, la hipótesis general planteada en la ecuación 1 no puede contrastarse hasta tanto no se haya identificado al modelo especificado. De allí, la discusión de este procedimiento en el próximo apartado.

IDENTIFICACIÓN DE MODELOS DE MEDICIÓN

Un primer aspecto a considerar en la identificación es el número de parámetros conocidos y desconocidos, así como también los criterios que deben seguirse para identificarlos (Bollen, 1989). En general, la identificación de un modelo consiste en asignar reglas de correspondencias numéricas que permitan identificar los valores específicos correspondientes a, por lo menos, uno de los términos incluidos en las matrices paramétricas involucradas en la especificación de un modelo (Carmines, 1990; Seidel y Eicheler, 1990). A todos aquellos parámetros que se le asignan dichas reglas, se les denomina parámetros identificados y sin ellos sería imposible la contrastación de la hipótesis general que aparece en la ecuación 1 (Bollen, 1989).

En general, hay cuatro criterios básicos para determinar cuáles son los parámetros identificados de un modelo. Raykov y Marcoulides (2000) sugieren la selección de uno o varios de los siguientes parámetros:

- Las varianzas de las variables exógenas (i.e., la varianza de los factores latentes y las varianzas de los residuales o términos de error).
- Las covarianzas o correlaciones entre las variables exógenas (i.e., la covarianza entre factores latentes o la covarianza entre residuales).
- Las cargas factoriales (coeficientes de regresión) que vinculan a una variable latente con la variable observada.
- Los coeficientes de regresión entre variables observadas o los coeficientes de regresión entre variables latentes.

Con el reconocimiento de estos cuatro criterios, puede emplearse un conjunto de reglas (como la regla-t, explicada más adelante), que aseguran la identificación de un modelo. A pesar de estas observaciones, conviene advertir las implicaciones de identificar a las variables latentes dentro del contexto de la especificación de modelos estructurales.

Una primera razón para identificar a las variables latentes consiste en que ello permite distinguir entre modelos con sólo variables observadas y modelos con variables observadas y variables latentes, a los que se les ha denominado modelos de ruta con variables latentes (McDonald, 1996). Usualmente, en los modelos que incluyen sólo variables observadas se asume que la variable latente es igual a su indicador o variable observada (i.e., $\xi = X$) y, a partir de esto, se estima la relación entre ellas por la vía de una regresión múltiple (e.g., Wright, 1960a, 1960b). Sin embargo, un aspecto implícito en el procedimiento anterior es la inexistencia de errores de medición para la variable latente (Bollen, 1989), lo cual conduce a un estimado sesgado de relación entre variables (McDonald, 1996). Lo anterior puede ilustrarse a partir de los diagramas que aparecen en la figura 2, en donde se muestra la diferencia entre aquellos modelos que sólo asumen variables observadas y aquellos que asumen tanto variables observadas como variables latentes. En la nota 5 del apéndice se discuten estas implicaciones.

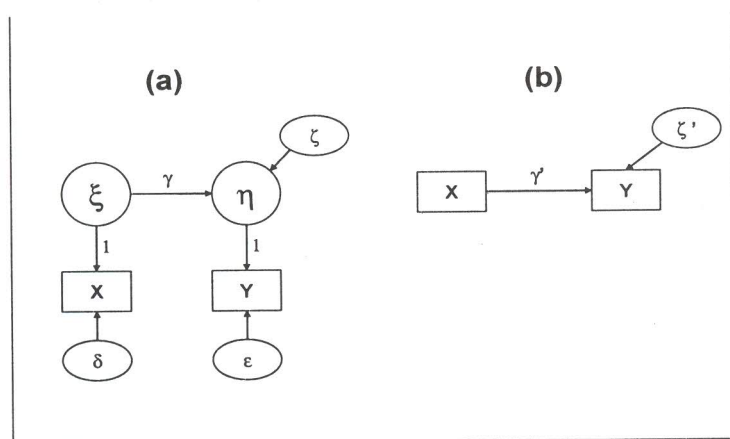


Figura 2. (a) Modelo con variables observadas y latentes y (b) modelo con sólo variables observadas

En el modelo de la figura 2(a) se estima la relación de influencia entre dos variables latentes (i.e., ξ y η), que poseen un indicador o variable observada (X y Y) con su correspondiente término de error asociado (δ y ϵ). Estos términos no están ni asociados entre sí, ni con las variables latentes (η y ξ) ni con el término de error o disturbancia de la variable endógena (ζ). Además, este último error (ζ) se ha especificado de forma tal que no esté asociado con la variable latente exógena ξ . Por un lado, en el modelo de la figura 2(b) la relación se estima directamente entre las variables observadas, por la vía de una regresión simple. En términos generales, ha sido documentado que la influencia γ' de X sobre Y en el modelo de la figura 2(b) no iguala a la influencia γ de ξ sobre η en el modelo 2(a) (Bollen, 1989). Por otro lado, en el modelo de la figura 2(b), el término de error (ζ') de la variable endógena observada, tampoco iguala al término de error (ζ) de la variable endógena latente del modelo 2(a). La principal diferencia entre estos dos modelos surge a partir de la consideración del rol que tiene el error de medición de las variables latentes en la estimación de sus relaciones con otras variables.

Para asignar una escala al factor latente, se puede fijar la varianza de éste a una constante (usualmente $\phi_{11} = 1$) o de forma equivalente, se puede fijar a una constante el primer coeficiente de regresión de la variable observada al factor latente (usualmente $\lambda_{11} = 1$). Cualquiera de estas convenciones representa la relación existente entre la variable observada y el factor latente, a saber: el cambio esperado en una unidad de ξ_1 es igual, en promedio, al cambio de una unidad γ en X_1 . (Bollen, 1989; p. 240). Esta relación sugiere el carácter fundamentalmente hipotético de la medición de variables latentes, las cuales no poseen una métrica propia y, por lo tanto, requieren el establecimiento de

las reglas por medio de las cuales se les asigna una métrica (en puntajes y distribución) (Lord y Novick, 1968). De allí, el rol fundamental de la identificación en la evaluación de modelos de medición.

Dado que los términos de disturbancia o errores para las variables latentes también son considerados como variables latentes (Arbuckle y Wothke, 2003a, 2003b; Bollen, 2002), una convención similar a la aplicada para el factor latente se aplica a estos términos. La convención más estándar es fijar la varianza de cada término de error a una constante (usualmente, $\delta_1 = \delta_2 = \delta_3 = 1$); aunque equivalentemente se pueda fijar la escala de cada término de disturbancia haciendo que sus coeficientes de regresión sobre la variable observada sea igual a la unidad.

De hecho, esta última convención es la más frecuente utilizada, pues su implicación consiste en asumir que los errores de medición siguen una distribución normal, constituida por valores tipificados en torno a su esperanza matemática o promedio, que en este tipo de distribución es la característica más representativa de todo el conjunto de valores incluidos en el rango de la distribución, la cual, al estar expresada en términos estándares, se asocia con la influencia mínima que tienen estos valores sobre la variable observada. Esto último es uno de los supuestos en los que se basa la función de estimación de mínimos cuadrados ordinarios que se emplea en la regresión lineal múltiple, técnica base para el desarrollo de modelos multivariantes (Cohen, 1968) como los que sustentan la validación de estructuras factoriales especificadas en modelos de medición (Bollen, 1989).

Con las dos convenciones señaladas anteriormente, se puede distinguir entre dos tipos de parámetros dentro de los modelos de medición. Al primero se le denomina parámetro restringido, que puede asumírsele como un parámetro identificado porque se le ha asignado algún valor específico. Un segundo tipo de parámetro es el denominado parámetro libre, que representa la característica distribucional de una variable cuyo valor debe estimarse. Para contrastar la hipótesis general, debe conocerse cuántos parámetros libres están propuestos en el modelo. Así, por ejemplo, mientras en el modelo de la figura 1(a) el total de parámetros libres es ocho, luego de haberse identificado un total de cinco parámetros que representan las varianzas de la variable latente y de los términos de error ($\phi_{11} = \delta_1 = \delta_2 = \delta_3 = \delta_4 = 1$); en el modelo de la figura 1(b) el total de parámetros libres es nueve, luego de haberse identificado un total de siete parámetros que representan las covarianzas entre las variables observadas y la varianza del término de disturbancia de la variable latente.

La determinación del número de parámetros libres de un modelo es útil para calcular sus grados de libertad, los cuales son el punto de referencia estadístico para concluir si existe o no una distancia (diferencia) significativa entre el modelo propuesto

(en términos de la matriz $\Sigma(\theta)$) y los datos observados (en términos de la matriz Σ) (ver nota 4 del apéndice). Así, los grados de libertad de un modelo se definen como la diferencia entre el número de términos no redundantes de la matriz de covarianza poblacional (ver nota 3 del apéndice) y el número de parámetros libres a estimar en el modelo. Formalmente, esto puede representarse como:

$$gl = \frac{1}{2} p(p+1) - t \quad (\text{Ecuación 4})$$

Donde p indica el número de variables observadas y t el número de parámetros libres a estimar en un modelo. Siguiendo como ejemplo los modelos de la figura 1, para el caso del modelo de indicadores reflectivos se tiene un total de 10 elementos no redundantes y un un total de ocho parámetros libres por estimar, mientras que para el caso del modelo con indicadores formativos se tiene un total de 10 elementos no redundantes y total de nueve parámetros libres por estimar. Así, el modelo de indicadores reflectivos tiene dos grados de libertad ($gl = 2$) y el de indicadores formativos tiene un grado de libertad, por lo cual pueden ser contrastados.

Por otro lado, también es común utilizar algunos modelos con cero grados de libertad, a los que se les denominan “modelos saturados”, empleados para efectuar una regla de identificación: la denominada regla- t (Bollen, 1989). De acuerdo con ésta, se establece que el término t de la ecuación 4 debe ser menor o igual a la cantidad de elementos no redundantes de la matriz de covarianza. Si t es igual a la cantidad de términos no redundantes de la matriz, entonces se obtendrá un “modelo saturado” que no podrá, en sí mismo, contrastarse estadísticamente, pero servirá de base para la comparación con otros modelos identificados, algo que tiene especial relevancia para la evaluación de modelos de medición.

ESTIMACIÓN, FUNCIONES DE AJUSTE Y CONTRASTACIÓN DE MODELOS

Como se ha mencionado, todas las hipótesis de un modelo se contrastan de manera global mediante la hipótesis general que aparece en la ecuación 1. Dado que los valores paramétricos pertenecientes a la matriz Σ usualmente se desconocen (Bollen, 1989), se hace necesaria su estimación a partir de una matriz, S , que contiene los valores estadísticos de una muestra de sujetos. Así, la matriz S es empleada en distintas funciones de estimación para obtener una matriz estimada, $\hat{\Sigma}$, que contiene a todos los parámetros estimados (Bentler y Bonnet, 1980). Las ecuaciones 2 y 3 presentadas para los modelos de la figura 1 se refieren a matrices paramétricas. Pero para contrastar las hipótesis de estos modelos en circunstancias reales, la hipótesis general se basa en la comparación de dos matrices estimadas. En otras palabras, se basa en la siguiente hipótesis: $\hat{\Sigma} = \Sigma(\hat{\theta})$. Por tal razón, es conveniente presentar una versión de las ecuaciones anteriores con la siguiente forma:

$$\Sigma(\hat{\theta}) = \hat{\Lambda} x \hat{\Phi} \hat{\Lambda} x' + \hat{\Theta} \delta \quad (\text{Ecuación 5})$$

$$\Sigma(\hat{\theta}) = \hat{\Gamma} \xi + \hat{\zeta} \quad (\text{Ecuación 6})$$

En estas ecuaciones el símbolo $\hat{}$ indica los parámetros que son estimados a partir de la matriz S. En general, la estimación de un modelo supone un contraste de la forma $(\hat{\Sigma} - S)$, que representa una matriz cuyos elementos surgen a partir de la diferencia entre los valores estimados y los observados. De allí que a la matriz $(\hat{\Sigma} - S)$ se le denomine matriz residual, la cual es empleada para contrastar el modelo especificado de dos maneras. La primera es por la vía que implica el contraste global del modelo, mientras que la segunda consiste en analizar los elementos de la matriz residual (ver nota 7 del apéndice) (Bollen, 1989).

El contraste global resume, a través de una medida “omnibus”, la significancia estadística de todas las hipótesis (rutas y variables latentes) especificadas en un modelo. A dicha medida se le reconoce como un valor “ ν ” (ny) con la propiedad de ser un estadístico con distribución χ^2 y gl grados de libertad obtenidos por la ecuación 4 anteriormente presentada.

La lógica del contraste de hipótesis global en los modelos de medición difiere de la que se sigue en el contraste de hipótesis estándar (Marsh y Balla, 1994), pues en este tipo de contraste (como en el caso del ANOVA o la regresión lineal múltiple), normalmente se desea rechazar la hipótesis nula (H_0) para tomar este resultado como indicador favorable de las hipótesis planteadas. En cambio, en los modelos de medición, normalmente se desea aceptar la H_0 (ecuación 1), pues ello indicaría que el modelo especificado “se ajusta” a las relaciones observadas entre las variables, lo cual permite argumentar que no hay diferencias significativas entre las relaciones observadas y las propuestas en la especificación del modelo.

El análisis de la matriz residual consiste en observar (parámetro por parámetro) la distancia o diferencia entre los valores esperados y los parámetros estimados. Así, cuando $\nu > gl$ tal que $p < \alpha \leq .05$, el investigador concluye que las relaciones observadas no pueden considerarse como representativas de las relaciones (poblacionales) propuestas por el modelo, mientras que “si el estadístico es pequeño comparado con sus grados de libertad [y el nivel de significancia adoptado (e.g., $\alpha > .05$)], uno concluye que el modelo provee una adecuada representación del sistema de influencias entre las variables” (Bentler y Bonett, 1980; p. 591).

Desafortunadamente, el valor que obtiene ν es dependiente del tamaño muestral y la función de ajuste empleada para la estimación de parámetros, lo cual ha dado pie al planteamiento de otras formas de evaluar las hipótesis planteadas en un modelo de

medición (Bentler y Bonett, 1980; Marsh, Balla y McDonald, 1988). La dependencia del estadístico ν con el tamaño muestral y la función de estimación puede formalizarse de la siguiente manera:

$$\nu = nF(\hat{\theta}) \quad (\text{Ecuación 7})$$

Donde n es el número de sujetos o tamaño de la muestra y $F(\hat{\theta})$ es una función de ajuste empleada para la estimación de los parámetros de un modelo. Como puede observarse, a partir de la ecuación 7 surgen dos consecuencias obvias que afectan el contraste de hipótesis en los modelos de medición.

La primera consecuencia surge de la relación directa entre el estadístico ν y el tamaño muestral. "La probabilidad de rechazar cualquier modelo aumenta a medida que [el tamaño de] la muestra aumenta" (Bentler y Bonett, 1980; p. 591), aun cuando en la matriz residual, la magnitud de la diferencia entre los valores estimados y los observados sea despreciable (Cudeck y Henly, 1991).

La segunda consecuencia es algo más elaborada que la anterior y tiene que ver con las propiedades de las distintas funciones de ajuste empleadas para la estimación de los valores paramétricos contenidos en la especificación de un modelo (ver nota 6 del apéndice). Dichas funciones están amparadas bajo la teoría de distribución asintótica (Bollen, 1989; pp. 466-470), según la cual el valor de un estimado paramétrico se acerca más al valor real de un parámetro poblacional en la medida que el tamaño de la muestra se acerque al tamaño de la población, siempre y cuando se cumplan tres supuestos, a saber: la consistencia, la eficiencia y la distribución subyacente a la estimación de los parámetros.

La consistencia de un parámetro estimado se refiere a la propiedad de que su valor converja con el valor cierto (preestablecido) a medida que la muestra a partir de la cual se estimó, tienda hacia el infinito (ver nota 6 del apéndice); en otras palabras, la probabilidad de la convergencia entre el valor estimado y el parámetro preestablecido (usualmente el promedio de la variable dependiente) alcanza el valor de uno a medida que el tamaño muestral se hace más grande. La eficiencia de un parámetro se refiere a la propiedad de la exactitud con la cual es estimado. Debe recordarse que toda estimación tiene asociada un término de error de estimación que se refiere a la varianza de la distribución de los posibles valores dentro de los cuales está contenido el valor del parámetro.

En general, se dice que un parámetro es eficiente si su error de estimación es mínimo (Bollen, 1989). Finalmente, el supuesto de distribución hace referencia a la necesidad de asumir que la convergencia y eficiencia de los valores paramétricos se

ampanan en las propiedades de máxima representación estadística para aquellas características descriptivas de una distribución; por ejemplo, el uso de los dos primeros momentos muestrales (promedio y varianza) para representar adecuadamente una distribución de tipo normal.

Dadas las limitaciones del χ^2 frente al tamaño muestral y la función de estimación, se ha recomendado el uso de diversos criterios de ajuste para juzgar la plausibilidad de las relaciones contenidas en un modelo propuesto (e.g., Bentler, 1990; Bentler y Bonett, 1980; Bentler y Mooijart, 1989; Bollen, 1986, 1989, 1990a, 1990b; Cudeck y Henly, 1991; Curran, West y Finch, 1996; MacCallum y Austin, 2000; McDonald y Marsh, 1990; Mulaik, et. al., 1989). La presentación de estos criterios se discute en el próximo apartado.

MODELOS DE MEDICIÓN Y SU EVALUACIÓN A TRAVÉS DE CRITERIOS O ÍNDICES DE AJUSTE

El uso de criterios o índices de ajuste para juzgar la plausibilidad de modelos estructurales ha sido propuesto desde hace décadas (Bentler y Bonett, 1980), con base en la clasificación general de “índices de ajuste absoluto”, “índices de ajuste incremental” e “índices de parsimonia”, expresados en su mayoría en forma estandarizada dentro del rango 0 – 1, de forma análoga a como se expresan los coeficientes de regresión (Mulaik, et. al., 1989). En la tabla 2 se presentan algunos de los índices más frecuentemente empleados.

Con el criterio de ajuste absoluto se estima “cuan bien se reproducen los datos de una muestra con la especificación *a priori* de un modelo” (Hu y Bentler, 1999; p. 2), lo que en otras palabras supone la medida en la cual los datos de una muestra pueden considerarse como “ajustados” a la especificación del modelo propuesto.

Con el criterio de ajuste incremental se obtiene una estimación del ajuste de un modelo propuesto con base en su comparación con otros “modelos anidados” (Bentler y Bonnet, 1980; p. 592), entendiéndose por modelo anidado un modelo que representa, en términos estadísticos, una hipótesis alternativa al modelo propuesto. Dentro de los modelos anidados se encuentran los modelos saturados, los nulos y los intermediarios.

Tabla 2. Criterios de evaluación e índices de ajuste para modelos estructurales

Criterios de evaluación	Significación	Índices pertenecientes
Ajuste absoluto	Señala la medida en que los datos de una muestra definida pueden considerarse como "ajustados" a la especificación de un modelo teórico propuesto.	<ul style="list-style-type: none"> • χ^2/gl, (No se expresa en rango 0 – 1) • Índice de bondad de ajuste (GFI). • Índice de parámetro no centralizado (NCP). • Índice estandarizado de parámetro no centralizado (SNCP). • Índice estandarizado de residuo cuadrático medio (SRMR). • Índice de error de aproximación cuadrático medio (RMSEA). • Índice de validación cruzada esperada (ECVI). • Índice de validación cruzada (CVI).
Ajuste incremental	Evalúan el incremento obtenido en el ajuste de un modelo al compararse con otros modelos similares teóricamente.	<ul style="list-style-type: none"> • Índice de ajuste normalizado (NFI). • Índice Tucker – Lewis o de ajuste no normalizado (TLI; NNFI) (Puede caer fuera de los límites del rango 0 – 1). • Índice de ajuste incremental (IFI). • Índice RNI, CFI, RFI.
Parsimonia	Evalúan el ajuste de un modelo al compararse con otros similares, tomando en consideración la cantidad de hipótesis especificadas y el total de grados de libertad.	<ul style="list-style-type: none"> • Índice ajustado de bondad de ajuste (AGFI). • Índice de ajuste normalizado de parsimonia (PNFI). • Índice de calidad de ajuste de parsimonia (PGFI). • Criterio de información de Akaike (AIC).

El modelo saturado (M_S), al ser uno con cero grados de libertad que siempre conduce a un "ajuste perfecto", se considera como una hipótesis alternativa que no es informativa en sí misma, pues "no hay un simple modelo que la represente [gráfica o conceptualmente] en el sentido de la H_0 " (Bollen, 1989; p. 265). La comparación entre el M_S y el modelo teórico propuesto (M_T) se establece en términos de una diferencia entre sus valores χ^2 . De tal modo, si dicha diferencia no es significativa, ello implicaría que el M_T puede considerarse como una representación plausible de las relaciones que existen entre las variables, dando paso a su aceptación provisional (Bentler y Bonett, 1980).

El denominado “modelo nulo” (M_0) supone la independencia (no relación) de las variables, dado que todos sus parámetros (hipótesis) se fijan a cero. El modelo nulo también se toma como una H_1 con la cual se contrasta al M_T . De tal manera, si la diferencia entre los valores χ^2 no es significativa cuando se comparan el M_0 y el M_T , ello implicaría que los parámetros estimados bajo la especificación M_T no mejoran el ajuste global cuando se compara con un modelo poco informativo como el nulo, de modo que M_T debe ser rechazado.

El modelo intermediario (M_K) recibe este nombre porque representa un modelo que se encuentra a medio camino entre el modelo nulo, que supone la independencia entre variables, y el modelo teórico que especifica diversas relaciones entre éstas. De tal manera, cuando la diferencia entre M_K y M_T es significativa, ello sería indicador de que el modelo teórico propuesto mejora o aumenta el ajuste global, de modo que puede aceptarse como una representación plausible de las relaciones entre variables.

Finalmente, otro criterio para la evaluación de modelos se basa en la idea de parsimonia (ver nota 8 del apéndice); razón por la cual los índices basados en esta alternativa llevan su nombre (Bentler y Mooijart, 1989; Mulaik, 1998; Mulaik, et. al., 1989; Raykov y Marcoulides, 1999). La lógica que sustenta este criterio también se basa en la comparación de modelos, pero tomando en consideración a los grados de libertad resultantes de la especificación de cada modelo a comparar (Arbuckle y Wothke, 2003a; Mulaik, et. al., 1989; Mulaik, 1998, 2001). De tal manera, si dos modelos se ajustan igualmente bien a las relaciones observadas entre las variables, se asume usualmente que el modelo más sencillo es probablemente el “cierto” y, en consecuencia, el más fácil de replicarse (contrastarse) en otras muestras pertenecientes a la misma población definida por el investigador (Mulaik, et. al., 1989).

La evaluación de modelos con base en los criterios mencionados, usualmente se sustenta en la observación del valor que alcanzan estos índices luego de haberseles comparado con ciertos puntos de corte (e.g., $CFI \geq 0.90$; $RNI \geq 0.90$; $RMSEA \leq 0.05$) sugeridos para apoyar la decisión de validar los modelos contrastados. Sin embargo, esta forma de evaluación no ha sido tan adecuada, pues conlleva a argumentar las razones que determinan la escogencia de un determinado “punto de corte” para un índice de ajuste prescrito (Boosma, 2000; Hu y Bentler, 1999), dada la sensibilidad de éstos frente a condiciones como la disposición de distribuciones no normales, en pequeñas muestras, con datos perdidos o distintas funciones de estimación y errores de especificación (Curran, West y Finch, 1996; Fan, Thompson y Wang, 1999; Green, Thompson y Poirier, 1999; Hu, Bentler y Kano, 1992; Jackson, 2001; Kaplan, 1988, 1990; La Du y Tanaka, 1989; MacCallum, Roznowski y Necowitz, 1992; Marsh y Balla, 1994; Marsh y Hau, 1996; Muthén, Kaplan y Hollis, 1987; Robles, 1992; Tanaka, 1987).

Como consecuencia de lo anterior, cada vez más se ha reconocido la necesidad de utilizar criterios alternativos de evaluación que, sustentados en argumentos de tipo estadístico, permitan comprender la racionalidad subyacente a la decisión de validar un modelo, que en última instancia representa la información de mayor interés para los propósitos de sustentar su validez o ajuste a la luz de las aseveraciones anticipadas por una teoría. Por ello, en el próximo apartado se discute el rol de criterios alternativos de evaluación sobre la decisión nominal de aceptar como válido o no un modelo estructural.

CONDICIONES PARA LA VALIDACIÓN DE MODELOS: CRITERIOS ALTERNATIVOS DE DECISIÓN

Un aspecto final asociado con la evaluación de un modelo, es el referido a la decisión de aceptarlo como válido o no para una muestra de sujetos pertenecientes a una población teóricamente definida (Bollen, 2002). En estas condiciones, la decisión de validar al modelo, según una estructura teórica propuesta, conlleva el riesgo asociado de un error de decisión de tipo estadístico. En tal sentido, el uso de los SEM mantiene el mismo problema asociado con el uso de modelos estadísticos tradicionales como el ANOVA o el análisis de regresión; en el sentido de cometerse un error estadístico tipo I (rechazar la hipótesis nula cuando es cierta) o un error tipo II (aceptar la hipótesis nula cuando es falsa).

Dada la especial atención que requiere la selección de los índices de ajuste para propósitos de evaluación de modelos, Hu y Bentler (1999) han ofrecido una estrategia alternativa con la cual se busca superar los inconvenientes asociados con el uso individual de los puntos de corte para estos índices. La racionalidad ofrecida por estos autores descansa en el simple razonamiento de que no importa cuál índice de ajuste sea el escogido para la evaluación de un modelo, siempre debe tomarse “un adecuado criterio de punto de corte que resultaría en proporciones mínimas de error tipo I y de error tipo II” (p. 5).

Este criterio, denominado “estrategia de evaluación de doble índice”, consiste en utilizar el índice estandarizado de residuo cuadrático medio (SRMR) bajo la función de estimación de máxima verosimilitud y acompañarlo con otros índices como el Tucker-Lewis (TLI), el índice de ajuste comparado (CFI) o el índice de error de aproximación cuadrático medio (RMSEA). De acuerdo con esta estrategia, se obtiene un mecanismo que en principio es más exigente para que la toma de decisión no resulte inadecuada. En otras palabras, se mantienen el uso de los puntos de corte para los índices de ajuste, con la combinación de doble índice mencionada, de modo que la evaluación de un modelo supone la satisfacción conjunta de los valores definidos como puntos de corte para cualquiera de las parejas de índices referidas.

El uso de la metodología propuesta desde la perspectiva de los SEM exige una alta racionalidad que sustente la especificación de un modelo en particular. En la práctica se hace frecuente encontrar, a pesar del uso de la estrategia de doble índice, que los resultados referidos al ajuste de un modelo no hacen sentido a la luz de los procedimientos efectuados en el marco de una investigación y las evidencias empíricas previas.

En estos casos, muchos investigadores recurren a la re-especificación del modelo (añadiendo u omitiendo rutas o variables) para luego encontrar resultados más satisfactorios (Bollen, 1990a). Esta práctica es recomendable siempre y cuando la re-especificación del modelo siga manteniendo sentido con el marco de referencia teórico tomado en consideración. Sin embargo, debe comprenderse el propósito general dentro del cual se inscribe el uso de esta metodología, pues ella puede estar orientada no sólo a fines confirmatorios, sino también a fines heurísticos o exploratorios (Boosma, 2000) y en cualquier caso, no existe un único modelo sujeto a contrastación que se ajuste razonablemente bien a los postulados teóricos que se desean investigar (Bollen, 2000).

APÉNDICE

Nota 1: la notación formal adoptada en los SEM presentada a continuación considera la distinción clásica de “modelos estructurales” y “modelos de medición”.

Tabla A1. Sistema de notación para modelos estructurales

Símbolo	Dimensión de la matriz	Representa
VARIABLES		
η (eta)	$m \times 1$	Variables latentes dependientes (endógenas)
ξ (xi)	$n \times 1$	Variables latentes independientes (exógenas)
ζ (zeta)	$m \times 1$	Errores latentes en la ecuación
COEFICIENTES		
B (beta)	$m \times m$	Coficiente de la matriz para las variables latentes endógenas. Los coeficientes de esta matriz se denotan en minúsculas; e.g., β_{11}
Γ (gamma)	$m \times n$	Coficiente de la matriz para las variables latentes exógenas. Los coeficientes de esta matriz se denotan en minúsculas; e.g., γ_{11}
MATRICES DE COVARIANZAS		
Φ (phi)	$n \times n$	Matriz de covarianza para variables latentes exógenas. Los coeficientes de esta matriz se denotan en minúsculas; e.g., ϕ_{11}
Ψ (psi)	$m \times m$	Matriz de covarianza para los errores latentes en las ecuaciones. Los coeficientes de esta matriz se denotan en minúsculas; e.g., ψ_{11}

Tabla A2. Sistema de notación para modelos de medición o CFA

Símbolo	Dimensión de la matriz	Representa
VARIABLES		
Y	$p \times 1$	Variable observada (indicador) de la variable latente endógena η
X	$q \times 1$	Variable observada (indicador) de la variable latente exógena ξ
ϵ (epsilon)	$p \times 1$	Errores de medida para Y
δ (delta)	$q \times 1$	Errores de medida para X
COEFICIENTES		
Λ_y (Lambda y)	$p \times m$	Coefficiente que relaciona Y con η . Los coeficientes de esta matriz se denotan en minúsculas; e.g., λ_{11}
Λ_x (Lambda x)	$q \times n$	Coefficiente que relaciona X con ξ . Los coeficientes de esta matriz se denotan en minúsculas; e.g., λ_{11}
MATRICES DE COVARIANZAS		
Θ_ϵ (thetaepsilon)	$p \times p$	$E(\epsilon\epsilon')$ (matriz de covarianza de ϵ). Los coeficientes de esta matriz se denotan en minúsculas; e.g., ϵ_{11}
Θ_δ (thetadelta)	$q \times q$	$E(\delta\delta')$ (matriz de covarianza de δ). Los coeficientes de esta matriz se denotan en minúsculas; e.g., δ_{11}

Una consideración establecida por Bollen (1989; 2000; 2002) con respecto a los dos tipos de modelos mencionados es la de asumir que los modelos de medición (CFA) también son modelos estructurales. Dado que la denominación de “modelos estructurales” y “modelos de medición” podría sugerir que en los CFA no se especifica alguna estructura particular. Por esta razón, Bollen (2002) sugiere considerar a los CFA como “modelos de variables latentes”. A los modelos de ecuaciones estructurales también se les conoce bajo la denominación de análisis de estructuras de covarianzas, debido a que se toma en consideración a las covarianzas como los indicadores de relación entre variables, pues las covarianzas poseen propiedades de mayor robustez para la estimación de parámetros en comparación con sus homólogas, las correlaciones.

Nota 2: según Bollen (2002), el término “variable latente” ha cobrado al menos tres connotaciones informales o intuitivas (i.e., variables hipotéticas, variables que no pueden ser directamente medibles y variables descriptivas de la relación entre datos) y cuatro definiciones formales (independencia local, valor esperado, función no determinista de variables observadas y realización muestral). Las diferencias que surgen entre estas definiciones tienen distintas implicaciones en diferentes modelos metodológico-estadísticos (por ejemplo, teoría clásica de los tests, teoría de respuesta al ítem, análisis factorial). Por ejemplo, según la definición de independencia local, una variable latente es aquella que al ser controlada en una correlación parcial entre dos variables observadas,

elimina la correlación de éstas; mientras que según la definición de valor esperado, una variable latente es aquella considerada como el puntaje verdadero en la teoría clásica de los tests, el cual representa al promedio de todos los puntajes que obtendría un sujeto bajo el supuesto de habersele evaluado un número infinito de veces. En este artículo no se adopta explícitamente ninguna de las definiciones formales señaladas, pues “no hay una definición correcta o errada de variables latentes” (Bollen, 2002; p. 631).

Nota 3: las matrices de covarianzas entre variables observadas así como también las matrices de correlación son simétricas, lo que implica que los términos por debajo y por encima de la diagonal principal de una matriz son exactamente iguales; es decir: $Covarianza(X1, X2) = Covarianza(X2, X1)$; $rx1, x2 = rx2, x1$. Cuando se emplean matrices de covarianzas, los términos de la diagonal principal representan la covarianza de una variable consigo misma y eso es igual a la varianza de la variable; es decir: $Covarianza(X1, X1) = Varianza(X1)$. En los SEM es suficiente con utilizar los términos no redundantes de la matriz; es decir, los términos de la diagonal principal más los términos por debajo de ésta.

En los SEM se prefiere el uso de matrices de covarianzas en lugar de matrices de correlaciones debido a la naturaleza de las covarianzas, como indicadores no estandarizados de asociación, para las funciones de estimación de ajuste entre matrices como la que se emplea en el cálculo de la función de máxima verosimilitud (e.g., Jöreskog, 1978; Werts, Rock, Linn y Jöreskog, 1976; p. 1007).

Nota 4: los términos de la matriz $\Sigma(\theta)$ contienen los parámetros desconocidos que son estimados por distintas funciones de estimación (e.g., máxima verosimilitud, mínimos cuadrados generalizados, mínimos cuadrados no ponderados). Las funciones de los parámetros que se dan en esa matriz conducen a resultados numéricos que pueden o no igualar a los de la matriz Σ . La contrastación de modelos se expresa en términos de igualdad de matrices. Matemáticamente dos matrices son iguales si cada término s_{ij} (ubicado en la fila i y columna j) de una matriz Σ , es exactamente igual a cada término σ_{ij} (ubicado en la fila i y columna j) de la matriz $\Sigma(\theta)$.

Estadísticamente, la igualdad de matrices flexibiliza la lógica de igualdad matemática entre matrices, ya que para cada término σ_{ij} existen unos grados de libertad que fijan los límites dentro de los cuales puede variar el valor específico del parámetro estimado y considerarle igual al valor observado S_{ij} . La obra de Green y Carroll (1976) puede ser útil para familiarizarse con las operaciones de álgebra matricial implicadas en los modelos estructurales y en las técnicas de análisis estadístico multivariado.

Nota 5: cuando en los SEM se asume que una variable observada es un perfecto indicador de la variable latente ($X = \xi$; $Y = \eta$), se asume que no hay error de medición

para la variable latente. Este supuesto es una convención estándar en los análisis de rutas (Asher, 1981), los cuales son considerados desde la perspectiva de los SEM como modelos con sólo variables observadas (Bollen, 1989). Sin embargo, en los modelos de variables latentes, los indicadores de éstas tienen términos de error asociados, que generalmente hacen que la estimación de los parámetros de influencia γ' en análisis de rutas sean siempre inferiores a los parámetros de influencia γ en modelos de variables latentes (i.e., $\gamma' < \gamma$). De allí que los parámetros de influencia en análisis de rutas con sólo variables observadas sean estimadores sesgados (subestimados). Toda esta argumentación descansa en el supuesto de la “construcción de indicadores válidos y confiables” de las variables latentes (Asher, 1981; p. 9). Para una discusión más detallada sobre las implicaciones de asumir a las variables latentes como variables observadas puede consultarse Werts y Linn (1970) o incluso Bollen (2002).

Nota 6: las funciones de ajuste (e.g., máxima verosimilitud, mínimos cuadrados generalizados, mínimos cuadrados no ponderados, estimación de distribución asintótica libre, etc), pueden considerarse como métodos de estimación paramétricos. Tales funciones están basadas en el análisis de las matrices observadas y estimadas a través de fórmulas que expresan una relación no lineal entre los parámetros a estimar en un modelo, el número de variables endógenas y exógenas, así como también distintas propiedades que estas matrices deben tener. Por ejemplo, cuando se tiene un modelo de regresión múltiple convencional, la función de ajuste que se especifica en el modelo, es la igualdad del valor de la variable predicha Y a un valor estimado Y' , para cada valor de X y Y . De esta manera, fijar como criterio de estimación, $Y = Y'$, hace que la relación entre X y Y se ajuste a un función normal lineal, para la cual pueden existir un conjunto de valores Y' normalmente distribuidos, de entre los cuales se escoge, como valor inicial, el promedio de la distribución de puntajes Y' estimados para cada par de valores X y Y . Las funciones de ajuste poseen dos aspectos cruciales que se muestran matemáticamente por el establecimiento de relaciones entre los valores paramétricos y los muestrales. El primer aspecto es la convergencia en probabilidad y, el segundo, la convergencia en distribución.

La convergencia en probabilidad consiste, brevemente, en hacer que el valor resultante de la diferencia entre un estimado paramétrico y su valor poblacional “cierto” sea menor a cualquier valor pequeño arbitrario y previamente definido, de forma tal que se aumente la confianza del estimado paramétrico a medida que se aumente el número de observaciones sobre las cuales ha sido estimado. La convergencia en distribución consiste, brevemente, en observar que la distribución de los parámetros estimados se va ajustando a una distribución de probabilidad conocida (e.g., normal, F o χ^2), a medida que los valores paramétricos sean estimados con mayores tamaños muestrales. Tanto el aspecto de convergencia en probabilidad como el de convergencia en distribución descansan en la teoría de distribución asintótica, la cual describe el comportamiento de

variables aleatorias a medida que su tamaño muestral se aproxima al infinito (Bollen, 1989; pp. 466-470).

Otro uso corriente que se hace de los valores paramétricos estimados es la observación de soluciones impropias y soluciones no convergentes (e.g., Anderson y Gerbing, 1984). Se dice que una solución es no convergente "cuando [una función] de estimación, dentro de un conjunto de iteraciones, es incapaz de llegar a valores numéricos que alcancen un criterio prescrito" (Anderson y Gerbing, 1984; p. 156), que puede ser el criterio de convergencia en probabilidad o el de convergencia en distribución. Se dice que una solución es impropia cuando los valores de los parámetros estimados muestran propiedades no posibles en la población; por ejemplo, una correlación mayor que uno o una varianza negativa (Anderson y Gerbing, 1984). En los experimentos monte carlo, la necesidad de observar la existencia de soluciones impropias y no convergentes es necesaria cuando se manipula el tamaño muestral (Paxton, et. al. 2001). En caso de que este tipo de soluciones ocurran, se ha recomendado cambiar el valor de la semilla inicial que genera los datos de una muestra, para de esta manera tener una muestra que no conduzca a este tipo de soluciones.

Nota 7: la matriz residual, $(\hat{\Sigma} - S)$, contiene tantos elementos como la matriz muestral (S) o la matriz estimada ($\hat{\Sigma}$) y los valores que toman cada uno de estos elementos son considerados uno por uno para observar en qué medida los estimados paramétricos en la matriz convergen o coinciden con los valores observados. La matriz de valores estimados, por ejemplo, es útil para observar la ocurrencia de soluciones no convergentes e impropias, que pueden ocurrir aun cuando se obtengan indicadores de un ajuste global del modelo.

En general, todas las funciones de estimación contienen funciones algebraicas de estas dos matrices (S y $\hat{\Sigma}$) y conducen a un simple resultado numérico que representa el ajuste entre el modelo y los datos observados. Este simple resultado numérico es el valor ν definido en la ecuación 7. Por ejemplo, la función de estimación basada en mínimos cuadrados no ponderados (ULS) es:

$$U(\theta) = \frac{1}{2} \text{tr}(\hat{\Sigma} - S)^2 \quad (\text{Ecuación 8})$$

En esta función, tr es la función algebraica denominada traza, la cual consiste en sumar los elementos de la diagonal principal de la matriz residual, lo cual viene a ser un indicador similar al de la varianza total a ser explicada en un modelo de regresión múltiple (Green y Carroll, 1976). La función de mínimos cuadrados generalizados (GLS) es:

$$G(\theta) = \frac{1}{2} \text{tr}[(\hat{\Sigma} - S)W]^2 \quad (\text{Ecuación 9})$$

Donde W es una matriz de ponderación como el inverso de la matriz estimada ($\hat{\Sigma}$). El inverso de la matriz estimada se denota con $\hat{\Sigma}^{-1}$ y es una matriz que multiplicada por $\hat{\Sigma}$ conduce a una matriz de identidad I (que contiene sólo unos en la diagonal y ceros fuera de ésta); así, se tiene que $\hat{\Sigma}\hat{\Sigma}^{-1} = I$. El inverso de una matriz es similar a la lógica de la división en la aritmética, y en el caso de la función de mínimos cuadrados ponderados viene a representar la proporción de varianza explicada por el modelo.

La función de máxima verosimilitud (MV), que es la más frecuentemente utilizada para el contraste de modelos estructurales es:

$$MV(\theta) = \log |\hat{\Sigma}| + \text{tr}(S\hat{\Sigma}^{-1}) - \log |S| - p \quad (\text{Ecuación 10})$$

Donde $\log|S|$ representa la función del logaritmo del determinante (suma de productos cruzados) de la matriz observada S , p es el número de variables observadas contenidas en la matriz S y $\log|\hat{\Sigma}|$ representa el logaritmo del determinante de la matriz de valores paramétricos estimados. El determinante de una matriz es un indicador similar a la sumatoria de cuadrados intra en un análisis de varianza (Green y Carroll, 1976). La función MV supone que la distribución de las variables observadas es una de tipo normal multivariado (Bollen, 1989).

La función de estimación basada en la concepción de distribución asintótica libre (ADF) es diferente a las anteriores funciones de estimación, por cuanto no supone la distribución normal de los parámetros estimados; de allí su nombre. La función ADF es:

$$ADF(\theta) = \Sigma[s - \hat{\sigma}]' \Sigma^{-1}[s - \hat{\sigma}] \quad (\text{Ecuación 11})$$

En esta función el término $\Sigma[s - \hat{\sigma}]'$ es un número que representa la diferencia global entre cada uno de los valores contenidos en la matriz observada S y cada uno de los valores contenidos en la matriz estimada $\hat{\Sigma}$ cuando se transpone la matriz residual (intercambiando las filas por las columnas). El término $\Sigma^{-1}[s - \hat{\sigma}]$ es un número que muestra la misma diferencia entre los elementos de la matriz observada y la estimada cuando se toma en consideración el inverso de la matriz estimada. La función ADF no supone ningún tipo de distribución para las variables observadas (Curran, West y Finch, 1996). Sin embargo, una condición requerida para su uso en el contraste de modelos es que el tamaño muestral sea igual o mayor al número de elementos no redundantes en la matriz de covarianza de forma que $N \geq p(p+1)/2$ donde p es el número de variables observadas. Así por ejemplo, para un modelo con 22 variables se requiere un tamaño muestral mínimo de $N = 253 = (22 \times 23)/2$.

Nota 8: de acuerdo con Mulaik (1998) y Mulaik et. al. (1989) hay un criterio racional que permite justificar la diferencia entre índices de ajuste e índices de parsimonia, a pesar de la íntima relación entre los conceptos de estimación de parámetros, parsimonia y bondad de ajuste. Brevemente, el ajuste hace referencia al tipo de función matemática que se adopta para describir o explicar las relaciones entre dos o más variables (e.g., lineal, cuadrática, cúbica). Y en general, la parsimonia hace referencia a la escogencia de la función más simple dentro del conjunto posible de funciones que describen o explican la relación entre variables. Pero dado que las funciones lineales son el estándar de la estadística multivariada en las ciencias del comportamiento (McDonald, 1986), en el contexto de los SEM la parsimonia asume otra connotación. Tal connotación es la históricamente denominada como “navaja de Occam”, que estuvo asociada con la concepción que Kant desarrollara: “[la parsimonia] es el principio regulador que nos impone la razón para unificar la experiencia con el menor número de conceptos tanto como sea posible” (Mulaik, et. al., 1989; p. 437).

A partir de esta concepción, si a las variables latentes se les considera como conceptos teóricos, entonces en los modelos estructurales la parsimonia se considera como el criterio que permite distinguir la sencillez entre dos o más modelos causales (lineales), a partir de la distinción de aquel modelo que tenga un menor número de variables latentes (Mulaik, et. al., 1989). Otra acepción de parsimonia también es posible. En los SEM, la especificación de variables latentes supone la estimación de sus parámetros libres (o su varianza o su primer coeficiente de regresión con alguna variable observada). Técnicamente, “para dos modelos con índices de ajuste global comparables, el modelo preferido será aquel que tenga pocos parámetros libres [o en otras palabras, aquellos con] más grados de libertad” (Raykov y Marcoulides, 1999; p. 293). A partir de esta última noción es que puede apreciarse la relación entre la especificación, la parsimonia y el número de parámetros libres a estimar.

REFERENCIAS BIBLIOGRÁFICAS

- Anderson, J. C. y Gerbing, D. W. (1984). “*The effect of sampling error on convergence, improper solutions, and goodness-of-fit indices for maximum likelihood confirmatory factor analysis*”. *Psychometrika*, 49, 155 – 173.
- Arbuckle, J. y Wothke, W. (2003a). *Amos 5.0. User's guide* [Programa de computación]. Chicago: Smallwaters.
- Arbuckle, J. y Wothke, W. (2003b). *An introduction to programming with Amos* (Versión 5.0) [Programa de computación]. Chicago: Smallwaters.

- Asher, H. B. (1981). *Causal modeling*. Sage University Paper series on Quantitative Applications in the social sciences, 07-003. Beverly Hills and London: Sage Publications.
- Bentler, P. M. (1990). "Comparative fit indexes in structural models". *Psychological Bulletin*, 107, 238-246.
- Bentler, P. M. y Bonett, D. G. (1980). "Significance tests and goodness of fit in the analysis of covariance structures". *Psychological Bulletin*, 88, 588-606.
- Bentler, P. M. y Dudgeon, P. (1996). "Covariance structure analysis: Statistical practice, theory and directions". *Annual Review of Psychology*, 47, 563-592.
- Bentler, P. M. y Mooijart, A. (1989). "Choice of structural model via parsimony: A rationale based on precision". *Psychological Bulletin*, 106, 315-317.
- Bollen, K. A. (1984). "Multiple indicators: Internal consistency or no necessary relationship?". *Quality and Quantity*, 18, 377-385.
- Bollen, K. A. (1986). "Sample size and Bentler and Bonett's nonnormed fit index". *Psychometrika*, 51, 375-377.
- Bollen, K. A. (1989). *Structural equation modeling with latent variables*. New York: Wiley.
- Bollen, K. A. (1990a). "A comment on model evaluation and modification". *Multivariate Behavioral Research*, 25, 181-185.
- Bollen, K. A. (1990b). "Overall fit in covariance structure models: Two types of sample size effects". *Psychological Bulletin*, 107, 256-259.
- Bollen, K. A. (2000). "Modeling strategies: In search of the holy grail". *Structural Equation Modeling*, 7, 74-81.
- Bollen, K. A. (2002). "Latent variables in psychology and the social sciences". *Annual Review of Psychology*, 53, 605-634.
- Bollen, K. A. y Lennox, R. (1991). "Conventional wisdom on measurement: A structural equation perspective". *Psychological Bulletin*, 110, 305-314.

- Boosma, A. (2000). "Reporting analyses of covariance structures". *Structural Equation Modeling*, 7, 461-483.
- Breckler, S. J. (1990). "Applications of covariance structure modeling in psychology: Cause of concern?". *Psychological Bulletin*, 107, 260-273.
- Carmines, E. G. (1990). "The statistical analysis of overidentified linear recursive models". *Quality and Quantity*, 24, 65-85.
- Cohen, J. (1968). "Multiple regression as a general data-analytic system". *Psychological Bulletin*, 70, 426-443.
- Cudeck, R. y Henly, S. J. (1991). "Model selection in covariance structure analysis and the 'problem' of sample size: A clarification". *Psychological Bulletin*, 109, 512-519.
- Curran, P. J., West, S. G. y Finch, J. F. (1996). "The robustness of test statistics to nonnormality and specification error in confirmatory factor analysis". *Psychological Methods*, 1, 16-29.
- Fan, X., Thompson, B. y Wang, L. (1999). "Effects of sample size, estimation methods, and model specification on structural equation modeling fit indexes". *Structural Equation Modeling*, 6, 56-83.
- Fisher, R. A. (1936). *Statistical methods for research workers*. 11° Ed. New York: Hafner.
- Green, P. E. y Carroll, J. D. (1976). *Mathematical tools for applied multivariate analysis*. New York: Academic Press.
- Green, S. B., Thompson, M. S. y Poirier, J. (1999). "Exploratory analyses to improve model fit: Errors due to misspecification and a strategy to reduce their occurrence". *Structural Equation Modeling*, 6, 113-126.
- Hu, L., & Bentler, P. M. (1999). "Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives". *Structural Equation Modeling*, 6, 1-55.
- Hu, L., Bentler, P. M. y Kano, Y. (1992). "Can test statistics in covariance structure analysis be trusted?" *Psychological Bulletin*, 112, 351-362.

- Jackson, D. L. (2001). "Sample size and number of parameter estimates in maximum likelihood confirmatory factor analysis: A monte carlo investigation". *Structural Equation Modeling*, 8, 205-223.
- Jöreskog, K. G. (1978). "Structural analysis of covariance and correlation matrices". *Psychometrika*, 43, 443-477.
- Kaplan, D. (1988). "The impact of specification error on the estimation testing, and improvement of structural equation models". *Multivariate Behavioral Research*, 23, 69-86.
- Kaplan, D. (1990). "Evaluating and modifying covariance structures models: A review and recommendation". *Multivariate Behavioral Research*, 25, 137-155.
- Kerlinger, F. N. y Pedhazur, E. (1973). *Multiple regression in behavioral research*. Holt, Rinehart and Winston, Inc.: New York.
- La Du, T. J. y Tanaka, J. S. (1989). "The influence of sample size, estimation method, and model specification on goodness-of-fit assessment in structural equation models". *Journal of Applied Psychology*, 74, 625-635.
- Lord, F. M. y Novick, M. R. (1968). *Statistical theories of mental tests scores*. Reading, MA: Addison-Wesley.
- MacCallum, R. C. y Austin, J. T. (2000). "Applications of structural equation modeling in psychological research". *Annual Review of Psychology*, 51, 201-226.
- MacCallum, R. C., Roznowski, M. y Necowitz, L. B. (1992). "Model modifications in covariance structure analysis: The problem of capitalization on chance". *Psychological Bulletin*, 111, 490-504.
- Marsh, H. W. y Balla, J. R. (1994). "Goodness-of-fit indices in confirmatory factor analysis: The effect of sample size and model complexity". *Quality and Quantity*, 28, 185-217.
- Marsh, H. W. y Hau, K. T. (1996). "Assessing goodness of fit: Is parsimony always desirable?" *The Journal of Experimental Education*, 64, 364-390.
- Marsh, H. W., Balla, J. R. y McDonald, R. P. (1988). "Goodness of fit indexes in confirmatory factor analysis: The effect of sample size". *Psychological Bulletin*, 103, 391-340.
- McDonald, R. P. (1986). "Describing the elephant: Structure and function in multivariate data". *Psychometrika*, 51, 513-534.

- McDonald, R. P. (1996). "Path analysis with composite variables". *Multivariate Behavioral Research*, 31, 239-270.
- McDonald, R. P. y Marsh, H. W. (1990). "Choosing a multivariate model: Noncentrality and goodness of fit". *Psychological Bulletin*, 107, 247-255.
- Mulaik, S. A. (1998). "Parsimony and model evaluation". *The Journal of Experimental Education*, 66, 266-273.
- Mulaik, S. A. (2001). "The curve-fitting problem: An objetivist view". *Philosophy of Science*, 68, 218-241.
- Mulaik, S. A., James, L. R., Van Alstine, J., Bennet, N., Lind, S. y Stilwell, D. (1989). "Evaluation of goodness-of-fit indices for structural equation models". *Psychological Bulletin*, 105, 430-445.
- Muthén, B. O., Kaplan, D. y Hollis, M. (1987). "Structural equation modeling for data that are not missing completely at random". *Psychometrika*, 51, 431-462.
- Paxton, P., Curran, P. J., Bollen, K. A., Kirby, J. & Chen, F. (2001). "Monte carlo experiments: Designs and implementation". *Structural Equation Modeling*, 8, 287-312.
- Raykov, T. y Marcoulides, G. A. (1999). "On desirability of parsimony in structural equation model selection". *Structural Equation Modeling*, 6, 292-300.
- Raykov, T. y Marcoulides, G. A. (2000). *A first course in structural equation modeling*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Robles, J. (1992). *Contrastación del modelo N: Mayor efecto del tamaño muestral que el método de estimación y la especificación sobre los índices de ajuste*. Escuela de psicología, trabajo de licenciatura, no publicado. Universidad Católica Andrés Bello. Caracas, Venezuela.
- Seidel, G. y Eicheler, C. (1990). "Identification structure of linear structural models". *Quality and Quantity*, 24, 345-365.
- Spearman, C. (1904). "General intelligence: Objectively determined and measured". *American Journal of Psychology*, 15, 201-293.
- Tanaka, J. S. (1987). "How big is big enough?: Sample size and goodness-of-fit in structural equation models with latent variables". *Child Development*, 58, 136-146.

- Werts, C. E. y Linn, R. L. (1970). "Path analysis: Psychological examples". *Psychological Bulletin*, 74, 193-212.
- Werts, C. E., Rock, D. A., Linn, R. L. y Jöreskog, K. G. (1976). "Comparison of correlations, variances, covariances, and regression weights with or without measurement error". *Psychological Bulletin*, 83, 1007-1013.
- Wright, S. (1960a). "The treatment of reciprocal interaction, with or without lag, in path analysis". *Biometrics*, 16, 423-445.
- Wright, S. (1960b). "Path coefficients and path regressions: Alternative or complementary concepts?". *Biometrics*, 16, 423-445.