REVISTA DE MÉTODOS CUANTITATIVOS PARA LA ECONOMÍA Y LA EMPRESA (33). Páginas 242-263.
Junio de 2022. ISSN: 1886-516X. D.L: SE-2927-06.
www.upo.es/revistas/index.php/RevMetCuant/article/view/5414

Modelado de Ecuaciones Estructurales en el campo de las Ciencias de la Administración

SORIANO, JOSÉ LUIS
Consejo Nacional de Ciencia y Tecnología (México)
Correo electrónico: jl77soriano@gmail.com

MEJÍA-TREJO, JUAN
Universidad de Guadalajara (México)
Correo electrónico: jmejia@cucea.udg.mx

RESUMEN

El presente artículo analiza el papel de los modelos de ecuaciones estructurales en el campo de las ciencias de la administración. Se presenta el enfoque por el que se incorpora el uso de la técnica al campo de la administración y, adicionalmente, se genera una comparativa entre los diferentes algoritmos y paquetes estadísticos disponibles para la aplicación de la técnica. Encontrando en código cerrado: LISREL, EQS, AMOS, SmartPLS, Mplus, CALIS (en SAS), SEPATH (en Statistica), RAMONA (en Systat), SEM (Stata), Matlab, Semopy (en Python). Y en código abierto: Scilab, Julia, R: lavaan, sem, lava, OpenMx y Strum. Identificando las tendencias actuales y enfoques futuros que enrolan la técnica de análisis multivariante dentro de las ciencias de la administración. Encontrando softwares sugeridos para el área de las ciencias de la administración, así como tendencias de mezcla en la técnica, como el uso de las redes neuronales que se comienza a incorporar como alternativa o complemento a la técnica SEM.

Palabras clave: Ecuaciones Estructurales (SEM), administración, algoritmos, software.
Clasificación JEL: C87; C50; M00.
MSC2010: 91C15; 90-04; 68-00; 15-04.

Artículo recibido el 10 de noviembre de 2020 y aceptado el 17 de agosto de 2021.

Structural Equations Modeling in the Management Sciences

ABSTRACT

This article analyzes the role of structural equation modeling in the field of management science. The approach by which the use of the technique is incorporated into the field of management. Additionally, a comparison is generated between the different algorithms and statistical packages available for the application of the technique. Finding in closed source, LISREL, EQS, AMOS, SmartPLS, Mplus, CALIS (in SAS), SEPATH (in Statistica), RAMONA (in Systat), SEM (Stata), Matlab, Semopy (in Python). And in open source: Scilab, Julia, R: lavaan, sem, lava, OpenMx, Strum. Identifying current trends and future approaches that embody the multivariate analysis technique within the management sciences. Finding suggested software for the area of management sciences, as well as mixing trends in the technique, such as the use of neural networks that is beginning to be incorporated as an alternative or complement to the SEM technique.

Keywords: Structural Equations Models (SEM), management, algorithms, software.
JEL classification: C87; C50; M00.
MSC2010: 91C15; 90-04; 68-00; 15-04.

1. Introducción

Es importante aclarar que el presente documento tiene como finalidad analizar el uso de la técnica referente a los Modelos de Ecuaciones Estructurales (Structural Equations Models, SEM), en el campo de las ciencias administrativas y la pertinencia de cada software disponible, partiendo de las características propias que se requieren en el área de investigación. Por tal motivo, el presente documento no pretende ser un estudio técnico en referencia a la comparativa entre software. El análisis se enfoca en analizar la pertinencia de la técnica en el campo de las ciencias administrativas, el estado actual, futuro y los softwares más apropiados para los análisis dentro del área de estudio, de acuerdo a las características propias de los datos, del perfil del investigador, del tipo de pruebas y compatibilidad, entre otros.

Los Modelos de Ecuaciones Estructurales (Structural Equations Models, SEM) es una técnica flexible que permiten modelar relaciones complejas entre variables que se observan (variables observables) y variables que no se manifiestan (variables latentes). Se puede pensar en una combinación de modelos de análisis factorial (llamados modelos de medición) y modelos de regresión (llamados modelos estructurales). En Igolkina y Meshcheryakov (2020) se menciona que los modelos de ecuaciones estructurales (SEM) son una técnica estadística multivariante para estimar relaciones complejas entre variables observadas y latentes. El enfoque detrás de la estructura es que las variables manifiestas generalmente se miden con análisis factorial y modelos de error pueden dar cuenta del error de medición, por lo que la verdadera relación estructural entre las variables sin error de medición, se modela mediante modelos de regresión entre variables latentes.

En el campo de las ciencias sociales, los modelos suelen denominarse como "causales". Sin embargo, hay un matiz importante en la terminología, ya que estos modelos no establecen causalidad, sino que solo prueban las relaciones propuestas entre variables en estudio (Rosseel, 2012). A pesar de que la gran parte del desarrollo de SEM ha tenido lugar en el ámbito social y la ciencia del comportamiento, la investigación y las revisiones SEM se han publicado en las principales revistas estadísticas y capítulos de libros (Yuan & Bentler, 1998; Steiger, 2007; Marcoulides, 2018).

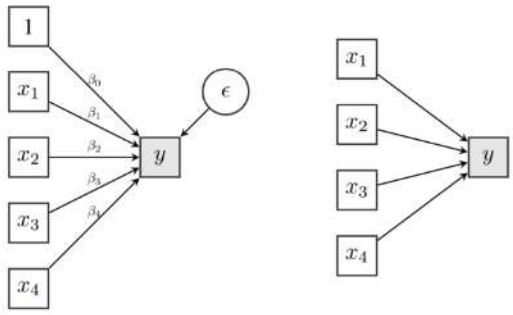
2.Tecnica estadística SEM

En términos generales, SEM es una técnica de modelado estadístico multivariante que permite probar una hipótesis (modelo sobre los datos), postulando un modelo generador de datos y este modelo puede ajustarse o no a los datos.

La característica principal de la técnica, destaca que puede contener variables latentes, que pueden ser "constructos" hipotéticos (por ejemplo, depresión) y medido por un conjunto de indicadores. Las variables latentes pueden ser efectos aleatorios (por ejemplo, intersecciones aleatorias). Los SEM permiten efectos indirectos (mediación) y efectos de ambas vías.

En la Figura 1 se presenta la lógica en la que se basa la técnica SEM. Se muestra una regresión lineal múltiple, donde cuatro variables independientes, más la suma del intercepto (1), generan un efecto en la variable dependiente, donde además se agrega el error.

Figura 1. Regresión lineal múltiple.

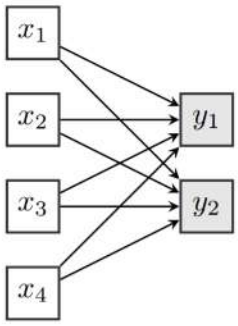


$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_3 x_{i3} + \beta_4 x_{i4} + \epsilon_i \quad (i = 1, 2, \ldots, n)$$

Fuente: Elaboración propia a partir de Rosseel (2012).

En la Figura 2 se plantea una regresión lineal múltiple donde las variables independientes generan un efecto en ambas variables observadas (Y_1, Y_2).

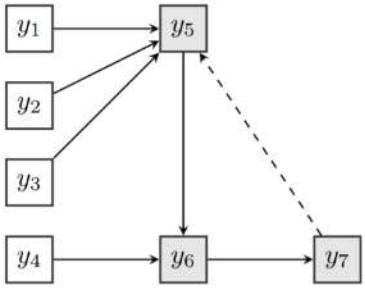
Figura 2. Efecto múltiple de variables independientes.



Fuente: Elaboración propia a partir de Rosseel (2012).

En la Figura 3 se muestra el efecto indirecto de la variable Y_5 en la variable Y_7 por medio del paso de una variable mediadora como la Y_6.

Figura 3. Efecto indirecto.

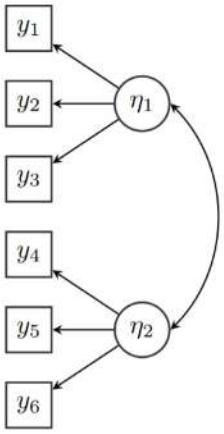


Fuente: Elaboración propia a partir de Rosseel (2012).

Todas las variables son observadas (manifiestas), se permite el efecto indirecto (por ejemplo, de Y_5, a través de Y_6 en Y_7), además se permiten los ciclos (por ejemplo, Y_7 podría influir en Y_5).

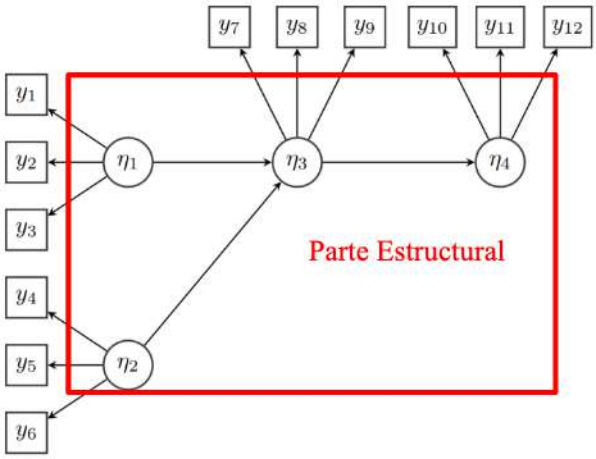
En la Figura 4 se observa una relación entre variables latentes, por medio de análisis factorial confirmatorio (CFA), donde se incorporan modelos de medición, que representa la relación entre una o más variables latentes y sus indicadores (observados). En la Figura 5 se presentan modelos de medida, más la parte estructural (dentro del recuadro rojo) y un análisis de ruta con variables latentes.

Figura 4. Variable latente e indicadores.



Fuente: Elaboración propia a partir de Rosseel (2012).

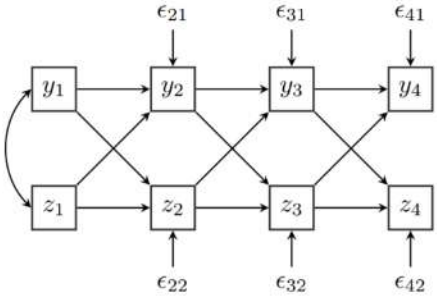
Figura 5. Parte estructural.



Fuente: Elaboración propia a partir de Rosseel (2012).

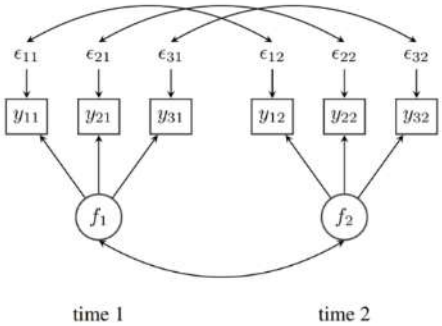
En la Figura 6 se plantea un análisis por medio de modelos de paneles (variables observables) con efectos cruzados, permitiendo contestar preguntas relacionadas a cómo se desarrollan dos variables de forma independiente o como ejerce Y sobre Z y viceversa. Por su parte, en la Figura 7 se plantea una comparativa entre modelos por medio del factor tiempo.

Figura 6. Modelo de paneles con efectos cruzados.



Fuente: Elaboración propia a partir de Rosseel (2012).

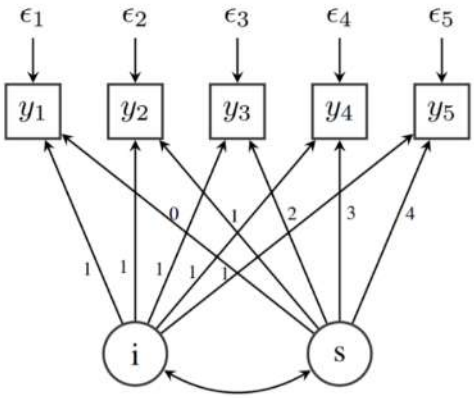
Figura 7. Dos puntos en el tiempo.



Fuente: Elaboración propia a partir de Rosseel (2012).

Para finalizar, en la Figura 8 se presenta un planteamiento que permite contestar preguntas relacionadas al conocer el efecto direcciones de una variable sobre otra.

Figura 8. Análisis de ruta con variables latentes

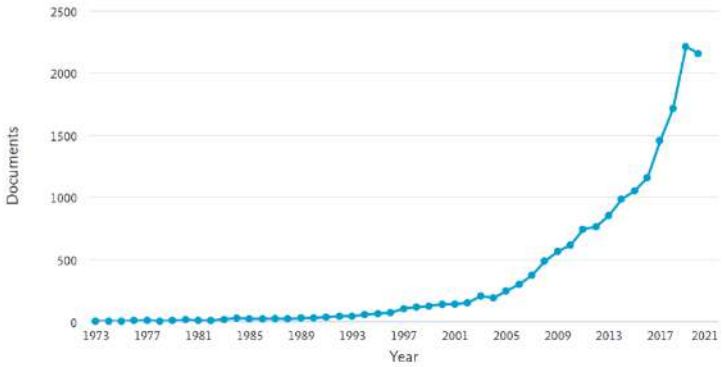


Fuente: Elaboración propia a partir de Rosseel (2012).

3. Uso actual de la técnica

Una vez que se ha mencionado en que consiste la técnica SEM y los principales planteamientos en la aplicación de la técnica, se hace un recuento de las áreas que utilizan la técnica para sus análisis. Conforme a lo mencionado por Rosseel (2017), el SEM es ampliamente utilizado en las ciencias sociales y es cada vez más "descubierto" en las ciencias médicas, biología, neuroimagen, ecología, investigación de operaciones, entre otras disciplinas. La mayoría del software SEM también se utiliza para realizar análisis estándar, por ejemplo, una simple regresión; no obstante, se debe lidiar con datos faltantes, errores estándar robustos, diagnósticos y restricciones de igualdad.

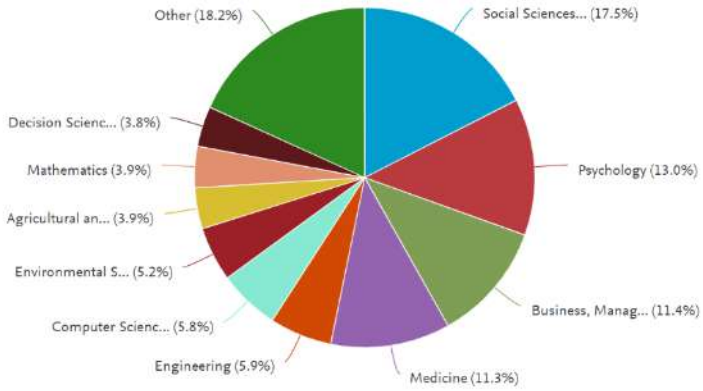
Figura 9. SEM como parte de la investigación.



Fuente: Scopus (2020).

Por medio de la plataforma "Scopus" se generó la búsqueda de todas las publicaciones disponibles donde se hace uso de la herramienta "Structural Equation Model", generando como resultado lo presentado en la Figura 9, donde se muestra el creciente uso de la técnica a partir de los años 90´s. Se utilizó la plataforma Scopus debido a que la misma plataforma argumenta contar con la mayor base de datos bibliográfica de resúmenes y citas de artículos de revistas científicas.

Figura 10. Áreas de investigación.

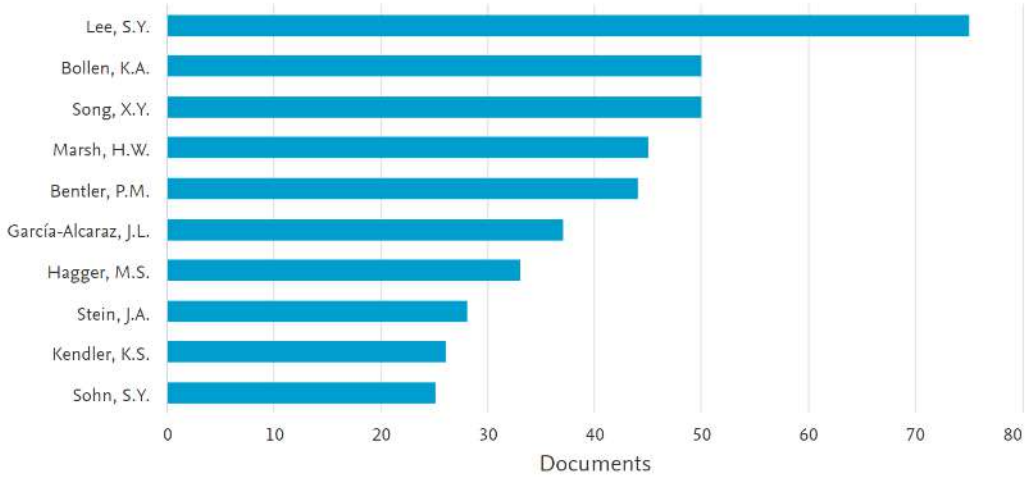


Fuente: Scopus (2020).

Por su parte en la Figura 10 se muestra el porcentaje de distribución del campo del conocimiento en el que se emplea la técnica. Destacando la participación de las ciencias sociales con un 17.5%, seguido de Psicología con el 13% y, en tercer lugar, con una participación similar, se encuentra el área de la administración junto al área de la medicina. En menor medida se utiliza en áreas como la ingeniería, ciencias computacionales, entre otros.

Por último, en la Figura 11 se muestran los autores con mayor producción científica relacionada al uso de la técnica SEM, destacando autores como Peter Bentler, creador del software EQS.

Figura 11. Producción científica relacionada al SEM.



Fuente: Scopus (2020).

Teniendo en cuenta los resultados obtenidos en la plataforma Scopus, se identifica que, dentro del área de las ciencias de la administración, el paquete estadístico que más se suele referenciar es el EQS, seguido de AMOS, Lavaan, Lisrel y SmartPLS.

4. Descripción de los Software

Una vez revisados los rubros más importantes donde se utiliza la técnica SEM, se presenta el contexto de los paquetes estadísticos que permiten el uso de la técnica. Los softwares para SEM se pueden dividir en código cerrado y código abierto. Dentro de los desarrollos iniciales se encuentra el paquete LISREL que surgió en los años 70 con Karl Joreskog; posteriormente en los años 80´s, nace el EQS con Peter Bentler; para la década de los 90´s surge AMOS por medio de James Arbuckle, seguido de Mplus en 1998 desarrollado por Bengt Muthen. Posteriormente surgen nuevos paquetes, como SmartPLS, CALIS (en SAS), SEPATH (en Statistica), RAMONA (en Systat), SEM (Stata), Matlab, Semopy (en Python) y en código abierto, Scilab, Julia y en paquetes en R: lavaan, sem, lava, OpenMx, Strum. En el presente apartado se describe de forma sintetizada los softwares disponibles con la herramienta SEM y posteriormente se presenta una tabla comparativa entre los diferentes paquetes y sus principales distinciones.

LISREL

LISREL es una aplicación de programa proporcionada por Windows para realizar el modelado de ecuaciones estructurales (SEM) y otro modelado de estructuras lineales relacionadas (por ejemplo, modelado de ecuaciones estructurales multinivel, modelado multinivel lineal y no lineal, etc.)

Conforme a lo mencionado por Jöreskog y Sörbom (2018) LISREL para Windows es útil para importar datos externos en varios formatos como SPSS, SAS, MS Excel, etc. como un archivo de sistema PRELIS (PSF). El software utiliza un archivo de gráficos con la extensión predeterminada llamada PTH para capturar el diagrama de ruta. También es útil para ajustar el modelo medido a los datos. Por ejemplo, supongamos que uno quiere usar LISREL mientras ajusta el modelo a los datos de los tutoriales de SPSS.

El paquete permite manejar modelos con error de medición, apoya al investigador en modelos no recursivos y se considera útil para resolver problemas de análisis multivariante (por ejemplo, MANOVA). Útil para trabajar en comparaciones de varios grupos (como desarrollar modelos separados para hombres, mujeres y pruebas de restricciones.

Una cualidad que es bastante común en el modelo LISREL es que los modelos ignoran las medias y consideran que todas las variables se centran en las medias de su grupo. Esto, a su vez, da como resultado que los modelos tengan medias cero. Esto se hace para reducir la complejidad asociada al análisis. Si se está trabajando en un modelo de varios grupos con la ayuda de LISREL, entonces dará el mismo resultado de ese proceso que se obtiene al ejecutar una regresión con variables ficticias en SPSS.

EQS

Conforme a Bentler (2006), el paquete EQS combina la sofisticación técnica con la facilidad de uso, EQS hace que el modelado de ecuaciones estructurales sea aún más fácil con Diagrammer, una herramienta de dibujo de modelos de EQS.

En términos generales EQS mantiene una interfaz amigable, con funciones de apuntar y hacer clic para un acceso rápido a cualquier prueba de EQS, sin necesidad de programar, como en Python o R. Asimismo, EQS prueba la gama completa de modelos de ecuaciones estructurales que incluyen regresión múltiple, análisis factorial confirmatorio, análisis de medias estructurado y comparaciones de poblaciones múltiples. EQS también incluye correlaciones poliseriales y policóricas para el tratamiento de datos categóricos, el estadístico chi-cuadrado escalado de Satorra-Bentler, mejoras en las opciones RETEST y WTEST, y la capacidad de importar matrices de peso. De acuerdo a Bentler (2006) el EQS sirve a miles de investigadores en psicología, educación, economía, investigación de mercado, sociología y otros campos que utilizan técnicas de modelado de ecuaciones estructurales.

AMOS

IBM SPSS Amos es un potente software de modelado de ecuaciones estructurales (SEM) que ayuda a respaldar las investigaciones y teorías al ampliar los métodos de análisis multivariante estándar, que incluyen regresión, análisis de factores, correlación y análisis de varianza. Hair et al. (2014) menciona la posibilidad de crear modelos de actitud y comportamiento que reflejen relaciones complejas con mayor precisión que con las técnicas de estadística multivariante estándar, utilizando una interfaz de usuario gráfica o programática intuitiva, facilitando la experiencia de usuario.

Smart PLS

Conforme a lo mencionado por Manley et al. (2020) SmartPLS permite generar un modelado de ruta de mínimos cuadrados parciales (PLS), Regresión de mínimos cuadrados ordinarios (OLS) basada en sumscores, PLS consistente (PLSc), PLS ponderado (WPLS), OLS ponderado (WOLS) y PLS consistente ponderado (WPLSc), Bootstrapping y el uso de opciones avanzadas de bootstrapping. Asimismo, permite análisis de mapas de importancia y rendimiento (IPMA), análisis multigrupo PLS (MGA) que analiza la diferencia y la importancia de las estimaciones del modelo de ruta PLS específicas del grupo, variables mediadoras y estimación de efectos indirectos en conjunto a sus pruebas de significación basadas en Bootstrap.

Aria (2020) menciona que, además de lo anterior, permite la estimación de los efectos de interacción y sus pruebas de significación basadas en bootstrapping, relaciones no lineales; estimación de efectos cuadráticos y sus pruebas de significación basadas en Bootstrap, análisis de tétrada confirmatorio (CTA): una técnica estadística que permite probar empíricamente la configuración del modelo de medición, segmentación de mezcla finita (FIMIX): un enfoque de clase latente que permite identificar y tratar la heterogeneidad no observada en los modelos de ruta, segmentación orientada a la predicción (POS): un enfoque para identificar grupos de datos, PLS Predict: una técnica para determinar la calidad predictiva del modelo de ruta PLS.

MPlus

Según Muthén y Muthén (2017), Mplus es un programa de modelado estadístico que proporciona a los investigadores una herramienta flexible para analizar sus datos. Mplus ofrece a los investigadores una amplia variedad de modelos, estimadores y algoritmos en un programa que tiene una interfaz fácil de usar y pantallas gráficas de datos y resultados de análisis. Mplus permite el análisis de datos transversales y longitudinales, datos de un solo nivel y multinivel, datos que provienen de diferentes poblaciones con heterogeneidad observada o no observada y datos que contienen valores faltantes. Se pueden realizar análisis para variables observadas que son continuas, censuradas, binarias, categóricas ordenadas (ordinales), categóricas no ordenadas (nominales), recuentos o combinaciones de estos tipos de variables. Además, Mplus cuenta con amplias capacidades para estudios de simulación Monte Carlo, donde los datos se pueden generar y analizar de acuerdo con cualquiera de los modelos incluidos en el programa.

CALIS (SAS)

En Yung (2010) se enfatiza que PROC CALIS parte del supuesto de que las variables aleatorias tienen una distribución normal aproximadamente multivariante. La no normalidad, especialmente la curtosis alta, puede producir estimaciones deficientes y errores estándar extremadamente incorrectos y pruebas de hipótesis, incluso en muestras grandes. En consecuencia, el supuesto de normalidad es mucho más importante que en modelos con variables exógenas no estocásticas. Por lo anterior, el paquete sugiere eliminar los valores atípicos y considerar transformaciones de variables no normales antes de usar PROC CALIS con máxima probabilidad (predeterminado) o estimación generalizada de mínimos cuadrados.

Alternativamente, hay varios enfoques disponibles para tratar el problema de la no normalidad. Si el número de observaciones es lo suficientemente grande, puede utilizar la distribución asintóticamente libre de Browne (ADF; Browne 1982) método de estimación. Sin embargo, no existe una directriz definida sobre el tamaño que debe tener una muestra.

SEPATH (Statistica)

Señalan Sotnikova-Meleshkina y Martynenko (2020) que el módulo incluye opciones de simulación que permite al investigador generar y guardar conjuntos de datos para modelos predefinidos, basados en distribuciones normales o sesgadas. Se pueden calcular estimaciones de Bootstrap, así como distribuciones para diversas estadísticas de diagnóstico, estimaciones de parámetros, etc., sobre los ensayos de Monte Carlo. Las principales aplicaciones del modelado de ecuaciones estructurales incluyen, modelado causal, o análisis de ruta, que hipotetiza relaciones causales entre variables y prueba los modelos causales con un sistema de ecuaciones lineales; los modelos causales pueden involucrar variables manifiestas, variables latentes o ambas, análisis factorial confirmatorio.

RAMONA (Systat)

Dentro del paquete RAMONA en Systat, Stapleton y Leite (2005) afirman que se pueden generar cuadros y gráficos en 2D y 3D. Las interfaces de comandos y menús de SYSTAT están interconectadas, lo que otorga a los usuarios la posibilidad de elegir el método que mejor les funcione. Permitiendo

analizar, matrices de covarianza o correlación, pruebas como MWL (máxima probabilidad de Wishart), GLS (mínimos cuadrados generalizados), OLS (mínimos cuadrados ordinarios), ADFG (estimación libre de distribución asintóticamente sesgada, Gramian) y ADFU (imparcial).

SEM (STATA)

Conforme a lo mencionado por Driver et al. (2017), el paquete SEM estima los efectos de mediación, analiza la relación entre un concepto latente no observado como la depresión y las variables observadas que miden la depresión, modela un sistema con muchas variables endógenas y errores correlacionados, o ajusta un modelo con relaciones complejas entre variables latentes y observadas. Ajusta modelos con resultados continuos, binarios, de conteo, ordinales, fraccionales y de supervivencia. Incluso ajuste de modelos multinivel con grupos de observaciones correlacionadas. Asimismo, permite evaluar el ajuste del modelo, calcular los efectos indirectos y totales, ajustar los modelos dibujando un diagrama de ruta o utilizando la sintaxis de comandos sencilla.

Matlab

Por su parte Williams (2020), menciona que MATLAB permite ver cómo funcionan diferentes algoritmos con los datos, permite realizar iteraciones hasta obtener los resultados deseados y, después, generar automáticamente un programa en el mismo MATLAB para reproducir o automatizar el trabajo. Asimismo, el paquete cuenta con la capacidad de escalación para la ejecución en clusters, GPUs y nubes únicamente con cambios menores en el código. No es necesario volver a escribir el código ni aprender programación para Big Data y técnicas de manejo de datos fuera de memoria; lo anterior se considera pertinente para análisis complejos, que requieran algoritmos específicos.

Semopy (Python)

Semopy significa “optimización de modelos de ecuaciones estructurales en Python” y está diseñado para ayudar a los estadísticos que emplean técnicas SEM a manejar su investigación de una manera más enfocada al código Python. El paquete busca llenar un nicho de herramientas SEM en Python que parece estar vacío, ya que de acuerdo a Igolkina y Meshcheryakov (2020) se encuentran opiniones negativas de los investigadores respecto al uso de software comercial o de lenguajes de computadora más antiguos. El paquete Semopy se proporciona de forma totalmente gratuita y todos pueden hacer contribuciones al proyecto.

Los desarrolladores de Semopy prometen estar atentos a los errores y considerar implementar nuevas funciones a pedido. El paquete se publicó en "Modelado de ecuaciones estructurales: una revista multidisciplinaria". Se demuestra que Semopy supera al popular paquete gratuito Lavaan en términos de rendimiento, estabilidad del proceso de optimización y precisión de estimaciones de parámetros. Desde entonces, Semopy ha enfrentado cambios positivos drásticos y ha desarrollado nuevas características, por lo tanto, un lector puede interpretar los resultados presentados en el artículo como un mínimo esperado.

Scilab

Scilab está disponible bajo la licencia GPL, lo que significa que se tiene la libertad de utilizar el software para cualquier propósito, cambiar el software para adaptarlo a las necesidades del investigador. De acuerdo a Campbell et al. (2006) el estilo de codificación está definido para todos los lenguajes utilizados en Scilab C, C ++, Java, XML, Scilab, Makefiles, Checkstyle está disponible para Java en Eclipse. Cuando se pretende desarrollar una nueva funcionalidad, el código debe estar documentado (comentarios usando Javadoc, Doygen), cada modificación en el código Scilab es revisada y aceptada por un desarrollador.

Cada nueva funcionalidad solo se integra si tiene una prueba correspondiente y esta ejecución de prueba se ejecuta correctamente en todas las plataformas compatibles con Scilab. Asimismo, para cada

corrección de error, se escribe una prueba para garantizar su corrección adecuada y evitar cualquier regresión futura. Las herramientas de cobertura de código se utilizan para verificar la calidad de las pruebas. Cada confirmación en el repositorio de Scilab GIT lanza un proceso de integración continúa basada en Jenkins. La compilación diaria del binario Scilab a través de un proceso de compilación, ejecuta todas las pruebas unitarias y las pruebas de no regresión en todas las plataformas compatibles.

Julia

Un paquete de Julia "Lavaan.jl" se puede utilizar para el modelado de ecuaciones estructurales inspirado en el paquete R del mismo nombre. El objetivo del proyecto es desarrollar un paquete de software en Julia que realice el modelado de ecuaciones estructurales, es la opción disponible para usuarios del lenguaje Julia. El proyecto hace crecer la infraestructura de paquetes de Julia y la intención del desarrollador es atraer a más investigadores aplicados al lenguaje. De acuerdo al desarrollador, proporcionar herramientas de código abierto que superan en rendimiento a las alternativas de código cerrado beneficiosas para la ciencia.

Lo anterior puede servir como una oportunidad de aprendizaje para los miembros del grupo interesados en aprender más sobre las mejores prácticas en ingeniería de software. Se puede utilizar el paquete R existente llamado Lavaan como referencia y también como referencia para el rendimiento.

Lavaan

El paquete Lavaan está desarrollado para proporcionar a los usuarios, investigadores y profesores un paquete gratuito de código abierto, pero de calidad comercial para el modelado de variables latentes. Se puede utilizar Lavaan para estimar una gran variedad de modelos estadísticos multivariantes, incluido el análisis de ruta, el análisis factorial confirmatorio, el modelado de ecuaciones estructurales y los modelos de curvas de crecimiento.

Los modelos de ecuaciones (con variables observadas y latentes) utilizan el enfoque RAM para ajustar ecuaciones estructurales en modelos de variables observadas, se realiza mediante mínimos cuadrados de dos etapas. Por lo que el paquete cuenta con funciones para ajustar modelos de ecuaciones estructurales lineales generales (con variables observadas y latentes) utilizando el enfoque RAM, y para ajustar ecuaciones estructurales en modelos de variables observadas mediante mínimos cuadrados de dos etapas.

Lava

En Holst y Budtz-Jørgensen (2013) se menciona que Lava es un paquete que permite una implementación general de modelos de ecuaciones estructurales con variables latentes (MLE, 2SLS y estimadores de verosimilitud compuestos) con resultados continuos, censurados y ordinales. Modelos de mezcla de variables latentes y modelos de variables latentes no lineales (Holst y Budtz-Joergensen, 2019). El paquete también proporciona métodos para la exploración de gráficos, separación de criterio de puerta trasera, simulación de modelos generales de variables latentes no lineales y estimación de funciones de influencia para una amplia gama de modelos estadísticos.

Open Mx

Conforme a lo mencionado por los desarrolladores Boker et al. (2021), OpenMx es un software gratuito y de código abierto para usar con R que permite la estimación de una amplia variedad de modelos estadísticos multivariantes avanzados. OpenMx consiste en una biblioteca de funciones y optimizadores que le permiten definir de forma rápida y flexible un modelo SEM y estimar parámetros dados los datos observados. OpenMx se ejecuta en MacOS, Windows y la mayoría de las variedades de Linux/GNU. Esto significa que los mismos scripts que escribe en Windows se ejecutarán en MacOS o Linux.

OpenMx puede ser utilizado por aquellos que piensan en términos de modelos de ruta o por aquellos que prefieren especificar modelos en términos de álgebra matricial. En OpenMx se pueden especificar y modificar modelos y conjuntos de datos complicados mediante el lenguaje R.

Strum

De acuerdo a Song et al. (2015), el paquete R Strum es una herramienta de software SEM adecuada para el análisis genético. Implementa un marco general para SEM dentro del contexto de datos genealógicos generales. Este contexto requiere consideraciones especializadas como correlaciones familiares y verificación. El paquete es una herramienta flexible capaz de modelar asociación genética, análisis de ligamiento, efectos poligénicos, ambiente compartido y verificación combinados con análisis factorial confirmatorio y SEM general. También proporciona una herramienta conveniente para la visualización de modelos e integra herramientas para simular datos de pedigrí. Las diversas características de este paquete se prueban mediante un estudio de simulación para evaluar el rendimiento, y los resultados muestran que strum es confiable y robusto en términos de precisión y cobertura de las estimaciones de parámetros.

5. Criterios comparativos

Una vez mencionados los paquetes disponibles para la aplicación de la técnica SEM, se presenta un cuadro comparativo, donde Segura et al. (2008), Gómez et al. (2010), Ríos y Idrobo (2017) designan criterios para evaluar distintos tipos de software, resaltando los más importantes, como la comunidad, información, compatibilidad, actualización, funcionalidad y facilidad de uso. En la Tabla 1 se presenta un cuadro comparativo con los criterios mencionados junto a dos columnas adicionales, de fortalezas y particularidades, con la finalidad de dar continuidad al trabajo de Narayanan (2012) donde se comparan ocho paquetes SEM.

El criterio de comunidad se relaciona con el número de usuarios, foros y discusiones relacionadas al paquete; por otra parte, en el criterio de información se toma en cuenta la información disponible por parte del proveedor, para el uso del paquete; al mismo tiempo la claridad en que se trasmite la información también se considera.

El criterio de compatibilidad, hace referencia a las versiones disponibles para los diferentes sistemas operativos y, al mismo tiempo, la posibilidad de enlazar el paquete SEM con paquetes estadísticos, gráficos o sistemas de composición de texto como Latex, que permite a los investigadores, una mayor versatilidad relacionada al software.

El rubro relacionado a la actualización, se enfoca en la revisión de las actualizaciones que cada software genera a lo largo del tiempo y cómo dichas actualizaciones generan mejoras e incluso resuelven problemas previos del propio código, permitiendo actualizar los algoritmos de acuerdo a los nuevos planteamientos, descubrimientos o discusiones académicas.

La funcionalidad, tomando en cuenta el objetivo del software, es el criterio fundamental, ya que se relaciona con la disponibilidad y capacidad del paquete para correr diferentes técnicas y pruebas, que permiten plantear un SEM de forma correcta, por lo que en el criterio relacionado a la funcionalidad, se hace referencia a las opciones disponibles a utilizar en el paquete, por lo que es notable, que paquetes de código abierto, cuenten con una mayor capacidad funcional, debido a que, en el caso de requerir aplicar pruebas experimentales, se puede crear el código específico para dicha prueba, acción que se complica en paquetes de código cerrado. Un ejemplo es Lavaan, donde el ecosistema se relaciona con paquetes adicionales que permiten incrementar la funcionalidad, como el paquete Lavaan (Ed Merkle, Yves Rosseel), Bayesian SEM (actualmente usando jags) con una interfaz de Lavaan, lavaan.survey (Oberski, 2016) que permite ponderaciones de la encuesta, agrupamiento, estratos y correcciones de muestreo finito en SEM.

No obstante, como ejemplo, dentro del paquete Lavaan, por mencionar uno de código abierto, dispone de más códigos que incrementan la funcionalidad, como el Ónix (Von Oertzen, Brandmaier & Tsang, 2015), la interfaz gráfica interactiva para SEM (escrito en Java), SemTools y colección de funciones útiles para SEMsimsem, simulación de modelos SEM, SemPlot, visualizaciones de modelos SEM, EffectLiteR, usando SEM para estimar efectos promedio y condicionales, Nlsem, estimación de modelos de ecuaciones estructurales con efectos no lineales y distribuciones no normales subyacentes, así como Bmem, Coefficientalpha, Eqs2lavaan, FSRM, Influence.SEM, MIIVsem, ProfileR, RAMpath, Regsem, RMediation, RSA, Rsem, Stremo, Faoutlier, Dame, Lavaan.shiny, Matrixpls, MBESS, NlsyLinks, Nonnest2, PiecewiseSEM, Pscore, Psytabs, Qgraph, Sesem, Sirt, TAM y Userfriendlyscience que son paquetes complementarios que incrementan la funcionalidad.

De acuerdo a Merkle et al. (2020), la metodología y el software están destinados a proporcionar a los usuarios un medio general de estimar los SEM bayesianos, clásico y novedoso, de una manera sencilla. Los usuarios pueden estimar versiones bayesianas de SEM clásicos con sintaxis de Lavaan, pueden obtener medidas de ajuste bayesiano de última generación asociados con los modelos, y pueden exportar el código JAGS para modificar los SEM como se desee. Por otra parte, el lenguaje de programación de Julia se desarrolló desde 2012 y es más rápido que Python ya que éste es un lenguaje interpretado en comparación del lenguaje compilado y de acuerdo al impulsor del lenguaje, recopila lo mejor de Python y R; no obstante, se considera más complicada la forma de programación, aunque es relativo, ya que se considera complicado por la novedad de ser un lenguaje proveniente del 2012.

Las desventajas relacionadas al lenguaje de Julia se enfocan en las escasas librerías provenientes del periodo corto que transcurre desde su creación; no obstante, se considera una opción disponible para investigaciones exploratorias donde se requiera una capacidad y velocidad mayor, como pudiera ser la integración de Big Data y los modelos de ecuaciones estructurales. El paquete disponible que se relaciona al modelado de ecuaciones estructurales es llamado “MixedModels” que se encuentra disponible de forma gratuita desde el 2016; no obstante, se ha actualizado constantemente, hasta llegar a 38 versiones (actualizaciones) llegando hasta la versión actual v2.3.0.

El paquete define modelos lineales mixtos (LinearMixedModel) y modelos lineales mixtos generalizados (GeneralizedLinearMixedModel). Los usuarios pueden usar la abstracción para la API del modelo estadístico para construir, ajustar y consultar los modelos ajustados. Un modelo de efectos mixtos es un modelo estadístico para una variable de respuesta en función de una o más covariables. Para una covariable categórica, los coeficientes asociados con los niveles de la covariable a veces se denominan efectos, como en "el efecto de usar el Tratamiento 1 frente al placebo". Si los niveles potenciales de la covariable son fijos y reproducibles, por ejemplo, los niveles de sexo pueden ser "F" y "M", se modelan con parámetros de efectos fijos. Si los niveles constituyen una muestra de una población, por ejemplo, el sujeto o el Ítem en una observación particular, se modelan como efectos aleatorios.

Un modelo de efectos mixtos contiene términos tanto de efectos fijos como de efectos aleatorios. Con efectos fijos, son los coeficientes en sí mismos o las combinaciones de coeficientes los que son de interés. En el caso de los efectos aleatorios, lo que interesa es la variabilidad de los efectos sobre la población. De acuerdo a Jacobucci et al. (2016) en este paquete, los efectos aleatorios se modelan como muestras independientes de una distribución gaussiana multivariante de la forma $\boldsymbol{\mathcal{B}} \sim \boldsymbol{\mathcal{N}}\ (0, \boldsymbol{\Sigma})$. Para el vector de respuesta, $\mathbf{y}$, solo la media de la distribución condicional, $\boldsymbol{\mathcal{Y}} \mid \boldsymbol{\mathcal{B}} = \mathbf{b}$ depende de $\mathbf{b}$ y lo hace mediante una expresión de predictor lineal, $\boldsymbol{\eta} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Zb}$, donde $\boldsymbol{\beta}$ es el vector de coeficientes de efectos fijos y $\mathbf{X}$ y $\mathbf{Z}$ son matrices modelo de los tamaños apropiados. En un modelo lineal mixto, la media condicional, $\boldsymbol{\mu} = \mathbb{E}\ [\boldsymbol{\mathcal{Y}} \mid \boldsymbol{\mathcal{B}} = \mathbf{b}]$, es el predictor lineal, $\boldsymbol{\eta}$, y la distribución condicional es gaussiana multivariante, $(\boldsymbol{\mathcal{Y}} \mid \boldsymbol{\mathcal{B}} = \mathbf{b}) \sim \boldsymbol{\mathcal{N}}\ (\boldsymbol{\mu}, \sigma^2\mathbf{I})$. Por lo que, en el paquete de Julia, la media condicional, $\mathbb{E}\ [\boldsymbol{\mathcal{Y}} \mid \boldsymbol{\mathcal{B}} = \mathbf{b}]$, está relacionada con el predictor lineal mediante una función de enlace. Las formas de distribución típicas son Bernoulli para datos binarios o Poisson para datos de recuento.

Asimismo, el criterio de facilidad de uso se encamina en el análisis relacionado a la experiencia técnica del usuario, donde es evidente que paquetes como el AMOS o SmartPLS, visualmente se

manejan con mayor facilidad en comparación con paquetes relacionados a R, Python o Julia; asimismo la lógica y la intuición de la plataforma, son considerados en el mismo apartado.

Por último, se añade una columna extraordinaria llamada Sugerido para análisis en Ciencias Administrativas (SCA), con la finalidad de sugerir el uso del software para los análisis relacionados al campo de las ciencias de la administración, donde el criterio principal se relaciona con la factibilidad de manipular los datos que se suelen utilizar en el área, por ejemplo, escalas que son muy comunes dentro de las ciencias de la administración e incluso adaptabilidad con Big Data. Un criterio secundario pero importante es el perfil del investigador, derivado de la factibilidad de programar en softwares de código abierto, ya que por lo regular el perfil de los investigadores en el área, no suele ser de programador o cuestiones relacionadas, por lo que se da prioridad a softwares que cuenten con versatilidad respecto al número de pruebas estadísticas de las que disponen, sin necesidad de programar, o incluso herramientas que permiten mejorar la interfaz para que el investigador se concentre en el análisis estadístico.

No obstante, si el investigador cuenta con un perfil enfocado al de programación o está altamente familiarizado con la lógica de los distintos softwares, no tiene relevancia el criterio. Se puede resumir que en el caso de contar con experiencia en la técnica y si se busca explorar nuevas incorporaciones relacionadas a pruebas o técnicas, los softwares de código abierto son los recomendados; por el contrario, para los iniciadores en el uso de la técnica, los softwares sugeridos pueden ser una opción, todo depende de experiencias previas, es decir cualquier paquete de R puede ser complicado de usar; no obstante si ya se cuenta con previo uso de R, no se tendría problema el uso de Lavaan. La casilla perteneciente a cada software se marca con la letra "S" indicando que se sugiere el software. Se enfatiza que la recomendación no es excluyente, ya que como se menciona posteriormente, todos los softwares permiten la aplicación de la técnica obteniendo resultados no diferenciados.

Tabla 1. Cuadro comparativo entre paquetes SEM para la aplicación en las Ciencias de la Administración.

Nombre del Paquete	Comunidad	Información	Compatibilidad	Actualización	Funcionalidad	Facilidad de uso	SCA	Fortalezas	Particularidades
LISREL	9	8	7	8	9	10		Software de modelado lineal general para SEM, modelado multinivel, modelado lineal generalizado y modelado recursivo	Estimación de residuos observacionales en análisis factorial y modelos de ecuaciones estructurales
EQS	9	8	8	9	9	10	S	El modelado estructural se simplifica al ofrecer diferentes formas de crear modelos; más capacidad de análisis exploratorio	Capacidad única para manejar variables no normales en su totalidad y situaciones de datos faltantes Capacidad para calcular el multiplicador de Lagrange multivariante y pruebas Wald. Estimaciones de confiabilidad para el modelo de medición
AMOS	10	10	8	10	9	10	S	Formato de salida bien organizado y de acceso rápido, Excelente interfaz gráfica	Estimación bayesiana de forma limitada, Varias opciones para bootstrapping Búsqueda de especificaciones en ausencia de teoría

Mplus	9	9	10	10	10	10		Marco de modelado integrado para manejar continuo, variables categóricas, observadas y latentes; incorporación de últimos métodos de manejo de datos faltantes, se puede integrar con R	Incorporación de variables auxiliares para el manejo de datos faltantes. Fácil implementación de efectos de interacción entre latentes variables. Capacidad única que incluye Bayesian SEM, SEM exploratorio, manejo de recuento y valores censurados, y datos faltantes en el mecanismo de no falta al azar (NMAR)
SmartPLS	10	10	9	10	9	10	S	Avance gráfico y extensión de algoritmos	Permite análisis de mapas de importancia y rendimiento (IPMA), análisis multigrupo PLS (MGA) que analiza la diferencia y la importancia de las estimaciones del modelo de ruta PLS específicas del grupo
CALIS (SAS)	9	10	9	10	9	10		Integrar otras capacidades de análisis disponibles en SAS, (por ejemplo, MI para el análisis de datos faltantes)	Flexibilidad de la especificación del modelo en muchos formatos diferentes, como COSAN, FACTOR, LINEQS, MSTRUCT y RAM
SEPATH (Statistica)	7	9	8	10	10	10		Plantillas prediseñadas, herramientas de inteligencia artificial	Permite guardar conjuntos de datos para modelos predefinidos, basados en distribuciones normales o sesgadas
RAMONA (Systat)	7	8	8	10	10	10		Flexibilidad en desarrollo visual	Se pueden generar cuadros y gráficos en 2D y 3D, ADFG (estimación libre de distribución asintóticamente sesgada, Gramian) y ADFU (imparcial)
SEM (Stata)	9	9	9	9	9	9		Ajuste modelos con resultados continuos, binarios, de conteo, ordinales, fraccionales y de supervivencia	Calcular los efectos indirectos y totales, ajustar los modelos dibujando un diagrama de ruta o utilizando la sintaxis de comandos sencilla
Matlab	6	7	10	9	9	9		Escalabilidad de análisis para la ejecución en clusters, GPUs y nubes únicamente con cambios menores en el código	No es necesario volver a escribir el código ni aprender programación para Big Data y técnicas de manejo de datos fuera de memoria, lo anterior se considera pertinente para análisis complejos, que requieran algoritmos específicos
Semopy (Python)	7	8	10	8	10	8		Se puede ajustar a las líneas de desarrollo (es decir, bioinformáticas)	Reciente incorporación para usuarios de lenguaje Python
Scilab	6	7	9	7	8	7		Se ajusta con los paquetes adicionales a Scilab. Cada confirmación en el repositorio de Scilab GIT lanza un proceso de integración continúa basada en Jenkins	El estilo de codificación está definido para todos los lenguajes utilizados en Scilab C, C ++, Java, XML, Scilab, Makefiles, Checkstyle está disponible para Java en Eclipse

Julia	8	8	10	8	8	8		Una oportunidad de aprendizaje para los miembros del grupo interesados en aprender más sobre las mejores prácticas en ingeniería de software	Reciente incorporación para usuarios de lenguaje Julia
Lavaan (R)	10	9	9	9	10	9	S	Entorno de código abierto con características completas de R que incluyen capacidades gráficas, Scripting compacto incluso para modelos complicados	Capacidad para calcular dos probabilidades diferentes para la estimación de ML (normal y Wishart), Capacidad para "imitar" los resultados de Mplus y EQS
SEM (R)	8	9	9	9	9	8		Entorno de código abierto con características completas de R, incluidas capacidades gráficas	Incorporación de mínimos cuadrados de dos etapas para modelado econométrico; capacidad para crear un diagrama de ruta a través de Graphviz
Lava (R)	8	9	9	9	9	8		El paquete proporciona métodos para la exploración de gráficos, separación de criterio de puerta trasera, simulación de modelos generales de variables latentes no lineales y estimación de funciones de influencia para una amplia gama de modelos estadísticos	MLE, 2SLS y estimadores de verosimilitud compuestos) con resultados continuos, censurados y ordinales
Open Mx (R)	9	9	9	9	9	8		Entorno de código abierto con características completas de R que incluyen capacidades gráficas Se integra con otros paquetes de R para permitir el procesamiento multinúcleo y procesamiento distribuido a gran escala	Capacidad para optimizar funciones objetivas especificadas por el usuario. Definición de modelo jerárquico para vincular una serie de modelos en un árbol. Estructura capacidad para utilizar el programa como optimizador y calculadora de álgebra matricial.
Strum (R)	7	8	9	9	9	8		Presentamos el nuevo paquete R Strum para satisfacer la necesidad de una herramienta de software SEM adecuada para el análisis genético	Reciente incorporación para usuarios de lenguaje Strum

Fuente: Elaboración propia.

Todos los paquetes producen estimaciones que están bastante cerca entre sí en precisión. La principal diferencia entre los paquetes viene en el usuario, interfaz y disponibilidad de diferentes opciones; mientras que paquetes como AMOS, LISREL y EQS se pueden utilizar en modo de programación y gráficos interfaz, el resto de los paquetes solo se pueden utilizar en un entorno de programación (R, Python, Julia).

Cada paquete tiene sus propias fortalezas y características especiales, como que se muestra en la Tabla 1. Los usuarios de SEM pueden elegir un paquete sobre otro basado en las características especiales necesarias para una particular aplicación o el entorno informático que ya están familiarizados. Mientras que los paquetes R son atractivos para los usuarios que prefieren entornos de

código abierto, los rápidos desarrollos también están en paquetes de software comerciales. Un enfoque unificado podría aprovechar de la fuerza de cada entorno y los paquetes de software SEM tienen cierto margen de mejora.

Como se menciona en Narayanan (2012), se identifican áreas de oportunidad que involucran ayuda con la identificación del modelo, especificación y comparación de modelos incorrectos. Una primera área es cuando no se identifica un modelo, muchos de estos programas ofrecen menor ayuda al usuario para resolver el problema. Mientras la identificación puede considerarse un requisito previo y fundamental para estimación en el ajuste de modelos de ecuaciones estructurales, los investigadores con pocas variables medidas pueden comprometer la fiabilidad del modelo.

Una segunda área de oportunidad es la especificación del modelo incorrecto, en ausencia de una teoría rigurosa, es probable que suceda con más frecuencia. Los programas SEM podrían implementar algunos de los algoritmos de modelado gráfico como se describe en Haughton, Kamis y Scholten (2006) al identificar modelos posiblemente equivalentes que sean consistentes con la estructura de covarianza de los datos.

La tercera área de oportunidad es la comparación de modelos, ya que, en muchos paquetes comerciales, no hay una manera fácil de almacenar los resultados de varios modelos junto con sus estadísticas de ajuste para la comparación. Esto es útil para análisis exploratorios y para la comparación de modelos anidados, el usuario tiene que hacer programación para calcular la diferencia en los valores de chi-cuadrado y grados de libertad. La misma idea también sería útil para comparar modelos en análisis de un solo grupo; la función Anova en R también es útil a este respecto.

El paquete MplusAutomation en R es útil para el pos-procesamiento de la extracción de resultados de Mplus y tabulación de parámetros e índices de ajuste. El paquete puede leer la salida de varios directorios, extraer información resumida y mostrarlos en forma de tabla concisa. Un ejemplo para una serie de modelos anidados que prueban la invariancia de medición. Los paquetes de R y EQS permiten llamar, ejecutar, y leer los resultados entre ellos para su posterior procesamiento, Mair, Wu y Bentler (2010) mostraron cómo combinar lo mejor de ambos mundos, mediante el estudio del efecto de las violaciones de la normalidad sobre las distribuciones de las estadísticas de prueba y el análisis de potencia para un parámetro estructural. Estos ejemplos muestran cómo combinar las potentes capacidades computacionales y gráficas de R con la implementación única en algunos de los paquetes de software SEM.

Lee (2015) menciona que en los paquetes como Mplus, LISREL, SAS PROC CALIS, y R package Lavaan se pueden automatizar los modelos para generar comparaciones con el cambio de datos por medio del Método de Monte Carlo; no obstante, se comenta que en el paquete R, por medio del paquete “simsem” se reducen los pasos y se simplifican los resultados.

En resumen, se identifican ciertas diferencias entre software (presentadas en la Tabla 1); no bastante, en términos de consistencia estadística, se puede establecer que los resultados de las pruebas determinaron estimaciones consistentes.

Sin embargo, existen diferencias en las capacidades analíticas de los paquetes (por ejemplo, estimadores disponibles). Se puede encontrar una comparación detallada, por ejemplo, en el "Handbook of Structural Equation Modeling" de Hoyle (2012); para la publicación, es más importante seguir los estándares de informes.

6. El futuro de la técnica SEM dentro de las ciencias de la administración

Mediante el uso de la plataforma Scopus, se observó la tendencia positiva respecto al uso de la herramienta SEM en investigaciones relacionadas al campo de la administración; no obstante, se

identificaron estudios donde se hace uso de la técnica SEM en combinación con técnicas alternativas como redes neuronales y m-learning.

En Liébana-Cabanillas et al. (2018) se hace el uso de la técnica SEM con la técnica relacionada de redes neuronales, que también se le puede denominar Deep-learning para predecir los determinantes de la aceptación del pago móvil. Asimismo, en Alshurideh (2020) se menciona que a pesar de la gran cantidad de estudios de aceptación del m-learning, pocos han abordado la importancia de examinar el uso real de los sistemas de m-learning desde la perspectiva de la influencia social, la confirmación de expectativas y la satisfacción. Además, la mayor parte de la literatura sobre adopción de tecnología anterior tiende a utilizar la técnica de modelado de ecuaciones estructurales (SEM) para analizar los modelos estructurales. Para abordar estas limitaciones, el estudio que presenta Alshurideh (2020) amplía el modelo de aceptación de tecnología (TAM) con el modelo de confirmación de expectativas (ECM) y la influencia social para predecir el uso real de los sistemas de aprendizaje móvil. Se empleó un enfoque comparativo que utiliza el modelado de ecuaciones estructurales de mínimos cuadrados parciales (PLS-SEM) y algoritmos de aprendizaje automático para probar el modelo propuesto con datos recopilados de 448 estudiantes.

Los resultados revelaron que ambas técnicas han apoyado con éxito todas las relaciones hipotéticas del modelo de investigación. Más interesante aún, el clasificador por medio de m-learning ha funcionado mejor que los otros clasificadores en la predicción de la variable dependiente en la mayoría de los casos, por lo que se cree que el empleo de un enfoque analítico comparativo agrega una contribución significativa a la literatura de Sistemas de Información (SI) en general, y al dominio del m-learning en específico.

7. Conclusiones

En el presente documento se analizó el papel del modelado de ecuaciones estructurales en el campo de las ciencias de la administración, resaltando que, en general, dentro de las ciencias sociales, la técnica se encuentra en constante crecimiento a partir de los años 80´s. Lo anterior proveniente de los desarrollos tecnológicos, ya sea por la capacidad de hardware y software que se ha incrementado a lo largo del tiempo, permitiendo incorporar análisis más complejos con relación al número de variables.

Asimismo, se presentó el enfoque por el que se incorpora el uso de la técnica al campo de la administración, destacando a los autores más relevantes en el planteamiento del SEM y, adicionalmente, se generó una comparativa entre los diferentes algoritmos y paquetes estadísticos disponibles para la aplicación de la técnica. Encontrando en código cerrado: LISREL, EQS, AMOS, SmartPLS, Mplus, CALIS (en SAS), SEPATH (en Statistica), RAMONA en (Systat), SEM (Stata), Matlab, Semopy (en Python) y en código abierto, Scilab, Julia y en paquetes en R: lavaan, sem, lava, OpenMx, Strum, obteniendo como resultado, que todos los paquetes producen estimaciones que están bastante cerca entre sí en precisión.

La principal diferencia entre los paquetes viene en el usuario, interfaz y disponibilidad de diferentes opciones; mientras que paquetes como AMOS, LISREL y EQS se pueden utilizar en modo de programación y gráficos interfaz, el resto de los paquetes solo se pueden utilizar en un entorno de programación (R, Python, Julia).

Asimismo, se identificaron enfoques futuros que enrolan la técnica de análisis multivariante dentro de las ciencias de la administración. Una tendencia es la generación de métodos mixtos donde se incorpora la técnica SEM y a su vez, se confirma o se contrasta con técnicas relacionadas a inteligencia artificial, como m-learning o Deep-learning. Se espera que, con el avance de los estudios, se permita comparar de mejor forma las técnicas alternativas al SEM, con la finalidad obtener mejores explicaciones dentro de las ciencias sociales.

Otra tendencia en el futuro puede ser una mayor integración del software SEM, por ejemplo, en R, ajustar paquetes con entorno informático y estadístico. Dicha integración permitiría desarrollos metodológicos, que ayuden a colocar rápidamente la información en paquetes de software comerciales. Un enfoque unificado podría aprovechar de la fuerza de cada entorno.

De la misma forma se espera que se desarrollen paquetes que contengan la técnica SEM en lenguajes de programación poco convencionales, como Erlang, Scala, Rust y D Haskell, Dart, Go, F# o Perl, Elixir, Clojure, OCaml, Groovy, lo que permite establecer un manejo de los datos de formas diversas, que a su vez amplíen el universo de posibilidades relacionadas a la técnica.

Las ventajas de SEM se relacionan con la posibilidad de obtener un enfoque confirmatorio, que permite probar la teoría, además de obtener medidas de bondad de ajuste y generar un enfoque de modelado estadístico flexible. De acuerdo a Tarka (2018), el SEM puede manejar, datos faltantes (imputación múltiple), restricciones de igualdad, datos categóricos (binarios, ordinales, recuento), variables latentes discretas, continuas, datos agrupados (multinivel) y muchos otros enfoques resultan ser casos especiales (por ejemplo, lineales generalizados, como puede ser modelos mixtos.

Entre las desventajas de SEM se encuentra que la flexibilidad del modelado puede ser abrumadora, se requiere un software dedicado ya que no está disponible en paquetes estadísticos de entrada, como SPSS; asimismo "especificar" el modelo o la opción de entrada (usando software) puede ser un desafío.

Por último, los desafíos para SEM como campo estadístico, se concentran en mejorar la inferencia para muestras pequeñas, resolver el tema de los métodos robustos, ya que para valores atípicos no forman parte de la caja de herramientas estándar de SEM, mejorar los enfoques semiparamétricos y no paramétricos; asimismo, la estimación de "información completa" es computacionalmente a menudo demasiado pesada (la necesidad de integrar las variables latentes); finalmente, se requiere integrar de mejor forma la técnica SEM con las demás ramas estadísticas, como los modelos gráficos, inferencia causal y modelos mixtos.

Referencias

Aria, M. (2020). *PLS-SEM Toolbox, MATLAB Central File Exchange*. https://www.mathworks.com/matlabcentral/fileexchange/54147-pls-sem-toolbox.

Alshurideh, M., Al Kurdi, B., Salloum, S., Arpaci, I., & Al-Emran, M. (2020). Predicting the actual use of m-learning systems: a comparative approach using PLS-SEM and machine learning algorithms. *Interactive Learning Environments, 28*(4), 1-15.

Bentler, M. (2006). *EQS Structural Equations Program Manual.* Encino, CA: Multivariante Software, Inc.

Boker, S., Neale, M., Maes, H., Wilde, M., Timothy, M., Estabrook, B., Paras, B. Ross, J. Michael, G., Hunter, D., Karch, H., Pritikin, J., Zahery, M., Kirkpatrick, R., (2021). *OpenMx User Guide*. https://vipbg.vcu.edu/vipbg/OpenMx2/docs//OpenMx/latest/OpenMxUserGuide.pdf.

Campbell, S., Chancelier, J., & Nikoukhah, R. (2006). *Modeling and Simulation in SCILAB*. Springer: New York.

Driver, C., Oud, J., & Voelkle, M. (2017). Continuous time structural equation modeling with R package ctsem. *Journal of Statistical Software, 77*(5), 24-56.

Gómez, O., López, R., & Bacalla, S. (2010). Criterios de selección de metodologías de desarrollo de software. *Industrial Data, 13*(2), 70-74.

Hair, J., Gabriel, M., & Patel, V. (2014). AMOS covariance-based structural equation modeling (CB-SEM): Guidelines on its application as a marketing research tool. *Brazilian Journal of Marketing*, *13*(2), 44-55.

Haughton, D., Kamis, A., & Scholten, P. (2006). A review of three directed acyclic graphs software packages: MIM, Tetrad, and WinMine. *The American Statistician, 60*(3), 272-286.

Hoyle, R. (2012). Handbook of structural equation modeling. New York: Guilford press.

Holst, K., & Budtz-Jørgensen, E. (2013). Linear latent variable models: the lava-package. *Computational Statistics*, *28*(4), 1385-1452.

Holst, K., & Budtz-Jørgensen, E. (2019). A two-stage estimation procedure for non-linear structural equation models. *Biostatistics, 21*(4), 676-691.

Igolkina, A, & Meshcheryakov, G. (2020). Semopy: A Python Package for Structural Equation Modeling. *Structural Equation Modeling: A Multidisciplinary Journal, 27*(6), 952-963, https://doi.org/10.1080/10705511.2019.1704289

Jacobucci, R., Grimm, K., & McArdle, J. (2016). Regularized structural equation modeling. *Structural Equation Modeling: A Multidisciplinary Journal, 23*(4), 555-566.

Jöreskog, K., & Sörbom, D. (2018). *LISREL 10 for Windows* [Computer software]. Skokie, IL: Scientific Software International, Inc.

Lee, S. (2015). Implementing a simulation study using multiple software packages for structural equation modeling. *SAGE Open, 5*(3), 1-15. https://doi.org/10.1177/2158244015591823.

Liébana-Cabanillas, F., Marinkovic, V., de Luna, I., & Kalinic, Z. (2018). Predicting the determinants of mobile payment acceptance: A hybrid SEM-neural network approach. *Technological Forecasting and Social Change, 129*, 117-130.

Mair, P., Wu, E., & Bentler, P. (2010). EQS goes R: Simulations for SEM using the package REQS. *Structural Equation Modeling: A Multidisciplinary Journal, 17*(2), 333-349.

Manley, S., Hair, J., Williams, R., & McDowell, W. (2020). Essential new PLS-SEM analysis methods for your entrepreneurship analytical toolbox. *International Entrepreneurship and Management Journal, 16*(2), 1-21.

Marcoulides, G., & Falk, C. (2018). Model specification searches in structural equation modeling with R. *Structural Equation Modeling: A Multidisciplinary Journal, 25*(3), 484-491.

Merkle, E., Fitzsimmons, E., Uanhoro, J., & Goodrich, B. (2020). *Efficient Bayesian Structural Equation Modeling in Stan*. https://arxiv.org/abs/2008.07733.

Muthén, B., & Muthén, L. (2017). *Mplus, Handbook of Item Response Theory Chapman and Hall/CRC*. New York: Routledge.

Narayanan, A. (2012): A Review of Eight Software Packages for Structural Equation Modeling. *The American Statistician, 66(*2), 129-138.

Oberski, D. (2016). Mixture models: Latent profile and latent class analysis. *In Modern statistical methods for HCI*, Springer, Cham, 275-287.

Ríos, M., & Idrobo, S. (2017). Análisis comparativo de software matemático para la formación de competencias de aprendizaje en cálculo diferencial. *Plumilla educativa, 19*(1), 98-113.

Rosseel, Y. (2012). Lavaan: an R package for structural equation modeling. *Journal of Statistical Software, 48*(2), 1-36.

Segura, A., Vidal, C., & Prieto, M. (2008). Evaluación de la Calidad del Software para el Aprendizaje. In *X Simposio Internacional de Informática Educativa SIIE*, Salamanca, 59-64.

Song, Y., Stein, C., & Morris, N. (2015). Strum: an R package for structural modeling of latent variables for general pedigrees. *BMC Genomic Data*, 16(35), 1-13. https://doi.org/10.1186/s12863-015-0190-3

Sotnikova-Meleshkina, Z., & Martynenko, O. (2020). A multivariate method of evaluating the effectiveness of optimizing the daily routine of schoolchildren. *Reports of Vinnytsia National Medical University, 24*(4), 659-664.

Stapleton, L., & Leite, W. (2005). Teacher's corner: A review of syllabi for a sample of structural equation modeling courses. *Structural Equation Modeling: A Multidisciplinary Journal, 12*(4), 642-664.

Steiger, J. (2007). Understanding the limitations of global fit assessment in structural equation modeling. *Personality and Individual differences, 42*(5), 893-898.

Tarka, P. (2018). An overview of structural equation modeling: its beginnings, historical development, usefulness and controversies in the social sciences. *Quality & Quantity, 52(*1), 313-354.

Von Oertzen, T., Brandmaier, A., & Tsang, S. (2015). Structural equation modeling with Ωnyx. *Structural Equation Modeling: A Multidisciplinary Journal, 22*(1), 148-161.

Williams, N. (2020). *Toolbox for Structural Equation Modelling (SEM),* MATLAB Central File Exchange. https://www.mathworks.com/matlabcentral/fileexchange/60013-toolbox-for-structural-equation-modelling-sem.

Yuan, K., & Bentler, P. (1998). Normal theory based test statistics in structural equation modelling, *British Journal of Mathematical and Statistical Psychology, 51*(2), 289-309.

Yung, Y. (2010). Introduction to structural equation modeling using the CALIS Procedure in SAS/STAT® software, Computer technology workshop presented at the Joint Statistical Meeting on August 4, 2010, Vancouver, Canada. https://support.sas.com/rnd/app/stat/papers/JSM2010_Yung.pdf