Modelos de ecuaciones estructurales con la librería SEM de R

Antonio Matas Terrón

Universidad de Málaga (España)

Versión 2 (marzo de 2023)

Depositado en ZENODO:
Matas-Terron, Antonio. (2023, April 11). Modelos de ecuaciones estructurales con la librería SEM de R. Zenodo. https://doi.org/10.5281/zenodo.7817028

Modelos de ecuaciones estructurales con la librería SEM de R

En estas páginas se hace un somera introducción a las Ecuaciones Estructurales y cómo realizarlas en R. El texto incluye enlaces y referencias para ampliar la exposición.

Concepto de ecuaciones estructurales

Las ecuaciones estructurales fueron desarrolladas por Sewall Wright sobre 1918 (http://en.wikipedia.org/wiki/Sewall_Wright). Wright trasladó un conjunto de correlaciones a un sistema de ecuaciones, utilizando un diagrama de vías (path analysis) para representar las relaciones (http://en.wikipedia.org/wiki/Path_analysis_(statistics).

Fotografía 1: Sewall Wright (http://en.wikipedia.org/wiki/File:Sewall_Wright.jpg)

El método de Wright fue mejorado posteriormente por economistas, sociológos y psicólogos, tales como Trygve Haavelmo, Herbert A. Simon, Judea Perl, transformando el análisis de vías en el actual método de modelos de ecuaciones estructurales.

Estos modelos se aplican para analizar las relaciones causales entre variables observables, asumiendo que entre estas variables existen relaciones lineales.

Regresión lineal frente a modelos estructurales

Los modelos de ecuaciones estructurales se basan en ecuaciones de regresión que se analizan al mismo tiempo como un todo. Estos modelos permiten incluir en el modelo errores de medida, tanto en las variables independientes como dependientes. Por tal motivo, son más flexibles y menos restrictivas que las ecuaciones de regresión. Otra diferencia, es que suelen incluir variables latentes como variables causales.

Por último, la regresión lineal trata de minimizar los errores de predicción en el modelo, es decir, acercar lo más posible los valores observados y los pronosticados por el modelo. Sin embargo, los modelos estructurales lo que tratan es de minimizar la covarianzas muestrales de las covarianzas pronosticadas.

Elementos de un modelo de ecuaciones estructurales

Un modelo de ecuaciones estructurales se divide en dos submodelos, por un lado, el submodelo de medida y por otro, el submodelo de relaciones:

- - Submodelo de medida: está constituido por las variables latentes, y las variables observadas que miden cada variable latente, así como sus errores de medida y las relaciones entre todas estas variables. Un ejemplo de modelos estructurales que sólo incluyen estos modelos son los modelos del Análisis Factorial Confirmatorio.
- - Submodelo de relaciones: formado por las variables latentes que son efecto de otras variables latentes, las variables observadas que son efectos de estas variables latentes, y los errores de predicción correspondientes. Los análisis de vías son un ejemplo de modelo estructural que sólo tiene este tipo de submodelo.

Según la funcionalidad dentro del modelo, los tipos de variables son los siguientes:

- - Variables observadas: datos registrados a través de distintos instrumentos. Pueden corresponder a indicadores de constructos predictores, o constructos predichos.
- - Variable latente o constructo: concepto no observado directamente, aunque estimado a partir de las variables observadas. Inicialmente se consideran que estas variables no están afectadas por los errores de medida.
- - Variables error. Se considera que hay errores de medida en las variables observadas, así como errores por el hecho de que el modelo sea deficitario y no incluya todas las variables necesarias para que el modelo refleje fielmente el comportamiento del fenómeno estudiado.
- - Error de las predicciones: son los errores asociados a las variables dependientes.

Además de estos tipos de variables, es posible identificar dos tipos de variables teniendo en cuenta su posición dentro del modelo:

- - Variable exógena: son aquellas variables que actúan con causas o predictoras, y que no reciben efecto de ninguna otra variable. Son como las variables independientes de los modelos de regresión.
- - Variable endógena: son aquellas que reciben efectos de otras. Estas variables tienen asociado su propio error. De hecho, cualquier variable que reciba el efecto de otra incluirá un término error.

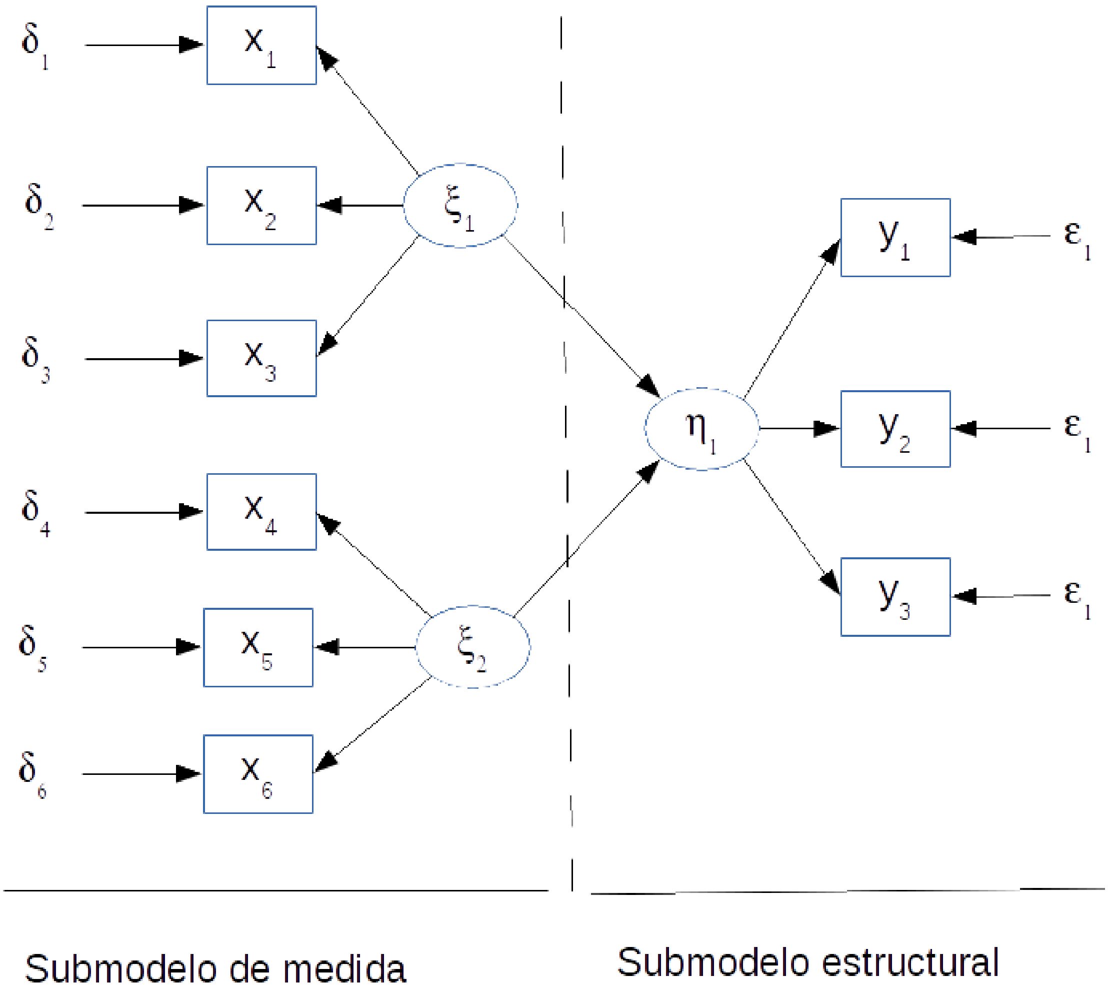


Ilustración 1: Representación de los submodelos

Representación de los diagramas de vías (path-diagrams)

Los diagramas de vías representan gráficamente el modelo estructural, incluyendo las relaciones, los parámetros a estimar, las variables observadas y los errores asociados. En las ilustraciones 1 y 2 se exponen dos ejemplos de diagramas de vías.

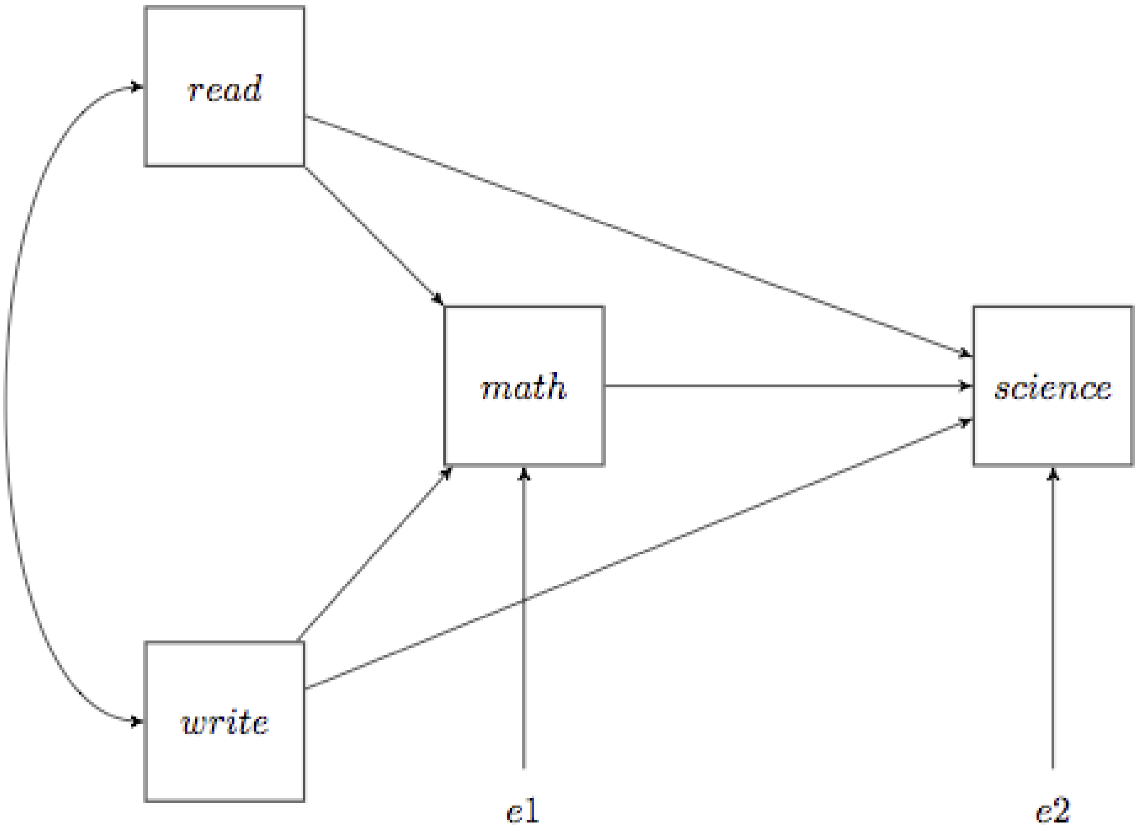


Ilustración 2: Ejemplo de diagrama de vías (http://www.ats.ucla.edu/stat/stata/faq/tikz5.png)

En el diagrama las variables observables se representan por un rectángulo, las variables latentes por una óvalos, las relaciones de causa-efecto por flechas unidireccionales rectas, y las covarianzas-correlaciones por flechas bidireccionales en forma de elipse. Los errores no se detallan con ningún tipo de figura (ilustración 3).

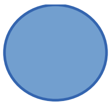
	Variable observable
	Variable latente
	Correlación o covarianza
	Efecto causal (regresión)

Ilustración 3: Resumen indicadores en los diagramas de vías

Los símbolos y letras que se utilizan para cada parte del diagrama se muestran en la tabla 1.

Tabla 1. Símbolos y letras usadas en los diagramas de vías y en las ecuaciones estructurales

Letra (griega o romana)	Nombre de la letra	Significado
η	eta	Variable latente endógena
ξ	psi	Variable exógena latente
ζ	theta	Errores en las ecuaciones, errores estructurales
x	equis	Indicadores de las variables latentes exógenas
y	y griega	Indicadores de las variables latentes endógenas
ε	épsilon	Errores de medida en los indicadores endógenos
δ	delta	Errores de medida en los indicadores exógenos
λ	Lambda	Parámetro estructural de la relación entre variable latente y variable observada.
β	Beta	Parámetro estructural de la relación entre variables latentes endógenas.
γ	Gamma	Parámetro estructural de la relación entre variables latentes exógenas y variables latentes endógenas.
ϕ	Fi	Covarianza entre las variables latentes exógenas.
ψ	Psi	Covarianzas entre los errores estructurales.
Γ	Gamma mayúscula	Matriz pxq de coeficientes que relacionan las variables latentes exógenas con la endógenas.
B	Beta mayúscula	Matriz qxq de coeficientes que relacionan las variables latentes endógenas entre sí.

En la ilustración 4 se representa un gráfico estructural con todos los símbolos incluidos.

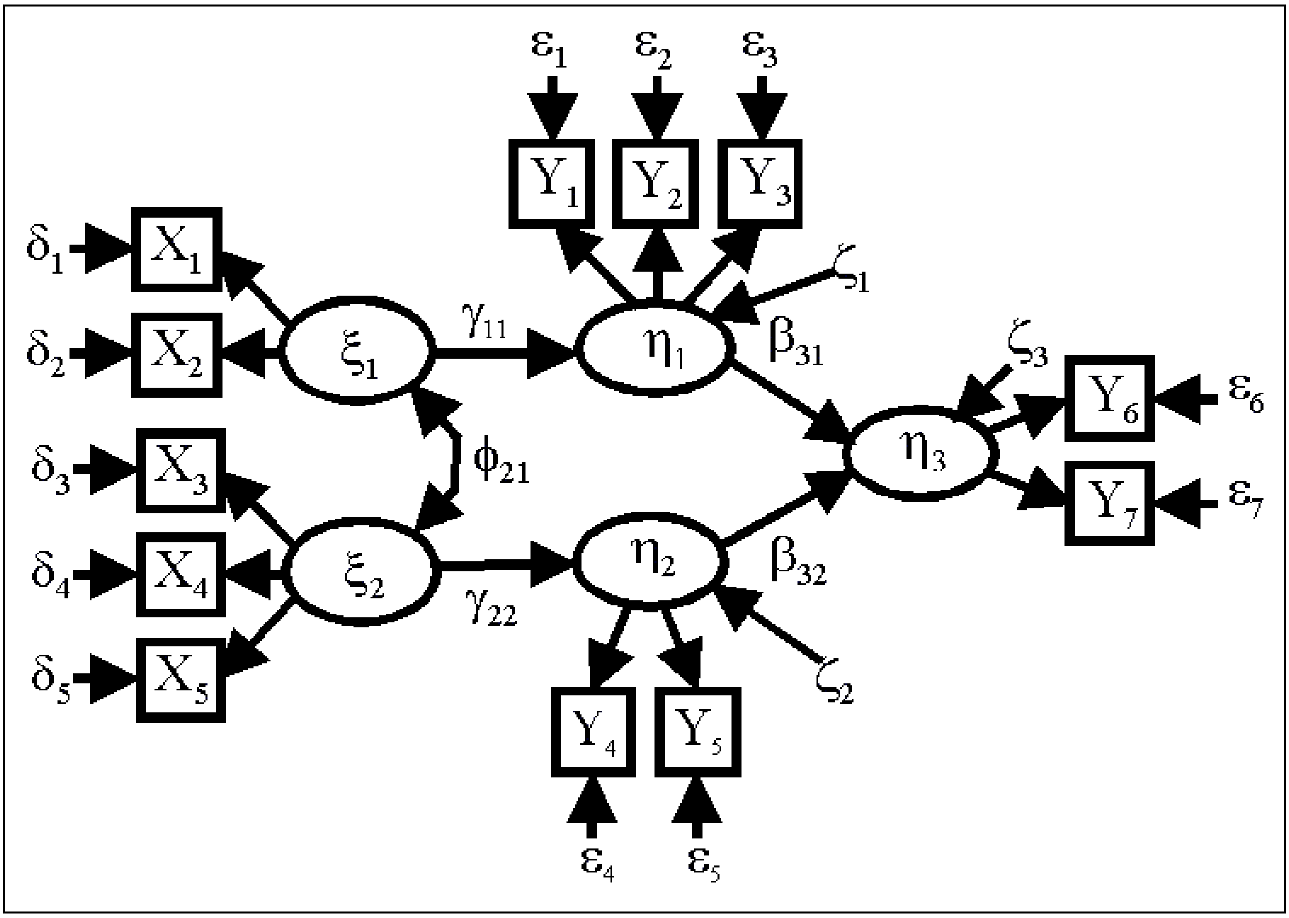


Ilustración 4. Gráfico estructural con la simbología habitual (fuente: http://www2.gsu.edu/~mkteer/sem2.html)

Tipos de modelos

En los diagramas de vías es posible identificar dos tipos de modelos:

- \- Modelos recursivos: son modelos donde 1) los errores no están relacionados entre sí, y 2) donde los efectos causales solamente van de una variable a otra, sin retroalimentación. Es decir, no hay dos variables que se relacionen entre sí de forma bidireccional.
- \- Modelos no recursivos: en estos modelos, existen lazos bidireccionales entre variable o errores correlacionados.

Inicialmente Wright asumía que en los modelos no había lazos, de forma que los diagramas no contenían ciclos.

Desarrollo de los modelos

1.- **Especificación**. La formulación del modelo se inicia realizando el gráfico correspondiente. A partir del gráfico se debe garantizar que cada parámetro se puede estimar a partir de la información de la matriz de varianzas-covarianzas generada con los datos observados.

2.- **Identificación**. Cuando los parámetros puede ser estimados a partir de esta información, se dice que el parámetro está "identificado".

3.- **Estimación**. Estimar los parámetros.

4.- **Evaluación de la calidad**. Implica evaluar la calidad del modelo obtenido a partir de los datos con relación a un modelo teórico perfecto.

En esta fase también hay que evaluar la calidad de los parámetro estimados: que los efectos sean significativos, que no tengan varianzas negativas (estimaciones impropias), que todos los parámetros incluidos sean compatibles con la simplicidad del modelo (parsimonia), etc.

5.- Mejora del modelo. En función de la evaluación del modelo, se pueden realizar modificaciones para mejorarlo. Los modelos pueden compararse entre sí a través de un simple ANOVA.

Fase de especificación

La especificación implica definir las ecuaciones estructurales. Hay tantas ecuaciones como constructos latentes explicados por las variables exógenas incluye el modelo, tanto variables exógenas observadas como latentes. Una ecuación general, considerando sólo una variable latente endógena que recibe el efecto de una variable latente exógena y de otra variable latente endógena, tiene la siguiente forma (expresión 1).

$$\eta = \beta\eta + \gamma\xi + \zeta \text{(expresión 1)}$$

Incluyendo más de una variable latente, la fórmula se puede expresar en términos matriciales (expresión 2).

$$\eta = B\eta + \Gamma\xi + \zeta \text{(expresión 2)}$$

Por otro lado, las variables observables, tanto las predictoras (x) como las predichas (y), se pueden definir en función de las variables latentes del modelo. En los siguientes ejemplos una variable predictora (x) se expresa en función de las variables latentes suponiendo que sólo existe una variable latente exógena, para la variable observable x, así como una única variable latente endógena para la variable observable y (expresiones 3 y 4).

$$x = \lambda\xi + \delta \text{(expresión 3)}$$

$$y = \lambda\eta + \varepsilon \text{(expresión 4)}$$

Las expresiones 3 y 4, cuando se tiene un conjunto de variables latentes, se pueden expresar de forma matricial (expresiones 5 y 6).

$$x = \Lambda\xi + \delta \text{(expresión 5)}$$

$$y = \Lambda\eta + \varepsilon \text{(expresión 6)}$$

Para establecer el modelo, deben darse relaciones significativas entre las variables observadas (ver cómo se hace en R).

Fase de identificación

Se dice que un modelo está identificado cuando la información que se tiene permite estimar todos sus parámetros. Para conseguir que los parámetros sean identificados, se pueden tener en cuenta los siguientes consejos:

- - Utilizar al menos tres indicadores por cada variable latente exógena.
- - Igualar la métrica de cada variable latente con uno de sus indicadores. Esto se consigue igualando a 1 la relación entre un indicador (una variable observada) y su variable latente exógena. Esta igualación es arbitraría, pudiéndose realizar en cualquiera de los indicadores.

A pesar de estas recomendaciones, es posible que el modelo no esté totalmente identificado. En tal caso, se pueden fijar más valores de los parámetros. Al hacer esto, lo que está ocurriendo es que se aumentan el número de restricciones del modelo.

Se puede saber si un modelo está identificado a través de la llamada “regla del conteo”. Está regla es la siguiente:

- Se parte del número total de parámetros (t).

- Se cuentan todos las variables del modelo, tanto endógenas como exógenas (s).

- Se decide en función de la siguiente regla:

$t>s(s+1)/2$ → modelo no identificados.

$t=s(s+1)/2$ → modelo identificado.

$t<s(s+1)/2$ → modelo sobreidentificado.

En esta fase deben tenerse claros los indicadores que se van a utilizar. Esta decisión, que puede realizarse antes del establecimiento del modelo, consiste en decidir qué instrumentos de recogida de datos se van a utilizar, así como qué tipo de transformación se aplicará a los datos directos que se obtengan.

Fase de estimación

En esta fase se pueden aplicar distintos métodos en función de las características de la muestra y de los datos. Los métodos habituales son el de máxima verosimilitud y el de mínimos cuadrados (que pueden ser mínimos cuadrados generalizados o mínimos cuadrados ponderados).

Las dos condiciones básicas para elegir entre un método u otro, son el cumplimiento o no de la normalidad en las distribuciones, así como contar con un tamaño muestral amplio.

- - Máxima verosmilitud. Las estimaciones son eficientes en condiciones de normalidad, aunque también pueden ser consistentes en violaciones leves de dicha normalidad, siempre que el tamaño muestral sea grande. No obstante, también se puede aplicar con muestras pequeñas, aunque garantizando en este caso que no se viola la normalidad. El procedimiento suele completarse con técnicas de remuestreo (como bootstrap) permitiendo determinar el intervalo de confianza para las estimaciones. Por otro lado, el remuestreo favorece el acercamiento asintótico a la normalidad.
- - Mínimos cuadrados generalizados. Se suele utilizar para muestras grandes con distribuciones normales. No obstante, es poco adecuado para modelos muy complejos.
- - Mínimos cuadrados ponderados. Se recomienda cuando no se cumplen las condiciones de normalidad. Es por tanto apropiado para datos ordinales. No obstante, es necesario contar con una muestra grande para obtener valores del Chi-cuadrado precisos. Este procedimiento no está implementado en la librería SEM, no obstante, se puede seguir un procedimiento basado en la correlación policórica (ver como se hace en R).

La comprobación de la normalidad se puede hacer a través de distintas pruebas (ver como se hace con R).

Por su parte, la comprobación del tamaño muestral se puede hacer con el índice CN de Hoeteler (ver como se hace con R).

Fase de evaluación

La evaluación se realiza en tres ámbitos: ajuste del submodelo de medidas, ajuste del submodelo estructural, y ajuste del modelo global.

Evaluación del ajuste del submodelo de medidas

Básicamente, es como evaluar la bondad en un análisis factorial. Se trata de valorar la pertinencia de las cargas de las variables observadas en las variables. De forma generalizada, aunque arbitraria, se considera que una variable debería tener al menos una carga de 0.50 (García Veiga), aunque se suele ser algo más flexible, admitiendo una carga de 0.40 en el factor para considerarla dentro de dicho factor.

En esta evaluación también se analiza la varianza explicada por el constructo. Esta varianza (R^2) indica también la fiabilidad de la medida.

Otro aspecto a analizar es la correlación que manifiestan entre sí los constructos, valorando la posibilidad de que puedan existir relaciones bidireccionales.

Evaluación del ajuste del submodelo estructural

Se trata de analizar la significación de los coeficientes estimados. Los parámetros deben ser estadísticamente significativos.

Evaluación del ajuste global

Existen tres tipos de medidas de ajuste global:

- - Medidas absolutas de ajuste: estas medidas hacen una valoración global del modelo. Los indicadores basados en Chi-cuadrado comparan el modelo observado con el modelo esperado si dicho modelo fuese perfecto. Aunque Chi-cuadrado es muy útil porque permite aplicar una prueba de significación, es muy sensible al tamaño muestra. Un elevado número muestral tiende a rechazar la hipótesis nula.
- - Medidas incrementales de ajuste: se basan en la comparación del modelo empírico con un modelo base nulo, donde las variables no están relacionadas.
- - Modelo de ajuste de la parsimonia: valoran el ajuste del modelo teniendo en cuenta el número de parámetros a estimar.

En la tabla 2 se presentan los índices más habituales, sus valores de referencia y algunas observaciones al respecto.

Tabla 2. Índices de ajuste

Índice de ajuste	**Ajuste alto**	**Ajuste aceptable**	**Observaciones**
Ajuste global			
Chi-cuadrado	0≤Chi-cuadrado≤2df	2df≤Chi-cuadrado≤3df	Chi-cuadrado es sensible al tamaño muestral. Es el único índice con test de significación. Déficits: es mejor usarlo en condiciones de normalidad; y a mayor complejidad del modelo, más posibilidad de que el modelo sea aceptado por el índice.
P-valor	.05<p≤1	.01≤p≤.05	
Chi-cuadrado/df	0≤Chi-cuadrado/df≤2	2<Chi-cuadrado/df≤3	
RMSEA	0≤RMSEA≤.05	.05≤RMSEA≤.08	Tienen en cuenta a la población.
P-valor del test	.1≤p≤1	.05≤p≤.1	
SRMR	0≤SRMR≤.05	.05≤SRMR≤.1	
RFGI			Tiene en cuenta el ajuste con la muestra
CN	CN>200	200>CN>75	Sólo cuando Chi-cuadrado es significativo y n>200
Índices de incremento			
NFI	.95≤NFI≤1	.9≤ NFI≤ .95	
NNFI	.97≤ NNFI≤ 1	.95≤ NNFI≤ .97	Índice de ajuste No Normalizado (NNFI) o Índice de Tucker-Lewis (TLI)
CFI	.97≤ CFI≤ 1	.95≤ CFI≤ .97	
GFI	.95≤ GFI≤ 1	.9≤ GFI≤ .95	
AGFI	.90≤ AGFI≤ 1	.85≤ AGFI≤ .9	
Índice de parsimonia			
AIC	Valores pequeños		
CAIC	Valores pequeños		

Modificación y comparación del modelo

Se incluyen las modificaciones en el modelo según las bondades de ajuste. El nuevo modelo se puede comparar con el anterior a través de un ANOVA.

Comprobación relaciones entre las variables

Para que un modelo tenga sentido, las variables observadas deben presentar correlaciones significativas entre sí. Estas correlaciones se pueden obtener solicitando a R la matriz de correlaciones.

```
> # Correlación entre las variables
> cor(datos)
```

O bien:

```
> cor(datos, method=”pearson”, use=”complete”)
```

Para datos no escalares se puede usar la correlación de Spearman y la de Kendall.

```
> cor(datos, method=”spearman”, use=”complete”)
> cor(datos, method=”kendall”, use=”complete”)
```

Comprobación de la normalidad

Comprobación de la normalidad de cada variable

Se examina la normalidad de cada variable. Para la normalidad de la variable debe estar medida a nivel escalar, o excepcionalmente, ser una variable ordinal con al menos 4 niveles.

También debe analizarse la asimetría y la curtosis:

asimetría >3 → asimietría extrema.

curtosis entre 8 y 20 → curtosis extrema.

curtosis > 20 → ausencia de normalidad.

En R se puede usar el test de Shapiro:

```
> shapiro.test(variable)
```

Comprobación de la normalidad multivariante

Se calcula a través de distintos test, como el de Mardia. Se considera que existen normalidad si el valor de Mardia es inferior a p(p+2), siendo p el numero de variables observables del modelo.

$$\text{Mardia} < p(p+2) \quad \text{(Expresión 7)}$$

También se puede tomar como criterio que un Mardia inferior a 70 indicaría normalidad (Rodríguez y Ruíz, 2008).

En R el procedimiento se inicia cargando la librería MVN:

```
> library(MVN)

> # Comprobación normalidad multivariante
> mardiaTest(datos, cov=T, qqplot=T)
```

Del resultado que genera la salida, se mira el gp2.

También se puede utilizar el test de Henze-Zirkler:

```
> hzTest(datos, cov=T)
```

Si cov=T (True) la matriz de covarianzas es normalizada por n. Si cov=F (False) la normalización se hace por n-1.

Adecuación del tamaño muestral

El estadistico número crítico (CN) de Hoelter (1983) es un estimador del tamaño muestral necesario para aceptar el ajuste de un modelo dado. Este indicador establece el tamaño muestral más pequeño para el cual Chi-cuadrado sería no significativo.

Hoelter consideraba como adecuados valores superiores a 200, y como indicio de un pobre ajuste valores inferiores a 75. En cualquier caso, la interpretación del índice de Hoelter sólo tiene sentido para valores muestrales n>200 al mismo tiempo que el Chi-cuadrado es estadísticamente significativo. Otros autores no recomiendan este indicador como criterio de ajuste por su sensibilidad al tamaño muestral (Hu y Bentler, 1998).

En R se puede calcular a través de la librería semGOF.

```
>summaryGOF(modelo).
```

Comprobación de la consistencia interna

Se calcula el alpha de Cronbach para cada instrumento o escala de medida, así como para cada dimensión. Se analiza si los datos son más consistentes si se elimina algún ítem. Esto se puede realizar en R con los siguientes comandos:

```
> # ANÁLISIS DE LA FIABILIDAD
> library(Rcmdr)
> fiabilidad<-reliability(cov(datos[,c("x1","x2","x3","y1","y2")], use="complete.obs"))
> fiabilidad
```

Desde otra libraría:

> library(epicalc)

> alpha(vars=c(var1, var2, var3, va4), dataFrame=Datos)

Permite obtener el mejor alfa:

> alphaBest(vars=c(var1, var2, var3, va4), standardized=F, dataFrame=Datos)

En esta libraría, el número mínimo de variables debe ser 3.

Desarrollo del análisis con R

Para datos escalares

```
> # CARGAMOS DATOS Y LIBRERÍAS NECESARIAS

> datos<-ejemplo1_datos
> library(sem)
> library(MVN)
> library(semGOF)

> # ANALISIS DESCRIPTIVO PREVIO DE LOS DATOS
> summary(datos)


> # Comprobación normalidad multivariante
> mardiaTest(datos, cov=T, qqplot=T)
> hzTest(datos, cov=T)

> # Correlación entre las variables
> cor(datos)


> # ANÁLISIS DE LA FIABILIDAD (requiere librería Rcmdr)
> fiabilidad<-reliability(cov(datos[,c("x1","x2","x3","y1","y2")], use="complete.obs"))
> fiabilidad
```

Ejemplo con la libraría Epicalc

```
> library(epicalc)
```

Loading required package: foreign

Loading required package: survival

Loading required package: splines

Loading required package: nnet

```
> alpha(vars=c(multip, comp_pali, restas, comp_arab, dictado, suma, ac_1secuencia), Datos)
```

Number of items in the scale = 7

Sample size = 78

Average inter-item correlation = 0.2434

Cronbach's alpha: cov/cor computed with 'pairwise.complete.obs'

unstandardized value = 0.6083

standardized value = 0.6925

```
Item(s) reversed:

New alpha if item omitted:
             Reversed Alpha  Std.Alpha r(item, rest)
multip          .    0.6443 0.6731    0.3487
comp_pali       .    0.6234 0.6983    0.1447
restas          .    0.5544 0.6364    0.4449
comp_arab       .    0.5529 0.6422    0.4179
dictado         .    0.5181 0.6585    0.4737
suma            .    0.5671 0.6499    0.4028
ac_1secuencia   .    0.5554 0.6473    0.4239
> alphaBest(vars=c(multip, comp_pali, restas, comp_arab, dictado, suma, ac_1secuencia), standardized=F,
dataFrame=Datos)
$best.alpha
[1] 0.6460616

$removed
 multip dictado
      1       5

$remaining
    comp_pali      restas    comp_arab         suma ac_1secuencia
            2           3            4            6             7

$items.reversed
character(0)

> # GENERACIÓN DEL MODELO
> # Se especifica el modelo

> modelo<-specifyModel("Modelo_ejemplo.txt)
        Nota: Archivo Modelo_ejemplo.txt:
        x1<-f1, gamma1, NA
        x2<-f2, gamma2, NA
        x3->f2, gamma3, NA
        f1<-f2, beta1, NA
        f1->y1, lambda1, NA
        f2->y1, lambda2, NA
        f2->y2, lambda3, NA
        x1<->x1,vx1, NA
        x2<->x2, vx2, NA
        x3<->x3, vx3, NA
        f1<->f1, vf1, NA
```

```
f2<->f2, vf2, NA
y1<->y1, vy1, NA
y2<->y2, vy2, NA
```

Nota: no hace falta especificar las variables error, ni los parámetros varianzas - covarianzas para las variables observadas exógenas, aunque en el ejemplo están descritos.

```
> # Se calcula la matriz de covarianzas
> C<-cov(datos)
```

```
> # Se genera el modelo ajustado
> modelo.fit<-sem(modelo, C, N=500)
```

```
> # Se calculan los indicadores de ajuste
> sumario<-summary(modelo.fit, conf.level=.95, fit.indices=c("RMSEA", "NFI", "GFI", "AGFI", "NNFI", "CFI", "RNI", "IFI", "SRMR", "AIC", "BIC"))
```

Se pueden calcular más indicadores de ajuste, incluyendo el CN para el tamaño de la muestra:
>summaryGOF(modelo.fit).

Se puede crear un modelo alternativo y ver cual es mejor a través de un anova:
> anova(modelo1, modelo2).

Para datos ordinales

```
> library(polycor)
> hetcor(CNES, ML=T)

Maximum-Likelihood Estimates

Correlations/Type of Correlation:
        MBSA2      MBSA7      MBSA8      MBSA9
MBSA2       1 Polychoric Polychoric Polychoric
MBSA7 -0.3028          1 Polychoric Polychoric
MBSA8  0.2826     -0.344          1 Polychoric
MBSA9 -0.2229     0.5469    -0.3213          1

Standard Errors:
        MBSA2   MBSA7   MBSA8
MBSA2
MBSA7 0.02737
MBSA8 0.02773 0.02642
MBSA9 0.02901 0.02193 0.02742

n = 1529

P-values for Tests of Bivariate Normality:
          MBSA2     MBSA7   MBSA8
MBSA2
MBSA7 1.277e-07
MBSA8 1.852e-07 2.631e-23
MBSA9 5.085e-09 2.356e-10 1.5e-19
```

```
>
> hcor<-function(data) hetcor(data, std.err=F)$correlations
> R.CNES<-hcor(CNES)
> R.CNES
           MBSA2      MBSA7      MBSA8      MBSA9
MBSA2  1.0000000 -0.3017953  0.2820608 -0.2230010
MBSA7 -0.3017953  1.0000000 -0.3422176  0.5449886
MBSA8  0.2820608 -0.3422176  1.0000000 -0.3206524
MBSA9 -0.2230010  0.5449886 -0.3206524  1.0000000
>
>
> model.CNES<-specifyModel()
1: F->MBSA2, lam1, NA
2: F->MBSA7, lam2, NA
3: F->MBSA8, lam3, NA
4: F->MBSA9, lam4, NA
5: F<->F, NA, 1
6: MBSA2<->MBSA2, the1, NA
7: MBSA7<->MBSA7, the2, NA
8: MBSA8<->MBSA8, the3, NA
9: MBSA9<->MBSA9, the4, NA
10:
Read 9 records
```

O bien: specifyModel(“modelo”)

```
> sem.CNES<-sem(model.CNES, R.CNES, N=1529)
> summary(sem.CNES)

 Model Chisquare =  33.2115   Df =  2 Pr(>Chisq) = 6.14066e-08
 AIC =  49.2115
 BIC =  18.54676

 Normalized Residuals
      Min.   1st Qu.    Median      Mean   3rd Qu.      Max.
-0.000003  0.030010  0.207800  0.847900  1.035000  3.830000

 R-square for Endogenous Variables
 MBSA2  MBSA7  MBSA8  MBSA9
0.1516 0.6052 0.2197 0.4717

 Parameter Estimates
      Estimate  Std Error  z value   Pr(>|z|)
lam1 -0.3893289 0.02875484 -13.53959  9.129470e-42 MBSA2 <--- F
lam2  0.7779157 0.02996521  25.96063 1.379394e-148 MBSA7 <--- F
lam3 -0.4686834 0.02839946 -16.50325  3.476850e-61 MBSA8 <--- F
lam4  0.6867992 0.02921502  23.50842 3.344853e-122 MBSA9 <--- F
the1  0.8484230 0.03281417  25.85539 2.116323e-147 MBSA2 <--> MBSA2
the2  0.3948472 0.03567529  11.06781  1.797436e-28 MBSA7 <--> MBSA7
the3  0.7803360 0.03152466  24.75319 2.864281e-135 MBSA8 <--> MBSA8
the4  0.5283069 0.03212698  16.44434  9.208259e-61 MBSA9 <--> MBSA9

 Iterations =  14

> system.time(boot.cnes<-bootSem(sem.CNES, R=1000, Cov=hcor, data=CNES))
   user  system elapsed
 54.431   2.237  55.994
```

```
> summary(boot.cnes, type="norm")
Call: bootSem(model = sem.cnes, R = 1000, Cov = hcor, data = CNES)

Lower and upper limits are for the 95 percent norm confidence interval

       Estimate          Bias  Std.Error      Lower      Upper
lam1 -0.3893289  3.685841e-03 0.03268788 -0.4570818 -0.3289476
lam2  0.7779157  3.467082e-04 0.03235698  0.7141504  0.8409875
lam3 -0.4686834  1.144803e-03 0.03277752 -0.5340709 -0.4055854
lam4  0.6867992  9.161340e-04 0.03036848  0.6263619  0.7454042
the1  0.8484230  1.798709e-03 0.02516248  0.7973067  0.8959418
the2  0.3948472 -1.576024e-03 0.04987111  0.2986777  0.4941688
the3  0.7803360  8.043225e-06 0.03094814  0.7196707  0.8409852
the4  0.5283069 -2.172296e-03 0.04160811  0.4489288  0.6120296
```

Otros softwares

Con JAMOVI: https://semlj.github.io/example2.html#Research_data

Referencias

Hoelter, J.W. (1983). The analysis of covarinace structures: Goodnes-of-fit indices. *Sociological Methods and Research, 11*, 325-334.

Hu, L., y Bentler, P.M. (1998). Fit indices in covarinacce structure modeling: sensitivity to underparmetrizaed model misspecification. *Psychological Methods*, *3*(4), 424-453.

García Veiga, M.A. (2011). *Análisis causal con ecuaciones estructurales de la satisfacción ciudadana con los servicios municipales (Proyecto Fin de Máster)*. Santiago de Compostela. Recuperado de http://eio.usc.es/pub/mte/descargas/proyectosfinmaster/proyecto_610.pdf

Rodríguez-Ayán, M. N. y Ruiz-Díaz, M. A., (2008). Atenuación de la asimetría y de la curtosis de las puntuaciones observadas mediante transformaciones de variables: Incidencia sobre la estructura factorial, *Psicológica, 29,* 205-227.