Clase 8: Estimación de confiabilidad con Omega

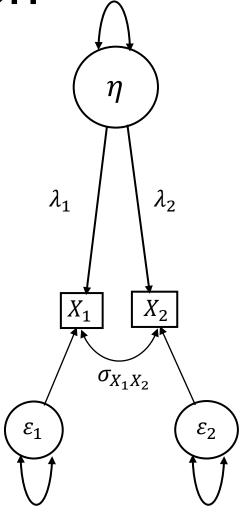
Dr. Héctor Nájera PUED-UNAM



Confiabilidad y modelos de medición

 Hemos visto que todas las mediciones deben de tener un modelo (Teoría y modelo estadístico)

- Tres modelos (reflexivos) estadísticos útiles:
 - Test paralelos
 - Equivalencia Tau
 - Modelos congéneres y variantes multidimensionales

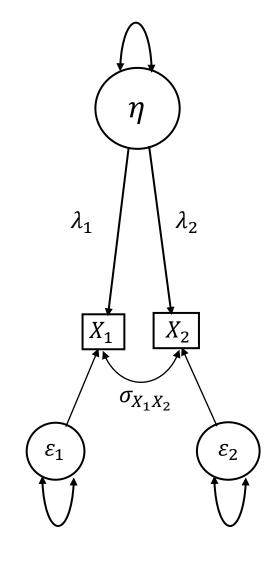




Modelos de medición razonables y testables

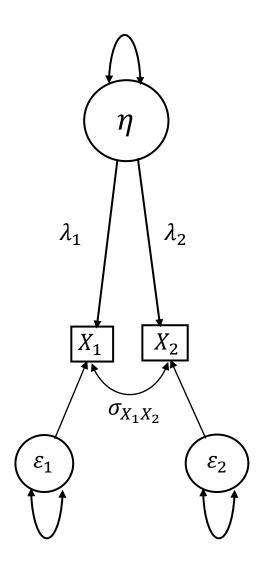
• Si mi modelo es una mala aproximación de la realidad - mundo observable- ya ni siquiera vale la pena estimar confiabilidad

 Desafortunadamente el modelo estándar de la teoría clásica del test durante mucho tiempo no era testeable





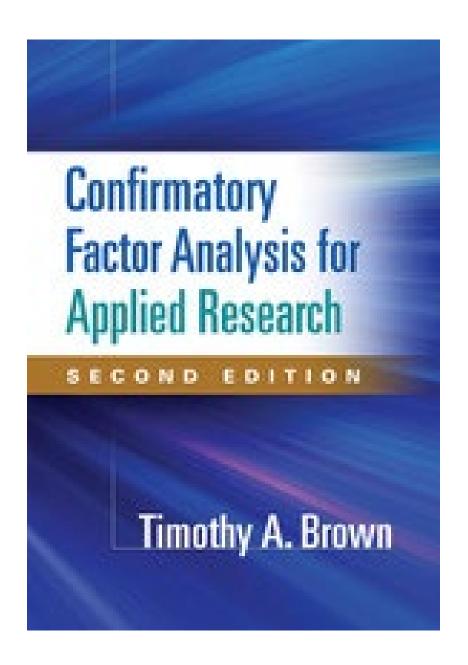
SEM: Modelos de medición testeables



SEM permitió la estimación de los tres modelos estadísticos de la teoría de la medición de variables latentes

Para ello se valió de un método en particular: Análisis factorial

- ¿Se sostiene el modelo de medición dados los datos si puedo reproducir el mecanismo generador de datos-?
- ¿Los indicadores capturan la señal de interés? ¿En qué cantidad?





Análisis factorial: Introducción

El análisis factorial es un método o caso especial de SEM

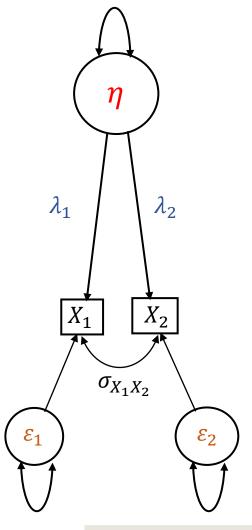
Es un método general de inferencia de varios tipos (populares) de modelos de medición: <u>inferir de los datos</u> si los indicadores x_i son producto de un factor latente η

Ecuación estructural de medición general.

$$\rightarrow$$

$$x_i = \lambda_i \eta + \varepsilon_i$$

En este caso λ_i se conocen como cargas factoriales (factor loadings) y capturan la relación entre las variables latentes (η) y los indicadores observados, y entre las dimensiones y el factor de mayor orden.

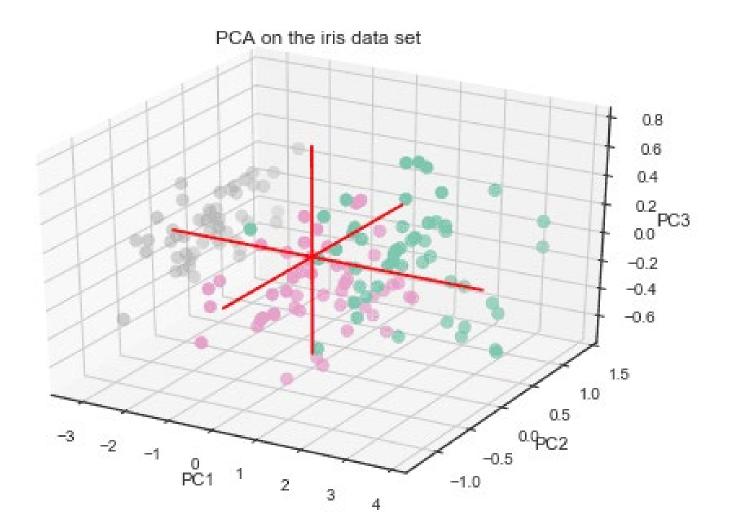




$$x_i = \lambda_i \eta + \varepsilon_i$$

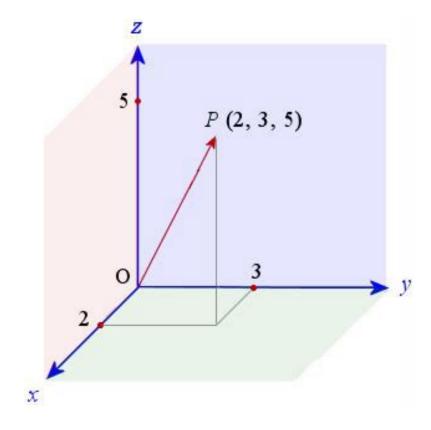
$$x_i$$

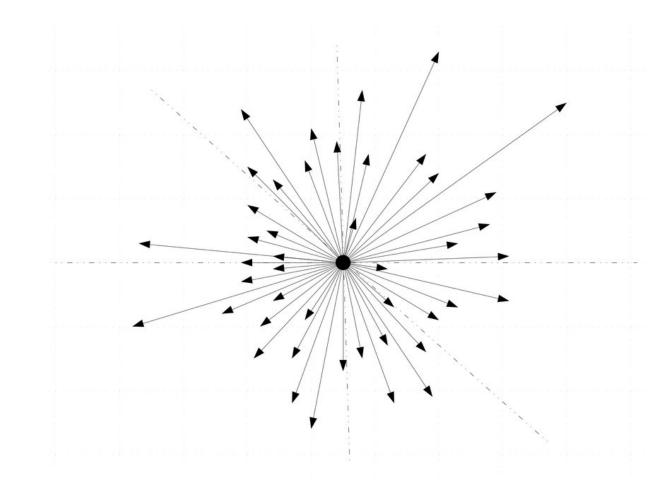






$$x_i = \lambda_i \eta + \varepsilon_i$$







Dos caminos en el análisis factorial

Análisis factorial confirmatorio: Examinar si mi modelo explícito reproduce las relaciones observadas

Modelo← explícito de medición

Análisis factorial exploratorio: Inferir a partir de los datos si existe evidencia de que los indicadores disponibles capturan una misma señal -factor-

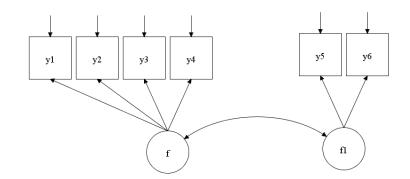
Sin modelo← explícito de medición

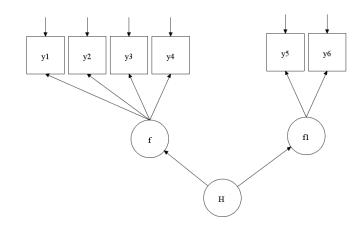
Los que no se vale es hacer un exploratorio y luego confirmatorio con los mismos datos.



Pasos de la estimación de confiabilidad el análisis confirmatorio (SEM)

- 1. Modelo de medición
- 2. Modelo estadístico
 - 1. Hipótesis pasadas a parámetros
- 3. Identificación estadística del modelo
- 4. Estimación
- 5. Inferencia
- 6. Cálculo de omega

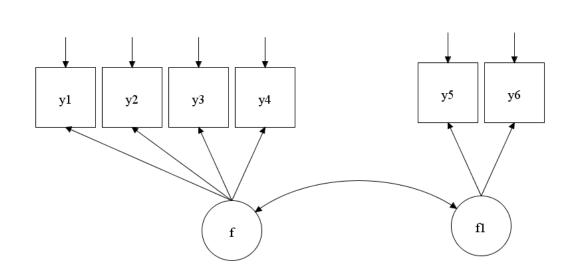


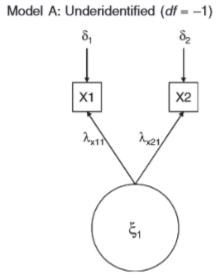




Identificación estadística del modelo

¿Cuántas variables observadas? ¿Cuántas incógnitas/parámetros?





Input Matrix (3 elements)

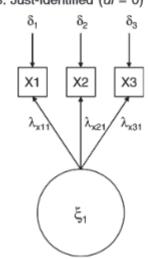
$$\begin{array}{cccc} & & X1 & & X2 \\ X1 & & \sigma_{\,11} & & \\ X2 & & \sigma_{\,21} & & \sigma_{\,22} \end{array}$$

Freely Estimated Model Parameters = 4 (e.g., 2 factor loadings, 2 error variances)



Identificación estadística del modelo

Model B: Just-Identified (df = 0)

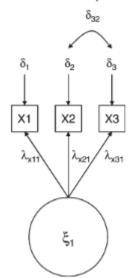


Input Matrix (6 elements)

$$\begin{array}{ccccc} & X1 & X2 & X3 \\ X1 & \sigma_{11} & & \\ X2 & \sigma_{21} & \sigma_{22} & \\ X3 & \sigma_{31} & \sigma_{32} & \sigma_{33} \end{array}$$

Freely Estimated Model Parameters = 6 (e.g., 3 factor loadings, 3 error variances)

Model C: Underidentified (df = -1)



Input Matrix (6 elements)

	X1	X2	X3
X1	σ 11		
X2	σ 21	σ_{22}	
Х3	σ ₃₁	σ_{32}	σз

Freely Estimated Model Parameters = 7 (e.g., 3 factor loadings, 3 error variances, 1 error covariance)

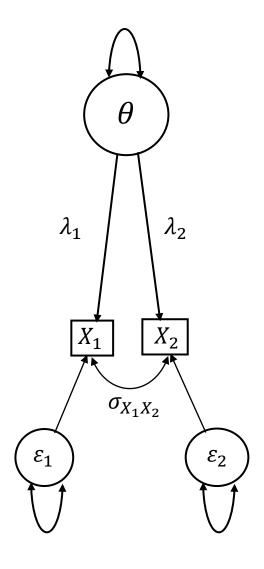


Estimación del modelo

- ¿Qué algoritmo? Máxima Verosimilitud, WLS -Mínimos cuadrados ponderados- Bayes (MCMC)
- Depende del número de variables latentes y tamaño de muestra
 - Modelo sencillo: Todas variables continuas, grados positivos de libertad, no más de dos variables latentes y hasta 100,000. Máxima Verosimilitud
 - Modelo complejo: Variables categóricas o mixtas, grados positivos de libertad y más de dos variables latentes y hasta 100,000. WLS o Bayes
- Escritura del modelo: Siguiendo las reglas de ecuaciones estructurales



SEM: Modelos de medición testeables

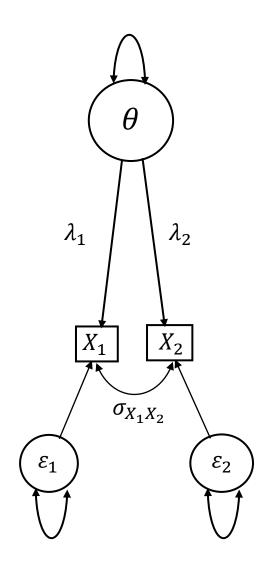


SEM permitió la estimación de los tres modelos estadísticos de la teoría de la medición de variables latentes

Para ellos se valió de un método en particular: Análisis factorial

- ¿Se sostiene el modelo de medición dados los datos -si puedo reproducir el mecanismo generador de datos-?
- ¿Los indicadores capturan la señal de interés? ¿En qué cantidad?

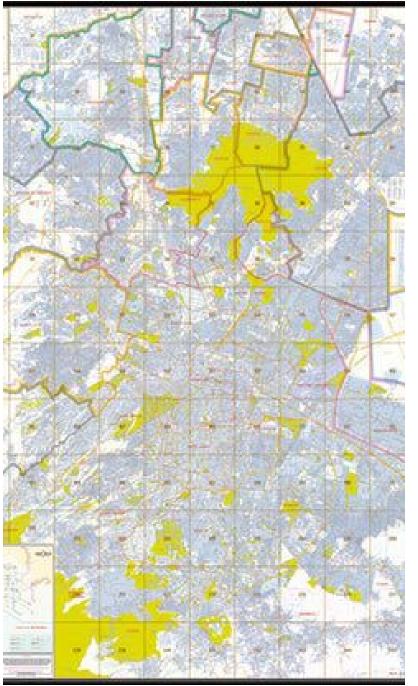




¿Se sostiene el modelo de medición dados los datos?



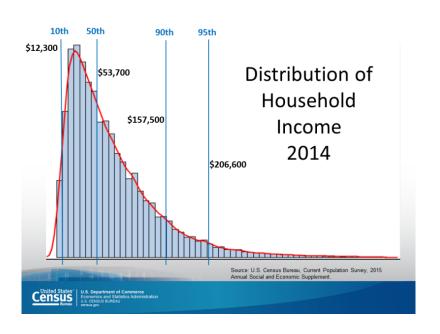




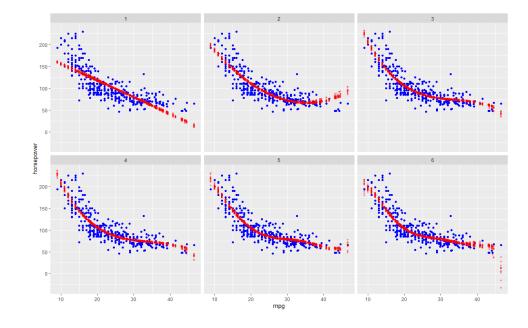


Tialten Tia

¿Se sostiene el modelo de medición dados los datos?



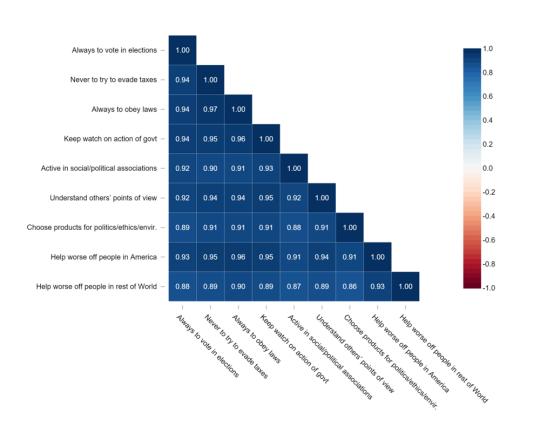
$$Y_i = \alpha + \beta_i X_i + \epsilon_i$$



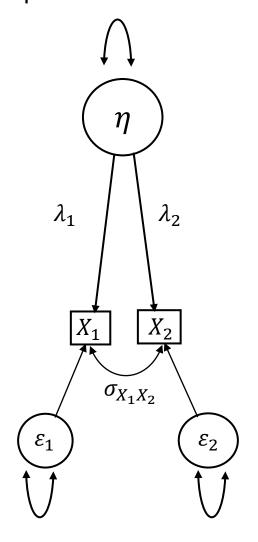


¿Se sostiene el modelo de medición dados los datos?

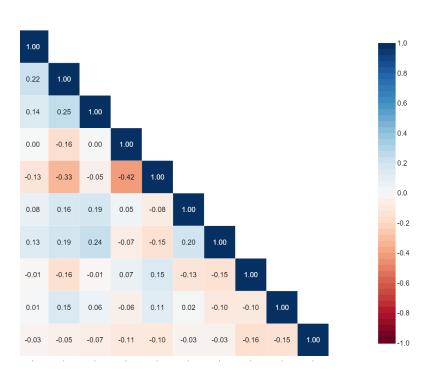
Correlaciones observadas



Supuestos del modelo



Lo que predice el modelo



Si la predicción es mala, entonces el modelo no sirve

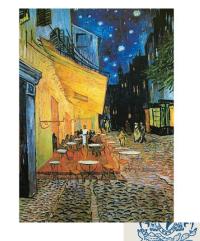


Evaluación formal y global de un SEM

Los CFA tienen una serie de estadísticos:

- La estimación de un CFA maximiza una función F_{ML} (qué tan bien ajusta el modelo a los datos)
- Con esta estimación se puede calcular un estadístico χ^2 cuya hipótesis nula es que $S=\Sigma$. Es decir, que si p<.05 rechazamos que el modelo sirve. S. En otras palabras, el número, tipo de dimensiones e indicadores no son una representación del construct.
- Un problema es que $\chi^2 = F_{ML}(N-1)$ y por tanto es sensible a N .
- Por esta razón hay otros estadísticos. Estos índices son relativos y comparan un modelo nulo con el modelo propuesto. Existen estadisticos como el Comparative Fit Index (CFI) o el Tucker-Lewis index (TLI) ambos varian entre cero y uno.







- Checar ajuste global del modelo
 - RMSEA, TLI, CFI, Chi-Square

Criterios para evaluar un CFA

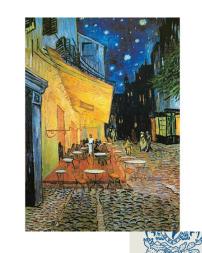
Summary of the suggested cut off for the goodness-of-fit statistics. The values of RMSEA, CFI and TLI need to be taken as an approximation.

	1.1	
Index	Range values	Poor model fit rule

χ^2	p-values 1-0	p < .05
RMSEA	p-values 1-0	p > .06
CFI	1-0	< .95
TLI	1 - 0	< .95



VS





Pasos adicionales

- Re-ajustar el modelo teórico
 - Correlaciones entre algunas variables -dada la teoría-
- Re-estimar el modelo

En ocasiones el modelo puede estar "ligeramente mal". TLI=.93, CFI=.92, RMSEA=.07

Debido a que comparamos la matriz observada y la estimada por el modelo, Podemos calcular las fuentes de la discrepancia: índices de modificación

Este es un tema de validez pero es importante tenerlo presente.

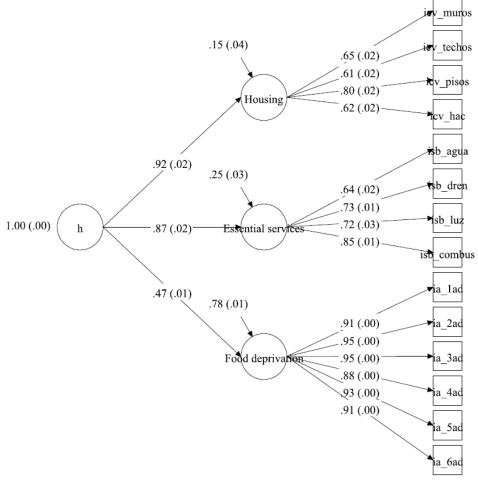


Análisis factorial confirmatorio 1.1

Se propone que la pobreza tiene tres dimensiones

Se propone que hay cuatro indicadores observados que son una manifestación de Vivienda

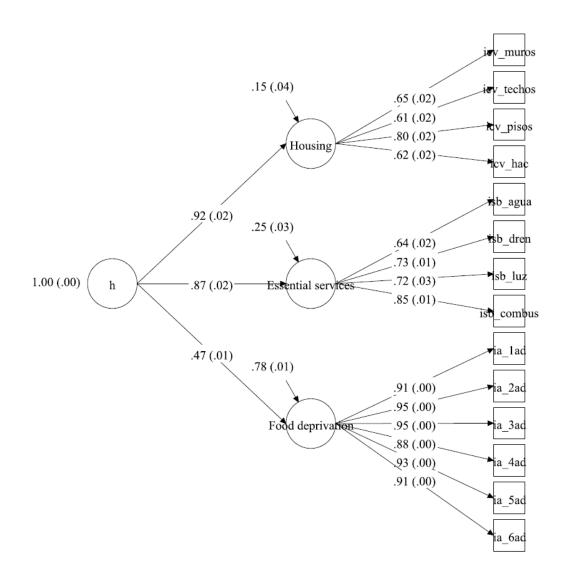
Se propone que hay seis indicadores observados que son manifestación de inseguridad alimentaria





PUED

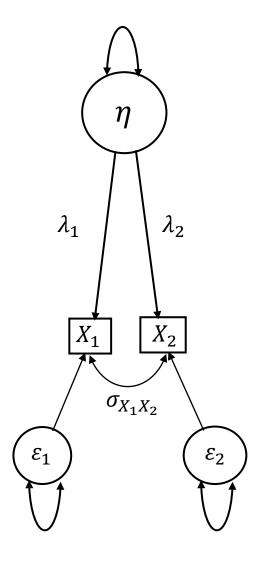
Dado un modelo razonable







SEM: Modelos de medición testeables



SEM permitió la estimación de los tres modelos estadísticos de la teoría de la medición de variables latentes

Para ellos se valió de un método en particular: Análisis factorial

- ¿Se sostiene el modelo de medición dados los datos -si puedo reproducir el mecanismo generador de datos-?
- ¿Los indicadores capturan la señal de interés? ¿En qué cantidad?



Inferencia en análisis factorial

 $x_i = \lambda_i \eta + \varepsilon_i$

Parto de que la estructura es unidimensional (Solo hay un factor latente)

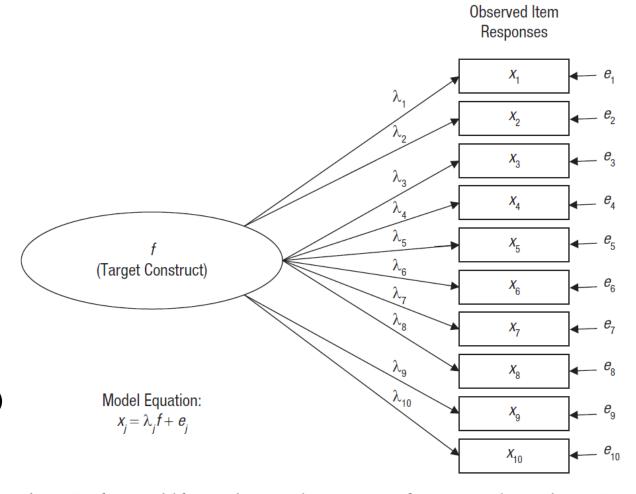


Fig. 1. One-factor model for a unidimensional test consisting of 10 continuously scored items. See the text for further explanation.

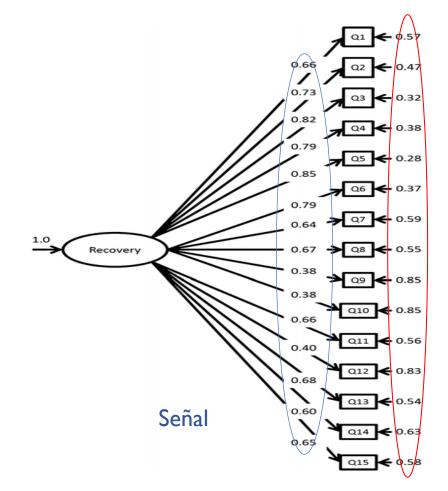


Continuous

Interpretación de las cargas factoriales

- Relaciones teóricas vs observadas
- Proporción explicada de la varianza de cada indicador

Noten que estos modelos me ayudan a saber de una si equivalencia tau se sostiene



Tenemos que Q1 el 0.66² = .43 es la varianza explicada por "Recovery" con un 0.57 de error.

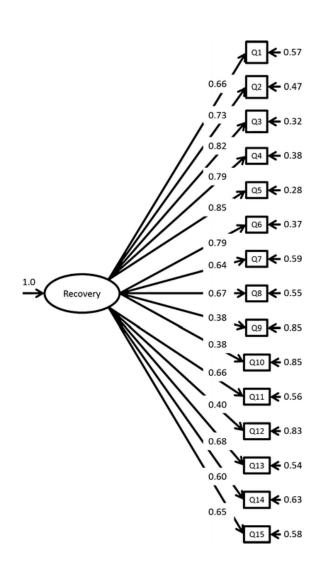
¿Qué proporción de la varianza es deseable?

Por estudios de Monte Carlo se estima que arriba de .4 o .5 es deseable. Esto depende de la precisión que se busque.

Ruido!



¿Cargas factoriales bajas?



Imagínense que con estas variables "omega" es alto.

Sin embargo, Q9 y Q10 no me ayudarían de mucho, tienen más ruido que señal.

¿Qué hacer?

Tirarlas

Dejarlas

Ponderar

¿Qué les parecería apropiado?

Además de ayudarnos a la estimación de confiabilidad global, el AFC nos permite identificar si debemos o no ponderar

PUED

Con alta confiabilidad ponderar no tiene sentido.

AFC y la inferencia sobre las cargas factoriales

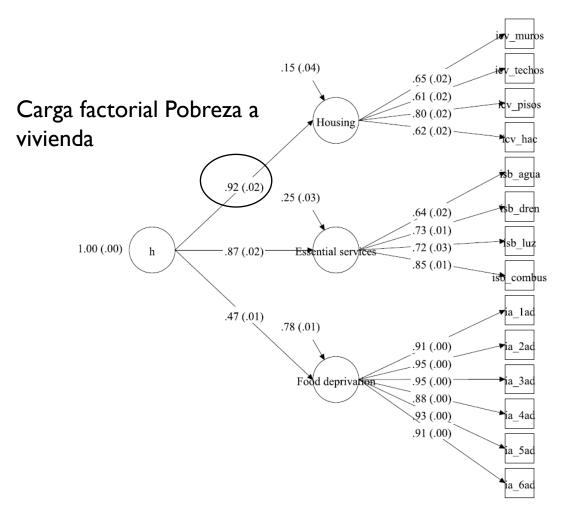
¿Què significan los números?

Los MFC estiman soluciones estandarizadas, parcialmente estandarizadas y no estandarizadas

Para confiabilidad es mejor utilizar resultados estandarizados: La varianza del factor es I

 $0.92^2 = 85\%$ de la varianza de la dimensión se explica por la variable latente

1 - .85 = 15% de la varianza de la dimensión se explica por error





PUED

AFC y exploración de cargas factoriales

¿Qué significan los números?

El efecto indirecto de pobreza en carencia de muros es:

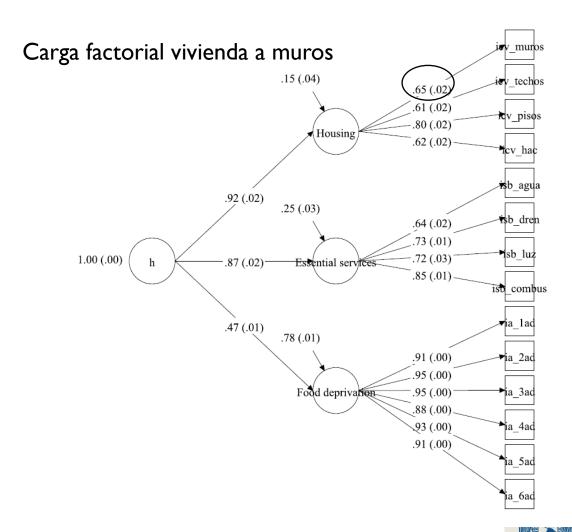
$$0.92 * 0.65 = 0.60$$

Es decir, el efecto está "mediado" por la dimensión de vivienda

 0.6^2 es la varianza explicada por la dimensión vivienda condicional en el factor general

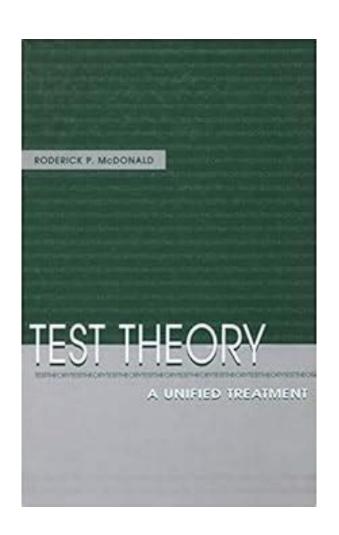
¿Pero qué si queremos saber el efecto directo?

Veremos que esto es importante para estimar Omega



PUED





Estadísticos de confiabilidad: Omega

- Los mejores estadísticos de confiabilidad son ω and ω_h
- \bullet es conocida como el techo de confiabilidad, i.e. valor más alto estimado.
- Calcula la proporción de la varianza de todos los indicadores que se explica por el factor.
- Tiene un valor máximo de 1 y un mínimo de 0.
- Valores > .8 son deseables. Explicaremos por qué.

$$\omega = rac{\sum\limits_{j=1}^k \left(\sum\limits_{i=1}^p \lambda_{ij}
ight)^2}{\sum\limits_{j=1}^k \left(\sum\limits_{i=1}^p \lambda_{ij}
ight)^2 + \sum\limits_{i=1}^p e_i}$$



Clasificación de población y confiabilidad

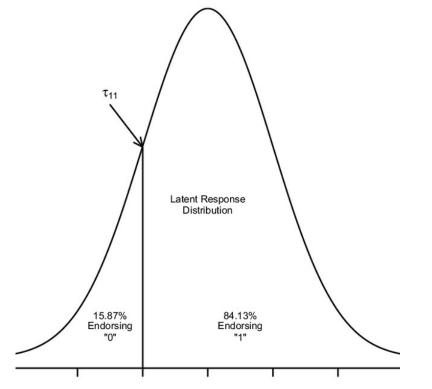
- Una pregunta es: Qué tan bajo es el estadístico de confiabilidad para ser aceptable?
- Una de las consecuencias de confiabilidad es que lleva a ordenamientos de la población consistentes
- Nájera (2018) hizo un estudio de Monte Carlo para ver la relación entre confiabilidad y entropía: Unidimensional,
 Multidimensional y multidimensional (débil).
- Hay una relación muy clara en los tres casos

Summary of the relations among β , α , ω and entropy depending on index dimensionality. Summarised from Nájera (2018). In this case, the unidimensional model seem to meet τ equivalence, i.e. equal loadings.

Reliability statistic	Leads to	lassification error (%)	Entropy value > .8		
$\alpha > .8$	\approx	< 5%			
$\omega > .8$	\approx	< 5%	> .8		
$\omega > .85$	\approx	< 5%	> .8		
$\omega_h > .65$	\approx	< 5%	> .8		
$\omega > .85$	\approx	< 5%	> .8		
$\omega_h > .70$	\approx	< 5%	> .8		
			750		

Scores latentes y observados

AFC estima una distribución a partir de las cargas factoriales estimadas- score óptimo/latente



Con alta confiabilidad

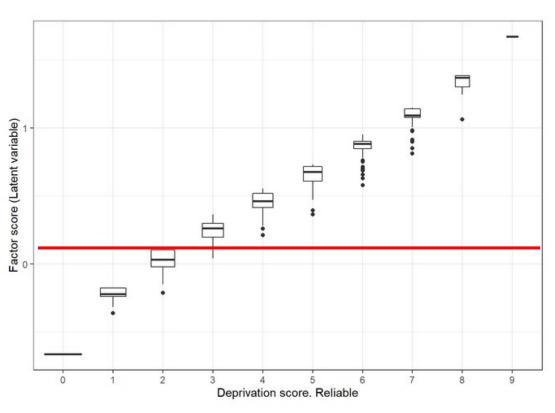
¿Cuál sería la relación entre los scores observados y los latentes?

##		Lavadora	Televisi	ón Agua	ent	ubada	Ref	rigera	ador	GasNati	ıral	Drenaj	e
##	1	0		0		0			0		0		0
##	2	0		0		0			0		0		0
##	3	0		0		0			0		0		0
##	4	0		0		0			0		0		0
##	5	0		0		0			0		0		0
##	6	0		0		0			0		0		0
##		Cisterna	Celular (Computad	lora	Netf	lix	Panel	Sola	r Jard	in E	stufa	
##	1	0	0		0)	0			0	0	0	
##	2	0	0		0)	0			0	0	0	
##	3	0	0		0)	0			0	0	0	
##	4	0	0		0)	0			0	0	0	
##	5	0	0		0)	0			0	0	1	
##	6	0	0		0		0			0	0	1	
##		AireAcon	dicionado	Bicicle	ta	Obs		Exp		z1		se.zl	
##	1		0		0	260 3	08.5	83613	-1.3	723183	0.6	710940	
##	2		0		1	78	26.8	15590	-0.9	565429	0.6	137046	
##	3		1		0	80	25.7	40288	-0.9	565429	0.6	137046	
##	4		1		1	31	3.3	39630	-0.6	034440	0.5	701628	
##	5		0		0	100	27.5	92368	-0.9	565429	0.6	137046	
##	6		0		1	32	3.5	79925	-0.6	034440	0.5	701628	

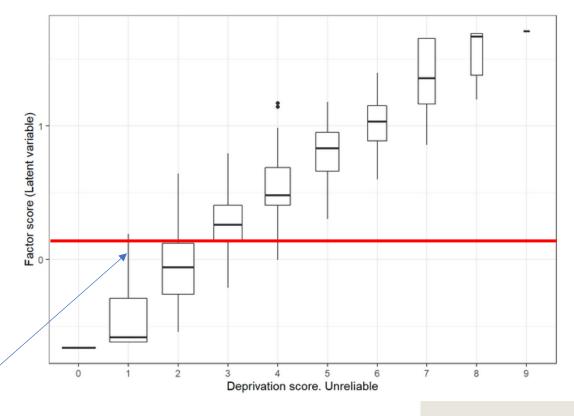


Correlación entre distintas versiones del mismo test

Alta confiabilidad



Moderada confiabilidad



Más error en la clasificación



Ejemplo: Cálculo de Omega AFC

- El cálculo de omega consiste en los siguientes pasos:
 - Estimo modelo factorial (Confirmatorio de preferencia)
 - Evalúo el ajuste del modelo
 - Estimo omega con la fórmula de McDonald

Leer la documentación

avevar 0.2461983

```
library(lavaan)
                                     + sinedbasica + accesosalud
                                            sinlavadora + sinrefrigerador'
I1 <- sem(model,data=D[,3:13],std.lv=TRUE)</pre>
 > reliability(fit1f)
 This call to reliability produces the following output:
   openness
                                                                Variables
 alpha 0.5999111
                                                                categóricas se
 omega 0.6079033
                                                                requiere
 omega2 0.6079033
                                                                especificar la
 omega3 0.6078732
                                                                opción "ordinal"
```



Distintos tipos de omega

- El omega relevante depende de la estructura y tipo de indicadores que se tiene:
 - Multidimensional o unidimensional
 - Categóricos o continuos (O una mezcla)



- Se tienen cinco indicadores: O_1 ... O_5
- Se estima un modelo unidimensional
- TLI=.88

The factor-loading estimates of the fit1f model are listed under the Latent Variables heading in the results summary as follows:

```
Latent Variables:

Estimate Std.Err z-value P(>|z|)

openness =~

O1 0.622 0.029 21.536 0.000

O2 0.684 0.042 16.466 0.000

O3 0.794 0.032 24.572 0.000

O4 0.361 0.031 11.779 0.000

O5 0.685 0.036 19.069 0.000
```



This call to reliability produces the following output:

openness

alpha 0.5999111

omega 0.6079033

omega2 0.6079033

omega3 0.6078732

avevar 0.2461983

Average extracted variance: Items not factors

Omega ((Bollen, 1980; see also Raykov, 2001) y omega2 (Bentler, 1972, 2009) se calculan utilizando la varianza "supuesta" bajo el modelo -casi siempre =1-

Omega3 se basa en la varianza observada en los datos (más conservador)

.61 es la proporción del score total que es atribuible al factor -61% de señal-

The first and the second coefficients omega will have the same value when the model has simple structure, but different values when there are (for example) cross-loadings or method factors. The first coefficient omega can be viewed as the reliability controlling for the other factors. The second coefficient omega can be viewed as the unconditional reliability (correlated errors).

This call to reliability produces the following output:

openness

alpha 0.5999111

omega 0.6079033

omega2 0.6079033

omega3 0.6078732

avevar 0.2461983

Omega ((Bollen, 1980; see also Raykov, 2001) y omega2 (Bentler, 1972, 2009) se calculan utilizando la varianza "supuesta" bajo el modelo -casi siempre =1-

Omega3 se basa en la varianza observada en los datos

.61 es la proporción del score total que es atribuible al factor -61% de señal-

Yang and Green (2015) asserted that applied researchers should be more interested in the reliability of an observed total score X than in the reliability of a latent total score X^* because observed scores, rather than latent scores, are most frequently used to differentiate among individuals in research and practice. Yang and Green established that, compared with ω_u , ω_{u-cat} produces more accurate reliability



Además de cargas factoriales, tenemos un parámetro adicional: El umbral

El umbral es el valor latente a partir del cual una persona cambia de la categoría 0 a 1

Por ejemplo, esperaríamos que las personas con carencia de agua entubada tuvieran el mismo umbral - severidad latente-

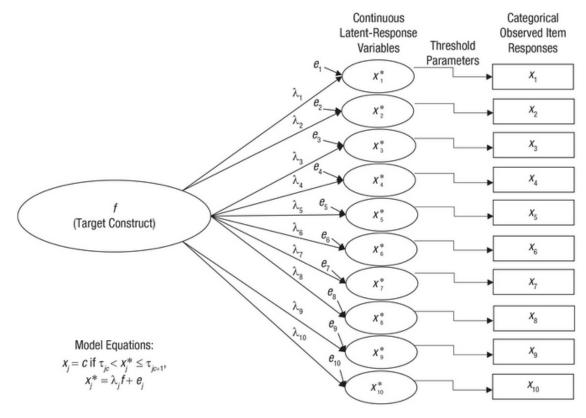


Fig. 2. One-factor model for a unidimensional test consisting of 10 ordinally scaled items. See the text for further explanation.



> reliability(fit1f)

This call to reliability produces the following output:

psyctcsm

alpha 0.8007496

omega 0 70020E2

omega2 0.7902953

omega3 0.7932682

avevar 0.5289638

Alpha es un Alpha Ordinal!

Omega y Omega2 usan la varianza que supone el modelo del factor

Omega3 se basa en la varianza observada en los datos



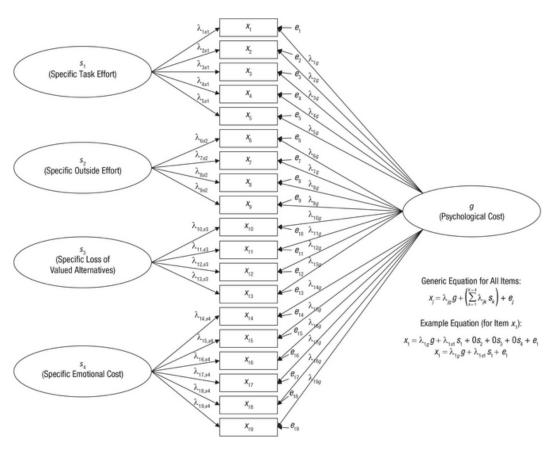
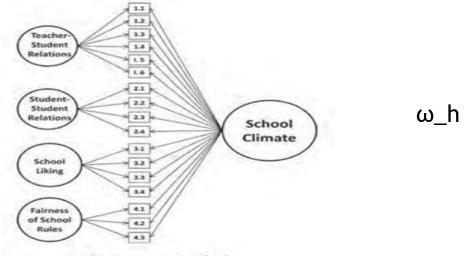


Fig. 3. Bifactor model for the psychological-cost scale. See the text for further explanation.

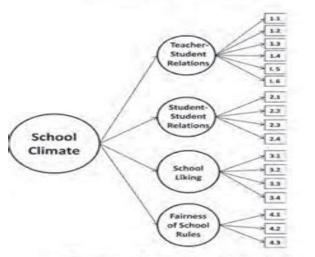
Nos interesa el omega total y el omega jerárquico:

La confiabilidad atribuible a las dimensiones





Bifactor Model



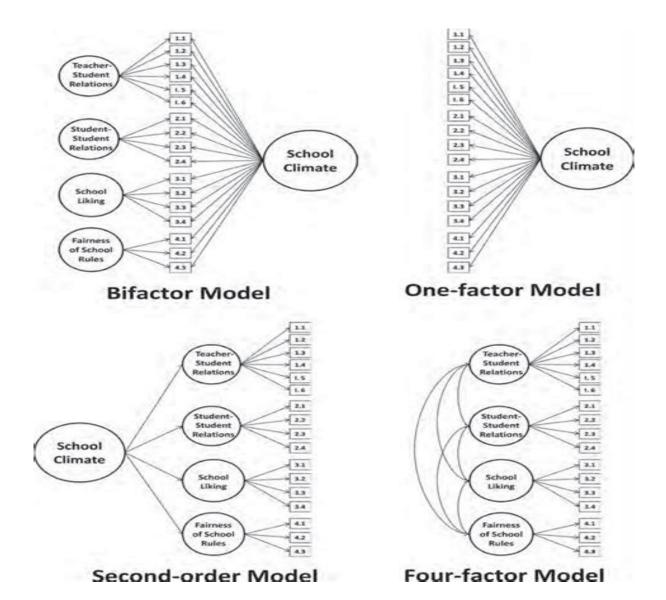
 ω_{ho}

% del score global que se atribuye al factor general o de alto orden

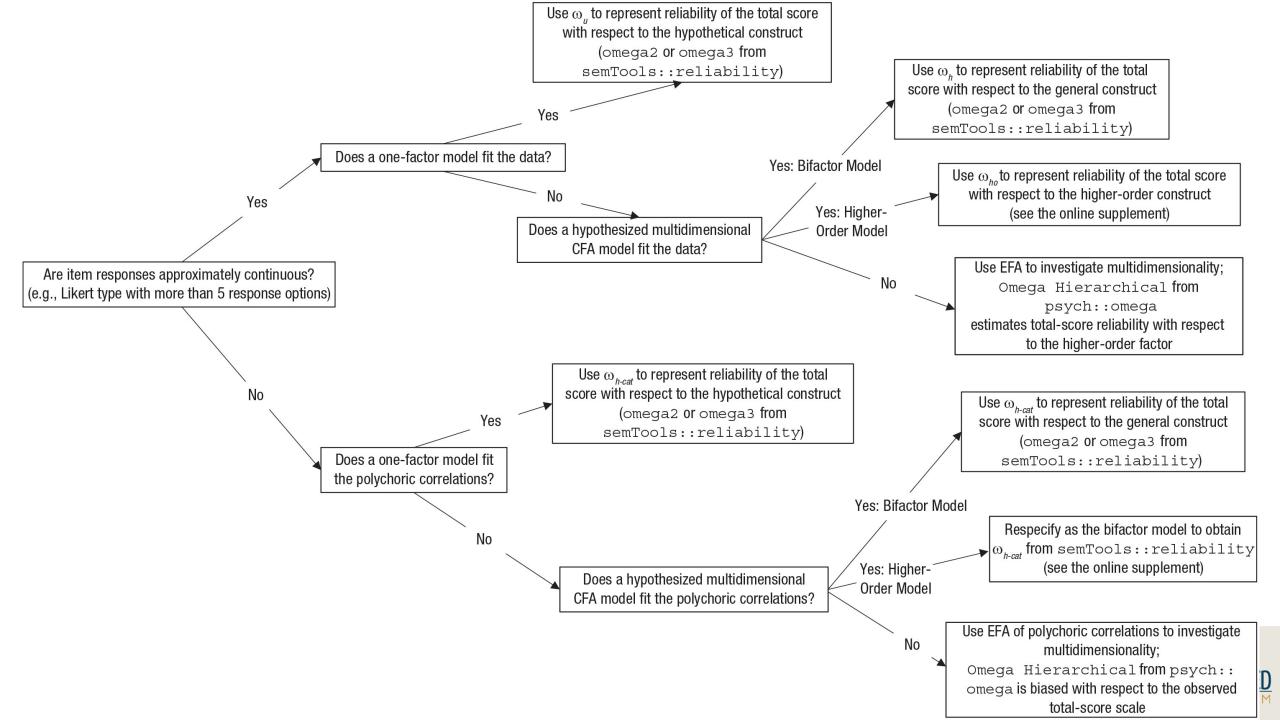


Second-order Model

Posibles estructuras multidimensionales







Análisis factorial exploratorio

- El modelo de medición está subdeterminado
- La relación entre los indicadores y aquello que quiero medir está pobremente especificada
- Las hipótesis de la relación entre los indicadores y el factor de interés son generales
- La probabilidad de que haya más dimensiones no es igual a cero pero desconozco su valor

Análisis factorial exploratorio

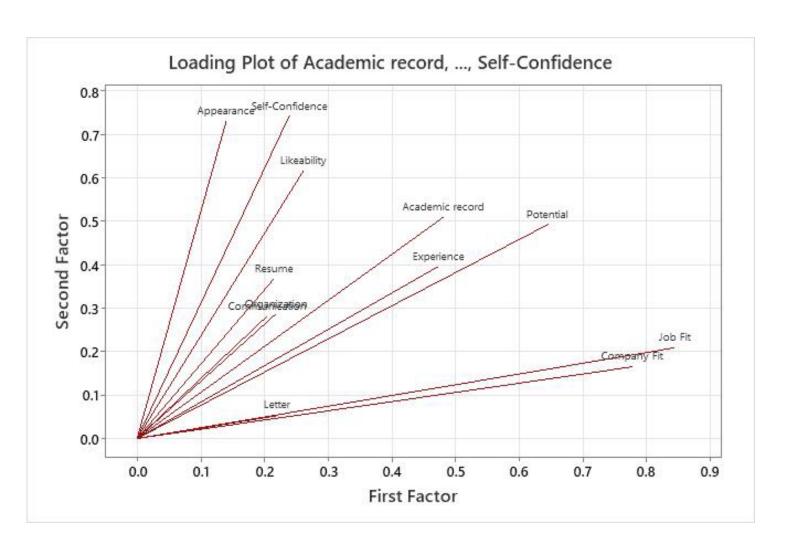


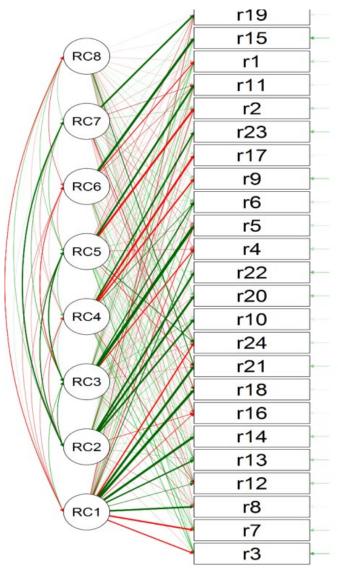
Análisis factorial exploratorio

Múltiples grados de libertad

- Se estima un modelo incondicional de una dimensión
- Se evalúa el ajuste y el error de medición por indicador
- Se seleccionan sólo aquellos indicadores que rastrean la misma señal
- ¿Cómo descarto la posibilidad de que los indicadores con alto error no sean parte de una segunda dimensión?







Múltiples especificaciones



Análisis factorial exploratorio

AFC vs Análisis de componentes principales

La discusión ya no es en términos de modelos de medición sino de métodos de estimación:

Los resultados son distintos porque usamos métodos distintos y no porque usamos modelos diferentes



Ejemplo: Cálculo de Omega AFE

- En el caso exploratorio:
 - Estimo un modelo factorial exploratorio con número creciente de factores
 - Elijo una solución
 - Aplico la fórmula de McDonald

Mal modelo: RMSEA>.06

```
## The df corrected root mean square of the residuals is 0.08
## RMSEA index = 0.165 and the 10 % confidence intervals are 0.16 0.17
## BIC = 2626.29
```

1 Podría usar la función omega () del paquete pysch

```
O<-omega(D[,3:13], nfactors = 1)
O</pre>
```

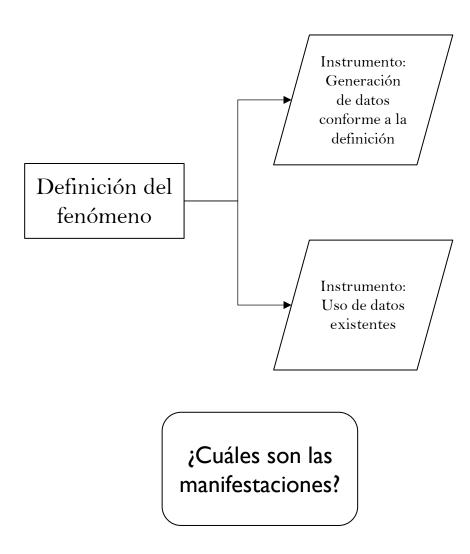
```
## Omega
## Call: omegah(m = m, nfactors = nfactors, fm = fm, key = key, flip = flip,
       digits = digits, title = title, sl = sl, labels = labels,
      plot = plot, n.obs = n.obs, rotate = rotate, Phi = Phi, option = option,
## Alpha:
                          0.86
                          0.89
## Omega Hierarchical:
## Omega H asymptotic:
## Omega Total
                          0.88
## Schmid Leiman Factor loadings greater than 0.2
## analfabetismo
                 0.85
                             0.72 0.28 1
## inasistencia
                  0.30
                             0.09 0.91 1
## sinedbasica
                             0.57 0.43 1
## accesosalud-
                             0.01 0.99 1
                   0.81
                             0.65 0.35 1
## pisotierra
## sinsanitario
                             0.09 0.91 1
## sinaguaentubada 0.49
                             0.24 0.76 1
## sindrenaje
                             0.53 0.47 1
## sinenergia
                   0.60
## sinlavadora
## sinrefrigerador 0.87
                             0.76 0.24 1
```





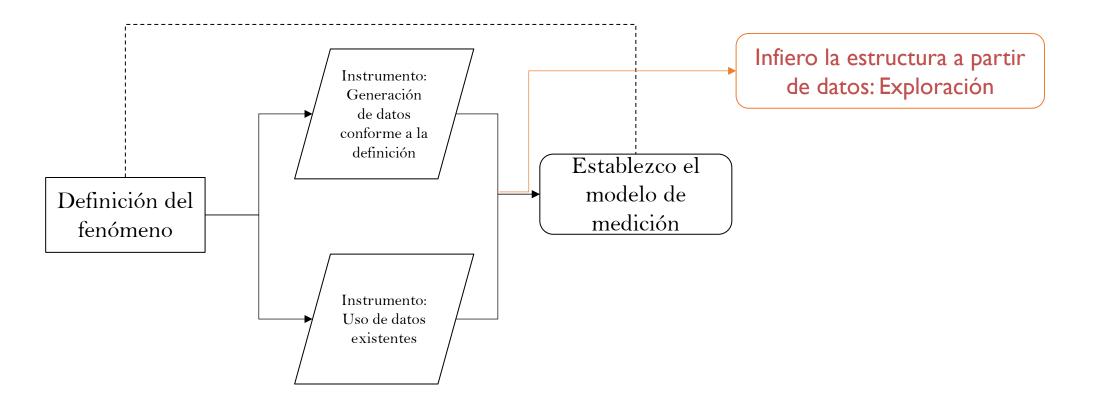
Definición del fenómeno



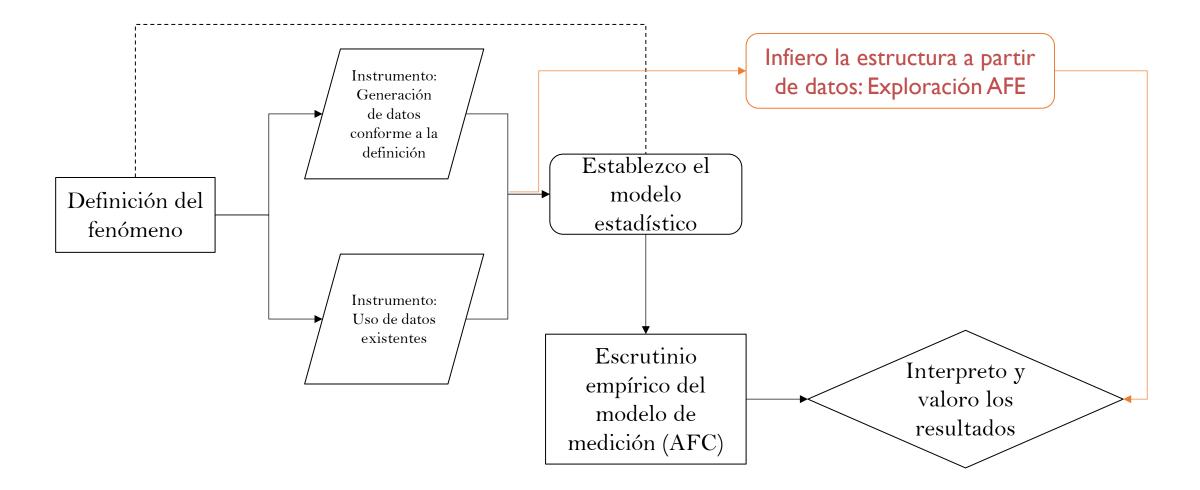


Flujo de trabajo en Medición

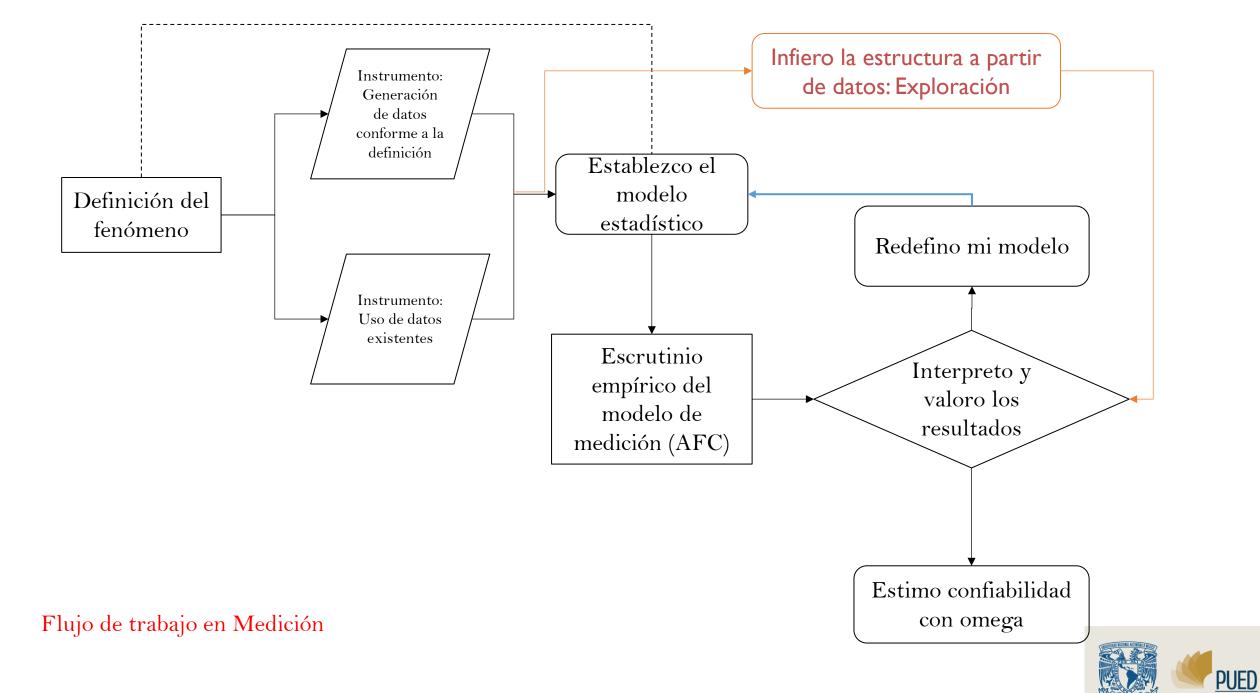


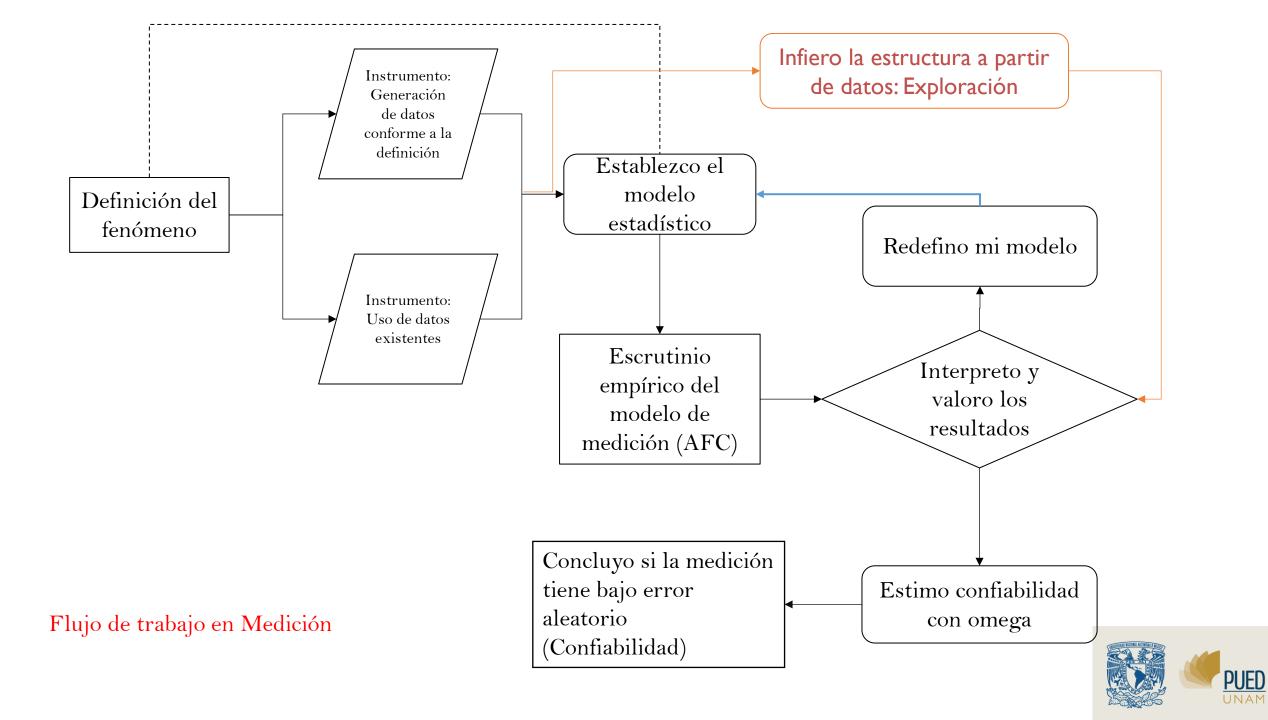












El paquete lavaan

lavaan is easy and intuitive to use

 the 'lavaan model syntax' allows users to express their models in a compact, elegant and useR-friendly way; for example, a typical CFA analysis looks as follows:

```
library(lavaan)
myData <- read.csv("/path/to/mydata/myData.csv")
myModel <- '
    f1 =~ item1 + item2 + item3
    f2 =~ item4 + item5 + item6
    f3 =~ item7 + item8 + item9
.

fit <- cfa(model = myModel, data = myData)
summary(fit, fit.measures = TRUE)</pre>
```

- you can choose between a user-friendly interface in combination with the fitting functions cfa() and sem() or a low-level interface using the fitting function lavaan() where 'defaults' do not get in the way
- convenient arguments (eg. group.equal="loadings") simplify many common tasks (eg. measurement invariance testing)
- lavaan outputs all the information you need: a large number of fit measures, modification indices, standardized solutions, and technical information that is stored in a fitted lavan object



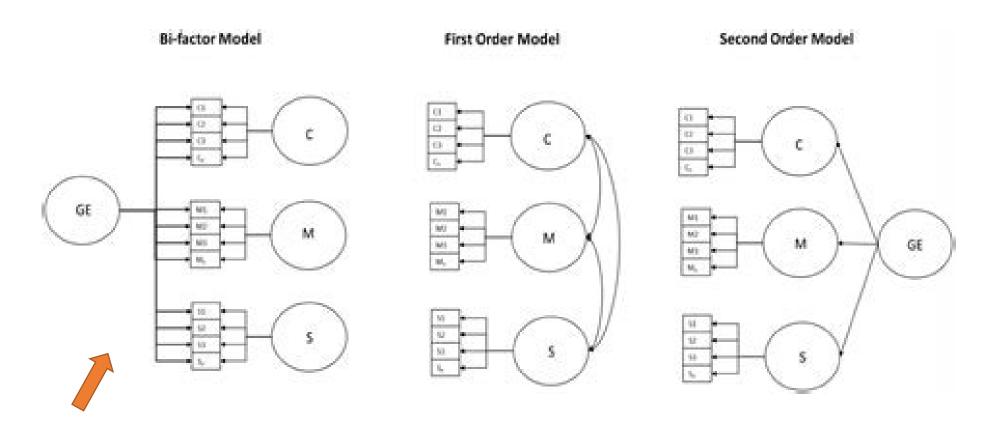
El paquete lavaan

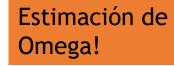
lavaan provides many advanced options

- · full support for meanstructures and multiple groups
- several estimators are available: ML (and robust variants MLM, MLMV, MLR), GLS, WLS (and robust variants DWLS, WLSM, WLSMV), ULS (ULSM, ULMV), DLS, and pairwise ML (PML)
- standard errors: standard, robust/huber-white/sandwich, bootstrap
- test statistics: standard, Satorra-Bentler, Yuan-Bentler, Satterthwaite, scaled-shifted, Bollen-Stine bootstrap
- missing data: FIML estimation
- · linear and nonlinear equality and inequality constraints
- full support for analyzing categorical data: lavaan (from version 0.5 onwards) can handle any mixture of binary,
 ordinal and continuous observed variables
- (from version 0.6 onwards): support for multilevel level SEM



Medición multidimensional







Próxima clase

- Instalar/actualizar R y Rstudio
- Instalar el paquete: lavaan install.packages ("lavaan")
- Instalar el paquete: semTools
- Instalar el paquete: psych

