EPIDEMIOLOGÍA PARA LA SALUD PÚBLICA

Análisis factorial confirmatorio

EPIDEMIOLOGÍA MODERNA: CONCEPTOS Y MODELOS AVANZADOS

ANÁLISIS CONFIRMATORIO:

Uso de la técnica estadística multivariada para probar (o confirmar o refutar) un conjunto preestablecido de relaciones.
En el caso del modelado de ecuaciones estructurales (MEE), el análisis confirmatorio se aplica tanto a la estimación (análisis de regresión) como a la medición (análisis factorial).
Los modelos que incluyen solo análisis de medición representan un subtipo de los modelos completos / híbridos (estructural + confirmatorio).

En el caso de los análisis de medición, el objetivo es solo medir constructos latentes a partir de variables observadas.

Las variables observadas para medir constructos (o variables) latentes que deben estar fuertemente asociadas, pero que no tienen relaciones determinantes entre ellas.

Los análisis de medición de los MEE requieren la aplicación de análisis factoriales confirmatorios. En este caso, el análisis factorial se usa para confirmar la expectativa de que dos o más variables observadas van a converger para formar un mismo constructo latente y que dicho constructo es consistente y confiable.

Sin embargo, debemos ser conscientes de que, en el análisis factorial exploratorio, los constructos (o factores) se forman solo en base a procedimientos estadísticos y sus constituciones son desconocidas para el epidemiólogo en momentos previos al análisis, de esta manera, el análisis factorial confirmatorio se presta a confirmación o no de una teoría de medición previa. Es decir, es necesario que el investigador tenga una expectativa previa de qué variables deben formar un constructo.

En el análisis factorial confirmatorio, los constructos no son directamente, puesto que se infieren indirectamente de los indicadores (variables observadas).

Comienza con una hipótesis a priori sobre las relaciones entre variables, con el objetivo de verificar la validez del constructo y la estructura factorial de las escalas.
En este tipo de modelo, solo hay un error de medición: variables exógenas no medidas.

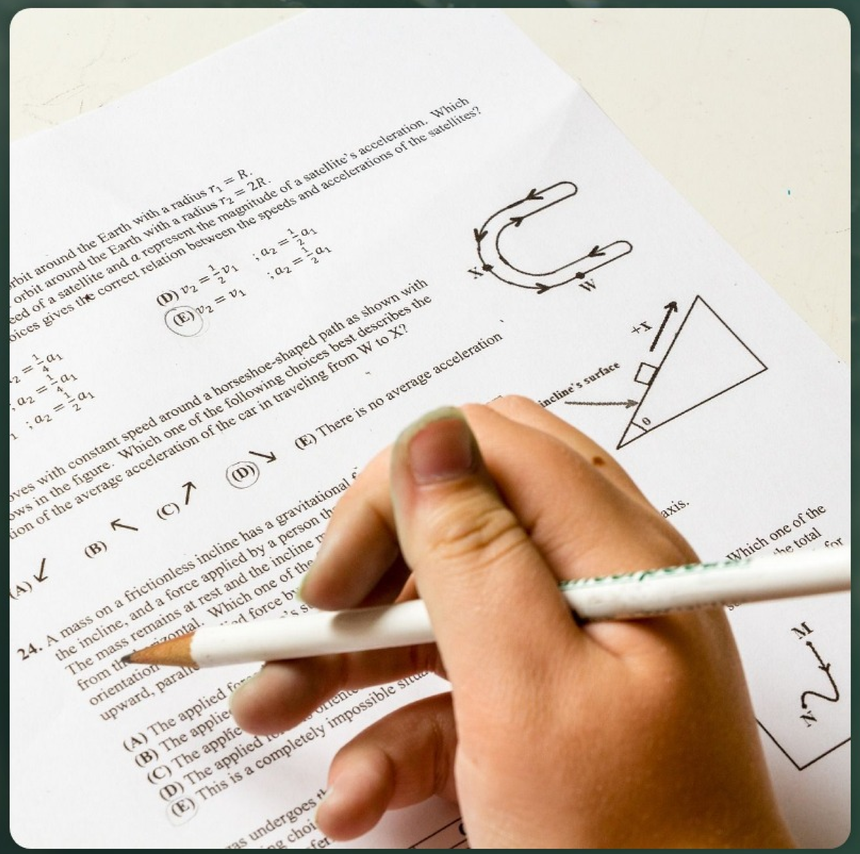
La variable latente (constructo) consta de múltiples indicadores: es más confiable y válido, ya que cada indicador mide un aspecto diferente del constructo. Es una variable no observada, deducida de la covarianza entre dos o más indicadores.
La variable latente debe estar compuesta por al menos 2 indicadores, sin embargo, lo recomendado es 3 indicadores, y lo ideal: 4 indicadores.

La definición de la adecuación del modelo de medición propuesto viene dada por análisis de confiabilidad y / o ajuste. Es común usar el coeficiente alfa de Cronbach para el análisis de confiabilidad de escalas ***(¿Recuerda que los modelos de ecuaciones estructurales, especialmente el análisis confirmatorio, se usan mucho en psicometría?)*** o constructos latentes.

Sin embargo, estos coeficientes se ven muy afectados por el número de elementos (variables observadas) en la escala (constructo).

Más específicamente, manteniendo todo más constante, cuanto mayor sea el número de elementos en la escala medida, mayor será el valor del coeficiente alfa de Cronbach (> 0,70).

Cuando el modelo de medición tiene una pequeña cantidad de ítems, el valor del alfa de Cronbach tiende a ser muy bajo y la utilidad de realizar mediciones desde MEE es aún más evidente, ya que es posible obtener un análisis de ajuste modelo más adecuado y amplio que con el coeficiente.



Sobre la validez del constructo

Para verificar si el constructo realmente mide lo que desea medir, tenemos la evaluación de la validez convergente (los indicadores que miden un constructo deben tener altas cargas estandarizadas en el factor> 0.70). Por lo tanto, se puede concluir que las variables observadas convergen al mismo lugar.

También existe la validez discriminante: las correlaciones entre los factores no deben ser excesivamente altas (> 0,90), después de todo, cada constructo mide una "cosa" diferente.

Ejemplo de análisis factorial confirmatorio

Para proponer un modelo de medición basado en un análisis factorial confirmatorio, se debe utilizar una teoría de medición.

Para nuestro ejemplo, utilizaremos las nueve variables observadas de PHQ9 (poner un pequeño globo con el texto: cuestionario validado para la detección del trastorno depresivo mayor) para componer el constructo de la depresión. Los datos provienen de una submuestra de la Encuesta Nacional de Salud (PNS, Brasil, 2013). Por lo tanto, el diagrama de ecuaciones estructurales propuesto es el siguiente:

Figura 1: Diagrama del modelo de ecuación estructural para medir la variable latente de depresión.

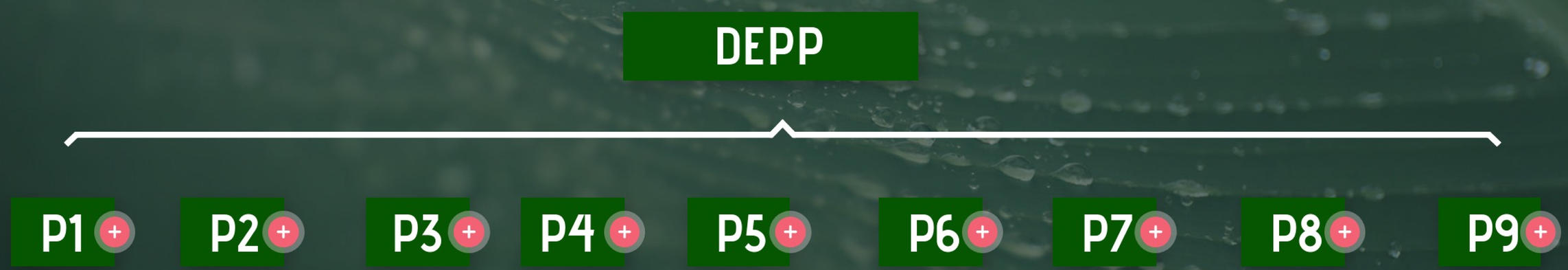


Figura 2: Diagrama del modelo de ecuaciones estructurales para la medición de la variable latente de depresión a partir de los datos de los trabajadores de PNS (2013).

Para la figura 2, inserte la misma figura 1, pero con los valores de carga factorial insertados en las flechas que conectan DEP a los cuadrados.

Sobre el modelo de medición para la depresión: de los 9 ítems, 6 tenían una carga factorial superior a 0,70 y significativa. Tres preguntas (poco interés o placer en hacer las cosas; falta de apetito o comer en exceso; pensar en lastimarse de alguna manera o en que sería mejor estar muerto) tenían un factor de carga entre 0.66 y 0.67, todo significativo (Tabla 1 , figura 2).

Variables	Carga factorial	valor p	Error de medida
P1	0,66	0,000	0.56
P2	0,78	0,000	0.39
P3	0,79	0,000	0.38
P4	0,80	0,000	0.36
P5	0,67	0,000	0.55
P6	0,77	0,000	0.41
P7	0,79	0,000	0.38
P8	0,75	0,000	0.44
P9	0,67	0,000	0.55

Tabla 1: Resultados del modelo de medición de depresión, 2018.

Medidas para ajustarse al modelo: ¿en qué medida el modelo hipotético se ajusta a los datos? Piense en la siguiente lógica:

- Datos = modelo + residuo
- Datos = mediciones de variables observadas
- Modelo = hipótesis de relaciones entre variables
- Residuo = discrepancia entre la matriz de covarianza del modelo hipotético y la matriz de covarianza de muestra (datos observados)

Cuanto más pequeño sea el residuo, mejor será el ajuste, es decir, el modelo se aproxima a los datos.

1- Prueba chi-cuadrado - X2

Una de las medidas de ajuste es el resultado de la prueba chi-cuadrado - X2 - modelo basal o nulo. Para esto, se supone independencia, es decir, todas las covarianzas entre las variables observadas son iguales a cero. El resultado de la prueba muestra la comparación entre el chi-cuadrado del modelo hipotético y el modelo de referencia - si el valor disminuye, hay una mejora en el ajuste.

Recordatorio! El valor alto y significativo muestra discrepancia, se rechaza H0. Valor bajo y no significativo, muestra un buen ajuste.

Limitaciones: muy sensible al tamaño de la muestra. Las pequeñas diferencias entre los datos observados y el modelo tienden a ser estadísticamente significativas.

2- CFI (Índice de ajuste comparativo) y TLI (Índice de Tucker-Lewis)

Índices de ajuste incremental: indican el porcentaje de ganancia en el ajuste comparando el modelo hipotético probado con el modelo nulo (modelo que supone independencia entre las variables). Si la relación de ganancia para estos indicadores es > 0,95, el ajuste se considera bueno.

3- RMSEA (error cuadrático medio de aproximación) y SRMR (raíz media cuadrática residual)

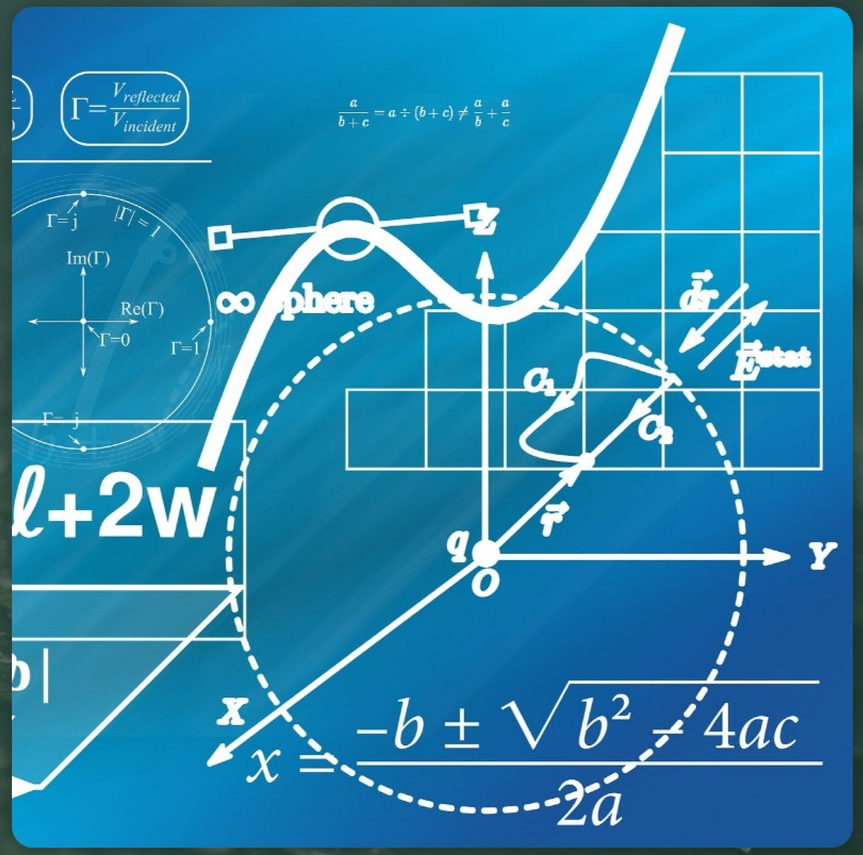
Índices de ajuste absoluto basados en residuos: cuanto más pequeños sean los residuos, mejor será el ajuste. Los valores < 0.05 son indicativos de un buen ajuste.

RMSEA está disponible con un intervalo de confianza del 90%. Interpretación:

Buen ajuste <0.05 (Browne y Cudeck, 1993)
Razonable <0.08 (Browne y Cudeck, 1993)
Buen ajuste <0.06 (Hu y Bentler, 1999)
Límite superior del intervalo de confianza <0.08

SRMR: aplicable a las variables continuas observadas.

Raíz Media Cuadrática Residual (SRMR) - varía de 0 a 1. Buen ajuste <0.05
Para las variables categóricas, se utiliza el estimador de mínimos cuadrados ponderados con media y varianza ajustada (WLSMV), sin embargo, su resultado es muy sensible para descartar el ajuste del modelo si las otras medidas así lo indican.



WRMR – Raíz cuadrada media residual ponderada: los valores <0.95 son indicativos de un buen ajuste.

Los siguientes son los resultados de ajustar el modelo en nuestro ejemplo. Un coeficiente que tiene un valor p < 0.05 indica que el primer modelo de la comparación (en este caso, el modelo nulo) tiene un ajuste más bajo que el segundo modelo (en este caso, el modelo saturado). Por lo tanto, dado que el modelo utilizado en el análisis es equivalente al modelo saturado, se concluye que tiene un ajuste apropiado, cuando el análisis de chi-cuadrado se toma como referencia. Otras medidas de ajuste del modelo utilizadas con frecuencia (CFI, TLI y CD), todas con un valor igual o cercano a 1.00, también presentan resultados que corroboran la calidad del ajuste del modelo de medición propuesto. Todas las variables son categóricas, por lo que se utilizó el estimador de mínimos cuadrados ponderados con media y varianza ajustada (WLSMV). En vista de una muestra muy grande (19,450 sujetos), la prueba chi-cuadrado y el WRMR se presentan como medidas muy sensibles y rutinariamente con valores altos, sin embargo, dado que las otras estimaciones son adecuadas, el ajuste del modelo es aceptable.

Hay muchos conceptos estadísticos involucrados en el Análisis Factorial Confirmatorio, sin embargo, para el epidemiólogo es importante que comprenda la estructura del análisis, sus conceptos básicos, su aplicación e interpretación. Después de todo, como cualquier análisis epidemiológico, las estadísticas no definen la relación causal per se, sino que ratifican o refutan una hipótesis teórica elaborada a priori.

En la próxima y última semana, estudiaremos el modelo estructural y también el híbrido, o completo, que combina el análisis factorial confirmatorio y el modelo estructural

Tabla 2: Resultados de ajuste del modelo

	Modelo DEP
X^2	510,569 (valor p<0)
CFI	0,996
TLI	0,995
RMSEA	0,013 (IC90% 0,012-0,015; valor p=1,0)
WRMR	1,67

Unidad basada en el libro: Neves, J. A. B. (2018). Modelo de equações estruturais: Uma introdução aplicada. ENAP, en el curso de modelación con ecuaciones estructurales ofrecidas por la Universidad Federal de Maranhão (Brasil), por el profesor Antonio Augusto Moura y en la tesis dela profesora Nágila Oenning.

PUCE Virtual

Quito
Av.12 de Octubre 1073 y Roca
Edificio de la Facultad de Comunicación,
Lingüística y Literatura. Primer Piso. Oficina 106.

Teléfono:
(593-2) 299 1592 / (593)09 8 851 2839

Correo:
soportevirtual@puce.edu.ec