

Estimando la heredabilidad del CI: Cómo analizar datos genéticos utilizando SEM



David Gallardo, Uwe Kramp, Carlos García-Forero, Albert Maydeu-Olivares y Antonio Andrés-Pueyo
GRC Personalitat: Avaluació i Predicció del Comportament
Departament de Personalitat, Avaluació i Tractament Psicològics.
Facultat de Psicologia, Universitat de Barcelona
Passeig de la Vall d'Hebrón, 171. 08035 Barcelona (Spain)

Quando se analizan datos genéticos, los Modelos de Ecuaciones Estructurales (SEM) proporcionan una metodología clara y sencilla para descomponer la varianza fenotípica utilizando una aproximación basada en diferentes modelos. Además, esta versatilidad permite que se puedan implementar, probar y comparar una gran multiplicidad de modelos utilizando SEM, permitiendo así al investigador información muy valiosa respecto a las fuentes de variabilidad. En este póster describimos brevemente la metodología necesaria y la aplicamos al reanálisis a un juego de datos de CI (Carrillo-Gistain, 1987) obtenido en una muestra de gemelos, utilizando el modelo ACE. Los resultados que hacen referencia a la heredabilidad son consistentes con la bibliografía existente.

1. INTRODUCCIÓN

El debate acerca de la importancia de la naturaleza versus la crianza en el desarrollo de las diferencias individuales parece interminable. Desde que Galton (1982) publicó su reconocido "Hereditary genius: an inquiry into its laws and consequences", ha sido muchos los intentos de un abordaje científico sobre el origen genético o ambiental de las diferencias individuales. También desde entonces ha habido diversas aproximaciones metodológicas a esta cuestión, desde la aproximación clásica de la fórmula de Falconer (1960) hasta el análisis de estructuras de covarianza propuesto por Martin e Eaves (1977; para una revisión exhaustiva, ver Neale y Maes, en prensa), basándose en la técnica propuesta por Jöreskog (1970). Esto ha llevado al establecimiento de varios registros de gemelos en el mundo para proporcionar a los investigadores las herramientas necesarias. Desafortunadamente, España se ha mantenido al margen de esta tradición, y sólo hay un estudio utilizando una muestra de gemelos en el campo de las diferencias individuales (Carrillo-Gistain, 1987).

2. OBJETIVO

El objetivo de la presente investigación es reanalizar los datos de Carrillo-Gistain empleando modelos de ecuaciones estructurales (SEM) para ejemplificar cómo éstos pueden ser utilizados para analizar datos genéticamente informativos (p.ej. Correlaciones entre familiares). Nuestro objetivo secundario es contrastar si nuestros resultados son similares a los reportados por la literatura existente.

3. MÉTODO

La muestra utilizada estaba compuesta por 140 sujetos, distribuidos en 2 grupos: 39 pares de gemelos monocigóticos (MZ) y 31 pares de gemelos dicigóticos (DZ), todos provenientes de la provincia de Lérida. Nos referimos a Carrillo-Gistain (1987) para datos más concretos.

Respecto al análisis, en primer lugar pondremos a prueba diferentes modelos teóricos (ACE, AE, CE y E; ver Figura 1) que incluyen diferentes factores latentes (varianza genética aditiva (A), varianza ambiental compartida (C) y varianza ambiental específica (E) y que explican la varianza observada (puntuaciones CI de un par de gemelos). Para ello, utilizaremos las reglas RAM (McArdle, 2005) para representar de forma matricial los modelos de la Figura 1. Como ejemplo, veamos como serían las matrices del modelo ACE:

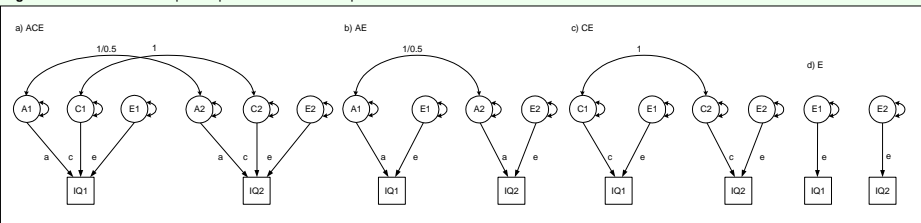
$$A = \begin{pmatrix} 0 & 0 & a & c & e & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & a & c & e \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \quad S = \begin{pmatrix} 0 & & & & & & & \\ 0 & 0 & & & & & & \\ 0 & 0 & 1 & & & & & \\ 0 & 0 & 0 & 1 & & & & \\ 0 & 0 & 0 & 0 & 1 & & & \\ 0 & 0 & r & 0 & 0 & 1 & & \\ 0 & 0 & 0 & 1 & 0 & 0 & 1 & \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \quad F = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Dónde a, c y e representan las rutas causales (flechas) en la figura 1. Para evaluar el ajuste del modelo se compara la matriz de covarianzas esperada calculada bajo la fórmula:

$$Cov \begin{pmatrix} IQ_1 \\ IQ_2 \end{pmatrix} = F(I - A)^{-1} S(I - A)^{-1} F'$$

con la matriz de covarianzas observada. Los modelos fueron estimados vía máxima verosimilitud (ML) utilizando las covarianzas como estadístico muestra. Como los modelos no son anidados, la selección del mejor modelo se hizo mediante el Criterio de Información Bayesiano (BIC: Schwarz, 1978). MPLUS (Muthén y Muthén, 1998) fue el software utilizado para estimar los modelos.

Figura 1. Modelos estimados para explicar la varianza fenotípica del CI. Ver detalles en el texto.



7. BIBLIOGRAFIA

Bouchard, T. J., & McGue, M. (2002). Genetic and environmental influences on human psychological differences. *Journal of Neurobiology*, 54(1), 4-45.
Carrillo-Gistain, F. J. (1987). *La inteligencia: Un estudio genético ambiental*. Universitat de Barcelona.
Falconer, D. S. (1960). *Introduction to quantitative genetics*. London: Oliver and Boyd.
Galton, F. (1869). *Hereditary genius: An inquiry into its laws and consequences*. London: Macmillan and Co.
Jöreskog, K. (1970). A general method for analysis of covariance structures. *Biometrika*, 57, 239-251.
Martin, N. G., & Eaves, L. J. (1977). The genetical analysis of covariance structure. *Heredity*, 38(79-95).
McArdle, J. J. (2005). The development of the ram rules for latent variable structural equation modeling. In A. Maydeu-Olivares & J. J. McArdle (Eds.), *Contemporary psychometrics: A festschrift for roderick p. McDonald*. Hillsdale, NJ: Lawrence Erlbaum.
Neale, M. C., & Maes, H. H. (in press). *Methodology for genetic studies of twins and families*. Dordrecht, The Netherlands: Kluwer Academic Publishers B.V.
Schwarz, G. (1978). Estimating the dimension of a model. *Ann Stat*, 6, 461-464.

4. RESULTADOS

La Tabla 1 muestra los resultados del ajuste de los modelos. El modelo que mejor se ajusta es el modelo ACE, y el que peor se ajusta es el modelo E.

Tabla 1. Resultados del ajuste de los diferentes modelos.

Modelo	χ²	gl	p-valor	BIC	RMSEA	90% IC de RMSEA
1. ACE*	0.56	3	0.91	1081.78	0	0.000 - 0.116
2. AE	5.09	4	0.28	1082.06	0.09	0.000 - 0.283
3. CE	10.08	4	0.03	1087.06	0.21	0.043 - 0.373
4. E	84.23	5	0	1156.95	0.67	0.551 - 0.803

*Modelo mejor ajustado. Notas: A=Varianza genética aditiva; C=Varianza ambiental compartida; E=Varianza ambiental específica; gl=Grados de libertad; BIC= Bayesian Information Criteria; RMSEA= Root Mean Square Error of Approximation.

La Tabla 2 indica la proporción de varianza explicada por cada uno de los factores latentes y el valor de los parámetros estimados.

Tabla 2. Estimaciones de los parámetros, E.E. y % de varianza..

	Parámetro	E.E.	% de varianza
A	9.31	1.89	40
C	10.46	2.14	50
E	4.83	0.54	10

Notas: A=Varianza genética aditiva; C=Varianza ambiental compartida; E=Varianza ambiental específica.

5. DISCUSIÓN Y CONCLUSIONES

Los resultados que hemos obtenido están en la línea de la bibliografía existente (Bouchard y McGue, 2002), aunque con alguna ligera discrepancia. Por ejemplo, respecto al elevado peso de la varianza ambiental compartida en nuestro estudio. Estas discrepancias pueden explicarse debido a la heterogeneidad de la muestra estudiada.

6. AGRADECIMIENTOS

Este trabajo ha sido posible gracias a las ayudas BSO2003-08507 (Albert Maydeu-Olivares, IP) y SEC2001-3821-C05-01 (Antonio Andrés-Pueyo, IP) ambas del Ministerio de Ciencia y Tecnología.

Figura 2. Ilustración representando a dos gemelos de las Crónicas de Nuremberg, por Hartmann Schedel (1440-1514).

