

Análisis multivariado causal de los trastornos de aprendizaje desde el diagnóstico clínico. Una aplicación del modelo log-lineal.

por
Maria Dolores Peris

INTRODUCCIÓN

Abordamos el tema, muy frecuente, del diagnóstico de los trastornos de aprendizaje durante los primeros años escolares. Los criterios y procesos de su evaluación son muy diversos, y para obtener cierta homogeneidad, acudimos a los Registros de los Servicios de Diagnóstico del INSERSO, donde un mismo equipo diagnostica a los niños con problemas escolares susceptibles de recibir subvención si se llega al diagnóstico de Deficiencia Mental (C.I. < 45). Seleccionamos los historiales con C.I. > 70, para eliminar de la muestra los trastornos asociados a las limitaciones en el desarrollo cognitivo. Los diagnósticos, a excepción de la edad y el C.I., estaban registrados en escalas nominales y en formato sin estructurar.

La estructuración de estas variables se realizó tomando como variable cada una de las afirmaciones y señalando en la matriz de datos la presencia (1) o ausencia (0) de cada variable en el informe de cada sujeto. La edad se dividió en 5 intervalos mutuamente exclusivos (uno para cada uno de los años desde 6 hasta 10), de modo que se registraba la «presencia» en la edad correspondiente al sujeto, y del mismo modo se dividió el C.I. en 4 intervalos (> 100, 90-99, 80-89, < 79). Las variables que exigían más de dos categorías, se dividieron en tantas variables como respuestas posibles (asi, hijo primogénito, hijo menor; parto con forceps, cesárea,...). Se acumuló un total de 57 variables fundamentalmente de dos clases: anamnesis y diagnósticos.

El análisis de estos datos dicotómicos ha sido tema de estudios en los últimos años, dada su importancia en las Ciencias sociales (Goodman, L.A. 1978; Reynolds, H.T. 1977; Feinberg, S.E. 1983; Knoke, D. y Burke, P.J. 1977; Nigel, G. 1981, entre otros). Los modelos loglineales aparecen en tales estudios como la alternativa más adecuada para llegar a establecer las relaciones entre variables, determinando las interacciones entre ellas, de forma equivalente

al modelo de ANOVA. En caso de que exista una relación temporal antes-después entre las variables analizadas, la causalidad puede ser evidenciada, en base a las relaciones entre las frecuencias en las tablas de contingencia. El análisis factorial de Bool, es otra alternativa para llegar a definir la estructura de las tablas del tipo «presencia-ausencia». Uno y otro análisis fueron realizados sobre la matriz de datos en que se registraron los historiales de 276 casos. Veamos los resultados de cada uno de ellos.

ANÁLISIS FACTORIAL DE BOOL

Obtenemos los siguientes resultados:

Factor I: varón; C.I. > 100; 6 años; DISLALIA

Factor II: varón; DISLEXIA; COMPRENSIÓN DISMINUIDA; ATENCIÓN DISMINUIDA; MEMORIA DISMINUIDA.

Factor III: varón; $80 < \text{C.I.} < 90$; COMPRENSIÓN DISMINUIDA.

Factor IV: lugar intermedio entre hermanos; DISLEXIA.

Factor V: C.I. < 79.

Los diagnósticos que aparecen en los factores, se corresponden con las variables que acumulan las mayores frecuencias en las características de «presencia». La dislexia (var. 15) se diagnostica en el 62% de los casos; la dislalia (var. 20) en el 36%, la disminución de la comprensión (var. 28), de la atención (var. 29) y de la memoria (var. 30) en el 68%, 58% y 52% respectivamente. El número superior de varones, como viene siendo habitual en estos estudios (66%), justifica su presencia en los tres primeros factores, si bien no se asocia con el último, referido a los cocientes más bajos. La dislexia aparece dividida en dos tipos: uno asociada con déficits de las actividades intelectuales (comprensión, atención y memoria), aunque no se asocia con ningún nivel concreto de C.I., y otro asociado con el lugar intermedio entre los hermanos, que generalmente se define, como el que menos atención recibe por parte de los padres. Si bien ambos tipos son independientes del C.I., parece estar vinculado el uno, con otros déficits cognitivos y el otro, con una posible falta de atención por parte de la familia.

En general, los tradicionales trastornos de aprendizaje diagnosticados (discalculia, disgrafía, disortografía,...) así como las variables causales más comunes, alcanzan tasas muy bajas, no presentando apenas asociaciones. Es de destacar que los valores del C.I., aun no apareciendo relacionados con trastornos concretos del aprendizaje, se ordenan por orden decreciente en tres de los factores obtenidos. Podría interpretarse este resultado como que la graduación del C.I. propuesta es la mejor clasificación, o diferenciación, del conjunto de variables analizadas. El primer factor, puede ser interpretado como un «pseudo-retraso», apareciendo los trastornos del lenguaje en los diagnósticos de los primeros años (var. 6), con un C.I. superior al medio (var. 2) y desapareciendo en años posteriores.

Tabla 1. – Resultados del Análisis Log-lineal. (Significación de efectos).

***** THE RESULTS OF FITTING ALL K-FACTOR MARGINALS.
THIS IS A SIMULTANEOUS TEST THAT ALL K-1 AND HIGHER FACTOR
INTERACTIONS ARE ZERO.

K-FACTOR	D.F.	LR CHISQ	PROB.	PEARSON CHISQ	PROB.	ITERATION
0-MEAN	255	614,11	.00000	1310,69	.00000	
1	247	265,12	.20435	422,53	.00000	2
2	219	86,70	1.00000	94,73	1.00000	7
3	163	45,13	1.00000	44,91	1.00000	3
4	93	15,10	1.00000	15,29	1.00000	4
5	37	3,99	1.00000	4,11	1.00000	4
6	9	,85	.99971	,86	.99970	3
7	1	,02	.87706	,02	.87706	2
8	0		1.	0,	1.	

***** A SIMULTANEOUS TEST THAT ALL K-FACTOR INTERACTIONS ARE
SIMULTANEOUSLY ZERO.
THE CHI-SQUARES ARE DIFFERENCES IN THE ABOVE TABLE.

K-FACTOR	D.F.	LR CHISQ	PROB.	PEARSON CHISQ	PROB.
1	8	348,98	.00000	888,16	.00000
2	28	178,42	.00000	327,80	.00000
3	56	41,57	.92476	49,82	.70637
4	70	30,03	.99999	29,62	.99999
5	56	11,11	1.00000	11,18	1.00000
6	28	3,14	1.00000	3,25	1.00000
7	8	,83	.99911	,84	.99909
8	1	,02	.87706	,02	.87706

***** ASSOCIATION OPTION SELECTED FOR ALL TERMS OF ORDER LESS
THAN OR EQUAL TO 8 PARTIAL ASSOCIATION MARGINAL ASSOCIATION

EFFECT	D.F.	CHISQUARE	PROB.	ITER	CHISQUARE	PROB.	ITER
1	1	15,15	.0001				
2,	1	105,70	.0000				
3,	1	22,07	.0000				
4,	1	68,54	.0000				
5,	1	10,18	.0014				
6,	1	14,38	.0001				
7,	1	88,95	.0000				
8,	1	24,01	.0000				
12,	1	,46	.4980	7	,58	.4477	2
13,	1	,89	.3449	7	,90	.3433	2
14,	1	,24	.6209	7	,12	.7259	2
15,	1	,24	.6243	7	,89	.3453	2
16,	1	5,60	.0179	7	5,37	.0205	2
17,	1	,30	.5831	7	,21	.6447	2
18,	1	,04	.8409	7	,22	.6355	2
23,	1	,25	.6191	7	2,59	.1076	2
24,	1	23,24	.0000	4	40,28	.0000	2
25,	1	,78	.3779	6	4,85	.0277	2
26,	1	,44	.5048	6	0,54	0.105	2
27,	1	4,69	.0303	4	12,03	.0005	2
28,	1	14,36	.0002	4	25,96	.0000	2
34,	1	2,68	.1018	7	5,16	.0231	2
35,	1	,20	.6521	7	,77	.3809	2
36,	1	,07	.7867	7	,67	4.134	2
37,	1	5,12	.0236	7	2,01	.1563	2
38,	1	9,36	.0022	7	11,19	.0008	2
45,	1	4,58	.0323	7	13,16	.0003	2
46,	1	8,82	.0030	6	20,03	.0000	2
47,	1	3,09	.0788	4	9,25	.0023	2
48,	1	1,80	.1798	4	12,00	.0005	2
56,	1	22,59	.0000	7	30,81	.0000	2
57,	1	,00	.9717	7	,88	.3491	2
58,	1	,16	.6935	7	,56	.4526	2
67,	1	1,36	.2441	7	3,95	.0468	2
68,	1	,68	.4093	7	3,43	.0638	2
78,	1	3,76	.0525	6	7,22	.0072	2

MODELO LOG-LINEAL.

Seleccionadas las 8 variables que presentan asociaciones más altas en los análisis de Kendal y de Spearman y cuyas frecuencias son más elevadas, realizamos con todas ellas una tabla de contingencia distribuyendo las frecuencias entre todas las combinaciones posibles de «presencia-presencia», «presencia-ausencia», «ausencia-presencia» y «ausencia-ausencia».

A continuación se estimó el logaritmo de la frecuencia esperada para cada celda, como una función aditiva de los logaritmos de los efectos e interacciones más elevadas, según el modelo lineal no saturado.

tabla

En la tabla 1 aparecen sucesivamente los siguientes resultados:

- A) Prueba de significación de los efectos más altos para cada modelo.
- B) Prueba de significación de la suma de efectos para cada modelo.
- C) Prueba de significación parcial y marginal de cada uno de los efectos.

Las variables analizadas han sido:

1: varón;	2: C.I. > 100	3: C.I. < 79	4: 6 años
5: DISLEXIA	6: DISLALIA	7: HIPERKINESIA	8: Comprensión disminuida.

Tanto en A) como en B) obtenemos que los modelos dejan de ser significativos a partir de las asociaciones de orden 3. Tan solo los efectos simples y algunas de las asociaciones entre dos variables son significativas. Por ello en C) tan sólo analizamos las asociaciones de orden 2 y las de las variables independientemente.

El modelo log-lineal no saturado (con sólo los efectos significativos) presenta las asociaciones señaladas en al fig. 1., según la ecuación:

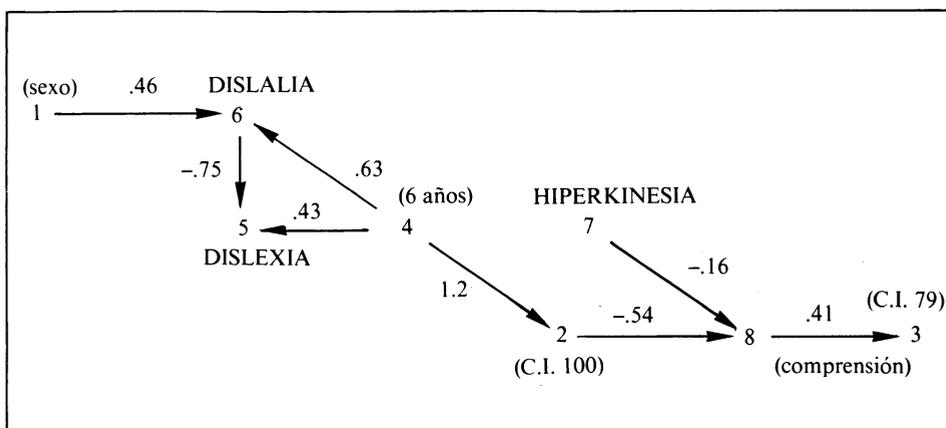
$$\theta + \lambda^1 + \lambda^2 + \lambda^3 + \lambda^4 + \lambda^5 + \lambda^6 + \lambda^7 + \lambda^8 + \lambda^{16} + \lambda^{24} + \lambda^{28} + \lambda^{33} + \lambda^{45} + \lambda^{46} + \lambda^{56} + \lambda^{78}$$

Este modelo, elegido como el mejor explica las relaciones entre las variables, permite establecer un diagrama causal direccionado las relaciones, en base a dos características de las variables:

a) Por ser las variables atributivas inmodificables y en consecuencia, origen y no término de la relación. Aplicando el modelo logit serían variables independientes que actuarían sobre las dependientes.

b) Por la distribución de las frecuencias entre las categorías de las variables, que muestran cuando la presencia en una variable está modulada por la presencia (relación positiva) o ausencia (relación negativa) en la otra variable.

Fig. 1. Modelo log-lineal



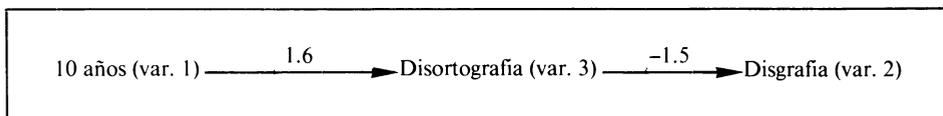
Algunas conclusiones a partir de este modelo son las siguientes:

1. La ausencia de Hiperkinesia aumenta el diagnóstico de Comprensión disminuida.
2. La presencia de Dislexia se ve especialmente favorecida por la ausencia de dislalia.
3. La relación entre Comprensión disminuida y C.I. < 79 es, como cabría esperar, positiva. La relación entre Hiperkinesia y C.I. > 100, es, sin embargo, positiva.
4. Los 6 años parecen decisivos para los diagnósticos que aquí se estudian, diferenciando dos grupos distintos de problemas: lenguaje y C.I.

Un segundo análisis log-lineal fue realizado sobre las variables disgrafia, disortografía y «10 años». Los resultados, como se muestran en la tabla 2, indican una significación equivalente, tanto en los efectos parciales como en los marginales, para las relaciones 13 y 23. El mejor modelo es el expuesto en la figura 2, según la ecuación

$$\theta + \lambda^1 + \lambda^2 + \lambda^3 + \lambda^{13} + \lambda^{23}$$

Fig. 2.- Modelo para las variables analizadas.



Obsérvese cómo en las tablas de frecuencias se evidencia el aumento del diagnóstico de disgrafia en ausencia del de disortografía, que a su vez aparece

Tabla 2. Resultados del análisis log-lineal. (Busqueda del mejor modelo).

***** OBSERVED FREQUENCY TABLE

3M	2M	1M			
			AUSENCIA PRESENCI TOTAL		
AUSENCIA	AUSENCIA		208	52	260
	PRESENCI		7	2	9
	TOTAL		215	54	269
PRESENCI	AUSENCIA		1	3	4
	PRESENCI		0	3	3
	TOTAL		1	6	7

TOTAL OF THE OBSERVED FREQUENCY TABLE IS 276

ALL CASES HAD COMPLETE DATA FOR THIS TABLE.

***** DELTAS .500 15 ADDED TO EACH CELL FOR ALL ANALYSES

***** THE RESULTS OF FITTING ALL K-FACTOR MARCINALS.
THIS IS A SIMULTANEOUS TEST THAT ALL K-1 AND HIGNER FACTOR INTERACTIONS ARE ZERO.

K-FACTOR	D.F.	LR. CHISQ	PROB.	PEARSON CHISQ	PROB.	ITERATION
D-HEAN	7	704,51	.00000	1043,37	.00000	
1	4	26,85	.00002	123,11	.00000	2
2	1	,19	.66500	,18	.67141	5
3	0	0,	1.	0,	1.	

***** A SIMULTANEOUS TEST THAT ALL K-FACTOR INTERACTIONS ARE SIMULTANOUSLY ZERO.
THE CHI-SQUARES ARE DIFFERENCES INT HE ABOVE TABLE.

K-FACTOR	D.F.	LR CHISQ	PROB.	PEARSON CHISQ	PROB.
1	3	677,66	.00000	920,27	.00000
2	3	26,67	.00001	122,93	.00000
3	1	,19	.66500	,18	.67141

***** ASSOCIATION OPTION SELECTED FOR ALL TERMS OF ORDER LESS THAN OR EQUAL TO 3
PARTIAL ASSOCIATION MARGINAL ASSOCIATION

EFFECT	D.F.	CHISQUARE	PROB.	ITER.	CHISQUARE	PROB.	ITER.
1,	1	92,06	.0000				
2,	1	276,99	.0000				
3,	1	308,58	.0000				
12,	1	,36	.5488	2	3,16	.0754	2
13,	1	10,32	.0013	2	13,12	.0003	2
23,	1	10,38	.0013	2	13,19	.0003	2
123,	1	,19	.6650				

***** ALL MODELS ARE REQUESTED

MODEL	D.F.	LIKELIHOOD. BATIO CHISQ.	PROB.	PEARSON CHISQ.	PROB.	ITERATIONS
1,	6	612,43	.0000	698,54	.0000	1
2,	6	427,52	.0000	436,81	.0000	1
3,	6	395,93	.0000	415,26	.0000	1
1,2	5	335,43	.0000	256,15	.0000	1
2,3	5	118,94	.0000	147,30	.0000	1
3,1	5	303,85	.0000	249,98	.0000	1
1,2,3,	4	26,85	.0000	123,11	.0000	1
12,	4	332,27	.0000	253,21	.0000	1
13,	4	290,73	.0000	233,00	.0000	1
23,	4	105,75	.0000	98,79	.0000	1
1,23,	3	13,67	.0034	17,22	.0006	1
2,13,	3	13,73	.0033	32,61	.0000	1
3,12,	3	23,69	.0000	64,71	.0000	1
12,13,	2	10,57	.0051	17,28	.0002	1
13,23,	2	,65	.7607	,54	.7650	1
23,12,	2	10,51	.0052	11,90	.0026	1
12,13,23,	1	,19	.6650	,18	.6714	5

después de los 10 años. Y cómo no puede mantenerse el modelo de independencia.

Algunos de los resultados obtenidos podrían plantear la hipótesis del «efecto del diagnóstico alternativo», de modo que cuando el sujeto carece de los síntomas de un diagnóstico determinado, aumenta la probabilidad de que se le dé un diagnóstico alternativo, apareciendo así las parejas de diagnósticos asociados con signo negativo, a la manera de constructos (Kelly) alternativos del diagnosticador.

REFERENCIAS BIBLIOGRÁFICAS

- COX, D.R. (1970): *Analysis of Binary Data*. Chapman and Hall. London.
- ÉVERIT, B.S. (1977): *The Analysis of Contingency Tables*. Chapman and Hall. London.
- FIENBERG, S.E. (1983): *The Analysis of Cross-Classified Categorical Data*. The MIT Press. Cambridge.
- GOODMAN, L.A. (1978): *Analyzing Qualitative/Categorical Data*. Addison-Wesley. London.
- KNOKE, D. y BURKE, P.J. (1977): *Log-linear Models*. Sage. Beverly Hills.
- Mc CULLAGH, P. y NELDER, J.A. (1983): *Generalized Linear Models*. Chapman and Hall. London.
- NIGEL GILBERT, G. (1981): *Modelling Society*. G. Allen & Unwin. London.
- PLACKETT, R.L. (1981): *The Analysis of Categorical Data*. Griffin & C. London.
- REYNOLDS, H.T. (1977): *Analysis of nominal data*. Sage. Beverly Hills.