Factores de dificultad en el análisis de ítems. Qué son, por qué aparecen y posibles soluciones

José Luis GAVIRIA SOTO

Departamento de Métodos de Investigación y Diagnóstico en Educación.
Universidad Complutense de Madrid

La aplicación del modelo de factores comunes a conjuntos de ítems dicotómicos, aunque es una práctica corriente supone una cierta violación del modelo, ya que las variables observadas no son continuas. Esto implica ciertas disfunciones, de las que la más importante es la aparición de lo que en la literatura científica se ha venido llamando factores de dificultad.

Ferguson (1941) demostró que cuando se analiza la matriz de intercorrelaciones de un conjunto de ítems de distintos índices de dificultad, el rango de esa matriz es mayor que uno, aunque todos los ítems sean homogéneos en cuanto a contenido. Se da por supuesto que la homogeneidad se define como la propiedad en función de la cual el rango de la matriz de correlaciones ha de ser uno. Estudiando los factores aparecidos se vió que éstos correlacionaban altamente con la media de los tests, cuando se aplicó a tests el análisis factorial, o con la media de los ítems, es decir, la proporción de sujetos que respondieron correctamente a cada uno o índice de dificultad, cuando se aplica a ítems.

Wherry y Gaylord (1944), reanalizando la investigación de Ferguson, llegan a la conclusión de que hay dos causas que originan la aparición de esos factores. Una es la ambigüedad en la definición de lo que se entiende por homogeneidad, si se refiere al contenido, a la dificultad o a ambas. Otra es el fallo en la determinación de la correlación utilizada para obtener la matriz de intercorrelaciones. En su estudio llegan a dos conclusiones:

1. Si los ítems son homogéneos tanto en contenido como en dificultad, no importa cuál sea la correlación utilizada, ya que no aparecerán factores espúreos.

2. Si los ítems son homogéneos en contenido pero no en dificultad, sólo puede utilizarse la correlación tetracórica, pues de otro modo aparecen factores de dificultad.

Estas afirmaciones están fundadas en la forma en que se obtiene cada uno de los coeficientes. Supongamos que poseemos un conjunto de ítems de distintos niveles de dificultad; Wherry y Gaylord suponen que pertenecen a una escala próxima al tipo «escala perfecta» de Guttman. El ejemplo es aplicable también a cualquier conjunto de ítems que respondan a un modelo IRT. Si ordenamos a los ítems por dificultad creciente tendremos:

$$b_1 \leqslant b_2 \leqslant b_3 \leqslant \ldots \leqslant b_n$$

Entonces las proporciones de sujetos que responden correctamente a cada ítem formarán una sucesión monótona decreciente, ya que la probabilidad de responder correctamente a cada ítem es una función inversa de la dificultad. Por tanto,

$$p_1 \geqslant p_2 \geqslant p_3 \geqslant \ldots \geqslant p_n$$

Una fórmula ordinariamente utilizada para el cálculo de la correlación ϕ viene dada por:

$$\phi = \frac{ad - bc}{\sqrt{(k \cdot l \cdot m \cdot n)}} \tag{1}$$

(Downie, N. M., y Heath, R. H., 1973: pp. 127-128), donde los valores a, b, c, d y k, l, m, n, son las frecuencias de la tabla 1.

		k	l	
	1	b	a	m
	0	d	c	n
ítem i	j	0	1	N

Tabla 1.

Si deseamos obtener la correlación entre dos ítems cualesquiera veremos que la proporción de sujetos que supera simultáneamente los dos ítems, es igual que la proporción de sujetos que superan el ítem más difícil. Si llamamos P_{ij} a esa proporción, P_i a la proporción de sujetos que responden correctamente al ítem i, y P_j la proporción de sujetos que responden correctamente al ítem j, y si $b_j > b_i$, entonces:

$$P_{ij} = P_j \ll P_i$$

si en la ecuación (1) dividimos numerador y denominador por N^2 tenemos:

$$\phi = \frac{\frac{a}{N}\frac{d}{N} - \frac{b}{N}\frac{c}{N}}{\sqrt{\left(\frac{k}{N}\frac{l}{N}\frac{m}{N}\frac{n}{N}\right)}}$$

Es fácil comprobar que:

$$\frac{a}{N} = P_{ij} \qquad \frac{k}{N} = 1 - P_i = Q_i$$

$$\frac{b}{N} = P_j - P_{ij} \qquad \frac{l}{N} = P_i$$

$$\frac{c}{N} = P_i - P_{ij} \qquad \frac{m}{N} = P_j$$

$$\frac{d}{N} = (1 - P_j) - \frac{d}{N} = (1 - P_j) - P_i + P_{ij}$$

$$\frac{n}{N} = 1 - P_j = Q_j$$

luego (1) puede escribirse como:

$$= \frac{P_{ij}[(1 - P_j) - P_i + P_{ij}] - (P_j - P_{ij})(P_i - P_{ij})}{\sqrt{(P_i\, P_j\, Q_i\, Q_j)}} =$$

$$= \frac{P_{ij} - P_{ij}P_j - P_{ij}P_i + P_{ij}^2 - P_jP_i + P_jP_{ij} + P_{ij}P_i - P_{ij}^2}{\sqrt{(P_i\, P_j\, Q_i\, Q_j)}} =$$

$$= \frac{P_{ij} - P_j\, P_i}{\sqrt{(P_i\, Q_i\, P_j\, Q_j)}}$$

Con el tipo de datos que hemos supuesto se cumplía que $P_{ij} = P_j$, por tanto, podemos escribir:

$$= \frac{P_j - P_j P_i}{\sqrt{(P_i Q_i P_j Q_j)}} = \frac{P_j (1 - P_i)}{\sqrt{(P_i Q_i P_j Q_j)}} = (\text{como } 1 - P_i = Q_i) =$$

$$= \frac{P_j Q_i}{\sqrt{(P_i Q_i P_j Q_j)}} = \frac{\sqrt{(P_j Q_i)}\sqrt{(P_j Q_i)}}{\sqrt{(P_j Q_i}\sqrt{P_i Q_j)}} = \sqrt{\left(\frac{P_j Q_i}{P_i Q_j}\right)} =$$

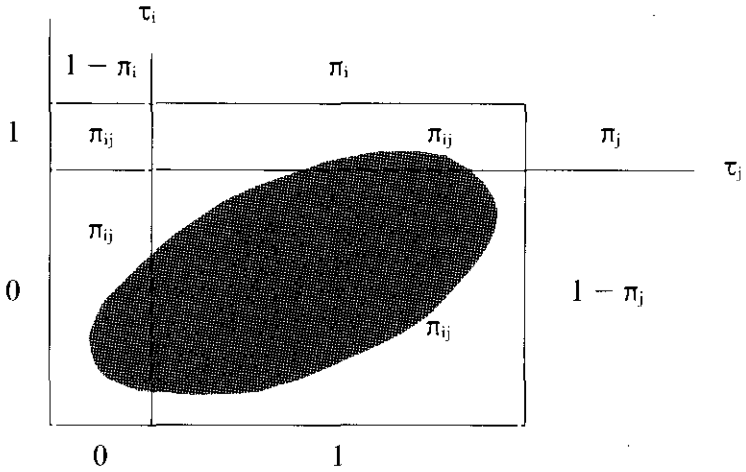
$$= \sqrt{\left(\frac{P_j}{P_i}\right)}\sqrt{\left(\frac{1 - P_i}{1 - P_j}\right)} = \sqrt{\left(\frac{P_j - P_i P_j}{P_i (1 - P_j)}\right)} =$$

$$= \sqrt{\left(\frac{P_j - P_i P_j + P_i - P_i}{P_i (1 - P_j)}\right)} = \sqrt{\frac{(P_i - P_i P_j) - (P_i - P_j)}{P_i (1 - P_j)}} =$$

$$= \sqrt{\frac{P_i (1 - P_j)}{P_i (1 - P_j)} - \frac{(P_i - P_j)}{P_i (1 - P_j)}} = \sqrt{\left(1 - \frac{(P_i - P_j)}{P_i (1 - P_j)}\right)}.$$

En esta expresión vemos claramente que ϕ sólo podrá tomar el valor 1 cuando $P_i - P_j = 0$, es decir, cuando la dificultad de los ítems sea exactamente igual. Cuando los ítems tengan índice de dificultad muy distintos, las correlaciones diferirán mucho de 1. Eso lleva a que aunque los tests estén formados por un conjunto de ítems unidimensionales en cuanto a contenido, la matriz de intercorrelaciones será de rango superior a 1 debido a que se agruparán los ítems de similares dificultades.

¿Por qué las correlaciones tetracóricas no dan lugar a factores de similares dificultades? La correlación tetracórica parte del supuesto de que a cada ítem subyace una variable cuya distribución es normal, y cuya distribución conjunta es una distribución binormal. Sean esas dos variables Y_i^* y Y_j^*(Lord y Novick, 1968: pp. 345-346). Supongamos que cada una de ellas se dicotomiza para los valores $Y_k^* = \tau_k$. A partir de ese momento los valores que se observan son los de dos variables dicotómicas, u_i y u_j que toman los valores $u_i = 1$ cuando $Y_i^* > \tau_i$ y $u_i = 0$ cuando $Y_i^* < \tau_i$, y $u_j = 1$ si $Y_j^* > \tau_j$ y $u_j = 0$ si $Y_j^* < \tau_j$.



El problema ahora consiste en estimar el valor de la correlación con los valores π_{ij}, π_i, y π_j, que son, respectivamente, la proporción de sujetos que responden correctamente los ítems i y j, los que responden correctamente al ítem i y los que responden correctamente al ítem j.

La distribución normal bivariante tiene como función de densidad:

$$f(Y_i, Y_j) = \frac{1}{2\pi\sqrt{(1-R^2)}} \exp\left[-\frac{1}{2(1-R^2)}(Y_i^2 + Y_j^2 - 2RY_iY_j)\right]$$

donde R = correlación entre Y_i e Y_j, y además con la condición de:

$$\varepsilon\{Y_i\} = \varepsilon\{Y_j\} = 0 \text{ y } \varepsilon\{Y_i^2\} = \varepsilon\{Y_j^2\} = 1.$$

(McFarlane Mood, 1965: p. 167).

La probabilidad de que un sujeto en general esté situado en una región del plano definido por Y_i, Y_j, viene dada por:

$$P\,[(Y_i, Y_j)\ \ \varepsilon R] = \int_R \int f(Y_i, Y_j)\, dY_i\, dY_j$$

los valores, por tanto, de π_{ij}, π_i y π_j están determinados por:

$$\pi_{ij} \equiv \int_{\tau_i}^{\infty} \int_{\tau_j}^{\infty} f(Y_i, Y_j; P)\, dY_i\, dY_j;$$

$$\pi_i \equiv \int_{\tau_i}^{\infty} \int_{-\infty}^{\infty} f(Y_i, Y_j; P)\, dY_i\, dY_j;$$

$$\pi_j \equiv \int_{-\infty}^{\infty} \int_{\tau_j}^{\infty} f(Y_i, Y_j; P)\, dY_i\, dY_j;$$

si τ_i y τ_j y R se conocen, los valores de π_{ij}, π_j y π_i están determinados. Y si se conocen π_{ij}, π_i y π_j, τ_i, τ_j y R también están determinados. La resolución de este sistema de ecuaciones implica la resolución de sistemas de integrales dobles, y acarrea bastantes problemas de cálculo.

Existe otra cuestión relacionada con los factores de dificultad, como es las relaciones de no-linealidad entre los ítems y el rasgo latente y entre los mismos ítems. Wherry y Gaylord (1944) plantearon por primera vez la cuestión en los siguientes términos. Supongamos que tenemos un test cuyos ítems están ordenados de más fáciles a más difíciles. Imaginemos que ese test es unidimensional, y que decidimos separarlo en dos mitades. La primera conteniendo a los ítems fáciles y la segunda a los difíciles. Si el test fuese suficientemente fiable, ocurriría que para que un sujeto tuviese en la mitad fácil una puntuación inferior al máximo, su puntuación en la mitad difícil debería ser necesariamente cero. Y para que un sujeto puntuase por encima de cero en la mitad superior del test, su puntuación en la otra mitad debería ser el máximo. Si dibujamos la regresión de una mitad sobre la otra, tendríamos una figura como 2a. Para obtener una figura como la 2b:

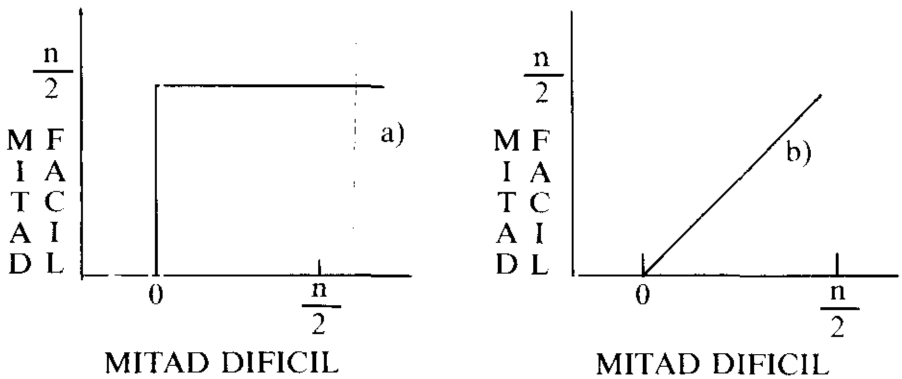


FIG. 2.

Cada mitad debería recoger ítems fáciles, intermedios y difíciles en la misma proporción que la otra mitad. El cálculo de una correlación que suponga la linealidad de regresión de las variables, en el caso a) llevará a errores de bulto, ya que es evidente que esa linealidad no se da. Esto puede aplicarse a los ítems, simplemente haciendo que cada mitad esté compuesta por un solo ítem.

McDonald (1967, 1987) y McDonald y Ahlawat (1974) mantienen que si la regresión de los ítems sobre el rasgo latente es una función polinomial con términos de segundo grado, un análisis factorial lineal proporcionará un factor de contenido y un segundo factor de dificultad. Al analizar la cuestión en profundidad, McDonald y Ahlawat (1974) llegan a las siguientes conclusiones:

1. Si la regresión de los ítems sobre los rasgos latentes responde realmente a un modelo lineal, al analizar los datos con un modelo factorial lineal no aparecerán factores de dificultad sea cual sea el tipo de correlaciones utilizadas.
2. Los factores para los que existe correlación entre las saturaciones y los índices de dificultad de los ítems no tienen por qué ser necesariamente espúreos. El análisis factorial tradicional no permite detectar cuándo lo son.
3. Los datos generados por una regresión polinomial cuadrática producen, al ser sometidos a análisis factorial tradicional un factor de curvatura cuyos coeficientes correlacionan con los índices de dificultad.
4. Los datos generados por el modelo de ogiva normal producen factores espúreos debido a la no linealidad.

Lumsden (1976) concuerda con estas conclusiones y propone los que él llama análisis en cascada. Los ítems se dividen en tres índices de dificultad y se analiza cada uno de los niveles por separado. Se comienza por el nivel más fácil, y los ítems más difíciles de ese nivel se pasan en la siguiente etapa al segundo nivel. Se analiza a continuación el segundo nivel procediendo en etapas sucesivas como con el primero. Este método garantiza una cierta homogeneidad de los niveles de dificultad de los ítems, y, por tanto, previene la aparición de los factores de dificultad.

ALGUNAS MODIFICACIONES PROPUESTAS AL MODELO FACTORIAL GENERAL CON APLICACION AL ANALISIS DE ITEMS

No podemos abandonar el tema sin hacer, al menos, una rápida mención de las alternativas al modelo general de análisis factorial desarrolladas para solventar algunos de los inconvenientes antes mencionados. Se trata, por una parte, del problema que plantea el nivel de medida de las variables que reflejan el rendimiento en los ítems y, por otra parte, la presencia de relaciones no lineales entre esas variables y entre esas variables y las variables latentes. En lo que sigue se hace una breve descripción de las alternativas propuestas acerca de estos temas.

ANALISIS FACTORIAL DE VARIABLES DICOTOMICAS

La mayoría de los métodos de análisis factorial suponen que las variables observadas están medidas en un nivel de intervalo. Cuando se aplica este modelo directamente a variables dicotómicas como los ítems de un test, se utilizan o bien los coeficientes ϕ, con el problema ya tratado de la posible aparición de factores de dificultad, o bien correlaciones tetracóricas, que producen muchos casos Heywood y matrices no definidas positivas en ocasiones.

Christofferson (1975) presentó una alternativa a los métodos de máxima verosimilitud corrientes, las estimaciones condicional e incondicional. Su método se basa en la utilización para la estimación de las *distribuciones marginales de los ítems por separado y por parejas.* Eso, evidentemente, supone no utilizar más que una parte de las 2^p marginales. Pero la pérdida de información se ha demostrado no ser importante.

La idea básica consiste en considerar que las variables observadas del modelo factorial general son, en realidad, variables continuas latentes en cada uno de los ítems. Es algo similar a lo que se hace para obtener las correlaciones tetracóricas. La diferencia es que aquí no se obtienen esas correlaciones. La variable manifiesta y* toma dos valores, 1 y 0, en función de si la variable latente y alcanza o no el punto de corte:

$$y_i^* = \begin{cases} 1 \text{ si } y_i \geqslant h_i \\ 0 \text{ si } y_i < h_i \end{cases}$$

la distribución de los valores que toma el vector de respuestas y* viene dada por:

$$g(y^*) = \int_{h_1}^{\infty} \int_{-\infty}^{h_2} \int_{h_3}^{\infty} \dots \int_{-\infty}^{h_m} f(x)\, dx \quad (2)$$

donde y* = (1, 0, 1, ..., 0) y f (x) es la normal multivariante, suponiendo que todas las variables están normalizadas, la distribución marginal de probabilidades de los valores de un item cualquiera i, serán:

$$P_i^* = P(y_i^* = 1) = \int_{h_i}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-(x^2/2)}\, dx \tag{3}$$

$$Q_i^* = P(y_i^* = 0) = 1 - P_i^*$$

la distribución marginal de probabilidades de los valores de dos ítems viene dado por:

$$P_{i.j}^* = P(y_i = 1, y_j = 1) = \int_{h_i}^{\infty} \int_{h_j}^{\infty} \frac{1}{2\pi (\Sigma_{ij})^{1/2}}\, e^{(x'\Sigma ij x/2)}\, dx \tag{4}$$

donde $\Sigma_{ij} = \begin{bmatrix} 1 & \sigma_{ij} \\ \sigma_{ji} & 1 \end{bmatrix}$ y $\sigma_{ij} = \sum_{r=1}^{k} \sum_{s=1}^{k} \mu_{ir}\mu_{js}\phi_{rs}$

donde ϕ_{rs} es el r, s-ésimo elemento de ϕ, la matriz de varianza-covarianza reproducida por los ítems i y j.

El método de Christofferson consiste en utilizar las ecuaciones (4) y (3) en lugar de la (2), que es más completa pero presenta más problemas de cálculo.

Para la estimación utiliza un eficiente procedimiento de mínimos cuadrados generalizados.

Muthén (1978) presenta un método similar al de Christofferson en el que se basa, pero utiliza para la estimación la información procedente de las proporciones de primer y segundo orden para ajustar el modelo. Se simplifica, en este caso, bastante el proceso de estimación con esas transformaciones de los datos originales. Muthén también utiliza un estimador de mínimos cuadrados generalizados que es tan eficiente asintóticamente como el anterior, aunque en este caso tiene la ventaja de ahorrar tiempo de cálculo.

Christofferson (1975) demostró que la utilización de la información de las marginales en lugar de las 2^p proporciones suponía una pérdida en eficiencia en la estimación muy pequeña. El estimador GLS (mínimos cuadrados generalizados) tenía la ventaja de exigir menos cálculos que el de máxima verosimilitud. La aportación de Muthén no es un cambio fundamental en este contexto, aunque permite simplificar aún más los cálculos.

Estos dos métodos tienen el inconveniente de que aunque permiten analizar más variables que con máxima verosimilitud en la ecuación 2, método en el que el máximo posible no pasaba de 10 ó 15 variables, no permiten sobrepasar las 20 ó 25.

Tanto estas dos soluciones como las que se basan en ULS (mínimos cuadrados no ponderados), pueden entenderse como soluciones de «información limitada», ya que utilizan sólo la información de las marginales de orden inferior de la tabla de contingencia completa de 2^p casillas. Una solución de «información completa» (Full-Information Factor Analysis) es desarrollada por Bock y Aitkin (1981), Bock, Gibbons y Muraki (1985), y cuyos antecedentes se remontan a Bock y Lieberman (1970). Esta solución ha terminado en el programa Testfact, de Wilson, Wood y Gibbons (1983). Este procedimiento, conocido como estimación marginal de máxima verosimilitud, ha sido detalladamente explicado en Gaviria Soto (1988). La función que se maximiza es:

$$P(r \mid \eta, \Gamma) = \frac{N!}{r_1!\, r_2! \ldots r_3!} P_1^{r1}\, P_2^{r2} \ldots P_s^{rs} \qquad (5)$$

donde N es el número de sujetos, r_l son las frecuencias con que se presentan cada una de las s distintas respuestas, y $P_l^{r_l}$ es la probabilidad de la respuesta l, η son las cargas factoriales y Γ los puntos de corte de las variables latentes que dan origen a las variables dicotómicas. Como puede verse la expresión (5) responde a la ley de distribución multinomial.

Recientemente, Mislevy (1986) ha dado un repaso a todos estos modelos planteando el estado actual de la cuestión. Es un buen punto de partida para quien desee profundizar en estas interesantes variantes del análisis factorial para variables dicotómicas.

EL MODELO DE ANALISIS FACTORIAL NO LINEAL

Este modelo fue desarrollado, principalmente, por McDonald (1962, 1965, 1967a, 1967b, 1967c), presentando una variante no lineal del modelo de factor común, en el que la regresión de las variables observadas sobre los factores comunes era no lineal. Para la estimación utilizó un procedimiento de mínimos cuadrados en dos etapas para modelos polinomiales y modelos con términos de interacción.

Posteriormente, Gnanadesikan y Wilk (1969) y Carroll (1969), describieron métodos de análisis de componentes principales polinomiales para representar conjuntos de datos multivariados en un espacio dimensional más pequeño que el original.

El modelo básico desarrollado por McDonald (1967a) puede expresarse como:

$$Z_i = \sum_{p=1}^{r} a_{ip} Q_p (x_1, ..., x_t) + u_i l_i$$

donde Z_i (i = 1, ..., n) es el conjunto de las *n* variables observadas, siendo $r \ll n$. Normalmente se considera que:

$$\varepsilon \{Z_i\} = 0, \ \varepsilon \{Z_i^2\} = 1, \ i = 1, ..., n.$$

los símbolos $x_i, ..., x_t$, en general x_m, son el conjunto de t variables latentes que son estadísticamente independientes entre sí y estadísticamente independientes de los factores únicos l_i, y los l_i son mutuamente independientes entre sí.

La función de densidad conjunta de las variables x_m y l_i en la población, dado que son estadísticamente independientes entre sí, viene dada por:

$$dF = \prod_{m=1}^{t} f_m(x_m)\, dx_m \prod_{i=1}^{n} g_i(l_i)\, dl_i.$$

como antes, se supone sin pérdida de generalidad que:

$$\varepsilon\{l_i\} = 0,\ \varepsilon\{l_i^2\} = 1,\ i = 1, ..., n.$$

las funciones $Q_p(x_1, ..., x_t)$, siendo $t \leqslant r \leqslant n$, son un conjunto de r funciones linealmente independientes de las variables latentes x_m, y tienen un número finito de parámetros. Se supone que la matriz de los parámetros a_{ip}, de orden $n \times r$, es de rango r.

En forma matricial el modelo es:

$Z = qA + eU$ donde:
$Z = [Z_1, ..., Z_n]$
$A' = a_{ip}$, matriz de orden $n \times r$ de rango r
$e = [e_1, ..., e_n]$
$U =$ matriz diagonal $[U_1, ..., U_n]$
$q = [Q_1(x_1, ..., x_t) ... Q_r(x_1 ... x_t)]$
$\varepsilon\{Z'Z\} = R$ ya que $\varepsilon\{Z\} = 0$ y $\varepsilon\{Z^2\} = 1$
y $R = A'\varepsilon\{q'q\}A + U^2$

que es análogo a lo obtenido en el modelo lineal de factores comunes.

El modelo puede expresarse en función de componentes ortogonales (McDonald, 1967a: p. 31).

Las distintas formas que adoptan las funciones Q_p determinan los modelos concretos en cada caso.

Etezadi-Amoli y McDonald (1983) han desarrollado un modelo en el que la función es:

$$Q_p = \Phi_j(\Theta_i) = \sum_{l=1}^{t} \sum_{q=1}^{q-1} \alpha_j l_q \Theta_{li}^q + \sum_{m=1}^{t-1} \sum_{l=m+1}^{t} \tau_{jml} \Theta_{mi} \Theta_{li} + \mu_j$$

para $j = 1, ..., n$; e $i = 1, ..., N$.

El modelo es lineal en sus coeficientes α_{jlq} y τ_{jml}, pero no lineal en los rasgos latentes $\Theta_1 ... \Theta_t$. Los términos de interacción se reducen a los productos dos a dos de los factores.

La experiencia parece recomendar no pasar de los términos de tercer grado en el polinomio, ya que los datos de las ciencias sociales rara vez justifica la utilización de términos más allá de este límite.

Como desarrollo de este modelo McDonald (1980) ha propuesto un modelo de análisis de estructuras de covarianza que sirve de base para estudios confirmatorios con modelos no lineales.

Cabe decir, como conclusión, que esta técnica es de gran interés teórico, aunque no se ha generalizado su uso. De hecho, hay una gran variedad de casos en los que sería aplicable con mejores resultados que su contrapartida lineal. La existencia de términos de interacción permite reflejar mejor la complejidad que, por lo general, presentan los datos de la realidad. En el caso de los modelos de respuesta al ítem, es claro que la regresión de las variables observadas sobre el rasgo latente es no lineal.

La desventaja más importante de este modelo estriba en su gran complejidad matemática y de cálculo. Así como de los modelos lineales hay una gran abundancia de textos divulgativos, de este modelo no hay muchas fuentes de información, son a veces de difícil acceso por estar sólo en publicaciones técnicas, y son de gran complejidad.

Por otra parte, aunque hay algún programa desarrollado para llevar a cabo análisis factorial no lineal, no están tan difundidos como los lineales. Y, además, mientras que los usuarios del análisis factorial tradicional tienen gran práctica en la interpretación de los resultados de sus investigaciones, no es fácil, sin embargo, en este modelo comprender el significado de algunos términos complejos, como los términos cuadráticos, cúbicos, cuádricos, etc., donde suelen reflejarse curvaturas, interacciones y otros fenómenos complejos.

La realidad es, sin embargo, compleja, y sería deseable que estos modelos alcanzasen en los próximos años mayor difusión y comprensión entre los investigadores en ciencias sociales.

CONCLUSIONES

1. El modelo de factores comunes no puede aplicarse directamente al análisis de ítems, aunque sea una práctica habitual, ya que algunos términos, como los factores específicos, no tienen fácil interpretación en el contexto de los ítems.
2. Cuando, a pesar de todo, se aplica el análisis factorial para el análisis de la dimensionalidad de un conjunto de ítems, no parece haber grandes diferencias entre la utilización del análisis de componentes principales o la estimación máximo-verosímil.
3. La aplicación indiscriminada de estas técnicas crea muchos problemas, el primero de ellos la aparición de factores de dificultad debido al uso de correlaciones ϕ y a la existencia simultáneamente de relaciones no lineales entre variables latentes e ítems.
4. De las modificaciones o variantes del análisis factorial desarrolladas en los últimos años, los modelos factoriales no lineales parecen los más prometedores, ya que permiten reflejar el tipo de relación que existe entre los ítems. Los que más se han desarrollado han sido modelos polinomiales de hasta tercer grado.

REFERENCIAS

BOCK, R. D., y AITKIN, M.: «Marginal maximum likelihood estimation of item parameters: Application of an EM algorithm». *Psychometrika*, vol. 46 (443-459), 1981.

BOCK, R. D.; GIBBONS, R. D., y MURAKI, E.: «Full-Information item factor analysis». CMRC Report, National Opinion Research Center Chicago.

BOCK, R. D., y LIEBERMAN, M.: «Fitting a response model for n dichotomously scored items». *Psychometrika*, vol. 35 (179-197), 1970.

CARROLL, J. D.: «Polynomial factor analysis». Proceedings of the 77th Annual Convention of the American Psychological Association, vol. 40 (103-104), 1969.

CHRISTOFFERSON, A.: «Factor analysis of dichotomized variables». *Psycometrika*, vol. 40 (5-32), 1975.

DOWNIE, N. M., y HEATH, R. W.: «Métodos estadísticos aplicados». Ed. Harla, México, 1973.

ETEZADI-AMOLI, J., y MCDONALD, R. P.: «A second generation nonlinear factor analysis». *Psychometrika*, vol. 48 (315-342), 1983.

FERGUSON, G. A.: «The factorial interpretation of test difficulty». *Psychometrika*, vol. 6 (323-329), 1941.

GAVIRIA SOTO, J. L.: «El supuesto de unidimensionalidad en la teoría del rasgo latente. Aportaciones metodológicas». Tesis doctoral no publicada, 1988.

GNANADESIKAN, R., y WILK, M. B.: «Data analytic methods in multivariate statistical analysis». En: Krishnaiah, P. R. (ed.): «Multivariate analysis: II». Academic Press, New York, 1969.

LORD, F. M., y NOVICK, M.: «Statistical theories of mental test scores». Addison-Wesley, Massachusets, 1968.

LUMSDEN, J.: «Test theory». En: Rosenzweig, M. R., y Porter, L. W. (Eds.). *Annual Review of Psychology*, Annual Reviews Inc. Palo Alto, CA, 1976.

MCDONALD, R. P.: «A general approach to nonlinear factor analysis». *Psychometrika*, vol. 27 (397-415), 1962.

MCDONALD, R. P.: «Difficulty factors and nonlinear factor analysis». *British Journal of Mathematical and Statistical Psychology*, vol. 18 (11-23), 1965.

MCDONALD, R. P.: «Nonlinear factor analysis». *Psychom. Monogr.*, vol. 15, 1967a.

MCDONALD, R. P.: «Numerical methods for polynomial models in nonlinear factor analysis». *Psychometrika*, vol. 32 (77-112), 1967b.

MCDONALD, R. P.: «Factor interaction in nonlinear factor analysis». *British Journal of Mathematical and Statistical Psychology*, vol. 20 (205-215), 1967c.

MCDONALD, R. P.: «A simple comprehensive model for the analysis of covariance structures: some remarks on applications». *British Journal of Mathematical and Statistical Psychology*, vol. 33 (161-183), 1980.

MCDONALD, R. P., y AHLAWAT, K. S.: «Difficulty factors in binary data». *British Journal of Mathematical and Statistical Psychology*, vol. 27 (82-99), 1974.

MCFARLANE MOOD, A.: «Introducción a la teoría de la estadística». Aguilar, Madrid, 1965.

MISLEVY, R. J.: «Bayes modal estimation in item response models». *Psychometrika*, vol. 51 (177-195), 1986.

WHERRY, R. J., y GAYLORD, R. H.: «Factor pattern of test items and tests as a function of the correlation coefficient: Content, difficulty, and constant error factors». *Psycometrika*, vol. 9 (237-244), 1944.

WILSON, D.; WOOD, R. L., y GIBBONS, R.: «Testfact: testing scoring and item factor analysis». Scientific Software, Chicago, 1983.

RESUMEN

En este trabajo se aborda uno de los problemas más importantes que aparecen cuando se analizan items dicotómicos. Una de las herramientas más utilizadas para descubrir la estructura interna de un conjunto de items es el análisis de factores comunes. En este estudio se demuestra que la utilización de esa técnica implica la adopción de ciertos supuestos referidos a la naturaleza de las variables y a la naturaleza de las relaciones entre ellas, cuya ignorancia tiene consecuencias graves para la determinación de la estructura latente de los items analizados. También se presentan algunas alternativas desarrolladas para evitar esos problemas.

SUMMARY

This work is concerned about one of the most important problems in the analysis of dichotomous items. The common factor analysis is one of the most used tools to discover the internal structure of an item bank. In this work it is shown that the use of this technic involves the adoption of some assumptions about the nature of variables and about the relationship among them, whose ignorance is a very serious problem to determine the latent structure of the analized bank. Some alternatives developed to prevent these problems are also presented.

Este trabajo forma parte de la investigación realizada por el autor para la tesis doctoral sobre el tema «El supuesto de Unidimensionalidad en la teoría del rasgo latente. Aportaciones metodológicas.»