

# Aplicación del estadístico $\pi^*$ de bondad de ajuste a modelos de estructura latente

Javier Revuelta y Dominique Kessel  
Universidad Autónoma de Madrid

Existen pocos estadísticos de bondad de ajuste aplicables a modelos de estructura latente. En este artículo se describe la aplicación del estadístico  $\pi^*$ , que se basa en el concepto de clases latentes, por lo que su interpretación resulta natural en este tipo de modelos. El estadístico  $\pi^*$  asume que la población está dividida en varias clases que siguen un modelo paramétrico, más una clase residual fuera del modelo. Se define  $\pi^*$  como la proporción de la población en la clase residual. En el artículo se describen los algoritmos de estimación de  $\pi^*$  en modelos de rasgo y clase latente, y se muestra un ejemplo con datos reales procedentes de una escala de hábitos de estudio. Se ha encontrado que hay una clase de malos y otra de regulares estudiantes, las cuales se relacionan con un criterio de responsabilidad.

*Application of the  $\pi^*$  goodness-of-fit index to latent structure models.* Testing model fit for latent structure models (latent trait models and latent class models) is difficult because of the lack of goodness-of-fit statistics with a known distribution. This paper describes the application of the  $\pi^*$  goodness-of-fit statistic to latent structure models. The statistic  $\pi^*$  is based on the concept of latent classes and has a natural interpretation when applied to these models. This statistic assumes that the population is made up of several classes that follow a parametric model, and a residual class outside the model. The value of  $\pi^*$  is the population proportion in the residual class. This paper describes the estimation algorithms of  $\pi^*$  for latent trait and latent class models and an empirical example with a scale of study habits. There are two latent classes in these data: bad and regular students, which are related to the student's responsibility.

El análisis de estructura latente es una teoría estadística que asume que las variables observadas pueden explicarse haciendo uso de variables no observadas o latentes (Langeheine y Rost, 1988; Heinen, 1996; Rost y Langeheine, 1997). Una variable latente es una variable perdida por diseño, es decir, cuyos valores no pueden observarse. Posteriormente, se utiliza la distribución condicionada de las variables observadas para estimar las variables latentes.

En los modelos de estructura latente lo realmente importante no son las variables observadas, sino las latentes, que se interpretan como niveles de conocimiento de los sujetos, actitudes, etc. El análisis de estructura latente fue sistematizado por Lazarsfeld y Henry (1968), y en términos generales comprende dos tipos de modelos. Los modelos de rasgo latente son aquellos que asumen que las variables latentes son continuas, por ejemplo, el análisis factorial o la teoría de respuesta al ítem. Los modelos de clase latente asumen que las variables latentes son discretas (Clogg, 1981, 1995; Elosua y López Jáuregui, 2005). Además, existen modelos híbridos que incluyen variables latentes discretas y continuas (Lindsay, Clogg, y Greco, 1991; Uebersax y Grove, 1993; Rost y von Davier, 1995; Rijmen y De Boeck, 2005).

Desde un punto de vista de teoría estadística, los modelos de clase latente se conocen como distribuciones de mezcla finita (McLachlan y Basford, 1988; McLachlan y Peel, 2000; Titterton, Smith, y Makov, 1985). Uno de los mayores problemas en la aplicación de estos modelos es la falta de estadísticos de bondad de ajuste. El estadístico de la razón de verosimilitudes ( $G^2$ ) se utiliza para contrastar la hipótesis de ajuste perfecto gracias a que su distribución asintótica es chi-cuadrado (Mood, Graybill, y Boes, 1974). Sin embargo, en modelos de mezcla finita, la distribución asintótica de  $G^2$  es desconocida, por lo que son necesarios otros métodos de bondad de ajuste.

El objetivo de este artículo es describir la aplicación del estadístico  $\pi^*$  a modelos de estructura latente. El estadístico  $\pi^*$  fue desarrollado por Rudas, Clogg y Lindsay (RCL; 1994) para evaluar la bondad de ajuste en modelos aplicables a tablas de contingencia. La interpretación de  $\pi^*$  es meramente descriptiva (su distribución es desconocida) e indica la proporción poblacional de individuos fuera de un determinado modelo estadístico. Hay una ventaja importante en utilizar  $\pi^*$  en lugar de  $G^2$  para evaluar descriptivamente la bondad de ajuste. El estadístico  $\pi^*$  tiene una interpretación clara en términos de clases latentes, cosa que no sucede con  $G^2$ . Además, al ser una probabilidad,  $\pi^*$  se define en el intervalo  $[0, 1]$ ; por el contrario,  $G^2$  no tiene un máximo definido.

El resto del artículo se organiza del siguiente modo. En primer lugar se describen los modelos de estructura latente. En segundo lugar, se describe el estadístico  $\pi^*$  de bondad de ajuste. A conti-

---

Fecha recepción: 31-3-06 • Fecha aceptación: 13-10-06

Correspondencia: Javier Revuelta  
Facultad de Psicología  
Universidad Autónoma de Madrid  
28049 Madrid (Spain)  
E-mail: javier.revuelta@uam.es

nuación se desarrollan los métodos de estimación. En cuarto lugar, estos métodos se aplican a una matriz de datos reales. La última sección son las conclusiones de esta investigación.

Modelos de estructura latente

En este artículo se utiliza como modelo psicométrico un modelo de mezcla finita de modelos logísticos de dos parámetros (MF-2PL). Supongamos que  $y_{ij}$  indica la respuesta del sujeto  $i$  al ítem  $j$ , y toma los valores 0 y 1. Según el 2PL, la probabilidad de  $y_{ij}$  condicionada en  $\theta$  es (Ferrando y Lorenzo Seva, 1993):

$$P(y_{ij}|\theta) = \frac{\exp(y_{ij}(\alpha_j\theta_i - \beta_j))}{1 + \exp(\alpha_j\theta_i - \beta_j)}$$

El vector de parámetros del ítem se denomina  $\epsilon_j = (\alpha_j, \beta_j)$ . En el modelo de Rasch se asume  $\alpha_j = \alpha$  (es decir, toma el mismo valor para todos los ítems; Prieto y Delgado, 2003). De acuerdo con el supuesto de independencia local, los elementos de un vector de respuestas  $y_i$  son estadísticamente independientes para un valor fijo de  $\theta$ , por lo que:

$$P(y_i|\theta) = \prod_{j=1}^J P(y_{ij}|\theta)$$

Un modelo de mezcla finita asume que la población está dividida en  $C$  subpoblaciones o clases, siendo  $\tau_c$  la probabilidad de muestrear a un sujeto de la clase  $c$ . En cada clase, la variable  $y_{ij}$  sigue el 2PL con parámetros  $\epsilon_{jc}$ ; y  $\theta$  tiene una función de densidad  $f(\theta)$  que en el resto del artículo será normal estándar. Entonces, la probabilidad marginal de  $y_i$  es

$$P(y_i) = \sum_{c=1}^C \tau_c \int_{-\infty}^{\infty} P(y_i|\theta) f(\theta) \quad (1)$$

La ecuación 1 es el modelo MF-2PL. El caso particular más obvio es el 2PL, en el cual  $C=1$  y  $\tau_c=1$ . Otro caso particular son los modelos de clase latente, en los cuales la variable latente es categórica. Un modelo de clase latente se obtiene fijando  $\alpha_{jc}=0$  para todo  $j$  y  $c$ . De este modo, el rasgo latente  $\theta$  desaparece de la función de probabilidad condicionada de  $y_{ij}$ , la cual se convierte en

$$P(y_{ij}|c) = \frac{\exp(y_{ij}(-\beta_{jc}))}{1 + \exp(-\beta_{jc})}$$

Finalmente, la función de probabilidad de  $y_i$  toma la forma

$$P(y_i) = \sum_{c=1}^C \tau_c P(y_i|c) = \sum_{c=1}^C \tau_c \prod_{j=1}^J P(y_{ij})$$

El modelo MF-2PL incluye variables latentes continuas y discretas, admitiendo como casos particulares los modelos de rasgo y clase latente.

El estadístico  $\pi^*$

El estadístico  $\pi^*$ , propuesto por RCL, asume que la población está compuesta por  $C+1$  clases. Las  $C$  primeras clases siguen un determinado modelo paramétrico, mientras que la clase  $C+1$  es una clase residual sin restricciones. La probabilidad de obtener un sujeto de las  $C$  primeras clases es  $(1-\pi)$ , por lo que  $\pi$  es la probabilidad de obtener un sujeto de la clase residual. Aplicado al MF-2PL, la probabilidad de obtener el vector de respuestas  $y$  es:

$$P(y) = (1-\pi) \sum_{c=1}^C \tau_c \int_{-\infty}^{\infty} P(y;\theta) f(\theta) + \pi R(y),$$

siendo  $R(y)$  la probabilidad residual del vector  $y$ . Como la clase residual no está sujeta a restricciones, este modelo ajusta perfectamente a los datos para un valor suficientemente alto de  $\pi$ . Se define el estadístico  $\pi^*$  como el valor más pequeño de  $\pi$  para el cual el ajuste es perfecto. Si  $p_y$  es la verdadera probabilidad de  $y$ , entonces:

$$\pi^*(p) = \inf\{\pi : p_y = P(y); \text{ para todo } y\}$$

En la aplicación práctica, el estimador de  $p_y$  es la proporción de sujetos que dan como respuesta el vector  $y$ . Sea  $n_y$  la frecuencia observada. Entonces,  $\pi^*$  es el valor más pequeño de  $\pi$  para el cual  $G^2=0$ , donde:

$$G^2 = 2 \sum_y n_y \log \left( \frac{n_y}{nP(y)} \right)$$

RLC demuestran que el intervalo confidencial al 95% para  $\pi^*$  toma la forma  $(\pi_L, 1)$ , donde  $\pi_L$  es el valor de  $\pi$  para el cual  $G^2=2.7$ .

Estimación de parámetros y de  $\pi^*$

La estimación de  $\epsilon$  y  $\pi^*$  se realiza mediante un algoritmo EM. La integral sobre  $\theta$  se evalúa numéricamente, aproximándola mediante una suma sobre  $G$  nodos de cuadratura ( $\theta_g$ ), cada uno con un peso  $f(\theta_g)$ . Entonces, la probabilidad posterior del par  $(\theta_g, c)$  es

$$f(\theta_g, c|y, \epsilon) \approx \frac{(1-\pi)\tau_c P(y|\theta_g, c) f(\theta_g)}{(1-\pi) \sum_{c=1}^C \tau_c \sum_{g=1}^G P(y|\theta_g, c) f(\theta_g) + \pi R(y)}$$

El algoritmo EM se aplica para estimar  $\epsilon$  con un valor fijo de  $\pi$ . El valor inicial de  $\epsilon$ , denominado  $\epsilon^{(0)}$ , se toma al azar, y los valores iniciales de  $R(y)$ , denominados  $R^{(0)}$ , son iguales al recíproco del número de posibles patrones de respuesta diferentes. El algoritmo se descompone en un paso E y uno M, que en la iteración  $s$  consisten en lo siguiente:

1. Calcular  $f(\theta_g, c|y, \epsilon)$  para cada  $\theta_g$  y  $c$ , y las frecuencias posteriores  $f_{ygc}^{(s)} = n_h f(\theta_g, c|y, \epsilon^{(s)})$ .
2. Resolver las ecuaciones de estimación de todos los parámetros. Las ecuaciones para los parámetros  $\alpha_j$  y  $\beta_j$  son:

$$\sum_y n_y \sum_{c'=1}^C \tau_{k'} \sum_{g=1}^G (y_j \theta_g - P(y_{j=1}) \theta_g) f_{ygc}^{(s)} = 0,$$

$$\sum_y n_y \sum_{c'=1}^C \tau_{k'} \sum_{g=1}^G (-y_j + P(y_{j=1})) f_{ygc}^{(s)} = 0,$$

De esta manera se obtiene un vector de parámetros actualizado  $\epsilon^{(s+1)}$ . El vector de probabilidades residuales se actualiza del modo:

$$R(y)^{(s+1)} = \frac{f(R(y)^{(s)} | y, \epsilon^{(s+1)}) n_y}{\sum_{y'} f(R(y')^{(s)} | y', \epsilon^{(s+1)}) n_{y'}}$$

donde

$$f(R(y)^{(s)} | y, \epsilon^{(s+1)}) \approx \frac{\pi R(y)^{(s)}}{(1 - \pi) \sum_{c=1}^C \tau_c \sum_{g=1}^G P(y_h | \theta_g, c) f(\theta_g) + \pi R(y)^{(s)}}$$

Ambos pasos se repiten hasta que  $G^2$  no disminuye en dos iteraciones consecutivas. El estimador máximo verosímil de  $\epsilon$  se obtiene aplicando este algoritmo con  $\pi = 0$  (Revuelta, Abad, y Ponsoda, 2006; López Pina, 1995). El estimador de  $\pi^*$  se obtiene aplicando el algoritmo con distintos valores de  $\pi$ , hasta que se encuentra el valor más pequeño que resuelve la ecuación  $G^2 = 0$ .

Una vez fijados los parámetros de los ítems, es posible obtener los estimadores de  $\theta$  y la clase de cada sujeto. El estimador de  $\theta$  es el valor que resuelve la ecuación de estimación:

$$\sum_{c=1}^C \tau_c \left( \sum_{j=1}^J y_j \alpha_j \right) f(\theta, c|y) = \sum_{c=1}^C \tau_c \left( \sum_{j=1}^J P(y_j = 1) \alpha_j \right) f(\theta, c|y)$$

donde la probabilidad posterior de  $\theta$  y  $c$  es

$$f(\theta, c|y) = \frac{P(y|\theta) f(\theta) \tau_c}{P(y)}$$

En el modelo de clase latente, la probabilidad posterior de  $c$  es

$$f(c|y) = \frac{\tau_c P(y|c)}{P(y)}$$

El estimador de la clase es el valor de  $c$  que maximiza  $f(c|y)$ .

### Análisis de una escala de hábitos de estudio

En esta sección se describe la aplicación del estadístico  $\pi^*$  para evaluar la bondad de ajuste de varios modelos de rasgo y clase latente. Se han utilizado datos procedentes de una escala de preguntas de hábito de estudio para estudiantes universitarios. La escala ha sido elaborada para esta investigación, y se ha escogido este rasgo por ser un tema relevante para todos los estudiantes (Rosario, Núñez, González-Pienda, Almeida, Soares, y Rubio, 2005; Pons-Salvador, Cerezo, y Bernabé, 2005). El objetivo de la

investigación es de carácter metodológico, en el sentido de que no se intenta contrastar una hipótesis previa acerca de los hábitos de estudio. Por el contrario, se investiga si el estadístico  $\pi^*$  es útil en la práctica para evaluar el ajuste de distintos modelos de estructura latente y compararlo con otros estadísticos de bondad de ajuste.

### Recogida y análisis de datos

*Material.* La escala consta de 10 preguntas, cada una con cinco alternativas de respuesta: 1) totalmente en desacuerdo; 2) en desacuerdo; 3) neutral; 4) de acuerdo; y 5) totalmente de acuerdo. La tabla 1 muestra los ítems, cinco son directos y cinco inversos. Para los ítems directos, las categorías 1 a 3 se codificaron como 0 y las categorías 4 y 5 como 1. Para los ítems inversos, las categorías 1 y 2 se codifican como 1 y las categorías 3 a 5 como 0. La escala se ha aplicado durante el tiempo de clase a una muestra de 100 estudiantes universitarios pertenecientes a dos grupos de una asignatura de psicometría, los cuales conocían que el propósito de la recogida de datos era la investigación y que sus respuestas no les acarrearían ninguna repercusión.

*Datos analizados.* Como la muestra es pequeña, se han analizado por separado los ítems directos y los inversos. De esta manera, el número de posibles respuestas diferentes a cada subescala es  $2^5 = 32$ . Al haber 100 sujetos, el número de sujetos es mayor que el de posibles patrones de respuesta, aunque la muestra de cada subescala sigue siendo pequeña para evaluar la bondad de ajuste utilizando el estadístico  $G^2$ . Además, para cada sujeto se han registrado las variables sexo, nota media en la licenciatura y puntuación en una escala de cinco ítems de responsabilidad.

*Modelos y análisis.* Se han aplicado los siguientes modelos: modelo de Rasch, modelo de clase latente con una clase a cinco clases latentes. La bondad de ajuste se ha evaluado mediante los estadísticos  $G^2$ ,  $\pi^*$ , AIC (Akaike Information Criterion) y BIC (Bayesian Information Criterion; Zucchini, 2000). El estadístico  $G^2$  únicamente sigue (asintóticamente) la distribución chi-cuadrada

Tabla 1  
Ítems del cuestionario de hábitos de estudio

Ítems directos	Enunciado
1	Acostumbro a organizar lo que tengo que estudiar por orden de importancia
2	Suelo tener mis apuntes de clase organizados
3	Acostumbro a marcar en mis apuntes lo que no he entendido para aclararlo más adelante
4	Suelo planificar la materia que voy a estudiar en un tiempo determinado
5	Suelo asistir a tutorías en el momento en el que me surgen las dudas
Ítems inversos	
1	Me gusta estudiar viendo la televisión o conectado al Messenger, aunque sé que me puede distraer
2	Cuando estudio, me levanto frecuentemente por cualquier motivo
3	Casi nunca estoy al día con los ejercicios de clase
4	Me distraigo fácilmente en clase
5	Con frecuencia preparo los exámenes pocos días antes

Nota: cada ítem se responde con las categorías: 1) totalmente en desacuerdo; 2) en desacuerdo; 3) neutral; 4) de acuerdo; y 5) totalmente de acuerdo.

do para el modelo de Rasch y en el modelo de una clase; en los modelos con más de una clase latente la distribución de este estadístico es desconocida, por lo que su interpretación es meramente descriptiva. Los estadísticos  $\pi^*$ , AIC y BIC se interpretan de manera descriptiva, y el modelo que los minimiza es el que obtiene mejor ajuste. Se ha escrito un programa informático en lenguaje C++ para realizar la estimación de modelos y el cálculo de  $\pi^*$  debido a que este último no está incluido en ningún programa comercial.

Resultados

*Bondad de ajuste*

La tabla 2 muestra los 32 posibles patrones de respuesta que pueden obtener los sujetos y los datos para la subescala de ítems directos. En concreto, la frecuencia observada de cada patrón, la frecuencia esperada con varios modelos y los estadísticos de bondad de ajuste. Los grados de libertad se calculan como  $gl = \text{número de patrones} - \text{número de parámetros estimados} - 1$ .

Las frecuencias observadas y residual se obtienen como parte del cálculo del estadístico  $\pi^*$ . La frecuencia esperada de cada patrón de respuestas es

$$n(1 - \pi^*) \sum_{c=1}^C \tau_c \int_{-\infty}^{\infty} P(\mathbf{y}; \theta) f(\theta),$$

por lo que la suma de las frecuencias esperadas es el número de sujetos que actúan según el modelo ( $n(1 - \pi^*)$ ). Puede verse que las discrepancias más importantes entre frecuencias observadas y esperadas se dan para el modelo de una clase y el modelo de Rasch, en concreto, en los patrones extremos (todo acuerdo o todo desacuerdo) y también en el patrón 01000.

Según  $G^2$ , el modelo de una clase debe rechazarse y el modelo de Rasch ajusta a los datos. El p-valor es desconocido para los modelos de dos y tres clases, aunque, descriptivamente, reducen en más de la mitad el valor de  $G^2$  del modelo de una clase. Los estadísticos AIC y BIC también favorecen el modelo de Rasch sobre cualquier otro; sin embargo, entre los modelos de clase latente, el AIC favorece el de tres clases y el BIC el de una.

La tabla 2 muestra también el resultado del estadístico  $\pi^*$ . Puede apreciarse que los modelos de Rasch y de una clase dejan fuera a más de un 50% de la población. Al introducir una segunda clase en el modelo, el valor de  $\pi^*$  baja un 22%, y con una tercera clase la reducción es del 8%. Con la cuarta y quinta clase la reducción es del 8 y 4%, respectivamente. Por tanto, el estadístico  $\pi^*$  indica que hay dos o tres clases de sujetos, como los demás estadísticos favorecen los modelos simples, parece razonable interpretar los datos utilizando el modelo de dos clases.

*Parámetros estimados de los ítems*

La tabla 3 muestra las probabilidades de acuerdo para los modelos con una a tres clases latentes. Los parámetros  $\beta_j$ , del modelo de Rasch son -0,38, -0,48, 0,00, -0,58 y 1,94; el valor de  $\alpha$  es 1.03.

A diferencia del modelo de Rasch, en los modelos de clase latente las probabilidades de acuerdo no están sujetas a ninguna restricción. En los modelos de dos y tres clases existe un orden natural entre las clases; la probabilidad de acuerdo se incrementa a lo largo de las clases, exceptuando para los ítems tres y cinco en el modelo de tres clases. Según el modelo de dos clases, más del 60%

de los sujetos se ubican en la segunda clase, que se caracteriza por tener una probabilidad de acuerdo mayor que la clase uno para todos los ítems.

En el modelo de dos clases latentes con ítems directos, la clase mayoritaria está formada por los alumnos con hábitos de estudio regulares; el resto tienen malos hábitos de estudio. Los ítems que más diferencian entre grupos son el 1 (Acostumbro a organizar lo que tengo que estudiar por orden de importancia), 3 (Acostumbro a marcar en mis apuntes lo que no he entendido para aclararlo más adelante) y el 4 (Suelo planificar la materia que voy a estudiar en un tiempo determinado). El ítem 2 (Suelo tener mis apuntes de clase organizados) tiene una probabilidad alta en ambos grupos, posiblemente porque la organización de los apuntes está implícita en el ritmo de explicación del profesor, y no depende tanto del trabajo del alumno. Resulta destacable que todos los alumnos, buenos o malos estudiantes, tienen una baja probabilidad de asistir a tutorías (ítem 5), prefiriendo el estudio individual a la consulta con los profesores. Por tanto, parece que las diferencias entre clases se deben al trabajo que realiza el alumno para organizar y planificar la materia de estudio.

Los resultados de bondad de ajuste para los ítems inversos son similares a los de los ítems directos, por lo que también se selecciona el modelo de dos clases. La tabla 4 contiene las probabilidades estimadas de acuerdo en cada ítem y los tamaños de las clases. La interpretación de las dos clases es similar a la de los ítems directos. Sin embargo, el resultado más destacable es que la clase mayoritaria (74%) es la de los sujetos con malos hábitos de estudio. Las diferencias entre grupos se deben fundamentalmente a los ítems 2 (Cuando estudio, me levanto frecuentemente por cualquier motivo) y 5 (Con frecuencia preparo los exámenes pocos días antes). Los ítems 1, 3 y 4 apenas diferencian entre clases. El 1 porque su probabilidad es baja en ambas, y el 3 y 4 porque es alta en ambas. La interpretación de los ítems inversos es que la diferencia entre clases está en el hecho de ponerse a estudiar. La mayoría los sujetos se distraen en clase y no llevan la materia al día. Sin embargo, hay sujetos que cuando se ponen a estudiar aprovechan el tiempo, y otros que estudian poco y distraen. Los parámetros  $\beta_j$ , del modelo de Rasch son -1.13, 0.49, 1.19, 1.19 y 0.74.

*Parámetros estimados para los sujetos*

En el modelo de Rasch hay seis valores estimados de rasgo diferentes. Cada uno se corresponde con una de las posibles puntuaciones en la escala (calculada como suma de las puntuaciones en los ítems) que es el estadístico suficiente para  $\theta$  en este modelo. Estos valores aparecen en la tabla 2, junto con su error típico y la clase en que se clasifica cada sujeto en el modelo de dos clases.

El estimador de  $\theta$  guarda relación con la clase estimada para cada sujeto ( $F_{1,98} = 125.57, p < 0.001$ ). La figura 1 muestra el diagrama de dispersión entre ambas variables. Los triángulos y los valores numéricos del gráfico son las medias de cada grupo. Los valores altos de  $\theta$  se relacionan con la clase 2.

En cambio, los valores de  $\theta$  que se calculan para cada sujeto mediante las dos escalas no se relacionan entre sí ( $r = 0.163, p = 0.106$ ). Asimismo, no existe relación entre la distribución de clases encontrada mediante las dos escalas. Aunque los significados de cada una de las clases parecen ser idénticos, la proporción de estudiantes agrupados en una u otra resulta muy diferente ( $n_0 = 22, n_1 = 78$  para directos y  $n_0 = 76, n_1 = 24$  para inversos), y hay independencia en las clasificaciones ( $\chi^2 = 1.661, p = 0.197$ ).

<i>Tabla 2</i> Frecuencias observadas y esperadas de cada patrón de respuestas, estadísticos de bondad de ajuste, clase y rasgo estimado. Datos para los ítems directos							
Patrón	<i>n</i>	<i>n</i> esperada			Estimadores		
		1 clase	2 clases	3 clases	Rasch	Clase	Rasch: $\theta$ (Se)
00000	7	0.00	7.17	7.33	0.09	1	-1.23(0.73)
00001	0	0.00	0.00	0.00	0.00	2	-0.72(0.70)
00010	5	2.82	3.76	4.71	5.53	1	-0.72(0.70)
00011	0	0.00	0.00	0.00	0.00	2	-0.23(0.69)
00100	3	0.00	3.02	0.00	0.01	1	-0.72(0.70)
00101	2	0.00	0.00	0.00	0.00	1	-0.23(0.69)
00110	2	2.39	2.21	0.00	1.46	2	-0.23(0.69)
00111	0	0.00	0.00	0.00	0.00	1	0.25(0.70)
01000	7	0.00	7.15	6.66	0.03	1	-0.72(0.70)
01001	0	0.00	0.00	0.00	0.00	1	-0.23(0.69)
01010	4	3.83	4.12	4.28	3.83	1	-0.23(0.69)
01011	0	0.00	0.00	0.00	0.00	1	0.25(0.70)
01100	3	0.00	3.01	3.05	0.01	2	-0.23(0.69)
01101	1	0.00	0.00	1.06	0.00	1	0.25(0.70)
01110	6	3.25	2.54	5.84	1.88	1	0.25(0.70)
01111	2	0.00	0.00	2.03	0.00	2	0.76(0.72)
10000	3	0.00	0.00	2.28	0.07	2	-0.72(0.70)
10001	0	0.00	0.00	0.00	0.00	1	-0.23(0.69)
10010	8	7.76	8.09	8.13	7.97	2	-0.23(0.69)
10011	0	0.00	0.00	0.00	0.00	2	0.25(0.70)
10100	0	0.00	0.00	0.00	0.02	2	-0.23(0.69)
10101	1	0.00	0.00	0.00	0.00	2	0.25(0.70)
10110	7	6.59	7.15	3.08	3.91	2	0.253(0.70)
10111	2	0.00	0.00	0.00	0.00	2	0.76(0.72)
11000	2	0.00	0.00	2.07	0.05	1	-0.23(0.69)
11001	0	0.00	0.00	0.00	0.00	2	0.25(0.70)
11010	10	10.56	10.29	9.85	10.28	2	0.25(0.70)
11011	4	0.00	0.00	0.00	0.00	2	0.76(0.72)
11100	8	0.00	0.00	2.79	0.02	1	0.25(0.70)
11101	1	0.00	0.00	0.97	0.00	2	0.76(0.72)
11110	9	8.97	9.09	9.28	9.49	2	0.76(0.72)
11111	3	0.00	0.00	1.85	0.00	2	1.30(0.75)
$\pi^*$		0.54	0.32	0.24	0.55		
$\pi_1^*$		0.45	0.23	0.15	0.38		
$G^2$		42.88	29.30	15.75	34.88		
<i>gl</i>		26	20	14	26		
<i>p</i>		0.021			0.070		
<i>AIC</i>		640.04	638.45	636.91	632.03		
<i>BIC</i>		653.06	667.11	681.20	645.06		

Nota: *n* es la frecuencia observada; *n* esp. es la frecuencia esperada y *n* res. es la frecuencia residual.  $G^2$  es el estadístico de la razón de verosimilitudes; el *p*-valor aparece entre paréntesis. *gl* son los grados de libertad. *AIC* es el criterio de información de Akaike, y *BIC* es el criterio de información bayesiano. La clase se ha estimado con el modelo de dos clases y  $\theta$  con el modelo de Rasch

Esta incongruencia, por un lado, puede ser debida a que las escalas están formadas por muy pocos ítems y, por otro, a que los estudiantes no hayan respondido de la misma forma a los ítems directos e inversos. Los ítems formulados inversamente son más difíciles de entender porque pueden resultar ambiguos por dar lugar

a una doble negación. Además, requieren una cierta atención por parte del sujeto que, de lo contrario, mantendrá el mismo patrón de respuesta de los ítems directos (por ejemplo siempre «de acuerdo»).

Respecto a las variables criterio se ha encontrado una relación significativa sólo para la Responsabilidad. Tanto para la escala di-

recta como para la inversa, las puntuaciones  $\theta$  se relacionan moderadamente con el rasgo de Responsabilidad. Existe una relación lineal positiva según la cual sería posible predecir  $\theta$  a partir de la Responsabilidad ( $r= 0.491$ ,  $F_{1,98}= 31.060$ ,  $p<0.001$  y  $r= 0.365$ ,  $F_{1,98}= 15.102$ ,  $p<0.001$ , para directos e inversos, respectivamente). En caso de la Nota Media la correlación con  $\theta$  no resulta significativa para los ítems directos ( $r= 0.160$ ,  $F_{1,95}= 2.510$ ,  $p= 0.016$ ). Para los inversos se asumirá lo mismo por ser muy débil la relación encontrada ( $r= 0.240$ ,  $F_{1,95}= 6.184$ ,  $p= 0.015$ ). Para la variable Sexo la diferencia de medias en  $\theta$  no resulta significativa para ninguna de las dos escalas, por lo que se puede concluir que  $\theta$  no guarda relación con el sexo ( $t_{82}= -0.326$ ,  $p= 0.745$  para directos y  $t_{82}= -0.712$ ,  $p= 0.478$  para inversos).

Respecto a la relación entre las variables criterio y las clases, las observaciones son similares a las anteriores. Se obtiene una relación significativa sólo para la Responsabilidad ( $t_{98}= -3.958$ ,

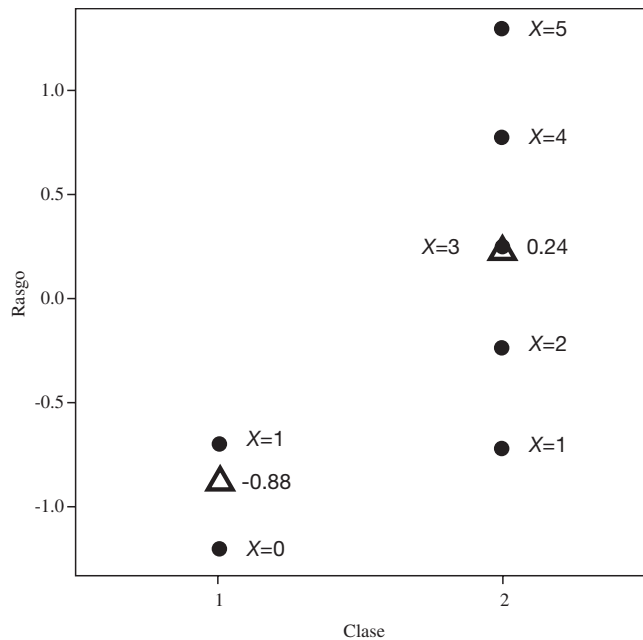
*Tabla 3*  
Probabilidades de acuerdo, tamaños de las clases y niveles estimados de riesgo. Ítems directos

Ítem	Modelo	Clase 1	Clase 2	Clase 3
1	C1	0.58		
	C2	0.09	0.72	
	C3	0.27	0.51	1.00
2	C1	0.60		
	C2	0.42	0.65	
	C3	0.46	0.70	0.62
3	C1	0.50		
	C2	0.19	0.59	
	C3	0.00	1.00	0.37
4	C1	0.62		
	C2	0.27	0.72	
	C3	0.40	0.51	1.00
5	C1	0.16		
	C2	0.00	0.27	
	C3	0.00	0.24	0.22
<i>Tamaño</i>	C1	1.00		
	C2	0.38	0.62	
	C3	0.38	0.32	0.30

Nota: C1 es el modelo de una clase, C2 el de dos clases y C3 el de tres clases. La tabla muestra las probabilidades de acuerdo para cada una de las clases latentes o grupos de sujetos. La fila *Tamaño* muestra las probabilidades estimadas de las clases.

*Tabla 4*  
Probabilidades de acuerdo para el modelo de dos clases y tamaño de las clases. Ítems inversos

Ítem	Clase 1	Clase 2
1	0.38	0.00
2	0.74	0.20
3	0.81	0.50
4	0.76	0.64
5	0.81	0.19
<i>Tamaño</i>	0.74	0.26



**Figura 1.** Diagrama de dispersión del nivel de rasgo (eje de ordenadas) y la clase para el modelo de dos clases (eje de abscisas). Los puntos negros son los datos de los sujetos individuales, los triángulos y valores numéricos son las medias de los grupos. Los valores de X son las puntuaciones en la escala. Resultados para los ítems directos

$p<0.001$  para directos y  $t_{98}= -2.546$ ,  $p= 0.012$  para inversos). No hay relación ni con la Nota Media ( $t_{95}= -1.652$ ,  $p= 0.102$  en directos y  $t_{95}= -2.250$ ,  $p= 0.027$  en inversos) ni con el Sexo ( $\chi^2= 1.400$ ,  $p= 0.237$  para directos e inversos).

### Discusión y conclusiones

Actualmente hay pocos estadísticos de bondad de ajuste aplicables a modelos de estructura latente. La utilización de  $G^2$  es problemática, bien porque la muestra es pequeña o porque su distribución es desconocida, como sucede en el caso de los modelos de mezcla finita. Con respecto al problema de la muestra pequeña, Maydeu-Olivares y Joe (2005) han propuesto los estadísticos de información limitada, que evalúan la bondad de ajuste utilizando frecuencias marginales de la tabla de contingencia de los patrones de respuesta.

Para evaluar los modelos con más de una clase latente sólo es posible utilizar estadísticos descriptivos, como el propio  $G^2$ , AIC, BIC y  $\pi^*$ . La ventaja de  $\pi^*$  es que tiene una interpretación clara en términos de clases latentes, como la proporción de individuos en el grupo residual. Su principal inconveniente es la complejidad de su cálculo, tanto analíticamente como en tiempo de cálculo, pues se basa en una modificación del procedimiento de estimación marginal-EM. Además, con muestras pequeñas es habitual encontrar sobrestimaciones del valor de  $\pi^*$ ; no obstante, este problema también puede afectar a los otros estadísticos, por lo que requeriría una investigación más amplia.

Se ha utilizado una muestra de datos reales para ejemplificar el uso de estos estadísticos. Se trata de una escala de hábitos de estudio, formada por ítems directos e inversos. Se ha encontrado que hay dos clases latentes, formadas por estudiantes malos y regulares. Con los ítems directos e inversos la interpretación de las cla-

ses es similar, y en ambos casos guardan relación con un criterio de responsabilidad. Sin embargo, las clasificaciones de los sujetos con los ítems directos e inversos no están relacionadas.

Como futura línea de investigación, RLC describen la aplicación de  $\pi^*$  a variables continuas que siguen una distribución normal bivariante. Podría intentarse la estimación de este estadístico

en modelos factoriales con variables normales, de modo que proporcionara información adicional sobre las características psicométricas de la muestra obtenida (Morales-Vives, Codorniz-Raga, y Vigil-Colet, 2005; Aciego, Domínguez, y Hernández, 2005; Ximénez y García, 2005). En concreto, acerca de la bondad de ajuste del modelo factorial.

## Referencias

- Aciego de Mendoza, R., Domínguez Medina, R., y Hernández Hernández, P. (2005). Consistencia interna y estructura factorial de un cuestionario sobre autorrealización y crecimiento personal. *Psicothema*, 17, 134-142.
- Clogg, C.C. (1981). Latent class models for measuring. En R. Langeheine y J. Rost (eds.): *Latent trait and latent class models*. New York: Plenum.
- Clogg, C.C. (1995). Latent class models. En G. Arminger, C.C. Clogg y M.E. Sobel (eds.): *Handbook of statistical modeling for the social and behavioral sciences*. New York: Plenum.
- Elosua, P., y López Jáuregui, A. (2005). Clases latentes y funcionamiento diferencial del ítem. *Psicothema*, 17, 516-521.
- Ferrando, P.J., y Lorenzo Seva, U. (1993). Algunas relaciones entre el modelo de un factor común y el modelo logístico de dos parámetros. *Psicothema*, 5, 403-412.
- Formann, A.K. (1985). Constrained latent class models. Theory and applications. *British Journal of Mathematical and Statistical Psychology*, 38, 87-111.
- Heinen, T. (1996). *Latent class and discrete latent trait models: Similarities and differences*. Newbury Park: Sage Publications.
- Langeheine, R., y Rost, J. (1988). *Latent trait and latent class models*. New York: Plenum.
- Lazarsfeld, P.F., y Henry, N. (1968). *Latent structure analysis*. New York: Houghton.
- Lindsay, B., Clogg, C.C., y Greco, J. (1991). Semiparametric estimation in the Rasch model and related exponential response models, including a simple latent class model for item analysis. *Journal of the American Statistical Association*, 86, 96-107.
- López Pina, J.A. (1995). Estimación de parámetros en la TRI. Una evaluación de Bilog en muestras pequeñas. *Psicothema*, 7, 173-185.
- Maydeu-Olivares, A., y Joe, H. (2005). Limited- and full-information estimation and goodness of fit testing in 2<sup>n</sup> contingency tables: An unified framework. *Journal of the American Statistical Association*, 100, 1009-1020.
- McLachlan, G.J., y Basford, K.E. (1988). *Mixture models. Inference and applications to clustering*. New York: Marcel Dekker.
- McLachlan, G.J., y Peel, D. (2000). *Finite mixture models*. New York: Wiley.
- Mood, A.M. Graybill, F.A., y Boes, D.C. (1974). *Introduction to the theory of statistics*. London: McGraw-Hill International.
- Morales-Vives, F., Josep Codorniu-Raga, M., y Vigil-Colet, A. (2005). Características psicométricas de las versiones reducidas del cuestionario de agresividad de Buss y Perry. *Psicothema*, 17, 96-100.
- Pons-Salvador, G., Cerezo, M.A., y Bernabé, G. (2005). Características psicológicas de adolescentes pertenecientes a comunidades educativas vulnerables. *Psicothema*, 17, 37-42.
- Prieto, G., y Delgado, A.R. (2003). Análisis de un test mediante el modelo de Rasch. *Psicothema*, 15, 94-100.
- Revuelta, J., Abad, F.J., y Ponsoda, V. (2006). *Modelos politómicos de TRI*. Madrid: La Muralla.
- Rosário, P., Núñez, J.C., González-Pienda, J.A., Almeida, L., Soares, S., y Rubio, M. (2005). El aprendizaje escolar examinado desde la perspectiva del «modelo 3P» de J. Biggs. *Psicothema*, 17, 20-30.
- Rijmen, F., y de Boeck, P. (2005). A relation between a between-item multidimensional IRT model and the mixture Rasch model. *Psychometrika*, 70, 1-16.
- Rost, J. (1997). Logistic mixture models. En W.J. van der Linden y R.K. Hambleton: *Handbook of modern item response theory*. New York: Springer.
- Rost, J., y Langeheine, R. (eds.) (1997). *Applications of latent trait and latent class models in the social sciences*. Muenster, Germany: Waxmann.
- Rost, J., y von Davier, M. (1995). Mixture distribution Rasch models. En G.H. Fischer, G., e I.W. Molenaar: *Rasch models. Foundations, recent developments and applications*. New York: Springer Verlag.
- Rudas, T., Clogg, C.C., y Lindsay, B. (1994). A new index of fit based on mixture methods for the analysis of contingency tables. *Journal of the Royal Statistical Society, Series B*, 56, 623-639.
- Titterton, D.M., Smith, A.F.M., y Makov, U.E. (1985). *Statistical analysis of finite mixture distributions*. New York: Wiley.
- Uebersax, J.S., y Grove, W.M. (1993). A latent trait finite mixture model for the analysis of rating agreement. *Biometrics*, 49, 823-835.
- Ximénez, C., y García, A.G. (2005). Comparación de los métodos de estimación de máxima verosimilitud y mínimos cuadrados no ponderados en el análisis factorial confirmatorio mediante simulación Monte Carlo. *Psicothema*, 17, 528-535.
- Zucchini, W. (2000). An introduction to model selection. *Journal of Mathematical Psychology*, 44, 41-61.