Papel de los modelos matemáticos en ciencias de la educación

Por Pedro SANCHEZ ALGARRA

En los últimos años, el papel del modelo como instrumento metodológico en las ciencias naturales, económicas y sociales se ha incrementado de forma tan espectacular que ha generado largas listas de obras o artículos sobre el tema; primero, para tratar de esclarecer el propio concepto de modelo, muy polémico, e indicador de una realidad epistemológica tal que, como dijo Paris (1972, p. 468), «los modelos constituirían semiteorías o conatos de teoría, última posibilidad allí donde aún no se ha conquistado un suficiente dominio técnico y metodológico»; y segundo, para estudiar con profundidad los distintos tipos de modelo, portadores de funciones específicas, y por tanto aplicables a situaciones diferentes.

Entre las múltiples clasificaciones que se han elaborado sobre los modelos, destaca la de Black (1962), en que se confiere carta de naturaleza a los modelos matemáticos, y que ha sido confirmada y ratificada por autores posteriores, si bien otros anteriores a él, como Arrow (1951), discuten acerca de la conveniencia del razonamiento matemático en las Ciencias Sociales. Pero en las Ciencias de la Educación ha escaseado la elaboración y construcción de modelos matemáticos, sin que hayamos reflexionado probablemente sobre su papel catalizador en la comprensión de una teoría, por parte deductiva, y estructurador de datos empíricos, por vía inductiva.

No pretendemos en este artículo otro fin que demostrar cómo a través de un modelo matemático de razonamiento se puede convertir (Bunge, 1980) una hipótesis sociopedagógica infundada en una hipótesis científica modesta pero fundada: «La mayoría de los alumnos que terminan su formación con el título de Graduado Escolar proceden de familias obreras», «La mayoría de los Licenciados son de familias burguesas», «La mayoría de los que terminan el BUP o una carrera de grado medio pertenecen a familias de pequeños burgueses».

Estas afirmaciones, planteadas así, son problemáticas para el pedagogo y quien planifique la enseñanza en sus distintos niveles. En primer lugar, ¿cómo caracterizamos a las familias según su clase social? Porque existen varios criterios-base en torno a los cuales sería posible una estratificación: prestigio ocupacional, ingresos, forma de vida, características y zona de la vivienda, etc., y con la presencia de una constante permeabilidad en la estructura social, donde la distribución de personas en torno a una serie de localizaciones significativamente diferentes varía en esta estructura, aunque tengamos que reconocer que la variación en la estructura ocupacional en las sociedades de algunos países, como Gran Bretaña y Dinamarca, e incluso España, en la década de los años 60, es pequeña en comparación con la variación entre todas las sociedades en el momento actual. Y, en segundo lugar, el haber cursado un nivel de estudios elemental o superior, ¿implica realmente una correspondencia posterior en el nivel de vida, por ejemplo?

Esquematizando las hipótesis iniciales, su forma general podría enunciarse:

Los hijos de familias F de la comunidad C tienden a realizar los estudios E (en donde apreciamos la existencia de un término impreciso: «tiende», que deberá aclararse).

Supongamos que nuestro universo es la comunidad C, que podemos dividir por circunscripciones territoriales (distritos, regiones, barrios, etc.), y en clases sociales, como profesiones liberales, obreros manuales, pequeños empresarios, etc. Llamemos F al grupo de familias incluido en la comunidad C, cuyo nivel de estudios E de los hijos nos interesa investigar. El resto de la comunidad será el complemento de F en C, es decir, $\overline{F}$. Si el número de zonas o circunscripciones en que se ha dividido el territorio que ocupa la comunidad C es n, tendremos que C es la unión de las n poblaciones C_i de dichos distritos o barrios. De forma parecida, el grupo de familias F es la unión de los n subconjuntos F_i de F que habitan tales circunscripciones. En resumen, tenemos:

$$C. = \bigcup_{i=1}^{n} C_i \qquad F. = \bigcup_{i=1}^{n} F_i, \qquad \text{donde } F_i \leqslant C_i$$

Se entiende que tenemos forma de saber si una persona de la comunidad C cursa los estudios E, a través de cuestionarios, encuestas, archivos, etc. Llamemos y_i a la fracción de individuos que habitan el i-ésimo distrito y estudian E (por ejemplo, y_i puede ser el número de universitarios que habitan la zona C_i). Según la hipótesis inicial, este número y_i es tanto menor cuanto más reducida sea la fracción x_i de los habitantes de la misma zona C_i que corresponden al grupo de familias F_i.

La frase «y_i es tanto menor cuanto más reducida sea la fracción x_i» puede dar lugar a un desdoblamiento de la hipótesis: a) hay una correlación positiva entre los valores de x y los de y; b) x e y están relacionadas funcionalmente entre sí, y la función que las relaciona es decreciente. Para decidir entre ambas son precisos datos empíricos que podemos obtener, y que formarán una nube de puntos en el plano x-y, más o menos dispersa, pero cuya tendencia general o línea de regresión podemos hallar.

La hipótesis más simple, y por esto sospechosa, es que dicha línea de regresión es una recta de pendiente a que corta al eje de las y en el punto (O, b). O sea,

$$H_1 : y_i = ax_i + b \quad \text{para todo} \quad 1 \leqslant i \leqslant n$$

Hay la posibilidad de que H_1 concuerde con los datos, aunque sea bastante lejano, en cuyo caso hay que interpretarlo, por ejemplo, atribuyendo a todos los individuos una propensión a cursar los estudios E; incluso podemos suponer (en primera aproximación) que esta propensión no depende de la zona, sino del tipo de familia a la que pertenece. Si el individuo pertenece al tipo de familia F de interés, le atribuimos la propensión p, siendo $O < p < 1$; y si pertenece a cualquier otro tipo de familia, es decir, $\overline{F}$, le atribuimos la propensión q, también $O < q < 1$, pero que según nuestra conjetura previa, es menor que p. Suponemos que p y q son probabilidades condicionales.

En resumen, postulamos:

$$H_2 : px_i + q \ (1—x_i), \quad \text{donde} \quad O < p, \ q \leqslant 1$$

y donde x_i es la fracción de la población de la comunidad C_i que pertenece al tipo de familia F, y $l\text{-}x_i$ es la fracción de los que no pertenecen a F. Hallamos que la pendiente es $a = p\text{-}q$, y la ordenada en el origen es $b = q$.

a y b han adquirido una clara interpretación sociopedagógica: $a = p\text{-}q$ es la ventaja que otorga a los estudios E la pertenencia al tipo de familia F, y b es la tendencia a estudiar E cuando no se pertenece a F.

La nueva hipótesis H_2, más profunda que H_1, es útil aunque sea rechazada por la evidencia empírica, es decir, si los datos muestran que la línea de regresión no es una recta. Es decir, si la tendencia general no es lineal, podemos ensayar la hipótesis de que p y q, lejos de ser constantes, son a la vez funciones de alguna variable. Hay tres posibilidades: a) p y q son funciones de densidad de las F en cada circunscripción (por ejemplo, cuantas más familias obreras hay en una zona mayor es la tendencia a que sus hijos sólo lleguen a la obtención del Graduado Escolar); b) p y q son funciones del número de estudiantes de E en cada distrito (es decir, la tendencia a realizar los estudios E es tanto mayor cuanto más estudiantes hay en el distrito); y c) p y q son funciones de una tercera variable por averiguar (por ejemplo, inteligencia, conocimiento de personas con aquel nivel de estudios, existencia de los Centros adecuados, u otras).

Para simplificar consideremos solamente los casos a) y b), y en ambos limitémonos al caso lineal, improbable pero simple.

La conjetura de que las proporciones p y q son funciones lineales de la densidad de los F en cada distrito se formularía:

$$p_i = a_1 x_i + b_1 \quad , \qquad q_i = a_2 x_i + b_2$$

Reemplazando en H_2 queda:

$$y_i = (a_1 x_i + b_1)\ x_i + (a_2 x_i + b_2)\ (1 - x_i)$$

Reordenando se obtiene una parábola:

$$H_3 : y_i = a_1 - a_2)\ x_i^2 + (b_1 + a_2 - b_2)\ x_i + b_2 \quad 1 \leqslant i \leqslant n$$

Si esta curva se ajusta bien a los datos, la damos de momento por buena. En caso contrario, ensayamos la alternativa b), es decir:

$$p_i = a_1 y_i + b_1 \quad , \qquad q_i = a_2 y_i + b_2$$

de modo que

$$y_i = (a_1 y + b_1)\ x_i + (a_2 y_i + b_2)\ (1 - x_i)$$

En definitiva queda la curva de regresión llamada homográfica:

$$H_4 : y_i = \frac{(b_1 - b_2)\ x_i + b_2}{(a_2 - a_1)\ x_i + 1 - a_2} \quad , \qquad 1 \leqslant i \leqslant u$$

Y en caso de que tampoco se ajustara bien a los datos tendríamos que ir ensayando sucesivamente las distintas hipótesis provinientes de c), una vez formuladas de forma explícita y rigurosa.

Ha quedado, creemos que claramente explicitado, el papel creativo del modelo (Harré, 1976), en el sentido de dotarnos progresivamente de un «corpus» teórico, o, siguiendo el patrón desarrollado por Bunge (1980), partiendo de una conjetura y/o de datos empíricos que sucesivamente vamos reorganizando en el camino lento, pero fructífero, de llegar a una explicación de unos hechos o eventos reales —en este caso de índole educativa o pedagógica— que sólo podrían ofrecer una descripción, pero nunca su comprensión.

REFERENCIAS

ARROW, K. J. **Social choice and individual values.** New York: Wiley, 1951.

BLACK, M. **Models and metaphors: Studies in language and philosophy.** Ithaca, New York: Cornell University Press, 1962.

BUNGE, M. **Epistemología. Curso de actualización.** Barcelona: Ariel, 1980.

HARRE, R. «The constructive role of models». In L. Collins (Ed.) **The use of models in the social sciences.** London: Tavistock Publications, 1976, 16-43.

PARIS, C. «Razón y experiencia en la metodología de los modelos». En J. Zubiri, **Homenaje** (vol. 2). Madrid: Moneda y Crédito, 1972, 463-473.