

El análisis de datos como útil y objeto de la didáctica de la matemática

C. Batanero, J. D. Godino y A. Vallecillos

Educación Matemática, 4(1), 46- 53 , 1992

Resumen

En los trabajos de índole etnoestadística sobre investigaciones en ciencias sociales y experimentales se ha puesto de manifiesto la existencia de dificultades y errores en la aplicación de los conceptos y procedimientos estadísticos. Debido a la naturaleza matemática de los conceptos y a los procesos didácticos implicados, en este trabajo analizamos el interés de abordar, desde la perspectiva de la Didáctica de la Matemática, estudios sistemáticos sobre esta problemática. El objetivo de este trabajo es atraer la atención de los investigadores en Didáctica de la Matemática hacia dos núcleos de interés.

- (1) La comparación del papel actual del análisis de datos en la investigación frente a sus posibilidades. De este estudio se deduce la necesidad de un cambio de actitudes hacia la Estadística y de una mayor participación de los investigadores en los procesos de diseño de la investigación y análisis de los datos, como medio de mejorar la calidad global de la investigación.
- (2) El diseño de planes de formación en razonamiento estadístico y análisis de datos para estudiantes de doctorado y maestría, basados en un estudio previo de dificultades, sesgos y obstáculos sobre este campo conceptual entre los propios investigadores.

En resumen, en este trabajo queremos resaltar el hecho de que el empleo del *útil estadístico* en las Ciencias experimentales y sociales, en general, y en la didáctica de la matemática, en particular, se convierte, en esta área de conocimiento, en *objeto específico de estudio*, debido a la naturaleza matemática de los conceptos y a los procesos didácticos implicados.

Introducción

La investigación en distintas ciencias experimentales y sociales y, en particular, en didáctica de la matemática necesita frecuentemente recoger y analizar datos sobre los fenómenos objeto de su estudio. En consecuencia, los conceptos y procedimientos estadísticos y las técnicas de análisis de datos, tanto cualitativos como cuantitativos, deben ser tema de consideración y debate en el seno de la comunidad de investigadores.

Entre los diferentes sistemas didácticos, un subsistema de enseñanza de especial importancia para la didáctica de la matemática lo constituyen los cursos de doctorado y maestría. Dentro de estos cursos, los conceptos y procedimientos de análisis de datos y, en general el razonamiento estadístico, constituyen una materia obligada. Como indican Harris y Kanji (1988), un conocimiento básico estadístico en los especialistas de las diferentes disciplinas es necesario para realizar por sí mismos las tareas estadísticas rutinarias específicas de sus disciplinas, para reconocer los errores estadísticos en los trabajos de su especialidad y para reconocer las situaciones en que precisan ayuda de un estadístico profesional. En el caso de la Didáctica de la Matemática, esta formación presenta además un carácter "recursivo": una mejor preparación de los investigadores inducirá una mejora en los

resultados de la investigación y estos resultados, a su vez, influyen en la mejora de la enseñanza.

Douady (1986) ha llamado la atención sobre dos modos de funcionamiento de un concepto matemático: como útil y como objeto. “Un concepto es útil cuando focalizamos nuestro interés sobre el uso que se hace de él para resolver problemas... Por objeto, entendemos el objeto cultural que tiene su lugar en un edificio más amplio que el saber sabio en un momento dado, reconocido socialmente” (pág. 9).

Nuestro propósito es llamar la atención en el modo particular de funcionamiento del objeto de saber que denominamos análisis de datos como herramienta para resolver problemas de investigación. Consideramos que el análisis de la dialéctica útil-objeto para este saber particular, en los cursos de formación de investigadores, es un nuevo “objeto” de estudio para la Didáctica. Este sentido de la palabra “objeto”, como tema o campo de estudio para la investigación didáctica, es el que deseamos destacar en este trabajo.

Estadística y análisis de datos

Como indica Benzecri (1982) los probabilistas del siglo XVIII fundan la estadística moderna a partir de las aplicaciones a la demografía y el desarrollo de la teoría de los errores. A pesar de su juventud, esta rama de las matemáticas es, sin embargo, una de las que mayor aplicación y desarrollo cuenta en la actualidad. A partir de las necesidades de muchas ciencias diferentes se han ido desgajando de ella ramas interdisciplinarias específicas como la biometría, la psicometría, la sociometría, etc. Los resultados teóricos en estadística obtenidos en estas diferentes ramas han sido de tal importancia, que han mostrado su utilidad general, pudiéndose hablar aquí realmente de una ciencia transdisciplinar en que las fronteras de lo que corresponde esencialmente a la estadística y lo que es propio de cada una de sus aplicaciones no quedan, con frecuencia, claramente delimitadas.

El advenimiento de los ordenadores ha permitido una rápida difusión y un uso intensivo de las técnicas estadísticas multivariantes y su aplicación al análisis de colecciones de datos progresivamente más diversos. Esto ha dado lugar a un nuevo enfoque que está recibiendo el nombre de Análisis de datos, especialmente en Francia, donde se han puesto a punto los métodos de análisis taxonómico y análisis de correspondencias. Detrás de este término hay toda una filosofía de investigación: análisis multivariante, extensión de las posibilidades de análisis a escalas de medida nominales y ordinales, enfoque exploratorio frente al confirmatorio, complementación de los aspectos probabilísticos con otros geométricos, topológicos y numéricos. Todo ello hace posible el estudio simultáneo de un gran número de hechos, con objeto de descubrir su estructura global, complementando, mediante la inducción y la síntesis, el análisis y la deducción posibilitados por el paradigma experimental.

El análisis de datos como útil en la investigación

Una primera faceta de estudio, que pretende encuadrarse en la investigación de carácter metodológico, es el análisis del papel que el análisis de datos está desempeñando como herramienta en la investigación experimental. Este papel está condicionado por las actitudes hacia la estadística por parte de los investigadores, que podemos clasificar en un continuo situado entre dos polos extremos.

El primero de estos extremos es la de los que creen que con unas técnicas apropiadas de análisis sería posible conseguir unos resultados extraordinarios a partir de cualquier conjunto de observaciones; es el análisis adecuado lo que hace una buena investigación. Esta creencia

puede llevar al empirismo exagerado y al olvido de la teoría, sin la cual ningún conjunto de datos puede obtener su sentido.

En la actualidad esta postura es especialmente peligrosa, debido a la existencia de paquetes de programas para el análisis de datos, que permiten la aplicación de estas técnicas de un modo rápido y sencillo. Ello ha llevado al empleo inadecuado de las mismas, como señala White (1980) para el caso de la educación. Este autor indica que las ciencias de la conducta tienen muchos problemas para identificar las relaciones entre las variables, al ser tantas y tan complejas, incluso cuando los métodos empleados están libre de errores. Cuando el empleo de las técnicas de investigación está plagado de errores evitables, la probabilidad de una conclusión errónea es bastante elevada.

Además, están comenzando a aparecer sistemas expertos que aconsejan sobre el tipo de análisis a realizar, y las hipótesis necesarias para su empleo, por lo cual, supuestamente podría evitarse tal empleo inadecuado mediante la utilización de estos nuevos recursos. Pero, como señala Hawkins (1990), el quitar al usuario la responsabilidad en el proceso de toma de decisiones esenciales impide el análisis y la interpretación sensibles e informados, dando la visión errónea de que el uso de la estadística es un proceso mecánico.

La dificultad del uso correcto de los conceptos y procedimientos estadísticos lleva en otros casos al rechazo precipitado del enfoque cuantitativo en la investigación. Esta visión constituye el otro extremo, compartido por un grupo de investigadores en el paradigma cualitativo, que prefieren prescindir de tales técnicas. Para ellos, el análisis de los datos ha de ser preferentemente interpretativo; la categorización, la codificación, la reducción de datos supone pérdida de una parte sustancial de la riqueza fenomenológica del mundo que se trata de explorar.

Sin embargo, esta postura supone en gran parte una ilusión de determinismo sobre el fenómeno observado; se argumenta que no se pretende la generalización, que el fenómeno observado es único. Pero todo proceso de toma de datos supone un muestreo; como indican Goetz y Lecompte (1988) incluso fijado el sujeto, se muestrean los tiempos, las circunstancias, las preguntas planteadas, las posibles reacciones, que provienen de poblaciones finitas y específicas.

Puede caerse también aquí en los que Kahneman y cols. (1982) denominan "creencia en la ley de los pequeños números", es decir, la creencia en que las características generales de un proceso estocástico han de reflejarse en el comportamiento del mismo, incluso con una muestra muy pequeña de ensayos. El "creyente en la ley de los pequeños números" sobreestima la potencia de sus métodos de investigación, estima a la baja la posibilidad de obtener resultados significativos por azar y tiene una confianza excesiva en la replicabilidad de sus primeros hallazgos.

Por otro lado, entre las técnicas aplicables al análisis cualitativo de datos, se van incorporando poco a poco algunas que clásicamente se han considerado como cuantitativas. Como ejemplo, Miles y Huberman (1984) en una exposición sistemática de tales técnicas señalan algunas de este tipo, como la aplicación de técnicas estadísticas a las frecuencias obtenidas en las diversas categorías, la clasificación o la factorización.

Afortunadamente se van abandonando estas posturas extremas y, cada vez con más frecuencia, las investigaciones se encuentran en un punto intermedio entre los paradigmas cuantitativo y cualitativo. En el caso de la educación matemática, esta idea de complementariedad de los métodos cualitativo y cuantitativo es sugerida por Kilpatrick (1981) que indica que en lugar de abandonar los métodos cuantitativos en favor de los cualitativos, "deberíamos dirigir nuestros esfuerzos en la dirección de enriquecerlos". En particular, sugiere que los investigadores en educación matemática deberían estudiar nuevas

técnicas de análisis exploratorio de datos y representación de resultados y considerar el uso de técnicas para el re-análisis de los datos y el meta-análisis de los estudios efectuados hasta la fecha.

La comunicación entre el investigador y el estadístico

Con frecuencia, el empleo de la estadística con fines de investigación requiere el trabajo en colaboración de los investigadores con estadísticos profesionales. En estas ocasiones, el estadístico es un resolutor de problemas planteados por otra persona, y se exige de él más de lo que razonablemente está capacitado para proporcionar. En palabras de Barnett (1988): "debe ser un traductor y un comunicador: necesita comprender lo suficiente de las otras disciplinas para apreciar sus problemas. Ha de expresarlos en términos estadísticos y, lo más importante, comunicar las respuestas en forma comprensible",... Todo lo que se espera de él es que sea maestro de su propio oficio estadístico, y conocedor de muchos otros".

El proceso global comprende, en el caso más general, desde la planificación de la forma en que se han de recoger los datos, hasta la interpretación de los resultados obtenidos en el análisis de los mismos. En este proceso hay aspectos exclusivamente estadísticos, como serían los relativos al diseño del muestreo, o la elección del método concreto de análisis y su realización. Pero hay otros cualitativos y específicos del problema como son la elección de las unidades de análisis, variables, escalas de medida y proceso de categorización y codificación. Incluso algunos aspectos del análisis, como son la elección de las medidas de similitud o distancias a emplear, que condicionan fuertemente los resultados, depende, grandemente, del problema planteado.

Es decir, entre la toma de datos y su tratamiento estadístico hay todo un proceso que no puede ser dejado a la responsabilidad exclusiva del estadístico, puesto que el significado de las categorías finales sólo se muestra a la luz de las preguntas que guían la investigación. Igualmente, como indica Dawis (1986), la interpretación de los resultados del análisis estadístico sólo tiene sentido profundo para la persona que ha planteado las preguntas.

Nuestra experiencia personal de colaboración como estadísticos en este tipo de trabajos con investigadores de diversas disciplinas nos ha hecho apreciar la enorme dificultad de comunicación entre el estadístico y el investigador en las fases iniciales y finales del proceso, que son, sin duda, las de mayor importancia. En este sentido se expresa también Speed (1988), quien indica que estas dificultades surgen por haber prestado una atención insuficiente al contexto no estadístico en que tiene lugar la discusión. Esta falta de acoplamiento del lenguaje estadístico y el lenguaje técnico, propio de la disciplina en la que se investiga, repercute en la calidad final del análisis efectuado que, en ocasiones, puede llegar a ser inadecuado por los motivos expuestos o bien puede ser interpretado incorrectamente en los informes finales de la investigación.

Formación de investigadores en el empleo de técnicas estadísticas

En esta sección deseamos llamar la atención sobre las particularidades que presenta la investigación sobre la enseñanza de la estadística y el análisis de datos en el seno de un sistema didáctico particular: los cursos de formación de investigadores, en aquellas disciplinas en que la formación previa en el tema sea insuficiente. La necesidad de un conocimiento básico en esta materia y, sobre todo, del desarrollo de un nivel adecuado de razonamiento estadístico es puesta de manifiesto por Schuyten (1990), que señala como básicas las siguientes competencias estadísticas para el desarrollo de la investigación.

(1) Competencias requeridas para la interpretación adecuada de la literatura de investigación.

- Conocimiento y comprensión de los métodos empleados en la investigación dada, su dominio de aplicación y las hipótesis en que se basan.
- Capacidad para juzgar las decisiones del investigador respecto a los aspectos metodológicos y evaluación frente a otras posibles.
- Capacidad para evaluar las interpretaciones dadas por el autor a los resultados del análisis, frente a otras interpretaciones alternativas.
- Capacidad para juzgar los rasgos del estudio y sus posibilidades de generalización de los resultados.

(2) Competencias requeridas para el desarrollo de la investigación propia:

Debido a la existencia de los paquetes de programas estadísticos, la destreza en el cálculo ya no es el punto principal necesario para emplear estas técnicas. Pero estos paquetes serán de poca ayuda en la planificación del estudio, en la elección de la técnica apropiada y en la interpretación de los resultados.

Como vemos, estos dos tipos de competencias resultan básicos para la calidad de la investigación. Las del primer grupo resultan indispensables para realizar un estudio crítico de la bibliografía de investigación y juzgar la validez y posibilidad de generalización de cada estudio particular. Sin esta capacidad crítica podremos juzgar cada resultado como más o menos elegante, como más o menos de acuerdo con la tendencia general de las investigaciones en una misma problemática. Seremos, sin embargo, incapaces de juzgar la mayor o menor evidencia experimental de las teorías desarrolladas, que aceptaremos, no por la fuerza de su fundamentación, sino por el prestigio del autor, o por nuestra afinidad con sus ideas. De este modo, corremos el peligro de que nuestro campo científico emergente derive en lo que Bunge (1985) denomina una pseudoconciencia o una ideología.

Este peligro es aún mayor, si se considera que, respecto a las competencias del segundo grupo, pueden verse estos vacíos conceptuales en los trabajos publicados en los que, con cierta frecuencia, se aprecian análisis incorrectos o incompletos, como es señalado por White (1980). Se corre, además, el peligro de que estos errores sean transmitidos y aumentados por la difusión de esta literatura y la imitación de dicha metodología incorrecta por parte de investigadores que, debido a una preparación insuficiente en este tema, confían en la adecuación de unos métodos empleados por otros investigadores.

La comunidad de investigadores constituye, en consecuencia, un “sistema didáctico” de características particulares ya que se identifican dos componentes del mismo: el saber -pensamiento estocástico- y los discentes -usuarios de dicho saber. Pero la función “docente” queda diluida en la dialéctica de las publicaciones científicas y en los expertos que imparten cursos de especialización o colaboran en el desarrollo de la misma.

Etnoestadística y Educación Matemática

Las cuestiones que hemos planteado están siendo objeto de consideración en un campo de investigación que Gephart (1988) ha denominado etnoestadística. “La preocupación fundamental de la etnoestadística es describir, analizar, explicar y comprender cómo se usa realmente la estadística en el proceso de investigación. La etnoestadística examina los aspectos cualitativos de la realización de análisis estadísticos, y utiliza este examen como fuente de comprensión de los procesos sociales que subyacen en el conocimiento científico”

(Gephart, 1988: p.11). El autor citado distingue tres órdenes o categorías de estudios etnoestadísticos:

- (1) Producción de datos estadísticos. En esta clase de estudios se investigan las actividades, significados y contextos implicados en la producción de la información estadística que ha de servir de base para el análisis. Estos etnógrafos se interesan por la observación empírica directa de la conducta y de las prácticas de los productores de datos estadísticos y de los entornos de investigación naturales en los que se producen y usan los mismos.
- (2) Adecuación del uso de los métodos e hipótesis. Se evalúa cómo el investigador, el sujeto, el instrumento de evaluación y el contexto social de medición afectan a los resultados de las investigaciones. Se trata pues de realizar análisis críticos de las propiedades técnicas del proceso de recogida y análisis de datos. Esta es una preocupación característica de las investigaciones metodológicas, si bien el rasgo distintivo de los estudios etnoestadísticos es la inclusión de hipótesis sobre usos y prácticas sociales.
- (3) Retórica de la estadística. Los estudios etnoestadísticos del tercer tipo tratan del empleo de la estadística como una clase de retórica. Examina los informes estadísticos como si se trataran de una clase de recurso literario que usa esta información para persuadir al lector.

Perspectiva didáctica de la etnoestadística

Los fenómenos que trata de poner de manifiesto la etnoestadística se refieren al uso de los conceptos y procedimientos estadísticos en la investigación, teniendo en cuenta el contexto social donde ésta tiene lugar. La hipótesis subyacente es que las restricciones personales e institucionales particulares donde estos saberes se usan les imponen significados específicos, con frecuencia inapropiados; en otras ocasiones por el desconocimiento de esta herramienta conceptual no se utiliza, suponiendo una pérdida de posibilidades en la obtención de conocimiento científico.

El investigador aplicado que recoge, analiza y comunica información estadística es un usuario de conceptos y procedimientos matemáticos con un conocimiento parcial y con frecuencia sesgado de los mismos. Este sujeto puede contemplarse como un “estudiante” (de hecho, incluso puede tener esta consideración académica) aunque esté inmerso en un “sistema didáctico” informal. La determinación de la relación personal al saber en cuestión en un momento y circunstancias dadas, así como de las condiciones de su evolución, constituyen un objeto de estudio para la investigación didáctica.

Esta clase de investigaciones constituirían una nueva categoría de estudios etnoestadísticos en la que se destacaría la dimensión didáctica de las relaciones institucionales y personales al saber (Chevallard, 1989), que en cierto modo tendría un carácter comprensivo de las otras tres categorías descritas por Gephart (1988). En este nivel el interés se centraría en la búsqueda de errores, dificultades y obstáculos en el uso de la estadística y la identificación de la naturaleza y condiciones de aparición de los mismos. Estos resultados pueden, a su vez, revertir en los escenarios tradicionales de enseñanza.

Conclusiones

En este trabajo hemos destacado la importancia del análisis de datos como útil en la investigación y, como consecuencia, hemos deducido la necesidad de desarrollar un razonamiento estocástico adecuado y unos conocimientos mínimos del tema en los cursos de formación de investigadores. De este modo, el estudio de las condiciones requeridas para dicho desarrollo se convierte en un objeto de investigación para la didáctica de la matemática.

Asimismo, la realización de estudios etnoestadísticos desde una perspectiva didáctica debe ser estimulada como un medio para mejorar y vitalizar la investigación y, en especial, la relativa a la Educación Matemática. Dado el uso tan extendido del Análisis de datos en las ciencias sociales y naturales, el estudio de las cuestiones planteadas sería también una contribución significativa al diálogo entre las ciencias, con el carácter transdisciplinar auspiciado por Steiner (1985).

Referencias

- BARNETT, V. (1988). Statistical consultancy. A basis for teaching and research. En R. Davison y J. Swift (Eds.), *Proceedings of the Second International Conference in Teaching Statistics*. University of Victoria, pp. 303-307.
- BATANERO, C., ESTEPA, A., GODINO, J. D. (1991). Análisis exploratorio de datos: sus posibilidades en la enseñanza secundaria. *Suma*, 9, 25-31.
- BATANERO, C., GODINO, J., STEINER, H. G., WELZELBURGER, E. (1992). The training of research in mathematics education. Results from a survey. *Educational Studies in Mathematics*, 26(1), 95-102.
- BENZECRI J.P. (1982). *Histoire et préhistoire de l'analyse des données*. Paris: Dunod.
- CHEVALLARD, Y. (1989). Le concept de rapport au savoir. Rapport personnel, rapport institutionnel, rapport officiel. *Séminaire de Didactique des Mathématiques et de l'Informatique*. Université Joseph Fourier. Grenoble.
- DAWID. A. P. (1986). Contribution to the discussion of "Statistical modelling issues in school effectiveness studies" by M. Aitki and N. Longford. *Journal of the Royal Statistical Society, Ser. A*, 149,1-43.
- DOUADY. R. (1986). Jeux de cadres et dialectique outil-objet. *Recherches en Didactique des Mathématiques*, 7(2), 5-31.
- GEPHART, R.P. (1988). *Ethnostatistics :Qualitative foundations for quantitative research*. Sage University Paper series in Qualitative Research Methods (Vol.2). Beverly Hills, CA: Sage.
- GOETZ, L. P. y LECOMPTE, M. D. (1988). *Etnografía y diseño cualitativo en la educación*. Madrid: MacGraw Hill.
- HARRIS, R. R. y KANJI, G. K. (1988). The challenges of statistical training. En: R. Davidson y J. Swift (Eds.). *Proceeding of the Second International Conference in Teaching Statistics*. University of Victoria, pp. 498-501.
- HAWKINS. A. (1990). Success and failure in statistical education. An U.K. perspective. Presentado en el *ICOTS III*. University of Otago.
- KAHNEMAN, D., SLOVIC. P. y TVERSKY, A. (1982). *Judgement under uncertainty: Heuristics and biases*. Amsterdam: North-Holland.
- KILPATRICK, J. (1981). Research in mathematical learning and thinking in the United States. *Recherches en Didactique des Mathématiques*, 2, pp. 363-379.
- MILES, M. B. y HUBERMAN, A. M. (1984). *Qualitative data analysis*. Beverly Hills, CA: Sage.
- SCHUYTEN. G. (1990). Statistical thinking in psychology and education. Presentado en el *ICOTS III*. University of Otago.

- SPEED. T. (1988). Questions, answers and statistics. En R. Davidson y J. Swift (Eds.). *Proceedings of the Second International Conference on Teaching Statistics*. University of Victoria, pp. 17-27.
- WHITE. A. L. (1980). Avoiding errors in educational research. En: R S. Shumway (Ed.). *Research in Mathematics Education*. Reston, VA: N.C.T.M., pp.47-65.