

¿Por qué no empezar a formular buenas preguntas?

Lourdes Figueiras Ocaña

Ilustración: Jaume Gubianas

Cuadernos de Pedagogía, Nº 438, Sección Tema del Mes, Octubre 2013, Editorial Wolters Kluwer España, ISBN-ISSN: 0210-0630

Una educación estadística de calidad debe conseguir que los alumnos sean capaces de formular buenas preguntas que les permitan comprender una situación o predecir tendencias. Para ello es necesario que dispongan de datos ricos y que dediquen más tiempo a conocer ejemplos de investigaciones, adaptados a su nivel, que utilizan la estadística de forma óptima. Aprender la ingeniería de la estadística puede ser más fácil si se tienen referentes adecuados.

Lourdes Figueiras Ocaña. Profesora de la Universitat Autònoma de Barcelona. Correo-e: Lourdes.Figueiras@uab.cat

Los decretos de enseñanza que regularon la antigua Educación General Básica en 1982 imponían como contenido que los niños y las niñas se habituaran a la recogida de datos y a su organización mediante tablas, cuadros numéricos, gráficos o representaciones a escala que les permitieran relacionarlos. Incluían también la preparación de encuestas y cuestionarios para el estudio de temas relacionados con la cultura, las profesiones, o la geografía del territorio próximo. En el año 2004 se abrió un camino explícito para la estadística en los documentos legales que regularon la Educación Primaria.

En el ámbito de las ciencias, y de la geografía y la historia la legislación se refirió a la importancia del papel del profesor para favorecer el desarrollo de técnicas de observación, representación y clasificación de datos, y para guiar un proceso continuo de reflexión sobre la información. En el ámbito de las matemáticas se enumeraron algunas técnicas básicas como la construcción de tablas de frecuencias, la iniciación intuitiva a las medidas de centralización, y la realización e interpretación de diagramas de barras, poligonales y sectoriales.

Al hilo de estas directrices quisiera detenerme a reflexionar sobre algunas cuestiones: ¿Cuál es el sentido que otorgamos a la estadística durante la Educación Primaria? ¿Cuáles son los aspectos en los que quisiéramos centrar la atención? ¿Deberíamos concretar su estudio en el ámbito de la matemática o en el ámbito más amplio del conocimiento social y natural?



El uso de la estadística en la escuela

Los argumentos utilizados para justificar la importancia de la estadística escolar pueden resumirse en dos: por una parte, su uso crucial para explorar cuestiones políticas, sociales o médicas; por otra, la formalización que ofrece de ciertos procesos más o menos intuitivos en los que todos nos vemos envueltos cuando repetimos una experiencia, buscamos regularidades o sacamos conclusiones basadas en observaciones.

Un ejemplo en el que se ponen de manifiesto ambos argumentos es la correlación, que mide hasta qué punto los cambios que se producen en una variable se asocian con el cambio en otra variable. Esta asociación, sin embargo, no nos dice a priori que los cambios en la primera sean la causa de los cambios en la segunda. Imaginemos que en una merienda en el jardín uno de los invitados se ha tragado una mariquita que había en la ensalada y una hora después le ha brotado un sarpullido en la espalda. Como resultado de esta observación puedo concluir que la ingestión de mariquitas provoca que aparezcan puntitos a ambos lados de la columna vertebral, aunque no parece muy riguroso establecer esta causalidad a partir de una única observación y sin tener en cuenta que pudo ser cualquier otro ingrediente lo que desencadenó una reacción alérgica.

Continuemos con otro ejercicio de fantasía un poco más elaborado: "Las personas que consumen caracoles caminan más lentamente que las que no lo hacen". La conclusión se basa en un estudio, llevado a cabo en una ciudad holandesa, en el que se comparó la velocidad a la que caminaban los comensales que no habían comido caracoles y los que sí, a la salida de un restaurante que ofrecía este plato en su carta. De nuevo, aunque los datos fueran reales y correctos, parece lógico pensar que la conclusión no es muy rigurosa, y que quienes acudieron al restaurante a comer caracoles fueron mayoritariamente originarios de algún país mediterráneo, habituados a su consumo, y de altura y zancada inferior a la de los holandeses.

Estos ejemplos caricaturizan relaciones falsas de causalidad entre variables relacionadas, frecuentes en los razonamientos intuitivos de la estadística, y que tienen consecuencias muy importantes en el ámbito científico. Por ejemplo, en el año 1999 la prestigiosa revista *Nature* publicó un estudio en el que se había detectado una correlación entre el hecho de dormir con una luz encendida durante los dos primeros años de vida y el desarrollo posterior de la miopía. La miopía se desarrolla en la niñez y puede que sea el producto de una combinación de factores genéticos y ambientales. Algunos experimentos con animales habían demostrado que los ciclos de luz podían tener influencia en el desarrollo de la enfermedad, de manera que los autores diseñaron un estudio que permitiera relacionar los ambientes nocturnos de los

niños y el desarrollo juvenil de la miopía. Para obtener los datos, los padres de niños que se citaban en una consulta de oftalmología pediátrica contestaron un cuestionario sobre los hábitos nocturnos de sus hijos hasta los dos años. La media de edad de los niños de los que hablaba el estudio era de ocho años, algunos miopes y otros no. En los resultados se encontró una alta proporción de niños miopes que habían dormido de pequeños con la luz encendida. Los autores fueron muy cautos en la redacción de sus conclusiones, exponiendo que, aunque estos resultados no establecían una relación causal, la fuerza estadística de la asociación entre esas dos variables sugería que no tener periodos de oscuridad durante los primeros años de vida podía ser un factor que precipitara la aparición de la miopía. Las réplicas a este estudio no se hicieron esperar y enfatizaron los factores más importantes a tener en cuenta a la hora de intuir relaciones causales: por un lado, la fiabilidad que ofrece la selección de la muestra y el procedimiento de recogida de datos; por otro, la existencia de terceras variables, no consideradas, y que pueden ser las causantes de la relación. Por ejemplo, que los padres tuvieran que recordar los hábitos de los niños cuando tenían dos años, la edad de estos niños, o el hecho de que hubiera una determinada proporción de miopes, en la muestra, pudo influir en la observación de la relación. Una réplica del estudio con menos proporción de miopes y una media de edad mayor arrojó valores diferentes y no encontró a niños miopes entre los que durmieron con la luz encendida. Otra réplica encontró una asociación importante entre el hecho de ser padres miopes y dejar las luces encendidas. Por tanto, una posible explicación de la asociación que se había encontrado en el primer estudio es que los niños que desarrollaron miopía juvenil fueran hijos de padres miopes y la aparición de la enfermedad se debiera a una causa genética y no ambiental.

Detectar valores "raros" en los datos

En el estudio estadístico de un problema intervienen muchas decisiones de las que dependen que las conclusiones sean correctas. Por ejemplo, existen cálculos que evalúan cómo se alejan los datos de la media para detectar valores "raros" en los datos. Para que estos cálculos sean efectivos es necesario que se tenga la certeza de que si repitiéramos el estudio con muchas otras muestras de datos, suficientemente variados, se obtendrían resultados muy similares. Sin embargo, puede ocurrir que solo sea posible tomar unas pocas muestras y, en consecuencia, que la decisión sobre si un dato es raro necesite ser mucho más cuidadosa. Este tipo de situaciones ponen de manifiesto que si se cometen errores o se dejan de tomar en consideración factores importantes en las primeras fases de un estudio estadístico, es fácil que los datos se interpreten de manera incorrecta o incompleta. Además, a menudo no disponemos de los datos originales que condujeron a las conclusiones de un estudio, o su validación es extremadamente tediosa. Todo ello conduce al diseño de un tipo de tareas escolares que es el análisis de pifias y engaños: cambios de escala en los gráficos, manipulaciones subjetivas de los datos, elecciones erróneas de la muestra, etc.

Aunque no cabe duda de que los niños y las niñas deberían adquirir una idea equilibrada sobre los usos y abusos de la estadística, opino que los malos ejemplos, o los estudios con errores o pobres de contenido deberían utilizarse en muy contadas ocasiones, y siempre con un objetivo de aprendizaje muy concreto, más cercano a estimular la discusión sobre los factores que pueden influir en lo que se analiza que a la desconfianza. Una opción alternativa es adaptar buenos ejemplos de estudios estadísticos en los que esté claramente definido cuál es el problema que se pretende analizar, o excelentes representaciones gráficas que permitan extraer lo realmente importante de los datos, y que estimulen a niños y niñas a hacerse preguntas sobre ellos. Las tareas escolares que tienen como objetivo enseñar a construir gráficos extremadamente sencillos suelen venir acompañadas de preguntas cerradas y de un nivel interpretativo muy bajo. Por ejemplo, un ejercicio típico es pedir que recojan los datos de altura o peso de los niños de la clase, los ordenen y hagan un histograma para después proponer preguntas del tipo: ¿Cuántos niños son más altos que Juan? ¿Quién es el más alto de la clase? ¿Cuál es la media de altura?, etc., cuyo objetivo es leer el gráfico literalmente sin avanzar en su interpretación. Si el objetivo de una educación estadística de calidad es lograr que los niños hagan buenas preguntas que permitan comprender una situación o predecir, entonces parece lógico que se enfatice mucho más el contacto con estudios de datos ricos, buenas representaciones y discusiones de calidad, porque el esfuerzo por buscar preguntas interesantes, cuando los datos son pobres, suele ser inútil. Los niños no podrán valorar la importancia de encontrar buenas preguntas, generar buenos instrumentos de información gráfica que recojan múltiples variables o conjeturar a partir de los datos si solo conocen los ejemplos, observaciones o tipos de gráficos sencillos que ellos pueden recoger y construir.

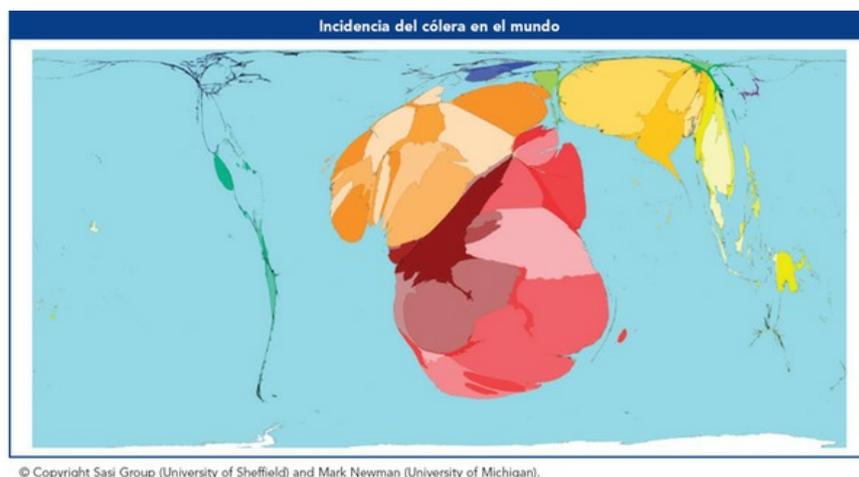
El médico John Snow y la epidemia del cólera

Un ejemplo bien conocido de razonamiento estadístico, que encuentro particularmente ilustrativo, es el que describe el *descubrimiento* de la causa de transmisión del cólera, una enfermedad que provocó la muerte de muchas personas en el siglo XIX. John Snow, un médico inglés, investigó una epidemia de cólera en Londres y registró todas las manifestaciones de una enfermedad que aparecía de repente y atacaba a muchísima gente. Los médicos de la época barajaban las hipótesis de que la enfermedad se transmitiera por el aire pero, para su desconcierto, algunas personas enfermaban y otras no, a pesar de vivir muy cerca. El mapa reproducido (véase el cuadro 1) es una réplica del original de 1854 en el que John Snow marcó las casas y el número de personas que habían enfermado.



Podemos pedir a los niños que examinen bien el mapa y traten de buscar fotografías o textos que describan la vida londinense de la época (¿Había agua corriente en las casas? ¿Y servicios? ¿Cómo se desplazaban los habitantes por la ciudad?). Esto los ayudará a poder conectar razonamientos, proponer hipótesis y responder a preguntas del tipo: ¿Crees que es una buena hipótesis afirmar que el cólera podría transmitirse por el aire? ¿Por agua contaminada? ¿Por los excrementos de caballos que quedaban en las calles? El doctor Snow también marcó en el mapa todas las fuentes públicas, ¿qué crees que pensaba para hacerlo? Además habló con residentes de la zona, enfermos y no enfermos, ¿qué preguntas crees que les hizo? En Internet existen muchos recursos y actividades, para analizar el caso de la epidemia de cólera londinense, que vale la pena rastrear y adaptar al contexto de la Educación Primaria.

Otro ejemplo excelente de representación de datos con un gran potencial para su uso en las aulas es el del Proyecto Worldmapper (<http://worldmapper.org>). En el mapa que reproducimos sobre la incidencia del cólera en el mundo (véase el cuadro 2), mucho más moderno que el de John Snow, la medida de los territorios coloreados muestra la proporción de datos de cólera en la actualidad. Podemos dejar que los niños observen el mapa durante un tiempo y se hagan preguntas para explicar la incidencia desigual de la enfermedad en el mundo. Pueden buscar en la misma página otros mapas que den respuesta a sus preguntas, o conjeturar cómo serán, por ejemplo, los mapas que mostrarán los datos mundiales de acceso al agua corriente o a la atención médica, antes de encontrarlos.



Una aproximación diferente a la estadística

Aún hay muchos textos escolares que hablan de "tipos de gráficos" y mencionan los diagramas de barras, de sectores, histogramas y pictogramas. Este reduccionismo limita el desarrollo de la creatividad y el estudio complejo de muchas variables (a menudo de carácter social y muy difíciles de medir) que a la hora de representar gráficamente la información prevalecen sobre las técnicas aritméticas de la estadística descriptiva. Sin embargo, las técnicas de ordenación y operatoria con los datos siguen siendo el foco de atención de los contenidos escolares de Matemáticas, cuando se estudia la estadística. Este hecho es, en realidad, un arma de doble filo: por una parte, si el argumento para trabajar contenidos de estadística en clase de Matemáticas es que las técnicas aritméticas encajan bien en esta disciplina, se refuerza una visión muy extendida de la matemática escolar que la identifica con tareas mecánicas en las que se opera con números sin saber muy bien por qué, mientras que la conjetura, la resolución de problemas y la demostración, que deberían ser el eje central de la matemática, se dejan de lado. Por otra parte, los estudios estadísticos tienen mucho que ver con problemas de índole social, que son complejos, extremadamente dependientes del contexto y a menudo impredecibles, y que requieren una aproximación muy diferente a la que caracteriza la resolución de un problema en una clase de Matemáticas. Es importante que los niños y niñas extraigan lo máximo posible de un gráfico, y que la mayoría de las preguntas descriptivas tengan una pregunta inferencial detrás, que puedan responder por ellos mismos.

No quisiera dar a entender que no considero importante que los niños y niñas *salgan* a recoger sus propios datos, los organicen en tablas y aprendan a hacer un diagrama de barras, pero creo que bastaría con la observación de una población pequeña y unos pocos ejercicios genuinos, en su contexto, fantástico o cotidiano, que ilustraran los aspectos esenciales en un estudio estadístico. En cambio, dedicaría mucho más tiempo a conocer buenos ejemplos de investigaciones que utilizan la estadística, adaptados a su nivel. Igual que es importante que los niños lean y aprendan a valorar la buena literatura para que después puedan disfrutar del placer de la lectura y acceder a la sociedad del conocimiento, es importante que tengan referencias estadísticas de calidad. Aprender la ingeniería de la estadística es una tarea muy lenta y compleja, que probablemente sea más sencilla si se tienen buenos referentes de adónde se quiere llegar.

Por todo lo expuesto hasta aquí, creo que el estudio de la estadística ha de permitir que los niños se hagan una idea del mundo social, científico, político y económico en el que viven, a través de estudios rigurosos y de calidad debidamente adaptados a su nivel. Para este fin, creo que centrar el estudio de la estadística en la matemática tiene más inconvenientes que ventajas. Un trabajo pendiente de los profesionales de la educación es generar documentos escolares, a partir de artículos de investigación y de divulgación, que destilen los datos y los hagan comprensibles para los niños y niñas, relacionándolos con lo que ellos ya saben.

Para saber más

Autoría compartida (1999): "Myopia and ambient lighting at night", en *Nature*, nº 339, p. 113.

Autoría compartida (2000): "Myopia and ambient night-time lighting" (Brief Communication), en *Nature*, nº 404, p. 144.

Hayes, Brian (2004): "Estadística de los conflictos bélicos", en *Investigación y Ciencia*, nº 332, pp. 9-14.

Tanur, Judith M. (1983): *La estadística, una guía a lo desconocido*. Madrid: Alianza.

Tufte, Edward R. (2001): *The visual display of quantitative information*. Cheshire, Connecticut: Graphic Press.