

EL SENTIDO DE LA DISPERSIÓN Y SU DESARROLLO EN EL CURRÍCULO

Ignacio González-Ruiz, Carmen Batanero, M. Mar López-Martín, J. Miguel Contreras

Universidad de Cantabria (España), Universidad de Granada (España)

ignaciogonzalezruiz@hotmail.com, batanero@ugr.es, mariadelmarlopez@ugr.es, jmcontreras@ugr.es

Palabras clave: Dispersión, sentido estadístico, currículo.

Key words: Spread; statistical sense, curriculum.

RESUMEN: La dispersión es un concepto básico en estadística, pues cuantifica la variabilidad en las distribuciones de datos y las distribuciones de probabilidad, estando ligada tanto al análisis exploratorio de datos, como a la probabilidad. Es también esencial en inferencia, pues permite valorar la precisión de las estimaciones o del riesgo asumido en los procesos de decisión. La finalidad de este trabajo es analizar la forma en que este concepto se contempla en las directrices curriculares españolas y establecer comparaciones con las de otros países. Todo ello para colaborar en una mejor adquisición del sentido de la dispersión en los estudiantes..

ABSTRACT: Spread is a basic concept in statistics, as it quantifies the variability in the data distributions and probability distributions. Consequently it is linked to both exploratory data analysis and probability, being also essential in statistical inference, since it allows valuing the estimates accuracy and the risk assumed in decision making processes. The aim of this work is to analyze the way in which the concept is taken into account in the Spanish curricular guidelines, as well as compare with other countries. All of this contributes to collaborate in the better acquisition of the sense of spread in the students.

■ INTRODUCCIÓN

Las medidas de dispersión son esenciales en una distribución estadística, complementando a las de posición central, al caracterizar la variabilidad de los datos respecto a las mismas. Wild y Pfannkuch (1999) incluyen la percepción de la variación en los datos (y por tanto de la dispersión) como un componente básico de su modelo de pensamiento estadístico. También Moore (1990) incluye tres aspectos de la misma relacionados con el razonamiento estadístico: a) la percepción de su omnipresencia en fenómenos de las ciencias y la vida diaria; a) la competencias para su explicación, identificando las fuentes y causas de la variación; y c) la habilidad para su cuantificación, lo que indirectamente implica la comprensión y cálculo de las medidas de dispersión. A pesar de la importancia que hemos señalado, la investigación sobre la comprensión de la dispersión y de sus medidas es escasa (Estepa y del Pino, 2013; Sánchez, Silva y Coutinho, 2011).

Para contribuir a rellenar este hueco, en este trabajo analizamos la forma en que el tema se considera en el currículo español, en las diversas etapas educativas. El estudio se realiza comparando las directrices curriculares españolas, con el fin de prever su desarrollo en el aprendizaje del estudiante y proporcionar recomendaciones que contribuyan a aumentar el sentido de la dispersión entre ellos.

■ COMPONENTES DEL SENTIDO ESTADÍSTICO

La comprensión de la dispersión es parte del sentido estadístico. Por ello, para comenzar nuestro trabajo, revisamos los componentes del sentido estadístico, descrito por Batanero (2013) como la unión de la cultura y el razonamiento estadístico, incidiendo en particular en la idea de dispersión.

En los últimos años se ha venido forjando el término cultura estadística para reconocer el papel del conocimiento estadístico en la formación elemental que requieren todos los ciudadanos para desenvolverse en su día a día (Murray y Gal, 2002). Pese a que en la literatura sobre este tema cohabitan distintos enfoques vinculados a la cultura estadística (Gal, 2002), existe un consenso generalizado que entiende necesario que una cultura estadística básica requiere de la comprensión y del dominio adecuado de las ideas estadísticas fundamentales. Dichas ideas contribuyen al desarrollo de la estadística como ciencia, están involucradas en la resolución de problemas estadísticos y pueden ser enseñadas con diversos niveles de formalización (siendo, por tanto asequibles en varios niveles educativos). Cada una de estas ideas se diferencia de otros objetos matemáticos próximos en características específicas que deben ser resaltadas en la enseñanza.

Un segundo componente del sentido estadístico, sería la habilidad de cálculo de estadísticos elementales, que no supone hoy día problema alguno, pues es facilitado por la tecnología, que también contribuye a la exploración de las ideas estadísticas fundamentales, por medio de la simulación. La mayor dificultad de aprendizaje es el tercer componente (precisamente el más específico), que consiste en el razonamiento a partir de los datos, para realizar inferencias de muestras a poblaciones y/ o tomar decisiones acertadas en situaciones inciertas.

Seguidamente ahondamos estos aspectos para el caso de la dispersión.

■ EL CASO DE LA DISPERSIÓN Y SUS MEDIDAS

Los distintos modelos que sientan las bases de la cultura estadística en la literatura, defienden necesidad de un conocimiento de las ideas estadísticas fundamentales. Para identificar cuáles de ellas organizan la noción de dispersión adaptamos la propuesta de Burrill y Biehler (2011), fundamentada en el estudio detallado de diversos marcos teóricos educativos y los currículos de estadística de distintos países. En relación a la dispersión distinguimos las siguientes ideas fundamentales: datos, variabilidad aleatoria, distribución, asociación y correlación, muestreo e inferencia. Asimismo, destacamos la importancia de las mismas en el desarrollo del pensamiento y el razonamiento estadístico.

Datos.

Moore (1991) definió la estadística como la ciencia de los datos y señaló que el objeto de la estadística es el razonamiento a partir de datos empíricos, subrayando la importancia del contexto de donde se han tomado los datos. Mientras que en otras ramas de la matemática con frecuencia los datos y contextos son imaginarios y el interés se centra en los conceptos, el contexto de los datos es esencial en estadística.

Además, los estudiantes no están acostumbrados a trabajar con datos de situaciones reales que frecuentemente requieren de interpretaciones y razonamientos de alto nivel. La aleatoriedad de las situaciones hace que los resultados no sean únicos, presentándose mayor variabilidad en los datos que otras áreas de las matemáticas (Sánchez y Batanero, 2012). Asumir la existencia de variación en los datos contribuye a asimilar su naturaleza y por ende, la de los fenómenos de los que proceden.

Variación

MacGillivray (2004) entiende la estadística como la ciencia de la variación, pues este concepto justifica su estudio, la modela y establece mecanismos para su cálculo; además de encargarse de su análisis y representación. Aunque en otras ramas de matemáticas se usan variables, se supone que los datos se ajustan perfectamente a un modelo y no suele haber estudio de la bondad de ajuste o de los residuos del modelo.

El estudio de la variabilidad es característico de la Estadística, que se interesa tanto por el modelo que describe a unos datos como por los residuos o diferencias entre los datos y el modelo (Engel y Sedlmeier, 2011). Por ello, la Estadística permite buscar explicaciones y causas de la variación para poder hacer predicciones. En consecuencia dos fines importantes de la enseñanza de la Estadística son que los estudiantes perciban la variabilidad y que manejen modelos que permitan controlarla y predecirla (Reading y Shaughnessy, 2004). Una de las formas de abordar el estudio de la variación de los datos es a través del cómputo de la dispersión de los mismos respecto de un valor central representativo.

Distribución

Una característica esencial del análisis estadístico es que trata de describir y predecir propiedades de los agregados de datos y no de cada dato aislado (Bakker y Gravemeijer (2004). La dispersión juega un papel especial en esta labor. Por ello la enseñanza de la Estadística ha de desarrollar la capacidad de leer, analizar, criticar y hacer inferencias a partir de distribuciones de datos (Shaughnessy, 2007).

El razonamiento distribucional implica también conectar los datos (distribución de datos), la población de donde se tomaron (distribución de probabilidad) y las posibles muestras de la misma (distribución muestral). Garfield y Ben-Zvi (2008) señalaron que la comprensión de las ideas de dispersión o variabilidad resulta clave para la comprensión del concepto de distribución, y es esencial para hacer inferencias estadísticas. A este respecto, Wild (2006) defiende que tras la noción de distribución subyacen todas las formas de razonamiento estadísticos acerca de la variación.

Asociación y correlación

Mientras que en una dependencia funcional a cada valor de una variable X (independiente) corresponde un solo valor de otra variable Y (dependiente), en el estudio de la asociación a cada valor de X corresponde una distribución de valores de Y , por lo que este concepto amplía el de dependencia funcional. Este hecho hace que tenga pleno sentido hacer extensivo a sendas variables el estudio de la dispersión.

En el estudio conjunto de dos variables se pueden diferenciar varios componentes de la dispersión: por un lado se puede calcular la dispersión conjunta, y la dispersión de cada una de las variables. Por otro se puede estudiar la dispersión de los datos respecto al modelo de regresión y la dispersión residual. Las medidas de dispersión (varianza respecto a cada variable y covarianza) intervienen además en el cálculo de los coeficientes de correlación y regresión.

Muestreo e inferencia

Relacionar las características de las muestras con las de la población que representan es el principal fin de la estadística y nos sirve para decidir qué datos recoger y para obtener conclusiones con algún grado de probabilidad (de la Fuente y Díaz, 2004). Varios autores sugieren que es posible una comprensión informal de la inferencia, desde la secundaria, que comenzaría por la discriminación entre la posición central y variabilidad en las distribuciones de datos y el uso de estas dos características para decidir cuándo dos distribuciones son iguales o diferentes (Rubin, Hammerman y Konold, 2006).

Además de la comprensión de las ideas anteriores Wild y Pfannkuch (1999) defienden la necesidad de desarrollar el razonamiento estadístico, necesario a la hora de proceder a resolución problemas de esta índole; como los vinculados a la noción de dispersión. En este sentido, subrayan la importancia de la percepción de la variación, como uno de los aspectos básicos del razonamiento estadístico.

■ DESARROLLO DEL SENTIDO DE LA DISPERSIÓN EN EL CURRÍCULO ESPAÑOL

El segundo punto a desarrollar es el análisis de la presencia de la idea de dispersión en el currículo español (MEC, 2006; 2007a; 2007b; MECD, 2014, 2015) y la forma en que este concepto se va ampliando a lo largo de toda la escolarización, con niveles progresivos de amplitud y complejidad.

La dispersión en la Educación Primaria

En el Decreto de Enseñanzas Mínimas de la Educación Primaria (MEC, 2006) se plantea, como uno objetivo fundamental el empleo de técnicas elementales de recogida de datos para obtener información sobre fenómenos y situaciones de su entorno; representarla de forma gráfica y numérica y formarse un juicio sobre la misma.

En este documento se incluye como criterio de evaluación valorar la capacidad de los niños para interpretar gráficos sencillos de situaciones familiares y verificar su habilidad para reconocer gráficamente informaciones cuantificables. Puesto que en toda representación de una variable estadística ha de aparecer cierta variabilidad, interpretamos que, aunque sólo implícitamente, se sugieren trabajar la idea de variabilidad y rango. Contenidos similares se recogen en MECD (2015).

La dispersión en la Educación Secundaria

En la etapa correspondiente a la Educación Secundaria Obligatoria (E.S.O.) se amplían estas ideas y se reformula el problema de medir la dispersión respecto a una medida de posición central.

En el Decreto de Enseñanzas Mínimas de la Educación Secundaria (MEC, 2007a), se arguye la importancia que supone en la toma de decisiones, la comprensión, modificación y producción mensajes en los que la información estadística aparece presentada, en tablas, gráficos y fórmulas, que demandan conocimientos matemáticos para su correcta interpretación. También se destaca la presencia de información estadística en los medios de comunicación y el uso que de ella se hace en las diferentes materias del currículo. Todo ello convierte a la estadística en un saber a destacar en la formación actual de los estudiantes. En este sentido se pide al profesor que acostumbre al alumno a analizar de forma crítica la información, las interpretaciones sesgadas y los abusos que, en ocasiones, caracterizan a la información estadística.

Estos problemas motivan los conceptos de desviación respecto a la media, varianza y desviación típica y recorrido intercuartílico, así como la terminología y simbolización asociada. Cuando se desea que la medida de dispersión sea relativa (a la medida de valor central) se introduce el coeficiente de variación.

Progresivamente, a partir de tercer curso, se trabaja con variables agrupadas, lo que amplía el repertorio de algoritmos y procedimientos; las directrices curriculares recomiendan el uso de la calculadora gráfica o la hoja de cálculo para facilitar estos algoritmos. Para dotar de mayor sentido la idea de dispersión se emplean los gráficos de frecuencias agrupadas, gráficos múltiples o de caja. También se recomienda el uso de las medidas de posición central y dispersión para la comparación de distribuciones.

Respecto a la probabilidad, donde aparecerá la idea de dispersión de una variable aleatoria, se sugiere una aproximación frecuencial. El estudio de fenómenos aleatorios sencillos mediante experimentación y el tratamiento, por medio de tablas y gráficas, de datos estadísticos ayuda a este enfoque y acá de nuevo aparece la dispersión en las variables estadísticas asociadas a estos experimentos.

La dispersión en Bachillerato

La enseñanza del Bachillerato comprende los cursos 1º y 2º. En relación a las Matemáticas, el Decreto 1467/2007 de 2 de noviembre, por el que se establece la estructura de estos estudios y se fijan sus enseñanzas mínimas (MEC, 2007), nos obliga a distinguir entre dos modalidades: 1) Ciencia y Tecnología y 2) Humanidades y Ciencias Sociales. En el Bachillerato, por un lado se introduce la correlación y regresión, donde la dispersión puede interpretarse desde diferentes puntos de vista: En primer lugar, se introducen la correlación y regresión, relacionadas directamente con la dispersión; cuanto mayor sea la dispersión de los puntos respecto a un modelo matemático de ajuste a los datos (línea o recta de regresión), menor será la intensidad de la relación (o de la correlación) entre las variables.

La covarianza y el coeficiente de correlación lineal serán nuevos estadísticos que permiten valorar la intensidad y signo de la relación, pero también son nuevos indicadores de la dispersión conjunta de los datos. Es precisamente el cuadrado del coeficiente de correlación (coeficiente de determinación) el que nos da la proporción de varianza de la variable dependiente Y explicada por el modelo línea de regresión (recta de regresión). Cuanto mayor sea este coeficiente más proporción de la varianza de Y es explicada por el modelo de regresión o, lo que es lo mismo, es mejor la bondad de ajuste. De este modo, se relaciona también la idea de dispersión con la bondad de ajuste de un modelo.

La introducción de algunas variables aleatorias sencillas y sus distribuciones (binomial, normal), en el 1º curso de Bachillerato para los alumnos de Ciencias Sociales y del 2º curso para los alumnos de Ciencias permite considerar un nuevo punto de vista para las medidas de posición central y dispersión: su carácter de parámetro de las distribuciones de probabilidad. La situación-problemática principal que motiva el estudio de estos conceptos será la de determinar modelos generales de distribuciones (familias) que permiten resolver una gama de situaciones probabilísticas y encontrar las expresiones de sus distribuciones de probabilidad. Así, en el caso de la distribución normal un cambio en su desviación típica no sólo produce una mayor menor dispersión de la función de densidad, sino que reduce el apuntamiento de la misma.

En Matemáticas aplicadas a las Ciencias Sociales II, asignatura opcional en el segundo curso del Bachillerato de Humanidades y Ciencias Sociales, se introducen contenidos propios de la inferencia estadística. En este contexto, un problema de interés radica en la estimación de los parámetros de las distribuciones de probabilidad en una población (por ejemplo, la media de una distribución normal) a partir de los datos de una muestra aleatoria tomada de dicha población. Un segundo problema consiste en contrastar hipótesis sobre los valores de dichos parámetros.

Estos problemas llevan a la introducción de muchos conceptos nuevos, como distribución muestral, error de estimación, intervalo de confianza, relacionados con la idea de dispersión. Son características propiedades como la relación entre tamaño de muestra, precisión y confianza o la relación entre la desviación típica de la variable en la población y la de la distribución de la media muestral.

■ CONCLUSIONES

A pesar de la brevedad de nuestro análisis, es posible observar cómo, cada etapa educativa da un papel relevante al concepto de dispersión, que se va enriqueciendo progresivamente, desde Educación Primaria a Bachillerato. En este último, se introduce la estadística bidimensional, variables aleatorias e inferencia que implican nueva terminología, conceptos y procedimientos relacionados con la dispersión. Propuestas muy similares aparecen en los currículos de otros países; por ejemplo en CCSI (2010).

Es por ello necesario que el profesor ayude al alumno a conectar el uso de la dispersión en los diferentes apartados de la estadística (análisis exploratorio de datos, probabilidad e inferencia) para que llegue a alcanzar un sentido adecuado de la dispersión que le ayude a afrontar situaciones de incertidumbre. Como indicó Moore (1990), la estadística es la ciencia de los datos y estos están caracterizados por la variabilidad. Por tanto la habilidad para percibir, medir y explicar la dispersión

de los datos y de los modelos que utilizamos para describirlos es la clave del razonamiento estadístico.

Esperamos finalmente que este estudio pretendemos orientar el trabajo del profesor en el aula y contribuir a que potencien el sentido de la dispersión en sus estudiantes.

Agradecimientos: Proyecto EDU2013-41141-P (MEC) y grupo FQM126 (Junta de Andalucía).

■ REFERENCIAS BIBLIOGRÁFICAS

Batanero, C. (2013). Sentido estadístico: Componentes y desarrollo. En J. M. Contreras (Ed.), *Actas de las I Jornadas Virtuales en Didáctica de la Estadística, Probabilidad y Combinatoria*, (pp. 55-61). Universidad de Granada, CCSSI (2010).

Common Core State Standards for mathematics. Washington, DC: National Governors Association for Best Practices and the Council of Chief State School Officers.

Bakker, A. y Gravemeijer, K. P. E. (2004). Learning to reason about distribution. En J. Garfield y D. Ben Zvi (Eds.), *The challenge of developing statistical literacy, reasoning and thinking* (pp 147-168). Dordrecht, The Netherlands: Kluwer.

Burrill, G., y Biehler, R. (2011). Fundamental statistical ideas in the school curriculum and in training teachers. En C. Batanero, G. Burrill y C. Reading (Eds.), *Teaching statistics in school mathematics. Challenges for teaching and teacher education - A joint ICMI/IASE study* (pp. 57-69). Dordrecht: Springer.

De la Fuente, I., y Díaz. C. (2004). controversias en el uso de la inferencia en investigación experimental. *Metodología de las Ciencias del Comportamiento*, número especial, 161-167.

Engel, J. y Sedlmeier (2011). Correlation and regression in the training of teachers. En C. Batanero, G. Burrill, C. Reading y A. Rossman (Eds.). *Teaching statistics in school mathematics-challenges for teaching and teacher education. A joint ICMI/IASE study* (pp. 247-258). New York: Springer.

Estepa, A. y del Pino, J. (2013). Elementos de interés en la investigación didáctica y enseñanza de la dispersión estadística. *Números 83*, 43-63.

Gal, I (2002). Adult's statistical literacy. Meanings, components, responsibilities. *International Statistical Review*, 70(1), 1-25.

Garfield, J., y Ben-Zvi, D. (2008). *Developing students' statistical reasoning. Connecting research and teaching practices*. New York: Springer.

MacGillivray, H. (2004). Coherent and purposeful development in statistics across the education spectrum. En G. Burrill, y M. Camden (Eds.), *Curricular Development in Statistics Education: International Association for Statistical Education 2004. Roundtable*. (pp. 230-243). Voorburg, The Netherlands: International Statistical Institute.

MEC (2006). *Real Decreto 1513/2006, de 7 de diciembre, por el que se establecen las enseñanzas mínimas de la Educación primaria*. Madrid: Autor.

- MEC (2007a). *Real Decreto 1631/2006, de 29 de diciembre, por el que se establecen las enseñanzas mínimas correspondientes a la Educación Secundaria Obligatoria*. Madrid: Autor.
- MEC (2007b). *Real Decreto 1467/2007, de 2 de noviembre, por el que se establece la estructura del bachillerato y se fijan sus enseñanzas mínimas*. Madrid: Autor.
- MECD (2015). *Real Decreto 1105/2014 de currículo básico de la Educación Secundaria Obligatoria y del Bachillerato*. Madrid: Autor.
- Ministerio de Educación, Cultura y Deporte, MECD (2014). *Real Decreto 126/2014 de currículo básico de la Educación Primaria*. Madrid: Autor.
- Moore, D. S. (1990). Uncertainty. En L. A. Steen (Ed.). *On the shoulders of giants. New approaches to numeracy*. Washinton, D. C.: National Academy Press.
- Murray, S. y Gal, I. (2002). Preparing for diversity in statistics literacy: Institutional and educational implications. En B. Phillips (Ed.), *Proceedings of the Sixth International Conference on Teaching of Statistics*. Ciudad del Cabo: IASE. CD ROM.
- Reading, C. y Shaughnessy, J. M. (2004). Reasoning about variation. En J. Garfield y D. Ben-Zvi (Eds.), *The challenge of developing statistical literacy, reasoning and thinking* (pp. 201-226). Dordrecht, The Netherlands: Kluwer.
- Rubin, A., Hammerman, J. K. y Konold, C. (2006). Exploring informal inference with interactive visualization software. En B. Phillips (Ed.), *Proceedings of the Sixth International Conference on Teaching Statistics*. Cape Town, South Africa: International Association for Statistics Education. Online: www.stat.auckland.ac.nz/~iase/publications.
- Sánchez, E. y Batanero, C. (2011). Manejo de la información. En E. Sánchez (Coord.), *Enseñanza y aprendizaje de las matemáticas, Casos y perspectivas* (pp. 64-92). México, D. F.: Secretaría de Educación Pública.
- Sánchez, E., da Silva, C. B. y Coutinho, C. (2011). Teachers' understanding of variation. En C. Catanero, G. Burrill, & C. Reading (Eds.), *Teaching statistics in school mathematics-challenges for teaching and teacher education* (pp. 211-221). New York: Springer.
- Wild, C. (2006). The concept of distribution. *Statistics Education Research Journal*, 5(2), 10–26.
- Wild, C. J. y Pfannkuch, M. (1999). Statistical thinking in empirical enquiry. *International Statistical Review*, 67(3), pp. 223-263.