

EL LENGUAJE DE LA PROPORCIÓN MUESTRAL EN LIBROS DE TEXTO DE BACHILLERATO

Juan J. Ortiz – Veronica Albanese – Nordin Mohamed
jortiz@ugr.es – vealbanese@ugr.es – nmohamed@ugr.es
Universidad de Granada, España

Resumen

En este trabajo analizamos el lenguaje utilizado para presentar la estimación de la proporción muestral en tres libros de texto españoles de bachillerato publicados el pasado año 2016. De las diferentes perspectivas teóricas para abordar el análisis de libros de texto, hemos optado por el Enfoque Onto-semiótico (EOS), por la importancia que otorga al lenguaje. Los resultados muestran la gran riqueza y diversidad de expresiones verbales, tanto del lenguaje ordinario usadas con sentido específico, como propias de la estadística y probabilidad y otras que se refieren a ejemplos de material que se utiliza en los juegos de azar. Hay predominio de lenguaje formal y lenguaje simbólico complejo y variado. El lenguaje numérico contempla todos los sistemas numéricos y se encuentra también amplio uso de representaciones tabulares y gráficas, algunas de ellas específicas del tema. Algunas diferencias en los libros indican el importante papel del profesor al seleccionar y usar estos libros en la enseñanza.

Introducción

La enseñanza de la inferencia es un tema fundamental de las Matemáticas Aplicadas a las Ciencias Sociales II, por la importancia que tiene en la sociedad actual (Batanero y Borovcnik, 2016). Además, en segundo curso de Bachillerato (17-18 años) de esta modalidad (MECD, 2015, p. 389), en el *Bloque 4. Estadística y probabilidad* se presentan los siguientes contenidos de inferencia:

Población y muestra. Métodos de selección de una muestra. Tamaño y representatividad de una muestra. Estadística paramétrica. Parámetros de una población y estadísticos obtenidos a partir de una muestra. Estimación puntual. Media y desviación típica de la media muestral y de la proporción muestral. Distribución de la media muestral en una población normal. Distribución de la media muestral y de la proporción muestral en el caso de muestras grandes. Estimación por intervalos de confianza. Relación entre confianza, error y tamaño muestral. Intervalo de confianza para la media poblacional de una distribución normal con desviación típica conocida. Intervalo de confianza para la media poblacional de una distribución de modelo desconocido y para la proporción en el caso de muestras grandes.

Para responder al cambio de directrices curriculares, en 2016 se han editado nuevos libros de texto. El libro de texto es uno de los principales recursos educativos, ya que muchas decisiones de los profesores sobre las tareas a realizar están mediadas por los mismos (Stylianides, 2009). Desde el currículo pretendido al implementado en el aula, una fase importante es el currículo escrito y la forma en que lo interpretan los profesores, a través de los libros de texto (Herbel-Eisenmann, 2007).

Una característica importante del libro de texto de matemáticas es el lenguaje, por ser un instrumento necesario en la representación y la actividad de matematización y por reflejar la complejidad conceptual de un tema. En este trabajo pretendemos analizar el lenguaje en el tema de estimación de la proporción muestral en tres libros de texto españoles, de segundo curso de Bachillerato de Ciencias Sociales, publicados según la nueva normativa. La finalidad es comparar con otros estudios previos y las directrices curriculares citadas.

Marco teórico

Un desafío en la enseñanza de las matemáticas es el uso de un lenguaje múltiple, que incluye el lenguaje verbal, los símbolos y expresiones algebraicas, las representaciones gráficas y las tablas (Schelepppegrell, 2007). Entre los diversos modelos disponibles para realizar un análisis de los libros de texto, hemos optado por el Enfoque Onto-semiótico (EOS) (Godino, Batanero y Font, 2007), por el papel predominante que otorga al lenguaje matemático, al que considera mediador de las prácticas personales o institucionales en la resolución de problemas, por su carácter representacional y operativo. En este marco teórico es también fundamental la idea de conflicto semiótico, que puede surgir al interpretar el lenguaje matemático (Godino, Batanero y Font, 2007).

Antecedentes

Son escasas las investigaciones sobre la presentación de la estadística y la probabilidad en los libros de texto, y menor aún la relacionada con el lenguaje. Respecto a la probabilidad, Ortiz, Serrano y Batanero (2001) estudiaron el lenguaje en dos libros de texto de Educación Secundaria, distinguiendo entre el lenguaje del azar y de la probabilidad. Observan mayor riqueza del lenguaje empleado respecto al azar en uno de los textos, con una gama más

variada de adjetivos y expresiones, ejemplos de generadores aleatorios asociados a una concepción frecuencial de la probabilidad y tablas de números aleatorios. El mismo texto presenta un vocabulario más rico respecto a la probabilidad, con gradaciones cualitativas, presentando las concepciones subjetivas y frecuencial y conectando con el estudio de la estadística.

Gómez, Ortiz, Batanero y Contreras (2013) analizaron el lenguaje utilizado en el tema de probabilidad en dos series de libros de texto españoles de Educación Primaria. Sus resultados muestran la gran riqueza y diversidad de expresiones verbales y predominio de lenguaje coloquial frente al formal; el lenguaje se asocia a diversos significados de la probabilidad (intuitivo, clásico, frecuencial y formal). El trabajo anterior es completado por Ortiz, Albanese y Serrano (2016) quienes analizaron el lenguaje de la probabilidad en tres libros de texto españoles de Educación Secundaria publicados en 2015. Los resultados muestran la gran riqueza y diversidad de expresiones verbales y predominio de lenguaje coloquial frente al formal; el lenguaje se asocia a diversos significados de la probabilidad (intuitivo, clásico, frecuencial y formal). El lenguaje numérico se desarrolla de acuerdo a la introducción de diferentes sistemas numéricos en la enseñanza y se encuentra también amplio uso de representaciones tabulares y gráficas.

García y García (2009) realizaron un estudio pormenorizado de los términos específicos relacionados con la inferencia estadística. Concluyen que el contexto de trabajo es determinante en el significado de los términos y que, en ocasiones, la definición de estos términos que aparece en los libros de texto no corresponde a la propia del contexto matemático, sino más bien a la del contexto cotidiano, lo que según el autor no es adecuado ya que el libro de texto, como herramienta de trabajo en el aula, debe presentar al estudiante los conceptos matemáticos de forma correcta. Nuestro trabajo trata de completar los anteriores, analizando tres libros de texto españoles de segundo curso de Bachillerato, en la modalidad de Ciencias Sociales, publicados en 2016.

Metodología

Se analizaron tres libros de texto, publicados en 2016, que se eligieron por ser editoriales de gran prestigio a nivel nacional. Se trata de una muestra intencional, puesto que el estudio es exploratorio, sin pretensiones de extender las conclusiones. Se incluyen como anexo y se

denotan con un código en el trabajo. En estos libros se ha realizado un análisis de contenido del capítulo dedicado a inferencia estadística y estimación de la proporción muestral, estudiando las variables determinadas en Gómez et al. (2013), que permiten lograr el objetivo de este estudio: a) expresiones verbales, según tipología; b) expresiones numéricas; c) símbolos; d) representaciones tabulares y gráficas.

Las categorías de cada una de estas variables se determinan mediante sucesivas revisiones de los textos de un modo cíclico e inductivo. A través de la comparación del contenido de estos textos, se establece la presencia o ausencia de cada una de las categorías en los libros de la muestra. Por último, se seleccionan ejemplos en los textos que ilustren las diferentes categorías y se elaboran unas tablas cuya lectura facilite la obtención de conclusiones sobre el uso del lenguaje en los libros analizados.

Resultados y discusión

Expresiones verbales

En primer lugar se analizaron las expresiones verbales. Siguiendo a Shuard y Rothery (1984), hemos tenido en cuenta, las palabras del lenguaje cotidiano, que se usan en el texto con sentido diferente al cotidiano, lo que puede crear problemas de ambigüedad al aplicarlas con un sentido diferente al conocido anteriormente por el estudiante (Barwell, 2005). Dentro de las específicas, siguiendo a Gómez et al. (2013), hemos diferenciado las que se refieren a juegos de azar y hemos separado las específicas de estadística y de probabilidad (ver Tabla 1).

Tabla 2. Expresiones distintas y frecuencia en los libros de texto según categoría

Tipo	[T1]	[T2]	[T3]
Expresiones cotidianas	14	12	8
Específicas probabilidad	21	14	12
Específicas estadística	42	27	35
Juegos de azar	3	4	0

En ella observamos una gran variedad de términos: encontramos palabras que se usan para indicar resumidamente un procedimiento o hacen alusión a conceptos o propiedades de estadística o probabilidad. El mayor número de expresiones diferentes son las específicas

de estadística y probabilidad, siendo muy escasas las referidas a juegos de azar, al contrario que en el estudio de Gómez et al. (2013) con textos de primaria, y en Ortiz et al. (2016) con textos de secundaria. Observamos un aumento en la formalización y variedad del lenguaje en Bachillerato. No obstante, todavía hay una gran variedad de expresiones del lenguaje ordinario usadas con sentido específico, lo que puede ocasionar conflictos semióticos (Godino et al., 2007), debido a problemas de ambigüedad (Barwell, 2005).

Entre las específicas de probabilidad, las más utilizadas están relacionadas con el cálculo de probabilidades, variable aleatoria, las distribuciones normal y binomial y las distribuciones muestrales. La distribución de las proporciones muestrales se trata en los tres textos. El texto [T3], es el único que destaca la idea de incertidumbre como una característica de los estudios sobre poblaciones. Respecto al estudio de Ortiz et al. (2001), con textos de secundaria dirigido a alumnos de 14 años, el lenguaje ha variado bastante ya que aparecen conceptos más complejos.

Los términos específicos de estadística son muy variados, destacando los conceptos de población y muestra, estimación de la proporción muestral, inferencia estadística e intervalos de confianza, que aparecen en los tres textos. En ningún texto se da una definición explícita del concepto de inferencia que es fundamental, lo que hacen, los textos [T1] y [T3], es utilizar la palabra inferir como sinónimo de deducir que, según García y García (2009), está más relacionado con el contexto cotidiano, ya que en matemáticas son dos términos opuestos con significados distintos, lo que puede generar obstáculos en el aprendizaje del alumnado.

Al comparar el contenido de los textos con las indicaciones del currículo, se observa que por ejemplo, el texto [T1] trata los test de hipótesis (T1, p.316), concepto no contemplado en los documentos curriculares. Otros textos omiten contenidos que sí están incluidos en el currículo, como el análisis de los elementos de una ficha técnica en un estudio estadístico, siendo el texto [T3] el único que lo trata. En el texto [T2] no se hace ninguna referencia a expresiones relacionadas con las tecnologías y la simulación tal y como se recomienda en dichos documentos.

Lenguaje numérico

En los tres libros de texto encontramos los mismos tipos de números: Números enteros que suelen expresar el tamaño de la muestra o el valor de los parámetros de las distribuciones; decimales y fracciones que a veces aparecen en la misma expresión: " $p = \frac{30}{150} = 0.2$ " (T3, p.303). Los porcentajes se utilizan para expresar probabilidades o para indicar el nivel de confianza. Los irracionales aparecen en el cálculo del error máximo admisible en la estimación de la proporción muestral, en el cálculo del tamaño mínimo de una muestra para que se cumplan las condiciones fijadas, o en la fórmula de la desviación típica de la distribución de la proporción muestral: " $\sigma = \sqrt{\frac{pq}{n}}$ " (T2, p.290), al contrario que en Ortiz et al. (2016) que no se presentaban.

Lenguaje simbólico

Se ha encontrado una gran variedad de lenguaje simbólico en todos los textos analizados. En todos ellos aparecen las operaciones aritméticas, igualdades y desigualdades. La notación conjuntista se presenta en todos los textos excepto en [T3]: el símbolo $\{ \}$, para presentar los elementos de un conjunto "*Si $\{x_1, x_2, \dots, x_n\}$ es una muestra*" (T2, p. 290) o el símbolo de pertenencia \in , para expresar la probabilidad de que un parámetro pertenezca a un intervalo de confianza, con nivel de confianza $(1-\alpha)$:

$$P \left[p \in \left(pr - z_{\alpha/2} \cdot \sqrt{\frac{pq}{n}}, pr + z_{\alpha/2} \cdot \sqrt{\frac{pq}{n}} \right) \right] = 1 - \alpha \quad (\text{T1, p. 324})$$

El símbolo \sim para indicar que una distribución binomial se aproxima a una normal: " $B(n, p) \sim N(np, \sqrt{npq})$ " (T1, p.312), o el símbolo \equiv para indicar que la distribución de la proporción muestral \hat{P} sigue una distribución normal: " $\hat{P} \equiv N\left(p, \sqrt{\frac{pq}{n}}\right)$ " (T2, p. 290). Los símbolos de implicación suelen indicar cálculos encadenados. El texto [T3] es el único que presenta una sumatoria (T3, p.303). Es muy amplia la presencia de símbolos literales en todos los textos.

En los tres textos queda reflejada la gran riqueza y complejidad del lenguaje simbólico, como en el trabajo de Ortiz et al. (2066), pero con mayor complejidad, indicador del alto grado de formalización que se pretende alcancen los estudiantes del Bachillerato.

Lenguaje tabular

El empleo de lenguaje tabular es muy escaso sino ausente en este tema de proporción muestral. El único texto que presenta tablas es el [T1]: una tabla resumen de niveles de confianza e intervalos de confianza (Figura 1.a), y otra tabla con datos de proporción en la población y tamaño de la muestra (Figura 1.b), ambas en la sección de ejercicios. En los otros dos textos hay algunas tablas, pero ninguna específica de proporción muestral.

RESUMEN								
	NIVEL DE CONFIANZA	INTERVALO DE CONFIANZA						
a)	90 %	(0,54; 0,70)						
b)	95 %	(0,525; 0,715)						
c)	99 %	(0,495; 0,745)						

a. Tablas de resultados (T1, p.317)

	a)	b)	c)	d)	e)	f)
PROPORCIÓN, p , EN LA POBLACIÓN	0,5	0,6	0,8	0,1	0,05	0,15
TAMAÑO, n , DE LA MUESTRA	10	20	30	50	100	100

b. Tabla de datos (T1, p.320)

Figura 1. Distintos tipos de tablas encontradas en los textos

Destacar que la presencia de tablas es casi nula, al contrario que en Ortiz et al. (2016), donde sí había una gran variedad de lenguaje tabular.

Lenguaje gráfico

En los tres textos aparecen las gráficas de la normal y de la binomial, aunque en algunos de ellos en un tema previo al de la estimación de la proporción muestral. En particular para la proporción solo hemos encontrado dos gráficas: En el texto [T1] se presenta un ejemplo del intervalo de confianza para la proporción, con un nivel de confianza $(1-\alpha)$ construido con una muestra de tamaño n (Figura 2.a), y en el texto [T3], en la sección de aplicaciones, hay una gráfica donde se muestra el intervalo de confianza para la proporción poblacional de parados en el tercer trimestre de 2015 (Figura 2.b).

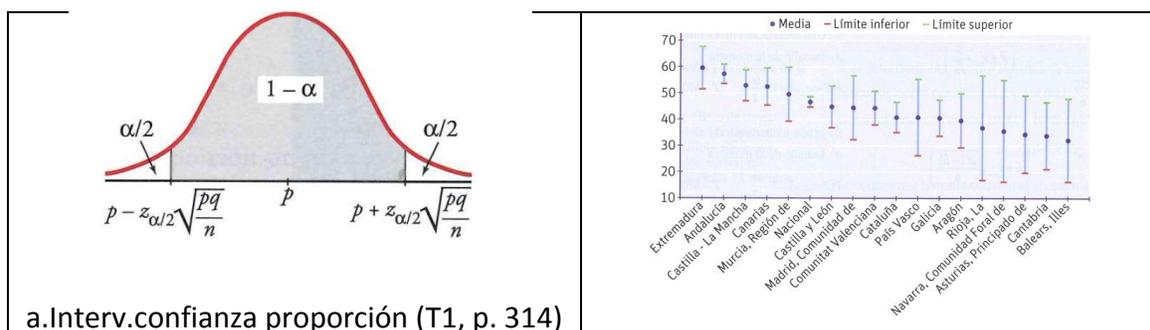


Figura 2. Ejemplos de gráficos en los textos

Se observa también que el empleo de gráficas disminuye con respecto a textos de niveles inferiores (Ortiz et al., 2016), aunque son de mayor complejidad, como es esperable en este nivel educativo.

Conclusiones

En este trabajo se ha mostrado la gran riqueza y diversidad de lenguaje en los textos analizados, que el profesor ha de tener en cuenta para valorar la dificultad que supone para los alumnos. Como indican Ortiz et al. (2001), a esta dificultad se añade el uso de algunas palabras del lenguaje cotidiano, con significado diferente, en el tema de probabilidad. Se encontraron mayor número de expresiones verbales específicas de la estadística con respecto a las de la probabilidad, y muy pocas relativas a los juegos de azar que si aparecen en el estudio de Ortiz et al. (2016). En contra de lo especificado en las orientaciones curriculares hay un texto que no hace referencia al uso de la tecnología.

Para que los estudiantes consoliden un lenguaje matemático más avanzado, los profesores deben cuidar el lenguaje formal que se utiliza en el aula, evitando dar definiciones incompletas o incorrectas que no se corresponden con el significado matemático y que pueden generar obstáculos en el aprendizaje del alumnado (García y García, 2009).

Esperamos con este trabajo contribuir a la mejora de la enseñanza de la estadística y la probabilidad, en el Bachillerato, así como facilitar la labor del profesorado, siendo necesario ampliar el estudio con otros textos.

Agradecimientos: Plan Propio Investigación Universidad de Granada: Programa 20, Proyectos EDU2013-41141-P, EDU2016-74848-P (AEI, FEDER), y Grupo FQMN-126 (Junta de Andalucía).

Referencias bibliográficas

- Barwell, R. (2005). Ambiguity in the mathematics classroom. *Language and Education* 19(2), 118–126.
- Batanero, C., & Borovcnik, M. (2016). *Statistics and probability in high school*. Rotterdam: Sense Publishers.

- Herbel-Eisenmann, B. A. (2007). From intended curriculum to written curriculum: Examining the "voice" of a mathematics textbook. *Journal for Research in Mathematics Education*, 38(4), 344-369.
- García, I. y García, J. A. (2009). Enseñanza de la estadística y lenguaje: un estudio en bachillerato. *Educación Matemática*, 21 (3), 95-126.
- Godino, J. D., Batanero, C. y Font, V. (2007). The onto-semiotic approach to research in mathematics education. *ZDM. The International Journal on Mathematics Education*, 39 (1-2), 127-135.
- Gómez, E., Ortiz, J. J., Batanero, C. y Contreras, J. M. (2013). El lenguaje de probabilidad en los libros de texto de Educación Primaria. *Unión*, 35, 75-91.
- MECD (2015). Real Decreto 1105/2014, de 26 de diciembre, por el que se establece el currículo básico de la Educación Secundaria Obligatoria y del Bachillerato. Madrid: *Boletín Oficial del Estado*.
- Ortiz, J. J., Albanese, V. y Serrano, L. (2016). El lenguaje de la estadística y probabilidad en libros de texto de Educación Secundaria Obligatoria. En J. A. Macías, A. Jiménez, J. L. González, M. T. Sánchez, P. Hernández, C. Fernández, F. J. Ruiz, T. Fernández y A. Berciano (Eds.), *Investigación en Educación Matemática XX* (pp. 397-406). Málaga: SEIEM.
- Ortiz, J. J., Serrano, L., Batanero, C. (2001). El lenguaje probabilístico en los libros de texto. *Suma*, 38, 5-14.
- Schleppegrell, M. (2007). The linguistic challenges of mathematics teaching and learning: A research review. *Reading and Writing Quarterly*, 23, 139-159.
- Shuard, H. & Rothery, A (Eds) (1984). *Children reading mathematics*. London: Murray.
- Stylianides, G. J. (2009). Reasoning-and-Proving in School Mathematics Textbooks. *Mathematical thinking and learning*, 11(4), 258-288.

ANEXO: Textos empleados en el análisis.

- [T1]. Colera, J., Oliveira, M. J., Colera, R. (2016). *Matemáticas aplicadas a las ciencias Sociales II. 2º Bachillerato*. Madrid: Anaya.
- [T2]. Gámez, J., Marín, S., Martín, A., Pérez, C. y Sánchez, D. (2016). *Matemáticas aplicadas a las ciencias Sociales II. 2º Bachillerato*. Madrid: Santillana.
- [T3]. Sanz, L., Alcaide, F., Hernández, J., Moreno, M. y Serrano, E. (2016). *Matemáticas aplicadas a las ciencias Sociales II. 2º Bachillerato*. Madrid: SM.