suma+ 76
Artículos
Julio 2014
pp. 37-45

# Análisis del lenguaje sobre la correlación y regresión en libros de texto de bachillerato

M. Magdalena Gea, Carmen Batanero, Pedro Arteaga, Gustavo R. Cañadas y J. Miguel Contreras

Presentamos un estudio sobre el lenguaje matemático utilizado en el tema de correlación y regresión en ocho libros de texto de Bachillerato. Se analizan los términos verbales, símbolos y expresiones algebraicas, representaciones tabulares y gráficas. Se concluye la complejidad del lenguaje matemático y su diferencia en los textos analizados, observando imprecisiones que podrían inducir conflictos semióticos en los estudiantes.

*Palabras clave:* lenguaje matemático, libros de textos, regresión y correlación, bachillerato.

**Analysis of correlation and regression language in high school textbooks**

We present a study of mathematical language in the topic correlation and regression in eight high school textbooks. We analyze the verbal terms, symbols and algebraic expressions, tabular and graphical representations. We conclude the complexity of mathematical language and the difference in the texts analyzed. We also observed inaccuracies that may induce semiotic conflicts among students.

*Key words:* mathematical language, textbooks, regression and correlation, high school.

La correlación y regresión extienden la dependencia funcional, por lo que su inclusión en el primer curso de Bachillerato de las modalidades de *Ciencias y Tecnología* y *Humanidades y Ciencias Sociales* (MEC, 2007) es claramente justificada.

La enseñanza de este tema no es simple, pues la investigación ha descrito sesgos de razonamiento y dificultades de comprensión, como no apreciar la correlación inversa, tener un sentido determinista o local de la correlación o identificar correlación con causalidad (Estepa y Batanero, 1995; Estepa, 2008; Zieffler y Garfield, 2009). Dichas creencias, en algunos casos, resisten al cambio incluso después de la enseñanza (Batanero, Estepa y Godino, 1997).

También se han observado errores al interpretar los coeficientes de correlación y regresión (Truran 1995, Sánchez Cobo, 1998; Sánchez Cobo, Estepa y Batanero, 2000).

En este trabajo completamos los anteriores analizando el lenguaje con que los libros de texto presentan las nociones de correlación y regresión. En lo que sigue analizamos los fundamentos, métodos y resultados del estudio, finalizando con algunas conclusiones para la enseñanza.



Artículo recibido en *Suma* en octubre de 2012 y aceptado en mayo de 2013

## Fundamentos

### *Marco teórico*

Nuestro análisis pretende observar algunos resultados de la transposición didáctica (Chevallard, 1991), esto es, los cambios del conocimiento matemático cuando es adaptado para la enseñanza. 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, 2007).

El lenguaje del libro de texto, que consta no sólo de vocabulario y símbolos, sino de representaciones complejas según Orton (1990), puede afectar al aprendizaje de las matemáticas, por ejemplo, si los alumnos tienen dificultad en su comprensión. Cordero y Flores (2007) indican que el discurso matemático escolar es determinado con frecuencia por el libro de texto, además de por las creencias de los actores del sistema didáctico y prácticamente regula las acciones de enseñanza y aprendizaje.

El lenguaje matemático es también fundamental en el Enfoque Onto-semiótico (EOS) (Godino, Batanero y Font, 2007), que postula que los objetos matemático emergen de las prácticas de un sujeto (persona o institución) al resolver problemas, y que estas prácticas están mediadas por el lenguaje, que es, a la vez, instrumento representacional y operativo. Los autores también indican la presencia de posibles conflictos semióticos al interpretar el lenguaje matemático, entendiendo por tales «cualquier disparidad o discordancia entre los significados atribuidos a una expresión por dos sujetos (personas o instituciones)» (Godino, Batanero y Font, 2007, p.133).

### Antecedentes

Aunque hay una amplia investigación sobre los libros de texto de matemáticas, esta tradición es mucho menor en el caso de la estadística y probabilidad, donde encontramos algunos ejemplos como los de Ortiz (1999), Ortiz, Batanero y Serrano (2001), Azcárate y Serradó (2006) y Cobo y Batanero (2004).

El primer antecedente relacionado con la correlación y regresión es el de Sánchez Cobo (1998) quien analiza once libros de texto de tercer curso de Bachillerato publicados desde 1987 hasta 1990. Como consecuencia, ofrece una taxonomía de definiciones y un análisis de las demostraciones, desde el punto de vista de la función que realizan y las componentes que la integran. Muestra una tendencia formalista en la presentación del tema, y el uso mayoritario de ejemplos basados en representaciones gráficas, y un fuerte sesgo en los ejemplos presentados hacia la correlación positiva. Más recientemente Lavalle, y cols. (2006) analizan la correlación y regresión en siete libros de texto argentinos de Bachillerato, observando un enfoque mayoritariamente socio-constructivista, con un nivel de profundidad adecuado, donde se plantean más actividades bajo una asociación directa que inversa.

Para complementar los citados trabajos analizaremos el lenguaje matemático utilizado en los textos, que fue estudiado por Ortiz, Batanero y Serrano (2001) para el caso de la probabilidad. En lo que sigue se presentan el método y resultados del estudio.

## Metodología

Se analizaron ocho libros de textos, todos ellos publicados recién implantado el currículo actual de Bachillerato (MEC, 2007) y no reeditados hasta la fecha. Son los más utilizados en la enseñanza pública en Andalucía, y corresponden a las editoriales de mayor tradición y prestigio en esta comunidad (ver Anexo 1).

Se partió de las variables utilizadas por Ortiz, Batanero y Serrano (2001): términos y expresiones verbales; notación simbólica y expresiones algebraicas y representaciones tabulares y gráficas. Para cada una de ellas, por un proceso inductivo y cíclico, se identificaron las categorías de análisis, cuya presencia se analiza en los textos, mostrando, cuando es necesario clarificar, ejemplos y resumiendo lo encontrado en tablas. En las siguientes secciones se presentan los resultados obtenidos.

## Resultados y discusión

### *Términos y expresiones verbales*

Se analizaron los términos y expresiones verbales clasificándolos en dos grupos: por un lado los que debe conocer el estudiante al iniciar el tema, como por ejemplo, intervalo (que se usa en el estudio de las tablas estadísticas de datos agrupados), y por otro los específicos de regresión y correlación, por ejemplo, covarianza. De cada tipo se ha encontrado una amplia variedad, que indica la riqueza conceptual y complejidad del tema y se presentan en la tabla 1.

Rothery (1980) diferencia tres tipos de expresiones en la enseñanza de las matemáticas: (a) Términos matemáticos específicos que, normalmente, no forman parte del lenguaje cotidiano; (b) Palabras usadas en matemáticas y el lenguaje ordinario, aunque no con el mismo significado y (c) Palabras con significados iguales o muy próximos en ambos contextos. Los problemas de aprendizaje estarían ligados sobre todo con las dos primeras categorías, aunque Pimm (1987) considera que la analogía (metáfora) por medio de palabras cotidianas es muy importante para la construcción del significado matemático. Un desafío es que los términos matemáticos tienen mayor precisión que el lenguaje ordinario, pues proporcionan definiciones necesarias y suficientes, mientras que el lenguaje ordinario es simplemente descriptivo (Schleppegrell, 2007).

En el estudio se encontraron términos del lenguaje ordinario utilizadas con diferente sentido para aludir a conceptos u objetos matemáticos (Tabla 2). Aunque la mayoría de estos términos son usados para disminuir la formalidad del enunciado matemático, podrían llevar, de acuerdo a Thompson y Rubenstein (2000), a imprecisiones en el uso de estas nociones por parte del estudiante o incluso a conflictos semióticos.

### *Notación simbólica y expresiones algebraicas*

Un segundo tipo de lenguaje es el simbólico, que se utiliza para referirse a conceptos o propiedades y permite una comunicación comprimida entre individuos, trabajando a un alto nivel de complejidad.

| *No específicos del tema* | *Específicos del tema* |
|---|---|
| Amplitud de intervalo; ángulo de dos rectas; área, baricentro, bisectriz; cambio de variable; coeficiente de variación; coordenada; crecimiento; cuadrante; desviación típica; diagrama de barras, barras adosadas, sectores; dispersión; distancia; distribución; ecuación, punto-pendiente: expresión algebraica; ejes cartesianos; escala; estimación; extrapolación; extremos de intervalo; fiabilidad; frecuencia absoluta, acumulada, relativa, relativa acumulada; función; histograma; individuo; intensidad, interpolación; intervalo de clase; marca de clase; máximo y mínimo; media aritmética; muestra; ordenada; paralelepípedo; parámetro ; pendiente de una recta; población; polígono de frecuencias; porcentaje; prisma; probabilidad; proporcionalidad; punto medio; raíz cuadrada; recta, perpendicular; subíndice; sumatorio; tabla de datos/frecuencias; tangente; tendencia; valor absoluto; variable estadística, cualitativa, cuantitativa continua, discreta; varianza y volumen. | Centro de gravedad; coeficiente de correlación de Pearson; coeficiente de determinación; coeficiente de regresión; correlación, dependencia, sentido, curvilínea, espuria, funcional, lineal, estadística; covarianza; desviación típica marginal; diagrama de barras tridimensional, de dispersión; distribución conjunta, marginal, condicional; error de estimación; frecuencia conjunta absoluta, relativa, condicionada, marginal; histograma tridimensional; incorrelada; independencia; media marginal; método de mínimos cuadrados; nube de puntos; pictograma tridimensional; recta de regresión; de mínimos cuadrados; de Tukey; regresión, lineal, exponencial, logarítmica, cuadrática, parabólica, potencial; tabla de doble entrada, de frecuencias marginal, bidimensional; valor esperado, predicción; valor/dato observado ; variable dependiente/independiente; variable estadística bidimensional; variación conjunta o varianza conjunta y varianza marginal. |

Tabla 1. Términos en los libros de texto

Al igual que Ortiz (1999), hemos encontrado notación funcional, subíndices y superíndices, que con frecuencia son variables (tabla 3).

Llamamos también la atención al uso de letras griegas, que en estadística representan variables aleatorias (estudio de poblaciones o inferencias sobre ellas). En el primer curso de Bachillerato, el estudio de la estadística es descriptivo y no hay una intención inferencial, por lo que los símbolos debieran utilizar letras latinas. Este punto no es intranscendente, pues en el segundo curso del Bachillerato de Ciencias Sociales, los mismos alumnos estudiarán inferencia y se encontrarán con una doble notación (letras griegas para referirse a las características de las poblaciones y latinas para nombrar las mismas características en las muestras), lo que les puede llevar a confusión.

| *Expresión del lenguaje habitual* | *Sentido matemático* |
|---|---|
| Estatura normalita ([T1], p.225) | Estatura media |
| Según lo apretados que estén los puntos ([T1], p. 227); los puntos de la nube están completamente en desorden. ([T2], p.222); los puntos del diagrama están esparcidos al azar ([T3], p.252); la nube de puntos es estrecha/ancha ([T5], p.247) | Dispersión |
| A ojo ([T1], p. 230, p. 232, p. 238) | Aproximación o ajuste |
| Rectas que «se acoplan bien» o «se amoldan» a la nube de puntos ([T1], p. 230, p. 231); la nube de puntos se condensa en torno a ([T3], p. 252) | Ajuste lineal a la nube de puntos |
| Hinchar los puntos proporcionalmente a su frecuencia ([T1], p. 233) | Representar circunferencias con diámetro proporcional a la frecuencia |
| Cómo se apartan a la vez las dos coordenadas de un dato respecto de la media ([T3], p. 251) | Covariación |
| Se puede apostar, suponer su estatura, con una certeza probable ([T6], p.250) | Se puede estimar su estatura |
| Una nube de puntos alargada indica correlación lineal. La estrechez de la nube expresa que la correlación es fuerte ([t6], p. 252) | La nube de puntos informa del tipo, intensidad y sentido de la correlación. |
| Siempre que no se exagere en la extrapolación de resultado ([T6], p. 264) | Siempre que la estimación se realice en valores próximos a la medias |

Tabla 2. Ejemplos de expresiones de lenguaje habitual utilizadas con sentido matemático en los textos analizados

| *Notación* | *Concepto representado* | *T1* | *T2* | *T3* | *T4* | *T5* | *T6* | *T7* | *T8* |
|---|---|---|---|---|---|---|---|---|---|
| $\sum_{i=1}^{n}$ $\sum_{i=1}^{n}$ $\sum_{j=1}^{m}$ | Sumatorio con subíndices | | x | x | | x | | | x |
| $F_i$; $H_i$; $h_i$; | Frecuencia absoluta; relativa acumulada y relativa | | | | | x | | | |
| $x_{max}$ $x_{min}$ | Valor máximo y mínimo de la variable $X$ | | x | | | | | | |
| $n_{ij}$ o $f_{ij}$; $h_{ij}$ | Frecuencia absoluta, relativa; dato bidimensional | | x | x | | x | | | x |
| $P(x_0, y_0)$ | Punto $P$ en el diagrama cartesiano | | | x | | | | | |
| $(x_i, y_i)$ | Valor de la variable bidimensional | x | | x | x | x | x | x | |
| $\overline{x}(\overline{x}, \overline{y})$ | Media, Centro de gravedad | x | x | x | x | x | x | x | x |
| $\sigma_x$ $S_x$ $\sigma_x^2$ $S_x^2$ | Desviación típica y varianza de una variable $X$ | x | x | x | x | x | x | x | x |
| $CV_x$ | Coeficiente de variación de la variable $X$ | | | | | | | x | |
| $\sigma_{xy}$ $S_{xy}$ $r_{xy}$ | Covarianza y Coeficiente de correlación lineal de las variables $X$ e $Y$ | x | x | x | x | x | x | x | x |
| $r_{xy}^2$ $R^2$ | Coeficiente de determinación | | | | | | x | | |
| $\lvert\cdot\rvert$ | Valor absoluto | x | | x | x | | x | | |
| $m_{yx}$ | Pendiente de la recta de regresión de $Y$ sobre $X$ | x | | | x | | | | |
| $d_i$ $d'_i$ | Distancia entre ordenada (abcisa) de un punto y una recta o el error cometido por el modelo | x | | x | x | x | | | |
| $\hat{x}_0(y_0)$;$\hat{x}_{y_0}$;$\hat{x}$ | Valor estimado de $x$ | x | x | x | | | x | | |
| ≈ | Aproximadamente igual a… | | x | x | | x | x | | |
| % | Porcentaje | | | | | x | x | x | x |
| ln $y$ o $e^x$ | Función logarítmica o exponencial | | | x | | | | | |

Tabla 3. Ejemplos de notación simbólica en los textos analizados

Además se incluyen numerosas expresiones algebraicas como por ejemplo

$$\frac{\sum f_i(x_i-\bar{x})\cdot(y_i-\bar{y})}{N}=\frac{\sum x_i y_i f_i}{N}-\bar{x}\cdot\bar{y}$$

donde los mismos símbolos se usan para representar variables o incógnitas.

### *Representación tabular y gráfica*

Las tablas estadísticas ofrecen una estructuración particular del espacio, presentando no sólo números, sino las diversas relaciones que entre ellos se pueden realizar (Ortiz, 1999). Todos los libros analizados reconocen su importancia, aunque su tratamiento varía, siendo la más utilizada el listado de datos en dos filas/columnas, denominada en algunos manuales tabla de frecuencias bidimensional simple ([T3], p. 248). Es común, avanzado el tema, añadir columnas por ejemplo, para el cálculo de la covarianza.

Aunque la tabla de doble entrada se define al comienzo del tema, no se suele utilizar en su desarrollo, sino sólo las tablas bidimensionales simples (textos [T1], [T3], [T4], [T6] y [T7]). Destacamos los textos [T1], [T6] y [T7] por la escueta definición que proporcionan, y el escaso uso de ellas (sólo un ejercicio resuelto en [T6] y dos en los textos [T1] y [T7]). Por el contrario, los textos [T2], [T5] y [T8], describen los pasos para construir una tabla de doble entrada, y hacen un uso generalizado de la misma. Incluso [T2] y [T5] presentan el procedimiento para agrupar los datos de la distribución en intervalos. Aún así, los ejercicios mayoritariamente se basan en tablas simples, debido a la necesidad de agilizar los cálculos, pues el tiempo disponible para impartir el temario es escaso.

En la tabla 4 observamos el tratamiento diferenciado de las tablas de datos en los textos, que es importante, pues según Arteaga (2011), los listados de datos no llegan a representar explícitamente la distribución de la variable bidimensional, y tendrían menor complejidad semiótica que la tabla bidimensional simple con frecuencias o la tabla de doble entrada, que han resumido las frecuencias. Es de prever un aprendizaje más completo y significativo en los alumnos que utilicen los textos basados en estas últimas. Cabe destacar que la mayoría de los textos analizados presentan la tabla bidimensional simple con frecuencias como conversión en filas o columnas de la tabla de doble entrada y sólo [T3] ofrece una definición explícita de esta representación.



A continuación se describen las representaciones gráficas, de uso privilegiado en el tema, como recurso didáctico.

#### Diagramas de dispersión y gráfico de burbujas

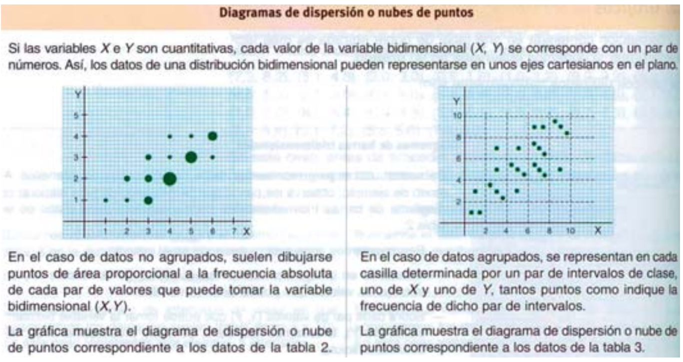
El diagrama de dispersión o nube de puntos (figura 1), representa los datos de una distribución estadística bidimensional mediante coordenadas cartesianas. Ayuda a deducir la intensidad de la relación (a través de la mayor o menor dispersión de la nube de puntos), visualizar su sentido (si la relación es directa o inversa) y el tipo (lineal o no), observando su tendencia (Sánchez Cobo, 1998). En ocasiones se añade el trazado de la recta que mejor se ajusta al mismo.

| | *Presencia en el tema* | *T1* | *T2* | *T3* | *T4* | *T5* | *T6* | *T7* | *T8* |
|---|---|---|---|---|---|---|---|---|---|
| Tabla bidimensional simple | Esencial en el desarrollo teórico y práctico | x | | | | | x | x | |
| | Uso mayoritariamente práctico | x | x | x | x | x | x | x | x |
| Tabla bidimensional simple con frecuencias | Conversión en columnas/filas de la tabla de doble entrada con uso eminentemente práctico | | | x | x | x | x | x | x |
| Tabla de doble entrada | Definición y uso en el tema | | x | x | x | x | | | x |
| | Definición y uso mínimo | x | | | | | | | |
| | Presencia anecdótica | | | | | | x | x | |

Tabla 4. Representación tabular en los textos analizados

En general se introducen mediante ejemplos y se definen posteriormente. Por ejemplo, [T1] lo describe como «conjunto de datos de una distribución bidimensional representados en ejes cartesianos» (p. 226) precisando, al final del tema, que la frecuencia absoluta de cada dato no tiene que ser necesariamente uno. Tan sólo los textos [T2], [T3] y [T8] describen su construcción a partir de una tabla de doble entrada.

Hacemos notar que los textos no diferencian el diagrama de dispersión y el *diagrama de burbujas*, en que el diámetro de cada punto es proporcional a su frecuencia absoluta (figura 1) y que podría utilizarse para representar simultáneamente tres variables (el diámetro) o incluso cuatro, mediante el color. Destacamos el texto [T5], que no precisa la importancia de la proporcionalidad del grosor del punto representado y el [T4] donde no se describe, dando por supuesta la sencillez de su construcción.


**Diagramas de dispersión o nubes de puntos**

Si las variables $X$ e $Y$ son cuantitativas, cada valor de la variable bidimensional $(X, Y)$ se corresponde con un par de números. Así, los datos de una distribución bidimensional pueden representarse en unos ejes cartesianos en el plano.

En el caso de datos no agrupados, suelen dibujarse puntos de área proporcional a la frecuencia absoluta de cada par de valores que puede tomar la variable bidimensional $(X, Y)$.

La gráfica muestra el diagrama de dispersión o nube de puntos correspondiente a los datos de la tabla 2.

En el caso de datos agrupados, se representan en cada casilla determinada por un par de intervalos de clase, uno de $X$ y uno de $Y$, tantos puntos como indique la frecuencia de dicho par de intervalos.

La gráfica muestra el diagrama de dispersión o nube de puntos correspondiente a los datos de la tabla 3.

Figura 1. Diagrama de burbujas y diagrama de dispersión ([T2], p. 220)

### Diagrama de barras tridimensional

Los textos [T1] y [T5] realizan una aproximación imprecisa a este gráfico, que no diferencian del histograma. En [T2] se define correctamente como un gráfico tridimensional utilizado para representar datos bidimensionales no agrupados en intervalos, donde para cada dato se levanta una barra de altura proporcional a su frecuencia absoluta.

### Pictograma tridimensional

Sólo [T2] define este gráfico, que es tratado como una variante del diagrama de barras donde cada barra es sustituida por dibujos. Explica su construcción y la importancia de la proporcionalidad del tamaño de los dibujos a la frecuencia.

### Histograma tridimensional

Tan sólo [T2] lo define correctamente como un gráfico tridimensional utilizado para representar datos bidimensionales agrupados en intervalos, donde para cada par de intervalos de clase se levanta un prisma de volumen proporcional a su frecuencia absoluta, explicando los pasos en su construcción.

En la tabla 5 podemos observar el uso mayoritario del diagrama de dispersión, constituyendo una herramienta indispensable para la enseñanza de la correlación y regresión, y el escaso uso de otras re-

| | *Presencia en el tema* | *T1* | *T2* | *T3* | *T4* | *T5* | *T6* | *T7* | *T8* |
|---|---|---|---|---|---|---|---|---|---|
| Diagramas de dispersión | Definición imprecisa con ejemplos | x | | | x | x | x | x | |
| | Definición correcta | | x | x | | | | | x |
| | Uso esencial en el desarrollo del tema | x | x | x | x | x | x | x | x |
| Gráfico de burbujas | Definición imprecisa | x | x | | | | | | |
| | No distingue del diagrama de dispersión | | x | | | x | | | x |
| Histograma tridimensional | Definición imprecisa | x | | | | x | | | |
| | Definición y representación correcta | | x | | | | | | |
| Pictograma tridimensional | Definición como variante del diagrama de barras y representación correcta | | x | | | | | | |
| Gráfico de barras tridimensional | Definición imprecisa | x | | | | | | | |
| | Definición y representación correcta | | x | | | | | | x |
| | Uso sin definición | | | | x | | | | |

Tabla 5. Representación gráfica en los textos analizados

presentaciones gráficas básicas como el diagrama de barras o el histograma tridimensional. Las imprecisiones en la definición del diagrama de dispersión suelen ser debidas a no considerar el caso en que la frecuencia de los datos sea distinta de uno, y que la mayoría de las veces se confunde con el diagrama de burbuja.

## Conflictos semióticos potenciales

En el análisis realizado hemos encontrado asignaciones imprecisas de significado a elementos del lenguaje, susceptibles de provocar en el estudiante un conflicto semiótico si el profesor no lo detecta. A continuación se describen los más importantes.

1. *Confusión de un concepto con su representación tabular o gráfica.* Duval (1993) indica el interés de manejar diferentes representaciones, ya que los objetos no son directamente accesibles a la percepción, aunque nunca deben ser confundidos con su representación. En el texto [T1], se define una distribución bidimensional, en dos momentos. En primer lugar se describe mediante una representación gráfica y tabular, por lo que el estudiante podría confundir el objeto (distribución) con su representación. No es hasta la siguiente página cuando se da una definición más precisa del concepto. Algo parecido ocurre con la distribución marginal que se introduce como simple etiqueta de una tabla de frecuencias ([T1], p. 237) y a diferencia del caso anterior no llega a definirse, a pesar de su relevancia.
2. *Uso inadecuado de representaciones gráficas*. Por ejemplo, en [T1] y [T5] se representan los datos de una tabla de doble entrada de una variable discreta mediante un histograma en vez de utilizar un diagrama de barras.
3. *Generalización indebida de términos.* Por ejemplo, algunos textos usan la palabra *correlación,* que sólo es aplicable a variables numéricas, como sinónimo de relación estadística (dependencia). Ello podría implicar un obstáculo didáctico (en el sentido de Brousseau, 1983) para el estudio posterior de la asociación estadística con variables cualitativas. Asimismo [T5], (p. 247) realiza un uso demasiado amplio del término *dependencia funcional* indicando que «la nube de puntos se puede ajustar también a una función que no sea una recta. A este tipo de dependencia se le denomina dependencia funcional» y añade, como ejemplo un diagrama de dispersión en que la dependencia es claramente aleatoria, pero su forma se ajusta a una parábola.
4. *Lenguaje algebraico confuso.* Al introducir una fórmula, textos como [T2] para la covarianza, aclaran el significado de los símbolos utilizados. Otros textos, como [T1] no los precisan lo que podría producir un conflicto en los estudiantes. Más aún, este texto usa fórmulas en que valores y frecuencias varían en función de un único índice *i,* mientras que los datos se dan en una tabla de doble entrada, que se representaría mejor con un índice doble (*i*, *j*), pues hay frecuencias distintas a la unidad.

Hay una gran variabilidad en los conflictos encontrados en los textos (tabla 6), desde libros que no presentan ninguno de estos conflictos hasta otros que los presentan todos. En consecuencia la posibilidad de interpretaciones incorrectas en el tema varía de uno a otro texto.

| | *T1* | *T2* | *T3* | *T4* | *T5* | *T6* | *T7* | *T8* |
|---|---|---|---|---|---|---|---|---|
| Confusión de concepto y representación | x | | | | | | | |
| Uso inadecuado de representaciones | x | x | | | x | | | x |
| Uso demasiado amplio de términos del lenguaje | x | | | x | x | x | | |
| Lenguaje algebraico confuso | x | | | | | | | |

Tabla 6. Conflictos semióticos inducidos por el lenguaje en los textos analizados

## Conclusiones

Nuestro análisis sugiere que la presentación de la correlación y regresión en los textos podría llevar un uso sesgado de diferentes representaciones (tabular, verbal, gráfica y numérica), con tendencia hacia el registro gráfico, pero sin prestar atención al proceso de construcción de estos gráficos. Más aún, el lenguaje en algunos textos podría inducir conflictos semióticos, como confundir un objeto con su representación gráfica, confundir gráficos entre sí, imprecisión del lenguaje simbólico, o generalización abusiva de conceptos. No es menos importante destacar el uso sesgado de la tabla de doble entrada en la mayoría de los textos analizados, a favor del uso casi generalizado del listado de datos, cuya complejidad semiótica, según Arteaga (2011) es insuficiente para visualizar las tendencias en los datos.

Todos estos resultados han de interpretarse con precaución, pues, de acuerdo a Lowe y Pimm (1996) el impacto del libro de texto depende no sólo del mismo libro, sino del lector, y del profesor, así como de las interacciones que determinan su uso en el aula.

## Agradecimientos

Proyecto EDU2010-14947, FPI-BES-2011-044684, FPU-AP2009-2807 (MICINN-FEDER) y grupo FQM126 (Junta de Andalucía).

## Referencias bibliográficas

Arteaga, P. (2011), *Evaluación de conocimientos sobre gráficos estadísticos y conocimientos didácticos de futuros profesores,* Tesis Doctoral, Universidad de Granada.

Azcárate, P., y A. Serradó (2006), «Tendencias didácticas en los libros de texto de matemáticas para la ESO», *Revista de Educación*, n.º 340, 341-378.

Brousseau, G. (1983), «Les obstacles épistémologiques et les problèmes en Mathématiques», *Recherches en Didactique des Mathématiques*, n.º 4 (2), 165-198.

Batanero, C., A. Estepa, y J. D. Godino (1997), «Evolution of students' understanding of statistical association in a computer based teaching environment», en J. B. Garfield y G. Burrill, (eds.), *Research on the Role of Technology in Teaching and Learning Statistics. IASE Round Table Conference Papers*, Voorburg, The Netherlands: Internacional Statistical Institute, 191-205.

Cobo, B., y C. Batanero (2004), «Significados de la media en los libros de texto de secundaria», *Enseñanza de las Ciencias,* n.º 22 (1), 5-18.

Cordero, F., y R. Flores (2007), «El uso de las gráficas en el discurso matemático escolar. Un estudio socioepistemológico en el nivel básico a través de los libros de texto», *Revista Latinoamericana de Matemática Educativa,* n.º 10 (1).

Chevallard, Y. (1991), *La transposición didáctica. Del saber sabio al saber enseñado*, Aique, Buenos Aires.

Duval, R. (1993), *Semiosis et Noesis. Lecturas en Didáctica de la Matemática: Escuela Francesa,* Sección de Matemática Educativa del CINVESTAV-IPN, México.

Estepa, A. (2008), «Interpretación de los diagramas de dispersión por estudiantes de Bachillerato», *Enseñanza de las Ciencias,* n.º 26 (2), 257-270.

Estepa, A., y C. Batanero (1995), «Concepciones iniciales sobre la asociación estadística», *Enseñanza de las Ciencias,* n.º 13 (2), 155-170.

Godino, J. D., C. Batanero y V. Font (2007), «The onto-semiotic approach to research in mathematics education», *ZDM. The International Journal on Mathematics Education*, n.º *39* (1-2), 127-135.

Herbel, B. A. (2007), «From intended curriculum to written curriculum: Examining the» voice» of a mathematics textbook», *Journal for Research in Mathematics Education*, n.º 38 (4), 344-369.

Lavalle, A. L., E. B. Micheli y N. Rubio (2006), «Análisis didáctico de regresión y correlación para la enseñanza media», *RELIME*, n.º 9 (3), 383-406.

Lowe, E., y D. Pimm (1996). 'This is so': a text on texts, en Bishop, A., Clements, K., Keitel, C., Kilpatrick, J. y Laborde, C. *In-*

*ternational Handbook of Mathematics Education*, Kluwer, Dordrecht, 371-410.

MEC (2007), *Real Decreto 1467/2007, de 2 de noviembre, por el que se establece la estructura de bachillerato y se fijan sus enseñanzas mínimas,* Madrid.

Ortiz, J. J. (1999), *Significado de los conceptos probabilísticos elementales en los textos de Bachillerato,* Tesis Doctoral, Universidad de Granada.

Ortiz, J. J., C. Batanero y L. Serrano (2001), «El lenguaje probabilístico en los libros de texto», *Suma,* n.º 38, 5-14.

Orton, A. (1990). *Didáctica de las matemáticas,* MEC y Morata, Madrid.

Pimm, D. (1987), *Speaking mathematically*, Routledge and Kegan Paul, Nueva York

Rothery, A. (1980), *Children reading mathematics*, College of Higher Education, Worcester.

Sánchez Cobo, F. T. (1998), *Significado de la correlación y regresión para los estudiantes universitarios*, Tesis doctoral no publicada, Universidad de Granada.

Sánchez Cobo, F. T., A. Estepa y C. Batanero (2000), «Un estudio experimental de la estimación de la correlación a partir de diferentes representaciones», *Enseñanza de las Ciencias,* n.º 18 (2), 297-310.

Schleppegrell, M. (2007), «The linguistic challenges of mathematics teaching and learning: A research review», *Reading and Writing Quarterly,* n.º *23,* 139-159.

Thompson, D. R. y R. N. Rubenstein (2000), «Learning mathematics vocabulary: Potential pitfalls and instructional strategies», *Mathematics Teacher,* n.º 93*,* 568-574.

Truran (1995), «Some undergraduates' understanding of the meaning of a correlation coefficient», en B. Atweh y S. Clavel, (eds.), *Proceedings of the Eigteenth Annual Conference of the Mathematics Education Research Group of Australasia (MERGA)*, Northern Territory University, Darwin, Australia, 524-529.

Zieffler, A, y J. Garfield (2009), «Modeling the growth of students' covariational reasoning during an introductory statistics course», *Statistics Education Research Journal,* n.º 8 (1)*,* 7-31.

## Anexo 1. Textos utilizados en el análisis

[T1] J. Colera, M. J. Oliveira, R. García y E. Santaella (2008), *Matemáticas aplicadas a las Ciencias Sociales I,* Grupo Anaya, Madrid.

[T2] J. Anguera, A. Biosca, M. J. Espinet, M. J. Fandos, M. Gimeno y J. Rey, (2008), *Matemáticas I aplicadas a las Ciencias Sociales*, Guadiel, Barcelona.

[T3] J. R. Vizmanos, J. Hernández, F. Alcaide, M. Moreno y E. Serrano (2008), *Matemáticas aplicadas a las Ciencias Sociales 1*, Grupo SM, Madrid.

[T4] J. M. Arias e I. Maza (2008), *Matemáticas aplicadas a las Ciencias Sociales 1,* Grupo Editorial Bruño, Madrid.

[T5] M. Antonio, L. González, J. Lorenzo, A. Molano, J. del Río, D. Santon, y M. de Vicente (2008), *Matemáticas aplicadas a las Ciencias Sociales I*, Santillana Educación, Madrid.

[T6] J. M. Martínez, R. Cuadra, A. Heras (2008), *Matemáticas aplicadas a las Ciencias Sociales. 1.º Bachillerato*, McGraw-Hill, Madrid

[T7] E. Bescós, y Z. Pena (2008), *Matemáticas aplicadas a las Ciencias Sociales. 1 Bachillerato,* Oxford University Press, Vizcaya

[T8] M. F. Monteagudo y J. Paz (2008), *1.º Bachillerato. Matemáticas aplicadas a las Ciencias Sociales,* Luis Vives, Zaragoza

M. Magdalena Gea
<mmgea@ugr.es>
Carmen Batanero
<batanero@ugr.es>

Pedro Arteaga
<parteaga@ugr.es>
Gustavo R. Cañadas
<grcanadas@ugr.es>
J. Miguel Contreras
<jmcontreras@ugr.es>
*Universidad de Granada*