

# LA REGRESIÓN EN LOS TEXTOS DE BACHILLERATO DE CIENCIAS SOCIALES

## Regression in high school social sciences textbooks

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

Universidad de Granada

### Resumen

*Se analizan los conceptos asociados a la regresión en ocho libros de texto de Bachillerato de la modalidad Humanidades y Ciencias Sociales. Se clasifican los conceptos según su definición se presente de modo formal o mediante ejemplo, y se estudian las propiedades asociadas. Encontramos algunas definiciones incompletas o parcialmente correctas, o que equiparan conceptos no equivalentes.*

**Palabras clave:** *Regresión, libros de texto, Bachillerato, conceptos y propiedades.*

### Abstract

*We study the concepts linked to regression in eight high school textbooks of the Humanities and Social Sciences modality. We classify these concepts according to the way (formally or using examples) that their definitions are presented and study the associated properties. We found some incomplete or partly correct definitions: other times non equivalent concepts are identified.*

**Keywords:** *Regression, textbooks, high school, concepts and properties.*

### INTRODUCCIÓN

La regresión es un concepto fundamental por su utilidad en la predicción en diversos campos científicos. Junto a la correlación, extiende la dependencia funcional a variables aleatorias. Se incluye en el currículo español en el primer curso de Bachillerato en las modalidades de *Ciencias y Tecnología* y *Humanidades y Ciencias Sociales* (MEC, 2007).

La investigación en psicología y didáctica de la matemática se centra preferentemente en la correlación, donde se describen dificultades 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).

El objetivo de este trabajo es analizar los conceptos y propiedades asociados a la regresión en los libros de texto del Bachillerato de Ciencias Sociales para sugerir posibles mejoras en esta presentación y completar algunos estudios previos. Forma parte de un proyecto de investigación más amplio (Gea, 2014; pendiente de presentación) en que se realiza un análisis más completo de estos textos, junto a los de la misma editorial y año en el Bachillerato de Ciencias y Tecnología para comparar la presentación del tema en las dos modalidades de Bachillerato.

Debido a la limitación de espacio, en este trabajo presentamos sólo una parte del proyecto global. Resultados complementarios se han publicado en Gea, Batanero, Cañadas y Contreras (2013), Gea, Batanero, Fernandes y Gómez (en prensa), así como en Gea, Batanero, Arteaga, Cañadas y Contreras (en prensa).

### Marco Teórico

En este trabajo nos centramos en los conceptos y propiedades que constituirían la parte conceptual de la enseñanza de la regresión. Conocimiento conceptual y procedimental son polos de un

continuo, siendo el primero más flexible y generalizable, ya que incluye la comprensión implícita o explícita de los principios de un dominio dado y sus interrelaciones (Rittle-Johnson, Siegler y Alibali, 2001). Sfard (1991) describe un concepto como constructo correspondiente al universo matemático formal y diferencia entre la definición estructural (describiendo sus condiciones o propiedades) y operacional (cuando se describe mediante una operación o fórmula) de un concepto.

Nos basamos en el enfoque Onto-semiótico (EOS) (Godino y Batanero, 1994; Godino, Batanero y Font, 2007) que considera los conceptos y propiedades como objetos matemáticos fundamentales cuyo significado puede variar en distintas instituciones. Más concretamente, Godino (2002) indica que un objeto matemático se caracteriza por su definición y propiedades; pero éstas pueden variar según las distintas instituciones en que se trate, y por tanto, hemos de concederles un carácter relativo. Puesto que estas definiciones y propiedades son evocadas por el estudiante cuando se enfrenta a la resolución de una situación problema, es importante analizar el tratamiento que se realiza de éstas en la enseñanza, ya que, la progresiva construcción de su significado depende directamente de los conceptos que se describan y utilicen.

### **Antecedentes**

Entre las investigaciones relacionadas con la comprensión de la regresión destacamos la de Estepa (1994), quien señala como puntos importantes la comprensión de la diferencia entre variable dependiente e independiente y la de Sánchez Cobo (1999) que indica dificultades en la interpretación de los coeficientes de regresión, y su relación con la pendiente de la recta y tipo de correlación. Agnelli y cols. (2009) indican que algunas de estas dificultades pueden estar ligadas al estudio previo de la función lineal que se generaliza excesivamente.

La investigación sobre libros de texto de matemáticas es amplia, siendo menor en estadística y probabilidad, donde encontramos ejemplos como los de Ortiz (1999), Cobo y Batanero (2004) y Azcárate y Serradó (2006). Respecto a la regresión destacamos las investigaciones de Sánchez Cobo (1999), quien estudia las definiciones relacionadas con la regresión en once libros de texto de Bachillerato, publicados entre 1977 y 1990, clasificándolas según se presenten apoyadas en otros conceptos, con un fin puramente procedimental, o bien una mezcla de las anteriores. En su estudio a veces no diferencia entre conceptos y propiedades por lo que éste será un punto original de nuestro trabajo.

Lavalle, Micheli y Rubio (2006) analizan la enseñanza de la regresión en siete textos, considerando los conceptos y procedimientos asociados, así como sus relaciones. Encuentran distintos niveles de profundidad en el tratamiento del tema. En cuanto a la regresión, solo un texto trata de modo estructural el concepto de recta de regresión; cuatro la ecuación de la recta, siendo únicamente dos los que además aluden a funciones de ajuste distintas a la lineal. En la mayoría se trata también la estimación de la variable dependiente.

Por su parte Gea y cols. (2013) identifican aquellas situaciones que permiten dotar de significado a la regresión al nivel que esta enseñanza se desarrolla en España, en una revisión de textos de Bachillerato de Ciencias Sociales. En Gea, Batanero, Fernandes y Gómez (en prensa) analizamos, en estos mismos libros de texto, la presentación de la distribución bidimensional y sus representaciones gráfica y tabular; tema que el estudiante aprende antes de comenzar el de la regresión propiamente dicha. En Gea, Batanero, Arteaga, Cañadas y Contreras (en prensa) estudiamos el lenguaje simbólico, expresiones algebraicas y verbales sobre la correlación y regresión. Queda pendiente el análisis complementario del modo en que los conceptos y propiedades se describen y utilizan en los textos mencionados, que es el objetivo del presente trabajo.

## METODOLOGÍA

Se analizaron ocho libros de textos de primer curso de Bachillerato en la modalidad de Humanidades y Ciencias Sociales, publicados recién implantado el currículo actual (MEC, 2007). Los libros se eligieron por ser los más utilizados en la enseñanza pública en la Comunidad Autónoma de Andalucía, y estar publicados en editoriales de gran tradición y prestigio (Anexo 1). Resaltamos el hecho de que todos han estado vigentes hasta final de 2013. Se realiza un análisis de contenido, con los siguientes pasos:

- En primer lugar se categorizan todos los conceptos implicados en el tratamiento de la regresión, analizando la forma en que se describe o presenta el concepto, que puede ser mediante ejemplos (E), formalmente (F) o bien una mezcla de ambos: proponiendo ejemplos y luego formalmente (EF), o al contrario (FE).
- Seguidamente se consideran las propiedades asociadas a cada uno de los anteriores conceptos, que no fueron consideradas en trabajos previos.

Este análisis difiere de otras técnicas de estudio documental (por ejemplo, del método histórico), porque sustituye en lo posible las interpretaciones y subjetividad del estudio de documentos o de comunicaciones por procedimientos estandarizados, con el fin de convertir en datos los contenidos analizados en los documentos (León y Montero, 2002).

## RESULTADOS Y DISCUSIÓN

Los conceptos ligados a la regresión en los libros analizados son los siguientes:

*C1. Variable dependiente e independiente.* Su diferencia ocupa un lugar central en la regresión (Estepa, 1994), con objeto de expresar una variable en función de otra. Una definición explícita se encuentra en [H8], (p.254): “*La variable dependiente es aquella que se quiere estimar, y la variable que se utiliza para ello se denomina variable independiente.*” Implícitamente se describe al resaltar que existen dos rectas de regresión.

*C2. Regresión.* Este concepto suele tratarse de modo implícito. Una definición explícita es: “*Al análisis que pretende determinar la curva que mejor aproxima un diagrama de dispersión se le llama regresión*” ([H3], p. 226).

*C3. Modelos de regresión.* El concepto de modelo es fundamental en el tratamiento de la regresión, aunque raramente aparece en los textos, que se limitan a la regresión lineal. En algunos casos, no se define pero queda implícita, por ejemplo:

Si existe una correlación fuerte entre las variables X e Y, el análisis de la regresión permite encontrar la ecuación de la función matemática que mejor se ajusta a la nube de puntos. Esta puede ser una recta, una parábola, una exponencial, una cúbica, etc. ([H4], p.226).

La situación es parecida a la descrita por Lavalle, Micheli y Rubio (2006), donde sólo dos de los siete textos incluyen a la regresión no lineal. Otros plantean tareas donde el modelo de ajuste más indicado es no lineal sin definirlo. En Sánchez Cobo (1999) sólo cinco libros analizados incluyen ejemplos de regresión distintos al lineal sin definirlo.

*C4. Coeficientes de regresión lineal.* En el modelo lineal existen dos coeficientes de regresión que corresponden a las pendientes de las rectas de regresión asociadas. Sánchez Cobo no los analiza. En nuestro estudio, algunos libros los definen en forma parcialmente correcta pues sólo definen el coeficiente de regresión de Y sobre X:

La recta que hace mínima la suma  $\sum d_i^2$  tiene por ecuación:  $y = \bar{y} + \frac{\sigma_{xy}}{\sigma_x^2}(x - \bar{x})$ . Se llama recta de regresión de Y sobre X. A la pendiente,  $\frac{\sigma_{xy}}{\sigma_x^2}$ , se la llama coeficiente de regresión ([H1], p. 230).

*C5. Rectas de regresión.* Todos los textos analizados la definen correctamente y precisan el método de mínimos cuadrados. Se encuentran también enunciados imprecisos como el siguiente donde s no se precisa también cuál es la variable dependiente e independiente.

La tabla adjunta da los alargamientos de una barra metálica por efecto de cambios en la temperatura. Calcular la recta de regresión y hacer algunas estimaciones ([H1], p.231).

Tan sólo [H8] alude al tratamiento de datos atípicos, y a la metodología de Tukey para obtener la recta de regresión respecto a la mediana. En cuanto a la utilidad predictiva de la recta de regresión, sólo la mitad de los textos analizados la resaltan. Estos resultados coinciden con los de Sánchez Cobo (1999), quien señala que se suele aludir al ajuste a los datos (siete de los once textos) más que a su carácter predictivo (tres de los once textos).

*C6. Bondad de ajuste. Coeficiente de determinación.* Estas nociones ayudan al estudiante a dotar de significado al análisis de regresión. A pesar de ello, los textos analizados no suelen incluirla; tan sólo las encontramos en [H5], [H6] siendo parcialmente correctas en [H6] ya que no menciona el coeficiente de determinación. Además, se pudiera inducir a error al, considerar la fiabilidad del modelo y la fiabilidad de la predicción como nociones equivalentes. Sin embargo no lo son, puesto que la primera se refiere a la fiabilidad global de la predicción y la segunda se reduce a la de un punto particular.

En la Tabla 1 se resume la forma de definición de estos conceptos, observando un predominio de definiciones acompañadas de ejemplos, sobre todo en [H6]. La recta de regresión se define en todos, y se suelen acompañar de ejemplos; en la mitad de los textos antes y en el resto después de su definición. Generalmente se introduce de modo estructural (Sfard, 1991), utilizando diagramas de dispersión que añaden la recta de mínimos cuadrados mediante ejemplos que muestran la distancia de los puntos de la nube a la recta, evidenciando la diferencia entre el valor observado (ordenadas del punto) y el predicho (punto sobre la recta); finalmente se ofrece la definición operacional de la misma. Igual ocurre en la investigación de Sánchez Cobo (1999).

Tabla 1. Conceptos asociados a la regresión y forma de introducción

Conceptos	H1	H2	H3	H4	H5	H6	H7	H8
C1.Variable dependiente e independiente					F			F
C2.Regresión			EF	FE				
C3.Modelos no lineales de regresión					FE	FE		
C4.Coefficientes de regresión lineal	F	F		F				
C5.Rectas de regresión	EF	EF	F	FE	FE	FE	FE	F
C6.Bondad de ajuste. Coef. de determinación					FE	FE		

Destacamos los pocos textos que incluyen otras definiciones, como los coeficientes de regresión, o la bondad del ajuste, induciendo un tratamiento meramente operacional de la regresión, pudiendo así olvidar la importancia de aspectos estructurales de la misma, y con ello, un aprendizaje más significativo.

### Propiedades consideradas

Los libros de texto suelen presentar atributos o propiedades de los conceptos definidos, cuya naturaleza es epistémica, y por tanto institucional. Se han encontrado los siguientes:

*P1. Modelos de regresión.* Puesto que los textos no suelen incluir modelos no lineales, estas propiedades se refieren a aspectos generales del análisis de regresión. Un ejemplo es el siguiente:

Si en una variable  $(X,Y)$  existe una correlación fuerte entre las variables  $X$  e  $Y$ , el análisis de la regresión permite encontrar la ecuación de la función matemática que mejor se ajusta a la nube de puntos. ([H4], p.226).

*P2. La recta de regresión hace mínima la suma de los cuadrados de las distancias de los puntos (abscisas y ordenadas) de la distribución bidimensional a la recta.* Todos los textos la incluyen, generalizando la propiedad equivalente de la media (la media es el valor que minimiza la suma de distancias de los datos).

*P3. Dos rectas de regresión diferentes.* En todos los textos se indica implícitamente que hay dos rectas de regresión diferentes, aunque algunos lo hacen explícito. Como Sánchez Cobo (1999), creemos que si se presenta esta propiedad se podría prevenir que los estudiantes obtuvieran una de las rectas despejando la variable de la otra recta calculada previamente. Un ejemplo es:

Las rectas de regresión de  $Y$  sobre  $X$  y de  $X$  sobre  $Y$  son distintas, por eso hay que saber qué variable es la dependiente, pues  $X$  e  $Y$  no son intercambiables. (Es posible que tenga sentido deducir la variable  $Y$  a partir de la  $X$ , mientras que deducir la  $X$  a partir de la  $Y$  carezca de significado) ([H5], p.258).

*P4. Propiedades del Coeficiente de determinación.* La principal es que el coeficiente de determinación informa de la fiabilidad del ajuste lineal. Otros textos analizan su relación con la proporción de varianza explicada:

En ocasiones, con el fin de calcular la calidad o bondad del ajuste realizado mediante la recta de regresión y, por tanto, la fiabilidad de las predicciones que con ella se puedan realizar, se utiliza la expresión  $(r^2 \cdot 100)\%$  que nos da el porcentaje en el que la variable  $Y$  se justifica por el valor de la variable  $X$  ([H6], p.185).

*P5. La recta de regresión permite estimar valores de la variable dependiente a partir de la independiente.* Pocos textos precisan esta propiedad. Un ejemplo es: “La recta de regresión nos permite predecir valores de una variable a partir de los de la otra” ([H3], p.227).

*P6. Las estimaciones obtenidas con la recta de regresión son aproximaciones, en términos de probabilidad, al valor real.* Esta propiedad no se suele incluir, aunque en su mayoría se usa de modo implícito, pues se indica que, al contrario del caso de la dependencia funcional, para cada valor de la variable independiente, corresponden varios valores de la dependiente, siendo el valor proporcionado por la recta de regresión el promedio de todos ellos. Así:

Las estimaciones siempre se realizan aproximadamente y en términos de probabilidad: es probable que si  $x = x_0$  entonces y valga, aproximadamente,  $\hat{y}(x_0)$ . ([H1], p.230).

*P7. La estimación de cada variable se realiza con su correspondiente recta de regresión.* Se incluye generalmente, y algunos textos señalan que si el coeficiente es próximo a 1 o -1, se puede utilizar una única recta de regresión para predecir cualquiera de las variables del estudio.

### **Relaciones entre conceptos**

Además de las propiedades que se refieren a un solo concepto se han determinado las siguientes relaciones:

*R1. La recta de regresión pasa por el centro de gravedad de la distribución.* En el estudio de Sánchez Cobo (1999), siete de los once textos citan esta propiedad, aunque dos lo hacen de modo implícito.

*R2. Las estimaciones con la recta de regresión son mejores en valores cercanos a la media de la variable independiente.* Un ejemplo es: “Las predicciones obtenidas para valores próximos al

*punto medio de la distribución son más fiables que las obtenidas para valores muy alejados*" ([H3], p.227). No suele aparecer en muchos textos ya que, por lo general, la fiabilidad de la predicción se juzga por la proximidad del coeficiente de correlación a 1 o -1, aunque esta condición es insuficiente.

*R3. La fiabilidad de la predicción con la recta de regresión aumenta con el tamaño de muestra.* Se presenta en pocos de los textos analizados a pesar de ser su utilidad. Un ejemplo es: "*La fiabilidad aumenta al aumentar los datos. Una recta obtenida a partir de pocos datos genera grandes riesgos, aunque r sea muy alto*". ([H5], p.260).

*R4. El tipo de dependencia informa de la posición de las rectas de regresión.* Así es que, dependiendo de si las variables presentan dependencia funcional o independencia, las rectas de regresión son coincidentes o perpendiculares, respectivamente.

*R5. Recta de regresión y signo de la covarianza.* Se interpreta la covarianza según la proximidad de los puntos a la recta de regresión, y se relaciona el signo y los coeficientes de regresión, o pendiente de las rectas de regresión. En el estudio de Sánchez Cobo (1999) ninguno de los textos los incluye:

Según donde esté situado  $(x_i, y_i)$  respecto a  $(\bar{x}, \bar{y})$ , el área  $(x_i - \bar{x}) \cdot (y_i - \bar{y})$  será positiva o negativa. Si los puntos están próximos a una recta de pendiente positiva, los sumandos son casi todos positivos y la covarianza es grande ([H1], p.228).

*R6. Recta de regresión y sentido de la correlación.* De la pendiente de la recta de regresión se puede deducir el signo de la correlación:

Si la pendiente de la recta de regresión es positiva o negativa, la correlación se llama positiva o negativa, respectivamente. ([H1], p.227).

*R7. El producto de los dos coeficientes de regresión es  $r^2$ .* Esta proposición se encuentra en la mayoría de los textos, y también en los de Sánchez Cobo (1999).

*R8. Signo de los coeficientes de regresión y correlación.* El coeficiente de correlación tiene el mismo signo que los coeficientes de regresión. Si no se definen los coeficientes de regresión, esta propiedad se formula en términos de la pendiente de la recta.

*R9. Relación entre las estimaciones con la recta de regresión y el coeficiente de correlación.* La mayoría de los textos relacionan ambos conceptos como por ejemplo:

La fiabilidad de los cálculos obtenidos mediante las rectas de regresión será tanto mayor cuanto mayor sea el valor absoluto del coeficiente de correlación lineal  $r$ .

Si  $r$  está muy próximo a cero, no tiene sentido realizar ninguna estimación o previsión.

Si  $r$  está próximo a 1 o a -1, los valores reales serán, probablemente, próximos a nuestras estimaciones.

Si  $r = 1$  o  $r = -1$ , los valores reales coincidirán con las estimaciones efectuadas ([H4], p.226)

*R10. Posición relativa de las rectas de regresión y coeficiente de correlación.* No es muy habitual encontrar esta propiedad geométrica sobre el ángulo que forman las rectas en función del valor del coeficiente.

En la Tabla 2 presentamos un resumen de las propiedades y relaciones que se incluyen en el tema, que se refieren al análisis de regresión. Observamos diferencia en los textos, desde algunos que apenas incluyen propiedades (como [H2]) hasta otros como [H5] y [H6] que presentan casi todas las analizadas. La propiedad más frecuentemente encontrada es la de mínimos cuadrados y el uso de la recta específica para hacer predicciones, así como que la recta permite hacer estimaciones. En todo caso, que se ha indicado, algunas otras propiedades se utilizan, sin mencionarlas explícitamente en los textos.

Tabla 2. Propiedades de la regresión

	H1	H2	H3	H4	H5	H6	H7	H8
P1. Modelos de regresión				x	x	x		x
P2. La recta de regresión minimiza la suma de los cuadrados de las distancias de los puntos de la distribución a la recta	x	x	x		x	x	x	x
P3. Dos rectas de regresión diferentes			x		x	x		
P4. Propiedades del Coeficiente de determinación					x	x		
P5. La recta de regresión permite realizar estimaciones		x	x	x	x	x	x	
P6. Una predicción es una aproximación al valor real de la variable que se estima	x		x	x			x	
P7. Uso de la recta específica para realizar predicciones	x		x	x	x	x	x	
Relaciones								
R1. Recta de regresión - centro de gravedad	x	x	x		x	x	x	x
R2. Predicción - centro de gravedad	x	x	x		x	x	x	x
R3. Tamaño de la muestra - recta regresión - predicción			x		x			
R4. Tipo de dependencia - recta de regresión			x	x				
R5. Covarianza - recta de regresión	x				x	x		
R6. Correlación - recta de regresión	x		x		x		x	
R7. Producto c. regresión - coef. de correlación lineal	x		x					
R8. Coeficiente de correlación - coeficientes de regresión	x		x	x				x
R9. Coeficiente de correlación - predicción	x	x	x	x	x	x	x	
R10. Posición rectas de regresión - coef. de correlación	x		x				x	x

## CONCLUSIONES

El análisis realizado permite cumplir el objetivo de este trabajo, que fue analizar los conceptos y propiedades asociados a la regresión en los libros de texto del Bachillerato de Ciencias Sociales, completando los estudios previos.

Respecto al estudio de Sánchez Cobo (1999), nuestro trabajo lo completa al analizar libros de texto actuales, pues los estudiados por el autor fueron publicados entre 1977 y 1990. Además estudiamos en forma diferenciada los conceptos y propiedades; clasificando también las definiciones de los conceptos siguiendo la propuesta de Sfard (1991). Tampoco Lavalle, Micheli y Rubio (2006) estudian las propiedades en los textos y además su estudio se ha realizado sobre textos argentinos.

Además, al estudiar los conceptos y propiedades de la regresión complementamos nuestros trabajos previos sobre las situaciones problema (Gea, Batanero, Cañadas y Contreras, 2013), la distribución de datos bidimensional (Gea, Batanero, Fernandes y Gómez, en prensa) y el lenguaje de la correlación y regresión (Gea, Batanero, Arteaga, Cañadas y Contreras, en prensa).

Para completar el objetivo, a continuación utilizamos los resultados para sugerir posibles mejoras en la presentación de las definiciones y propiedades de la regresión en los libros de texto.

En primer lugar, los resultados muestran la escasa presencia de definiciones relativas a la regresión, exceptuada la de recta de regresión, posiblemente por la amplitud que se dedica al concepto de correlación, que en algunos casos se confunde con el de regresión. Pensamos que es importante que los textos den la relevancia debida a la regresión, pues una vez aceptada una correlación entre las variables, la verdadera utilidad de este tema está en tratar de ajustar un modelo de regresión a los datos. Recomendamos consecuentemente reforzar las definiciones de regresión, modelo de regresión, modelos no lineales, rectas de regresión y coeficientes de regresión.

En general, los textos cubren todas las propiedades importantes relativas a la regresión; la variedad y cantidad de estas propiedades y relaciones muestra la gran complejidad semiótica del tema, debido a la riqueza conceptual y de propiedades de la configuración epistémica asociada, que el

alumno ha de aprender. Sin embargo, podemos apreciar también un escaso tratamiento estructural de estas propiedades y relaciones. La mayoría de los textos enfatizan la presentación operacional, mostrando a un énfasis excesivo en el cálculo, sin poner atención al significado y aspectos interpretativos. Así es que se presentan los cálculos para determinar cada una de las rectas de regresión, pero no se analizan a fondo las propiedades que relacionan la regresión con la correlación. Además, la relación de la recta de regresión con el coeficiente de correlación lineal se suele establecer sólo para determinar la fiabilidad de la predicción, sin analizar e interpretar la conexión entre correlación y regresión lineal, propiamente dicha. Por otra parte, sería importante que los libros resalten esta relación para que los estudiantes lleguen a comprender que es posible un coeficiente de correlación bajo o moderado en un diagrama de dispersión de dependencia alta pero no lineal (por ejemplo, parabólica).

En otros casos, las relaciones que se establecen entre la recta de regresión y el coeficiente de correlación se presentan en forma muy apresurada, por lo que el profesor en la clase debiera organizar un debate y reflexión en torno a las mismas. De otro modo, los estudiantes pudieran realizar interpretaciones incorrectas que les lleven a error; por ejemplo, si al estudiar la relación entre el coeficiente de correlación y la recta de regresión no se diferencia la existencia de dos rectas. Los estudiantes pueden llegar a pensar que con una misma recta se pueden realizar predicciones para cada una de las variables; deducción que es cierta si el valor de  $|r|$  es próximo a 1, pero no en caso general.

Finalizamos recomendando interpretar nuestros resultados con precaución pues el impacto del libro de texto sobre el estudiante depende no sólo del mismo libro, sino del alumno y sus conocimientos previos, así como de la forma que el profesor lo usa en el aula (Lowe y Pimm, 1996)

### Agradecimientos

Proyecto EDU2010-1494 y grupo FQM126 (Junta de Andalucía).

### Referencias

- Agnelli, H., Konic, P., Peparelli, N. Z. y Flores, P. (2009). La función lineal obstáculo didáctico para la enseñanza de la regresión lineal. *Revista Iberoamericana de Educación Matemática*, 17, 52-61.
- Azcárate, P. y Serradó, A. (2006). Tendencias didácticas en los libros de texto de matemáticas para la ESO. *Revista de Educación*, 340, 341-378.
- Cobo, B. y Batanero, C. (2004). Significados de la media en los libros de texto de secundaria. *Enseñanza de las Ciencias*, 22 (1), 5-18.
- Estepa, A. (1994). *Concepciones iniciales sobre la asociación estadística y su evolución como consecuencia de una enseñanza basada en el uso de ordenadores*. Tesis doctoral. Universidad de Granada.
- Estepa, A. (2008). Interpretación de los diagramas de dispersión por estudiantes de Bachillerato. *Enseñanza de las Ciencias*, 26 (2), 257-270.
- Estepa, A. y Batanero, C. (1995). Concepciones iniciales sobre la asociación estadística. *Enseñanza de las Ciencias*, 13 (2), 155-170.
- Gea, M. M. (2014). *Evaluación y desarrollo del conocimiento matemático para la enseñanza de la probabilidad en futuros profesores de educación primaria*. Tesis doctoral, pendiente de lectura.
- Gea, M. M., Batanero, C., Cañadas, G. R. y Contreras, J. M. (2013). Un estudio empírico de las situaciones-problema de correlación y regresión en libros de texto de bachillerato. En A. Berciano, G. Gutiérrez, A. Estepa y N. Climent (Eds.), *Investigación en Educación Matemática XVII* (pp. 293-300). Bilbao: Sociedad Española de Investigación en Educación Matemática (SEIEM).
- Gea, M. M., Batanero, C., Arteaga, P., Cañadas, G. R., Contreras, J. M. (En prensa). Análisis del lenguaje sobre la correlación y regresión en libros de texto de bachillerato. *SUMA*,



- Gea, M. M., Batanero, C., Fernandes, J. A. y Gómez, E. (en prensa). La distribución de datos bidimensionales en los libros de texto de matemáticas de Bachillerato. *Quadrante*. Godino, J. D. y Batanero, C. (1994). Significado institucional y personal de los objetos matemáticos. *Recherches en Didactique des Mathématiques*, 14 (3), 325-355
- Godino, J. D. (2002). Un enfoque ontológico y semiótico de la cognición matemática. *Recherches en Didactique des Mathématiques*, 22 (2-3), 237-284.
- 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.
- Lavalle, A. L., Micheli, E. B. y Rubio, N. (2006). Análisis didáctico de regresión y correlación para la enseñanza media. *Revista Latinoamericana de Investigación en Matemática Educativa*, 9 (3), 383-406.
- León, O. G. y Montero, I. (2002). *Métodos de Investigación*. Madrid: Mc Graw Hill.
- Lowe, E. y Pimm, D. (1996). 'This is so': a text on texts. En A. Bishop, K. Clements, C. Keitel, J. Kilpatrick y C. Laborde (Eds.). *International Handbook of Mathematics Education*, (pp. 371-410). Dordrecht: Kluwer.
- M.E.C. (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: Autor.
- Ortiz, J. J. (1999). *Significado de los conceptos probabilísticos elementales en los textos de Bachillerato*. Tesis Doctoral. Universidad de Granada.
- Rittle-Johnson, B., Siegler, R. S., y Alibali, M. W. (2001). Developing conceptual understanding and procedural skill in mathematics: An iterative process. *Journal of educational psychology*, 93 (2), 346-362.
- Sfard, A. (1991). On the dual nature of mathematical conceptions: Reflections on processes and objects as different sides of the same coin. *Educational Studies in Mathematics*, 22 (1), 1-36.
- Sánchez Cobo, F.T. (1999). *Significado de la correlación y regresión para los estudiantes universitarios*. Tesis doctoral no publicada. Universidad de Granada.
- Zieffler, A, y Garfield, J. (2009). Modeling the growth of students' covariational reasoning during an introductory statistics course. *Statistics Education Research Journal*, 8 (1), 7-31.

### **Anexo 1: Textos utilizados en el análisis**

- [H1]. Colera, J., Oliveira, M.J., García, R. y Santaella, E. (2008). *Matemáticas aplicadas a las Ciencias Sociales I*. Madrid: Grupo Anaya.
- [H2]. Anguera, J., Biosca, A., Espinet, M. J., Fandos, M.J., Gimeno, M. y Rey, J. (2008). *Matemáticas I aplicadas a las Ciencias Sociales*. Barcelona: Guadiel.
- [H3]. Vizmanos, J. R., Hernández, J. y Alcaide, F. Moreno, M. y Serrano, E. (2008). *Matemáticas aplicadas a las Ciencias Sociales I*. Madrid: Grupo SM.
- [H4]. Arias, J. M. y Maza, I. (2008). *Matemáticas aplicadas a las Ciencias Sociales I*. Madrid: Grupo Editorial Bruño.
- [H5]. Antonio, M., González, L., Lorenzo, J. Molano, A., del Río, J., Santon, D. y de Vicente, M. (2008). *Matemáticas aplicadas a las Ciencias Sociales I*. Madrid: Santillana Educación.
- [H6]. Martínez, J. M., Cuadra, R., Heras, A. (2008). *Matemáticas aplicadas a las Ciencias Sociales. 1º Bachillerato*. Madrid: McGraw-Hill.
- [H7]. Bescós, E. y Pena, Z. (2008). *Matemáticas aplicadas a las Ciencias Sociales. I Bachillerato*. Vizcaya: Oxford University Press.
- [H8]. Monteagudo, M. F. y Paz, J. (2008). *1º Bachillerato. Matemáticas aplicadas a las Ciencias Sociales*. Zaragoza: Luis Vives.