

# ANÁLISIS DE LOS CONCEPTOS ASOCIADOS A LA CORRELACIÓN Y

REGRESIÓN EN LOS TEXTOS DE BACHILLERATO

Carmen Batanero, M<sup>a</sup> Magdalena Gea, M<sup>a</sup> del Mar López-Martín y Pedro Arteaga

Didacticae 1, 60-76, 2017.

DOI: 10.1344/did.2017.1.60-76

## RESUMEN

El objetivo de esta investigación fue analizar las definiciones de los conceptos ligados a la correlación y regresión en 16 libros de texto de Bachillerato. Se identificaron los conceptos de dependencia funcional y estadística, covarianza, correlación, coeficiente de correlación, regresión, modelos de regresión, regresión lineal, recta de regresión, coeficiente de determinación y bondad de ajuste. Para cada uno de ellos se estudia el modo en que se definen (en forma operacional o estructural, o a través de ejemplos) y su uso a lo largo del tema. Se concluye que la presentación es, en ocasiones, incompleta o parcialmente correcta y otras se equiparan conceptos no equivalentes.

**Palabras clave:** correlación y regresión, conceptos relacionados, definiciones, libros de textos, bachillerato.

## ANALYSIS OF CONCEPTS LINKED TO CORRELATION AND REGRESSION IN HIGH SCHOOL TEXTBOOKS

### ABSTRACT

The aim of this study was analysing the definitions linked to correlation and regression in 16 Spanish high school textbooks. The concepts of functional and statistical dependence, covariance, correlation, correlation coefficient, regression, regression model, lineal regression, regression line, determination coefficient and goodness of fit were identified. For each of these concepts we study the way in which they are defined (operational or structural definition or using examples) and its use along the theme. We concluded that the presentation is sometimes incomplete or in part erroneous, also some of them use concepts with the same status as other but are not equivalent.

**Keywords:** correlation and regression, related concepts, definitions, textbooks, high school.

## 1. INTRODUCCIÓN

El estudio de la relación entre dos variables y la posible predicción de una de ellas (variable respuesta o dependiente) a partir de la otra (variable explicativa o independiente), es decir, de la correlación y regresión, es fundamental en estadística y en la investigación empírica. Por un lado, este estudio se vincula a la toma de decisiones y, por otro, extiende el concepto de dependencia funcional a situaciones aleatorias (Crocker, 1981; Moritz, 2004; Lavalle, Micheli y Rubio, 2006).

Debido a esta importancia, el tema se incluye en España en el primer curso de Bachillerato en las modalidades de *Ciencias y Tecnología*, y *Humanidades y Ciencias Sociales* en el currículo anterior (MEC, 2007), donde se fijaron los siguientes contenidos: “*Distribuciones bidimensionales. Interpretación de fenómenos sociales y económicos en los que intervienen dos variables a partir de la representación gráfica*

*de una nube de puntos. Grado de relación entre dos variables estadísticas. Regresión lineal. Extrapolación de resultados.”* (MEC, 2007, p. 45475). En forma muy similar se introduce en las especialidades de *Ciencias y Humanidades* y *Ciencias Sociales* en el nuevo currículo (MECD, 2015).

La enseñanza del tema no es simple, pues se basa en la comprensión de múltiples conceptos como los de variable estadística y distribución bidimensional; dependencia estadística, funcional e independencia; covarianza y correlación; regresión; variable dependiente e independiente; modelo y modelo lineal; bondad de ajuste y coeficiente de determinación. Investigaciones previas, como por ejemplo el estudio de Estepa (2007), muestran que algunos alumnos confunden estos conceptos o tienen dificultades en la realización de tareas relacionadas con los mismos.

Para facilitar el aprendizaje de los estudiantes sería necesario asegurar una correcta presentación de estos conceptos en los libros de texto, que se convierte en un recurso de gran importancia para el alumno y el profesor (Herbel, 2007; Cordero y Flores, 2007), y es un tema poco investigado hasta la fecha. Aunque algunos autores han analizado la presentación de la correlación y regresión en libros de texto (Sánchez Cobo, 1999; Sánchez Cobo, Estepa y Batanero, 2000; Lavalle, Micheli y Rubio, 2006), no se centraron específicamente en las definiciones de los conceptos y se realizaron hace tiempo o sobre textos argentinos. En trabajos previos hemos analizado los problema de correlación y regresión (Gea, Batanero, Cañas y Contreras, 2013) y el lenguaje de la correlación y regresión (Gea, Batanero, Arteaga, Cañas y Contreras, 2014) en una muestra de libros de texto españoles recientes.

El objetivo del estudio que se presenta en este artículo es completar los anteriores, analizando las definiciones de los citados conceptos en una muestra de libros de texto recientes de Bachillerato. En lo que sigue analizamos los fundamentos, métodos y resultados del estudio, finalizando con algunas conclusiones para la enseñanza.

## 2. FUNDAMENTOS

### 2.1. MARCO TEÓRICO

Un elemento fundamental en la construcción del conocimiento matemático son los conceptos involucrados en la resolución de los problemas. Conocimiento conceptual y procedimental son polos de un continuo, siendo el primero más flexible y generalizable, ya que incluye la comprensión de los principios de un dominio dado y sus

interrelaciones (Rittle-Johnson y Alibali, 1999; Rittle-Johnson, Siegler y Alibali, 2001).

La importancia de las definiciones de los conceptos ha sido analizada por Zazkis y Leikin (2008), quienes señalan la necesidad de determinar el conjunto de propiedades y justificaciones que se usarán en la enseñanza y la forma en que se presentará el tema a los estudiantes. Por ello, su análisis es un problema relevante de investigación, ya que modelan la relación entre esta definición y la imagen del concepto que crean los estudiantes (Vinner, 1991).

Sfard (1991) describió los conceptos como constructos teóricos que corresponden a la matemática formal y diferenció entre la definición estructural (describiendo sus condiciones o propiedades) y operacional (cuando se describe mediante una operación o fórmula) de un concepto, sugiriendo que ambas son complementarias.

Skemp (1993) indica que los conceptos no pueden ser comunicados simplemente a través de una definición; al definir por primera vez un concepto sería necesario proporcionar una colección apropiada de ejemplos, apoyados en otros conceptos que ya se conocen. Joyce, Weil y Calhoun (1986), por su parte, detallan diversos principios a considerar cuando se define un concepto: (a) la denominación que se asocia al concepto (palabra o símbolo); (b) sus atributos o propiedades específicas que lo diferencian de otros conceptos; (c) los ejemplos y contraejemplos del concepto; y (d) la regla de definición, que es una afirmación sintética y precisa elaborada a partir de sus atributos esenciales. Vergnaud (1982), asimismo, considera el concepto matemático como una tripleta formada por el conjunto de situaciones que lo hacen significativo; el conjunto de invariantes que lo constituyen; y las representaciones simbólicas asociadas.

Leikin y Winicki-Landman (2001) indican que las definiciones dan nombre a los conceptos y establecen sus condiciones necesarias y suficientes; deben incluir un número mínimo de tales condiciones y suelen incluir otros conceptos previamente definidos. Mientras que, desde el punto de vista matemático, definir supone dar un nombre a un objeto que previamente existía y estaba caracterizado por unas propiedades, en la enseñanza, la definición da vida a algo que no existía para el estudiante (Mariotti y Fischbein, 1997).

Godino (2003) indica que un objeto matemático se caracteriza por su definición y el enunciado de sus propiedades que pueden variar según las instituciones en que se trate, y por tanto, tienen un carácter relativo. En el marco teórico del enfoque onto-semiótico propuesto por este autor, se consideran, además de los conceptos o definiciones asociadas, otros objetos matemáticos como el lenguaje, las propiedades, los

procedimientos y los argumentos. En este trabajo nos centramos específicamente en los conceptos asociados a la correlación y regresión, para complementar nuestros trabajos previos en los que hemos analizado, dentro de este mismo tema, los problemas y el lenguaje. Puesto que las definiciones de conceptos son evocadas por el estudiante cuando se enfrenta a la resolución de problemas, es importante analizar el tratamiento que se realiza de éstas en la enseñanza.

## 2.2. ANTECEDENTES

Aunque hay una amplia investigación sobre los libros de texto de matemáticas, esta tradición es menor en estadística y menor aun las que se centran en las definiciones de los conceptos, que se resumen a continuación.

Cobo y Batanero (2004) analizan las definiciones de la media aritmética en 22 textos de secundaria, identificando las propiedades que se les atribuyen y las representaciones asociadas. Las autoras aprecian que algunos de los textos estudiados desarrollan significados parciales de la media al omitir propiedades básicas. García-Alonso (2011) compara las definiciones de conceptos de inferencia en cuatro libros de texto de Bachillerato de Ciencias Sociales, mostrando, en ocasiones, sesgos en las definiciones; además, algunos libros usan ejemplos para definir un concepto sin llegar a definirlo en forma estructural, mediante sus propiedades.

El primer antecedente directo de nuestro trabajo es el de Sánchez Cobo (1999), que estudia las definiciones relacionadas con la correlación y regresión en 11 textos de Bachillerato publicados entre 1977 y 1990, clasificándolas según describan sus propiedades (definición estructural, según Sfard, 1991), en forma operacional o bien una mezcla de las anteriores. El autor observa que sólo dos libros presentan la definición de los conceptos antes de dar ejemplos de ellos y la mayoría se reduce a una definición operacional. Analiza con detalle los contenidos expuestos, indicando que sólo tres libros incluyen la diferenciación entre dependencia funcional, aleatoria e independencia. La mayoría de los textos incluyen la correlación y la diferenciación de las dos rectas de regresión. Más recientemente, Lavalle, Micheli y Rubio (2006) analizan la correlación y regresión en siete libros de texto argentinos de Bachillerato, pero no se centran en las definiciones.

En lo que sigue se presenta nuestro trabajo, que complementa los anteriores, comenzando por la descripción de la metodología. Seguidamente se presentan los resultados, y se finaliza con algunas implicaciones para la enseñanza.

### 3. METODOLOGÍA

Se analizaron 16 libros de texto de Bachillerato, de publicación reciente, 8 de cada una de las modalidades en que se enseñan estos contenidos (*Humanidades y Ciencias Sociales* y *Ciencia y Tecnología*). Para cada editorial se estudian los textos correspondientes a las dos modalidades citadas (ver Anexo) que actualmente se utilizan y corresponden a editoriales de prestigio, elegidas por su amplia difusión en Andalucía. En ellos se identificaron los conceptos presentados en el tema de correlación y regresión, clasificándolos respecto a los dos campos de problema principales que se determinaron en un estudio histórico previo del tema (Estepa, Gea, Cañadas y Contreras, 2012): análisis de dependencia entre variables y predicción.

Sobre cada uno de los conceptos identificados en estos dos campos de problemas se analizan dos variables:

V1.*Tipo de definición del concepto*, que puede ser mediante ejemplos (E), usualmente utilizando implícitamente la definición, o mediante una definición explícita formal (F); que de acuerdo a Sfard (1991) puede ser operacional (O) o bien estructural (S), o bien mezclando varios de estos tipos de definiciones. Se ha tenido también en cuenta si los ejemplos se proponen antes de la definición, como es sugerido por Skemp (1993) y Joyce, Weil y Calhoun (1986).

V2.*Uso* en el tema. Para analizar la importancia relativa de cada concepto, se ha tenido en cuenta si el uso del concepto es continuado a lo largo de todo el tema (S), si se usa esporádicamente (P), o se define pero no se usa (N).

En lo que sigue presentamos y discutimos los resultados. En primer lugar se realiza un análisis cualitativo de las características específicas de las definiciones de los conceptos identificados, mostrando cuando sea preciso ejemplos de la forma en que se definen. Seguidamente se realiza una síntesis del tipo de definición y su uso a lo largo del tema.

### 4. RESULTADOS Y DISCUSIÓN

#### 4.1. CONCEPTOS LIGADOS AL ESTUDIO DE LA DEPENDENCIA ENTRE DOS

## VARIABLES

Comenzamos nuestro análisis por los conceptos ligados a la dependencia estadística pues, como indica Sánchez Cobo (1999), su estudio es un paso previo al análisis de regresión, ya que la existencia de dependencia entre variables es la motivación en la búsqueda de modelos que permitan expresar relaciones funcionales entre las mismas. A continuación se analizan los conceptos asociados, asignándoles un código que se utiliza posteriormente en las diferentes tablas de resumen de resultados.

### 4.1.1. Características de las definiciones

*C11. Dependencia funcional.* Conviene que al comenzar el estudio de la correlación y regresión, los libros recuerden la definición de dependencia funcional. Al contrario que en el estudio de Sánchez Cobo (1999), todos nuestros textos la describen, diferenciándola de la estadística, aunque prácticamente se utiliza sólo al inicio del tema para introducir otros conceptos como la correlación, covarianza y el coeficiente de correlación lineal. Es frecuente presentarla a partir de ejemplos, a veces presentando como contrapunto otros de dependencia aleatoria. Por ejemplo, en [T1] (p. 331), se incluye como ejemplo la altura que alcanza una piedra al lanzarla hacia arriba con diferente fuerza, resaltando que a cada valor de la variable independiente (fuerza) se da sólo uno de la dependiente (altura).

Algunos textos ([H2], [H8] y [T2], [T8]) consideran sinónimos los conceptos de dependencia y correlación, lo que muestra una falta de precisión en los mismos pues, el concepto de dependencia es más amplio. Por ejemplo, [H2] y [T2] definen *correlación funcional* del siguiente modo: “*Correlación funcional: todos los puntos están situados sobre una recta o una curva. Existe una relación funcional entre las variables X e Y*” ([H2], p.248). Por otra parte, en [H7] y [T7] se identifica la relación funcional con la existencia de relación no lineal, sin tener en cuenta que puede haber dependencia estadística de tipo no lineal: “*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*” ([T7], p. 305).

*C12. Dependencia estadística o aleatoria.* Al definir este nuevo tipo de dependencia se amplía el concepto de dependencia funcional, ya conocido por los estudiantes. La dependencia estadística se define en todos los textos, salvo en [H2] y [T2], siempre con ejemplos para facilitar su comprensión. En ocasiones se identifica con correlación, aunque el concepto de dependencia estadística es más amplio, pues la correlación se

limita al tratamiento de variables cuantitativas, mientras que también existe la dependencia estadística entre variables cualitativas o mixtas. Sólo [H4] y [T4] precisan las diferencias entre ambas nociones:

En ocasiones se observa que existe una relación entre las variables, pero dicha relación no puede expresarse como una función matemática. En este caso se dice que entre las variables X e Y existe una dependencia estadística, que podrá ser fuerte o débil” ([T4], p. 338).

*C13. Independencia.* Este concepto no siempre es bien comprendido por los estudiantes, como se mostró en Cañas (2012). Sirve de apoyo para el desarrollo de las ideas de correlación y regresión y, de manera implícita o explícita, fundamenta el diseño de muchas tareas.

A pesar de su uso generalizado en este tema, las definiciones que se presentan en los textos suelen ser imprecisas o parcialmente correctas, ya que se identifica correlación nula e independencia. Estos términos no son idénticos pues, en caso de dependencia no lineal (por ejemplo parabólica), pudiera darse un coeficiente de correlación igual a cero, siendo las variables dependientes. Una definición correcta se incluye en [H4] y [T4]: “*Cuando no existe dependencia funcional ni estadística, se dice que hay independencia estadística entre las variables*” ([H4], p. 224).

En [H1], [H2], [H8] y [T1], [T2], [T8] no se comenta la diferencia entre variables independientes e incorreladas. Por otro lado, [H7] y [T7] no definen la independencia, utilizando en su lugar expresiones como “dependencia débil”, aunque incluyen tareas donde la relación es prácticamente de independencia como se muestra en la Figura 1.

A partir de los diagramas de dispersión, decide si hay o no dependencia lineal y, en su caso, si es fuerte o débil, y si es positiva o negativa.

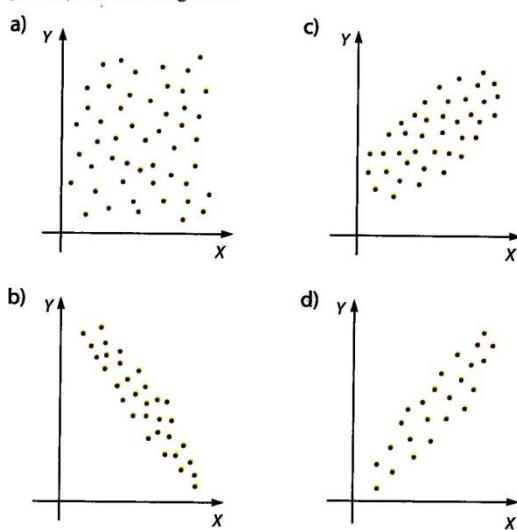


Figura 1. Tarea de interpretación de dependencia a partir del diagrama de dispersión ([T7], p.316)

Suponemos que el concepto de independencia es considerado sencillo, quizá motivado por su fácil visualización en el diagrama de dispersión. Aun así, nuestra opinión es que los libros debieran definir con precisión este concepto, pues es la base de muchos temas estadísticos posteriores. Por ejemplo, en inferencia, un supuesto básico de aplicación de la mayor parte de contrastes estadísticos es admitir la independencia estadística de los datos de la muestra.

*C14. Correlación.* Puesto que puede haber diferentes intensidades de dependencia estadística, se crea la necesidad de medir la relación entre las variables de estudio. Este concepto se define en todos los textos, menos en [H7] y [T7], que, sin definirlo, plantean tareas de cálculo del coeficiente de correlación: “*Estudia la correlación entre estas variables, utilizando la calculadora para realizar las operaciones. [...]*” ([H7], p. 253). Como se ha indicado, es común identificar correlación y dependencia estadística, sin requerir que las variables sean numéricas. Esta situación se presenta en cuatro de los textos que analiza Sánchez Cobo (1999), aunque nuestro caso es mucho más frecuente.

Destacamos los textos [H3], [H5], [H8], [T3] y [T5], que incluyen la definición de correlación espuria o casual (Barbancho, 1973), aunque apenas la usan posteriormente; [H8] incluye esta definición al final del tema, en una sección denominada “correlaciones engañosas”. Además, [H3], [T3] y [H8] presentan la correlación espuria incorrectamente, como complementaria a la relación de causalidad, como se muestra en el siguiente ejemplo en el que [H8], a pesar de considerar otras relaciones de dependencia como la dependencia causal unilateral y la dependencia indirecta, indica la siguiente propiedad:

Hoy en día se encuentran muchos ejemplos de variables entre las que hay una fuerte correlación sin que exista una relación causa-efecto entre ellas. Cuando esto ocurre nos encontramos frente a una correlación espuria ([H8], p.265).

Por otro lado, los ejemplos que proponen estos textos tratan de relaciones de dependencia indirecta, que las denominan espurias, como el siguiente:

Son muchos los casos en los que dos caracteres varían a la vez, sin que por ello estén correlacionados. Por ejemplo: las canas y la miopía de las personas. Es posible que entre los canosos haya más miopes, pero no por ser canosos sino por ser mayores. Este tipo de falsas correlaciones se llaman espurias ([H5], p. 252).

*C15. Covarianza.* La covarianza proporciona una primera medida del tipo de

dependencia entre variables cuantitativas y se define formalmente en todos los textos analizados, al igual que en la mayoría de los de Sánchez Cobo (1999). Es muy utilizada en los textos en el cálculo del coeficiente de correlación lineal; principalmente desde listado de datos y sólo en casos aislados mediante el uso de tablas de doble entrada. Ello explica la poca precisión en la notación asociada a su fórmula, que, en general, utiliza un sumatorio sencillo. Por lo general, la definición se acompaña de ejemplos.

Para introducir el concepto los textos [H1], [H2], [T1] y [T2] utilizan un recurso didáctico propuesto en Holmes (2001) (Figura 2). Consiste en la división en cuatro cuadrantes de la nube de puntos, utilizando rectas paralelas a los ejes que pasan por el centro de gravedad formado por las medias de cada variable. Esta presentación permite al estudiante desarrollar una comprensión más significativa de la misma. [H6], [H8] y [T6] ofrecen la misma interpretación, aunque no presentan la gráfica.

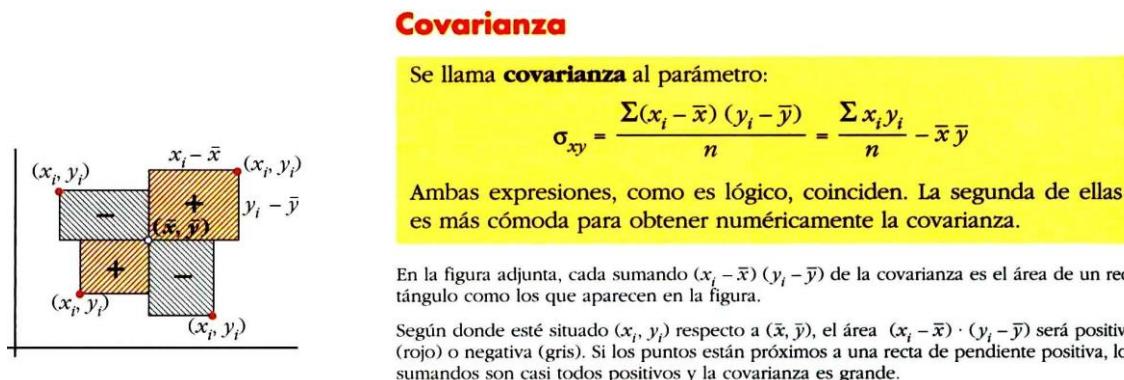


Figura 2. Definición e interpretación de la covarianza ([H1], p. 228)

*C16. Coeficiente de correlación de Pearson.* Puesto que puede haber diferentes intensidades de dependencia estadística, donde la covarianza en sí misma no facilita la comparabilidad entre diferentes relaciones, se crea la necesidad de definición del coeficiente de correlación. El coeficiente de correlación de Pearson informa de la intensidad y el sentido de la dependencia lineal de las variables, y además es adimensional. Es definido en todos los textos, generalmente desde el punto de vista operacional y acompañado de ejemplos, que en su mayoría, implícitamente, ofrecen una definición estructural del mismo. Así, en [H4] se indica (p. 225):

Con la observación de la nube de puntos correspondiente a una distribución (X,Y) podemos identificar la correlación existente entre las variables X e Y. Además, podremos cuantificar esta correlación utilizando coeficientes; en particular, para cuantificar la correlación de tipo lineal usaremos el coeficiente de correlación lineal de Pearson, cuya expresión es:

$$r = \frac{S_{xy}}{S_x S_y}$$

Las definiciones suelen ser correctas; sólo los textos [H1], [H2], [T1] y [T2] pudieran inducir confusión en el estudiante, pues no se indica que este coeficiente pudiera no ser apropiado en otro tipo de dependencia; por ejemplo, curvilínea. Este hecho se resalta en [H5] y [T5], que completa su definición, añadiendo:

Este coeficiente [coeficiente de correlación lineal] mide exclusivamente la correlación lineal entre variables. Por tanto, puede haber otro tipo de correlación no detectada por  $r$ . Por ejemplo,  $r$  no detectaría la correlación exponencial perfecta que hay entre los puntos  $(-1, 0.5)$ ,  $(0,1)$ ,  $(1,2)$ , y  $(4,16)$ , que pertenecen todos a la gráfica de  $y = 2^x$  ([H5], p.255).

#### 4.1.2. Síntesis de la presentación de conceptos ligados a la dependencia

Prácticamente todos los conceptos considerados en este apartado se contemplan en la mayoría de los libros (Tabla 1) y los textos de la misma editorial, dirigidos a las dos modalidades de Bachillerato, introducen los mismos conceptos e incluso los presentan del mismo modo. Por tanto, no se tiene muy en cuenta las recomendaciones de las orientaciones curriculares de disminuir la formalidad de los temas en el Bachillerato de Humanidades y Ciencias Sociales.

Tabla 1. Forma de introducción de los conceptos ligados a la dependencia

Modalidad	Conceptos	H1	H2	H3	H4	H5	H6	H7	H8
Humanidades y CC Sociales	C11. Dependencia funcional	ES	ES	ES	SE	SE	SE	SE	SE
	C12. Dependencia estadística o aleatoria		ES		ES	SE	SE	SE	SE
	C13. Independencia	ES	ES	ES	SE	SE	SE	SE	
	C14. Correlación	ES	SE	ES	SE	SE	SE	SE	
	C15. Covarianza	OSE	OSE	OE	OE	OSE	OSE	OE	OSE
	C16. Coeficiente de correlación lineal	OSE	OSE	OSE	OSE	SOE	OSE	OSE	OSE
		T1	T2	T3	T4	T5	T6	T7	T8
Ciencia y Tecnología	C11. Dependencia funcional	ES	ES	ES	SE	SE	SE	SE	SE
	C12. Dependencia estadística o aleatoria		ES		ES	SE	SE	SE	SE
	C13. Independencia	ES	ES	ES	SE	SE	SE	SE	
	C14. Correlación	ES	SE	ES	SE	SE	SE	SE	
	C15. Covarianza	OSE	OSE	OE	OE	OSE	OSE	OE	OE
	C16. Coeficiente de correlación lineal	OSE	OSE	OSE	OSE	SOE	OSE	OSE	OSE

E=mediante ejemplos; O=definición operacional; S=definición estructural

Prácticamente se presentan todos los conceptos que hemos identificado y se acompañan siempre con ejemplos. La definición suele ser estructural, es decir, indicando las propiedades características del concepto; se completa con una definición operacional para la covarianza y coeficiente de correlación; tan sólo en [H5] y [T5] se introduce en primer lugar la definición operacional. Esta presentación difiere de la encontrada por Sánchez Cobo (1999), donde los libros que él analizó presentaban

generalmente sólo la definición operacional de los conceptos.

Otra diferencia es que, para el resto de los conceptos [H1], [H2], [H3], [T1], [T2] y [T3] presentan los ejemplos antes de la definición, lo que indica una orientación más constructivista del texto. El resto de textos lo hace al contrario; es decir muestran mayor orientación teoría-práctica. En el estudio de Sánchez Cobo (1999) sólo dos libros presentan las definiciones antes de dar ejemplos de los conceptos, según recomienda Skemp (1983) y Joyce, Weil y Calhoun (1986).

En la Tabla 2 podemos observar que estos conceptos son muy utilizados a lo largo del tema principalmente la covarianza, correlación y coeficiente de correlación. Las relativas a la dependencia y sus tipos tienen poco uso en algunos textos; se ven sustituidas en gran medida por el de correlación, planteando tareas al estudiante en las que se utilizan, aunque no aparezcan como tales en sus enunciados, como es el caso de los textos [H6], [T6], [T8] y [H8]. Por ello, el profesor deberá estar atento a la adquisición de la idea de dependencia y sus tipos por parte del estudiante, junto a su uso apropiado y lenguaje preciso, ya que el libro de texto se centra más en la medida de la intensidad de la dependencia que en su interpretación, como base al estudio posterior de la regresión.

Tabla 2. Uso en el tema de los conceptos ligados a la dependencia

Modalidad	Conceptos	H1	H2	H3	H4	H5	H6	H7	H8
Humanidades y CC Sociales	C11.Dependencia funcional	S	P	S	P	P	S	S	P
	C12.Dependencia estadística o aleatoria	S		S	P	P	P	S	P
	C13.Independencia	P	P	S	P		P		P
	C14.Correlación	S	S	S	P	S	S		S
	C15.Covarianza	P	S	P	S	S	P	S	S
	C16.Coeficiente de correlación lineal	S	S	S	S	S	S	S	S
		T1	T2	T3	T4	T5	T6	T7	T8
Ciencia y Tecnología	C11.Dependencia funcional	S	P	S	P	P	S	S	P
	C12.Dependencia estadística o aleatoria	S		S	P	P	P	S	P
	C13.Independencia	P	P	S	P		P		P
	C14.Correlación	S	S	S	P	S	S		S
	C15.Covarianza	P	S	P	S	S	P	S	S
	C16.Coeficiente de correlación lineal	S	S	S	S	S	S	S	S

S=se usa a lo largo de todo el tema; P=uso esporádico.

#### 4.2. CONCEPTOS LIGADOS AL ANÁLISIS DE LA REGRESIÓN

Una vez aceptada la dependencia entre variables, el siguiente problema es buscar un modelo que permita la predicción de la variable dependiente, conocido un valor de la independiente. A continuación se analizan los conceptos asociados, asignándoles un código que se utiliza posteriormente en las diferentes tablas de resultados.

#### 4.2.1. Características de las definiciones

C21. *Variable dependiente e independiente.* La primera idea importante en regresión es la diferencia entre variable dependiente e independiente ya que, una vez aceptada la dependencia entre las variables, se necesita seleccionar qué variable servirá como respuesta o predictora para expresar, en forma de ecuación, una variable en función de otra. Estos conceptos se suelen utilizar implícitamente en los textos al exponer que existen dos rectas de regresión y advertir de su uso adecuado, al igual que ocurre en el estudio de Sánchez Cobo (1999). Un ejemplo se muestra a continuación ([H4], p.226):

Si  $X$  se considera la variable independiente e  $Y$  la variable dependiente, la ecuación de la recta de regresión es:  $y - \bar{y} = \frac{S_{xy}}{S_x^2} (x - \bar{x})$ . Esta recta se denomina *recta de regresión de Y sobre X*. A partir

de ella, conocidos los valores de  $X$ , y sustituyéndolos en la ecuación, se pueden calcular con *una cierta aproximación* los valores de  $Y$ .

Si se considera  $Y$  como variable independiente y  $X$  como variable dependiente, se obtiene la *recta de regresión de X sobre Y*, cuya ecuación es:  $x - \bar{x} = \frac{S_{xy}}{S_y^2} (y - \bar{y})$ . Igual que en el caso

anterior, conocidos los valores de  $Y$ , y sustituyéndolos en la ecuación de la recta, se obtienen con cierta aproximación los valores de  $X$ .

La definición explícita se encuentra sólo en [H5], [T5] y en [H8] del siguiente modo ([H8], p.254): *La variable dependiente es aquella que se quiere estimar, y la variable que se utiliza para ello se denomina variable independiente.*

C22. *Regresión.* Como señala Moore (2005), la regresión es un método general para modelizar las relaciones entre variables. Al contrario que en el estudio de Sánchez Cobo (1999), hay un escaso tratamiento de la definición general de regresión, lo que nos lleva a pensar que la enseñanza se centra más en la regresión lineal que en el propio concepto general de regresión. Cuando se incluye suele ser de modo implícito; sólo [H3], [H4], [T3] y [T4] introducen una definición explícita como por ejemplo: “*Al análisis que pretende determinar la curva que mejor aproxima un diagrama de dispersión se le llama regresión.*” ([H3], p. 226). En [H4] y [T4] se resalta su utilidad predictiva ([T4], p. 340):

La regresión consiste en tratar de encontrar una función matemática que relacione las variables  $X$  e  $Y$  de una distribución bidimensional, de forma que, si se conoce el valor de una variable, se puede calcular el correspondiente de la otra, con mayor o menor aproximación.

En todos los textos donde se define la regresión se define también regresión lineal, a veces en forma imprecisa; otros como [H5], [H6], [H8] y [T5], [T6], [T8], aluden a la

regresión lineal sin definirla.

*C23. Modelos no lineales de regresión.* Los textos analizados raramente definen modelos no lineales de ajuste a los datos, limitando así el significado de la regresión. La proporción de textos que la incluyen es parecida a la encontrada por Lavalle, Micheli y Rubio (2006), donde sólo dos de los siete textos que analizan la incluyen. En nuestro estudio la hemos encontrado en seis de los dieciséis textos analizados ([H4] [H5], [H6] y [T4] [T5], [T6]). Un ejemplo sería el siguiente:

Aunque aquí nos limitaremos a la regresión lineal, existen otros modelos de regresión para ajustar otros tipos de correlación. Por ejemplo, regresión cuadrática si la nube de puntos tiene forma parabólica, o regresión exponencial cuando los puntos se ajustan a funciones exponenciales. ([T5], p. 364)

En otros textos su definición queda implícita, como en el siguiente ejemplo:

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. Esta puede ser una recta, una parábola, una exponencial, una cúbica, etc. A partir de aquí nos centraremos en el estudio de la regresión lineal ([H4], p.226).

Otros, a pesar de no definirla, plantean tareas donde el modelo de ajuste más indicado sea curvilíneo. Los casos más extremos son [H2] y [T2], que no definen los modelos no lineales de regresión y proponen un 37% de las tareas cuyos datos se ajustan a un modelo polinómico de segundo o tercer grado. No obstante, la situación es mejor que en la investigación de Sánchez Cobo (1999), donde sólo en cinco de los 11 libros analizados hubo ejemplos de regresión no lineal.

#### *C24. Rectas de regresión.*

Todos los textos analizados definen las rectas de regresión, aunque [T8] lo hace en forma imprecisa: “*La recta que mejor se aproxima a la nube de puntos se llama recta de regresión de  $Y$  sobre  $X$ .*” ([T8], p. 324). Como en Sánchez Cobo (1999), el apartado sobre la regresión se conforma principalmente alrededor de la determinación de las dos rectas:  $Y$  sobre  $X$  y  $X$  sobre  $Y$ , según qué variable se considere como dependiente e independiente. Además, todos los textos describen el método de mínimos cuadrados como el procedimiento que permite obtener aquella recta que minimiza los cuadrados de las diferencias entre los datos teóricos y los observados. La definición se usa para motivar los cálculos de la pendiente y ordenada en el origen.

A pesar de diferenciar a nivel teórico las dos rectas de regresión, como se indicó anteriormente, [H1], [H2], [H5], [H6], [H7], y [T1], [T2], [T5], [T6], [T7] presentan

tareas donde no se indica cuál es la variable dependiente e independiente en el modelo:

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).

El tratamiento de datos atípicos y la recta de regresión respecto a la mediana sólo se incluyen en [H8] y [T8]. En cuanto a la utilidad predictiva de la recta de regresión, sólo la mitad de los textos analizados ([H1], [H2], [H4], [H6] y [T1], [T2], [T4], [T6]) la resaltan. Estos resultados coinciden con los de Sánchez Cobo (1999), quien señala que se suele resaltar más el cálculo y ajuste de la recta de regresión a los datos (siete de los once textos en su estudio) que su carácter predictivo (tres de los once textos), siendo ésta última su mayor utilidad.

*C25. Coeficientes de regresión.* Para determinar la recta específica de regresión que se ajusta a unos datos, deben calcularse sus parámetros. La pendiente de la recta se denomina coeficiente de regresión, y es posible calcular dos diferentes, dependiendo de qué variable se considere como dependiente. Sólo [H1], [H2], [H4] y [T1], [T2], [T4] definen los coeficientes de regresión. La definición es parcialmente correcta en [H1] y [T1] pues sólo se considera el coeficiente de regresión de  $Y$  sobre  $X$ , lo que puede llevar a pensar al alumno que la misma recta se puede utilizar para predecir  $X$  conocido  $Y$ . Un ejemplo en [H1] (p. 230) es el siguiente:

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.

*C26. Coeficiente de determinación. Bondad de ajuste.* Una vez determinadas las rectas de regresión, se utiliza el coeficiente de determinación (cuadrado del coeficiente de correlación de Pearson) para valorar la bondad del ajuste. Si este coeficiente toma el valor 1 el ajuste es perfecto, siendo peor conforme se acerca a 0. Este coeficiente también puede interpretarse como la proporción de varianza de la variable dependiente que es explicada por el modelo de regresión.

Este coeficiente sólo se define en [H5], [H6], [T5], [T6] y [T8], mientras que el resto de textos tratan este concepto implícitamente, cuando valoran las estimaciones que se realizan con la recta de regresión. En [H6] y [T6] se define de manera operacional la bondad de ajuste, dando la fórmula de este coeficiente, sin indicar cómo se denomina a

esta expresión al definirlo:

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).

Obviamente, la fiabilidad del modelo influye en la fiabilidad de la predicción, pero también influyen otros aspectos, como por ejemplo, que el valor utilizado en la predicción se encuentre en el intervalo de valores utilizados para el modelo y su proximidad a la media.

#### 4.2.2 Síntesis de la presentación de la regresión en el tema

Vemos en la Tabla 3 que todos los textos definen las rectas de regresión y además incluyen tanto la definición estructural como la operacional, aunque no siempre en el mismo orden. Hay predominio de las definiciones acompañadas de ejemplos, siguiendo entonces las recomendaciones de Skemp (1993) y Joyce, Weil y Calhoun, (1986), aunque también varía el lugar donde se introducen ya que, en la mitad de los textos se incluyen los ejemplos antes de indicar la definición y en el resto posteriores a la definición. Los ejemplos se suelen incluir utilizando diagramas de dispersión que añaden la recta de mínimos cuadrados y mostrando la distancia de los puntos de la nube a la recta, la tendencia de la variación conjunta y la diferencia entre el valor observado y el predicho. Esto mismo ocurre en la investigación de Sánchez Cobo (1999).

Tabla 3. Forma de introducción de los conceptos ligados a la regresión

Modalidad	Conceptos	H1	H2	H3	H4	H5	H6	H7	H8
Humanidades y CC Sociales	C21.Variable dependiente e independiente C22.Regresión C23.Modelos no lineales de regresión C24.Coeficientes de regresión C25.Rectas de regresión C26.Bondad de ajuste. Coef. de determinación					S ES SOE O ESO	SE SOE O SO SOE SOE		S SOE SO
	Conceptos	T1	T2	T3	T4	T5	T6	T7	T8
Ciencia y Tecnología	C21.Variable dependiente e independiente C22.Regresión C23.Modelos no lineales de regresión C24.Coeficientes de regresión C25.Rectas de regresión C26.Bondad de ajuste. Coef. de determinación					ES O ESO	SE SO SOE	S SOE SOE	
									OE SOE SO

E=mediante ejemplos; O=definición operacional; S=definición estructural

El resto de conceptos analizados aparece tan sólo en textos aislados; así

encontramos pocos textos que definen la diferencia entre variable dependiente e independiente y sólo en forma estructural, sin mostrar ejemplos. El concepto de regresión es definido sólo en forma estructural y por pocos textos. Igualmente los de modelos no lineales, coeficiente de regresión y coeficiente de determinación son escasamente definidos, aunque se definen tanto estructuralmente como operacionalmente.

En cuanto a las modalidades de bachillerato, encontramos diferencias sólo en los textos [H8] y [T8], ya que en este último no se incluye la definición de variable dependiente e independiente, aunque sí la del coeficiente de determinación.

Tabla 4. Uso en el tema de los conceptos ligados a la regresión

Modalidad	Conceptos	H1	H2	H3	H4	H5	H6	H7	H8
Humanidades y CC Sociales	C21.Variable dependiente e independiente					N			N
	C22.Regresión			N	N				
	C23.Modelos no lineales de regresión					S	P		
	C24.Coeficientes de regresión	P	N		N				
	C25.Rectas de regresión	S	S	S	S	S	S	S	S
	C26.Bondad de ajuste. Coef. de determinación				S	S			
Conceptos		T1	T2	T3	T4	T5	T6	T7	T8
Ciencia y Tecnología	C21.Variable dependiente e independiente					N			
	C22.Regresión			N	N				
	C23.Modelos no lineales de regresión					S	P		
	C24.Coeficientes de regresión	P	N		N				
	C25.Rectas de regresión	S	S	S	S	S	S	S	S
	C26.Bondad de ajuste. Coef. de determinación				S	S			S

S=se usa a lo largo de todo el tema; P=uso esporádico; N=no se usa

Al analizar el uso de estos conceptos, podemos observar (Tabla 4) cómo la definición que más se usa es la de recta de regresión; así es que, en general, otros conceptos tienen un uso muy pobre. Este hecho es debido a que algunos textos plantean la regresión desde un enfoque algorítmico, donde el estudiante se limita a calcular las rectas de regresión (principalmente la de  $Y$  sobre  $X$ ) y hacer predicciones a través de ellas, pero no se incluyen actividades interpretativas ni de reflexión sobre los conceptos o sus propiedades tal y como se recomienda en las orientaciones internacionales para el desarrollo de la competencia matemática en el proyecto PISA/OECD (OECD, 2004). Más aún, algunos textos plantean tareas donde el diagrama de dispersión muestra una clara tendencia curvilínea de los datos, y se pide realizar un análisis de regresión lineal.

## 5. CONCLUSIONES E IMPLICACIONES PARA LA ENSEÑANZA

El análisis realizado permite cumplir el objetivo de este trabajo, que fue caracterizar las definiciones de los conceptos ligados a la regresión y correlación en una muestra

amplia de libros de texto españoles de Bachillerato. Para ello se han analizado el tipo de definición y su uso a lo largo del texto, comparando entre textos de la misma modalidad de Bachillerato según diferente editorial y entre las dos modalidades en una misma editorial.

Estos resultados completan los estudios previos ya que, además de actualizar el estudio de Sánchez Cobo (1999) y Lavalle, Micheli y Rubio (2006), un punto original en nuestro trabajo es la clasificación de las definiciones los conceptos siguiendo la propuesta de Sfard (1991) (estructural u operacional). Igualmente, se han completado nuestros trabajos previos sobre las variables incluidas en las situaciones problema de correlación y regresión (Gea, Batanero, Cañadas y Contreras, 2013) y el lenguaje de la correlación y regresión (Gea, Batanero, Arteaga, Cañadas y Contreras, 2014).

Una primera conclusión del estudio es la escasa presencia de definiciones relativas a la regresión, exceptuada la de la recta de regresión, en comparación con las que se incluyen referidas a la dependencia entre las variables. Esta falta de atención a la regresión creemos que se debe a la amplitud que en el tema se dedica al concepto de correlación, con lo que la enseñanza de estos tópicos queda descompensada. Como se ha indicado, la principal finalidad del estudio de la dependencia es ajustar posteriormente un modelo que sirva para la predicción. Recomendamos consecuentemente, incluir las definiciones de regresión, modelo de regresión y coeficientes de regresión, que sólo son incluidos en algunos libros.

Los conceptos incluidos en los textos se suelen definir acompañados de ejemplos, aunque no siempre se presentan los ejemplos antes de la definición, como recomienda Skemp (1993) y Joyce, Weil y Calhoun (1986). Sería también admisible la presentación formal seguida por ejemplos, que aparece en otros textos, aunque es preferible lo contrario.

En los casos en que hay una definición explícita de los conceptos, observamos un predominio de la presentación operacional (Sfard, 1991). En este sentido, recomendamos, cuando sea posible, dar relevancia a la definición estructural de los conceptos pues encontramos, en los conceptos que se definen, un énfasis en los aspectos operacionales, sin prestar mucha atención al significado y aspectos interpretativos. En este sentido, recordamos que para conseguir una comprensión conceptual adecuada de cualquier tópico se deben incluir los principios del dominio dado y sus interrelaciones (Rittle-Johnson, Siegler y Alibali, 2001).

Autores como Ouvrier-Buffet (2004) han mostrado las dificultades que tienen los

estudiantes para realizar definiciones o comprender definiciones de los conceptos y, por lo tanto, es importante que los libros de texto presenten definiciones asequibles y completas. Por ello los autores de los libros de texto debieran recoger las recomendaciones sugeridas por los autores que hemos descrito en el marco teórico y de otros que se han interesado por el proceso de definición y se sintetizan en Escudero, Gavilán y Sánchez-Matamoros (2014).

Otro resultado del estudio, aunque dirigidos a alumnos con diferente preparación y motivación, es que no se han observado grandes diferencias en los libros de texto de las dos modalidades de Bachillerato. De hecho, la misma editorial suele hacer una presentación muy parecida o idéntica del tema a ambos tipos de alumnos; e incluso se incluyen menos definiciones en los textos de la modalidad de Bachillerato en Ciencias y Tecnología, a pesar que para estos alumnos las directrices curriculares oficiales (MEC, 2007; MECD, 2015) sugieren presentar el estudio de las matemáticas con mayor grado de formalidad que en el Bachillerato de Humanidades y Ciencias Sociales.

Para finalizar, hacemos notar que 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 EDU2013-41141-P (MEC) y grupo FQM126 (Junta de Andalucía).

## ANEXO

Tabla A1. Libros de texto utilizados en el análisis

Código	Referencia
H1	Colera, J., Oliveira, M. J., García, R. y Santaella, E. (2008). <i>Matemáticas aplicadas a las Ciencias Sociales I</i> . Madrid: Anaya.
H2	Arias, J. M. y Maza, I. (2011). <i>Matemáticas aplicadas a las Ciencias Sociales I</i> . Madrid: Bruño.
H3	Anguera, J., Biosca, A., Espinet, M. J., Fandos, M.J., Gimeno, M. y Rey, J. (2008). <i>Matemáticas I aplicadas a las Ciencias Sociales</i> . Barcelona: Guadriel.
H4	Monteagudo, M. F. y Paz, J. (2008). 1º Bachillerato. <i>Matemáticas aplicadas a las Ciencias Sociales</i> . Zaragoza: Edelvives.
H5	Martínez, J. M., Cuadra, R., Heras, A. (2008). <i>Matemáticas aplicadas a las Ciencias Sociales. 1.º Bachillerato</i> . Madrid: McGraw-Hill.
H6	Bescós, E. y Pena, Z. (2008). <i>Matemáticas aplicadas a las Ciencias Sociales</i> . Vizcaya: Oxford University Press.
H7	Antonio, M., González, L., Lorenzo, J., Molano, A., del Río, J., Santos, D. y de Vicente, M. (2009). <i>Matemáticas aplicadas a las Ciencias Sociales I</i> . Madrid: Santillana.
H8	Vizmanos, J. R., Hernández, J., Alcaide, F., Moreno, M. y Serrano, E. (2008). <i>Matemáticas aplicadas a las Ciencias Sociales I</i> . Madrid: SM.
T1	Colera, J., Oliveira, M. J., García, R. y Santaella, E. (2008). <i>Matemáticas I</i> . Madrid: Anaya.
T2	Arias, J. M. y Maza, I. (2011). <i>Matemáticas I</i> . Madrid: Bruño.

- 
- T3 Biosca, A., Doménech, M., Espinet, M. J., Fandos, M. J. y Jimeno, M. (2008). *Matemáticas I*. Barcelona: Guadiel.
- T4 Monteagudo, M. F. y Paz, J. (2008). *1º Bachillerato. Matemáticas. Ciencias y Tecnología*. Zaragoza: Edelvives.
- T5 Martínez, J. M., Cuadra, R., Barrado, F. J. (2007). *Matemáticas 1º Bachillerato*. Madrid: McGraw-Hill.
- T6 Bescós, E. y Pena, Z. (2009). *Matemáticas. I Bachillerato*. Navarra: Oxford University Press.
- T7 Antonio, M., González, L., Lorenzo, J., Molano, A., del Río, J., Santos, D. y de Vicente, M. (2008). *Matemáticas I. I Bachillerato*. Madrid: Santillana.
- T8 Vizmanos, J. R., Hernández, J., Alcaide, F., Moreno, M. y Serrano, E. (2008). *Matemáticas I*. Madrid: SM.
- 

## Referencias

- Barbancho, A. G. (1973). *Estadística elemental moderna*. Barcelona: Ariel.
- Cañas, G. (2012). *Comprensión intuitiva y aprendizaje formal de las tablas de contingencia en alumnos de psicología*. Tesis doctoral. Universidad de Granada.
- 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.
- Cordero, F. y Flores, R. (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*, 10(1), 7-38.
- Crocker, J. (1981). Judgment of covariation by social perceivers. *Psychological Bulletin* 90(2), 272-292.
- Escudero, I., Gavilán, J.M. y Sánchez-Matamoros, G. (2014). Una aproximación a los cambios en el discurso matemático generados en el proceso de definir. *Revista Latinoamericana de Investigación en Matemática Educativa*, 17(1), 7-32.
- Estepa, A. (2007). Caracterización del significado de la correlación y regresión de estudiantes de Educación Secundaria. *Zetetiké*, 15(28), 119-151.
- Estepa, A., Gea, M. M., Cañas, G. R. y Contreras, J. M. (2012). Algunas notas históricas sobre la correlación y regresión y su uso en el aula. *Números*, 81, 5-14.
- García-Alonso, I. (2011). Análisis de los términos de inferencia estadística en Bachillerato. *Números. Revista de Didáctica de las Matemáticas*, 77, 51-73.

- 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.
- Gea, M. M., Batanero, C., Arteaga, P., Cañadas, G. R., Contreras, J. M. (2014). Análisis del lenguaje sobre la correlación y regresión en libros de texto de bachillerato. *SUMA*, 76, 37-45
- Godino, J. D. (2003). *Teoría de las funciones semióticas. Un enfoque ontológico-semiótico de la cognición e instrucción matemática*. Granada: Departamento de Didáctica de la Matemática.
- Herbel, 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.
- Holmes, P. (2001). Correlation: From picture to formula. *Teaching Statistics*, 23(3), 67-71.
- Joyce, B. R., Weil, M. y Calhoun, E. (1986). *Models of teaching*. Englewood Cliffs, NJ: Prentice-Hall.
- Lavalle, A. L., Micheli, E. B. y Rubio, N. (2006). Análisis didáctico de regresión y correlación para la enseñanza media. *Revista Latinoamericana de Matemática Educativa*, 9(3), 383-406.
- Leikin, R. y Winicky-Landman, G. (2001). Defining as a vehicle for professional development of secondary school mathematics teachers. *Mathematics Teacher Education and Development*, 3, 62–73.
- 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.
- Mariotti M. A. y Fischbein, E. (1997). Defining in classroom activities. *Educational Studies in Mathematics*, 34, 219-248.
- MEC (2007). *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 26 de diciembre, por el que se establece el currículo básico de la Educación Secundaria Obligatoria y del Bachillerato*. Madrid: Autor.

Moore, D. S. (2005). *Estadística aplicada básica*. Barcelona: Antoni Bosch.

Moritz, J. (2004). Reasoning about covariation. En D. Ben-Zvi y J. Garfield (Eds.), *The challenge of developing statistical literacy, reasoning and thinking* (pp. 221-255). Dordrecht: Kluwer.

OECD (2004). *Learning for tomorrow's world: First results from PISA 2003*. Paris: OECD.

Ouvrier-Buffet, C. (2004). Construction of Mathematical Definitions: An epistemological and didactical study. In M. J. Hoines & A.B. Fuglestad (Eds.), *Proceedings of the 28 th Conference of the International Group for the Psychology of Mathematics Education* (Vol. 3, pp. 473-480).

Rittle-Johnson, B., y Alibali, M. W. (1999). Conceptual and procedural knowledge of mathematics: Does one lead to the other? *Journal of educational psychology*, 91, 175-189.

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.

Sánchez Cobo, F. T. (1999). *Significado de la correlación y regresión para los estudiantes universitarios*. Tesis doctoral. Universidad de Granada.

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

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.

Skemp, R. R. (1993). *Psicología del aprendizaje de las matemáticas* (Vol. 15). Ediciones Morata.

Vergnaud, G. (1982). Cognitive and developmental psychology and research in

mathematics education: some theoretical and methodological issues. *For the learning of Mathematics*, 3(2), 31-41.

Vinner, S. (1991). The role of definitions in the teaching and learning of mathematics. En: Tall, D. O. (Ed.), *Advanced mathematical thinking* (pp. 65–81). Dordrecht: Kluwer.

Zazkis, R. y Leikin, R. (2008) Exemplifying definitions: a case of a square. *Educational Studies in Mathematics*, 69, 131–148.