ISSN 2177-9309
Revista de Educação Matemática e Tecnológica Iberoamericana
Universidade Federal de Pernambuco
Pós-Graduação em Educação Matemática e Tecnológica
EDUMATEC

JUSTIFICACIONES EN EL TEMA DE CORRELACIÓN Y REGRESIÓN EN TEXTOS ESPAÑOLES DE BACHILLERATO

María M. Gea, Carmen Batanero, Pedro Arteaga y Gustavo R. Cañadas
Profesores de Didáctica de la Matemática, Universidad de Granada
mmgea@ugr.es, batanero@ugr.es, parteaga@ugr.es, grcanada@ugr.es

Resumen

Presentamos un estudio de las justificaciones usadas en el tema de correlación y regresión en una muestra de dieciséis libros de texto españoles de Bachillerato, de las dos especialidades en que se incluye el tema. Se observa el uso de diversos tipos de justificación informal, tal como ejemplos y contra ejemplos, o apoyo exclusivo de un gráfico como modo de justificación. La principal diferencia entre editoriales se muestra en la inclusión o no de demostraciones algebraicas deductivas. No se detectan diferencias en los textos de la misma editorial dirigidos a diferente tipo de Bachillerato, lo que implica que no se consideran las orientaciones metodológicas curriculares específicas para cada uno de ellos.

Palabras clave: Correlación y regresión, justificaciones, libros de texto, Bachillerato.

Abstract

In this paper we analyse the arguments used in the topic correlation and regression in a sample of sixteen Spanish high school textbooks in the two specialties that include this topic. We observe the use of different types of informal justifications, such as examples and counter examples or exclusive support of graphs as a way of justification. The main difference between editorials lies in the use or absence of algebraic deductive proofs. We found no differences in the books from the same editorial directed to different high school specialty; this suggests that the methodological curricular orientations are not taken into account.

Keywords: Correlation and regression; justifications, textbooks; High school.

1. INTRODUCCIÓN

La correlación y regresión son conceptos estadísticos fundamentales, pues extienden la idea de dependencia funcional, y se relacionan con muchos otros como como los de variación, distribución, centralización o dispersión. Otra razón que justifica el interés de este tema es que el razonamiento sobre la correlación y regresión se vincula a la toma de decisiones en ambiente de incertidumbre (ESTEPA; GEA; CAÑADAS; CONTRERAS, 2012).

La investigación sobre la comprensión del tema, desarrollada básicamente desde la psicología, muestra que los adultos no emplean reglas matemáticas, sino estrategias intuitivas incorrectas, al estimar la correlación, llegando a pobres resultados. Un ejemplo es el sesgo denominado "correlación ilusoria" por Chapman y Chapman (1969), que consiste en realizar estimaciones sesgadas de la correlación debidos a las expectativas personales sobre las variables en estudio.

La investigación didáctica se ha centrado en la comprensión de algunas propiedades de la correlación y regresión por alumnos universitarios. Así, Estepa y Batanero (1995) describieron casos de estudiantes que no consideran la correlación inversa, tienen un sentido determinista de la correlación o identifican correlación con causalidad (ver también ESTEPA, 2008; ZIEFFLER; GARFIELD, 2009). Sánchez Cobo (1999) analiza el efecto de la intensidad y signo de la correlación sobre la precisión de la estimación de la correlación y describe errores relacionados con esta estimación. La creencia infundada en la transitividad del coeficiente de correlación es descrita por Castro-Sotos et al. (2009).

En España se estudia este tema en el primer curso de Bachillerato (16-17 años) en las modalidades de Ciencias y Tecnología y Humanidades y Ciencias Sociales (MEC, 2007), que, aunque incluyen un contenido similar, sugiere orientaciones metodológicas diferentes, pues mientras el primero de estos Bachilleratos está dirigido a estudiantes futuros de ciencias y matemáticas, los estudiantes que cursan el segundo seguirán carreras de humanidades, como la pedagogía, psicología o sociología.

Así, en relación al uso de justificaciones, en las directrices oficiales el Decreto de Enseñanzas Mínimas de Bachillerato (MEC, 2007) establece como objetivo para el Bachillerato en Ciencias y Tecnología utilizar argumentaciones razonadas y algunos ejemplos de demostraciones rigurosas, indispensables para el avance de la ciencia y la tecnología. Se matiza que:

> Las definiciones formales, las demostraciones (reducción al absurdo, contraejemplos) y los encadenamientos lógicos (implicación, equivalencia) dan validez a las intuiciones y confieren solidez a las técnicas aplicadas. Sin embargo, este es el primer momento en que el alumno se enfrenta con cierta seriedad al lenguaje formal, por lo que el aprendizaje debe ser equilibrado y gradual. El simbolismo no debe desfigurar la esencia de las ideas fundamentales, el proceso de investigación necesario para alcanzarlas, o el rigor de los razonamientos que las sustentan. Deberá valorarse la capacidad para comunicar con eficacia esas ideas aunque sea de manera no formal. Lo importante es que el estudiante encuentre en algunos ejemplos la necesidad de la existencia de este lenguaje para dotar a las definiciones y demostraciones matemáticas de universalidad, independizándolas del lenguaje natural (MEC, 2007, p. 45449)

Por el contrario, se indica el menor peso de la demostración en las matemáticas dirigidas a Humanidades y Ciencias Sociales, resaltando la importancia del apoyo de la tecnología para facilitar el trabajo con las matemáticas en esta especialidad.

> En este contexto, la fuerte abstracción simbólica, el rigor sintáctico y la exigencia probatoria que definen el saber matemático, deben tener en esta materia una relativa presencia. Las fórmulas, una vez que se las ha dotado de significado, adoptan un papel de referencia que facilita la interpretación de los resultados pero, ni su obtención, ni su cálculo y mucho menos su memorización, deben ser objeto de estudio. (Matemáticas aplicadas a las Ciencias Sociales I, MEC, 2007. p. 45474).

La finalidad de este trabajo fue describir las justificaciones que sirven para probar las propiedades de la correlación y regresión o las soluciones de los problemas en los libros de texto de Bachillerato en las dos especialidades mencionadas. Con ello completamos la investigación respecto a la correlación y regresión, que es muy escasa en lo que concierne a los libros de texto, como se verá en los antecedentes. En lo que sigue analizamos los fundamentos, métodos y resultados del estudio.

2. FUNDAMENTOS

2.1. Marco teórico

Nuestro estudio 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).

Son muchos los autores que analizan la importancia de las demostraciones; entre ellos Davis y Hersh (1988) indican que generalmente se tiende a demostrar una propiedad cuando se está convencido de su certeza; por tanto se utiliza para profundizar en el porqué de la proposición o para clarificarla. Igualmente indica que se puede utilizar como ayuda al descubrimiento, pues es posible obtener nuevos resultados durante un proceso de demostración. Por su parte, Krygowska (1971) resalta que la demostración puede utilizarse para sistematizar resultados dentro de un sistema de conceptos fundamentales y teoremas relacionados; por tanto es una ayuda en la axiomatización.

Una característica de los libros de texto de matemáticas es la presencia de justificaciones, que se usan con fines de verificación o convicción para establecer la validad de una propiedad o proposición, convencer de un resultado o método en la resolución de un problema, sistematizar conocimientos, descubrimiento de nuevos resultados o comunicación de los mismos (DE VILLIERS, 1993). Estos argumentos son útiles para validar y hacer comprensibles a los estudiantes los procedimientos, propiedades, definiciones, así como las representaciones que se enlazan en la resolución de problemas.

La demostración es esencial en matemáticas y debiera ser enseñada a los alumnos a partir de la educación secundaria (12-13 años) (CRESPO; ARFÁN, 2005). Su papel en el currículo se destaca en los estándares del NCTM (2000) que indican la necesidad de contemplarlos a lo largo de la escolaridad para desarrollar, con diferentes contextos, la capacidad de argumentación de los estudiantes, incluyendo la elaboración de conjeturas, formulación de contraejemplos o elaboración de pruebas informales, incluyendo el principio de inducción y las demostraciones indirectas.

Siguiendo a Godino y Recio (2001), en este trabajo hemos considerado necesario ampliar la visión de la demostración; teniendo en cuenta, no sólo las demostraciones formales deductivas características de la matemática. Estos autores indican que en los textos de matemáticas las propiedades y teoremas se consideran necesariamente verdaderos y que las argumentaciones que se usan para justificarlos suelen ser informales, no deductivas e incluso basadas en criterios de autoridad:

> Estas diferencias en las situaciones y prácticas argumentativas indican sentidos distintos del concepto de *demostración* –o bien diversos *objetos demostración* según el modelo onto semántico adoptado (GODINO; RECIO, 2016, p. 206).

2.2. Antecedentes

Los libros de texto son un recurso didáctico muy utilizado en la enseñanza y aprendizaje, aportando un apoyo al profesor y al alumno y determinando, en parte, la enseñanza. En base a una revisión de la literatura Ortiz (1999) enumera varios aspectos relacionados con su uso: son una fuente de datos y actividades para el aula; resultan de un gran esfuerzo de planificación y síntesis; se asumen como un conocimiento que hay que transmitir; y el alumno lo considera como una autoridad del conocimiento y guía del aprendizaje.

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, Batanero y Serrano (2001), Cobo y Batanero (2004) y Azcárate y Serradó (2006).

El primer antecedente relacionado con nuestro trabajo es el de Sánchez Cobo (1999), quien analiza la correlación y regresión en once libros de texto españoles de bachillerato publicados entre 1987 y 1990. El autor analizó los contextos, el contenido matemático implicado, el tipo de tarea, y el tipo e intensidad de la dependencia. Muestra una tendencia formal en la presentación del tema, y el uso mayoritario de ejemplos basados en representaciones gráficas, así como un fuerte sesgo en los ejemplos presentados hacia la correlación positiva. Este autor analiza también cuáles propiedades específicas se justifican en los textos estudiados, encontrando que la que principalmente se demuestra es el uso del método de mínimos cuadrados para construir la recta de regresión (9 de 11 textos); por el contrario solo uno o dos textos justifican la fórmula de la covarianza o del coeficiente de regresión.

Más recientemente, Lavalle et al. (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 también se plantean más actividades bajo una asociación directa que inversa.

Nuestra finalidad es completar estos estudios con el estudio de los tipos de justificaciones empleadas, que es punto no tratado por los autores citados.

3. MÉTODO

3.1. Muestra utilizada

En este estudio se analizaron dieciséis libros de texto, todos ellos correspondientes al currículo actual español de Bachillerato en las dos modalidades en que se incluye el tema: ocho libros de la modalidad de *Humanidades y Ciencias Sociales* y otros ocho de la modalidad de *Ciencias y Tecnología*, tomando las mismas editoriales para los dos bachilleratos. Corresponden a editoriales de gran prestigio y tradición en España y la mayoría siguen vigentes en el momento actual, pues no se han reeditado.

Todos los textos son posteriores al actual decreto en vigor en España (MEC, 2007), y se usan actualmente. Presentamos la lista de estos textos en apéndice, junto con un código, que será utilizado a lo largo del trabajo para referirse al mismo. Los códigos se han asignado por orden alfabético según las editoriales, utilizando [H] para los textos de la modalidad de Humanidades y Ciencias Sociales y [T] para la modalidad en Ciencias y Tecnología.

De acuerdo a López Noguero (2002), nuestro análisis se sitúa entre los métodos intensivos, que estudian con detenimiento algunos documentos, en lugar de recurrir a una muestra más amplia, pero analizada someramente (métodos extensivos). Siguiendo a este mismo autor, se trata de un análisis interno de los documentos, procurando destacar su sentido y características fundamentales. Al tratarse de una investigación cualitativa, la muestra es intencional, tratando que incluya los casos más representativos y paradigmáticos para capturar la mayor riqueza posible de la realidad analizada (MARTÍNEZ, 2006). Para asegurar la comparabilidad de nuestros resultados con otros estudios similares (GOETZ; LECOMPTE, 1998) se describe con detalle la terminología y marco analítico utilizado.

3.2.Análisis

La metodología empleada se basa en el análisis de contenido, que asume que un texto puede dividirse en unidades que pueden clasificarse en un número reducido de categorías en función de variables subyacentes, y que permiten realizar inferencias sobre su contenido (KRIPPENDORFF, 1997). Seguimos el mismo método utilizado en la investigación de Cobo y Batanero (2004), el cual consiste en los siguientes pasos:

1. Seleccionados los libros, y el capítulo correspondiente a la correlación y regresión, se efectuaron varias lecturas cuidadosamente, para determinar los párrafos que constituirían la primera unidad de análisis.
2. Mediante un proceso cíclico e inductivo se analizaron las justificaciones utilizada en dichos párrafos, que constituirían nuestras unidades secundarias de análisis.
3. Seguidamente se categorizan estas justificaciones siguiendo la clasificación propuesta por Recio (2002), quien realiza una revisión de los estudios más importantes relacionados con el tema de la demostración.
4. Se elaboran tablas indicando la presencia o ausencia de cada tipo de justificación en los textos, con el fin de resumir los resultados.

A continuación se describen los resultados del análisis; se comienza describiendo las características de los diferentes tipos de justificación encontrados y posteriormente se realiza una síntesis y discusión de la presencia de los diferentes tipos de justificaciones en los textos, comparando los de las dos modalidades de Bachillerato. Se finaliza con algunas implicaciones para la docencia y la investigación.

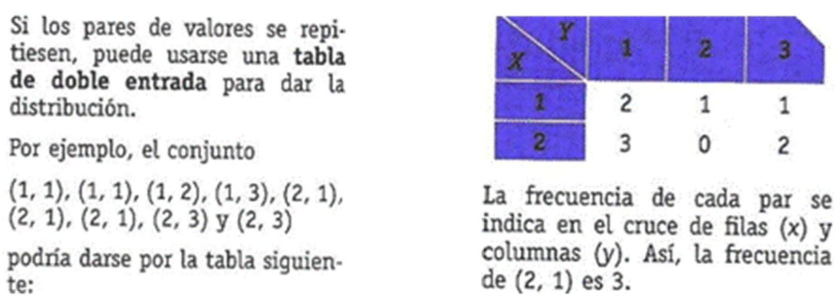
4. RESULTADOS Y DISCUSIÓN

4.1. Características de las justificaciones encontradas en los textos

Como se ha indicado, se ha encontrado una gran variedad de tipos de justificación. A continuación describimos cada una de las categorías encontradas, mostrando algunos ejemplos en los textos que sirvan para clarificarlas.

A1. Uso de ejemplos o contraejemplos como medio de justificación. Es frecuente que los textos se limiten a presentar ejemplos o contraejemplos, como argumento para validar una determinada propiedad, o bien negarla, así como para justificar la solución de un problema. Un posible interés de este tipo de "prueba" es que ayuda a desarrollar el pensamiento inductivo, pero se debería utilizar sólo como primer paso para posteriormente generalizar. Recordemos que la combinación de razonamiento inductivo y deductivo en la clase de matemáticas es recomendada por el NCTM (2000). Además, según Crespo y Farfán (2005) en la mayoría de las ciencias, y en particular en la mayoría de los campos de la matemática, se parte de la inducción como método para enunciar sus proposiciones. Sin embargo son pocos los textos que finalizan este proceso de generalización, limitándose en la mayoría de los casos a dar por "validada" la propiedad o solución, simplemente mostrando ejemplos.

Por ejemplo, varios textos al introducir la representación tabular de los datos bidimensionales usan ejemplos como el mostrado en la Figura 1, donde se explica el significado de una celda de la tabla y el estudiante deberá generalizar el caso particular que se explica, al resto de celdas.


Si los pares de valores se repitiesen, puede usarse una **tabla de doble entrada** para dar la distribución.

Por ejemplo, el conjunto

(1, 1), (1, 1), (1, 2), (1, 3), (2, 1), (2, 1), (2, 1), (2, 3) y (2, 3)

podría darse por la tabla siguiente:

X \ Y	1	2	3
1	2	1	1
2	3	0	2

La frecuencia de cada par se indica en el cruce de filas (*x*) y columnas (*y*). Así, la frecuencia de (2, 1) es 3.

Figura 1. Justificación mediante ejemplo de las ventajas de la representación tabular ([H5], p. 250)

Este mismo tipo de argumentos se encuentran para explicar la construcción de definiciones de las distribuciones marginales en los textos [H7] y [T7]. En los textos [H5] y [T5] al definir la dependencia funcional, como correspondencia que a cada valor de la variable independiente corresponde uno sólo de la dependiente, lo hace a través del siguiente ejemplo:

> Si se dejan caer dos piedras desde 1 metro de altura ambas tardan el mismo tiempo en llegar al suelo. Y también, si otra piedra se deja caer desde 1,5 metros siempre tarda más en llegar al suelo que las anteriores. La ley de la gravitación se cumple siempre. Por tanto, conocida la altura desde la que se deja caer un objeto se puede saber cuánto tardará en llegar al suelo, con una certeza absoluta. ([T5], p. 357).

De igual modo, [H5] y [T5] incluyen un ejemplo, para justificar que el coeficiente de correlación lineal tan sólo cuantifica la dependencia lineal entre las variables:

> 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).

Finalmente, algunos textos justifican la diferencia entre correlación y causalidad mediante ejemplos como el siguiente:

> Se dice que existe correlación espuria entre dos variables estadísticas cuando estas aumentan o disminuyen de manera conjunta sin que exista una relación causa-efecto entre ellas. Por ejemplo, es muy posible que exista una cierta correlación entre el número de restaurantes de una ciudad y el número de profesores que trabajan en ella. Esto se debe a que ambas variables están relacionadas con el número total de habitantes de la ciudad. ([T3], p. 271)

A2. Uso de representaciones gráficas para apoyar una argumentación verbal o simbólica. Un caso particular del uso de ejemplos y contraejemplos es utilizar un gráfico para apoyar la verdad o falsedad de una afirmación o de una propiedad, pues este tipo de argumentaciones son muy intuitivas. Sánchez Cobo (1999) encontró en su estudio que los textos suelen utilizar la representación gráfica como medio de argumentación. Este tipo de prueba se incluiría según Godino y Recio (2001) en las pruebas deductivas informales que incluyen argumentaciones lógicas apoyadas en analogías. Suelen emplearse para el desarrollo de explicaciones y “demostraciones” informales de propiedades relativas a la correlación y regresión.

Encontramos este tipo de argumentaciones con mucha frecuencia en el análisis de la dependencia entre dos variables aleatorias (tipo, intensidad y signo). En este sentido, en [H3] y [T3] se analiza la existencia de dependencia mediante diagramas de dispersión indicando: “*La relación existente entre dos variables queda reflejada en los diagramas de dispersión o nubes de puntos*”. Igualmente en [H4] y [T4], donde se muestran tres ejemplos de diagramas de dispersión (Figura 2) para justificar las definiciones de independencia y dependencia estadística, así como el grado de dependencia entre las variables. Observamos que se pretende que el alumno generalice a partir de estos ejemplos.

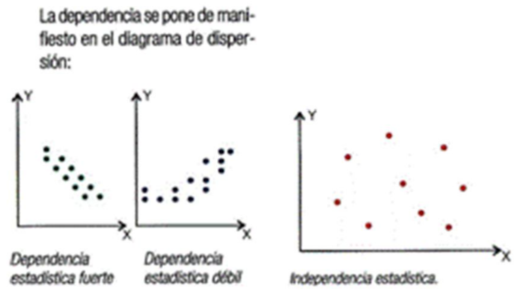


Figura 2. Justificación de la dependencia estadística y el grado de la relación ([T4], p. 338)

También encontramos textos en que se relaciona gráficamente el coeficiente de correlación, las rectas de regresión, y el ángulo que forman (Figura 3); observamos que la única justificación de la relación entre la intensidad de la correlación y el ángulo que forman las dos rectas de regresión es el gráfico presentado.

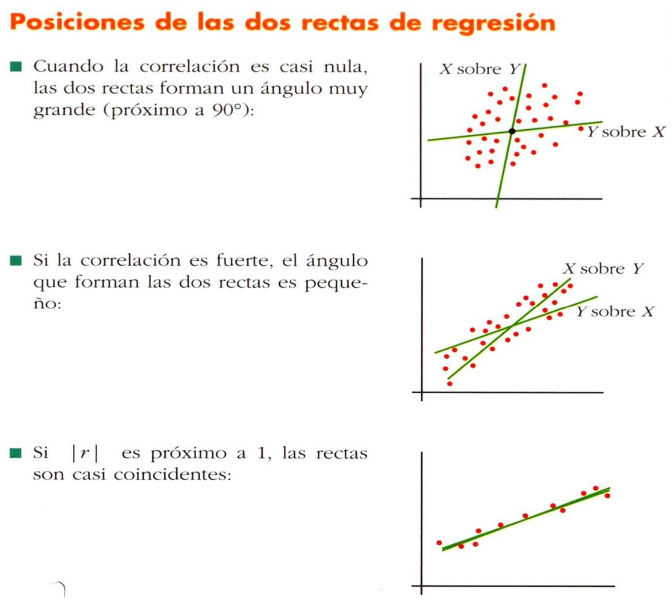


Figura 3. Justificación gráfica de relación entre correlación y ángulo de las rectas de regresión ([H1], p. 232).

Otros textos usan la división del plano en cuatro cuadrantes como medio de apoyo gráfico a sus demostraciones de las propiedades de la covarianza y correlación (Figura 4). Mostrando gráficamente la posición de diferentes puntos del diagrama de dispersión respecto a dos rectas paralelas a los ejes y que pasan por el centro de gravedad de la distribución se suele justificar la relación entre el signo de la covarianza, la forma creciente o decreciente de la nube, y la dispersión de los puntos respecto a la recta de regresión Esta representación fue propuesta en Holmes (2001), y se suele acompañar de de una demostración deductiva informal con poco apoyo algebraico donde se comprueba los signos del producto de las diferencias de cada punto al centro de gravedad, dependiendo de su posición en los mencionados cuadrantes.

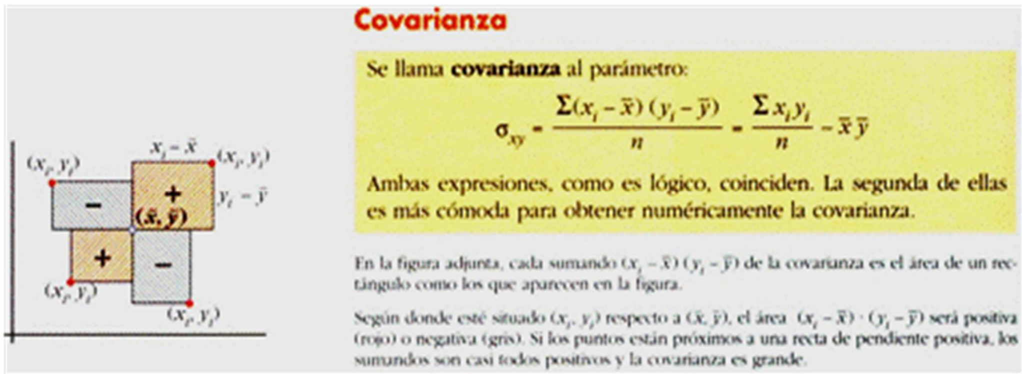


Figura 4. Uso de apoyo gráfico para la interpretación de la covarianza ([H1], p. 228)

En cuanto al tratamiento de regresión, la mayoría de los textos justifican la pertinencia del método de mínimos cuadrados para el ajuste de la recta de regresión en forma gráfica, utilizando un diagrama de dispersión en el que se visualizan las distancias entre los valores reales y los estimados en la recta de regresión. Quizá éste sea uno de los casos más representativos de este tipo de argumento (Figura 5).

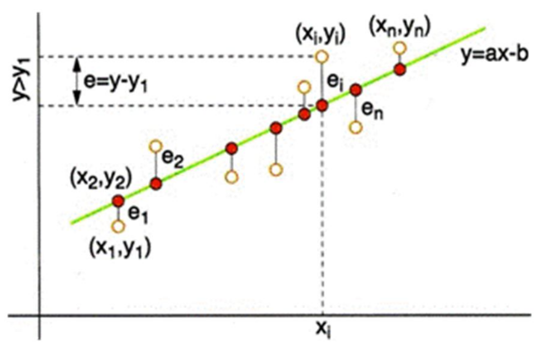


Figura 5. Relación entre el valor estimado y el valor real en el diagrama de dispersión ([T5], p. 364)

A3. Razonamientos verbales deductivos. Para otras propiedades se llega a enunciar o demostrar una propiedad en forma deductiva utilizando propiedades, axiomas o teoremas que el alumno conoce previamente. Se trata, generalmente, de razonamientos no excesivamente formalizados que, de forma deductiva, justifican una propiedad utilizando principalmente el lenguaje verbal, más que el simbólico. Este tipo de argumentación es la más utilizada en los textos analizados, como ocurre en el trabajo de Sánchez Cobo (1999).

Es muy común encontrar este tipo de argumentación para relacionar la covarianza y el coeficiente de correlación. Por ejemplo, en [H7] y [T7] se indica: "*Como* σ_x *y* σ_y *son siempre positivos, el signo del coeficiente de correlación viene determinado por el signo de* σ_{xy}" ([H7], p. 248). Por lo cual se justifica que el signo de correlación y covarianza es el mismo, en forma verbal, sin necesidad de realizar cálculos algebraicos. Otros ejemplos que encontramos de justificación de la misma propiedad son los siguientes, en algunos de los cuáles también se advierte que la covarianza no informa sobre la intensidad de la correlación:

> El signo del coeficiente de correlación y el del coeficiente de regresión coinciden, pero aquí termina la coincidencia: puede ser que la recta de regresión tenga pendiente alta y, sin embargo, el coeficiente de correlación sea bajo. O al contrario. ([H1], p. 230).
>
> La covarianza indica cómo es la correlación entre dos variables; es decir, cómo se orienta la nube de puntos, pero este parámetro no indica de una forma precisa la medida de esa relación. Para resolver este problema, se definen los conceptos de correlación y coeficiente de correlación. ([H2], p. 248).
>
> Como las desviaciones típicas son siempre positivas, el signo de r viene dado por el signo de la covarianza. ([H4], p. 225).

En el caso de la regresión, es habitual encontrar este tipo de argumentaciones para justificar la utilidad del método de mínimos cuadrados. Por ejemplo, en el texto [H3] (p. 226) se menciona la necesidad de utilizar un método alejado del subjetivismo de un trazado por ajuste manual. En algunos casos además, se utilizan para motivar el estudio de la regresión lineal ([H4], p. 226; [T4] p. 340).

A4. Demostración por reducción al absurdo. Este tipo de demostración parte de la situación contraria a la que se quiere demostrar y por medio de un razonamiento deductivo se muestra una contradicción, que lleva a la necesidad de negar la premisa inicial. El fundamento lógico básico de la misma consiste en que, al no poder ser cierta la negación de la tesis, ya que conduce a una contradicción, a un absurdo, se infiere la necesidad de que la tesis sea verdadera (CRESPO; FARFÁN, 2005).

En el análisis realizado tan sólo encontramos este tipo de demostración en los textos [H3] y [T3]. Un ejemplo aparece en la Figura 6, para demostrar que existen dos rectas de regresión diferentes. Partiendo de lo contrario (se supone que las rectas coinciden), se calculan los coeficientes de regresión cuyo producto es el coeficiente de correlación, llegándose a un absurdo pues se obtiene un coeficiente cuyo valor absoluto es mayor que la unidad.

Las rectas de regresión de una distribución bidimensional son las siguientes:

$$r:\ 7x - 5y = 8$$
$$s:\ 4x - 5y = -19$$

Demuestra que r *es la recta de regresión de* X *sobre* Y, *y* s, *la recta de regresión de* Y *sobre* X.

Para efectuar esta demostración utilizaremos el método de *reducción al absurdo:*

— Consideramos como cierta la opción contraria a la que pretendemos demostrar; es decir, que s es la recta de regresión de X sobre Y, y r, la recta de regresión de Y sobre X.

— Expresamos las rectas de regresión en la forma $y = A + Bx$ y $x = A' + B'y$.

$$r:\ y = -\frac{8}{5} + \frac{7}{5}x \qquad s:\ x = -\frac{19}{4} + \frac{5}{4}y$$

— Determinamos los coeficientes B y B', identificándolos en las expresiones de las rectas de regresión.

$$B = \frac{7}{5} \quad B' = \frac{5}{4}$$

— Calculamos el coeficiente de Pearson a partir de la expresión que lo relaciona con los coeficientes B y B':

$$r^2 = B \cdot B'$$

Así pues:

$$r = \sqrt{\frac{7}{5} \cdot \frac{5}{4}} = 1{,}323$$

— El valor de r que acabamos de hallar no es posible, puesto que el coeficiente de Pearson no puede tomar nunca valores superiores a 1.

Vemos que la opción contraria a la que pretendemos demostrar nos lleva a una situación absurda, por lo que no puede ser cierta.

Luego la opción cierta es la del enunciado.

Figura 6. Resolución de una tarea por el método de reducción al absurdo ([T3], p. 279)

A5. Argumento algebraico deductivo. Rara vez algún texto incluye también algunos argumentos deductivos básicamente realizados a partir de lenguaje algebraico. Se trata de manipulaciones algebraicas para tratar de argumentar una propiedad o mostrar la equivalencia de dos expresiones algebraicas. Estas son las justificaciones que tienen mayor dificultad, pues el alumno, además de seguir el razonamiento deductivo, ha de comprender todos los pasos en las operaciones con los símbolos. En el tratamiento de la correlación y regresión se suelen utilizar para demostrar algunas fórmulas de cálculo de de la covarianza, coeficiente de correlación lineal y coeficientes de regresión.

Un ejemplo de argumento deductivo se muestra en la Figura 7 donde, mediante operaciones simbólicas, se prueba la equivalencia entre dos fórmulas de cálculo de la covarianza. El mismo ejemplo aparece en [H8] (p. 251). Observamos que, además de desarrollar el producto de dos expresiones en paréntesis, descompone a continuación en suma de fracciones algebraicas, desarrollando el producto del factor común en cada sumando.

Además, se vuelve a sacar factor común cada una de las medias de X e Y y utiliza la propiedad de que la suma de frecuencias es igual a N para simplificar en el último término. En el tercer paso se ha de identificar las fórmulas de las dos medias, para finalmente simplificar de nuevo y llegar a la expresión. Observamos además, que el alumno ha de adivinar que el sumatorio recorre los valores 1 a N, pues no se indica el rango de variación. Todo ello es de gran complejidad

$$s_{XY} = \frac{\sum f_i(x_i - \bar{x})(y_i - \bar{y})}{N} \quad \text{o} \quad s_{XY} = \frac{\sum f_i x_i y_i}{N} - \bar{x}\,\bar{y}$$

El paso de la primera a la segunda expresión se hace de la siguiente forma:

$$s_{XY} = \frac{\sum f_i(x_i - \bar{x})(y_i - \bar{y})}{N} = \frac{\sum f_i(x_i y_i - \bar{y}x_i - \bar{x}y_i + \bar{x}\,\bar{y})}{N} =$$

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

Figura 7. Demostración de la equivalencia de dos expresiones de la covarianza ([T8], p. 321)

En cuanto a la relación entre el coeficiente de correlación lineal y los coeficientes de regresión, habitualmente se encuentra la siguiente cadena de igualdades, como ocurre en los textos [H1] (p. 232) y [T1] (p. 338):

$$\frac{\sigma_{xy}}{\sigma_x^2}\cdot\frac{\sigma_{xy}}{\sigma_y^2} = \left(\frac{\sigma_{xy}}{\sigma_x\sigma_y}\right)^2 = r^2$$

4.2. Síntesis de resultados

En las Tablas 1 y 2 presentamos la presencia o ausencia de cada uno de los tipos de justificación descritos en los textos analizados.

Todos los textos utilizan ejemplos y contraejemplos y gráficos como soporte de argumentación, resultado también encontrado en otros trabajos sobre libros de texto, como los de Alvarado (2007) sobre el teorema central del límite y Olivo (2008), sobre intervalos de confianza, incluso cuando en dichos estudios los textos son de nivel universitario Igualmente encontramos en todos los textos el uso de argumentaciones verbales deductivas, en general, con poca formalización.

Tabla 1. Argumentos en los textos analizados (Humanidades y Ciencias Sociales)

Argumentos	H1	H2	H3	H4	H5	H6	H7	H8
A1.Ejemplos/contraejemplos	x	x	x	x	x	x	x	x
A2.Gráficos auxiliares	x	x	x	x	x	x	x	x
A3.Verbales deductivos	x	x	x	x	x	x	x	x
A4.Reducción al absurdo			x					
A5.Algebraicos deductivos	x					x	x	x

Tabla 2 Argumentos en los textos analizados (modalidad Ciencias y Tecnología)

Argumentos	T1	T2	T3	T4	T5	T6	T7	T8
A1.Ejemplos/contraejemplos	x	x	x	x	x	x	x	x
A2.Gráficos auxiliares	x	x	x	x	x	x	x	x
A3.Verbales deductivos	x	x	x	x	x	x	x	x

A4.Reducción al absurdo			x					
A5.Algebraicos deductivos	x		x		x	x	x	x

Destacamos los textos [H3] y [T3], por ser los únicos en tratar el método de reducción al absurdo, proponiendo, además, a los estudiantes que realicen alguna demostración utilizándolo y explicando los pasos utilizados. Una parte de los textos utilizan también argumentos algebraicos deductivos; siendo [T3] el único texto que utiliza todas las formas de argumentación que hemos analizado y el más formalizado en sus argumentos.

Encontramos algunas diferencias en el tipo de argumentación utilizada por la misma editorial, dependiendo de la modalidad de bachillerato para el cual se diseñen cada uno. Así es que, algunos textos de la modalidad de ciencias y tecnología tratan mucho más la argumentación de tipo simbólico que los correspondientes a la misma editorial de la modalidad de humanidades y ciencias sociales.

Los textos [T3] y [T5] de la modalidad científico-tecnológica, se diferencian de sus respectivos [H3] y [H5] por incluir tareas donde se pide al estudiante realizar alguna demostración utilizando un tipo específico de argumento, a pesar de no incluir argumentos simbólicos deductivos desarrollados (un ejemplo se presenta en la Figura 8). Un ejemplo es el siguiente ([T3], p. 278):

> Demuestra las siguientes proposiciones:
>
> a) Las dos formas alternativas de expresar las rectas de regresión de *X* sobre *Y* y de *Y* sobre *X* son equivalentes.
>
> b) El signo de las pendientes *B* y *B'* de las rectas de regresión y el signo del coeficiente de correlación de Pearson son iguales
>
> c) Si existe dependencia funcional entre las variables *X* e *Y*, las rectas de regresión de *X* sobre *Y* y de *Y* sobre *X* coinciden.
>
> d) Si existe independencia entre las variables *X* e *Y* las rectas de regresión de *X* sobre *Y* son perpendiculares entre si y paralelas a los ejes.

Este tipo de tareas no son exclusivas de la modalidad científico-tecnológica ya que, los textos [H6] y [H7] incluyen este tipo de situaciones, al igual que sus respectivos [T6] y [T7]. Un ejemplo se presenta en la Figura 8, que se muestra tanto en [H6] como en [T6], donde el estudiante no sólo ha de argumentar utilizando un lenguaje algebraico deductivo, sino incluso ha de descubrir la propiedad que se quiere demostrar.

Considerando el conjunto de datos:

X	x_1	x_2	....	x_n
Y	y_1	y_2	....	y_n

De una distribución bidimensional de variables X e Y, explicar qué es la regresión lineal.

Supongamos que $y = a + bx$ es la recta de regresión de la variable y sobre la variable x. Indica la relación de los coeficientes a y b con la expresión siguiente:

$$(y_1 - a - bx_1)^2 + (y_2 - a - bx_2)^2 + \dots + (y_n - a - bx_n)^2$$

Figura 8. Tarea para justificar mediante lenguaje simbólico una relación ([T6], p. 321).

En el caso de [H7] y [T7] se pide al estudiante que investigue sobre dos tareas, para las que necesitará hacer uso de lenguaje simbólico, sobre todo en la tarea (b); como en el anterior ejemplo, en estas dos actividades el alumno ha de descubrir qué es lo que tiene que demostrar.

> 56. Investiga sobre las siguientes cuestiones.
> a) ¿Es cierto que el signo de las pendientes de las dos rectas de regresión de una variable bidimensional es siempre igual?
> b) ¿Qué sucede si las dos rectas de regresión tienen la misma pendiente? ¿Cómo es la correlación? ([H7], p. 263).

5. CONCLUSIONES

El análisis de las justificaciones, en la muestra de textos de bachillerato, se ha llevado a cabo por su relevancia en la construcción de la competencia matemática de los estudiantes. La investigación sobre este tema contribuye, por tanto, a proporcionar criterios para la mejora de los textos y por ello a facilitar el aprendizaje de los estudiantes. Como hemos indicado en los antecedentes, la investigación sobre el libro de texto de matemáticas es una línea importante, pues nos ayuda a comprender mejor el fenómeno de la transposición didáctica (CHEVALLARD, 1991).

En general, los textos poseen un nivel de argumentación aceptable, e incluyen los diferentes tipos de argumentos considerados por Godino y Recio, ya que incluyen las argumentaciones gráficas y mediante ejemplos o contraejemplos (que a veces coinciden pues una misma demostración a través de un ejemplo es de tipo gráfico), acompañado del uso verbales deductivos, aunque son mucho menos frecuentes.

Aunque dirigidos a alumnos con diferente preparación y motivación, no se han observado grandes diferencias en las dos modalidades de bachillerato; de hecho la misma editorial suele hacer una presentación muy parecida e la justificación ambos tipos de alumnos. Una excepción son los argumentos simbólicos o por reducción al absurdo y aparecen, que en general son escasos, pero aparecen con mayor frecuencia en los textos de la modalidad de

Ciencia y Tecnología, preferentemente como tarea propuesta. Pensamos que este mayor énfasis se debe a las directrices relacionadas con el uso de la demostración para *Matemáticas I* y *Matemáticas aplicadas a las Ciencias Sociales I* como se muestra en el Decreto de Enseñanzas Mínimas.

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. En todo caso coincidimos con Ortiz (1999) en la necesidad de dar más atención a la actividad de formación de argumentos matemáticos, posiblemente y dar oportunidad a los alumnos de realizar actividades que impliquen la producción de sus propios textos matemáticos.

REFERENCIAS

ALVARADO, H. *Significados del teorema central del límite en la enseñanza de la estadística en ingeniería*. Tesis doctoral. Universidad de Granada, 2007.

AZCÁRATE, P.; SERRADÓ, A. Tendencias didácticas en los libros de texto de matemáticas para la Eso. *Revista de Educación*, *340*, 341-378, 2006.

CASTRO-SOTOS, A. E., VANHOOF, S., VAN DEN NOORTGATE, W.; ONGHENA, P. The transitivity misconception of Pearson's correlation coefficient. *Statistics Education Research Journal 8* (2), 33-55, 2009. Disponible en: www.stat.auckland.ac.nz/~iase/serj/.

CHAPMAN, L. J.; CHAPMAN, J. P. Illusory correlation as an obstacle to the use of valid psychodiagnostic signs. *Journal of Abnormal Psychology*, *74* (3), 271-280, 1969.

CHEVALLARD, Y. *La transposición didáctica. Del saber sabio al saber enseñado*. Buenos Aires: Aique, 1991

COBO, B.; BATANERO, C. Significados de la media en los libros de texto de secundaria. *Enseñanza de las Ciencias, 22* (1), 5-18, 2004.

CRESPO, C.; FARFÁN, R. Una visión socio epistemológica de las argumentaciones en el aula. El caso de las demostraciones por reducción al absurdo. *Relime*, 8(3) 287-317, 2005.

DAVIS, P. J.; HERSH, R. *Experiencia matemática*. Barcelona: Labor, 1988.

DE VILLIERS, M. El papel y la función de la demostración en matemáticas. *Épsilon*, 26, 15-30, 1993.

ESTEPA, A. Interpretación de los diagramas de dispersión por estudiantes de Bachillerato. *Enseñanza de las Ciencias, 26* (2), 257-270, 2008.

ESTEPA, A.; BATANERO, C. Concepciones iniciales sobre la asociación estadística. *Enseñanza de las Ciencias, 13* (2), 155-170, 1995.

ESTEPA, A., M; GEA, CAÑADAS; G. CONTRERAS, M. Algunas notas históricas sobre la correlación y regresión y su uso en el aula. *Números, 81*, 5-14, 2012.

GODINO, J. D.; RECIO, Á. M. Significados Institucionales de la demostración: Implicaciones para la educación matemática. *Enseñanza de las Ciencias* 19(3), 405-414, 2001.

GOETZ, J. P.; LECOMPTE, M. D. *Etnografía y diseño cualitativo en educación*. Morata: Madrid, 1998.

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

HOLMES, P. Correlation: From picture to formula. *Teaching Statistics, 23* (3), 67-71, 2001.

KRIGOWSKA, A. Z. Treatment of the asiomatic method in class. En W. Servais y T. Vargas (Eds.), Teaching school mathematics (pp. 124-150). London: Penguin, 1971.

KRIPPENDORFF, K. *Metodología de análisis de contenido. Teoría y práctica*. Barcelona: Paidos, 1997.

LAVALLE, A. L.; MICHELI, E. B.; RUBIO, N. Análisis didáctico de regresión y correlación para la enseñanza media. *Revista Latinoamericana de Investigación en Matemática Educativa*, 9 (3), 383-406. 2006.

LÓPEZ NOGUERO, F. El análisis de contenido como método de investigación. *Revista de Educación, 4*, 167-180, 2002.

LOWE, E.; PIMM, D. '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, 1996.

MARTÍNEZ, M. La investigación cualitativa (síntesis conceptual). *Revista de Investigación en Psicología 9* (1), 123-146, 2006.

MEC. *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, 2007.

NCTM. *Principles and standards for school mathematics.* Reston: VA, 2000.

OLIVO, E. *Significados de los intervalos de confianza para los estudiantes de ingeniería en México*. Tesis doctoral. Universidad de Granada, 2008.

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

ORTIZ, J. J.; BATANERO, C.; SERRANO, L. El lenguaje probabilístico en los libros de texto. *Suma, 38*, 5-14, 2001.

RECIO, A. M. La demostración en matemática: Una aproximación epistemológica y didáctica. *Investigación en Educación Matemática V*, 27-44, 2002.

SÁNCHEZ COBO, F. T. *Significado de la correlación y regresión para los estudiantes universitarios*. Tesis doctoral. Universidad de Granada, 1999.

ZIEFFLER, A.; GARFIELD, J. Modeling the growth of students' covariational reasoning during an introductory statistics course. *Statistics Education Research Journal, 8* (1), 7-31, 2009.

ANEXO. Textos analizados

Código	Referencia
H1	COLERA, J.; OLIVEIRA, M. J.; GARCÍA, R.; SANTAELLA, E. *Matemáticas aplicadas a las Ciencias Sociales I*. Madrid: Grupo Anaya, 2008.
H2	ARIAS, J. M.; MAZA, I. *Matemáticas aplicadas a las Ciencias Sociales 1*. Madrid: Grupo Editorial Bruño, 2011.
H3	ANGUERA, J.; BIOSCA, A.; ESPINET, M. J.; FANDOS, M. J.; GIMENO, M.; REY, J. *Matemáticas I aplicadas a las Ciencias Sociales*. Barcelona: Guadiel - Grupo Edebé, 2008.
H4	MONTEAGUDO, M. F.; PAZ, J. *1º Bachillerato. Matemáticas aplicadas a las Ciencias Sociales.* Zaragoza: Edelvives (Editorial Luis Vives), 2008.
H5	MARTÍNEZ, J. M.; CUADRA, R.; HERAS, A. *Matemáticas aplicadas a las Ciencias Sociales. 1.º Bachillerato*. Madrid: McGraw-Hill, 2008.
H6	BESCÓS, E.; PENA, Z. *Matemáticas aplicadas a las Ciencias Sociales.* Vizcaya: Oxford University Press España, 2008.
H7	ANTONIO, M.; GONZÁLEZ, L.; LORENZO, J.; MOLANO, A.; DEL RÍO, J.; SANTOS, D.; DE VICENTE, M. *Matemáticas aplicadas a las Ciencias Sociales I*. Madrid: Santillana Educación, 2009.
H8	VIZMANOS, J. R.; HERNÁNDEZ, J.; ALCAIDE, F.; MORENO, M.; SERRANO, E. *Matemáticas aplicadas a las Ciencias Sociales 1*. Madrid: Ediciones SM, 2008.
T1	COLERA, J.; OLIVEIRA, M. J.; GARCÍA, R.; SANTAELLA, E. *Matemáticas I.* Madrid: Grupo Anaya, 2008.
T2	ARIAS, J. M.; MAZA, I. *Matemáticas 1*. Madrid: Grupo Editorial Bruño, 2011.
T3	BIOSCA, A.; DOMÉNECH, M.; ESPINET, M. J.; FANDOS, M. J.; JIMENO, M. *Matemáticas I*. Barcelona: Guadiel - Grupo Edebé, 2008.
T4	MONTEAGUDO, M. F.; PAZ, J. *1º Bachillerato. Matemáticas. Ciencias y Tecnología.* Zaragoza: Edelvives (Editorial Luis Vives), 2008.
T5	MARTÍNEZ, J. M.; CUADRA, R.; BARRADO, F. J. *Matemáticas 1º Bachillerato*. Madrid: McGraw-Hill, 2007.
T6	BESCÓS, E.; PENA, Z. *Matemáticas. 1 Bachillerato.* Navarra: Oxford University Press España, 2009.
T7	ANTONIO, M.; GONZÁLEZ, L.; LORENZO, J.; MOLANO, A.; DEL RÍO, J.; SANTOS, D.; DE VICENTE, M. *Matemáticas I. 1 Bachillerato*. Madrid: Santillana Educación, 2008.
T8	VIZMANOS, J. R.; HERNÁNDEZ, J.; ALCAIDE, F.; MORENO, M.; SERRANO, E. *Matemáticas 1*. Madrid: Ediciones SM, 2008.