

## Una experiencia de formación para futuros profesores en correlación y regresión

*A prospective teacher training experience on  
correlation and regression*

**María M. Gea Serrano**

Universidad de Granada, España  
[mmgea@ugr.es](mailto:mmgea@ugr.es)

**Nuria Begué Pedrosa**

Universidad de Zaragoza, España  
[nbegue@unizar.es](mailto:nbegue@unizar.es)

Recibido: 29 de junio de 2020

Aceptado: 29 de septiembre de 2020

DOI: 10.5377/ryr.v53i53.10892



## RESUMEN

En este trabajo se presenta una experiencia de formación en estadística, dirigida a futuros profesores de Educación Secundaria y Bachillerato, implementada en un curso del Máster en formación de profesorado en matemáticas en España, que es una titulación obligatoria para impartir docencia en dicha etapa educativa. La experiencia está basada en los contenidos de correlación y regresión y su diseño metodológico está fundamentado en herramientas del marco teórico del enfoque ontosemiótico del conocimiento y la instrucción matemáticos, que propone un modelo de investigación sobre Conocimientos y Competencias Didáctico Matemáticas (CCDM) del profesor. Se describen brevemente resultados de la experiencia, que consideramos será de utilidad tanto a profesores como a formadores de profesores en el tema.

**Palabras clave:** formación de futuros profesores, correlación y regresión, idoneidad didáctica.

## ABSTRACT

*This work describes a training experience aimed at prospective teachers of secondary education and high school in statistics, implemented in the Master's in secondary teacher training in mathematics in Spain, which is a compulsory degree to teach at that educational stage. The experience is based on correlation and regression through methodological tools from the theoretical framework of the ontosemiotic approach to mathematical knowledge and instruction, which proposes a research model on Didactic-Mathematical Knowledge and Competencies (DMKC). Results of experience are briefly presented that we consider there will be support for teacher and training teacher.*

**Keywords:** *prospective teacher training, correlation and regression, didactical suitability.*

## 1. Introducción

Los avances en ciencia y tecnología han permitido que, en la sociedad actual, tengamos acceso a gran cantidad de datos por parte de instituciones y oficinas de información. Estos datos se pueden descargar y analizar con software específico, o bien utilizar aplicaciones o software online que incluso las propias instituciones ponen a nuestro servicio para el tratamiento de la información.

La competencia en el análisis de los datos nos permitirá tomar decisiones ajustadas a la información disponible, pero como sugieren Batanero, Díaz, Contreras y Roa (2013), esta competencia es difícil de transmitir a un ciudadano, ya que requiere un adecuado razonamiento estadístico sobre los resultados obtenidos por dichas herramientas. En base al trabajo de Burrill y Biehler (2011), los autores identifican ideas que resultan fundamentales para enfrentarse a la toma de decisiones en ambiente de incertidumbre. Entre otras destacan, como componente fundamental del razonamiento estadístico, el análisis de la dependencia entre variables aleatorias.

Para determinar si hay sincronía más o menos intensa entre dos variables se debe poner en marcha una actitud crítica ante los datos. La correlación se refiere al análisis de la intensidad de la dependencia entre variables estadísticas cuantitativas, lo que informa de la utilidad de determinar, mediante el análisis de regresión, un modelo predictivo que relacione dichas variables. El estudio de la estadística bidimensional, además, es la base para entender otros conceptos y procedimientos estadísticos más avanzados.

En la literatura de investigación se han identificado diferentes concepciones erróneas cuando se razona ante datos bivariados (Estepa, 1994), siendo algunas de ellas muy resistentes al cambio después de un diseño específico de enseñanza. Según Castro-Sotos, Vanhoof, Van Den Noortgate y Ongghena (2009), algunos estudiantes extienden propiedades deterministas a este tipo de situaciones, lo que nos muestra la gran dificultad en comprender que los conceptos de correlación y regresión extienden el estudio del análisis funcional a situaciones aleatorias.

La enseñanza del tema en España se inicia en cuarto curso de Educación Secundaria (14-15 años), mediante el estudio de la representación tabular y gráfica de una variable estadística bidimensional (tablas de doble entrada y nube de puntos). En el primer curso de Bachillerato (16 años) se estudia el tema de una manera más formal, mediante el análisis de las distribuciones unidimensionales (marginales y condicionadas) asociadas a una variable estadística bidimensional, el cálculo e interpretación de la covarianza y el coeficiente de correlación lineal, así como ajustar un modelo lineal a los datos para hacer predicciones (cálculo de coeficientes de regresión y coeficiente de determinación) (MECD, 2015).

Al tratarse de contenidos incluidos en la normativa curricular vigente en la etapa de secundaria en España (MECD, 2015), y dada la relevancia que posee el tema de correlación y regresión en el razonamiento estadístico, nos planteamos como eje fundamental de actuación diseñar un plan de

formación dirigido a una muestra de estudiantes que se preparan para ser profesores en el Máster de Formación del Profesorado de Educación Secundaria y Bachillerato, en la especialidad de matemáticas, en España, que a su vez nos informe del nivel de conocimiento didáctico y matemático sobre correlación y regresión que posee la muestra participante del estudio.

Consideramos que una buena enseñanza requiere de una adecuada preparación del profesorado y, según las recomendaciones actuales de enseñanza de la estadística, se debe favorecer que el estudiante aplique los conocimientos y comprenda su utilidad. Siguiendo las directrices propuestas por Batanero, Díaz y Gea (2011), planteamos una situación de enseñanza basada en correlación y regresión mediante el trabajo con proyectos. Las autoras evidencian la gran riqueza y posibilidades de enseñanza que posee la estadística como es la interdisciplinariedad, por lo que escogemos un contexto cercano al estudiante como es la esperanza de vida al nacer en un país y su relación con otros indicadores de desarrollo humano como es el gasto en salud pública, el PIB, etc. La planificación del diseño metodológico responde por una parte al plan de formación del máster de profesorado y, por otra, al estudio exploratorio de evaluación del conocimiento didáctico y matemático del profesor, siguiendo herramientas del marco teórico del enfoque ontosemiótico del conocimiento e instrucción matemáticos (Godino, 2009; 2013; 2017; Godino, Batanero y Font, 2019; Godino, Giacomone, Batanero y Font, 2017).

En las siguientes secciones se describen los principios que fundamentan el diseño del plan de formación, algunos resultados de su implementación y finalizamos con algunas implicaciones para la formación de profesores.

## **2. Fundamentos**

### **2.1. El modelo de conocimiento y competencias del profesor en matemáticas**

La investigación sobre el conocimiento del profesor para la enseñanza de las matemáticas ofrece diferentes modelos teóricos en los que se caracterizan sus diferentes componentes (Godino *et al.*, 2017; Hill, Ball, y Schilling, 2008; Shulman, 1986). Para nuestro trabajo utilizamos el marco teórico del enfoque ontosemiótico del conocimiento e instrucción matemático (Godino, 2009; 2013; 2017; Godino *et al.*, 2019; Godino *et al.*, 2017), pues los principios en que se fundamenta responden a nuestra problemática sobre cómo un sujeto construye el conocimiento matemático y nos proporciona herramientas teóricas de gran utilidad para el desarrollo y evaluación del conocimiento didáctico matemático del profesor de matemáticas.

Desde el enfoque ontosemiótico se considera que el significado de un objeto matemático surge de las prácticas que un sujeto realiza cuando resuelve una situación problema, en la que intervienen otros muchos objetos, entendidos como primarios, como es el lenguaje (términos, expresiones, notaciones, gráficos, etc.), conceptos (elementos en la tarea que se puede definir), propiedades (enunciados o

proposiciones que regulan las prácticas en la tarea), la diversidad de procedimientos que pueden emplearse (operaciones, algoritmos, técnicas de cálculo, etc.), así como el tipo de argumentación empleado para demostrar o justificar las prácticas desempeñadas (Godino, 2017).

En cuanto al diseño, análisis e intervención didáctica, se distinguen distintos niveles o sub-competencias que conforman la competencia general de un profesor de matemáticas (Figura 1). En un primer nivel se identifica la competencia de análisis de significados globales implicados en la práctica matemática, explicado anteriormente, así como disponer de la competencia de análisis ontosemiótico de dichas prácticas, en un nivel más avanzado, entendido en términos de configuraciones implicadas en el proceso de enseñanza y aprendizaje. La gestión del profesor en el aula, pues las diferentes interacciones entre el alumnado y el material empleado requiere también atención, dada la compleja trama de normas, de diferente origen y naturaleza que influyen en el proceso de enseñanza y aprendizaje. Así es que el análisis normativo ocupa un rol fundamental en el desempeño de la competencia del docente, por su influencia en alcanzar niveles óptimos de idoneidad, que se refiere al último nivel de competencia del docente. La competencia de análisis de la idoneidad didáctica ha sido objeto de atención en muchos trabajos de investigación en educación estadística, dado que muchos procesos formativos para desarrollar la competencia y conocimientos específicos del profesor de matemáticas se han basado en su uso: "se trata de que el profesor reflexione sobre el diseño, implementación y evaluación de una experiencia de enseñanza de un tema específico en un contexto educativo fijado" (Godino *et al.*, 2017, p. 107).



Figura n.º 1. Competencia de análisis e intervención didáctica del profesor de matemáticas. Fuente: Godino *et al.* (2017, p. 103)

El modelo del conocimiento del profesor propuesto en el enfoque ontosemiótico se establece según tres dimensiones como es la dimensión matemática (conocimiento matemático no necesariamente

orientado a la enseñanza, referido al nivel educativo concreto al que se refiere el proceso de enseñanza o niveles educativos posteriores); dimensión didáctica (conocimiento de los factores que influyen en la planificación e implementación del proceso de enseñanza y aprendizaje del tema) y dimensión meta didáctica matemática (conocimiento de la propia práctica docente, referido al contexto de la enseñanza en el aula y su entorno, con objeto de establecer mejoras en el mismo). En particular, el conocimiento didáctico se estructura en seis facetas (Pino-Fan y Godino, 2015):

- Faceta epistémica, que se refiere a la tipología de objetos primarios implicados en la situación problema, por lo que constituye el conocimiento especializado de la dimensión matemática del profesor;
- Faceta cognitiva, que se refiere al conocimiento de la evolución del razonamiento y aprendizaje de los estudiantes;
- Faceta afectiva, referida al conocimiento del grado de implicación y motivación del estudiante en el proceso de enseñanza y aprendizaje del tema, sus estados afectivos en cuanto a actitudes, emociones, creencias, etc.;
- Faceta mediacional, que se refiere al conocimiento de los materiales y recursos (didácticos, tecnológicos, manipulativos, etc.) favorecedores de la enseñanza y aprendizaje del tema, así como la administración del tiempo en las diferentes acciones del proceso de enseñanza y aprendizaje;
- Faceta interaccional, referida al conocimiento de los patrones o tipos de interacción en el aula y su secuenciación, que condicionan la comunicación entre profesor y alumnos o entre alumnos en el proceso de enseñanza y aprendizaje. Esta faceta es muy importante en el conocimiento del profesor, pues le permite identificar posibles desajustes en la adquisición del significado de un objeto matemático por parte del estudiante;
- Faceta ecológica, referida al conocimiento del profesor en las relaciones del proceso de enseñanza y aprendizaje con el proyecto educativo (de la institución, el centro y el departamento), el entorno natural, social, económico, político y cultural en que se desarrolla.

## **2.2. Sentido de la correlación y regresión**

La idea de sentido estadístico definida por Batanero *et al.* (2013) ofrece un marco útil al docente para promover en sus estudiantes el desempeño y comprensión adecuada de los contenidos estadísticos. Este marco es adaptado por Gea, Batanero y Roa (2014) para los contenidos de correlación y regresión, por lo que los autores determinan los componentes de cultura y razonamiento estadístico que se requieren para comprender estos contenidos.

En cuanto a los componentes de la cultura estadística, se destaca la importancia de disponer de una posición crítica ante los datos y un componente conceptual adecuado, que asegure el dominio de contenidos o ideas fundamentales en el tema, que permitan al estudiante afrontar las situaciones problema con éxito. En cuanto a la actitud del estudiante, debe valorar la utilidad de estos conceptos así como plantear distanciamiento cuando se analicen datos bivariados, puesto que, por ejemplo, puede presentarse la existencia de otras variables que afecten a la dependencia de las variables de estudio (efecto conocido como paradoja de Simpson) o bien podemos dejarnos llevar por creencias equivocadas más que por la evidencia que aportan los datos (Chapman, 1967).

En cuanto a los conceptos fundamentales que se deben dominar se destacan:

- Datos y distribución. Los datos bivariados proceden de dos variables estadísticas unidimensionales, consideradas en un mismo individuo muestral, y su estudio aborda no sólo la distribución bidimensional, sino otros tipos de distribuciones como las marginales o condicionadas.
- Representación tabular y gráfica. Las tablas y gráficos estadísticos son esenciales para la visualización de la información. La representación de los datos más utilizada es el diagrama de dispersión o nube de puntos, aunque también encontramos otros tipos de gráficos de gran uso en nuestro día a día como es el gráfico de burbuja, que permiten representar hasta cuatro variables estadísticas).
- Variabilidad. La dispersión de los datos se observa fácilmente en la nube de puntos, donde se puede visualizar conceptos como la covarianza y su relación con el modelo de ajuste lineal de los datos.
- Dependencia funcional, aleatoria e independencia. La dependencia funcional difiere de la aleatoria en cuanto a que a cada valor de la variable independiente corresponde una distribución de valores de la variable dependiente. Conviene diferenciar estos tipos de dependencia, así como aquellos casos en que exista independencia, pues se suele confundir con correlación lineal nula.
- Covarianza y correlación. Estos conceptos permiten medir la intensidad y dirección de la dependencia entre dos variables. El coeficiente de correlación lineal, a diferencia de la covarianza, está acotado entre los valores  $-1$  y  $1$ .
- Regresión. El modelo lineal es uno de muchos modelos que se pueden ajustar a los datos, por lo que el análisis del modelo más adecuado permitirá asegurar un buen ajuste en la predicción de una variable en función de otra.
- Modelos de regresión y sus parámetros. Es importante el dominio de conceptos como variable dependiente e independiente en la determinación del modelo de ajuste, así como valorar la influencia en dicho ajuste de datos atípicos.

- Estimación y bondad de ajuste. La posibilidad de aplicar el modelo de ajuste para predecir valores de la variable dependiente requiere del dominio de conceptos como fiabilidad del modelo y centro de gravedad de la distribución de datos.

Con todo ello, en lo que se refiere al razonamiento con datos bivariados, se requiere dominar las diferentes representaciones (gráfica, tabular, textual o numérica) que se manejan en el estudio de la dependencia entre variables, para determinar los parámetros que conforman el modelo de ajuste a los datos. Así es que, se necesita razonar en torno al signo e intensidad del coeficiente de correlación lineal; plantear una actitud crítica en el razonamiento ante los datos y su contexto, cuando no concuerden con las teorías previas; disponer de suficientes datos e información para no limitar la toma de decisiones a un subconjunto de la información; y valorar la posibilidad de encontrar otras situaciones de dependencia entre las variables diferentes de la causa y efecto (Barbancho, 1992).

### 3. Antecedentes

La literatura previa en formación e investigación del profesorado en correlación y regresión es escasa. Se destaca la investigación desarrollada por Estepa y colaboradores (Estepa, 1994; Batanero, Estepa, y Godino, 1997; Batanero, Godino y Estepa, 1998), al tratarse del trabajo más completo sobre el aprendizaje de la correlación y regresión, no sólo con futuros profesores, sino en general. En el caso particular de la formación del profesorado, se describe el diseño e implementación de un proceso de enseñanza dirigido a una pequeña muestra de futuros profesores de educación primaria (22 futuros profesores) haciendo uso de la tecnología en un curso de análisis exploratorio de datos. Se observa cambio en muchas de las estrategias y concepciones erróneas de los participantes en los ítems relacionados con los juicios de correlación en diagramas de dispersión (casi el doble, en la mejora de estrategias empleadas), aunque el cambio no fue homogéneo y, en general, se observa una gran dificultad en la mejora de la concepción causal de la correlación, considerando la dependencia entre variables solo cuando se aprecia relación causa-efecto entre las mismas (Batanero *et al.*, 1997).

La línea de investigación desarrollada por Casey (2008; 2010) también se centra en investigar el conocimiento del profesor para la enseñanza sobre correlación y regresión, pero su atención se dirige a ofrecer un marco de referencia de conocimientos requeridos para la enseñanza del tema. En su investigación analiza el conocimiento en torno a la planificación, la implementación y la gestión del proceso de instrucción de una muestra de 19 profesores en ejercicio. Como resultados se observa el limitado conocimiento que poseen los profesores a la hora de resolver problemas de dependencia estadística, empleando generalmente sus teorías previas como estrategia en la resolución de las situaciones planteadas; no proponen situaciones que promuevan el análisis de la posible independencia entre variables y se sienten poco seguros al resolver situaciones de dependencia entre variables aleatorias (Casey y Wasserman, 2015). Resultados similares se encuentran en otras investigaciones donde se observa el conocimiento didáctico y matemático puesto en práctica en la enseñanza del tema (Quintas, Ferreira y Oliveira, 2015).

#### 4. Diseño de la experiencia

La formación del profesor en educación estadística ocupa una agenda de investigación de gran importancia, aunque su presencia es relativamente reciente en el panorama de investigación en educación matemática. Se observa la necesidad de impulsar esta línea de investigación a nivel internacional en la publicación sobre la educación y desarrollo profesional de profesor para enseñar estadística en el *Joint ICMI/LASE Study "Teaching Statistics in School Mathematics. Challenges for Teaching and Teacher Education"* (Batanero, Burrill y Reading, 2011).

Por abordar nuestro objetivo de diseñar un plan de formación dirigido a una muestra de estudiantes que se preparan para ser futuros profesores en el Máster de Formación del Profesorado de Educación Secundaria y Bachillerato, en la especialidad de matemáticas, en España, que a su vez nos informe del nivel de conocimiento didáctico y matemático que poseen sobre correlación y regresión, utilizamos el desarrollo de diferentes trabajos de investigación desarrollados en el Grupo de Investigación de Educación Estadística de la Universidad de Granada (GEEUG). Carmen Batanero ha dirigido diferentes proyectos de investigación y tesis doctorales centrados en el campo de la educación estadística (estadísticos de centralización, posición y dispersión; probabilidad; aleatoriedad; distribución normal, binomial; gráficos y tablas estadísticas; asociación y correlación, inferencia, etc.) y junto a ser co-protagonista de la elaboración del enfoque ontosemiótico, todo esto ha potenciado el uso de diversas herramientas del marco teórico (Godino, 2019) en las investigaciones sobre educación estadística, muchas otras desarrolladas también en el marco teórico del enfoque ontosemiótico aunque no en el grupo GEEUG (Vásquez y Alsina, 2015; 2017).

De entre las investigaciones desarrolladas en el grupo GEEUG destacamos la tesis doctoral de Arteaga (2011), dirigida a desarrollar y evaluar el conocimiento didáctico matemático de futuros profesores sobre gráficos estadísticos. Nosotros empleamos el mismo método para la correlación y regresión, considerando una situación de enseñanza adaptada al nivel educativo de Bachillerato, que es donde se regula la enseñanza del tema en el sistema educativo español.

Siguiendo el enfoque metodológico empleado por Arteaga (2011), se parte del diseño de un proyecto estadístico centrado en el tema de correlación y regresión, según las indicaciones curriculares de la etapa escolar de Bachillerato (MECD, 2015) y la fundamentación teórica descrita en este trabajo, y se emplea la pauta de valoración de la idoneidad didáctica (Godino, 2013) como recurso en la evaluación y desarrollo de las facetas del conocimiento del profesor.

#### **Desde este diseño metodológico, se trabaja con los futuros profesores en tres fases diferentes:**

- Trabajo con una situación didáctica orientada a la enseñanza de correlación y regresión en forma de proyecto, que el futuro profesor debe resolver, pero que, además, posteriormente puede utilizar con sus estudiantes de Bachillerato.

Se trata de desarrollar el razonamiento estadístico de los futuros profesores al recorrer las diferentes fases del ciclo de investigación descrito por Wild y Pfannkuch (1999): problema, plan, datos, análisis y conclusión; respondiendo a diferentes interrogantes (generar, buscar, interpretar, criticar y juzgar) y poniendo a prueba su disposición al tema (escepticismo, perseverancia, compromiso e imaginación, entre otras). Tal como se recomienda en el trabajo por proyectos, la interpretación de la información haciendo uso de la tecnología y disponer de datos reales contextualizados es una de las premisas fundamentales del diseño de la situación problema. De este modo, se invita al futuro profesor a plantear y responder preguntas en la realización del proyecto propuesto, que potenciarán su capacidad para argumentar sus decisiones y conclusiones, así como potenciar su creatividad y postura crítica ante la información que recibe en su día a día.

- Discusión en clase con el formador de profesores de las soluciones en la fase anterior.
- Reflexión y discusión final sobre el concepto de idoneidad didáctica, sus componentes, cómo se ponen de manifiesto en el proceso formativo y cómo se puede aumentar. Para llevar a cabo esta actividad se plantea el análisis por parte de los futuros profesores de la idoneidad didáctica del proyecto desarrollado (Godino, 2013). El estudio cualitativo sistemático de las producciones de los futuros profesores al analizar las facetas de la idoneidad didáctica dará información para realizar el análisis cualitativo de las producciones de los participantes, que se completa con la asignación de puntuaciones numéricas a las respuestas, lo que permite valorar la calidad del conocimiento del futuro profesor en el tema.

#### 4.1. El proyecto

En Gea (2014) se encuentra una descripción más detallada del proyecto diseñado sobre correlación y regresión dirigido a los futuros profesores de Educación Secundaria y Bachillerato que cursan el Máster de Formación del Profesorado de Educación Secundaria y Bachillerato, en la especialidad de matemáticas, en España. Se enmarca en una asignatura de innovación e investigación didáctica en matemáticas, que incluye tareas con datos reales basados en el Informe de Desarrollo Humano por las Naciones Unidas, que están disponibles en su servidor (<http://hdr.undp.org/es/data>). El formador de profesores ofrece el material necesario (herramientas y medios) para secuenciar y apoyar el aprendizaje del tema, aunque también muestra cómo obtener los datos desde el servidor para ofrecer la posibilidad de que ellos mismos completen el proyecto con el análisis de otras variables diferentes a las seleccionadas en las tareas.

El proyecto se enfoca en la pregunta “¿Cuáles son los factores que más influyen en la esperanza de vida al nacer en un país?”, desde la que se plantean diferentes tareas con los datos de los Informes sobre Desarrollo Humano por las Naciones Unidas, que ellos pueden representar con Excel, pero que en las hojas de trabajo vienen ya preparados para visualizar de manera rápida la información (Figura 2).

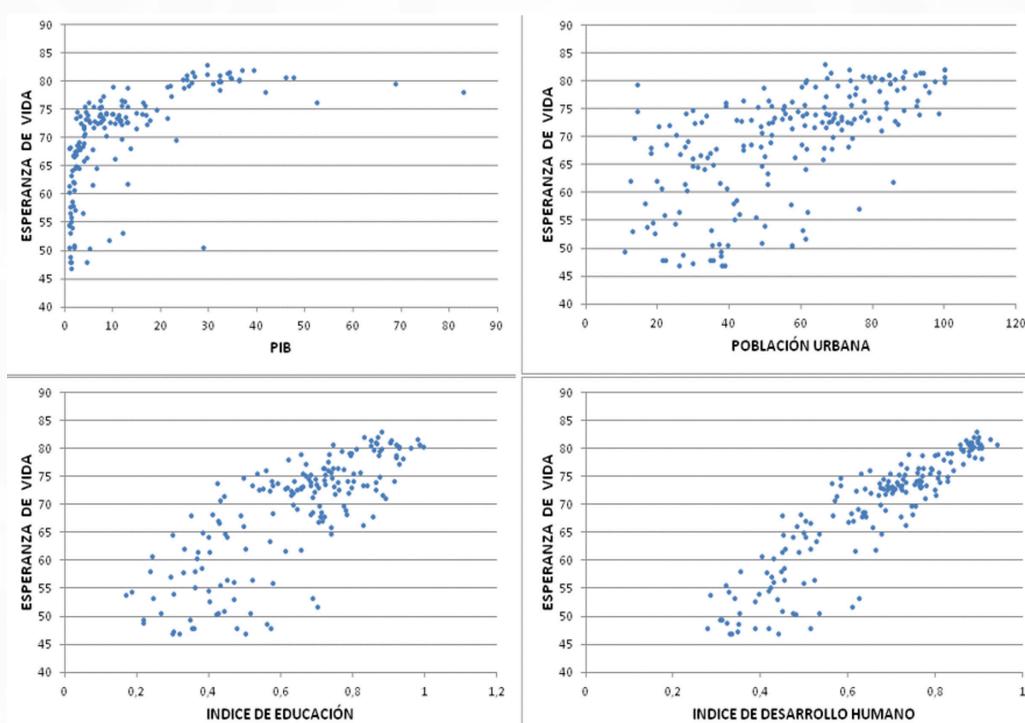


Figura n.º 2. Diagrama de dispersión de la variable esperanza de vida en función de algunas de las variables explicativas incluidas en el fichero de trabajo del proyecto. Fuente: Elaboración propia

Las variables utilizadas en el proyecto son indicadores internacionales de desarrollo humano (Índice de Desarrollo Humano, PIB per cápita, Tasa de fertilidad en adolescentes, Tasa de mortalidad en menores de cinco años, Gasto en salud pública, Índice de educación, Población total, Población urbana) y su selección responde a criterios tales como el signo y la intensidad de la correlación, el tipo de relación de dependencia (lineal, así como otros modelos de ajuste), explicación de la dependencia de los datos (Barbancho, 1992), así como la posibilidad de explicar la dependencia en función de teorías previas sobre el contexto de los datos.

Al comienzo se discute el significado de las diferentes variables que conforman el fichero Excel que se utiliza en el proyecto y la forma en que se recogen los datos. Se centra la atención en el análisis del significado de la variable esperanza de vida, por ser la variable dependiente del proyecto, según el análisis de diferentes representaciones gráficas y estadísticas calculados de dicha variable. Cada futuro profesor dispone de una hoja de respuestas a las tareas del proyecto, que se entregan en diferentes etapas. En la primera etapa se responde por escrito a la interpretación de la variable dependiente y, antes de comenzar la segunda etapa, se debaten las soluciones con el fin de resolver posibles dificultades y desarrollar el conocimiento matemático de los participantes. En la segunda etapa se plantea analizar

la dependencia de la variable esperanza de vida en función de las variables escogidas previamente para el fichero Excel, como se explicó anteriormente. En particular, se pide estimar el coeficiente de correlación, ordenar las variables analizadas según su poder de predicción de la variable esperanza de vida, analizar si la correlación observada en cada caso es o no causal y ajustar una función a los datos, indicando el tipo de función que podría emplearse para dicha finalidad. Estas etapas del proyecto se desarrollaron en dos sesiones de dos horas cada una. En la última etapa, llevada a cabo en una sesión (dos horas de duración) se realizó la evaluación de la idoneidad didáctica del proyecto desarrollado.

Cabe decir que, una vez finalizadas las tareas de la segunda etapa del proyecto, se ofreció la posibilidad de realizar tareas de ampliación con otras variables disponibles en el servidor de las Naciones Unidas, así como actividades de interpretación de los parámetros de las funciones de regresión o uso de *applets* útiles en la enseñanza del tema.

## 5. Resultados y discusión

El análisis de las producciones de los futuros profesores permitió responder a nuestra problemática por investigar el conocimiento didáctico matemático en correlación y regresión de una muestra de futuros profesores en el Máster de Formación del Profesorado de Educación Secundaria y Bachillerato, en la especialidad de matemáticas, en España. Además, las diferentes fases del diseño metodológico de la experiencia ofreció a los futuros profesores un espacio de desarrollo profesional orientado a la enseñanza de la correlación y regresión, en forma de proyecto, en base a las tareas que ellos mismos habían resuelto y, por tanto, disponían de un contexto apropiado para su propia reflexión sobre la práctica.

En cuanto al conocimiento matemático sobre correlación y regresión, los resultados muestran buen desempeño de los futuros profesores en el proyecto implementado. Presentan buenas estimaciones del valor absoluto y signo de la correlación; ordenan de modo correcto o parcialmente correcto las variables, según su poder de predicción de la esperanza de vida; identifican la tendencia de los datos en la situación de ajuste de una función a los mismos; muestran buena identificación de la independencia estadística en diagramas de dispersión, en comparación con resultados en estudios previos; aunque muestran cierta dificultad cuando argumentan el tipo de dependencia entre las variables, como se explica con más detalle a continuación.

La actividad de argumentar la posible correlación en los datos nos permite observar las concepciones de los futuros profesores (por ejemplo, la concepción causal), así como sus creencias sobre la relación de la esperanza de vida con otras variables. Esta actividad requiere un conocimiento avanzado del tema, puesto que en la justificación se debe relacionar el contexto de los datos con la correlación y discriminar que correlación no siempre implica causalidad; es decir, que cuando dos variables covarían puede ser por dependencia indirecta, correlación casual o interdependencia. También es

necesario identificar los objetos matemáticos implicados en la situación, fundamentalmente la variable estadística bidimensional y su distribución, y distinguir la variable dependiente e independiente en la interpretación de la covariación.

En los resultados de nuestro estudio, se observa que las variables que los futuros profesores relacionan con mayor frecuencia de manera causal con la esperanza de vida fueron el Índice de desarrollo humano (que en realidad se trata de una relación de interdependencia, pues la esperanza de vida se incluye en el cálculo de esta variable) y la Tasa de mortalidad, seguida del PIB *per cápita* (cuya relación es indirecta y no causal, pues un país con PIB elevado tendrá mayor gasto en salud y educación y esto hace variar la esperanza de vida). Todos los futuros profesores consideran la Población total independiente de la variable esperanza de vida, como ocurrió en la tarea de estimar y asignar el signo al coeficiente de correlación entre estas variables.

En el análisis de las justificaciones de los futuros profesores a las variables que escogen encontramos, mayoritariamente, argumentos referidos a terceras variables que influyen en la esperanza de vida. En estos casos, los futuros profesores podrían no diferenciar la relación causal con la dependencia indirecta, como se muestra en el siguiente ejemplo de argumento, en el que además se relaciona incorrectamente la superpoblación con la esperanza de vida: "A mayor fecundidad de la población, más población habrá. Menor será la esperanza de vida por superpoblación."

Es igualmente frecuente el argumento enfocado en el crecimiento/decrecimiento de la variable independiente en la dependiente, mayoritariamente identificado en la variable Tasa de mortalidad. Sin embargo, esta explicación no es válida en general, pues una causa puede actuar sobre un efecto en forma no lineal. Como ejemplo, en el siguiente argumento el futuro profesor relaciona el crecimiento en el Índice de Desarrollo Humano con la esperanza de vida: "Cuanto más avances tiene una población o mayor desarrollo humano posee, mayor será la esperanza de vida."

En menor medida encontramos otros tipos de argumentos como: identificar el modo en que se calcula la esperanza de vida (en relación con la Tasa de mortalidad), la descripción de la influencia de la variable independiente en la dependiente, o mencionar de modo explícito la correlación en la justificación de la relación causal.

En cuanto al conocimiento didáctico de los futuros profesores, siguiendo la pauta de valoración de la idoneidad didáctica (Godino, 2013) como recurso en la evaluación y desarrollo de las facetas del conocimiento del profesor, concluimos que fue adecuado. En particular, en el análisis de sus respuestas a la valoración de la idoneidad epistémica del proyecto desarrollado se identifican los objetos matemáticos implícitos y explícitos en el proyecto y se aportan ejemplos en sus respuestas de distintos objetos matemáticos presentes en el estudio de la correlación y regresión. Se identifica la contextualización de las situaciones problema propuestas en el proyecto y la posibilidad de plantear otras nuevas; la riqueza de lenguaje implicado; y las posibilidades que ofrece el proyecto para relacionar objetos matemáticos implicados en el tema.

El análisis de las respuestas de los futuros profesores en la valoración de la idoneidad didáctica del proyecto desarrollado se completa con la asignación de puntuaciones numéricas, pudiendo así establecer niveles de calidad de respuesta como se describe a continuación:

- Nivel 0. Cuando no se responde a la valoración del descriptor o no se hace referencia al mismo.
- Nivel 1. Se hace referencia al descriptor de la idoneidad que se valora, pero se copia casi literalmente el descriptor, sin aplicarlo al proyecto analizado.
- Nivel 2. Se hace referencia y se aplica el descriptor, pero no se centra específicamente en el proyecto analizado sino en aspectos anecdóticos. También consideramos en esta categoría el caso en que el futuro profesor aplica una parte del descriptor correctamente y otra incorrectamente.
- Nivel 3. El futuro profesor hace referencia y aplica el descriptor al proyecto analizado, pero la aplicación se centra en un aspecto no recogido por el descriptor. Por ejemplo, si se le pide valorar la tipología de situaciones-problema propuestas en el proyecto y, en su lugar, valora propiedades o conceptos implicados.
- Nivel 4. Se hace una aplicación correcta del descriptor en forma consistente con el proyecto analizado, pero se razona mediante un único ejemplo. Por ejemplo, ser capaz de identificar correctamente un tipo de situación problema propuesto en el proyecto cuando se le pide valorar todas las situaciones-problema incluidas en el mismo y ejemplificar algunas.
- Nivel 5. El futuro profesor hace una aplicación correcta y consistente del descriptor según el proyecto analizado, y razona mediante dos o más ejemplos. Sería idéntico, pero más completo al caso anterior.

En todas las componentes de la idoneidad didáctica, el nivel medio alcanzado se sitúa (en una escala de 5 puntos) sobre el valor medio teórico (2,5). La media global obtenida por el total de participantes (3,6) supera en un punto el nivel medio teórico (2,5) y es mucho mejor que el mostrado por los futuros profesores de educación primaria en el trabajo de Arteaga (2011). Una revisión más extensa de los resultados de la investigación en el conocimiento didáctico matemático de la muestra de futuros profesores se encuentra en Gea (2014).

## **6. Conclusiones e implicaciones didácticas**

El interés por el desarrollo del sentido estadístico en la formación de nuestros estudiantes, tal como sugieren las directrices curriculares actuales (MECD, 2015), motiva el diseño de la propuesta de enseñanza que se presenta en este trabajo. Hoy en día existe una gran variedad de recursos tecnológicos

como la calculadora, hoja de cálculo, applets y programas de ordenador específicos, que facilitan la realización de cálculos y gráficos (Pratt, Davies y Connor, 2011) y contribuyen a la adquisición de ideas fundamentales. Así, en este trabajo se proporciona un ejemplo de enseñanza con proyectos haciendo uso de la hoja de cálculo, que podrán utilizar los futuros profesores con sus estudiantes para recorrer un ciclo estadístico completo (Batanero y Borovcnick, 2016).

El diseño del proyecto resultó de gran interés a los futuros profesores, quienes pudieron desarrollar el razonamiento estadístico al trabajar con datos reales, manejar la correlación y regresión con el recurso tecnológico de la hoja de cálculo Excel, así como poner a prueba su disposición ante el tema, pues se trabaja sobre el contexto de los datos y se plantean preguntas que potencian la capacidad de argumentar decisiones, así como potenciar la creatividad en plantear sus propias conjeturas e hipótesis sobre nuevas variables relevantes en la sociedad actual y ponerlas a prueba. La experiencia formativa resultó exitosa, puesto que conforme se recogía la información de sus respuestas, se iniciaba un debate en el que los futuros profesores compartían sus ideas. En este escenario es donde se observó un gran desarrollo del conocimiento de los futuros profesores, puesto que por lo general desconocían investigaciones didácticas sobre el tema, lo que permitió que fueran discutidas junto a la reflexión del formador, en base a las tareas que ellos mismos habían resuelto y, por tanto, disponían de un contexto apropiado para su propia reflexión sobre la práctica.

**Agradecimiento:** Proyecto de investigación PID2019-105601GB-I00 (AEI, FEDER) y grupo de investigación FQM126 de la Junta de Andalucía.

## Referencias

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

Arteaga, P., Batanero, C., Cañadas, G. R., y Gea, M. M. (2012). Evaluación del conocimiento especializado de la estadística en futuros profesores mediante el análisis de un proyecto estadístico. *Educação Matemática Pesquisa*, 14(2) 279-297.

Barbancho, A. G. (1992). *Estadística elemental moderna*. Barcelona. Ariel.

Batanero, C. y Borovcnik, M. (2016). *Statistics and probability in high school*. Londres: Sense Publishers.

Batanero, C., Burrill, G. y Reading, C. (2011). *Teaching statistics in school mathematics challenges for teaching and teacher education: A Joint ICMI/LASE study*. New York: Springer.

Batanero, C., Díaz, C., Contreras, J. M. y Roa, R. (2013). El sentido estadístico y su desarrollo. *Números*, 83, 7-18.

Batanero, C., Díaz, C. y Gea, M. M. (2011). Estadísticas de la pobreza y desigualdad. En C. Batanero y C. Díaz (Eds.), *Estadística con proyectos* (pp. 97-124). Granada: Departamento de Didáctica de la Matemática.

Batanero, C., Estepa, A. y Godino, J. D. (1997). Evolution of students' understanding of statistical association in a computer based teaching environment. En J. B. Garfield y G. Burrill, (Eds.), *Research on the role of technology in teaching and learning statistics. IASE Round Table Conference papers* (pp. 191-205). Voorburg, The Netherlands: Internacional Statistical Institute.

Batanero, C., Godino, J. D., y Estepa, A. (1998). Building the meaning of statistical association through data analysis activities (Research Forum). En A. Olivier y K. Newstead, (Eds.), *Proceedings of the 22nd Conference of the Internacional Group for the Psychology of Mathematics Education*, (pp. 221-236). Stellembosch, South Africa: Universidad de Stellenbosh.

Burrill, G. y Biehler, R. (2011). Fundamental statistical ideas in the school curriculum and in training teachers. En C. Batanero, G. Burrill y C. Reading (Eds.), *Teaching statistics in school mathematics. Challenges for teaching and teacher education - A joint ICMI/IASE study* (pp. 57-69). Dordrecht: Springer.

Casey, S.A. (2008). *Subject matter knowledge for teaching statistical association*. Tesis doctoral. Universidad de Illinois.

Casey, S.A. (2010). Subject matter knowledge for teaching statistical association. *Statistics Education Research Journal*, 9(2), 50-68. Disponible en <http://iaseweb.org/Publications.php>.

Casey, S. A. y Wasserman, N. H. (2015). Teachers' knowledge about informal line of best fit. *Statistics Education Research Journal*, 14(1), 8-35.

Castro-Sotos, A. E., Vanhoof, S., Van Den Noortgate, W. y Onghena, P. (2009). *The transitivity misconception of Pearson's correlation coefficient*. *Statistics Education Research Journal*, 8(2), 33-55.

Chapman, L. J. (1967). Illusory correlation in observational report. *Journal of Verbal Learning and Verbal Behavior*, 6(1), 151-155.

Estepa, A. (1994). *Concepciones iniciales sobre la asociación estadística y su evolución como consecuencia de una enseñanza basada en el uso de ordenadores*. Tesis doctoral. Universidad de Granada.

Gea, M. M. (2014). *La correlación y regresión en bachillerato: análisis de libros de texto y del conocimiento de los futuros profesores*. Tesis doctoral. Departamento de Didáctica de la Matemática. Universidad de Granada. Disponible en <http://www.ugr.es/~batanero/pages/librotesis.html>

- Gea, M., Batanero, C. y Roa, R. (2014). El sentido de la correlación y regresión. *Números*, 87, 25-35.
- Godino, J. D. (2009). Categorías de análisis de los conocimientos del profesor de matemáticas, *UNION*, 20, 13-31.
- Godino, J. D. (2013). Indicadores de la idoneidad didáctica de procesos de enseñanza y aprendizaje de las matemáticas. *Cuadernos de Investigación y Formación en Educación Matemática*, 11, 111-132.
- Godino, J. D. (2017). Construyendo un sistema modular e inclusivo de herramientas teóricas para la educación matemática. In J. M. Contreras, P. Arteaga, G. R. Cañadas, M. M. Gea, B. Giacomone, & M. M. López-Martín (Eds.), *Actas del Segundo Congreso Internacional Virtual sobre el Enfoque Ontosemiótico del Conocimiento y la Instrucción Matemáticos*. Recuperado de <http://enfoqueontosemiotico.ugr.es/civeos.html>
- Godino, J. D. (2019). Perspectiva de las investigaciones sobre educación estadística realizadas en el marco del Enfoque Ontosemiótico. En J. M. Contreras, M. M. Gea, M. M. López-Martín y E. Molina-Portillo (Eds.), *Actas del Tercer Congreso Internacional Virtual de Educación Estadística*. Disponible en [www.ugr.es/local/fqm126/civeest.html](http://www.ugr.es/local/fqm126/civeest.html)
- Godino, J. D., Batanero, C. y Font, V. (2019). The onto-semiotic approach: implications for the prescriptive character of didactics. *For the Learning of Mathematics*, 39(1), 37- 42.
- Godino, J. D., Giacomone, B., Batanero, C., y Font, V. (2017). Enfoque ontosemiótico de los conocimientos y competencias del profesor de matemáticas. *Bolema*, 31(57), 90-113.
- Hill, H. C., Ball, D. L., y Schilling, S. G. (2008). Unpacking pedagogical content knowledge: Conceptualizing and measuring teachers' topic-specific knowledge of students. *Journal for Research in Mathematics Education*, 39(4), 372-400.
- 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.
- Pino-Fan, L. y Godino, J. D. (2015). Perspectiva ampliada del conocimiento didáctico-matemático del profesor. *Paradigma*, 36(1), 87-109.
- Pratt, D., Davies, N. y Connor, D. (2011). The role of technology in teaching and learning statistics, En C. Batanero, G. Burrill, y C. Reading (Eds.), *Teaching statistics in school mathematics. Challenges for teaching and teacher education. A joint ICMI and IASE study* (pp. 97-107). New York: Springer.

Quintas, S., Ferreira, R. y Oliveira, H. (2015). O conhecimento didático de estatística de duas professoras de matemática sobre dados bivariados. *Bolema*, 29(51), 284-306.

Shulman, L. S. (1986). Those who understand: Knowledge growth in teaching. *Educational Researcher*, 15(2), 3-14.

Vásquez, C. y Alsina, A. (2015). El conocimiento del profesorado para enseñar probabilidad: Un análisis global desde el modelo del Conocimiento Didáctico-Matemático. *Avances de Investigación en Educación Matemática*, 7, 27-48.

Vásquez, C. y Alsina, A. (2017). Aproximación al conocimiento común del contenido para enseñar probabilidad desde el modelo del Conocimiento Didáctico-matemático. *Educación matemática*, 29(3), 79-108. Disponible en: <http://dx.doi.org/10.24844/em2903.03>

Wild, C. y Pfannkuch, M. (1999). Statistical thinking in empirical enquiry. *International Statistical Review*, 67(3), 221-248.