



VOL. 23, Nº2 (ABRIL-JUNIO, 2019)

ISSN 1138-414X, ISSNe 1989-6395

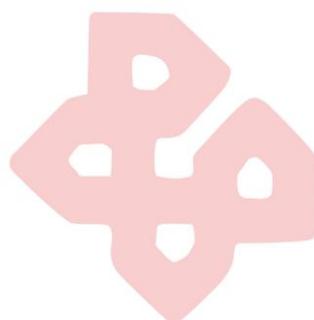
DOI: 10.30827/profesorado.v23i2.9693

Fecha de recepción: 26/01/2018

Fecha de aceptación: 13/07/2018

CONOCIMIENTO ESPECIALIZADO DE CORRELACIÓN Y REGRESIÓN EN FUTUROS PROFESORES DE EDUCACIÓN SECUNDARIA

Prospective secondary school teachers' knowledge of correlation and regression



*María M. Gea
Carmen Batanero
Pedro Arteaga
Antonio Estepa
Universidad de Granada
E-mail: mmgea@ugr.es; batanero@ugr.es;
parteaga@ugr.es; aestepa@ujaen.es
ORCID: 0000-0002-4189-7139*

Resumen:

La finalidad del trabajo fue evaluar el conocimiento especializado del contenido sobre correlación y regresión de 65 futuros profesores de Educación Secundaria Obligatoria y Bachillerato. Después de trabajar con un proyecto estadístico basado en datos de las Naciones Unidas, los participantes analizaron el componente epistémico de la idoneidad didáctica (Godino, 2013) de dicho proyecto. El análisis de las producciones escritas de los participantes permite asignarles un nivel en cada uno de los indicadores, comparar sus conocimientos en dichos indicadores e identificar ejemplos de su conocimiento especializado sobre la correlación y regresión.

Palabras clave: conocimiento especializado del contenido, correlación y regresión, evaluación.

Abstract:

The aim of this research was to evaluate the specialised content knowledge of correlation and regression in 65 prospective secondary and high school teachers. After working in a statistical workshop based on data from the United Nations, the participants analysed the epistemic component of didactic suitability (Godino, 2013) for this project. The analysis of the participants' written productions in this task served to assign them a level in each indicator, compare their knowledge of these indicators and identify examples of participants' specialised knowledge of correlation and regression.

Key Words: assessment, correlation and regression, specialised content knowledge.

1. Introducción

La correlación y regresión amplían la estadística descriptiva unidimensional, proporcionando la oportunidad de aplicar las funciones conocidas por el estudiante en una actividad de modelización, son contenidos ampliamente utilizados y constituyen la base de muchas técnicas estadísticas (Engel y Sedlmeier, 2011).

A pesar de su relevancia, la investigación previa advierte de errores y sesgos en su comprensión (Autor, 2001), siendo el más conocido la *correlación ilusoria* (Chapman, 1967), consistente en que los sujetos se guían por el contexto, más que por los datos, al estimar la correlación. Otras investigaciones describen concepciones erróneas (Autor, 1998; Castro-Sotos, Vanhoof, Van Den Noortgate y Onghena, 2009), como la *concepción causal*, donde el sujeto equipara correlación con causa-efecto; la *concepción unidireccional*, donde no se acepta una correlación inversa o la *concepción local*, cuando se estima la correlación con sólo una parte de los datos (Autor, 2008). Por otra parte, en el análisis de regresión, algunos estudiantes confunden la variable dependiente e independiente (Autor, 2001).

La enseñanza de este tema en España se realiza en el primer curso de Bachillerato, en las modalidades de Ciencias y de Humanidades y Ciencias Sociales. Además, se introduce la construcción e interpretación de diagramas de dispersión en cuarto curso de la Educación Secundaria Obligatoria, tanto en las Matemáticas orientadas a las enseñanzas aplicadas como en las orientadas a las enseñanzas académicas (MECD, 2015).

Puesto que una buena enseñanza requiere una preparación adecuada del profesorado, en esta investigación nos hemos interesado por los conocimientos de los futuros profesores sobre este tema. Para ello llevamos a cabo un estudio de evaluación, dentro de un taller formativo orientado también al desarrollo de dichos conocimientos.

En las siguientes secciones se desarrollan los fundamentos y metodología del estudio, se analizan sus resultados y se discuten las implicaciones para la formación de profesores.

2. Fundamentos

2.1. Marco teórico

En su análisis del conocimiento requerido por los profesores para abordar con éxito la enseñanza, Shulman (1986) describió el conocimiento del contenido, el conocimiento pedagógico del contenido y el conocimiento del currículo. Nuestro estudio se centra en uno de los componentes del conocimiento del contenido, considerado por Ball y sus colaboradores (Ball, Lubienski y Mewborn, 2001; Hill, Ball y Schilling, 2008) dentro de su modelo de conocimiento matemático para la enseñanza (MKT). Específicamente, consideramos el *conocimiento especializado del contenido*, que es el aplicado por el profesor para articular tareas de enseñanza (Hill, Ball y Schilling, 2008); es un conocimiento cualitativamente diferente del que poseen otros profesionales matemáticos o los profesores de otras materias (Scheiner et al., 2017). Un ejemplo se presenta al examinar y comprender procedimientos de resolución de problemas no usuales (Verdú, Callejo y Márquez, 2014).

La *idoneidad epistémica* de un proceso de estudio es parte de la idoneidad didáctica (Godino, 2009; 2013; Godino, Contreras y Font, 2006), que los autores introducen para evaluar situaciones de aprendizaje y enseñanza de las matemáticas y tiene los siguientes componentes: idoneidad epistémica, cognitiva, interaccional, mediacional, emocional y ecológica. La idoneidad epistémica consiste en la adecuación de los contenidos matemáticos enseñados respecto a un significado o contenido de referencia, fijado en las directrices curriculares y determinado por el significado del tema dentro de la matemática. Para valorarla, Godino (2009; 2013) propone una serie de indicadores que se describen en la Sección 6. En nuestro estudio se pedirá a los participantes utilizar dichos indicadores para valorar la idoneidad epistémica de un proyecto implementado en el aula; utilizaremos sus respuestas escritas para evaluar su conocimiento especializado del contenido.

La teoría de idoneidad didáctica ha sido ampliamente utilizada en la investigación educativa para evaluar programas didácticos y situaciones de aprendizaje y enseñanza de las matemáticas, a través de sus diversas facetas y componentes (por ejemplo, en Autor, 2012 y Godino, Rivas y Arteaga, 2012).

2.2. Antecedentes

Distintas investigaciones (ej., Nortes y Nortes, 2017) señalan la necesidad de evaluar el conocimiento matemático del profesorado. Sin embargo, son pocas las investigaciones previas que han analizado los componentes del conocimiento de los profesores sobre correlación y regresión. Sobre el conocimiento del contenido, el primer trabajo fue realizado por Autor (1998). Los autores realizaron una experiencia de enseñanza (incluyendo la correlación y regresión) con una muestra de 22 futuros profesores de Educación Primaria. Entre otras dificultades citan la identificación de la correlación inversa, y la confusión entre la variable dependiente e independiente en la regresión. También encuentran estrategias incorrectas en la interpretación de los diagramas de dispersión, por ejemplo, usar sólo parte de los datos.

Casey (2010) usa el marco teórico del MKT para analizar las clases desarrolladas por tres profesores en ejercicio. Realiza un análisis muy completo del conocimiento del contenido, abarcando incluso los aspectos inferenciales, que no tenemos en cuenta en este trabajo. En Casey y Wasserman (2015) se analiza el conocimiento de 19 profesores en ejercicio sobre la enseñanza del ajuste informal de una recta de regresión. Sin embargo, en estos trabajos no se diferencian los componentes común, avanzado y especializado de este contenido. Tampoco tienen como finalidad evaluar el conocimiento de los profesores, sino solo identificar el conocimiento que les sería necesario en la enseñanza del tema.

Quintas, Ferreira y Oliveira (2015), analizan el conocimiento pedagógico del contenido (en el modelo MKT) de dos profesoras de secundaria, con amplia experiencia, mientras explican la correlación y regresión en un curso de secundaria. Describen ejemplos del uso, por estas profesoras, de su conocimiento del contenido y la enseñanza y del conocimiento del contenido y el estudiante. Se observan también errores en el conocimiento común del contenido por parte de las profesoras; más concretamente, ninguna de ellas es consciente de la existencia de dos rectas diferentes de regresión y piden a los estudiantes usar siempre la recta de regresión de Y sobre X, independientemente de cuál sea la variable que se quiere predecir.

Aunque los trabajos citados fundamentan el nuestro y aportan puntos de comparación, ninguno realiza un estudio de evaluación con futuros profesores de secundaria y bachillerato o el análisis de los contenidos sobre correlación y regresión de las tareas por parte de los participantes. Estos serán puntos originales en nuestro trabajo. Nos apoyamos también en el estudio de Arteaga et al. (2012), quienes utilizan el análisis de la idoneidad epistémica de un proyecto estadístico realizado por futuros profesores de Educación Primaria para valorar su conocimiento especializado del contenido. Nosotros utilizaremos el mismo método para evaluar el conocimiento especializado sobre la correlación y regresión de los participantes en el presente estudio.

3. Método

Utilizamos una metodología exploratoria y mixta, común en otros trabajos de formación de profesores, como el de Pino-Fan, Godino y Font (2018). Puesto que se trabaja con futuros profesores, el estudio estuvo orientado no sólo a la evaluación de los conocimientos de los participantes, sino también a mejorar su formación. En consecuencia, los datos se recogen siguiendo el ciclo formativo propuesto por Godino, Ortiz, Roa y Wilhelmi (2008), que comienza con el trabajo de los futuros profesores en un proyecto estadístico, que pueden utilizar con sus futuros estudiantes, y continúa con una fase de análisis por parte de los futuros profesores de la idoneidad didáctica del proceso de enseñanza-aprendizaje basado en su experiencia personal con dicho proyecto.

3.1. Muestra y contexto

El estudio se desarrolló con 65 estudiantes que se preparaban para ser futuros profesores, dentro del Máster de Formación del Profesorado de Educación Secundaria y Bachillerato, especialidad de Matemática. Sólo el 56% de ellos eran licenciados en Matemáticas o Estadística y el resto habían cursado ingenierías, arquitectura u otras carreras de ciencias. Todos habían cursado una o más asignaturas de estadística y el 57% tenían experiencia de enseñanza. Los estudiantes estuvieron divididos en dos grupos de aproximadamente el mismo tamaño, que trabajaron con el mismo profesor y método.

La evaluación se llevó a cabo como parte de un taller formativo realizado dentro de una asignatura de Innovación docente en Matemáticas, en tres sesiones de 2 horas cada una. Las dos primeras sesiones estaban dirigidas a desarrollar y evaluar el conocimiento del contenido de la correlación y regresión de los participantes. En la tercera sesión, dedicada a desarrollar su conocimiento didáctico, los estudiantes analizaron la idoneidad didáctica del proyecto trabajado en las dos sesiones anteriores; en este trabajo sólo analizamos su análisis de la idoneidad epistémica del proyecto.

3.2. Taller previo a la toma de datos

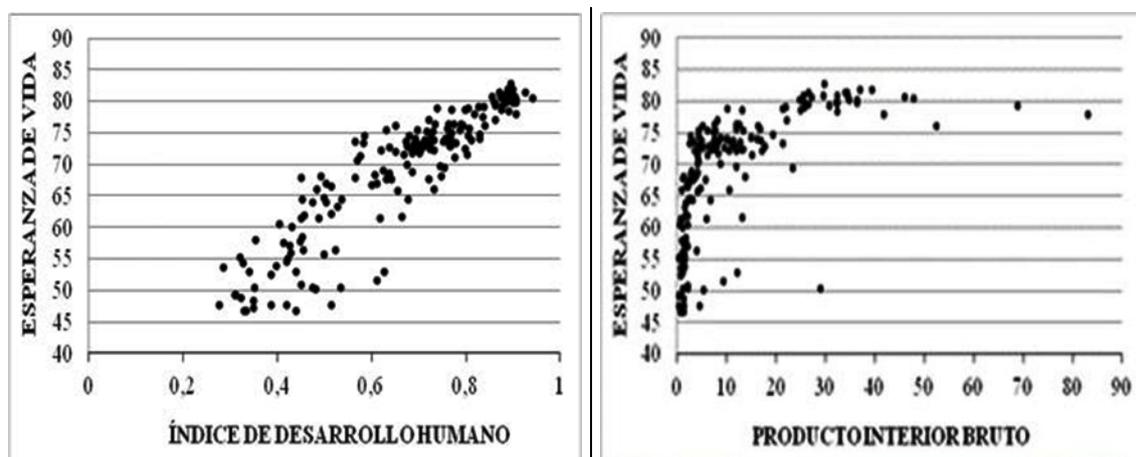
En la experiencia se trabajó con un proyecto estadístico, alrededor de una pregunta de investigación (*¿Cuáles son los factores que más influyen en la esperanza de vida al nacer en un país?*), que permitió plantear diferentes tareas. Los datos para elaborar los gráficos mostrados a los participantes se tomaron de los utilizados en la elaboración de los Informes sobre Desarrollo Humano por las Naciones Unidas, que están disponibles en su servidor (<http://hdr.undp.org/en/data>).

Las variables utilizadas en el proyecto son indicadores internacionales de desarrollo humano y se seleccionaron teniendo en cuenta los resultados de la investigación previa sobre comprensión de la correlación y regresión y los factores que inciden en la dificultad de estas tareas (Autor, 1998; Autor, 2001). No se limitó la dependencia entre las variables a la regresión lineal, sino se tomaron otros modelos de ajuste con funciones que se estudian en la Educación Secundaria. Igualmente se brindó la posibilidad de observar el signo positivo y negativo y los diferentes grados de intensidad en la correlación. Por otro lado, se utilizaron relaciones que pueden ser explicadas por causa y efecto, y otras debidas a interdependencia o dependencia indirecta. Finalmente, se consideraron variables en que la correlación observada en los datos contradice o confirma las posibles expectativas de los participantes.

Al comienzo del taller, el formador de profesores discutió con los futuros profesores el significado de las diferentes variables y la forma en que se recogen los datos. Se analiza en profundidad el significado de la variable esperanza de vida (variable dependiente en el proyecto) a través de diferentes representaciones gráficas (histograma, diagrama de frecuencias porcentuales acumuladas y gráfico de caja) y estadísticos de la misma. Una vez comprendido su significado, se fue proporcionando a los participantes, sucesivamente, cada una de las tareas del taller que aparecen resumidas en la Figura 1, en un formulario con espacio suficiente para

incluir la solución. Cada participante la respondió por escrito, individualmente, dejando para ello un tiempo suficiente. Una vez recogidas las respuestas a la tarea, se debatieron las soluciones, con el fin de resolver posibles dificultades y desarrollar el conocimiento del contenido de los participantes. Finalizada la recogida de datos, se dio la posibilidad, mediante el uso de Excel, para calcular las funciones de ajuste en la tarea 4, realizar nuevos análisis de correlación y regresión con otras variables y usar recursos de Internet, antes de pasar a la tercera sesión, donde se realizó la evaluación que se analiza en este trabajo y que se describe a continuación.

En las gráficas siguientes se representan los diagramas de dispersión de la esperanza de vida al nacer en función de ocho variables explicativas¹:



Utilizando cada una de estas gráficas, resuelve las siguientes tareas:

1. Asigna a cada gráfica una puntuación entre 0 (si no hay relación) y 1 (máxima intensidad de la relación) según la intensidad de esta relación. Asigna también un signo (+ o -) según la relación sea directa o inversa.
2. Ordena las ocho variables explicativas del proyecto, según su potencia (mayor a menor) para predecir mejor la esperanza de vida.
3. Indica cuáles de las variables explicativas tienen una relación causal con la esperanza de vida y explica por qué.
4. ¿Podríamos, para alguna de estas variables, hallar una función matemática para predecir, aproximadamente, la esperanza de vida a partir de la otra variable? Indica para cada gráfica un tipo de función que podría usarse con esta finalidad.

¹Se proporcionaron un total de ocho diagramas de dispersión, representando la variable esperanza de vida en función de cada una de las ocho variables explicativas. Además se proporcionó una hoja Excel con todos los datos del proyecto.

Figura 1. Tareas en el taller previo a la toma de datos

3.2. Material y categorías de análisis

Para valorar el conocimiento especializado de la correlación y regresión de los participantes se proporcionó a los futuros profesores una pauta de análisis de la idoneidad epistémica adaptada de Godino (2009), utilizada por Arteaga et al. (2012) y refinada por Godino (2013), pidiéndoles que respondiesen por escrito a cada una de

las preguntas para valorar el proyecto estadístico en el que habían trabajado en las sesiones anteriores.

Dicha tarea de reflexión es específica del trabajo del profesor, quien ha de ser capaz de valorar la idoneidad de cualquier actividad planteada a sus estudiantes. Para responderla, ha de identificar los contenidos matemáticos trabajados en el proyecto, lo que implica aplicar su conocimiento especializado del contenido (Verdú, Callejo y Márquez, 2014). Una vez recogidas las respuestas, se asignó en cada una un nivel, atendiendo al conocimiento mostrado en la aplicación del indicador. Estos niveles desarrollan los tres propuestos por Arteaga et al. (2012) para evaluar el conocimiento especializado sobre gráficos estadísticos de futuros profesores de Educación Primaria y son los siguientes, comunes a todos los indicadores:

- Nivel 0: Cuando el futuro profesor deja en blanco la respuesta a un indicador, **0** no es capaz de aplicarlo.
- Nivel 1: Se responde a la pregunta, aunque el participante se limita a copiar casi literalmente el indicador.
- Nivel 2: Se responde a la pregunta, aplicado el indicador, pero no se centra específicamente en el proyecto desarrollado, sino en aspectos anecdóticos o no estrictamente matemáticos. También consideramos en esta categoría el caso en que el futuro profesor aplica una parte del descriptor correctamente y otra es incorrecta.
- Nivel 3: Se aplica el indicador a contenidos matemáticos del proyecto, pero se centra en un aspecto no recogido por el indicador.
- Nivel 4: Se hace una aplicación correcta del indicador, utilizando contenidos matemáticos del trabajo en el proyecto, en forma consistente con la pregunta, pero se razona mediante un único ejemplo.
- Nivel 5. Se aplica correcta y consistentemente el indicador, utilizando contenidos matemáticos y razonando mediante dos o más ejemplos.

Tabla 1

Pauta de análisis de la idoneidad epistémica proporcionada a los participantes

Indicadores de la idoneidad epistémica	
Situaciones-problemas	11. ¿Te parece que los problemas que se presentan en el proyecto son útiles para contextualizar, aplicar y ejercitar los contenidos de correlación y regresión? ¿Por qué? 12. ¿Se proponen situaciones para que el estudiante de Bachillerato invente nuevos problemas? ¿Cuáles?
Lenguaje matemático	13. ¿Se usan diferentes modos de expresión matemática (verbal, gráfica, simbólica...), traducciones y conversiones entre los mismos? ¿Cuáles? 14. ¿Hay actividades de representación e interpretación del lenguaje matemático? ¿Cuáles?
Conceptos, propiedades, procedimientos	15. ¿Qué conceptos, propiedades y procedimientos sobre la correlación y regresión habría que explicar previamente para trabajar este proyecto? 16. ¿Se proponen tareas donde los estudiantes tengan que reconocer definiciones propiedades o procedimientos? ¿Cuáles?

Argumentos	17. ¿Son las explicaciones, comprobaciones y demostraciones adecuadas para Bachillerato?
	18. ¿Se incluyen situaciones donde el estudiante tenga que argumentar? ¿Cuáles?
Relaciones	19. ¿Se relacionan y conectan entre si los contenidos matemáticos (problemas, definiciones, propiedades, etc.) en el proyecto?

Los datos así obtenidos se codificaron y se calculó el porcentaje de estudiantes cuya respuesta corresponde a cada nivel en los diferentes indicadores, representando los resultados de los indicadores que corresponden a un mismo componente en gráficos de líneas que se presentan a continuación.

4. Resultados

4.1. Situaciones-problemas

Sobre las situaciones-problemas se propuso a los futuros profesores los dos indicadores I1 e I2 mostrados en la Tabla 1. En el proyecto aparecen varios problemas, en primer lugar, la interpretación de los diagramas de dispersión, que involucra la organización de los datos bidimensionales que originariamente vienen dados en forma de listado en la hoja Excel, entregada como parte del proyecto. Aunque el futuro profesor no ha elaborado dichos diagramas, al finalizar el proyecto se le propone explorar otras posibles relaciones entre las variables del fichero, lo que le lleva a trabajar este campo de problemas. Otros problemas serían analizar la existencia de relación entre variables y predecir una variable en función de otra, tanto en el análisis de la relación de la esperanza de vida con otras variables (Figura 1), como en las actividades de ampliación del proyecto.

En la respuesta al indicador I1 se espera que los futuros profesores indiquen que el proyecto planteado permite generar una muestra representativa de problemas de correlación y regresión o bien, que permite contextualizar otros contenidos estadísticos, como los de variable y distribución, medidas de posición central y dispersión y diferentes tipos de gráficos. Respecto al indicador I2, en el proyecto se pueden identificar momentos en que el estudiante podría inventar nuevos problemas. Por ejemplo, un estudiante que trabajase con este proyecto podría preguntarse cuáles variables afectan a otra de las variables consideradas. Asimismo, al trabajar con la hoja Excel, para obtener los gráficos y la línea de regresión, se plantean nuevos problemas, como por ejemplo: ¿Cómo elegir las escalas del gráfico de dispersión para visualizar mejor los datos?; ¿Sería mejor usar una función polinómica en vez de lineal? A continuación mostramos la forma en que se aplicaron los niveles 2 a 5, junto a algunos ejemplos.

Nivel 2. En el indicador I1 (problemas útiles para contextualizar los contenidos), justifican la pertinencia de las situaciones problema que se proponen en el proyecto, pero las que describen no son estrictamente matemáticas. En el indicador I2 (situaciones en las que el estudiante inventa nuevos problemas) se presentan ejemplos en forma muy genérica, sin indicar problemas específicos:

Creo que sí, porque al tomar los datos de la página web de la ONU, trabajamos en el contexto público que los estudiantes están acostumbrados a ver en los medios de comunicación (ME, I1).

En el proyecto de ayer no; se ofrece en el trabajo de ampliación (ChC, I2).

Nivel 3. Se aplica el indicador a contenidos matemáticos del proyecto y a la forma en que se trabajó con el mismo en el aula, pero sin centrarse en lo recogido en el descriptor. Por ejemplo, ChC, en lugar de valorar las situaciones-problema, describe una propiedad (grado de correlación) que se visualiza mediante el diagrama de dispersión. Por su parte, JMV analiza aspectos procedimentales, y MJG valora positivamente la argumentación. No se encuentran respuestas de futuros profesores al indicador I2 en este nivel:

Sí, porque la nube de puntos representa el grado de correlación entre dos variables (ChC, I1).

Sí, son útiles porque el estudio de fenómenos reales que afectan al estudiante, se hace de forma visual e interactiva en el ordenador (JMV, I1).

Sí, pues con estos problemas podemos hacer que el estudiante interprete los resultados que obtiene y los datos que te dan las diferentes gráficas, sacando conclusiones y observando las conclusiones de otros compañeros, enriqueciendo su propio conocimiento (MJG, I1).

Nivel 4. Se hace una aplicación correcta y consistente del indicador, utilizando situaciones problemas del proyecto, aunque se razona mediante un único ejemplo. Un ejemplo en el indicador I1 es MC, que indica la existencia de problemas de correlación. EGA, por su parte, es capaz de identificar situaciones-problema en las que un estudiante de Bachillerato pudiera inventar nuevos problemas (I2):

Sí, porque tienes que aplicar todos los conceptos por diferentes caminos. Tienes unas nubes de puntos para aplicar las cosas que sabes de correlación. Además está muy contextualizado y te permite ver su utilidad (MC, I1).

Sí, porque hay distintas variables que el estudiante puede analizar y relacionar como actividades complementarias (EGA, I2).

Nivel 5. Varios futuros profesores dan dos o más ejemplos de aplicación correcta y consistente de estos indicadores. Por ejemplo, en el indicador I1, PJ valora positivamente la contextualización de las situaciones-problema propuestas en el proyecto, e indica que permiten adquirir gradualmente los contenidos pretendidos. Otro ejemplo en el indicador I2 es ME, quien señala varias situaciones que permiten al estudiante inventar nuevos problemas:

Sí, porque comienzan hablándonos de una situación de la vida real y después se matematiza el contexto, aplicando la mayoría de los contenidos de correlación y regresión tales como nubes de puntos, coeficiente de correlación lineal, regresión, rectas de regresión, predicción. Y es un buen

ejemplo de que con una situación determinada hacer un estudio completo de variables bidimensionales y su relación entre ellas (PJ, I1).

Si. Por ejemplo en la actividad en la que se habla de qué variables sirven para predecir mejor la esperanza de vida. O en las actividades de ampliación cuando se pide estudiar “otras variables que te interesen e influyan en el desarrollo humano (ME, I2).

En la Figura 2 se presentan los resultados obtenidos en los dos indicadores correspondientes a las situaciones problemas. Observamos mejor resultado en el indicador I2 (invención de nuevos problemas) donde el 50,8% de los estudiantes llega a aplicar el indicador a nivel 4 y el 21,5% a nivel 5. Ello se debe a que la mayoría indicó la posibilidad de elegir nuevas variables para desarrollar otras nuevas situaciones problemas. Por el contrario, fue más deficiente la aplicación del indicador I1 (los problemas contextualizan el tema) puesto que, aunque el 44,6% alcanzaron los niveles 4 y 5, el 33% quedó en nivel 2, al hacer referencia al contexto de los datos y no tanto a las situaciones propuestas para trabajar la correlación y regresión.

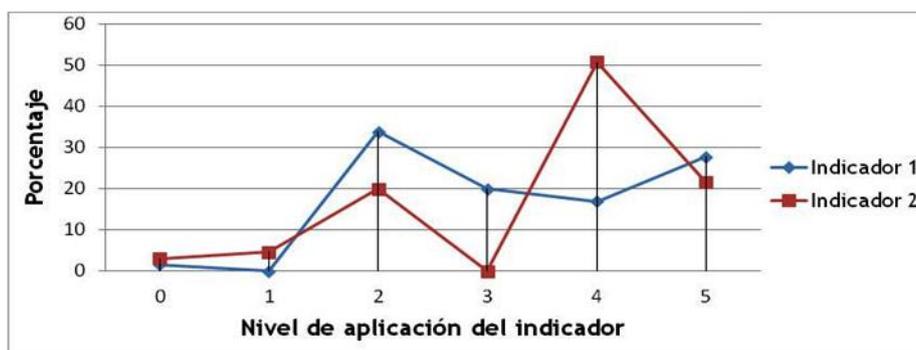


Figura 2. Porcentaje de participantes según nivel de valoración de situaciones- problema

4.2. Lenguaje

Sobre este componente de la idoneidad epistémica se plantearon dos preguntas correspondientes a los indicadores I3 e I4. Respecto al indicador I3 (variedad de lenguaje), en el proyecto se utilizaron términos como variable, valor, distribución, frecuencia, población, muestra, correlación, signo, intensidad, coeficiente de correlación, recta de regresión, pendiente o centro de gravedad. Se utilizó también lenguaje simbólico para expresar las líneas de regresión o los diferentes estadísticos que intervienen en el proyecto. Aparecen tablas de frecuencias absolutas y porcentajes, y diversos gráficos: histogramas, gráficos acumulativos, diagramas de dispersión, gráficos de caja, diagramas de burbuja.

Respecto al indicador I4 (actividades de interpretación y representación), Garfield y Ben-Zvi (2008) resaltan la importancia que el trabajo con proyectos tiene para desarrollar la capacidad discursiva de los estudiantes, que también se desarrolla en el proyecto que implementamos. Los futuros profesores tuvieron que interpretar diferentes gráficos estadísticos, símbolos y términos

estadísticos, así como utilizarlos al exponer sus resultados. Al trabajar con la hoja Excel se ha de interpretar el lenguaje específico, especialmente si se trabaja con algunas funciones; también la disposición tabular de la hoja donde cada fila corresponde a un país, y las columnas corresponden a las variables. Mostramos a continuación ejemplos de respuestas en que los futuros profesores aplican estos indicadores en los niveles 2 a 5.

Nivel 2. El futuro profesor hace referencia al lenguaje en el proyecto, aunque de modo anecdótico o parcial, en especial en el indicador I4, pues limita el uso del lenguaje matemático a la representación gráfica. Por ejemplo, MVC echa en falta actividades de representación en el proyecto (refiriéndose en exclusiva a las representaciones gráficas) y no considera otras sí incluidas al finalizar el proyecto como las realizadas con Excel para obtener el coeficiente de determinación (tipo numérico) y la función de ajuste a los datos (tipo simbólico), entre otras. Otros participantes, como ChC aplican el indicador con inconsistencias, pues identifican correctamente las actividades de interpretación y representación en los diagramas de dispersión y otros gráficos usados en el proyecto; no obstante, se responde (inconsistentemente con la primera parte) que no hay actividades de interpretación del lenguaje matemático. En otros casos, se refieren a aspectos particulares del proyecto, sin ceñirse matemáticamente a lo que el indicador pregunta, o se refieren al uso del lenguaje sin centrarse en el proyecto, como por ejemplo, MPF:

Sí, sobre todo de interpretación (interpretar diagramas de puntos, histogramas, diagramas de caja y bigotes,...) Quizás faltarían actividades de representación (sería una buena idea que el estudiante representara las gráficas) (MVC, I4).

Si. Todas las representaciones gráficas son actividades de representación e interpretación de los resultados estadísticos obtenidos. De interpretación del lenguaje matemático no hay (ChC, I4).

Para interpretar un diagrama, podemos pedirle que nos lo explique verbalmente ó con papel y lápiz (MPF, I3).

Nivel 3. Encontramos una respuesta de este nivel en un futuro profesor (FJM), quien aplica el indicador I4 al proyecto pero se refiere a aspectos procedimentales o propiedades de la correlación y regresión y no se centra en aspectos del lenguaje:

Si nos dan la ecuación de la recta de regresión sí que habría que hacer una interpretación matemática. Además, cuando nos dan la nube de puntos tenemos que interpretar si hay correlación o no (JFM, I4).

Nivel 4. Se aplica correcta y consistentemente el indicador mediante un único ejemplo. Las respuestas que generalmente dan los futuros profesores se refieren a la actividad 2, en la que se pedía interpretar distintos gráficos, o a la actividad inicial de análisis del significado de las variables del proyecto, donde se centra la atención en interpretar diversos estadísticos calculados de la variable dependiente, se manejan distintas representaciones, las traducciones entre las mismas, y se necesita

interpretarlas. A pesar de ello, tan sólo se alude a una actividad, sin precisar otras que aparecen en el proyecto:

Sí, sobre todo gráficas para representar lo que de forma verbal queremos transmitir sobre la parte de estadística que estamos tratando en el proyecto (PJB,I3).

Sí, a partir de la representación gráfica de la esperanza de vida con respecto a otra variable de la nube de puntos se realizó una interpretación sobre ellas, indicando los posibles valores del coeficiente lineal, por ejemplo (PP,I4).

Sí, a la hora de ver a qué función se aproxima a la gráfica con su determinada nube de puntos (MIC,I4).

Nivel 5. Cuando se dan dos o más ejemplos de aplicación correcta y consistente del indicador. Los estudiantes proporcionan varios ejemplos de lenguaje matemático en el indicador I3 (ChC, PJ) y varias actividades de interpretación y traducción en el I4 (PJ y PJB):

Sí, para los mismos datos se ofrecen gráficos como histogramas, caja, tabla y gráfico de frecuencias acumuladas (ChC,I3).

Sí se usan diferentes modos de expresión, por ejemplo los gráficos de histograma, de cajas, curva de frecuencias acumuladas. Además los datos concretos nos aportan rigor, con lo que podemos discernir qué modo de representación es más objetivo (PJ,I3).

Sí, por ejemplo cuando nos dan las nubes de puntos y los coeficientes de correlación lineal, tenemos que decidir qué tipo de correlación existe entre las variables a partir de la inspección visual y la interpretación del valor del coeficiente. Otro caso se da en la recta de regresión y la predicción de un valor a partir de otro, pues si no entendemos lo que representa dicha recta no podemos proceder al cálculo de predicciones (PJ,I4).

Sí, primeramente actividades donde representamos en gráficas los datos recogidos sobre las diferentes variables que, a priori, afectan a la esperanza de vida y, sobre todo, a través de las preguntas sobre las gráficas, tenemos actividades de interpretación (PJB,I4).

En la Figura 3 podemos observar el alto porcentaje de respuesta en los niveles 4 y 5 en los dos indicadores relacionados con la valoración del lenguaje matemático, en especial el I3 (diferentes modos de expresión matemática). Aun así, los futuros profesores muestran cierta dificultad al identificar tareas de representación e interpretación del lenguaje en el proyecto (I4), dándose un alto índice de respuesta en el nivel 2 (33,8%), pues los futuros profesores, en su mayoría, consideran únicamente como tareas de representación aquellas en que se pide al estudiante realizar o interpretar un gráfico.

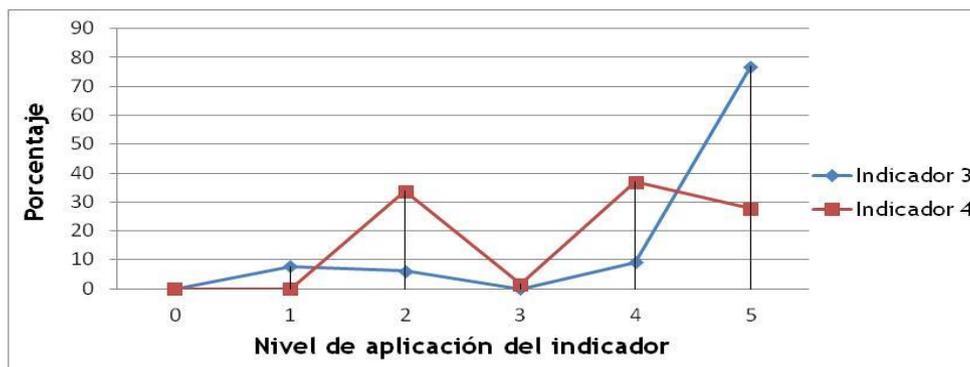


Figura 3. Porcentaje de participantes según nivel de valoración del lenguaje matemático

4.3. Conceptos, propiedades y procedimientos

El indicador 15 pregunta sobre los objetos matemáticos que el estudiante conoce antes del estudio de la correlación y regresión y debe emplear al trabajar con el proyecto. Es claro que se debe tener familiaridad con la variable estadística unidimensional, sus representaciones, las medidas de centralización y dispersión, las coordenadas cartesianas y funciones sencillas (polinómica, exponencial, logarítmica), así como las propiedades de estos conceptos. Los procedimientos requeridos son la lectura e interpretación de gráficos, elección de un valor representativo para una distribución, interpretación de las funciones citadas y manejo de la hoja Excel para el cálculo y la producción de gráficos. Respecto al indicador 16, hay varias situaciones en que los estudiantes pueden reconocer o aplicar definiciones, propiedades y procedimientos. Por ejemplo, elegir la medida de centralización más adecuada para representar la esperanza de vida o asignar el signo de la correlación, según la tendencia de un gráfico. Mostramos a continuación ejemplos de respuestas en que los futuros profesores aplican estos indicadores en los niveles 2 a 5, omitiendo el nivel 3 al no encontrarse respuestas en este.

Nivel 2. Los participantes en este nivel no suelen diferenciar los objetos matemáticos previos al estudio de la correlación y regresión de los propios del tema, como por ejemplo JPA. Otros, como MA, no señalan aspectos específicos del proyecto en sus respuestas. También encontramos respuestas inconsistentes en la valoración del indicador 16, como por ejemplo LT:

Distribuciones bidimensionales, Nube de puntos (diagrama de dispersión), recta de regresión, coeficiente de correlación (JPA, 15).

Yo creo que es necesaria una clase de repaso o explicación antes de realizar el proyecto, para poder afrontar todos los problemas de forma adecuada (MA, 15).

No se proponen tareas pero sí que tienen que reconocer definiciones y procedimientos de cálculo por ejemplo las propiedades de la recta de regresión y otros parámetros (LT, 16).

Nivel 4. Aplicación correcta y consistente utilizando un solo ejemplo de conceptos, propiedades o procedimientos que deben conocerse previamente para desarrollar el proyecto. En I5, la mayoría se refieren a propiedades o procedimientos relativos a la interpretación de la nube de puntos, como EGO, o a las medidas de centralización y dispersión como MM. Al aplicar I6, generalmente se hace referencia a las tareas iniciales del proyecto, como CM:

La nube de puntos (conceptos). Propiedades: $-1 \leq r \leq 1$. Procedimientos: interpretación de la nube de puntos (EGO, I5).

Como procedimientos se tendría que explicar los cálculos de los estadísticos (MM, I5).

Sí, por ejemplo, el cuadro de datos dado en la clase anterior de la media, moda, mediana... Para saber las relaciones entre ellos deben saber su significado (CM, I6).

Nivel 5. Se hace una aplicación correcta y consistente del indicador mediante dos o más ejemplos. Así por ejemplo, en el indicador I5, DG se refiere a la construcción de la nube de puntos; MRS, además, indica un procedimiento. En el indicador I6 algunos, como JG, se limitan a justificar las tareas mediante la riqueza de conceptos que se trabajan; otros se refieren a procedimientos y propiedades como ATL:

Por ejemplo, qué es una nube de puntos y cómo se construye, distribuciones marginales, tabla de frecuencias, media, varianza, ... (DG, I5).

Las definiciones y usos de la media, la moda, los cuartiles, los percentiles, la frecuencia absoluta, relativa y acumulada, la desviación típica y deben conocer y diferenciar los tipos de gráficos (diagrama de líneas, de barras, de cajas, ...) (MRS, I5).

Sí, reconocer las definiciones de las medidas de centralización, dispersión, coeficiente de correlación, regresión (JG, I6).

Sí, al tener que asignar una relación positiva-negativa, fuerte-débil e incluso el coeficiente de correlación, se le está pidiendo que reconozca los conceptos relacionados con esto. También lo es la tarea de ajustar nubes de puntos con un tipo de función (lineal, exponencial, logarítmica) (ATL, I6).

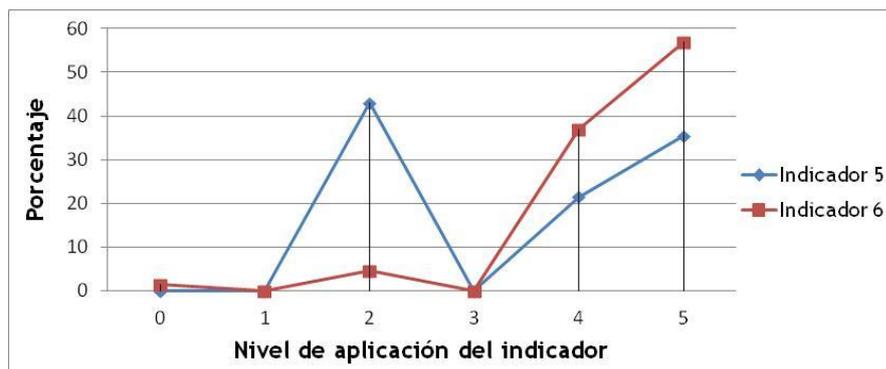


Figura 4. Porcentaje de participantes según nivel de valoración de reglas

Vemos, en la Figura 4, que la mayoría de los futuros profesores identifican correctamente los conocimientos, propiedades y procedimientos implicados en las tareas del proyecto, pues sus valoraciones en el indicador 16 se encuentran principalmente en los niveles 4 y 5.

4.4. Argumentos y relaciones

Sobre este componente se plantearon los indicadores 17 a 19. El proyecto en sí mismo no contiene argumentaciones (indicador 17), aunque en su desarrollo sí se hace uso de las mismas. Los futuros profesores han de valorar la claridad de las explicaciones que se dieron en las sesiones, así como en la corrección de las soluciones al proyecto. Las actividades planteadas son ricas en argumentación (indicador 18), pues los futuros profesores deben proponer una conclusión sobre la relación de diferentes variables con la esperanza de vida. También espera que lleven a cabo comprobaciones de propiedades mediante applets y se hace uso del programa Excel para añadir la línea de tendencia a la nube de puntos y comprobar, mediante el coeficiente de determinación, cuál es la función que mejor se ajusta a los datos.

De acuerdo con Godino (2013), los objetos matemáticos aparecen ligados entre sí en la resolución de problemas. Por ello, se incluye una última pregunta (19), que se refiere a las relaciones establecidas en el proyecto entre los diversos objetos matemáticos. En el proyecto se conectan los diferentes problemas planteados, por ejemplo, el problema de análisis de la relación entre variables, con el lenguaje gráfico (diagramas de dispersión), numérico y simbólico (coeficientes de correlación), tabular (tabla de estadísticos y/o frecuencias) y verbal, así como con los conceptos y propiedades asociados. Mostramos a continuación ejemplos de respuestas en que los futuros profesores aplican estos indicadores en los niveles 2 a 5.

Nivel 2. Encontramos este nivel principalmente en la valoración del indicador 17 (explicaciones y comprobaciones), pues los participantes aluden al nivel educativo sin centrarse en el proyecto: algunos se refieren al currículum (CMA); otros a la formación del estudiante (JMA) o su nivel de desarrollo cognitivo (PJ). También encontramos este nivel de respuesta, aunque en menor medida, en las valoraciones de los indicadores 18 e 19:

Considero que sí, ya que se adecuan con el currículo de Bachillerato. Aunque en muchas ocasiones no hay tiempo de dar los contenidos de Estadística y en esos casos no serían adecuadas para el Bachillerato (CMA, 17).

Sí, por supuesto. Es necesario que argumenten sus conclusiones con rigor matemático, en relación a la formación en la que están (JMA, 17).

Sí, puesto que a esta edad los adolescentes están en el último nivel de desarrollo cognitivo según Piaget, y han adquirido las competencias para que se les puedan mostrar estas comprobaciones y demostraciones y puedan asimilarlas (PJ,17).

Sobre todo las preguntas de “explicar justificadamente”, pero no las suelen hacer por la dificultad que conlleva (MI,18).

Sí. Porque al abordar el tratamiento estadístico de variables correspondientes a aspectos sociales interrelacionados en el marco institucional de las Naciones Unidas nos muestran un amplio abanico de tratamiento matemático y estadístico de dichas variables (GO,19).

Nivel 3. La respuesta se centra en otros aspectos y no en la argumentación o las relaciones entre objetos. Sólo lo encontramos en el indicador 17, donde los participantes valoran generalmente conceptos (CL) o situaciones problema (AL), en vez de argumentaciones:

Sí, son conceptos básicos que deben saber de años anteriores (media, varianza y desviación típica por ejemplo) (CL,17).

Sí, la tarea presentada es adecuada al nivel, ya que se encuentran en el Decreto de Enseñanzas Mínimas (AL,17).

Nivel 4. Aplicación correcta y consistente del indicador, indicandoun solo ejemplo de explicaciones, comprobaciones o demostraciones del proyecto o relaciones entre objetos matemáticos. En el indicador 17 estas respuestas suelen referirse a los conocimientos estadísticos que poseen los estudiantes (HB), así como al rigor matemático, como por ejemplo AMC. Lo encontramos también en el indicador 18, indicando tareas iniciales del proyecto (PP), o específicas de correlación y regresión (ME):

Sí, porque los estudiantes ya cuentan con una importante base matemática. Puede que no se haya profundizado mucho en Estadística durante la Secundaria, pero sí cuentan ya con unos sólidos conocimientos de funciones (HB,17).

Las explicaciones dadas eran poco matemáticas, pero creo que para un estudiante de Bachillerato podrían ser más formales (AMC, 17).

Sí, por ejemplo donde se pide explicar qué variables consideras que tengan relación causa-efecto con la esperanza de vida (ME,18).

Sí, por ejemplo, tenemos que tener en cuenta la moda y la mediana para justificar cuál es un valor representativo de nuestro estudio (PP, I9).

Nivel 5. Aplicación correcta y consistente de los indicadores referidos a la argumentación o relaciones entre objetos matemáticos trabajados en el proyecto mediante dos o más ejemplos. Los siguientes ejemplos indican la importancia de las comprobaciones visuales (BH), la existencia de ejercicios de argumentación (EGO) o de puesta en relación de varios objetos matemáticos (MC):

Sí en cuanto a que son aplicables a la realidad, fácilmente manipulables con las herramientas tecnológicas y muy visuales para facilitar su comprensión. Sin embargo, requiere el conocimiento de conceptos previos (BH, I7).

Sí, ya que no son ejercicios de cálculo sino de relacionar y deducir, por tanto hay que argumentar. Serían los de la nube de puntos o la interpretación de gráficos (EGO, I8).

Sí porque todo se puede relacionar, la nube de puntos con la recta de regresión y el coeficiente de correlación así como la dependencia entre variables (MC, I9).

En la Figura 5 se presentan los resultados obtenidos en la valoración de la argumentación y las relaciones entre contenidos matemáticos, con un alto porcentaje de respuesta en los niveles 4 y 5 en los indicadores I8 e I9 (cerca del 70% de los futuros profesores). En I7, las respuestas son menos precisas, ya que más de la mitad de las valoraciones (63,1%) no se centran en aspectos propios del proyecto o no valoran las explicaciones o los argumentos. Un aspecto positivo es que identifican correctamente tareas endonde se pide al estudiante que argumente (I8).

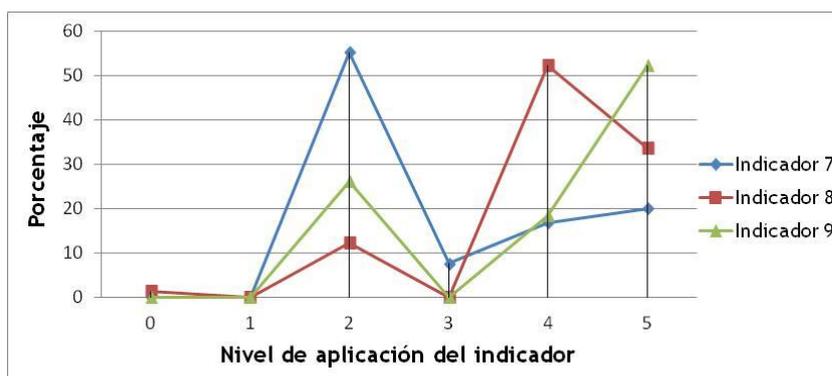


Figura 5. Porcentaje de participantes según nivel de valoración de argumentos y relaciones

4.5. Síntesis de resultados

Un resumen de todos los resultados planteados se presenta en la Tabla 2. Los resultados son buenos, ya que la media global (3.8) supera en un punto el nivel medio teórico (2.5) y es mucho mejor que la obtenida por Arteaga et al. (2012) (puntuación media 1.11) en su trabajo con futuros profesores de Educación Primaria.

Tabla 2
Medias y desviaciones típicas en los indicadores de la idoneidad epistémica

Indicador	Contenido	Media	D, Típica
I1.	Problemas contextualizados	3.3	1.3
I2.	Invencción de nuevos problemas	3.6	1.3
I3.	Variedad de lenguaje	4.4	1.2
I4.	Interpretación /traducción	3.6	1.2
I5.	Conocimientos previos requeridos	3.5	1.4
I6.	Tareas de reconocimiento de objetos	4.4	0.9
I7.	Nivel adecuado de argumentación	3	1.2
I8.	Argumentación por el estudiante	4	1
I9.	Relaciones entre objetos	3.9	1.4
Total		3.8	1.3

Otra aportación del análisis es la identificación, a partir de las respuestas de los estudiantes, de ejemplos de conocimientos especializados requeridos por parte del profesor para una buena enseñanza de la correlación y regresión, que mostramos en la Tabla 3, construida con ejemplos específicos de conocimientos manifestados por los participantes en su análisis de la idoneidad epistémica (niveles 4 y 5), algunas de las cuáles hemos presentado a lo largo del trabajo.

Tabla 3
Ejemplos de conocimientos especializados sobre la correlación en los futuros profesores

Objetos matemáticos	Ejemplos de conocimientos especializados
Situaciones-problemas	<ul style="list-style-type: none"> - Identifica ejemplos de problemas donde se contextualiza la correlación y regresión en el proyecto, como la búsqueda de la variable que mejor explique la variabilidad de la esperanza de vida. - Identifica situaciones en que los estudiantes tengan que proponer o modificar problemas relacionados con la correlación y regresión; en particular, el análisis de otras variables dependientes en el fichero de datos.
Lenguaje matemático	<ul style="list-style-type: none"> - Reconoce que los diagramas de dispersión, las ecuaciones de las funciones de ajuste, y los términos en el proyecto, son distintos modos de expresión matemáticas ligadas a la correlación y regresión. - Reconoce las representaciones gráficas que permiten visualizar mejor la correlación y regresión; por ejemplo, ve la necesidad de un cambio de escala en los datos. - Identifica situaciones de interpretación y traducción de representaciones de la correlación, como la búsqueda de un coeficiente de correlación o la estimación del mismo, a partir del diagrama de dispersión.
Conceptos, propiedades, procedimientos	<ul style="list-style-type: none"> - Reconoce los conceptos, propiedades y procedimientos asociados a la correlación y regresión que el estudiante ha de aplicar al resolver cada problema, relacionando, entre otros, la dependencia estadística y funcional, intensidad y sentido de la correlación, ajuste y modelo. - Reconoce cuáles objetos matemáticos han de ser explicados previamente y cuáles pueden ser construidos por el estudiante con la ayuda del profesor. - Identifica y propone nuevas tareas en las que el estudiante haya de reconocer o recordar conceptos, propiedades y procedimientos asociados a la correlación y regresión.
Argumentos	<ul style="list-style-type: none"> - Identifica argumentaciones adecuadas e inadecuadas para justificar una solución o una propiedad para el nivel del estudiante. - Identifica situaciones que requieran una argumentación por parte del estudiante, como por ejemplo, las referidas a la justificación de situaciones causales y no causales en el proyecto.

Relaciones	<ul style="list-style-type: none">- Identifica relaciones entre objetos matemáticos en el contexto del proyecto, como la relación entre la dirección de la dependencia y el signo del coeficiente de correlación o la dispersión de la nube de puntos y la intensidad de la correlación.- Es capaz de proponer nuevas tareas en que los estudiantes hayan de relacionar diferentes objetos matemáticos.
------------	--

7. Discusión

Los resultados de la evaluación del conocimiento de los futuros profesores en nuestro estudio fueron mucho mejores que los obtenidos por Arteaga (2011), quien también planteó el análisis de la idoneidad didáctica de un proyecto estadístico, experimentado por ellos mismos a futuros maestros de Educación Primaria. En su caso, sólo el 26,8% y 21,3% de los participantes llegan a aplicar los indicadores I1 e I2, correspondientes a las situaciones-problemas a nivel 3 (que acumularía nuestros niveles 3, 4 y 5).

Respecto al lenguaje, también fue más sencillo el indicador referido a la variedad del lenguaje (I3), con un 31,1% de futuros maestros que alcanza el nivel 3, que el referido a situaciones de interpretación y representación (I4) con sólo 17,6% en este nivel y en los dos indicadores con peores resultados que en nuestro estudio.

Nuestros participantes obtuvieron buenos resultados en la identificación de conceptos propiedades y procedimientos (I5 e I6), aunque algunos de ellos hacen referencia a los conceptos de correlación y regresión y no a los conocimientos previos. Los conceptos más citados fueron los de correlación, coordenadas cartesianas y representación de funciones, seguido de las medidas de centralización y dispersión, todos ellos básicos en modelización y otras técnicas estadísticas (Engel y Sedlmeier, 2011). Los conceptos de distancia y proyección de un punto a una recta, necesarios para la comprensión de la recta de regresión (Casey, 2010), sólo son identificados por dos participantes. Al contrario que en trabajo de Arteaga (2011), cuyos sujetos obtuvieron bajas puntuaciones en este componente, no se incluyen objetos no matemáticos u objetos no presentes en la misma, evidenciando los buenos conocimientos especializados de los futuros profesores en la identificación de los mismos (Verdú, Callejo y Márquez, 2014).

Igualmente fue buena la capacidad de identificar las situaciones que requieren argumentación y de establecer relaciones, superando también los resultados de Arteaga (2011), donde pocos sujetos logra el nivel 3 en estos indicadores.

El mejor desempeño de los futuros profesores *en nuestro estudio, razonable por su mayor preparación matemática, se confirma en la puntuación media global en el conjunto de indicadores, mucho más alta que la obtenida por Arteaga et al. (2012) en su trabajo con futuros maestros de Educación Primaria. En resumen, los participantes en nuestro estudio muestran su capacidad para identificar los objetos matemáticos implícitos o explícitos en una cierta situación de enseñanza, lo que

sería parte de su conocimiento especializado del contenido, según Hill, Ball, y Schilling (2008).

8. Implicaciones para la formación de profesores

La investigación que hemos descrito en este artículo informad el conocimiento especializado de la correlación y regresión de futuros profesores de Educación Secundaria y Bachillerato, complementando los antecedentes sobre el tema. Más concretamente, la actividad de valoración de la idoneidad didáctica permitió no solo evaluar, sino también mejorar su conocimiento especializado en correlación y regresión.

La información proporcionada sobre el nivel alcanzado por los participantes en la valoración de cada uno de los componentes de la idoneidad epistémica sobre este contenido permite identificar los puntos en que es conveniente mejorar su conocimiento especializado y orienta en este sentido la labor de los formadores de profesores.

Por otro lado, en el taller previo a la actividad de evaluación, descrito en forma resumida en el trabajo, se mostróa los participantes un ejemplo del trabajo con proyectos, que pueden utilizar con sus estudiantes para recorrer un estudio estadístico completo: problema, datos, análisis y conclusiones (Autor, 2016). Proporcionó igualmente una actividad de modelización, donde el interés se encuentra en analizar la dependencia y en caso de ser alta, encontrar el modelo de ajuste más adecuado (lineal, logarítmico, etc.). Estos últimos puntos forman parte del conocimiento didáctico del tema, que era escaso en los futuros profesores de la muestra, pues casi ninguno había trabajado previamente con datos reales o con la regresión utilizando Excel, ni la metodología de enseñanza por proyectos. Desconocían asimismo las investigaciones didácticas sobre el tema, que fueron discutidas con ellos a la vez que se debatieron las soluciones. Todas estas actividades posteriores se describen con detalle en Autor (2014).

Nuestra reflexión final es que la mejora de la enseñanza de la correlación y regresión depende de la formación de los profesores. Talleres y actividades de evaluación como las descritas en este trabajo, pueden contribuir a reforzar dicha formación.

Referencias bibliográficas

- 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., & Gea, M. M. (2012). Evaluación del conocimiento especializado de la estadística en futuros profesores mediante