

# Evaluación de conocimientos sobre el contenido de estadística descriptiva de futuros profesores de matemáticas

Felipe Ruz, Universidad de Granada (España)

Elena Molina-Portillo, Universidad de Granada (España)

José Miguel Contreras, Universidad de Granada (España)

---

## Evaluación de conocimientos sobre el contenido de estadística descriptiva de futuros profesores de matemáticas

### Resumen

*En este trabajo presentamos el proceso de construcción e implementación piloto de un instrumento para evaluar el conocimiento sobre estocástica de futuros profesores de matemática, en lo referido al contenido de estadística descriptiva. Las pruebas se llevaron a cabo en una muestra de 126 futuros profesores de matemática en España y Chile. Los resultados se organizan comenzando por indicadores sobre las propiedades psicométricas de los ítems analizados y de los resultados globales del cuestionario. Posteriormente, estudiamos conjuntamente los ítems que conforman cada categoría del conocimiento del contenido considerada, entre común o ampliado, buscando diferencias entre los grupos de la muestra. Se concluye que el instrumento cuenta con características aceptables, al igual que los reactivos analizados y se proyecta su aplicación reformada en el futuro.*

**Palabras clave.** Conocimiento del contenido; estadística descriptiva; formación de profesores de matemáticas; estudio piloto.

## Assessment of descriptive statistical knowledge of prospective mathematics teachers

### Abstract

*In this work we present the process of construction and pilot implementation of an instrument to evaluate the knowledge about stochastic of prospective mathematics teachers, in relation to descriptive statistics content. The tests were carried out in a sample of 126 prospective mathematics teachers, from Spain and Chile. The results are organized starting with some indicators on the psychometric properties of the items analysed and the overall results of the questionnaire. From here, we jointly study the items that conform each category of knowledge of the content in focus, between common or expanded, looking for differences amongst the sample groups. We conclude that the instrument has acceptable characteristics, as well as the items analysed, and its reformed application is projected into the future.*

**Keywords.** Content knowledge; descriptive statistics; mathematics teacher education; pilot study.

## 1. Introducción

Durante los últimos 30 años, la estocástica ha sido crecientemente valorada como un conocimiento necesario para que todo ciudadano educado pueda desenvolverse eficazmente en la sociedad de la información (Batanero y Borovcnik, 2016; Ben-Zvi y Makar, 2016). En consecuencia, muchos países como España, Estados Unidos y Chile, se han unido al movimiento reformista en torno a la enseñanza de la estocástica en la escuela, promoviendo su aprendizaje dentro del currículo de matemáticas durante toda la trayectoria escolar. De esta forma, se han establecido nuevas exigencias y desafíos para el profesorado responsable de esta tarea. Sin embargo, algunas investigaciones han reportado que los profesores no son lo suficientemente competentes, ni se sienten bien preparados para abordar esta temática en el aula de clase, lo que plantea una

problemática vigente con la formación estocástica de estos profesionales (Batanero, Burrill y Reading, 2011; Groth y Meletiou-Mavrotheris, 2018).

El profesor de matemáticas tiene que conocer las estocásticas escolares del nivel educativo donde imparte docencia, pero también debe poder articular esos conocimientos con los de niveles posteriores, es decir, manejar el conocimiento del contenido *per se* (Godino, Giacomone, Batanero y Font, 2017). Si bien resultados de investigación han mostrado que los conocimientos solo matemáticos/estocásticos no son suficientes para que el profesor organice, implemente y evalúe los procesos de enseñanza y aprendizaje, su formación debe tener en cuenta el conocimiento estocástico *per se*, ya que los conocimientos didácticos involucran al contenido disciplinar. Por tanto, nos interesamos en abordar el problema de evaluar el conocimiento sobre aspectos relevantes del contenido de estocástica de futuros profesores de matemáticas. Dada la extensión máxima permitida y las distintas etapas desarrolladas, en este trabajo profundizamos en lo referido a la estadística descriptiva.

La denominación “estadística descriptiva” se utiliza principalmente para organizar la difusión de los resultados de esta investigación, pero es también usada con frecuencia para diferenciar la etapa de exploración de los datos de aquella con fines inferenciales. En nuestro caso, al referirnos al *Contenido de Estadística Descriptiva* consideramos aquellos conceptos que emergen del proceso de análisis de datos donde se busca describir, organizar, representar y razonar sobre el centro, la dispersión y la distribución de los datos (Chan, Ismail y Sumintono, 2015) sin interés de generalizar o estimar alguna característica a un conjunto más grande o población.

De acuerdo con las consideraciones anteriores, en este artículo, nos proponemos el objetivo de describir el proceso de construcción y validación de un instrumento que evalúe el conocimiento del contenido de estadística descriptiva de futuros profesores de matemática. Con ello, buscamos describir el conocimiento declarado por una muestra piloto de 126 futuros docentes españoles y chilenos, destacando las principales diferencias obtenidas entre ellos.

## **2. Marco teórico**

Para evaluar el conocimiento de futuros profesores de matemáticas, es necesario establecer las bases teóricas que respaldan la perspectiva adoptada en la investigación. Tomamos un sistema teórico denominado Enfoque Onto-Semiótico del conocimiento y la instrucción matemáticos (EOS) (Godino, Batanero y Font, 2007), cuyo eje central es la modelización del conocimiento matemático/estocástico en su doble faceta epistémica (institucional) y cognitiva (personal), basada en una aproximación antropológica/pragmatista y ontosemiótica. En lo que sigue detallamos herramientas de interés para este estudio, entre ellas la noción de Configuración Ontosemiótica de prácticas, objetos y procesos matemáticos y el Modelo de Conocimientos y Competencias Didáctico-Matemáticas del profesor (CCDM).

### **2.1. Sistema de prácticas y configuración de objetos y procesos matemáticos/estocásticos**

Para el EOS, la actividad matemática es entendida como un sistema de prácticas (acciones de las personas ante cierto tipo de tareas o situaciones problemáticas) donde emergen e intervienen entidades u objetos matemático/estocásticos (Godino y Batanero, 1994). Además, dichas prácticas pueden ser idiosincrásicas de una persona y

nos referiremos a su *significado personal*; o compartidas en el seno de una institución, cuando hablaremos del *significado institucional*.

La distinción entre significado personal e institucional permite describir las diferencias que se podrían identificar entre el significado asignado por un estudiante a los objetos estocásticos y el acordado en la institución. Una institución (educativa o no) dirá que un sujeto entiende el significado de un objeto o que ha comprendido el significado de un concepto, si es capaz de llevar a cabo las diferentes prácticas que conforman el significado institucional del objeto. Así, el entendimiento personal de un objeto estocástico es deducido del análisis de las prácticas realizadas por la persona en la resolución de tareas problemáticas que son características de ese objeto, por lo que la selección de los ítems para diseñar un instrumento de evaluación es primordial.

## **2.2. Modelo de Conocimientos y Competencias Didáctico-Matemáticas (CCDM) del profesor de matemáticas**

Con herramientas del EOS, se establece el modelo para la formación de profesores de matemáticas denominado Modelo de Conocimientos y Competencias Didáctico-Matemáticas (CCDM) (Godino, 2009; Godino et al., 2017), sobre la base de modelos clásicos como el Modelo de Conocimiento del Contenido Pedagógico (PCK) de Shulman, el Modelo de Conocimiento Matemático para la Enseñanza (MKT) de Ball y colaboradores, y la Teoría de competencia en la enseñanza de la matemática de Schoenfeld y Kilpatrick. En el modelo CCDM, el conocimiento didáctico-matemático/estocástico del profesor se cimienta sobre el conocimiento del contenido *per se*, formado por conocimiento común (compartido con los estudiantes del nivel educativo donde el profesor enseña) y ampliado (compartido con los estudiantes de etapas educativas posteriores). Este conocimiento didáctico-matemático se articula y relaciona en torno a seis facetas: (1) Epistémica que abarca los conocimientos didáctico-matemáticos sobre el contenido; (2) Cognitiva que considera el conocimiento de cómo se aprenden y entienden las matemáticas; (3) Afectiva referida al conocimiento sobre las emociones, actitudes y creencias de los estudiantes sobre las matemáticas y su enseñanza; (4) Interaccional que incluye el conocimiento sobre la enseñanza de las matemáticas, organización de tareas e interacciones del aula; (5) Mediacional que abarca el conocimiento sobre los recursos y tiempos para potenciar el aprendizaje de los estudiantes; y (6) Ecológica referida al conocimiento sobre las relaciones de las matemáticas con otras disciplinas y los factores socio-profesionales, políticos y económicos que condicionan los procesos de instrucción.

## **3. Metodología**

Para evaluar la comprensión sobre los contenidos de estadística descriptiva en una muestra de futuros profesores de matemática, usamos herramientas y fundamentos del EOS, como el sistema de prácticas y la configuración de objetos y procesos, para describirlo semánticamente en términos de comportamientos observables o especificaciones entre él y la conducta (Batanero y Díaz, 2005). Estas especificaciones podrán describirse como campos de problemas representativos de las exigencias institucionales que caracterizan la comprensión sobre los contenidos de estocástica.

### **3.1. Instrumento**

A fin de determinar una definición semántica del constructo de interés, revisamos estudios sobre comprensión de conceptos de estadística descriptiva en el nivel universitario, con atención a los centrados en formación de profesores. Recopilamos

los ítems usados para la conformación de un banco de preguntas que sirvió de base para la construcción del instrumento. En Ruz, Díaz-Levicoy, Molina-Portillo y Ruiz-Reyes (2018) revisamos los aspectos a considerar para desarrollar la alfabetización, razonamiento y pensamiento estadístico en estos profesionales, donde identificamos que en términos del conocimiento del contenido se debe desarrollar la comprensión de conceptos estocásticos clave como variabilidad, distribución, representación y asociación, en estrecha relación con la estadística descriptiva. Sobre dichas nociones, destacamos los trabajos de Batanero et al. (2011), Ben-Zvi y Makar (2016) y Ben-Zvi, Makar y Garfield (2018) por organizar en algunos de sus capítulos los principales resultados de la investigación sobre el conocimiento de estos temas en futuros profesores. En consecuencia, podemos notar que la investigación se ha centrado en la comprensión de objetos específicos de la estadística descriptiva (promedio, gráficos, desviación estándar, etc.), como se refleja también en la revisión de Chan et al. (2015), lo que motivó a diseñar un instrumento que considerara varios tópicos de la estocástica y sus posibles interacciones, al estilo de los instrumentos Statistical Reasoning Assessment (SRA) (Garfield, 2003), Comprehensive of Assessment of Outcomes in a First Statistical Course (CAOS4) (DelMas, Garfield, Ooms y Chance, 2007) y Levels of Conceptual Understanding in Statistics (LOCUS) (Jacobbe, Case, Whitaker y Foti, 2014), pero dirigido específicamente a futuros profesores de matemáticas.

Dado el foco de este estudio en la realidad chilena, hemos realizado un análisis de contenido (Krippendorff, 1990) de las bases curriculares que orientan la formación matemática de estudiantes chilenos de los niveles 7 al 12 (entre 12 y 18 años), donde se incluyen aspectos de estocástica en el eje “Probabilidad y Estadística” (Ministerio de Educación Chile, 2009, 2015) que determinan el conocimiento común del futuro profesor con respecto a este contenido en la escuela. Por medio de la misma técnica se han analizado los principales documentos que orientan la formación estadística de profesores, determinando por medio de una colección de indicadores la Guía de Valorización de la Idoneidad Didáctica de procesos de Instrucción en Didáctica de la Estadística (GVID-IDE) (Ruz, Molina-Portillo y Contreras, 2019). Dicha guía, en su dimensión epistémica incluye la faceta contenido estadístico, donde se detallan aspectos del conocimiento de estocástica que la comunidad de educadores estadísticos y el Ministerio de Educación chileno considera para que el futuro profesor comprenda la disciplina (Ruz, Contreras, Molina-Portillo y Godino, 2018).

Tabla 1. Especificaciones sobre el contenido de estadística descriptiva del cuestionario

Contenido (C)	CC <sup>1</sup>
C1. Relación entre representaciones para explorar, resumir y describir patrones en datos unidimensionales	Común
C2. Relación entre representaciones para describir patrones de asociación entre dos variables	Ampliado
C3. Comparación de variabilidad entre grupos mediante análisis exploratorio de datos	Común
C4. Cuantificación de dispersión de los datos mediante indicadores estadísticos	Ampliado
C5. Deducción del comportamiento de distribución de datos a partir de indicadores estadísticos	Común
C6. Interpretación de coeficientes de asociación y de regresión lineal	Ampliado

<sup>1</sup> Categoría de conocimiento del contenido *per se* del modelo CCDM

De esta forma, aunque limitadamente, nos acercamos al significado institucional de referencia sobre lo que un futuro profesor de matemática debiese dominar tras culminar su proceso formativo en esta materia, es decir, luego de haber superado su etapa escolar y aquellas asignaturas disciplinares que fueron parte de su plan de

estudio. La Tabla 1 presenta especificaciones referidas a la comprensión del contenido de estadística descriptiva, donde diferenciamos si cada componente puede ser resuelto por medio de conocimiento común o ampliado del contenido, según el modelo CCDM.

Tras la planificación del contenido del instrumento, se retomó la revisión de la literatura para robustecer el banco de ítems y así abarcar aquellas especificaciones que no hubiesen sido cubiertas. Recopilamos preguntas de respuesta abierta y cerrada, sin embargo, según lo mencionado por Davies y Marriott (2010) dadas las restricciones del tiempo para responder, la longitud total del cuestionario y la amplia gama de ideas a evaluar, la mejor opción es utilizar preguntas cerradas. Así, para las preguntas abiertas se establecieron opciones de respuesta a partir de los resultados obtenidos en la investigación donde fueron utilizadas previamente y para las de respuesta cerrada se ajustó el número de opciones a cuatro. Por tanto, respecto a las especificaciones referidas a la comprensión de la estadística descriptiva (Tabla 1) comenzamos con un conjunto inicial de tres ítems por contenido, es decir, 18 en total.

Posteriormente, Osterlind (1998) recomienda analizar el grado en que la medición efectivamente represente la variable que se pretende medir (validez de contenido). Para ello, convocamos un panel de expertos con la intención de establecer un consenso acerca de la congruencia entre un ítem con el dominio de contenido que se pretende evaluar con él y además conocer cuál entre los ítems propuestos escogería cada uno para evaluar el contenido consultado. Seleccionamos un total de 10 investigadores iberoamericanos (incluidos España y Portugal) en educación estadística, todos doctores y cada uno con más de siete años de experiencia en investigación sobre los temas de estocástica y su enseñanza. Como resultado de este proceso, se estableció una primera versión del cuestionario con aquellos ítems con mayor puntuación media y menor dispersión entre las respuestas de los expertos. En este sentido, para cuantificar el grado de concordancia entre las puntuaciones otorgadas por los evaluadores, calculamos el índice Kappa de Fleiss que resultó de 0.6049, pudiendo concluir que la fuerza de acuerdo de esta evaluación es *considerable* (Landis y Koch, 1977, p. 165).

Tabla 2. Características de los ítems sobre la comprensión de la estadística descriptiva

Ítem-Contenido	Referencia	Conceptos	Lenguaje	Conocimiento
Ítem 1-C1	DelMas et al. (2007)	Medidas de centro Asimetría y sesgo Distribución y simetría	Natural- gráfico	Común
Ítem 2-C2	Allen (2006)	Correlación	Gráfico	Ampliado
Ítem 3-C3	DelMas et al. (2007)	Promedio Distribución y simetría	Gráfico	Común
Ítem 4-C4	DelMas et al. (2007)	Medidas de dispersión (desviación estándar) Histogramas	Gráfico	Ampliado
Ítem 5-C5	Elaboración propia	Medidas de posición (percentiles)	Natural	Común
Ítem 6-C6	Garfield (2003)	Correlación Causalidad	Natural	Ampliado

La Tabla 2 caracterizan los seis ítems, incluidos en el Anexo, seleccionados para evaluar el conocimiento sobre estadística descriptiva, detallando el contenido que evalúa cada uno según la Tabla 1, la referencia del ítem original, los conceptos estadísticos que emergen en su resolución, el lenguaje en que se presentan los datos (tabular, gráfico o natural) y la categoría de conocimiento del contenido necesario para resolverlo según el modelo CCDM.

### 3.2. Muestra

Una vez finalizada esta versión del cuestionario, estuvimos en condiciones de realizar una aplicación piloto a fin de comprobar empíricamente su utilidad para satisfacer los objetivos pretendidos e identificar aspectos que puedan ser mejorados. Estas pruebas se llevaron a cabo en una muestra no probabilística o intencional, de 126 futuros profesores de matemática, organizados en dos grupos según el país de residencia de los participantes. Nos interesamos en seleccionar dos realidades con modelos de formación diferentes, el consecutivo como en el caso de España y el concurrente en el de Chile (Muñiz-Rodríguez, Alonso, Rodríguez-Muñiz y Valcke, 2016). La principal diferencia es que en el primer modelo a la formación disciplinar le sucede en el tiempo la didáctica y pedagógica, y en el segundo estas convergen de manera simultánea en un mismo proceso. Así, nos preguntamos si existen disparidades entre ambas aproximaciones respecto al conocimiento del contenido, aun cuando en la literatura se ha reportado que no existen diferencias significativas en la aplicación de ambos modelos de formación (Valle, Jesús y Ayuso, 2011).

El primer grupo (G1) estuvo conformado por 84 estudiantes (35 mujeres y 49 hombres, con edades entre 21 y 50 años) del máster universitario de enseñanza secundaria obligatoria y bachillerato de una universidad española, pertenecientes a la especialidad de matemáticas o alguna doble titulación que considerase también matemáticas, a comienzos del curso académico 2018-2019. En cuanto a la formación estocástica previa de estos participantes, Muñiz-Rodríguez et al. (2016) destacan que en España esta es notablemente distinta según la universidad donde se haya estudiado, por lo que decidimos sintetizarla según el número de asignaturas cursadas y aprobadas durante el grado de ingreso al máster, como se muestra en la Tabla 3.

Tabla 3. Número (%) de asignaturas aprobadas en la formación inicial por G1

Nº asignaturas aprobadas	Nº (%) de participantes
1	31 (37.0)
2	8 (9.5)
3	9 (10.7)
4	10 (11.9)
5	6 (7.1)
No responde	20 (23.8)

El segundo grupo (G2) estuvo compuesto por 42 estudiantes (21 mujeres y 21 hombres, de entre 19 y 38 años) de pedagogía en matemática de una universidad chilena, que durante el segundo semestre de 2019 cursaban el cuarto de nueve semestres que considera su plan de estudios. Al tratarse de un programa concurrente, la formación disciplinar de estos participantes corresponde a dos asignaturas, una sobre probabilidad (Estadística 1) y otra sobre inferencia estadística (Estadística 2), cuyas tasas de aprobación se detallan en la Tabla 4.

Tabla 4. Número (%) de asignaturas aprobadas en la formación inicial por G2

Asignatura	Nº (%) participantes
Estadística 1	6 (14.3)
Estadística 2	36 (85.7)

Así aseguramos que los participantes de esta etapa de la investigación hayan cursado todas las asignaturas sobre el contenido de estocástica consideradas en su formación inicial. Además, acerca de la aplicación, todos los participantes recibieron y

completaron el cuestionario voluntaria e individualmente, dentro de una sesión planificada exclusivamente para este fin a mediados del segundo semestre de 2018.

## 4. Resultados y discusión

Analizamos características psicométricas de los reactivos presentados, como el grado de dificultad y discriminación, para luego referirnos a la consistencia interna del cuestionario completo. Posteriormente, agrupamos los ítems según la categoría del conocimiento del contenido que evalúan, entre común o ampliado, donde exponemos las frecuencias para cada opción de respuesta y evaluamos si existen diferencias entre los grupos de la muestra. También analizamos conjuntamente los ítems de cada categoría para comprobar si hay diferencias entre las puntuaciones obtenidas.

### 4.1. Algunas características psicométricas

Presentamos algunos indicadores sobre las propiedades psicométricas de los ítems analizados en este artículo y de los resultados globales del cuestionario. Calculamos el *índice de dificultad* de cada ítem, definido como la proporción de participantes que responde correctamente a un ítem particular (mientras menor, más difícil) y también su *índice de discriminación*, que describe la relación entre la dificultad de un ítem y las habilidades cognitivas de los participantes (Osterlind, 1998). Asignamos el valor 1 a aquellos ítems bien contestados y 0 a los equivocados, para así determinar el coeficiente de correlación biserial puntual entre la puntuación total del cuestionario y la obtenida en cada ítem. Escogimos este coeficiente por las condiciones del instrumento de evaluación, ya que cuando el ítem es inherentemente dicotómico y la puntuación total del cuestionario es una variable continua, el coeficiente de correlación del punto biserial es la estadística más útil para analizar el rendimiento relativo de un elemento entre dos grupos, los de alto y bajo rendimiento (Osterlind, 1998). Los resultados se recogen en la Tabla 5, acompañados del p-valor de haber contrastado la hipótesis nula de que las correlaciones obtenidas son iguales a 0.

Tabla 5. *Índices de dificultad y discriminación de los ítems analizados*

Ítem	Índice de Dificultad	Índice de Discriminación	p-valor
1	0.41	0.374	0.000*
2	0.39	0.410	0.000*
3	0.68	0.596	0.000*
4	0.37	0.483	0.000*
5	0.68	0.437	0.000*
6	0.25	0.202	0.001*

\*Resultados estadísticamente significativos a un nivel de 0.05

En cuanto al índice de dificultad, su valor oscila entre 0.25 (ítem 6) y 0.68 (ítems 3 y 5), de lo que concluimos que los ítems cuentan con un grado de dificultad moderada, es decir, ningún ítem es demasiado difícil (índice menor a 0.25), ni tampoco muy fácil (índice mayor a 0.75). Sobre los resultados de la capacidad de discriminación de los ítems, notamos que en los seis casos el p-valor es menor a 0.05, por lo que podemos rechazar la hipótesis nula y concluir que la correlación entre la puntuación total y la de cada ítem es significativa, es decir, distinta de cero. Además, según la clasificación propuesta por (Ebel y Frisbie, 1991, p. 232), los ítems 2, 3, 4 y 5 tienen una fuerza de discriminación *excelente*, por tener un índice mayor a 0.40, mientras que el ítem 1 es *adecuado* por estar entre 0.30-0.39 y el ítem 6 es *regular* por pertenecer al rango 0.20-0.29, sugiriendo una revisión del contenido de este último.

Ahora bien, dado que la fuerza de discriminación de un ítem se relaciona con su grado de dificultad en función del conocimiento sobre el contenido evaluado que es declarado por los examinados, es interesante analizar estos índices en conjunto, al menos para los casos conflictivos como el ítem 6. Este reactivo cuenta con el menor índice de dificultad y discriminación de la Tabla 3, es decir, la mayoría de los participantes responden incorrectamente y quienes lo hacen bien, no se corresponden con aquellos con mejor rendimiento en el cuestionario completo. Sin embargo, al observar el contenido que se busca evaluar, notamos que corresponde al sesgo de confundir una alta correlación entre dos variables con una relación de causalidad, aspecto que ha sido reportado en Estepa y Gea (2011) y explica este comportamiento. Entonces lo consideramos un ítem valioso para identificar la presencia de este sesgo en futuros profesores, por lo que decidimos mantenerlo en su estado actual.

Por otro lado, con el interés de dar información sobre la fiabilidad o consistencia interna del cuestionario, calculamos el alfa de Cronbach, que representa el grado en que covarían las puntuaciones de cada ítem particular con el resto de los ítems que componen el cuestionario. Obtuvimos un alfa de 0.564, que si bien no alcanza niveles ideales, Santiesteban (2009) reconoce como suficientes valores de alfa superiores a 0.5 en cuestionarios de conocimiento específico como en este caso. Además, Pedhazur y Pedhazur (1991) analizan el comportamiento del alfa de Cronbach para distintos tipos de cuestionarios, concluyendo que en aquellos casos donde la puntuación de los ítems es dicotómica, este valor se reduce considerablemente, estableciendo como aceptables valores entre 0.5 y 0.7, respaldando lo propuesto por Nunnally (1967). Por tanto, podemos afirmar que el instrumento cuenta con una consistencia interna adecuada para ser utilizado en futuros profesores de matemática, dado que la estimación de alfa obtenida con nuestros resultados pertenece a ese rango de valores.

#### 4.2. Evaluación del conocimiento común del contenido de estadística descriptiva

En esta sección, presentamos los resultados para los tres ítems que buscan evaluar el conocimiento sobre la estadística descriptiva escolar (significado institucional) que debería ser compartido entre el profesor de matemáticas y sus estudiantes (del nivel educativo donde se enseñe). Estos ítems evalúan los contenidos C1, C3 y C5 de la Tabla 1 respectivamente, de los que en la Tabla 6 mostramos las frecuencias (y porcentajes) de respuestas para cada alternativa, diferenciando entre los grupos que conforman la muestra y destacando con negrita la opción correcta en cada caso.

Tabla 6. Frecuencia (y porcentaje) de respuestas en ítems sobre el conocimiento común

Opción	Ítem 1 (C1)		Ítem 3 (C3)		Ítem 5 (C5)	
	G1	G2	G1	G2	G1	G2
A	18 (21.4)	7 (16.7)	2 (2.4)	1 (2.4)	3 (3.6)	1 (2.4)
B	<b>35 (41.7)</b>	<b>17 (40.5)</b>	<b>57 (67.9)</b>	<b>28 (66.7)</b>	3 (3.6)	3 (7.1)
C	9 (10.7)	1 (2.4)	4 (4.8)	1 (2.4)	4 (4.8)	7 (16.7)
D	11 (13.1)	8 (19.0)	10 (11.9)	7 (16.6)	<b>64 (76.2)</b>	<b>21 (50.0)</b>
No responde	11 (13.1)	9 (21.4)	11 (13.1)	5 (11.9)	10 (11.9)	10 (23.8)
Total	84	42	84	42	84	42

En general, notamos que en los tres ítems que conforman esta categoría las mayores frecuencias se obtienen en las opciones correctas, las que son bastante cercanas en términos relativos salvo en el ítem 6, donde un 76.2% de participantes de G1 responde bien, mientras que solo un 50% lo hace de la misma forma en G2.



El ítem 1 fue construido por DelMas et al. (2007) para evaluar la capacidad de relacionar las medidas de tendencia central con la representación gráfica de un conjunto de datos en estudiantes universitarios estadounidenses. La versión original de este ítem consideraba solo tres opciones de respuesta (por lo que fue adaptado levemente para ajustarle una cuarta alternativa, opción D) y en sus resultados, de un total de 730 respuestas, un 41.5% de ellas es correcta. En consecuencia, nuestros resultados coinciden con los reportados, ya que un 41% de nuestra muestra es capaz de relacionar distintas representaciones para identificar patrones en los datos (C1).

El ítem 3 corresponde a la fusión de las preguntas 11, 12 y 13 del test CAOS 4 (DelMas et al., 2007), que comparten un enunciado común y proponen distintas afirmaciones que deben ser catalogadas como válidas o no, para evaluar la capacidad de (1) comparar dos grupos considerando donde está la mayoría de los datos y centrándose en la distribución como entidades individuales; (2) comparar grupos mediante la comparación de las diferencias en promedio y (3) entender que para comparar dos grupos no es necesario que los tamaños muestrales sean iguales, respectivamente. En este caso, tomamos la situación propuesta y adaptamos que cada opción de respuesta evalúe alguno de los aspectos considerados inicialmente. En cuanto a nuestros resultados, estos son más bajos, ya que entre un 66.7% (G2) y un 67.9% (G1) responde correctamente, es decir son capaces de comparar la variabilidad existente entre grupos mediante un análisis exploratorio de datos (C3), mientras que en DelMas et al. (2007) de 753 respuestas válidas un 85.3% de ellas es correcta.

Finalmente, el ítem 5 fue propuesto por los autores para evaluar la capacidad de deducir el comportamiento de una distribución de datos a partir de indicadores estadísticos (C5). En este caso, si bien la mayoría de los participantes se inclinan por la alternativa correcta (opción D), un 76.2% de los estudiantes de G1 y un 50.0% de G2 responden bien, lo que al menos en términos relativos, presume posibles diferencias entre ambos, situación que pasaremos a estudiar a continuación.

Tabla 7. Estadísticos de resumen de la variable CCED

Variable	Cuartil 1		Cuartil 2		Cuartil 3		Prueba U M-W	
	G1	G2	G1	G2	G1	G2	Hip. Nula	p-valor
CCED	1	1	2	2	3	2	Igual distribución en G1 y G2	0.083

Con el interés de analizar conjuntamente los ítems que conforman esta categoría, asignamos 1 punto a las respuestas correctas y 0 puntos a las incorrectas, y definimos la variable *Conocimiento Común sobre Estadística Descriptiva* (CCED) como la suma de las puntuaciones obtenidas en los ítems anteriores. De esta forma, dicha variable tomará el valor 0 cuando un estudiante responda incorrectamente los tres ítems o 3 cuando lo haga correctamente en los tres casos. En la Tabla 7 presentamos estadísticos de resumen y contrastamos si la distribución de la variable CCED es la misma en ambos grupos, que, dados sus tamaños muestrales, se realiza por medio de la Prueba no paramétrica U de Mann-Whitney (M-W) para muestras independientes.

Con una significación de  $0.083 > 0.05$  no rechazamos la hipótesis nula, es decir, no hay evidencia muestral para asegurar que la distribución de la variable CCED es distinta en G1 y G2. Este resultado puede deberse a que el conocimiento que se evalúa es el compartido con estudiantes que han terminado su etapa escolar pre-universitaria y comparten, salvo excepciones menores, unos mismos requerimientos curriculares.

### 4.3. Evaluación del conocimiento ampliado del contenido de estadística descriptiva

Ahora presentamos los resultados de los tres ítems que buscan evaluar el conocimiento del contenido que es compartido con los estudiantes de etapas educativas posteriores a la escolar, sobre estadística descriptiva. Estos evalúan los contenidos C2, C4 y C6 de la Tabla 1 respectivamente. En la Tabla 8 organizamos las frecuencias (y porcentajes) de respuestas para cada opción, donde diferenciamos entre los grupos que conforman la muestra y destacamos con negrita la alternativa correcta en cada caso.

Tabla 8. Frecuencia (y porcentaje) de respuestas en ítems sobre el conocimiento ampliado

Opción	Ítem 2 (C2)		Ítem 4 (C4)		Ítem 6 (C6)	
	G1	G2	G1	G2	G1	G2
A	0 (0)	1 (2.4)	<b>36 (42.9)</b>	<b>10 (23.8)</b>	10 (11.9)	11 (26.2)
B	<b>24 (28.6)</b>	<b>12 (28.6)</b>	3 (3.6)	7 (16.7)	7 (8.3)	3 (7.1)
C	36 (42.9)	14 (33.3)	9 (10.7)	4 (9.5)	<b>29 (34.5)</b>	<b>3 (7.1)</b>
D	9 (10.7)	4 (9.5)	28 (33.3)	16 (38.1)	23 (27.4)	18 (42.9)
No responde	15 (17.9)	11 (26.2)	8 (9.5)	5 (11.9)	15 (17.9)	7 (16.7)
Total	84	42	84	42	84	42

En general y a diferencia del caso anterior, en los tres ítems de esta categoría no se observan resultados similares entre grupos. Mientras G1 acierta mayormente a las opciones correctas (salvo en el ítem 2), G2 tiene un comportamiento opuestos con más respuestas inclinadas a los distractores que a las opciones correctas en los tres casos.

En términos específicos, el ítem 2 fue propuesto por Allen (2006) para evaluar la capacidad de relacionar el valor numérico del coeficiente de correlación entre dos variables X e Y a partir de su representación en un gráfico de dispersión en estudiantes universitarios. La versión original de este ítem consideraba cinco opciones de respuesta. Fue adaptado levemente eliminando la opción que proponía un valor de -0.3, por considerarla poca atractiva para ser elegida (la nube de puntos tiene una clara tendencia al alza, lo que justifica que haya sido la menos seleccionada en su aplicación inicial). Entre los resultados originales, de un total de 305 respuestas, un 53% es correcta, mientras que en nuestro caso los resultados son más bajos, ya que solo un 28.6% de los participantes de G1 y G2 logran acertar a la mejor opción para relacionar representaciones para describir patrones de asociación entre dos variables (C2).

El ítem 4 fue construido por DelMas et al. (2007) para evaluar la capacidad para estimar correctamente y comparar las desviaciones estándar para diferentes histogramas, es decir, comprender que la más baja desviación estándar podría ser para un gráfico con la menor dispersión (típicamente), en estudiantes universitarios estadounidenses. La versión original de este ítem también consideraba cinco opciones de respuesta, por lo que fue adaptado eliminando la opción con menor frecuencia de elección en su implementación original. Entre sus resultados, de un total de 746 respuestas, un 34.3% de ellas es correcta, mientras que en nuestro caso G1 supera este valor con un 42.9% y G2 es menor con un 23.8% de los resultados obtenidos.

El ítem 6 fue propuesto por Garfield (2003) para analizar la capacidad de distinguir entre correlación y causalidad; así, evalúa el conocimiento de que una fuerte correlación entre dos variables no significa que una cause la otra. Con la intención de comparar la aplicación del instrumento SRA (de donde es parte este ítem) entre varias muestras, Tempelaar (2004) estandariza los resultados según porcentaje de respuestas correctas de cada ítem, obteniendo que este ítem es abordado correctamente por un 52% de 245 estudiantes universitarios americanos y un 65% de 267 estudiantes universitarios taiwaneses. Nuestros resultados son más desalentadores ya que por un lado G1 alcanza un 34.5% de respuestas correctas y G2 logra un 7.1%, inclinándose

mayormente por el distractor D (42.9%), que es el error de confundir una alta correlación entre dos variables con una relación de causalidad. Lo anterior parece suponer que existen diferencias entre G1 y G2 respecto al conocimiento ampliado.

Al igual que en el apartado previo, con la intención de analizar conjuntamente los ítems que conforman esta categoría, definimos la variable *Conocimiento Ampliado sobre Estadística Descriptiva* (CAED) como la suma de las puntuaciones obtenidas en los ítems previos. De esta forma, en la Tabla 9 presentamos algunos estadísticos de resumen y contrastamos si la distribución de la variable CAED es la misma en ambos grupos, por medio de la Prueba U de M-W para muestras independientes.

Tabla 9. Estadísticos de resumen de la variable CAED

Variable	Cuartil 1		Cuartil 2		Cuartil 3		Prueba U M-W	
	G1	G2	G1	G2	G1	G2	Hip. Nula	p-valor
CAED	0.25	0	1	0	2	1	Igual distribución en G1 y G2	0.000*

\* $p < 0.05$

Con una significación de  $0.000 < 0.05$  rechazamos la hipótesis nula, es decir, hay evidencia muestral para asegurar que la distribución de la variable CAED es distinta en G1 y G2. Este resultado debe interpretarse según el contexto de los participantes. La formación universitaria de los estudiantes de G1 (España) y de G2 (Chile) difiere; mientras que los primeros han finalizado algún grado antes de la realización del máster que los habilita para hacer clases de matemática en secundaria y bachillerato, los segundos cursan la especialización de profesor de matemáticas como parte del grado.

#### 4.4. Evaluación del conocimiento del contenido *per se* sobre estadística descriptiva

Finalmente, a modo de síntesis, analizamos globalmente estos seis ítems para comprobar si existen diferencias entre la puntuación obtenida según ambas categorías de conocimiento. Para ello, contrastamos si las medianas de las diferencias entre CCED y CAED calculadas en toda la muestra ( $n = 126$ ) son iguales a 0, por medio de la Prueba de rangos de Wilcoxon para muestras relacionadas. Sus resultados, junto a algunos estadísticos de resumen se presentan en la Tabla 10.

Por tanto, con un p-valor de  $0.000 < 0.05$  rechazamos la hipótesis nula, es decir, existen diferencias significativas entre conocimiento común y ampliado declarado por los participantes. Notamos que aparentemente se obtiene mejor resultado en el CCED, ya que desde el cuartil 2 (mediana) la diferencia CCED-CAED es 1, esto es, mayor a cero, reflejando una mayor deficiencia en el conocimiento ampliado del contenido.

Tabla 10. Estadísticos de resumen de las diferencias entre CCED y CAED

Variable	Cuartil 1	Cuartil 2	Cuartil 3	Prueba de Wilcoxon	
				Hip. Nula	p-valor
CCED-CAED	0	1	1	La mediana de las diferencias es 0	0.000*

\* $p < 0.05$

## 5. Proyecciones

En este trabajo, hemos reportado el proceso de construcción e implementación piloto de un instrumento para evaluar aspectos sobre el conocimiento del contenido de estocástica para futuros profesores de matemática, en lo referido a la estadística descriptiva. Motivados por la problemática vigente de la formación estadística deficiente de profesores, hemos elaborado una herramienta para evaluar el conocimiento del contenido en estos profesionales, con el cual esperamos recoger

información sobre el estado de esta etapa formativa. Posterior a su construcción, lo hemos implementado en una muestra piloto de estudiantes para profesor de matemática de España y Chile, cuyos resultados para aquellos ítems que evalúan el conocimiento sobre estadística descriptiva se detallaron a lo largo de este trabajo.

En cuanto a las características psicométricas del instrumento y los ítems reportados, pudimos concluir un nivel de dificultad medio y una fuerza de discriminación adecuada para cada reactivo, lo mismo para garantizar como aceptable la consistencia interna del cuestionario completo. Al agrupar aquellos ítems que evalúan el conocimiento común y los que hacen lo mismo respecto al conocimiento ampliado, observamos que existen diferencias significativas entre los resultados obtenidos por ambos grupos para el segundo de ellos, lo que puede deberse a las diferencias en la formación recibida previamente por los participantes de cada grupo. Aunque nuestro interés es simplemente describir el conocimiento declarado por los participantes y no hacer ningún tipo de generalización, estos resultados son opuestos a aquellos donde se asegura la inexistencia de diferencias significativas entre los modelos concurrente y recurrente para la formación del profesorado (Valle et al., 2011). Además, al comparar las respuestas obtenidas en toda la muestra piloto, obtuvimos que las puntuaciones de los ítems que evalúan ambas categorías del conocimiento del contenido *per se* son estadísticamente distintas, mostrando mayor deficiencia en lo referido al conocimiento del contenido ampliado. Esto establece un desafío de mejora en la formación disciplinar del profesorado, ya que el conocimiento ampliado es el desarrollado durante la estadía de estos sujetos en la universidad.

De esta forma, proyectamos una nueva implementación del instrumento en una muestra de futuros profesores de matemática chilenos, con la cual esperamos contar con información sobre este aspecto de la formación del profesorado. Lo anterior se vuelve de particular interés en esta población, ya que desde el año 2017 las autoridades chilenas organizan anualmente la denominada Evaluación Nacional Diagnóstica (END) dirigida a todos los estudiantes para profesor que se encuentren en los doce meses anteriores al egreso, con el objetivo de diagnosticar la formación de los nuevos docentes y generar información útil para los programas académicos (Centro de Perfeccionamiento Experimentación e Investigaciones Pedagógicas, 2018). Sin embargo, en los últimos resultados de la END, de 612 estudiantes para profesor de matemática que rindieron la prueba de Conocimientos Disciplinarios y Didácticos, los temas de geometría y datos y azar (estocástica) son lo más deficientes. Por tanto, consideramos que, en la próxima aplicación del instrumento, este nos entregará información específica sobre los conocimientos del contenido de estocástica con que cuentan actualmente estos profesionales de la educación matemática. Y con ella, se espera identificar aquellos tópicos más descendidos, de manera que en un futuro no muy lejano puedan desarrollarse mejoras a la formación específica del profesorado, que les permita responder satisfactoriamente a las exigencias de su campo profesional.

### **Agradecimientos**

Proyecto B-SEJ-063-UGR18, Grupo de investigación PAIDI 622 (Junta de Andalucía, España) y Comisión Nacional de Investigación Científica y Tecnológica (CONICYT) de Chile a través de beca de doctorado en el extranjero (Folio 72170025).

### **Referencias**

Allen, K. (2006). *The statistics concept inventory: Development and analysis of a cognitive assessment instrument in statistics*. Norman, EEUU: University of

Oklahoma.

- Batanero, C. y Borovcnik, M. (2016). *Statistics and probability in high School*. Rotterdam, Holanda: Sense Publishers.
- Batanero, C., Burrill, G. y Reading, C. (Eds.) (2011). *Teaching statistics in school mathematics: Challenges for teaching and teacher education. A Joint ICMI/IASE Study*. Nueva York: Springer.
- Batanero, C. y Díaz, C. (2005). Análisis del proceso de construcción de un cuestionario sobre probabilidad condicional. Reflexiones desde el marco de la TFS. En D. Ordoñez, C. Batanero y A. Contreras (Eds.), *I Congreso Internacional sobre Aplicaciones y Desarrollos de la Teoría de las Funciones Semióticas* (pp. 13-36). Jaén: Universidad de Jaén.
- Ben-Zvi, D. y Makar, K. (Eds.) (2016). *The teaching and learning of statistics. International perspectives*. Cham, Suiza: Springer.
- Ben-Zvi, D., Makar, K. y Garfield, J. (Eds.) (2018). *International handbook of research in statistics education*. Cham, Suiza: Springer.
- Centro de Perfeccionamiento, Experimentación e Investigaciones Pedagógicas (2018). *Resultados Nacionales Evaluación Nacional Diagnóstica de la Formación Inicial Docente 2017*. Santiago, Chile: CPEIP.
- Chan, S. W., Ismail, Z. y Sumintono, B. (2015). Assessing statistical reasoning in descriptive statistics: A qualitative meta-analysis. *Jurnal Teknologi*, 72(2), 1-6.
- Davies, N. y Marriott, J. (2010). Assessment and feedback in statistics. En P. Bidgood, N. Hunt y J. Flavia (Eds.), *Assessment methods in statistical education: An international perspective* (pp. 3-19). Londres: John Wiley & Sons.
- DelMas, R., Garfield, J., Ooms, A. y Chance, B. (2007). Assessing students' conceptual understanding after a first course in statistics. *Statistics Education Research Journal*, 6(2), 28-58.
- Ebel, R. y Frisbie, D. (1991). *Essentials of educational measurement* (5th ed.). Nueva Jersey, NJ: Prentice-Hall.
- Estepa, A. y Gea, M. M. (2011). Conocimiento para la enseñanza de la asociación estadística. En J. Ortiz (Ed.), *Investigaciones actuales en educación estadística y formación de profesores* (pp. 23-40). Granada: Departamento de Didáctica de la Matemática.
- Garfield, J. (2003). Assessing statistical reasoning. *Statistics Education Research Journal*, 2(1), 22-38.
- Godino, J. D. (2009). Categorías de análisis de los conocimientos del profesor de matemáticas. *Revista Iberoamericana de Educación Matemática*, 20, 13-31.
- Godino, J. D. y Batanero, C. (1994). Significado institucional y personal de los objetos matemáticos. *Recherches en Didactique des Mathématiques*, 14(3), 325-355.
- Godino, J. D., Batanero, C. y Font, V. (2007). The onto-semiotic approach to research in mathematics education. *ZDM-Mathematics Education*, 39(1-2), 127-135.
- 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.

- Groth, R. y Meletiou-Mavrotheris, M. (2018). Research on statistics teachers' cognitive and affective characteristics. En D. Ben-Zvi, K. Makar y J. Garfield (Eds.), *International handbook of research in statistics education* (pp. 327-355). Cham, Suiza: Springer.
- Pedhazur, E. y Pedhazur, L. (1991). *Measurement, design, and analysis*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Jacobbe, T., Case, C., Whitaker, D. y Foti, S. (2014). Establishing the validity of the LOCUS assessments through an evidenced-centered design approach. En K. Makar, B. de Sousa y R. Gould (Eds.), *Proceedings of the Ninth International Conference on Teaching Statistics* (pp. 1-6). Flagstaff, AZ: IASE.
- Krippendorff, K. (1990). *Metodología de análisis de contenido: Teoría y práctica*. Barcelona: Paidós.
- Landis, J. R. y Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, 33(1), 159-174.
- Ministerio de Educación Chile (2009). *Currículum. Objetivos fundamentales y contenidos mínimos obligatorios de la Educación Básica y Media*. Santiago: MINEDUC.
- Ministerio de Educación Chile (2015). *Bases curriculares 7° a 2° medio*. Santiago: MINEDUC.
- Muñiz-Rodríguez, L., Alonso, P., Rodríguez-Muñiz, L. y Valcke, M. (2016). ¿Hay un vacío en la formación inicial del profesorado de matemáticas de secundaria en España respecto a otros países? *Revista de Educación*, 372, 111-140.
- Nunnally, J. (1967). *Psychometric theory*. Nueva York: McGraw-Hill.
- Osterlind, S. J. (1998). *Constructing test items: Multiple-choice, constructed-response, performance, and other formats*. Dordrecht, Holanda: Kluwer.
- Ruz, F., Contreras, J. M., Molina-Portillo, E. y Godino, J. D. (2018). Idoneidad epistémica de programas formativos sobre didáctica de la estadística. En L. Rodríguez-Muñiz y otros (Eds.), *Investigación en Educación Matemática XXII* (pp. 515-524). Gijón: SEIEM.
- Ruz, F., Díaz-Levicoy, D., Molina-Portillo, E. y Ruiz-Reyes, K. (2018). Ways to strengthen the statistical literacy, reasoning and thinking in the mathematics teachers training. En M. Sorto, A. White y L. Guyot (Eds.), *Proceedings 10th International Conference on Teaching Statistics (ICOTS 10)*. Kyoto, Japón: ISI/IASE.
- Ruz, F., Molina-Portillo, E. y Contreras, J. M. (2019). Guía de valorización de la idoneidad didáctica de procesos de instrucción en didáctica de la estadística. *Bolema*, 33(63), 135-154.
- Santiesteban, C. (2009). *Principios de psicometría*. Madrid: Síntesis.
- Schoenfeld, A. H. y Kilpatrick, J. (2008). Toward a theory of proficiency in teaching mathematics. En D. Tirosh y T. Wood (Eds.), *International handbook of mathematics teacher education* (Vol. 2, pp. 321-354). Rotterdam, Holanda: Sense Publishers.
- Tempelaar, D. (2004). Statistical reasoning assessment: An analysis of the SRA instrument. *Proceedings ARTIST Roundtable Conference* (pp. 1-29). Appleton,

WI: University of Lawrence.

Valle, J. M., Jesús, L. y Ayuso, M. (2011). La nueva formación inicial del profesorado de educación secundaria: modelo para la selección de buenos centros de prácticas. *Revista de Educación*, 354, 267-290.

### **Referencias de los autores**

Felipe Ruz, Universidad de Granada (España). felipe.ruz.angel@gmail.com

Elena Molina-Portillo, Universidad de Granada (España). elemo@ugr.es

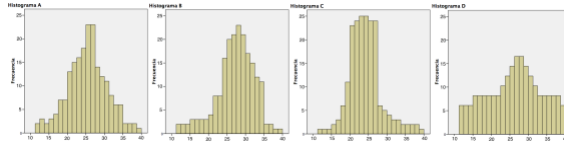
José Miguel Contreras, Universidad de Granada (España). jmcontreras@ugr.es

**Anexo: Ítems sobre el contenido de estadística descriptiva**

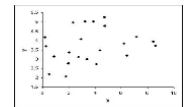
**Ítem 1)** Un estudio examinó la longitud de una especie de pez. El plan era tomar 200 peces y examinar los resultados. En la tabla se muestran estadísticos de resumen sobre la longitud en milímetros. ¿Cuál de los histogramas corresponde a los datos mostrados?

Promedio	Mediana	Desv. Estándar	Mínimo	Máximo
27,3 mm	28,6 mm	5,1 mm	12,1 mm	38,0 mm

A. Histograma A \*B. Histograma B C. Histograma C D. Histograma D

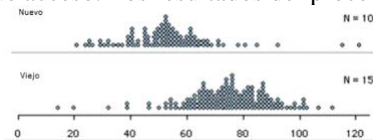


**Ítem 2)** El gráfico de dispersión muestra la relación entre puntuaciones de una escala de ansiedad (X) y calificaciones de un examen de ciencias (Y). En base al gráfico, deduzca el coeficiente de correlación entre X e Y.



A. 0 \*B. 0,3 C. 0,7 D. 1,3

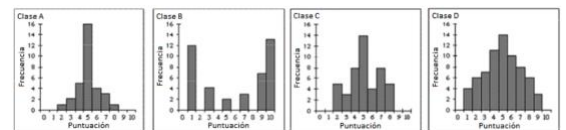
**Ítem 3)** Una compañía creó una fórmula de fibra óptica para mejorar la velocidad de navegación en internet. Para probar la efectividad de la fórmula, se realizó un análisis exploratorio de datos en un pueblo de 250 habitantes. 100 recibieron la nueva conexión y los otros 150 siguieron con conexión por cable. Luego se solicitó a cada persona visitar una página de prueba, registrando el tiempo (en segundos) de demora de acceso. Los resultados del procedimiento se muestran en los siguientes gráficos:



Las próximas declaraciones fueron realizadas por estudiantes de estadística. ¿Cuál de ellas es correcta?

- A. La antigua conexión funciona mejor. Dos personas tuvieron acceso en menos de 20 segundos. Además, la mayor demora - dos minutos - fue con la nueva fórmula.
- \*B. La nueva conexión funciona mejor. El tiempo promedio de la nueva fórmula de acceso a internet es menor que el de la antigua conexión, tarda unos 20 segundos menos.
- C. Ambas conexiones alcanzan igual velocidad de navegación, ya que la diferencia entre los resultados obtenidos no es estadísticamente significativa.
- D. No concluiría nada de estos datos. El número de habitantes en los dos grupos no es lo mismo, así que no hay una manera justa de comparar las dos fórmulas.

**Ítem 4)** A continuación se presentan los resultados de una prueba, en escala de 0 a 10 puntos, para cinco clases de estadística. ¿En cuál de ellas esperarías tener la dispersión más baja, y por qué?



- \*A. Clase A, porque tiene la mayoría de los valores cercanos a la media.
- B. Clase B, porque tiene el menor número de puntuaciones distintas.
- C. Clase C, es de las con menor rango y puntuaciones más homogéneas.
- D. Clase D, porque se la forma de su distribución se asemeja más a la curva normal.

**Ítem 5)** La puntuación obtenida por Ana en un examen corresponde al percentil 35 respecto a sus compañeros de clase, mientras que el resultado de Elena en la misma prueba corresponde al percentil 70. ¿Qué afirmaciones son correctas?

- A. Ana obtuvo mejor o igual puntuación que 35 compañeros de su clase.
- B. Ana responde bien aproximadamente una tercera parte del examen.
- C. Elena obtiene el doble de puntuación que Ana en el examen.
- \*D. El 30% de los compañeros de su clase obtuvieron mejor o igual puntuación que Elena.

**Ítem 6)** En tu vecindario, mediante un análisis exploratorio de datos se determinó un coeficiente de correlación de 0,9 entre nivel de ingresos y cantidad de plástico reciclado en una semana. ¿Qué indica este valor?

- A. En este vecindario ganar más dinero provoca que las personas reciclen más.
- B. Los vecinos que menos reciclan son los que ganan más dinero.
- \*C. Aunque los vecinos que menos dinero ganan son los que menos reciclan, esto no necesariamente significa que ganar menos sea la causa de reciclar menos plástico.
- D. Las personas que más ganan son las que más plástico reciclan.



## **Assessing descriptive statistical knowledge of prospective mathematics teachers**

Felipe Ruz, Universidad de Granada

Elena Molina-Portillo, Universidad de Granada

José Miguel Contreras, Universidad de Granada

In this paper, we present the process of design, construction and pilot implementation of an instrument to evaluate the knowledge about stochastic of mathematics teachers, specifically for descriptive statistics content. Tests were carried out in a sample of 126 prospective mathematics teachers from Spain and Chile, with results organized into two parts: on the psychometric properties of the items analysed and on the knowledge declared by the study participants. Regarding the psychometric characteristics of the instrument and the items reported, we were able to conclude a medium level of difficulty and an adequate discrimination force for each item, the same to guarantee as acceptable the internal consistency of the questionnaire. Regarding the answers declared by the participants, after grouping the items that evaluate common and extended knowledge, we observed significant differences between the results of both groups of participants for the second of them, what is justified by differences in previous training. When comparing the answers of the whole sample, we obtained that the scores of the items for the evaluation of each category of *per se* content knowledge are statistically significant, showing lower results in relation to extended knowledge. We conclude that the instrument has acceptable characteristics, as well as the items analysed, and it is hence projected as a useful tool to collect information regarding the state of disciplinary teacher training. By means of future applications we will be able to diagnose those more descended topics, and to contribute with specific antecedents that complement policy contributions of the National Diagnostic Evaluation (END) in Chile. We will thus be able to guide improvements in specific teacher education that satisfactorily respond to institutional demands on school tasks of teaching stochastic.