

El Examen de Estadística. Contenido, Tendencias y Reflexiones

Audy Salcedo
audy.salcedo@ucv.ve
Amalio Sarco Lira B.
asarcollira@gmail.com
Jesús González U.
jagonza05@gmail.com
Universidad Central de Venezuela

Recibido: 19 /03/2013
Aprobado:26/09/2013

Resumen

Este estudio presenta los resultados del análisis de un conjunto de exámenes utilizados para evaluar el aprendizaje de los estudiantes en estadística. A los profesores de la asignatura estadística aplicada a la educación se les solicitó los exámenes aplicados en los dos últimos semestres para evaluar el aprendizaje del estudiante en sus cursos, ya que se considera que su análisis debe proporcionar información sobre las habilidades cognitivas que se pretenden lograr en dicha asignatura. Las preguntas se clasificaron considerando el puntaje asignado, el área de estadística que evalúa, el contexto donde se formula, el tipo de pregunta y el nivel de dificultad. Además se clasificaron de acuerdo a la taxonomía SOLO (Biggs y Collis, 1982) y a las definiciones de alfabetización, razonamiento y pensamiento estadístico (Delmas, Garfield, Ooms y Probabilidad, 2007), como una forma de estimar las habilidades cognitivas que se pretenden evaluar. Los resultados sugieren que los exámenes analizados están orientados a evaluar habilidades cognitivas pertenecientes a los niveles inferiores de las taxonomías utilizadas.

Palabras clave: examen de estadística, taxonomía SOLO, alfabetización, razonamiento y pensamiento estadístico.

The Exam of Statistic to Evaluate Student Learning. Contents, Trends and Reflections

Abstract

This study presents the results of the analysis of a set of tests used to evaluate student learning in statistics. The professors who teach the course Statistics Applied to Education at the School of Education (UCV), were asked to provide the researchers the achievement tests they had applied in the last two semesters in order to measure students' learning in their courses, because we considered that such analysis could provide information about the cognitive skills involved in the answers. The questions

were classified considering the score assigned to the statistical area evaluated, the context in which they are formulated and their type and level of difficulty. The questions were also classified according to the SOLO taxonomy (Biggs and Collis, 1982) and the definitions of literacy, reasoning and statistical thinking (Delmas, Garfield, Ooms and Chance, 2007), as a way to estimate the cognitive skills which were intended to assess. The results suggest that the tests are designed to evaluate only cognitive skills belonging to the lower levels of the taxonomy used.

Key words: *statistical exam, SOLO taxonomy, literacy, reasoning and statistical thinking*

Introducción

En el sistema educativo el término evaluación alude por lo menos a tres procedimientos claves: el trabajo desarrollado por el docente, el currículo y el trabajo desarrollado por el estudiante. No obstante lo antes dicho, la evaluación de los aprendizajes de los estudiantes es considerada particularmente relevante, sin menospreciar el valor de las otras dos áreas, debido a que representa uno de los puntos fundamentales de todo proceso de enseñanza y aprendizaje.

En el marco de la evaluación de los aprendizajes, las pruebas escritas o exámenes son los instrumentos más utilizados por los docentes, sobretudo en la Educación Universitaria. Rodríguez Gómez (2008) señala que el examen es el instrumento de evaluación dominante en la Educación Universitaria y en muchos casos el único. Indica que aun en aquellos casos en que se utilizan otras técnicas e instrumentos de evaluación, el examen siempre tiene un peso mayor en la conformación de la calificación final. Salcedo Galvis (2010) coincide con el planteamiento anterior y señala que, a pesar de los cambios que han acompañado a la evaluación, el examen es el instrumento de evaluación predominante en Venezuela.

Es indudable que las pruebas escritas tienen una importante influencia para la toma de decisiones como medio de promoción y acreditación de los alumnos en la Educación Universitaria, particularmente en las asignaturas del área cuantitativa, en las que, por tradición, el examen es el principal instrumento de evaluación. Es por ello que esta investigación se centra en los exámenes escritos de Estadística como objeto de estudio, esto sin subvalorar otros instrumentos de evaluación.

Los exámenes son elaborados por el docente, quien como experto decide cuál es el plan de evaluación que utilizará durante el lapso académico, eso también incluye cuáles aspectos del programa evaluará mediante pruebas escritas y en qué momento se evaluarán. Es también el docente quien elabora las preguntas o reactivos que utilizará para conocer los cambios alcanzados por el estudiante durante un lapso determinado. Ese proceso lo lleva a cabo en función de los objetivos planteados, pero también considera los contenidos trabajados y las experiencias de aprendizaje desarrolladas, además de las características generales de sus estudiantes. La elaboración del examen no es un proceso que realizan los docentes en atención solo a aspectos técnicos, sino que también intervienen sus ideas respecto a la enseñanza, el aprendizaje, la evaluación y sobre la propia disciplina que enseñan.

En el examen se manifiestan, en las distintas preguntas, aquellos contenidos, procedimientos e ideas que el docente valora en un grado mayor que los demás, aquellos en los que está interesado en conocer lo que sabe el estudiante o la calidad de su progreso. Cada examen contiene una muestra de las habilidades cognitivas que el docente desea conocer si el estudiante logró desarrollar y, como ya se ha indicado, influirá en el aprendizaje de los estudiantes. Por tanto, el conjunto de exámenes utilizados en una asignatura, durante un lapso académico, puede proporcionar información sobre las habilidades cognitivas que se pretenden desarrollar en esa materia.

Por todo lo antes expuesto es que en esta investigación se analizan los exámenes aplicados en la asignatura Estadística de la carrera Educación, Escuela de Educación de la Universidad Central de Venezuela; en otras palabras, el objetivo de esta investigación fue analizar los exámenes utilizados para evaluar el aprendizaje de los estudiantes en la asignatura Estadística Aplicada a la Educación que forma parte del plan de estudio de la carrera Educación en la institución antes señalada.

Con ello se pretende explicitar las decisiones tomadas por un grupo de profesores en el diseño de sus exámenes para identificar tendencias en cuanto a las habilidades cognitivas que intentan evaluar. Se parte de la vinculación proceso de enseñanza y evaluación, puesto que,

se considera que el análisis de los exámenes debe proporcionar información sobre las destrezas y habilidades que se pretenden lograr en la enseñanza de la Estadística en el campo de la Educación.

Referente teórico

La importancia del conocimiento que tiene el profesor en el hecho educativo

Shulman (1986) destacó la necesidad de considerar en los procesos de enseñanza el papel que desempeña el conocimiento que tiene el profesor de la materia que imparte. En tal sentido, el autor señala dos componentes de ese conocimiento profesional del profesor: el conocimiento de la materia y el conocimiento pedagógico del contenido. Hill, Ball y Schilling (2008), hablan del *Conocimiento Matemático para la Enseñanza* (Mathematical Knowledge for Teaching, MKT), como el conocimiento matemático que es utilizado por los profesores en sus clases para realizar la enseñanza de la matemática. Schoenfeld y Kilpatrick (2008) introducen el concepto de *Proficiency in Teaching Mathematics*, para referirse a la competencia profesional del profesor de matemáticas, mientras que Godino, Contreras y Font (2006) hablan de *idoneidad didáctica*.

Todos esos modelos buscan caracterizar los conocimientos del profesor para destacar su importancia en el hecho educativo. El profesor de Estadística no escapa de esa situación, por ejemplo, Eichler (2008) realizó un estudio cualitativo con cuatro profesores alemanes de Estadística y encontró que, aunque los cursos no diferían en cuanto a metas y contenido, los planes de estudio puestos en práctica eran considerablemente distintos en cuanto a sus objetivos. Este investigador le atribuye las diferencias a la visión de cada profesor respecto a la Estadística, su enseñanza y evaluación.

La evaluación en Educación Estadística

El tema de la evaluación en Educación Estadística (EE) ha recibido mayor atención en los últimos años, particularmente desde la perspectiva de las respuestas que ofrecen los estudiantes en las actividades de evaluación. Educadores estadísticos han propuesto taxonomías para caracterizar la ejecución de los estudiantes en

tópicos específicos de la Estadística, por ejemplo, Jones, Langrall, Thornton & Mogill (1999), Watson y Moritz (2000), Vallecillos y Moreno (2006), Aoyama (2007), Reading y Reid (2007) y Landín y Sánchez (2010). Todos estos modelos buscan describir el nivel cognoscitivo alcanzado por el estudiante a partir de sus respuestas, no obstante, ellas dependen del tipo de pregunta que el profesor formule.

La taxonomía *Student Observed Learning Outcomes* (SOLO), propuesta por Biggs y Collis (1982), es una de las más utilizadas para categorizar el desempeño cognoscitivo de los estudiantes y es aplicada a diferentes áreas de contenido.

Taxonomía para la clasificación de las preguntas del examen de acuerdo con las habilidades cognitivas que se pretenden evaluar

La taxonomía “Resultados de Aprendizaje Observados en el Estudiante” -Student Observed Learning Outcomes (SOLO)-.

Si bien la taxonomía SOLO fue creada para evaluar el desempeño cognoscitivo de los estudiantes, también puede ser utilizada para definir los objetivos del currículo, incluyendo los niveles de aprendizajes requeridos (Biggs, 1999). En este sentido se realizó una adaptación de la taxonomía SOLO (Biggs y Collis, 1982) para estimar el nivel de desempeño cognoscitivo, que aspira el profesor que proporcione el estudiante. Las definiciones utilizadas son:

Nivel uniestructural (NU). Preguntas que contienen datos informativos explícitos, que al ser extraídos directamente del enunciado, permiten encontrar la respuesta. La comprensión de la información sugiere el procedimiento a utilizar para la solución de la pregunta.

Nivel multiestructural (NM). La pregunta requiere de la utilización de dos o más informaciones, que se obtienen directamente del enunciado. Se utiliza un procedimiento conocido, que surge a partir de la comprensión de la información del enunciado, para generar la respuesta. Los datos se procesan de forma secuencial para generar una nueva información.

Nivel relacional (NR). La pregunta requiere del análisis de la información, establecer relaciones entre los diferentes aspectos usados para integrarla y generar implicaciones, consecuencias, predicciones o conclusiones a partir de elementos del contexto. La forma de solución de la pregunta no es evidente.

Nivel de abstracción extendida (NAE). La pregunta exige la utilización de un principio general y abstracto que puede ser inferido del análisis de la información presente en el enunciado. Requiere integrar información adquirida para generar una nueva que permite solucionar la situación.

Clasificación específica de las preguntas de acuerdo a las definiciones de delMas et al (2007).

Alfabetización estadística (AE). La pregunta indaga sobre el conocimiento del significado de términos estadísticos, utilización del lenguaje básico y los instrumentos de la estadística, la comprensión del uso de símbolos, reconocer e interpretar la representación de datos (gráficas y tablas). Indaga por qué son necesarios los datos y la forma cómo se pueden producir. Procuran la comprensión de las nociones básicas de probabilidad y de cómo se llega a conclusiones o inferencias estadísticas. Apuntan a conocer el desarrollo de la capacidad para interpretar y evaluar críticamente la información estadística, así como comunicar ideas estadísticas

Razonamiento estadístico (RE). Se pretende investigar si el estudiante comprende las ideas estadísticas, explica los procesos estadísticos e interpreta cabalmente los resultados estadísticos. Implica la conexión de dos o más conceptos estadísticos y da sentido a la información estadística. Se trata de hacer interpretaciones basadas en conjuntos de datos, las representaciones de los datos o resúmenes estadísticos de los datos.

Pensamiento estadístico (PE). Se pregunta sobre el conocimiento del proceso de investigación estadística; entender cómo los modelos son usados para simular fenómenos aleatorios y por qué sirven para estimar probabilidades; implica una correcta comprensión de la naturaleza del muestreo, cómo hacer inferencias a partir de las muestras y cómo diseñar experimentos con el fin de establecer

causalidad; la capacidad de entender y utilizar el contexto de un problema de investigación, así como planificar, evaluar investigaciones y extraer conclusiones. Entender cómo, cuándo y por qué los instrumentos de inferencia pueden usarse para fundamentar los procesos de investigación y criticar y evaluar los resultados de investigaciones, lo que implica la comprensión del por qué y cómo se realizan las investigaciones.

En síntesis, la taxonomía SOLO permite catalogar las preguntas de una forma general; las definiciones de alfabetización, razonamiento y pensamiento estadístico permiten clasificarlas en un marco más específico de la Educación Estadística. Para este proceso se optó por las definiciones de delMas et al (2007):

Metodología

Tipo y nivel de la investigación

El estudio realizado es una investigación de campo de carácter exploratorio. A profesores que dictan asignaturas de Estadística, en la carrera Educación de la Universidad Central de Venezuela, se les solicitó el conjunto de exámenes utilizados en los dos últimos semestres para la evaluación de los aprendizajes en sus asignaturas. Se señaló que el propósito era estudiar las evaluaciones escritas realizadas como conjunto, que el interés no se centraba en el docente, sino en las evaluaciones y que no se realizarían análisis individuales por profesor. Asimismo se les garantizó la confidencialidad del material proporcionado y que su uso era estrictamente para la investigación.

El corpus analizado

Se recibió respuesta de 10 profesores, quienes proporcionaron 97 exámenes diferentes, para un total de 978 preguntas o situaciones de evaluación.

Las evaluaciones corresponden a las asignaturas Estadística Aplicada a la Educación, Matemática y Estadística I, Estadística II y Estadística III. La primera se dicta en la modalidad presencial en un régimen anual, mientras que las otras tres se dictan en los Estudios Universitarios Supervisados (EUS); por lo tanto pertenecen a la

modalidad semipresencial o mixta de la Escuela de Educación que se dicta en cinco centros del país. Los profesores que proporcionaron evaluaciones tienen entre 4 y 23 años de experiencia dictando las asignaturas.

Categorías de análisis y procedimiento

Las preguntas o situaciones de evaluación se clasificaron considerando el puntaje asignado, el área de estadística que evalúa, el contexto donde se formula la pregunta, el tipo de pregunta o actividad y el nivel de dificultad, estimado por juicio de experto.

- De acuerdo con el área de la estadística que evalúa cada pregunta, las áreas que más tradicionalmente se identifican en ella son: estadística descriptiva, probabilidad y estadística inferencial.
- La clasificación del contexto donde se formula la pregunta se realizó sobre la base de su relación con la carrera de educación: contextos asociados a educación, no asociados a educación y sin contextos. Respecto al tipo de pregunta o actividad, las preguntas se clasificaron en: preguntas de desarrollo, selección simple, verdadero – falso y de respuesta corta (completar).
- Se estimó el nivel de dificultad de las preguntas, para lo cual se resolvieron las preguntas y se valoró el procedimiento requerido para su resolución. Se trata de una estimación, de juicio de experto a partir de la consideración del porcentaje teórico de respuestas correctas que se podría lograr en la aplicación del examen. Las categorías utilizadas fueron: muy fácil (100% – 81%), fácil (80% – 61%), promedio (60% – 41%), difícil (40% – 21%) y muy difícil (20% – 1%).
- Para clasificar las preguntas, de acuerdo con las habilidades cognitivas que se pretenden evaluar, se utilizó una adaptación de la taxonomía SOLO y las definiciones de alfabetización, razonamiento y pensamiento estadístico de delMas, Garfield, Ooms, y Chance (2007).

Todas las preguntas fueron resueltas para poder establecer con claridad los procesos cognoscitivos que implicaban y poder clasificarlas. Cada investigador realizó de forma separada su propia clasificación y luego se compararon las tres clasificaciones. Se encontró que en sólo 126 preguntas hubo acuerdo total entre las tres clasificaciones, en el

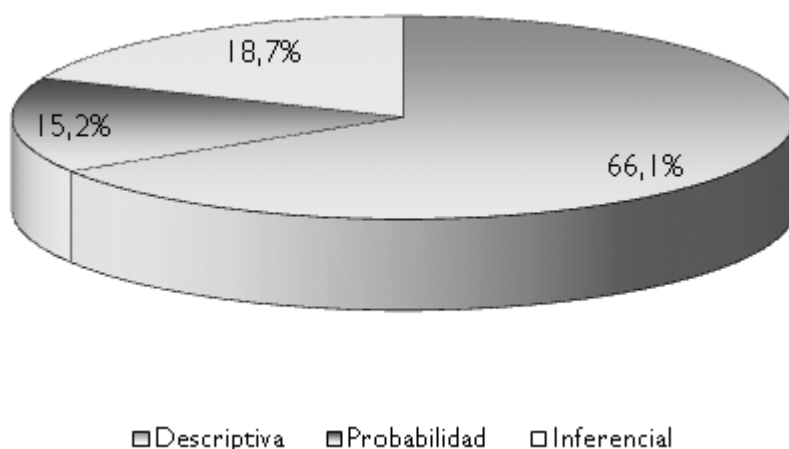
resto hubo por lo menos una diferencia, en alguna de las categorías consideradas. Se realizaron reuniones para discutir las divergencias hasta generar una única clasificación.

Resultados

La mayoría de las preguntas consideradas son de desarrollo (76,8%), el resto son preguntas denominadas objetivas o de respuesta breve: selección simple (8,4%), verdadero – falso (3,8%), completar oraciones (11,0%). Este resultado se corresponde con una larga tradición en los exámenes del área cuantitativa (por ejemplo, Matemática, Estadística) en los distintos niveles de educación, donde se prefiere que el estudiante “produzca” una respuesta y donde las preguntas objetivas o de respuesta breve suelen ser un complemento en el examen. Es importante destacar que este tipo de preguntas no se utiliza para evaluar el manejo de información o recordar hechos, sino para evaluar aprendizajes más complejos, en los que el estudiante debe mostrar tanto lo que sabe como su habilidad para argumentar y justificar sus respuestas.

En el gráfico 1 se muestran las preguntas clasificadas por áreas de la Estadística.

Gráfico 1
Preguntas clasificadas por área de la Estadística



Dos tercios de las preguntas analizadas corresponden a estadística descriptiva, mientras que las áreas de probabilidad e inferencia estadística se reparten de forma casi equitativa (el tercio restante). Que la mayoría de las preguntas sean de estadística descriptiva puede asociarse a la “pirámide escolar”, ya que esta parte de la estadística se ubica al inicio de la carrera Educación. En el primer semestre suelen abrirse una mayor cantidad de cursos, por una mayor presencia de estudiantes, que en los semestres segundo y tercero, donde se encuentran los temas de probabilidad e inferencia estadística. También podría deberse al hecho de que en los programas se haga un mayor énfasis en la estadística descriptiva, ya que la mayor parte del programa del plan anual y de las Estadísticas de la modalidad EUS se corresponden con esa rama de la estadística. En todo caso, la mayor cantidad de preguntas de Estadística descriptiva parece corresponderse con las necesidades de la asignatura, y no a una preferencia particular de los profesores.

En el gráfico 2 se muestran las preguntas clasificadas según el contexto.



En cuanto a los contextos donde se formulan las preguntas, el 56% de ellas se ubica en el campo de la Educación, lo cual es esperado en un curso de Estadística en la carrera Educación. El

contexto es un referente de la aplicación de la estadística a campos específicos, por lo que trabajar en estos cursos con una importante presencia de contextos asociados a educación ayuda al estudiante a visualizar posibles aplicaciones de la estadística en su vida como profesional de la educación. El contexto de educación de preferencia es el *rendimiento académico*, preguntas que tratan la actividad de evaluación (calificaciones, resultados de pruebas, notas, etc.) y su tratamiento mediante la estadística. Los otros contextos asociados a Educación que destacan por el porcentaje que logran son *características individuales de los estudiantes*, *ingreso a la educación* e *investigación en educación*, no obstante, ninguno de ellos supera el 5% del total de preguntas analizadas. Llama la atención la poca presencia de contextos asociados a la investigación en educación, cuando se puede suponer que un profesional de la educación debe estar relacionado con resultados de investigaciones.

Algunos ejemplos de preguntas formuladas en contextos asociados a Educación:

Rendimiento Académico:

Pregunta 51. *Se sabe que el 30% de los estudiantes de estadística acostumbran estudiar en grupo la asignatura. Si queremos realizar un estudio sobre todos los estudiantes de estadística, fijamos un error máximo admisible del 2,5% y una confianza del 90% ¿Cuán grande debe ser la muestra (3 puntos)*

Pregunta 104. *Un examen consta de 10 preguntas de selección simple, cada una con seis respuestas, donde solo una es la correcta. Si un alumno contesta al azar, ¿Cuál es la probabilidad de que conteste correctamente a lo sumo 1 pregunta? (2 puntos)*

Características Individuales de los Estudiantes

Pregunta 532. *Si estamos interesados en aplicar el Coeficiente de Correlación (Φ), para conocer la relación entre las variables genero y lugar de procedencia de los estudiantes, ¿Qué tendríamos que hacer con esta última variable para poder realizar los cálculos respectivos. (2 puntos)*

Pregunta 960. *Para estimar el promedio de edad de los 354 estudiantes del Centro Regional Bolívar, se determina el mismo en una muestra aleatoria de 65 personas obteniéndose una distribución aproximadamente normal, con*

media 32 y desviación típica 7. Se requiere que usted: Construya el intervalo de confianza para estimar el promedio de edad en toda la población con un nivel de Confianza de 95%. (3 puntos)

Que el 15% de las preguntas se formulen en contextos no asociados a Educación, se considera que es una forma de brindar oportunidad para que el estudiante aplique la estadística en contextos distintos a la de su área natural de trabajo. El contexto más usado fue lo que se denominó *Características de personas u objetos* seguido por *Tiempo en general*, sin embargo, la participación de ambos contextos es pequeña. En el primer caso apenas supera el 5% mientras que el segundo supera por muy poco el 3%. Todos los demás contextos se encuentran por debajo del 2%. Ejemplos de preguntas formuladas en contextos asociados a Educación son:

Pregunta 201. Indique en los siguientes casos cuál es la frecuencia y cuál es la variable: 35 años es la edad de 10 de los participantes.

Pregunta 854. Un grupo de estudiantes seleccionan 50 animales experimentales y les suministran una cierta clase de alimentación por un periodo de 2 semanas. El peso de estos arrojo un promedio de 42 grs. con una dispersión de 5 grs. Hallar un intervalo de confianza del 95% para estimar el aumento promedio que se espera tenga la población e interpretarlo. (3 puntos)

Llama la atención que el 29% de las preguntas se formulen sin contexto específico, cuando en Estadística el contexto es fundamental para interpretar y analizar los resultados logrados.

Ejemplos de este tipo de preguntas son:

Pregunta 3. ¿Cómo interpretaría usted un coeficiente de Kurtosis de -0.216 puntos?

Pregunta 30. En una variable en escala nominal. ¿Cuál sería la medida de tendencia central más adecuada?

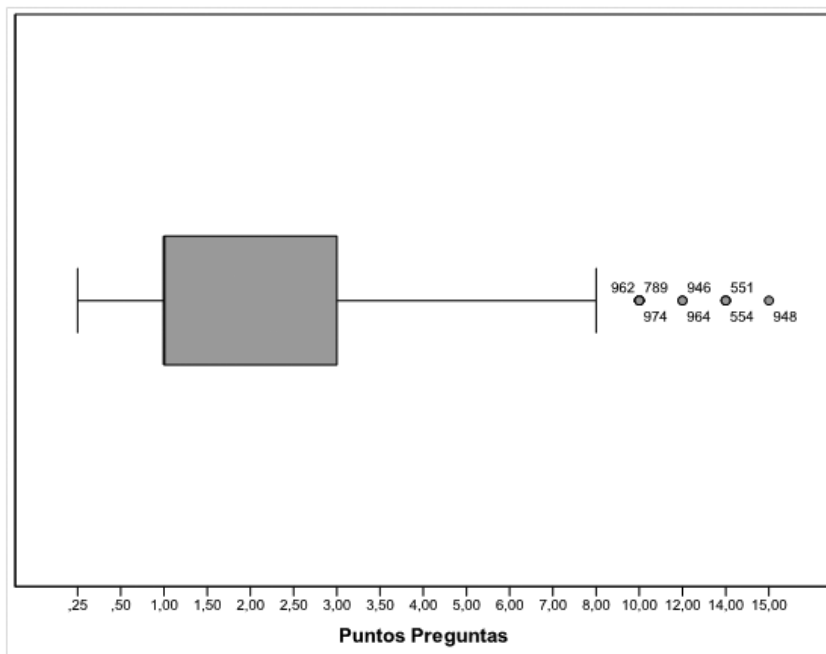
Pregunta 132. Si hemos construido un intervalo con un 95% de confianza esto nos indica que:

Pregunta 187. ¿Qué tipo de medidas podemos utilizar para estudiar la heterogeneidad de un conjunto de datos?

La mayoría de las preguntas sin contexto son del tipo de respuesta breve y buscan examinar elementos teóricos de la Estadística, se dirigen básicamente a evocar información o a realizar cálculos sencillos. La casi totalidad de ellas podrían transformarse y ubicarse en un contexto específico, y mejorar así la pregunta, incluso brindarle un nivel superior al de evocación de información.

En el gráfico 3 se muestran las preguntas clasificadas por valoración en puntos.

Gráfico 3
Preguntas clasificadas por la valoración en puntos



En el gráfico se puede apreciar que el 50% de las preguntas son valoradas con un punto o menos, mientras que el 50% restante oscila entre 1,5 y 8 puntos. Es probable que las preguntas con una valoración menor o igual a 1 punto formen parte de un ítem con varias partes o que sean preguntas de respuesta breve. El 75% de las preguntas tiene 3 puntos o menos, lo cual indica que las preguntas con alta puntuación son poco frecuentes. Es claro, que a la mayor parte de las preguntas se le asigna una baja puntuación, por lo que se puede

pensar que su respuesta no requiere de gran elaboración, pueden ser de baja exigencia cognitiva, por ende, a un muy bajo nivel de dificultad. Asimismo, en el gráfico anterior se observan valores atípicos (outlier), son preguntas cuyas puntuaciones van desde 10 a 15 puntos, las cuales difieren de forma considerable del resto de la distribución de los puntajes. Un valor outlier es un dato que está alejado por lo menos 1,5 rango intercuartil (RIC) del cuartil 1 o el 3 (Q_1 o Q_3). Los outlier pueden ser valores atípicos o valores extremos, los atípicos son aquellos que se alejan entre 1,5 RIC y 3 RIC de Q_3 o Q_1 y se señalan con un círculo, mientras que valores extremos se alejan al menos en 3 RIC de Q_3 o Q_1 y se indican con un asterisco. En este caso, se marcan 4 círculos y se destacan 8 números, los cuales indican la presencia de 8 preguntas consideradas como atípicas por la puntuación asignada. De acuerdo con la puntuación, este tipo de pregunta podría considerarse como muy difícil, posiblemente le exigen al estudiante altos niveles de razonamiento o habilidades cognitivas superiores. El hecho de ser representados en el gráfico como valores atípicos, permite identificarlas como preguntas poco frecuentes en los exámenes analizados.

Ejemplo de ese tipo de preguntas son:

Pregunta 948. Durante los últimos 10 años se ha registrado el número de docentes realizando estudios de postgrado en las Instituciones de Educación Superior de un Municipio:

AÑOS	N° DE DOCENTES
2000	309
2001	280
2002	271
2003	328
2004	310
2005	383
2006	365
2007	482
2008	567
2009	693

En base a esta información se requiere: Determinar la ecuación de la recta de Regresión que se ajusta a los datos (15 puntos).

Pregunta 946. Se dispone de las puntuaciones en un test de Habilidad Numérica y Verbal de un grupo de aspirantes a ingresar a la Facultad de Farmacia de la U.C.V.:

Test Habilidad Numérica	Test Habilidad Verbal
72	82
83	99
83	81
90	91
92	82

Determine e interprete la relación existente entre ambas variables mediante el cálculo de 2 índices diferentes. (12 puntos)

Los ítems presentados no parecen exigir altos niveles de razonamiento. Estos ejemplos tienen en común una considerable labor para producir la respuesta, esto es, realizar varios cálculos antes de llegar a la respuesta correcta, pero esta se produce por intermedio de algoritmos conocidos. Es importante acotar que en ocasiones este tipo de actividades se realiza con apoyo de calculadoras científicas, que contienen rutinas de cálculo de esas medidas estadísticas. Pareciera entonces que la puntuación se corresponde con la cantidad de pasos necesarios para resolverla y no con la exigencia cognitiva de la pregunta.

En la tabla 1 se muestra la clasificación de las preguntas según el nivel de dificultad estimado.

Tabla 1
Preguntas clasificadas por nivel de dificultad estimado

Nivel de dificultad	Frecuencia	Porcentaje
Muy fácil (100% – 81%)	363	37,1
Fácil (80% – 61%)	319	32,6
Promedio (60% – 41%)	269	27,5
Difícil (40% – 21%)	23	2,4
Muy difícil (20% – 1%)	4	0,4
Total	978	100,0

De acuerdo con los expertos, menos del 3% de las preguntas pueden calificarse como *Difícil* o *Muy difícil*, por lo tanto, la casi totalidad de las preguntas tiene un nivel de dificultad igual o inferior de

Promedio. Aproximadamente al 70% de las preguntas se le estimó su nivel de dificultad entre *Fácil* y *Muy fácil*, por lo tanto, se podría esperar que entre el 61% y 100% de los estudiantes resuelva correctamente el 70% de las preguntas que se les formulan, por lo que se podría estar en presencia de exámenes con bajo nivel de dificultad. Este resultado parece consistente con el encontrado en relación al puntaje asignado a las preguntas, donde el 50% se le asigna 1 punto o menos.

Habilidades cognitivas

Se utilizó la definición de delMas et al (2007) para clasificar las preguntas en función del aprendizaje estadístico esperado. En la tabla 2 se muestran los resultados de la clasificación realizada en la discusión con los expertos:

Tabla 2
Preguntas clasificadas por el aprendizaje esperado

Aprendizaje esperado	Frecuencia	Porcentaje
Alfabetización Estadística	814	83,2
Razonamiento Estadístico	148	15,1
Pensamiento Estadístico	16	1,6
Total	978	100,0

De acuerdo con la clasificación, el 83,2% de las preguntas examinadas evalúa logros en la *Alfabetización Estadística*, mientras que un 15,1% evalúa logros en *Razonamiento Estadístico* y un 1,6% en *Pensamiento Estadístico*. La mayoría de las preguntas se ubica en el nivel más bajo de la taxonomía utilizada, es decir, son preguntas que buscan conocer si el estudiante comprende y utiliza de forma adecuada el lenguaje básico de la Estadística y sus instrumentos. Este resultado parece corresponderse con las recomendaciones internacionales en las que se indica que los cursos iniciales de Estadística deben hacer énfasis en la *AE*, no obstante, parecen dejar de lado los otros niveles de la taxonomía, por lo menos en las evaluaciones escritas examinadas.

Para ilustrar el tipo de pregunta clasificada en cada nivel, a continuación se encuentran algunos ejemplos.

Alfabetización Estadística

Pregunta 4. A continuación se presentan las calificaciones obtenidas por los alumnos de las sección 12 de Estadística III, en la primera prueba parcial del semestre 2004-II.

Notas	4-6	7-9	10-12	13-15	16-18
Alumnos	3	8	9	15	12

Calcule e interprete el puntaje más frecuente obtenido por el curso. (2 ptos)

Pregunta 11. La rama de la estadística que solo describe las características de un grupo, sin hacer inferencias o predicciones de las características poblacionales a partir del análisis de estas se denomina: _____ (0,5 ptos)

Pregunta 95. Dos sucesos A y B se dice que son incompatibles cuando:

- La probabilidad de A condicionada por B es igual a la probabilidad de A.
- La probabilidad de la intersección es igual al producto de las probabilidades
- A y B no pueden ocurrir simultáneamente
- La unión de A y B es el suceso seguro.

Pregunta 232. Los siguientes datos corresponden a las calificaciones obtenidas por un grupo de 24 estudiantes de la escuela de educación, en los exámenes finales de Estadística I y II durante el año escolar 2003. (9 ptos)

ESTADÍSTICA I											
6	12	16	10	10	09	10	15	10	11	19	12
ESTADÍSTICA II											
9	11	16	13	12	08	11	13	10	15	20	14

A partir de los datos presentados:

Dibuje el diagrama de dispersión y analice el tipo de relación existente (3 puntos)

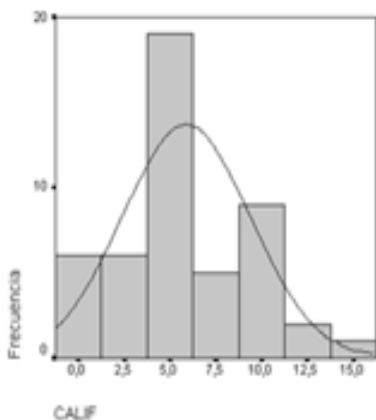
Pregunta 265. Para la construcción de un histograma se colocan en el eje X los valores de _____

Todas estas preguntas tienen en común que, para su solución, se necesita información o conocimiento básico de Estadística, aunque en algunos casos sea necesario hacer más de una actividad para resolverlas.

Razonamiento Estadístico

Pregunta 76. Se desea contrastar la hipótesis de que no existe relación entre el sexo y el rendimiento en la asignatura Estadística Aplicada a la Educación. Para esto se toman muestras al azar entre los estudiantes del profesor Sanadie en los últimos cinco años. De las mujeres seleccionadas se sabe que 10 aprobaron y 40 aplazaron, mientras que de la muestra de hombres 20 aplazaron y 30 aprobaron. Usando un nivel de significación de 5%, establezca las conclusiones correspondientes al caso.

Pregunta 307. Se presentan los resultados de las calificaciones de los Estudiantes de EUS de Estadística I del Núcleo Barquisimeto, para la primera evaluación parcial. Describa cómo es el rendimiento del grupo a partir de cada uno de los estadísticos de tendencia central y posición calculados para la distribución de puntajes. Determine la concordancia entre los resultados obtenidos a partir de los estadísticos calculados y el gráfico presentado.



N	48	
Media	5,88	
Mediana	6,00	
Moda	6	
Percentiles	10	1,00
	25	3,25
	75	8,75
	90	10,00

Pregunta 340. A continuación se presentan los resultados de la Prueba de Aptitud Académica de 36 estudiantes de quinto año de bachillerato del Colegio San José para las secciones A y B:

Concluya con respecto al rendimiento de los estudiantes en ambas secciones en la Prueba de Aptitud Académica.

	A	B
N		
Media		
Mediana	60,00	61,50
Modo		
Desviación		
Varianza	96,879	103,820
Asimetría (α^3)	0,386	
Kurtosis (α^4)		-1,650
Valor mínimo	46	44
Valor máximo	75	73

A	B
65	70
73	62
52	50
48	72
74	66
61	66
60	51
75	48
47	45
73	50
48	44
53	73
46	61
60	66
56	46
60	71
50	66
60	55

Como se puede observar, para resolver estas preguntas hace falta algo más que recordar terminología o realizar cálculos, el estudiante debe discurrir las ideas estadísticas involucradas, relacionarlas e interpretarlas. Lo importante aquí no es el cálculo, que cuando está suele ser sencillo, sino en vincular dos o más conceptos y producir una información estadística.

Pensamiento Estadístico

Pregunta 546. Con los datos abajo colocados Plantéese un estudio educativo donde establezca claramente población, muestra, estadísticos y parámetros; de modo que se pueda aplicar los siguientes procedimientos: Intervalo de confianza para la media poblacional. (Se incluye una tabla con información de las calificaciones finales de 20 estudiantes de estadística III de 4 centros regionales de los EUS).

Pregunta 576. Plantee una situación de investigación en la cual usted utilizaría la prueba de hipótesis para la Proporción Poblacional (P) (5 puntos).

Pregunta 582. Plantee una situación de investigación en la cual usted utilizaría la prueba de hipótesis para la diferencia de Medias Poblacionales ($\mu_1 - \mu_2$) (5 puntos).

Son pocas las preguntas que se clasificaron como de *Pensamiento Estadístico*, los ejemplos expuestos son casos característicos, los otros son preguntas semejantes a las aquí señaladas, tienen la misma estructura, pero cambian los datos o los parámetros a los que se refieren. Estos tres ejemplos tienen en común que le piden al estudiante plantear una “situación de investigación” o “estudio educativo”. Para hacer esta actividad, es necesario que el estudiante conozca bien el proceso de investigación estadística, pero además debe saber cómo, cuándo y por qué pueden ser utilizados los instrumentos de inferencia. El estudiante podría generar una situación, tipo ejercicio semejante al que aparece en algunos libros o manuales de estadística, para ilustrar el uso de los instrumentos estadísticos, pero se considera que la pregunta implica plantear un problema y planificar una investigación con el uso de la Estadística.

Al clasificar las preguntas de acuerdo con la Taxonomía SOLO, para estimar el nivel de respuesta que aspira el profesor que proporcione el estudiante, se encontró que 66% de las preguntas se ubicaron en el *Nivel Uniestructural*, mientras que 27,9% se consideran del *Nivel Multiestructural*, 5,7% se ubicaron en el *Nivel Relacional* y 0,4% en el *Nivel Abstracción Extendida*.

En la tabla 3 se muestran los resultados obtenidos en las preguntas clasificadas según la taxonomía SOLO.

Tabla 3
Preguntas clasificadas según la taxonomía SOLO

Nivel	Frecuencia	Porcentaje
Uniestructural	645	66,0
Multiestructural	273	27,9
Relacional	56	5,7
Abstracción extendida	4	0,4
Total	978	100,0

Los resultados anteriores indican que, en la mayoría de las preguntas analizadas se requiere de habilidades cognitivas bajas para su solución. El estudiante puede proporcionar la respuesta con el uso de la información que se encuentra explícita en el enunciado y en la mayoría de ellas la respuesta se logra al aplicar un procedimiento o algoritmo, previamente conocido, a la información proporcionada

en el enunciado. Un grupo importante de preguntas se soluciona al aplicar dos o más procedimientos conocidos a la información suministrada en el enunciado. Pareciera, entonces, que este grupo de preguntas son fundamentalmente para conocer si sus estudiantes pueden recordar o reproducir información proporcionada en clase (*Nivel Uniestructural*), así como clasificar o describir la comprensión de hechos sin sistematizarlos (*Nivel Multiestructural*).

A continuación, algunos ejemplos de las preguntas clasificadas en cada nivel de la taxonomía.

Nivel Uniestructural

Pregunta 1. *¿Qué escalas están asociadas con el nivel cualitativo de medición?*

Pregunta 55. *Algunas características de un experimento binomial son:*

- a) *Distribución simétrica; $p \neq q$ y dos resultados posibles en cada ensayo.*
- b) *Variable discreta; distribución simétrica y esperanza matemática igual a 1.*
- c) *Variable continua; $p = q = 0,5$ y $p + q = 1$*
- d) *$p + q = 1$; ensayos independientes y asintótica*
- e) *n ensayos idénticos; cada ensayo tiene dos resultados posibles donde p y q permanecen constantes.*

Pregunta 243. *El Decano de una Facultad desea realizar una encuesta para determinar la proporción de estudiantes que está a favor del cambio de sede. Ya que entrevistar 3540 estudiantes en un lapso razonable es casi imposible, determine el tamaño de la muestra (N° de estudiantes a entrevistar) necesario para estimar la proporción de estudiantes que están a favor, con error máximo admisible de 0,05 y un nivel de confianza del 95%. (4 puntos)*

Pregunta 460. *Las siguientes, son las calificaciones de un grupo de 20 estudiantes de un curso de Estadística: 12 15 10 13 16 12 18 18 20 8 6 17 14 1 8 6 7 5 4 12*

Se requiere que Ud. Construya una distribución de frecuencias de 5 clases.

Las preguntas anteriores tienen explícita tanto la actividad a realizar como la información necesaria para llevarla a cabo, por esa razón fueron clasificadas como del *Nivel Uniestructural*. En general,

son preguntas que solo exigen comprender la información y tener un conocimiento básico para lograr la respuesta.

Nivel Multiestructural

Pregunta 6. *A continuación se presentan las calificaciones obtenidas por los alumnos de la sección 12 de Estadística III, en la primera prueba parcial del semestre 2004-II.*

Notas	4-6	7-9	10-12	13-15	16-18
Alumnos	3	8	9	15	12

Calcule e interprete el coeficiente de variación (3 puntos).

Pregunta 42. *Una población está formada por los siguientes datos: 3,7,11 y 15 considerando todas las muestras posibles de tamaño 2 sin reemplazo que pueden extraerse de dicha población determinar: La media poblacional.*

Pregunta 195. *En un grupo de 60 estudiantes admitidos en la Escuela de Educación de la UCV, para el periodo 2003-04, se observó que el promedio de los índices académicos de dichos estudiantes fue de 66,253 puntos con una desviación típica de 8,435 puntos; y además, el promedio general de notas hasta 4º año, fue de 13,57 puntos con una desviación de 2,31 puntos. ¿En qué grupo de observaciones son más heterogéneos el conjunto de datos? 2 puntos.*

Pregunta 562. *Un alumno responde al azar un examen de verdadero-falso, que consta de 3 preguntas. Construya la distribución de probabilidad para la variable aleatoria "Número de preguntas respondidas correctamente". Calcule e interprete la Esperanza Matemática.*

En todas las preguntas anteriores, es necesario realizar por lo menos dos pasos u operaciones para resolverlas, pero esos pasos son conocidos previamente por el estudiante. Por ejemplo, en la pregunta 6 se pide calcular e interpretar el coeficiente de variación; para dar una respuesta correcta el estudiante debe: calcular la media aritmética, luego calcular la desviación típica y, por último, hallar el coeficiente de variación. Una vez obtenido el coeficiente debe interpretarlo en función de la homogeneidad o heterogeneidad del conjunto de valores respecto a su media aritmética. Todas esas medidas se calculan con algoritmos conocidos, cuyas fórmulas tiene el estudiante como parte del material que puede utilizar en el examen.

Nivel Relacional

Pregunta 75. Un profesor de Estadística Aplicada a la Educación está interesado en evaluar la efectividad de una nueva estrategia de enseñanza. Una vez que el curso ha presentado el primer examen escrito, el profesor comienza la aplicación de la experiencia. Para su posterior evaluación, decide seleccionar una muestra aleatoria de 8 alumnos, a los cuales les pregunta las calificaciones obtenidas en el primer y segundo examen respectivamente. La información recabada por él, se presenta en la siguiente tabla:

Alumno	Examen I	Examen II
A	10	15
B	5	13
C	12	15
D	16	20
E	8	14
F	11	12
G	18	17
H	20	19

Ayude al profesor a evaluar estadísticamente su propuesta, utilizando un nivel de significación del 5%.

Pregunta 175. El Liceo Julio Calcaño garantiza que sus estudiantes del 2do año del ciclo diversificado obtienen en la parte de habilidad verbal de la prueba de aptitud académica una calificación promedio superior a los 56 puntos. Un estudio realizado en forma aleatoria en los últimos cuatro años a 52 alumnos revela un promedio de 49 puntos con una desviación de 7,8 puntos. ¿Dicho estudio valida la garantía que proporciona el colegio? Concluya basado en que el nivel de significación es del 5% (3 puntos).

Pregunta 538. En una muestra de 150 escolares de educación básica en escuelas públicas en la población de Guanta. Se observó que 87 de ellos presentan dificultad en el desarrollo del pensamiento matemático elemental. Expertos del ministerio de educación dedicados a este tipo de investigación siempre han sostenido que el porcentaje de niños que presentan esta característica a nivel nacional es de menos del 55%. ¿Soporta la información obtenida en la muestra de los niños de Guanta la opinión de los expertos al 1%? (3 puntos).

Aunque en muchas de las preguntas ubicadas en el Nivel Relacional es posible que el estudiante pueda identificar de forma rápida el procedimiento que se debe utilizar para hallar la solución, la

pregunta le exige generar una implicación o conclusión, para lo cual debe establecer relaciones entre distintos elementos presentes en la pregunta. Luego de hallar un resultado, debe analizar la información, relacionarla y producir una predicción o conclusión, tomando en cuenta el contexto. Por ejemplo, obsérvese que en el caso de la pregunta 862, el estudiante podría incluso producir un juicio de valor, ya que se le pide su opinión respecto a la actuación del nuevo profesor, claro está sobre la base de la evidencia proporcionada. Es importante destacar que la casi totalidad de las preguntas ubicadas en este nivel corresponde a contenidos de Inferencia Estadística.

Nivel Abstracción Extendida

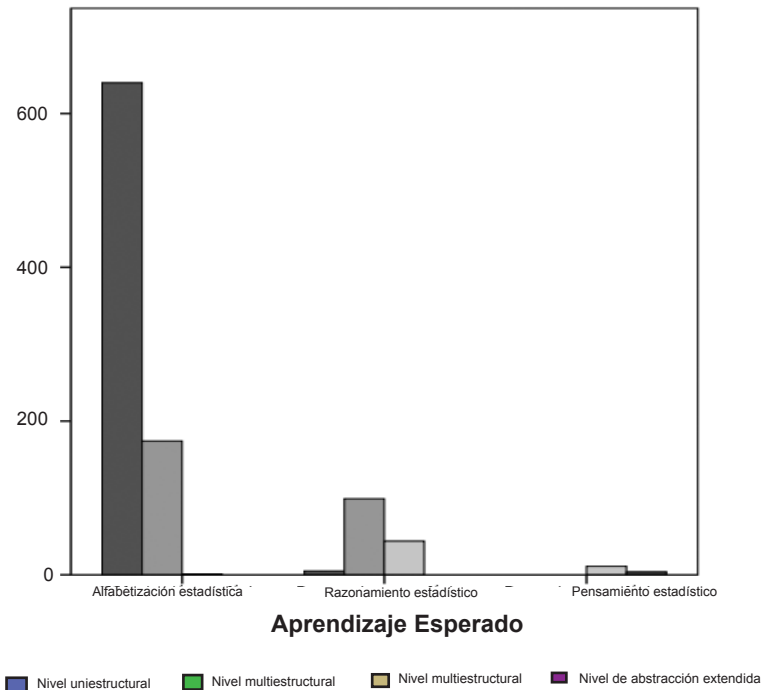
Pregunta 579. *Plantee una situación de investigación en la cual usted utilizaría la prueba de hipótesis para la Media Poblacional (μ) (5 puntos)*

Pregunta 613. *Plantee una situación de investigación en la cual usted utilizaría la prueba de hipótesis para la diferencia de Proporciones Poblacionales ($P_1 - P_2$) (5 puntos)*

Al igual que en el caso de las preguntas de *Pensamiento Estadístico*, las preguntas catalogadas en el *Nivel Abstracción Extendida* son poco frecuentes. Estas preguntas fueron incluidas en este nivel, ya que se considera que para plantear una situación de investigación real, es necesario que el estudiante integre los conocimientos adquiridos en el curso para generar un ejemplo de un problema auténtico, que requiera de técnicas particulares de inferencia Estadística para su solución.

En el gráfico 4 se muestran los resultados del análisis de las preguntas clasificadas en ambas taxonomías

Gráfico 4
Preguntas clasificadas en ambas taxonomías



Los resultados indican que los exámenes están conformados básicamente por preguntas de *Alfabetización Estadística*, a las cuales se le puede hallar solución al comprender la información del ítem y aplicar un procedimiento o algoritmo estudiado en clases.

En consecuencia, la mayoría de las preguntas se consideran de *Alfabetización Estadística* en los niveles *Uniestructural* y *Multiestructural*, por tanto, se puede afirmar que las preguntas evalúan básicamente las habilidades referidas a la reproducción de conocimiento y se busca incrementar los conocimientos sobre Estadística, a los que se les puede hallar solución realizando una o dos procedimientos. Son escasas las preguntas que buscan evaluar el conocimiento profundo y la aplicación de las ideas Estadísticas.

La tabla 4 presenta información sobre los tipos de pregunta que conforman los exámenes.

Tabla 4
Exámenes clasificados según su tipo

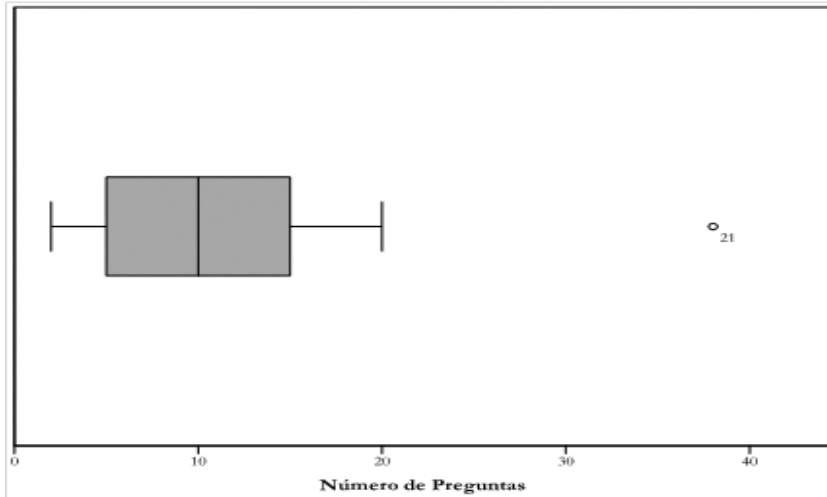
Tipo de examen	Frecuencia	Porcentaje
Desarrollo	59	60,1
Desarrollo y completar	15	15,5
Desarrollo y selección simple	18	18,6
Desarrollo y Verdadero – Falso	4	4,1
Selección simple	1	1,0
Total	97	100,0

En la tabla anterior, se puede apreciar que el tipo de examen que se utiliza con mayor frecuencia es el denominado de Desarrollo, aquel que está conformado exclusivamente por preguntas donde el estudiante debe dar respuesta con “sus propias palabras”. Este tipo de examen se utiliza con frecuencia para evaluar aprendizajes complejos, en los que el estudiante debe comprender la información que se le presenta, organizarla, procesarla y, luego, expresar sus ideas con la integración de diversos elementos del planteamiento expuestos en la pregunta. Pareciera que este no es el caso de los exámenes de Estadística, ya que el análisis realizado indica que se utilizan preguntas fundamentalmente de *Alfabetización Estadística*, en el nivel *Uniestructural* y *Multiestructural* de la taxonomía SOLO, es decir, preguntas de información y de procesos cognitivos sencillos.

Asimismo se observó que un poco más del 38% de los exámenes son una combinación de preguntas de desarrollo con preguntas de otro tipo (selección simple, verdadero y falso, etc.). En tal sentido, se podría pensar que con esta mezcla se busca un examen equitativo y justo, con una parte teórica y una práctica. Los exámenes con esa combinación, teóricamente, potencian la oportunidad de aprobar a los estudiantes de menor rendimiento académico, pero también incluyen preguntas reto para los más aventajados.

En el gráfico 5 se presentan los exámenes clasificados según el número de preguntas

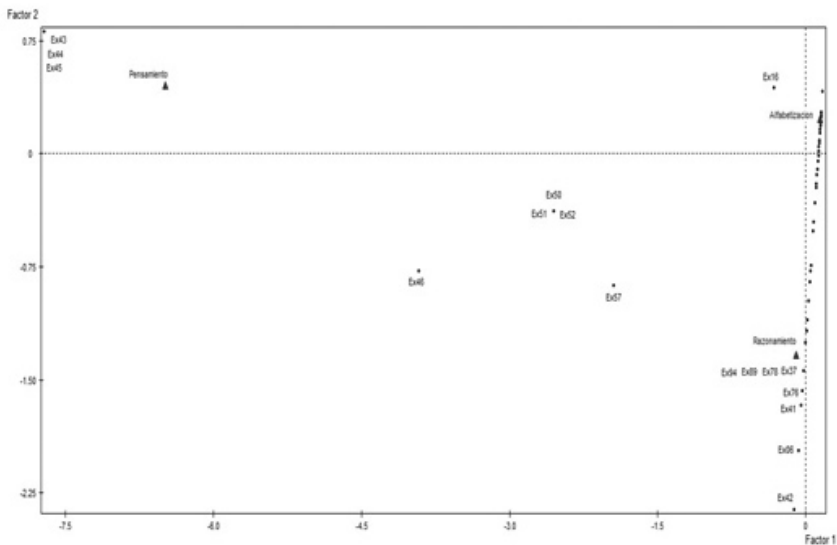
Gráfico 5
Exámenes según el número de preguntas



En el gráfico anterior, se aprecia que el número de preguntas oscila entre 2 y 20 preguntas y se observa un valor atípico, el examen número 21, que contiene 38 preguntas. Ese examen presenta una combinación de preguntas de Desarrollo (11) y de Completar (27), casi todas de tipo *Alfabetización Estadística* y *Uniestructural*. La media y la mediana del número de preguntas por examen está en 10; se puede pensar entonces que ese es el número de preguntas que representa mejor a un examen de Estadística, pero la elevada desviación estándar (5,66) indica que la distribución del número de preguntas por examen es heterogénea. Esa apreciación se ratifica al conocer el modo de la distribución, ya que 13 exámenes están conformados por 3 preguntas.

El gráfico 6 ilustra el análisis de correspondencia entre los exámenes y el aprendizaje esperado

Gráfico 6
Análisis de correspondencia entre exámenes y aprendizaje esperado



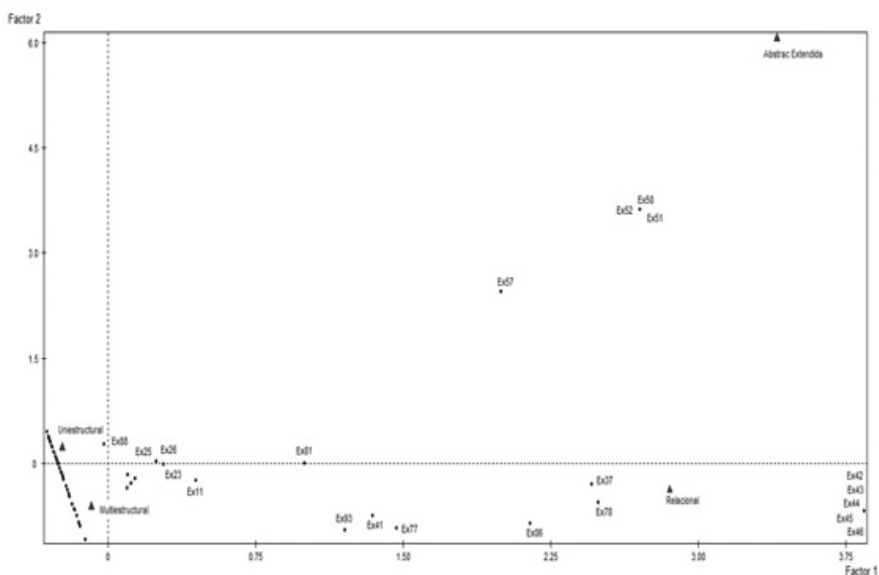
Tal y como se evidencia en el gráfico anterior una parte importante de los exámenes se agrupa alrededor de la *Alfabetización Estadística*, lo cual indica la asociación con esa categoría, *Afabetización Estadística*, lo que nos conduce a pensar que la mayoría de los exámenes están orientados a evaluar aspectos de ese nivel. Este resultado es congruente con la gran cantidad de preguntas ubicadas en esa categoría, como lo reveló el análisis realizado sobre las preguntas. Existe otro grupo de exámenes, menor en cantidad a los que se agrupan cerca de *Alfabetización Estadística*, que se encuentra en el espacio entre el *Razonamiento Estadístico* y la *Alfabetización Estadística*. Esos son exámenes que contienen preguntas de los dos tipos, mientras más cerca estén de una categoría indican que hay un mayor número de preguntas clasificadas en esa categoría. Se puede observar que los puntos cercanos a *Razonamiento Estadístico* son pocos, solo los exámenes 94, 89, 78 y 37 tienen mayoría de preguntas en esa categoría.

Los exámenes identificados con los números 43, 44 y 45 son muy semejantes, cada uno está conformado por 3 preguntas, todas iguales en cuanto a su estructura pero tratando contenidos distintos de inferencia estadística, clasificadas todas como de *Pensamiento Estadístico*, de allí que en el gráfico se le ubique cerca de la categoría *Pensamiento*. Un grupo pequeños de exámenes se encuentra entre el *Pensamiento Estadístico* y el *Razonamiento Estadístico*, debido a que reúne preguntas de las dos categorías, por ejemplo, el examen 46 tiene 4 preguntas, 2 clasificadas en *Pensamiento* y 2 en *Razonamiento*, de allí su ubicación a medio camino entre esas categorías.

En el gráfico 7 se muestra el nivel de correspondencia entre los exámenes y la taxonomía SOLO.

Gráfico 7

Análisis de correspondencia entre exámenes y la taxonomía SOLO



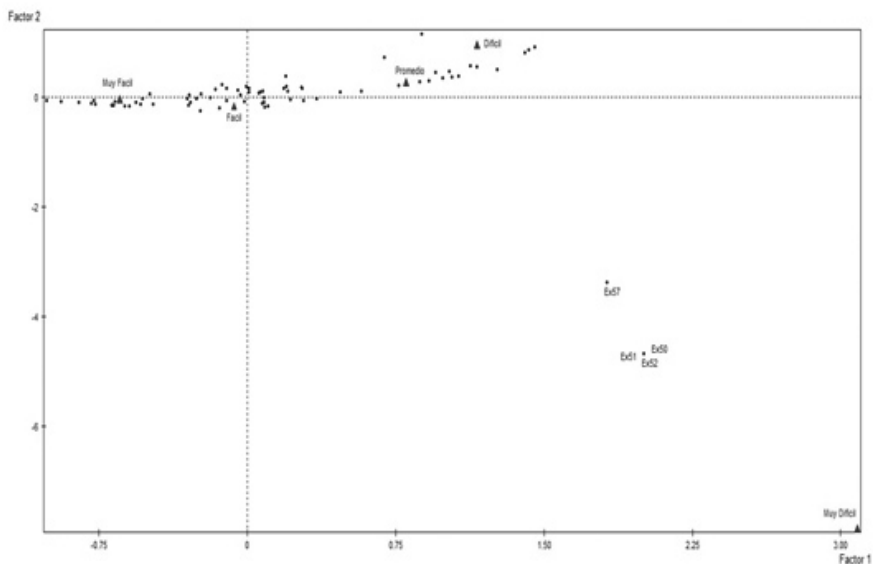
Como se puede apreciar en el gráfico 7, la mayor cantidad de exámenes se encuentra alrededor de los niveles *Uniestructural* y *Multiestructural*, lo cual indica que están conformados por preguntas que se ubican en esos dos niveles de la taxonomía SOLO. El conjunto de exámenes asociados a estos dos niveles se encuentra básicamente

ubicado detrás de los íconos que representan los dos niveles de la taxonomía formando una línea recta. Eso indica que hay un conjunto de exámenes que contienen principalmente preguntas clasificadas como *Uniestructurales* o *Multiestructurales* (las categorías más cercanas a cada uno de esos íconos), pero también hay un conjunto de exámenes que están conformados por preguntas de los dos tipos (los que están entre los dos íconos). Eso significa que los exámenes evalúan básicamente la reproducción del conocimiento estadístico, dejando poco espacio para la evaluación de procesos cognitivos complejos.

En el gráfico 8 se muestra la correspondencia existente entre los exámenes y los niveles de dificultad

Gráfico 8

Análisis de correspondencia entre exámenes y niveles de dificultad



En el gráfico 8, se aprecia que la mayoría de los puntos representativos de los exámenes se ubican alrededor de las categorías *Muy Fácil* y *Fácil*, lo cual indica que la mayoría de los exámenes están conformados por preguntas de bajo nivel de dificultad y se le consideran exámenes que deberían ser respondidos correctamente por una alta proporción (60% o más) de los estudiantes que los presenten.

Diversos manuales de construcción de pruebas (por ejemplo, APA, AERA y NCME, 1999) recomiendan que el nivel medio de dificultad del examen oscile entre 0,5 y 0,6; distribuyendo las preguntas de la siguiente forma: 5% de preguntas muy fáciles, 20% fáciles, 50% con una dificultad promedio, 20% difíciles y 5% muy difíciles. Sería interesante conocer si el sesgo que presentan los exámenes analizados es casual o responde a alguna premisa de los profesores de Estadística, por su conocimiento de la población a la cual van dirigidos o por la utilización de otras técnicas de evaluación que tiene mayores exigencias en cuanto a nivel de dificultad.

Conclusiones

Las características principales de las preguntas que conforman los exámenes analizados son: evalúan aspectos de Estadística Descriptiva, el estudiante debe desarrollar su propia respuesta (preguntas de desarrollo), en general evalúan los objetivos de los programas. El contexto de preferencia para formular las preguntas es el educativo, destaca el vinculado con el *Rendimiento Estudiantil* (resultados de pruebas, calificaciones, etc.). La mayoría de las preguntas fueron calificadas como *Muy Fácil* o *Fácil* y se valoran con puntuaciones de máximo 3 puntos.

Los exámenes son fundamentalmente de desarrollo, con un promedio de 10 preguntas por prueba. Los distintos análisis efectuados, con las preguntas y exámenes, sugieren que con estos últimos se evalúan aspectos parciales de la *Alfabetización Estadística*, con énfasis en procesos cognitivos básicos (niveles *Uniestructural* y *Multiestructural*), y con un nivel de dificultad que varía entre *Muy Fácil* y *Fácil*.

Partiendo de la premisa que el examen es una forma de recabar información sobre cuánto y con qué calidad han aprendido los estudiantes, de acuerdo con los resultados, los exámenes analizados están elaborados para saber si los estudiantes comprenden los conceptos básicos de Estadística y si son capaces de reproducir el contenido estudiado en las clases. Si como afirman Brown y Glasner (2003) y Watts, García-Carbonell y Llorens Molina (2006), las prácticas y criterios de evaluación orientan a los estudiantes hacia donde deben

dirigir su actividad de aprendizaje, en los cursos de Estadística de la carrera de Educación, ellos deben dirigir sus esfuerzos al conocimiento del significado de términos estadísticos, la utilización del lenguaje y los instrumentos básicos de la Estadística, así como la representación de datos en tablas y gráficas y el cálculo (mediante algoritmos conocidos) de medidas de Estadística descriptiva. La presencia mayoritaria de preguntas de *Alfabetización Estadística* se corresponde parcialmente con los planteamientos de GAISE (ASA, 2010), con relación a que en los cursos iniciales de Estadística, para no especialistas, se haga énfasis en ese tipo de habilidad cognitiva pero, prácticamente, deja de lado el *Razonamiento Estadístico* y el *Pensamiento Estadístico*, con lo cual es imposible cumplir con la segunda parte de esas recomendaciones, en lo que concierne al desarrollo del *Pensamiento Estadístico*.

El hecho de que la mayoría de las preguntas analizadas se ubica en los niveles bajos de las taxonomías sugiere que el centro del aprendizaje de esos cursos se encuentra en el contenido y los procedimientos, no en las ideas fundamentales de la estadística, su significado y aplicación en otros contextos. Las dos taxonomías utilizadas apuntan hacia la necesidad de definir el nivel de comprensión que se aspira logre el estudiante, alineando la evaluación con los objetivos de aprendizaje y los métodos de enseñanza.

Por tratarse de un curso básico de estadística es lógico que se haga énfasis en la Alfabetización Estadística, pero no deberían descuidarse los niveles más profundos de comprensión; se debe pensar siempre en promover el desarrollo del Pensamiento Estadístico. De la misma forma, los estudiantes deben ser expuestos a situaciones de aprendizaje donde puedan, al menos, interrelacionar ideas fundamentales de la estadística e ir más allá de las definiciones, fórmulas y cálculos.

Referencias

- American Educational Research Association, American Psychological Association y National Council of Measurement in Education. (1999). *Standards for educational and psychological testing*. Washington, DC: American Psychological Association.
- American Statistical Association. (2010). Guidelines for Assessment and Instruction in Statistics Education – GAISE. College Report. [En línea], http://www.amstat.org/education/gaise/GaiseCollege_Full.pdf

- Aoyama, K. (2007). Investigating a hierarchy of students' interpretations of graphs. *International Electronic Journal of Mathematics Education*, 4 (3), pp. 298 – 318. [En línea], <http://www.iejme.com/032007/d10.pdf>.
- Biggs, J. B. y Collis, K. F. (1982). *Evaluating the quality of learning: The SOLO taxonomy*. New York: Academic Press.
- Biggs, J.B. (1999). *Teaching for Quality Learning at University*. Buckingham. Open University Press
- Brown S. y Glasner A. (Eds.). (2003) *Evaluar en la Universidad. Problemas y nuevos enfoques*. Madrid: Narcea.
- 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), pp. 28 – 58.
- Eichler, A. (2008) Teachers' Classroom Practice and Students' Learning. En Batanero, C., Burrill, G., Reading, C. y Rossman, A. (Eds.) (2008). *Joint ICMI/IASE Study: Teaching Statistics in School Mathematics. Challenges for Teaching and Teacher Education. Proceedings of the ICMI Study 18 and IASE 2008 Round Table Conference*. Monterrey, Mexico, International Commission on Mathematical Instruction and International Association for Statistical Education.
- Godino, J. D., Contreras, A. y Font, V. (2006). Análisis de procesos de instrucción basado en el enfoque ontológico-semiótico de la cognición matemática. *Recherches en Didactiques des Mathematiques*, 26 (1), pp. 39 – 88.
- Hill, H. C., Ball, D. L. y Schilling, S. G. (2008). Unpacking pedagogical content knowledge: Conceptualizing and measuring teachers' topic-specific knowledge of students. *Journal for Research in Mathematics Education*, 39, 372 – 400.
- Jones, G. A., Langrall, C. W., Thornton, C. A. y Mogill, A. T. (1999). Student's probabilistic thinking in instruction. *Journal for Research in Mathematics Education*, 30(5), 487 – 519.
- Landín, P. R. y Sánchez, E. (2010). Niveles de razonamiento probabilístico de estudiantes de bachillerato frente a tareas de distribución binomial. *Educação Matemática Pesquisa*, 12 (3). [En línea], <http://revistas.pucsp.br/index.php/emp/article/viewArticle/4842>.
- Reading, C. y Reid, J. (2007). Reasoning about variation: Student voice. *International Electronic Journal of Mathematics Education*, 2 (3), pp. 110-127. [En línea], <http://www.iejme.com/032007/d1.pdf>
- Rodríguez Gómez, G. (2008). Realidades y retos de la evaluación del aprendizaje en la universidad. Conferencia presentada en Encuentros sobre la calidad en la Educación Superior 2008. Vigo. [Video en línea]: <http://tv.uvigo.es/es/video/26799.html>
- Salcedo Galvis, H. (2010). La evaluación educativa y su desarrollo como disciplina y profesión: presencia en Venezuela. *Revista de Pedagogía*, 31 (89), pp. 331 – 378.

- Schoenfeld, A. H. y Kilpatrick, J. (2008). Towards a theory of proficiency in teaching mathematics. En D. Tirosh & T. Wood (eds.), *Tools and Processes in Mathematics Teacher Education*, pp. 321 – 354. Rotterdam: Sense Publishers.
- Shulman, L. S. (1986). Those who understand: Knowledge growth in teaching. *Educational Researcher*, 15(2), pp. 4 – 14.
- Vallecillos, A. y Moreno, A. (2006). Estudio teórico y experimental sobre el aprendizaje de conceptos y procedimientos inferenciales en secundaria. *Tarbiya: revista de investigación e innovación educativa*. 38, 61 – 78.
- Watson, J. M. y Moritz, J. B. (2000). The longitudinal development of understanding of average. *Mathematical Thinking and Learning*, 2, 11 – 50.
- Watts, F., García-Carbonell, A. y Llorens Molina, J.A. (2006). Introducción. En F. Watts y A. García-Carbonell (Eds.). *La evaluación compartida: investigación multidisciplinar*, pp. 61 – 76. Valencia: UPV.