

Las preguntas del examen de estadística descriptiva

Salcedo, Audy

Universidad Central de Venezuela

Resumen

En este estudio se presentan los resultados del análisis de un conjunto de preguntas pertenecientes a exámenes de estadística descriptiva elaborados por distintos profesores. Se les pidió a profesores los exámenes utilizados en los dos últimos semestres para la evaluación de los aprendizajes en la asignatura Estadística Aplicada a la Educación. Se recibieron 58 exámenes que contenían 648 preguntas, las cuales fueron clasificadas utilizando la taxonomía SOLO (Biggs y Collis, 1982) y las definiciones de alfabetización, razonamiento y pensamiento estadístico (delMas, Ooms, Garfield y Chance, 2006); además fueron catalogadas por el puntaje asignado en la evaluación y el nivel de dificultad estimado. Los resultados indican que las preguntas de estadística descriptiva analizadas son en su mayoría valoradas con puntuaciones bajas, tienen un bajo nivel de dificultad y son de baja exigencia cognitiva, según las taxonomías utilizadas. Parece razonable suponer que preguntas como las analizadas ayudan a promover una visión mecanicista de la estadística descriptiva.

Palabras clave: Preguntas estadística descriptiva, Taxonomía SOLO, alfabetización estadística, razonamiento estadístico y pensamiento estadístico

1. Introducción

Este trabajo tiene como objetivo el análisis de un conjunto de preguntas pertenecientes a exámenes de estadística descriptiva elaborados por distintos profesores de la asignatura Estadística Aplicada a la Educación.

En la evaluación de los aprendizajes, el examen (o prueba escrita) es el instrumento de evaluación dominante en la educación universitaria. Salcedo Galvis (2010) señala que, a pesar de los cambios que han acompañado a la evaluación, el examen es el instrumento de evaluación predominante en Venezuela. Entonces, es innegable el peso que tiene el examen para la toma de decisiones en la promoción y acreditación de los alumnos en la educación universitaria, particularmente en las asignaturas del área cuantitativa, donde, por tradición, es el principal instrumento de evaluación. Es por ello que esta investigación se centra en los exámenes de Estadística como objeto de estudio, sin subvalorar otros instrumentos de evaluación.

La investigación en evaluación en Educación Estadística (EE) ha recibido mayor atención en los últimos años, fundamentalmente para conocer el tipo de respuesta que proporcionan los estudiantes. La taxonomía Student Observed Learning Outcomes (SOLO), propuesta por Biggs y Collis (1982), es una de las utilizadas para categorizar las habilidades cognitivas alcanzadas por los estudiantes, y se aplicada a diferentes áreas de conocimientos. Educadores estadísticos han propuesto taxonomías para caracterizar la ejecución de los estudiantes en tópicos específicos de la Estadística, por ejemplo, Watson y Moritz (2000), Vallecillos y Moreno (2006) 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 formule el profesor. Cada examen, expresado en las distintas preguntas que lo conforman,

contiene una muestra de las habilidades cognitivas que el docente desea conocer si el estudiante logró desarrollar. Por lo 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. De acuerdo con la recomendación de GAISE (ASA, 2010), los cursos iniciales de Estadística para no especialistas, como es el caso que nos ocupa, deben hacer énfasis en la Alfabetización Estadística con miras a desarrollar el Pensamiento Estadístico. Es por ello que en esta investigación se analizan los exámenes aplicados en la asignatura Estadística, se trata el caso particular de la carrera de Educación. 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 pretenden evaluar.

2. La investigación

A profesores que dictan asignaturas de Estadística, en la carrera de Educación, se les solicitó el conjunto de exámenes utilizados en dos últimos semestres, para la evaluación de los aprendizajes en esa asignatura. Se recibió respuesta de 10 profesores, quienes proporcionaron 58 exámenes que contenían 648 preguntas. Los profesores tienen entre 4 y 20 años de experiencia dictando la asignatura. Todas las preguntas fueron resueltas para poder establecer con claridad los procesos cognoscitivos que implicaban y luego poder clasificarlas.

Las preguntas se clasificaron de acuerdo con el nivel de dificultad estimado según juicio de experto, y las habilidades cognitivas que pretenden evaluar, para lo cual se utilizó una adaptación de la taxonomía SOLO y las definiciones de alfabetización, razonamiento y pensamiento estadístico de delMas, Garfield, Ooms, y Chance (2007). Además, se consideró la puntuación asignada por el profesor a cada pregunta. En el caso de la taxonomía SOLO (Biggs y Collis, 1982), se realizó una adaptación para estimar el nivel de respuesta que aspira el profesor que proporcione el estudiante. Las definiciones utilizadas son:

- Nivel uniestructural. 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. 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. La pregunta requiere del análisis de la información, establecer relaciones entre los diferentes elementos 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. La pregunta exige la utilización de un principio general y abstracto que puede ser inferido del análisis de la información presente en el enunciado, e integrar información adquirida para generar una nueva que permite solucionar la situación.

Las preguntas también se clasificaron desde un marco más específico de la Educación Estadística, a partir de las definiciones de alfabetización, razonamiento y pensamiento estadístico 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, reconocimiento e interpretación

de la representación de datos, conocer el desarrollo de la capacidad para interpretar y evaluar críticamente la información estadística, así como comunicar ideas estadísticas.

- Razonamiento estadístico. Se dirige a conocer 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. Reconocer cómo, cuándo y por qué pueden ser utilizados los instrumentos de inferencia estadística existentes. 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 planificar y evaluar investigaciones y extraer conclusiones. Implica la comprensión de por qué y cómo se realizan las investigaciones.

Las preguntas también fueron clasificadas por la puntuación asignada y se estimó su nivel de dificultad, mediante el 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 son: muy fácil (100% – 81%), fácil (80% – 61%), promedio (60% – 41%), difícil (40% – 21%) y muy difícil (20% – 1%).

Para validar la calificación realizada se les proporcionó a dos profesores de estadística el conjunto de preguntas, las respuestas y las definiciones utilizadas para la clasificación. Los profesores realizaron de forma individual su propia clasificación, se encontró divergencia en aproximadamente el 38% de las preguntas respecto a la realizada por el autor. Se discutió con los dos profesores sobre las preguntas donde hubo divergencias, hasta lograr un acuerdo sobre su clasificación.

3. Los resultados

La casi totalidad de las preguntas tienen 3 o menos puntos, en exámenes con 20 como puntuación máxima. Es importante destacar que se encontraron casos atípicos y extremos que tienen puntuaciones entre 4 y 15 puntos. El 75% de las preguntas tienen 2 puntos o menos, por lo tanto, se puede pensar que la mayoría de las preguntas son de baja exigencia cognitiva, probablemente referidas a la reproducción de información o aplicación de procedimientos básicos de cálculo.

En relación con el nivel de dificultad, estimado por juicio de experto, según el porcentaje de estudiante que se espera responda correctamente. El 81,1% de las preguntas se catalogaron como Muy fácil o fácil, por lo tanto se espera que sean respondidas correctamente por al menos el 80% de los estudiantes. El 17,3% de las preguntas se clasificó como de dificultad promedio (60% – 41% respuestas correctas esperadas) y el 1,5% como difícil (40% – 21%), no se encontraron preguntas Muy difíciles. Este resultado parece consistente con el referido a las puntuaciones con que son valoradas las respuestas.

La Tabla 1 muestra los resultados de las preguntas clasificadas mediante la adaptación de la taxonomía SOLO, con lo cual se puede estimar el máximo nivel de respuesta que podría evidenciar el estudiante al proporcionar sus respuestas.

Tabla 1. Preguntas clasificadas por taxonomía SOLO

Nivel	Frecuencia	Porcentaje
Uniestructural (UN)	498	77,1
Multiestructural (NM)	141	21,8
Relacional (NR)	7	1,1
Abstracción extendida (NAE)	0	0,0

Total	646	100
-------	-----	-----

En la casi totalidad de las preguntas analizadas el estudiante solo requiere de habilidades cognitivas bajas para dar la respuesta correcta. En más del 75% de las preguntas se puede ofrecer la respuesta recordando información o al aplicar un procedimiento o algoritmo, previamente conocido, con el uso de la información que se encuentra explícita en el enunciado. Pareciera, entonces, que este grupo de preguntas son fundamentalmente para conocer si los estudiantes pueden recordar o reproducir información proporcionada en clase (NU), así como clasificar o describir la comprensión de hechos sin sistematizarlos (NM). Obsérvese que las preguntas del NR y el NAE son prácticamente inexistentes, lo cual indica que lo prioritario es el cálculo y los procesos mecánicos. A continuación, presentamos ejemplos de preguntas clasificadas en cada uno de los niveles.

Uniestructural

Pregunta 392. ¿Qué medidas de posición deberíamos calcular si deseamos obtener los puntajes de la prueba para el 50% central? (1 pto.)

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. (1,5 ptos.)

Multiestructural

Pregunta 6. 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 coeficiente de variación (3 ptos)

Pregunta 241. Un investigador está interesado en conocer si existe relación entre el género y el gusto por cocinar. Se pregunto a 90 personas con los siguientes resultados: 40 pertenecen al sexo masculino y 20 de los opinan que les gustan cocinar, de las mujeres encuestadas a 15 no les gusta cocinar. Calcula el coeficiente de correlación apropiado para establecer si existe relación lineal entre las variables e intérprete los resultados. (3 ptos).

Relacional

Pregunta 305. A continuación se presentan las calificaciones del primer parcial de Estadística aplicada a la Educación: 17 13 08 15 14 11 13 18 19 20 14 17 16 18 12 16 14 15 19 17 02 04 19 10 06 18. Si un estudiante obtuvo una calificación de 14 puntos, ¿cómo se encuentra él con respecto a su grupo? (2 ptos.)

De acuerdo con la clasificación, se considera que el 91,5% de las preguntas examinadas evalúa logros en la *Alfabetización Estadística*, mientras que un 8,5% lo hace en *Razonamiento Estadístico*, no se clasificaron preguntas en *Pensamiento Estadístico*.

Tabla 2. Preguntas clasificadas por el aprendizaje esperado

Nivel de Dificultad	Frecuencia	Porcentaje
Alfabetización (AE)	591	91,5
Razonamiento (RE)	55	8,5
Pensamiento (PE)	0	0,0
Total	646	100

La mayoría de las preguntas se ubica en el nivel más bajo de la taxonomía utilizada, 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 de GAISE (ASA, 2010), en las que se indica que en los cursos iniciales de Estadística se debe hacer énfasis en la *AE*, no obstante, deja de lado los otros niveles de la taxonomía, por lo menos en las evaluaciones escritas examinadas. Por tratarse de un curso básico, es lógico que se haga énfasis en la *AE*, pero no debería descuidarse los niveles más profundos de habilidades cognitivas. A continuación ejemplos de las preguntas en categoría:

Alfabetización

Pregunta 3. ¿Cómo interpretaría usted un coeficiente de Kurtosis de -0.216 puntos? (1 pto.)

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 pts)

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. A partir de los datos presentados, dibuje el diagrama de dispersión y analice el tipo de relación existente. (3 puntos)

Estadística I: 06, 12, 16, 10, 10, 09, 10, 15, 10, 11, 19, 12

Estadística II: 09, 11, 16, 13, 12, 08, 11, 13, 10, 15, 20, 14

Razonamiento

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. (3 puntos)

	A	B	A	B
N			65	70
Media			73	62
Mediana	60,00	61,50	52	50
Modo			48	72
Desviación			74	66
Varianza	96,879	103,820	61	66
Asimetría (α^1)	0,386		80	51
Kurtosis (α^2)		-1,650	75	48
Valor mínimo	46	44	47	46
Valor máximo	75	73	75	50
			48	44
			53	73
			46	61
			80	66
			56	46
			80	71
			50	66
			60	55

4. Conclusiones

Las preguntas analizadas son en su mayoría valoradas con puntuaciones bajas, casi todas tienen 3 puntos o menos, exámenes con 20 como puntuación máxima, con un bajo nivel de dificultad. Las habilidades cognitivas evaluadas mediante las preguntas se corresponden con los niveles más bajos de las taxonomías utilizadas.

La mayoría de las preguntas se consideran de *AE* en el *NU*, por lo tanto, se puede afirmar que buscan evaluar habilidades referidas a la reproducción de conocimiento y procedimientos algorítmicos sencillos. Se está preguntando aspectos muy básicos de la *AE*. Son escasas las preguntas que buscan evaluar el conocimiento profundo y la aplicación de las ideas Estadísticas. Estos resultados sugieren que en general las preguntas de estadística descriptiva analizadas son de bajo nivel de dificultad y baja exigencia cognitiva. Es posible

que ese sea el nivel del curso previsto en los programas de estudio, que sean decisiones particulares de los profesores o que simplemente es la práctica institucional.

Es probable que la ausencia de preguntas de *PE* se deba a que su definición tiene un sesgo hacia los procesos de inferencia estadística, pero también puede deberse a que en estos exámenes de estadística descriptiva, no se incluyan aspectos como: por qué y cómo se realizan las investigaciones, utilizar el contexto de un problema, evaluar resultados de investigaciones.

Parece razonable suponer que las preguntas de bajo nivel cognitivo ayudan a promover una perspectiva mecanicista de la estadística descriptiva. Si esos cursos no tienen previstas actividades de aprendizaje y evaluación de alto nivel cognitivo, muy posiblemente los estudiantes no logren una visión más integral de esta rama de la Estadística y de su utilidad para la investigación en Educación.

Referencias

- 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
- Biggs, J. B. y Collis, K. F. (1982). *Evaluating the quality of learning: The SOLO taxonomy*. New York: Academic Press.
- 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.
- 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>.
- Salcedo Galvis, H. (2010). La evaluación educativa y su desarrollo como disciplina y profesión: presencia en Venezuela. *Revista de Pedagogía*, 31 (89), 331 – 378.
- 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.