Análisis de las respuestas de estudiantes de bachillerato a problemas sobre la distribución normal

Analysis of high school students' responses to problems on the normal distribution

Julio César Valdez Monroy y Jesús Salinas Herrera

Colegio de Ciencias y Humanidades, UNAM, México

Resumen

En este trabajo se utiliza la taxonomía SOLO para analizar las respuestas de dos grupos de estudiantes de bachillerato a tres problemas sobre la distribución normal. Para ello, se consideran dos ciclos de aprendizaje UMR: el primero relacionado con la estandarización, y el segundo con el cálculo de probabilidades. Entre los resultados se observa que la mayoría de los estudiantes determina y relaciona de forma adecuada los valores *z*, los *p* valores asociados y las áreas correspondientes bajo la curva normal, pero sólo a nivel procedimental. Conceptualmente, presentan dificultades para reconocer el significado de algunas etapas en el procedimiento. Como conclusión se menciona la necesidad de diseñar actividades para desarrollar el entendimiento conceptual de los estudiantes, en particular, mediante la integración adecuada del trabajo escrito y el recurso computacional.

Palabras clave: estudiantes de bachillerato, distribución normal, taxonomía SOLO

Abstract

In this work, we use the SOLO taxonomy to analyse the responses provided by two groups of high-school students to three problems concerning the normal distribution. Two UMR learning cycles are considered: the first one is related to the standardization process and the second one, to the use of the normal standard distribution to calculate probabilities. Among the results, we observed that most students determine and adequately relate z values, associated p values, and corresponding areas under the normal curve, but only at procedural level. Conceptually, they show difficulties to identify the meaning of some stages in the procedure. In conclusion, we consider the need to design activities to develop the students' conceptual understanding, particularly through the adequate integration of written activities and the computational resource.

Keywords: high school students, normal distribution, SOLO taxonomy

1. Introducción

La distribución normal es la primera distribución continua que comúnmente se ve en los cursos de estadística y probabilidad. Su importancia radica en que es un modelo que describe diversos fenómenos físicos y sociales, además de que dentro de la misma materia es clave para la transición del análisis de datos a la inferencia estadística (Batanero, Tauber y Sánchez, 2004). No obstante, tal parece que es poca la literatura de investigación en la que se analice cómo es la enseñanza y el aprendizaje de este tema, en particular, con estudiantes de bachillerato.

Los pocos trabajos encontrados hasta el momento se ubican en el nivel universitario, en cursos introductorios de estadística (Batanero et al., 2001; 2004; Carpio, Gaita, Wilhelmi y Sáenz, 2009), o se llevan a cabo con profesores que participan en cursos de actualización (Bansilal, 2014). De esta manera, con el presente trabajo se busca obtener información acerca de cómo es el razonamiento de los estudiantes de bachillerato al resolver problemas sobre la distribución normal, para lo cual se utiliza la taxonomía SOLO (structure of the observed learning outcome) (Biggs y Collis, 1982), ya que es

Valdez Monroy, J. C. y Salinas Herrera, J. (2019). Análisis de las respuestas de estudiantes de bachillerato a problemas sobre la distribución normal. En J. M. Contreras, M. M. Gea, M. M. López-Martín y E. Molina-Portillo (Eds.), *Actas del Tercer Congreso Internacional Virtual de Educación Estadística.* Disponible en www.ugr.es/local/fqm126/civeest.html

una herramienta de análisis que permite tener una imagen amplia del nivel de sofisticación de la respuestas.

2. Antecedentes

El concepto de distribución es fundamental para el razonamiento estadístico (Wild, 2006). De acuerdo con Reading y Canada (2011), un mejor conocimiento de este concepto "está estrechamente relacionado con el conocimiento de los conceptos de los que depende y también con el conocimiento de los conceptos que dependen de él" (p. 233). Bajo este escenario, la ausencia, o falta de consolidación, de los conocimientos previos relacionados con la distribución normal es una de las principales dificultades que se han detectado en la literatura. Al respecto, Batanero et al. (2004) analizaron los resultados de un experimento de enseñanza llevado a cabo con 117 estudiantes universitarios y encontraron que la complejidad de la distribución normal se debe a que su comprensión requiere la integración y la relación de muchos y diferentes conceptos e ideas estadísticas que resultan igualmente complejas; probabilidad, curva de densidad, dispersión, asimetría e histograma. En esta misma línea, Carpio et al. (2009) analizaron la desadaptación de los significados personales de 45 estudiantes universitarios respecto de los significados institucionales del concepto de distribución normal. Encontraron que los significados pretendidos no siempre se correspondieron con los implementados debido a que los alumnos no manejaban algunos conceptos previos relacionados con la estadística y con el cálculo.

Por otro lado, Bansilal (2014) llevó a cabo un estudio exploratorio con 290 profesores de bachillerato en servicio que participaban en un programa de actualización destinado al desarrollo de su conocimiento sobre la distribución normal. Entre los resultados, observó cómo sólo un pequeño porcentaje de profesores fue capaz de responder de forma correcta a dos preguntas de un problema tipo sobre distribución normal (27% y 14%). Las principales dificultades detectadas fueron el reconocimiento de la probabilidad como el área bajo la curva normal e identificar la proporción de área limitada por dos valores z. La ausencia de computadoras es señalada por la autora como una limitante del programa de formación. Al respecto menciona: "Las experiencias con la simulación de datos y la exposición a muchos applets interesantes puede ayudar a estos profesores a experimentar el fenómeno abarcado por el formalismo, tal vez llevándolos a un conocimiento más profundo" (p. 55).

La idea de apoyarse en el recurso computacional también es mencionada por Batanero et al. (2004), quienes consideran que puede facilitar la exploración de las propiedades y las representaciones relacionadas con la distribución normal, lo que puede proporcionar a los estudiantes una experiencia didáctica difícil de alcanzar en el mundo real. Además, mencionan que las tareas de evaluación apoyadas en este recurso pueden proporcionar una imagen completa de su entendimiento y formas de razonar. Asimismo, Carpio et al. (2009) señalan que "el uso de un software reduciría el tiempo empleado en procedimientos rutinarios, lo que permitiría introducir más preguntas relacionadas con la comprensión de la distribución normal y de sus propiedades, así como preguntas donde se deban interpretar los resultados" (2009, sección de Discusión de resultados, párr. 5). A partir de esta breve revisión de la literatura, se pueden prefigurar las posibles respuestas que pueden proporcionar los estudiantes de bachillerato. En particular, el trabajo de Bansilal (2014) es una guía, ya que el instrumento que se empleó para la recolección de datos, así como el instrumento de análisis, son similares a los utilizados en el presente estudio.

3. Marco de referencia

En el presente estudio se utiliza la taxonomía SOLO (*structure of the observed learning outcome*) (Biggs y Collis, 1982) para analizar las respuestas de los estudiantes con la finalidad de conocer 'qué tan bien', en vez de 'qué tanto', han aprendido (Pegg, 2002) sobre la distribución normal, después de haber recibido una enseñanza sobre el tema. Esta taxonomía permite categorizar las respuestas con base en los siguientes niveles jerárquicos: Pre-estructural (P), puede haber un intento de vincular la indicación con la respuesta mediante un elemento irrelevante; Uni-estructutal (U), sólo se toma en cuenta un elemento relacionado con la tarea; Multi-estructural (M), se consideran varios elementos, pero de forma aislada; Relacional (R), se manejan e integran varios elementos relacionados con la tarea; y Abstracción extendida (AE), hay una manipulación de lenguaje simbólico y conceptos formales.

Los niveles anteriores conforman un ciclo de aprendizaje que se presenta a través de los modos de representación que caracterizan el desarrollo cognitivo: sesoriomotriz, icónico, concreto/simbólico, formal y postformal (Biggs y Collis, 1991). El ciclo conformado por los niveles uni-estructural, multi-estructural y relacional (UMR) se repite en cada modo. El nivel pre-estructural (P) indica que el sujeto aún no cuenta con los elementos necesarios para acceder al siguiente modo. Por su parte, el nivel de abstracción extendida (AE) indica una transición al modo superior. En otros estudios se ha identificado que dentro de cada modo se presentan al menos dos ciclos UMR (Pegg, 2002), donde una respuesta de nivel relacional en un ciclo evoluciona a un nuevo nivel de respuesta uniestructural en un siguiente ciclo dentro del mismo modo.

3.1. Jerarquía para el análisis de las respuestas

Con base en el trabajo de Bansilal (2014), quien identifica la presencia de dos niveles (capas) de respuesta a problemas tipo sobre la distribución normal, en el presente estudio se consideran dos ciclos de aprendizaje UMR: el primero relacionado con el proceso de estandarización, y el segundo con la utilización de la distribución normal estándar para el cálculo de probabilidades.

Ciclo 1: estandarización

U_1: Se identifican algunos de los elementos involucrados en el problema, pero se establece una relación no relevante entre ellos.

M_1: Se identifican todos los elementos involucrados en el problema. Sin embargo, se relacionan de forma incorrecta mediante la expresión $z = (x - \mu) / \sigma$.

R1: Se identifican todos los elementos involucrados con el problema y se relacionan de forma correcta mediante la expresión $\mathrm{z} = \mathrm{x} - \mu/\sigma$. Sin embargo, no van más allá del cálculo de *z*, o le dan un significado distinto.

Ciclo 2: cálculo de la probabilidad

U_2: Se calcula correctamente el valor *z*, pero el *p* valor correspondiente no es identificado de forma apropiada.

M_2: Se asocia un *p* valor correcto a *z*, pero no logran utilizar *p* para encontrar la probabilidad solicitada.

R_2: Se asocia un *p* valor correcto a *z*, el cual es utilizado para encontrar la probabilidad solicitada.

4. Método

En el estudio participaron dos grupos de bachillerato (17-18 años) conformados por 20 (A) y 33 (B) alumnos, quienes se encontraban tomando el curso de estadística y probabilidad II. Previo a la evaluación, ambos grupos vieron el tema de distribución normal, aunque con un enfoque distinto; el principal recurso empleado por el profesor que atendió al grupo A fue el software fathom, mientras que la profesora que trabajó con el grupo B se apoyó principalmente en actividades escritas. En Salinas, Valdez y Salinas-Hernández (2018) se hace una descripción amplia de ambos tipos de enseñanza.

Para la recolección de datos se utilizó un cuestionario conformado por siete problemas, los cuales fueron tomados de un instrumento de evaluación institucional denominado *examen de diagnóstico académico* (EDA). La elección de este instrumento se debió a que los resultados que arroja muestran un nivel bajo en el rendimiento de los estudiantes sobre el tema de distribución normal. Debido a la similitud entre algunos problemas, en el presente estudio sólo se analizan los siguientes tres.

> **Problema 1.** Se sabe que una persona elegida al azar requiere de 555 horas para completar un trabajo; si se conoce que estos tiempos se distribuyen normalmente con media de 500 horas y con desviación estándar de 100 horas, ¿cuál es el valor estandarizado que corresponde a este tiempo?

Para dar respuesta al problema 1, primero se identifica la variable x (*horas para completar el trabajo*) y cómo se distribuye (*normal con* $\mu = 500$ *horas y* $\sigma = 100$ *horas*). Enseguida, se identifica el valor objetivo de la variable ($x = 555$ *horas*), para posteriormente transformarlo en un valor z mediante la expresión $z = (x - \mu) / \sigma$. De esta manera, $z = (555 - 500) / 100 = 0{,}55$. Aquí es necesario reconocer que la estandarización permite transitar de una situación particular a una general, en la que la nueva variable z también se distribuye de forma normal con $\mu = 0$ y $\sigma = 1$ (Figura 1).

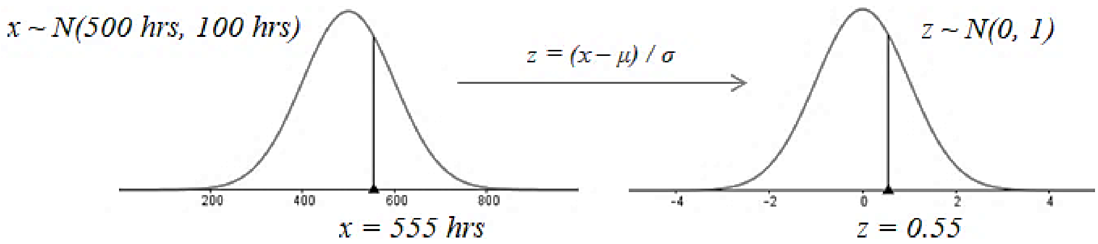

Figura 1. Representación gráfica del proceso de solución del problema 1

> **Problema 2**. El porcentaje de grasa de un tipo de queso es una variable aleatoria que sigue una distribución normal con media 4,30% y desviación estándar 0,52%. Para una porción de este queso, elegida al azar, la probabilidad de que su contenido de grasa sea menor que 4,82% es:

La respuesta normativa al problema 2 consiste en identificar la variable x (*porcentaje de grasa*) y cómo se distribuye (*normal con* $\mu = 4{,}30$ *y* $\sigma = 0{,}52$). Después, se identifican los valores de la variable que satisfacen la condición dada ($x < 4{,}82$) y se asocia la probabilidad con el área bajo la curva definida por los valores que satisfacen dicha condición ($P[x < 4{,}82]$) (Figura 2a).

En seguida, se transforma el valor objetivo x en un valor z mediante la expresión $z = (x - \mu) / \sigma$. De esta manera, $P[x < 4{,}82] = P[(x - \mu) / \sigma < (4{,}82 - 4{,}30) / 0{,}52] = P[z < 1]$, donde la nueva variable z se distribuye de forma normal con $\mu = 0$ y $\sigma = 1$. Así, la probabilidad de esta variable corresponde al área bajo la curva definida por los valores z que satisfacen la condición $z < 1$ (Figura 2b).

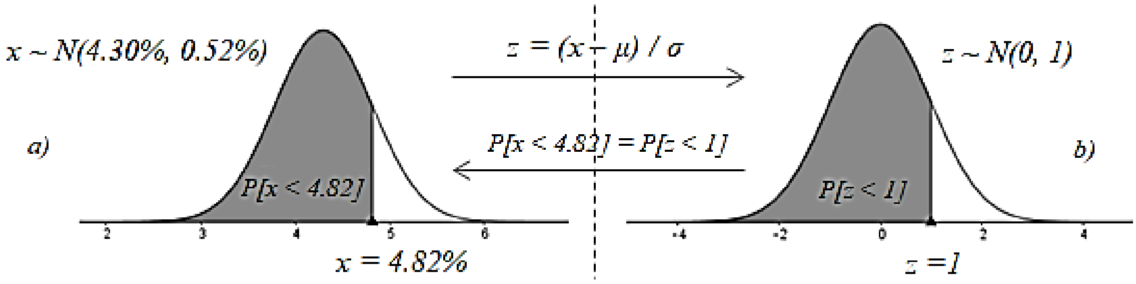

Figura 2. Representación gráfica del proceso de solución del problema 2

Después, a partir de la tabla de la distribución normal estándar, se asocia la probabilidad correspondiente al valor z ($z_{(1)} = 0{,}8413$, con la tabla completa, o $z_{(1)} = 0{,}3413$, con la mitad de la tabla) y ésta es empleada para determinar la probabilidad que se solicita; $P[z < 1] = 0{,}8413$ o $P[z < 1] = 0{,}3413 + 0{,}5 = 0{,}8413$. Finalmente, el resultado se asocia con la distribución original ($P[x < 4{,}82] = 0{,}8413$) (Figura 2).

> **Problema 3.** Si los contenidos de grasa de las piezas de carne de una bodega siguen una distribución aproximadamente normal con media de 100 gramos y desviación estándar de 15 gramos. ¿Cuál es la probabilidad de que una pieza de carne elegida aleatoriamente tenga un contenido de grasa entre 70 y 112 gramos?

La solución al problema 3 requiere la estandarización de dos valores de la variable x (*contenidos de grasa*) a través de la expresión $z = (x – \mu) / \sigma$. De esta manera, $P[70 < x < 112] = P[(70 – 100) / 15 < (x – \mu) / \sigma < (112 – 100) / 15] = P[-2 < z < 0{,}8]$. Después, a partir de la tabla, se asocian las probabilidades correspondientes a los valores z ($z_{(-2)} = 0{,}0228$ y $z_{(0,8)} = 0{,}7881$, tabla completa, o $z_{(-2)} = 0{,}4772$ y $z_{(0,8)} = 0{,}2881$, mitad de la tabla) y estas son empleadas para calcular la probabilidad que se solicita; $P[-2 < z < 0{,}8] = 0{,}7881 – 0{,}0228 = 0{,}7653$ o $P[-2 < z < 0{,}8] = 0{,}4772 + 0{,}2881 = 0{,}7653$. Finalmente, el resultado se asocia con la distribución original ($P[70 < x < 112] = 0{,}7653$) (Figura 3).

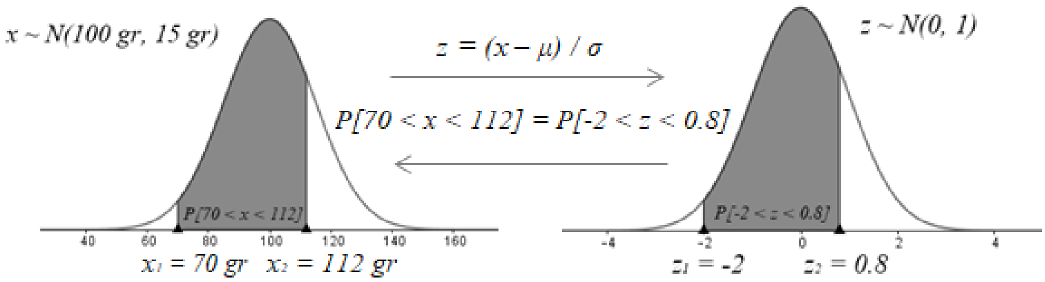

Figura 3. Representación gráfica del proceso de solución del problema 3

La aplicación del cuestionario se hizo dentro del horario de clase (en una sesión de 120 minutos para cada grupo), como evaluación del tema correspondiente a la distribución normal. En el caso del grupo A, el número de estudiantes era de 20, pero se redujo a 16 debido a que se decidió considerar sólo a quienes estuvieron presente en las cuatro sesiones (de 120 minutos cada una) en las que se trabajó el tema.

5. Análisis de resultados

En las tablas 1, 2 y 3 se muestra el número de respuestas dadas por los estudiantes a cada uno de los problemas, las cuales se clasifican de acuerdo con las jerarquías elaboradas mediante la taxonomía SOLO. Asimismo, solo se proporcionan algunos ejemplos con la finalidad de destacar aspectos relevantes de las respuestas.

Problema 1. De acuerdo con lo que se pide en el problema 1, el nivel más alto que se puede alcanzar en la jerarquía es U_2, ya que sólo es necesario calcular el valor *z*.

Tabla 1. Niveles de respuesta al problema 1

	U_1	M_1	R_1	U_2	M_2	R_2
Grupo A	1	2	12	1	–	–
Grupo B	0	1	25	7	–	–
Total	1	3	37	8	–	–

En la Tabla 1 se observa que la mayoría de las respuestas se ubican en el nivel R_1, en el cual se identifican todos los elementos involucrados en el problema y se relacionan de forma correcta mediante la expresión $z = (x - \mu) / \sigma$. Sin embargo, en todas se va más allá del cálculo de *z* (Figura 4).

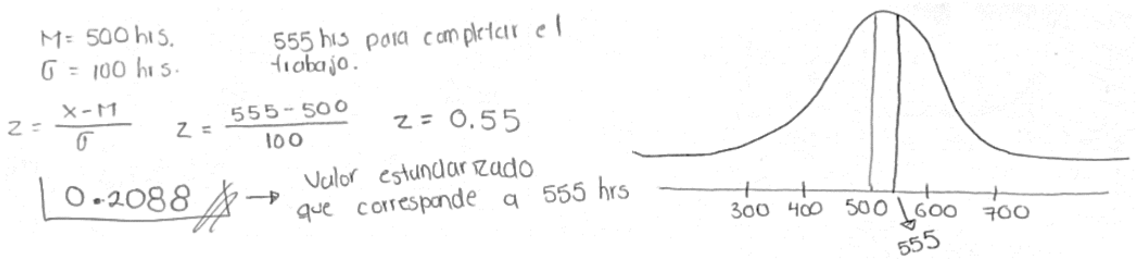

Figura 4. Respuesta de nivel R_1 al problema 1

En la respuesta se aprecia cómo el estudiante identifica la variable *x* así como el valor objetivo *x = 555 horas*. Incluso lo hace de forma gráfica, indicando también el comportamiento de la variable (*normal*). Enseguida, determina el valor *z* correspondiente ($z = 0{,}55$) al valor objetivo. Sin embargo, asocia una probabilidad a *z* (*0,2088*), la cual considera es el valor estandarizado.

Una respuesta similar fue proporcionada por la mayoría de los 37 estudiantes que se ubicaron en el nivel R_1, quienes no se percataron que su solución en realidad correspondía a $P[500 < x < 555]$ o a $P[555 < x]$. No obstante, ocho dieron la respuesta correcta (Figura 5), la cual se ubica en el nivel U_2, ya que el no asociar un *p* valor a *z* es indicio de la relación que guardan estos dos elementos.

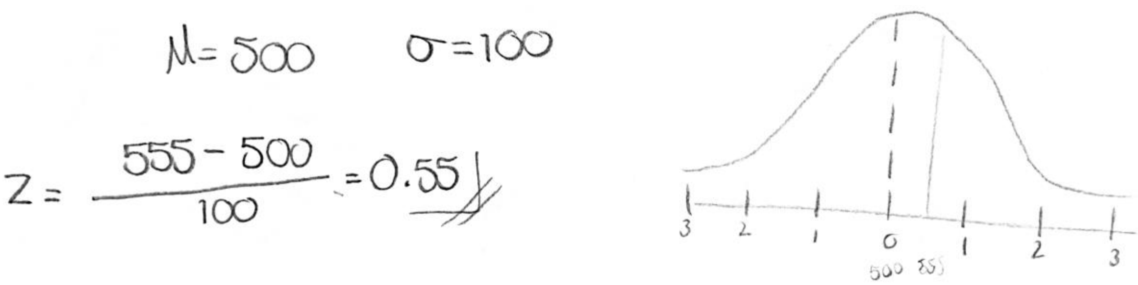

Figura 5. Respuesta de nivel U_2 al problema 1

Problema 2. En la Tabla 2 se muestra un rendimiento adecuado por parte de los estudiantes, ya que la mayoría de las respuestas (31) alcanzan el nivel R_2; se establece una relación correcta entre *z*, el *p* valor y el área bajo la curva normal estándar correspondiente a este último valor.

Tabla 2. Niveles de respuesta al problema 2

	U_1	M_1	R_1	U_2	M_2	R_2
Grupo A	3	1	1	1	3	5
Grupo B	1	0	0	0	5	26
Total	4	1	2	1	8	31

Una respuesta ubicada en el nivel R_2 es la que se muestra en la Figura 6. Además de seguir el procedimiento normativo para responder el problema, el estudiante ofrece otra solución que se apoya en la regla empírica. Al parecer observa que el valor $x = 4{,}82$ se encuentra a una desviación estándar (*0,52*) de la media (*4,30*), por lo que se apoya en la propiedad $\mu \pm \sigma \rightarrow 68\%$. De esta manera, puede ser que sólo haya sumado al 68% el porcentaje correspondiente a la cola izquierda (*16%*), lo que da como resultado la respuesta que destaca con el término 'Regla empírica' (*84%*).

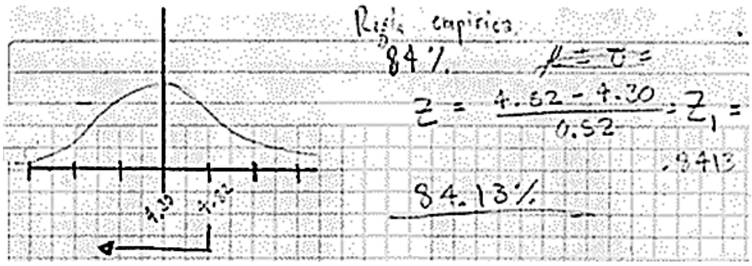

Figura 6. Respuesta de nivel R_2 al problema 2

Razonamientos como el anterior fueron característicos de varios de los estudiantes del grupo A, ya que al parecer trataban de reproducir la manera en cómo se llevó a cabo la enseñanza del tema; primero trabajaron la regla empírica y después, cuando esta regla no era aplicable, pasaron a la distribución normal estándar. Una respuesta en la que se aprecia de forma clara esta afirmación es la que se muestra en la Figura 7.

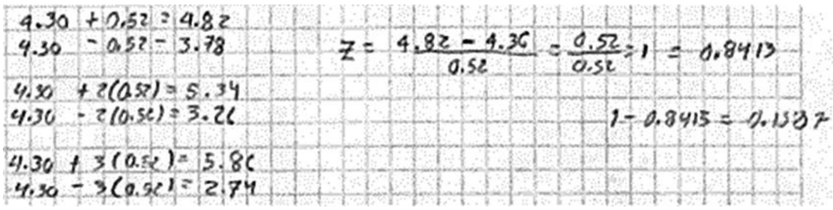

Figura 7. Respuesta de nivel M_2 al problema 2

En la respuesta se observa que el estudiante analiza los tres casos de la regla empírica ($\mu \pm \sigma \rightarrow 68\%$, $\mu \pm 2\sigma \rightarrow 95\%$, $\mu \pm 3\sigma \rightarrow 99{,}7\%$) y al no poder utilizarlos para dar respuesta al problema, decide resolverlo mediante el procedimiento normativo. Sin embargo, no utiliza el *p* valor de forma adecuada para encontrar la probabilidad que se solicita, por lo que la respuesta se ubica en el nivel M_2. Aunque respuestas como esta no se presentaron en el grupo B, el rendimiento de este grupo fue mejor, pues la mayoría alcanzó los niveles R_2 y M_2.

Problema 3. En la Tabla 3 de nueva cuenta se aprecia un buen rendimiento de los estudiantes, ya que la mayoría de las respuestas se ubican en el nivel R_2 (30). En general, los resultados son similares a los de la Tabla 2, a pesar de que en el problema 3 intervienen dos valores de la variable *x*, lo que posiblemente lo haría más complejo. Asimismo, los resultados del grupo A son más bajos y dispersos que los del grupo B.

Tabla 3. Niveles de respuesta al problema 3

	U_1	M_1	R_1	U_2	M_2	R_2
Grupo A	3	2	1	2	2	5
Grupo B	0	2	0	0	5	25
Total	3	4	1	2	7	30

Nuevamente hay respuestas en el grupo A que tratan de aprovechar la propiedad de la regla empírica para resolver el problema. Por ejemplo, en la Figura 8 se observa cómo la alumna combina dos casos de esta regla ($\mu \pm \sigma \rightarrow 68\%$, $\mu \pm 2\sigma \rightarrow 95\%$), pero cambia uno de los valores objetivo de la variable; $x = 112$ por $x = 115$. De acuerdo con el gráfico que proporciona, puede ser que, como señala, trató de dar una respuesta ‘aproximada’, ya que el área sombreada bajo la curva, al parecer, está limitada por $x = 70$ y $x = 112$. No obstante, de acuerdo con la categorización, esta respuesta se ubica en el nivel M_1, ya que no determina el valor z adecuado, correspondiente a $x = 115$.

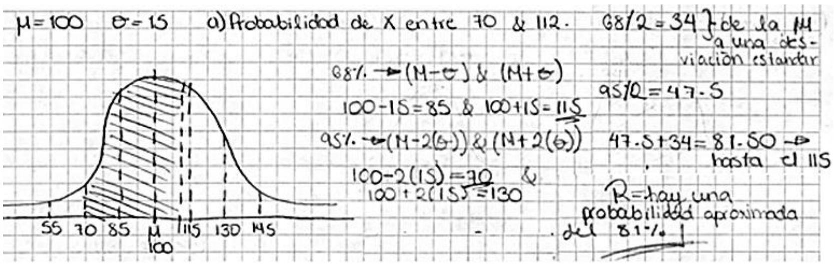

Figura 8. Respuesta de nivel M_1 al problema 3

Así como en la respuesta anterior, hay otras en las que la representación gráfica que se utiliza permite entender mejor el razonamiento involucrado. Por ejemplo, en la Figura 9 se observa un procedimiento analítico correcto, lo que supondría una respuesta de nivel R_2. Sin embargo, el área sombreada en la representación gráfica no corresponde con la probabilidad calculada, por lo que la respuesta se ubica en el nivel M_2, ya que no relaciona de forma correcta el área con la probabilidad. Además, como la mayoría de los estudiantes, utiliza la misma representación para la curva de densidad empírica y la teórica, lo que puede ser un indicio de la poca claridad de la diferencia entre ambas curvas (Batanero et al., 2004).

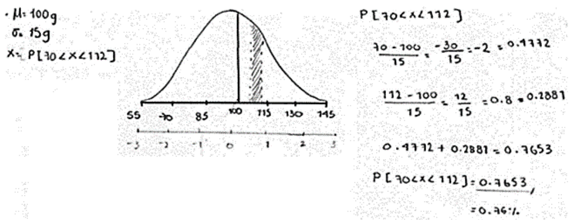

Figura 9. Respuesta de nivel M_2 al problema 3

En general, el rendimiento del grupo B fue mejor que el del grupo A. En el primero, la mayoría de las respuestas a los tres problemas se ubicaron en los niveles más altos de la jerarquía. En el caso del grupo A, el rendimiento fue más disperso, con una ligera ventaja del número de respuestas ubicadas en los niveles superiores. No obstante, en ambos grupos se presentaron respuestas que cayeron fuera de los dos ciclos UMR, las cuales se ubicaron en un nivel preestructural, ya que se consideraban algunos elementos involucrados en los problemas, pero no se estableció relación alguna entre ellos. De este tipo de respuesta hubo dos en el grupo A y una en el grupo B en el problema 2, y una en cada grupo en el problema 3.

6. Conclusiones

La taxonomía SOLO permitió tener una imagen amplia del tipo de respuestas que expusieron los estudiantes a problemas sobre la distribución normal. En general, son capaces de aplicar el procedimiento normativo de forma adecuada (Carpio et al., 2009), algunos incluso sin apoyarse en una representación gráfica. Sin embargo, tal parece que no tienen claro el significado de algunas etapas en el procedimiento, ni la relación que guardan varios de los elementos involucrados.

La ausencia de significado fue evidente en el problema 1, en el que la mayoría interpretó a *p* como el valor estandarizado que se solicitaba, lo que es un indicio de la complejidad que existe para entender la estandarización (Batanero et al., 2004). Al parecer, la mecanización del procedimiento fue la causa de esta confusión y, posiblemente, también sea una de las causas de los bajos resultados que se obtienen en la evaluación de la cual se tomaron los problemas (EDA), ya que la falta de significado puede hacer que dicho procedimiento se olvide con facilidad.

Por otro lado, el desempeño de los estudiantes que participaron en el presente estudio fue mejor que el de los profesores que participaron en el estudio de Bansilal (2014). A pesar de esta diferencia, el rango de respuestas que se presentaron en ambos estudios es similar, así como las dificultades detectadas, siendo la principal el reconocer que existe una relación entre la probabilidad y el área bajo la curva normal, y una vez reconocida, establecer la relación adecuada entre estos dos elementos. Dificultad también detectada en Batanero et al. (2004).

Llama la atención que en la mayoría de las respuestas del grupo A no se utilizó una representación gráfica como apoyo; sólo en una respuesta al problema 1, en dos en el problema 2, y en cinco en el problema 3. Caso contrario en el grupo B, donde la frecuencia de respuestas en las que se usó una gráfica fue de 27, 33 y 32, respectivamente. Esta situación posiblemente contribuyó al bajo rendimiento del grupo A, ya que los valores numéricos son menos intuitivos que las representaciones gráficas y requieren de un alto nivel de abstracción (Batanero et al., 2004). Asimismo, el uso de una sola representación para referirse tanto a la curva empírica como a la teórica pone en duda si los estudiantes reconocen la estandarización como una transformación que permite pasar de casos particulares a uno general.

Entre las respuestas, pocos estudiantes consideraron la regla empírica para resolver los problemas. En el grupo A, la mayoría de quienes la utilizaron lo hicieron de forma adecuada, pero sin llegar a la solución, otros la aplicaron de forma incorrecta, y sólo uno la utilizó de forma precisa en un problema. En el grupo B, a pesar de que también se vio esta regla, no se consideró en ninguna solución. De acuerdo con Batanero et al. (2004), esto pudo deberse a que la aplicación de esta propiedad requiere de una alta complejidad semiótica.

Respecto a los problemas utilizados, estos no permitieron profundizar en qué es lo que los estudiantes entienden por estandarización, para qué sirve y bajo qué condiciones es factible utilizarla. Sobre este último punto, sería interesante observar si son capaces de apoyarse en el hecho de que la distribución normal es una buena aproximación para otras distribuciones, como la binomial, para resolver problemas. Bajo este escenario, un siguiente paso es generar un instrumento en el que se consideren estos aspectos, el cual permita describir y predecir el razonamiento probabilístico de los estudiantes de bachillerato sobre la distribución normal.

Finalmente, las respuestas son el reflejo del tipo de enseñanza que recibió cada grupo. La enseñanza del grupo B se caracterizó por apoyarse principalmente en el trabajo con lápiz y papel. De las pocas ocasiones en las que se trató de utilizar otro recurso, como el software Fathom, fue al tratar el tema de la regla empírica. Sin embargo, la falta de experiencia de la profesora para utilizar el software como un recurso para la enseñanza, así como la poca familiaridad de los alumnos con el mismo, generó mucha confusión en ellos (Salinas et al., 2018). De igual forma, el bajo rendimiento del grupo A, cuya enseñanza estuvo apoyada en el software Fathom, pudo deberse a que no hubo suficiente tiempo para trabajar problemas escritos sobre la estandarización (ibíd.). De esta manera, resulta importante que durante el diseño de las lecciones haya un equilibrio entre las actividades que se desarrollarán por escrito y las que se llevarán a cabo mediante el recurso computacional, además de que exista una integración adecuada entre ambas.

Referencias

Bansilal, S. (2014). Using an APOS framework to understand teachers' responses to questions on the normal distribution. *Statistics Education Research Journal*, *13*(2), 42-57.

Batanero, C., Tauber, L. y Sánchez, B. (2001). Significado y comprensión de la distribución normal en un curso de análisis de datos. *Quadrante*, *10*(1), 59-92

Batanero, C., Tauber, L. y Sánchez, V. (2004). Students' reasoning about the normal distribution. En D. Ben-Zvi and J. Garfield (Eds.), *The challenge of developing statistical literacy, reasoning and thinking*, (pp. 257–276). Dordrecht: Kluwer.

Biggs, J. y Collis, K. (1982). *Evaluating the quality of learning: The SOLO Taxonomy.* New York: Academic press.

Biggs, J. y Collis, K. (1991). Multimodal learning and the quality of intelligent behavior. En H. Rowe (ed.). *Intelligence: Reconceptualization and Measurement* (pp. 57-76). Hillsdale, NJ: Lawrence Erlbaum.

Carpio, M., Gaita, C., Wilhelmi, M. y Sáenz de Cabezón, A. (2009). Significados de la distribución normal en la universidad. En M. J. González; M. T. González y J. Murillo (Eds.), Trabajo presentado en los grupos de trabajo del *XIII Simposio de la SEIEM.* Santander: SEIEM. Disponible en: http://www.seiem.es/docs/comunicaciones/GruposXIII/depc/Carpio_Gaita_Wilhelmi_Saenz_R.pdf.

Pegg, J. (2002). Assessment in mathematics. A developmental approach. En J. Royer (Ed.), *Mathematical cognition* (pp. 227 – 259). Greenwich, CT: Information Age.

Reading, C. y Canada, D. (2011). Teachers' knowledge of distribution. En C. Batanero, G. Burrill, y C. Reading (Eds.), *Teaching statistics in school mathematics: challenges for teaching and teacher education* (pp. 223–234). Dordrecht: Springer.

Salinas, J., Valdez, J. y Salinas-Hernández, U. (2018). Un acercamiento a la metodología lesson study para la enseñanza de la distribución normal. En L. J. Rodríguez-Muñiz, L. Muñiz-Rodríguez, A. Aguilar González, P. Alonso, F. J. García García y A. Bruno (Eds.), *Investigación en Educación Matemática XXII* (pp. 525- 534). Gijón: SEIEM.

Wild, C. (2006). The concept of distribution. *Statistics Education Research Journal*, *5*(2), 10-26.