e-ISSN: 2595-0967

Razonamiento de estudiantes de bachillerato ante una situación binomial

High school students' reasoning when facing a binomial distribution

Raciocínio de estudantes do ensino médio em situação binomial

Nuria Begué
Universidad de Zaragoza , Didáctica de la Matemática
Zaragoza, España
e-mail nbegue@unizar.es
ORCID: 0000-0003-1369-8711

Carmen Batanero
Universidad de Granada , Didáctica de la Matemática
Granada, España
e-mail: batanero@ugr.es
ORCID: 0000-0002-4189-7139

María M. Gea
Universidad de Granada , Didáctica de la Matemática
Granada, España
e-mail: : mmgea@ugr.es
ORCID: 0000-0002-0970-9775

Danilo Díaz-Levicoy
Universidad Católica del Maule (UCM), Educación
Talca, Chile
e-mail: dddiaz01@hotmail.com
ORCID: 0000-0001-8371-7899

Enviado:12/01/2020

Aceito:18/06/2020

DOI: 10.30612/tangram.v3i2.10888

Resumen: El objetivo del trabajo fue evaluar la comprensión intuitiva de la distribución binomial de estudiantes de Bachillerato. Se analizan las respuestas de 127 estudiantes españoles a una tarea en

Tangram
Revista de Educação Matemática
e-ISSN: 2595-0967

que deben dar cuatro valores probables de la distribución binomial y justificar los valores dados. Se analizan la media y el rango de los cuatro valores en sus respuestas y los argumentos se clasifican siguiendo una metodología de análisis de contenido. La mayoría de los estudiantes proporciona valores cuya media y rango se encuentran en el intervalo aceptable, mostrando un razonamiento distribucional. Entre los argumentos correctos, destacamos los basados en la estimación correcta de la probabilidad, mediante una asignación frecuencial (56,7% de los estudiantes) y la observación de la convergencia y variabilidad de la variable (40,2%). Encontramos argumentos incorrectos como el sesgo de equiprobabilidad o creencias erróneas sobre la aleatoriedad. Alrededor del 22% no son capaces de justificar la respuesta.

Palabras clave: Distribución binomial. Razonamiento. Estudiantes de Bachillerato.

Abstract: The aim of this study we to analyze the students' intuitive understanding of the binomial distribution. We analyze the responses by 127 students to a task where students should provide four values of the binomial distribution and justify the values provided. We analyze the mean, rank of the four values in the response, and classify the arguments using content analysis. Most students provided values with mean and rank in the acceptable interval suggesting a distributional reasoning. Among the correct arguments, the correct estimation of probability stood out (56.7% of students), as well as the observation of convergence and variability of the distribution (40.2%) and some students provide more than an argument. We also found biases, such as the equiprobability bias or erroneous beliefs about randomness. Finally, 22% of students were unable to provide a justification.

Keywords: Binomial distribution. Reasoning. High school students.

Resumo: O objetivo do trabalho foi avaliar a compreensão intuitiva da distribuição binomial de estudantes do ensino médio. São analisadas as respostas de 127 estudantes de espanhol a uma tarefa na qual eles devem fornecer quatro valores prováveis da distribuição binomial e justificar os valores fornecidos. A média e a faixa dos quatro valores em suas respostas são analisadas e os argumentos são classificados seguindo uma metodologia de análise de conteúdo. A maioria dos estudantes fornece valores cuja média e faixa estão na faixa aceitável, mostrando o raciocínio distributivo. Entre os argumentos corretos, destacamos aqueles baseados na estimativa correta de probabilidade, por meio de uma atribuição de frequências (56,7% dos estudantes) e a observação da convergência e variabilidade da variável (40,2%). Encontramos argumentos incorretos, como viés de equiprobabilidade ou crenças errôneas sobre a aleatoriedade. Cerca de 22% não conseguem justificar a resposta.

Palavras-chave: Distribuição binomial. Raciocínio. Alunos do ensino médio.

Introducción

En la actualidad, nos encontramos con nuevos contenidos de probabilidad incluidos en el currículo de Bachillerato, debido a la importancia que el razonamiento probabilístico alcanza en la toma de decisiones y en la predicción en gestión, política o economía. La cultura probabilística es necesaria no solo en la vida profesional sino en la vida diaria, en situaciones como las votaciones, la inversión o la toma de un seguro, así como para comprender el muestreo y la inferencia (Gal, 2005). En consecuencia, la sociedad de la información exige que el ciudadano adquiera una adecuada comprensión de la probabilidad con la finalidad de superar la tendencia al pensamiento determinista y reconocer la presencia del azar, adquiriendo las formas de razonamiento necesarias para resolver las situaciones aleatorias (Batanero, Chernoff, Engel, Lee y Sánchez, 2016).

Por otro lado, cuando trabajamos con probabilidad, realizamos un intenso trabajo de modelización de la realidad, por lo que la enseñanza de la probabilidad, también es enseñanza de la modelización (Chaput, Henry y Girard, 2011). En este sentido, en probabilidad contamos con modelos diferentes que se aplican para estructurar la realidad y poder estudiarla.

En este trabajo nos interesamos por la distribución binomial, que es uno de los modelos probabilísticos de mayor aplicación en múltiples situaciones de la vida diaria y profesional (Landín y Sánchez, 2010). En dicho modelo nos interesamos por un experimento aleatorio donde únicamente consideramos dos posibles sucesos: A, de probabilidad p y su contrario, de probabilidad $q=1-p$. Generalmente se considera que se ha producido un éxito, si al realizar el experimento obtenemos el suceso A, y que se ha producido un fracaso en caso contrario.

El interés de la distribución binomial es analizar la variable aleatoria *número de éxitos en n repeticiones independientes del experimento*. Por ejemplo, podemos considerar el número de recién nacidos niñas en 100 recién nacidos, o el número de personas vacunadas en un grupo de 1000, u otra situación semejante. Dicha variable aleatoria toma valores entre 0 y n, siendo n el número de repeticiones del experimento, se suele denominar como $B(n, p)$, donde n y p son los parámetros de la distribución.

En España, el estudio de la distribución binomial se lleva a cabo en el primer curso de Bachillerato, que es la etapa educativa dirigida a estudiantes de entre 17 y 18 años de edad que desean ingresar en la universidad y consta de dos cursos. Además, este contenido está presente tanto en las modalidades de Ciencias como en la de Ciencias Sociales. Más concretamente, para los alumnos de 1º curso de Ciencias Sociales y en el 2º curso de Ciencias se incluyen los contenidos, criterios de evaluación y estándares de aprendizaje evaluables presentados en la Tabla 1 (MECD, 2015, p. 385 y 422).

Tabla 1: Orientaciones sobre la distribución binomial en Bachillerato (MECD, 2015, p. 385 y 422)

Contenidos	**Criterios de evaluación**	**Estándares de aprendizaje**
Distribución binomial. Caracterización e identificación del modelo. Cálculo de probabilidades	Identificar los fenómenos que pueden modelizarse mediante las distribuciones de probabilidad binomial y normal calculando sus parámetros y determinando la probabilidad de diferentes sucesos asociados	Identifica fenómenos que pueden modelizarse mediante la distribución binomial, obtiene sus parámetros y calcula su media y desviación típica. Calcula probabilidades asociadas a una distribución binomial a partir de su función de probabilidad, de la tabla de la distribución o mediante calculadora, hoja de cálculo u otra herramienta tecnológica y las aplica en diversas situaciones. Calcula probabilidades de sucesos asociados a fenómenos que pueden modelizarse mediante la distribución binomial a partir de su aproximación por la normal valorando si se dan las condiciones necesarias para que sea válida.

El objetivo de este trabajo es analizar la comprensión intuitiva de una muestra de estudiantes de 2º curso de Bachillerato de una situación, para ellos familiar, que puede modelizarse mediante una distribución binomial. Para ello, se les plantea una tarea en la que se pide proporcionar cuatro valores probables de la distribución binomial que, además, deben argumentarse. A partir de las respuestas se evalúa su comprensión del valor esperado y la variabilidad en dicha distribución.

En lo que sigue, se describen los antecedentes de nuestro trabajo, su metodología y sus resultados.

Investigación previa

En nuestro trabajo tendremos en cuenta otros previos que se interesan por la comprensión que tienen los estudiantes de la distribución binomial. Uno de dichos trabajos es el de Van Dooren, De Bock, Depaepe, Janssens y Verschaffel (2003), que analizan las soluciones de 225 estudiantes de Bachillerato de dos cursos diferentes. Plantean siete ítems de opción múltiple de situaciones binomiales en el contexto de lanzar un dado, variando los valores n, p y x (número de éxitos). En unos ítems donde pide comparar probabilidades obtiene porcentajes de éxito en torno al 80%, mientras que cuando pide calcular probabilidades o ver si un cálculo es correcto el porcentaje desciende entre el 16% y 30%, según la tarea. Muchos de los estudiantes de su muestra resuelven los problemas simplemente utilizando proporciones, lo que los autores denominan *ilusión de linealidad*, por tratarse de una generalización incorrecta de este concepto. Los autores atribuyen el resultado a la potencia del modelo lineal para analizar muchas situaciones cotidianas y la familiaridad del estudiante con el mismo, lo que le lleva a no diferenciar las situaciones en que puede y no puede aplicarse.

Landín y Sánchez (2010), por su parte, estudian el proceso mediante el cual 66 estudiantes de bachillerato en México llegan a comprender esta distribución e investigan sus sesgos cognitivos. Identifican los siguientes componentes de razonamiento intuitivo en la distribución binomial: 1) reconocimiento de las situaciones en que se aplica y representación de las secuencias de éxitos y fracasos; 2) reconocimiento de la variable aleatoria; 3) reconocimiento del carácter combinatorio y necesidad de contar las combinaciones; 4) uso de la definición clásica de probabilidad y regla del producto; y 5) relación entre las combinaciones y la probabilidad de un valor de la variable.

Los autores utilizan dichos componentes para iniciar la construcción de una jerarquía de razonamiento intuitivo en torno a la definición que hacen los estudiantes ante una serie de problemas de esta distribución, de la definición clásica de probabilidad, regla del producto, combinaciones y fórmula de la distribución binomial. Clasifican las soluciones de un grupo de 66 estudiantes en dicha jerarquía concluyendo que la mayoría se queda en el primer nivel de razonamiento o en el segundo, siendo pocos los que alcanzan los niveles superiores.

El trabajo anterior se continúa en Sánchez y Landín (2011), quienes reanalizan las respuestas de 66 estudiantes a una serie de problemas binomiales y en base a las soluciones de los estudiantes, los autores, con ayuda de 10 expertos, finalizan la construcción de la jerarquía de razonamiento en torno a la distribución binomial, con los siguientes niveles:

1. *Nivel subjetivo*. El estudiante presenta sesgos de razonamiento cuando trata de calcular probabilidades binomiales. Uno de los más frecuentes es el sesgo de equiprobabilidad (Lecoutre, 1992), donde los estudiantes consideran equiprobables los resultados de cualquier experimento aleatorio. Por tanto, en la distribución binomial considerarían equiprobables todos sus valores posibles.
2. *Nivel transicional*. El estudiante es capaz de listar los valores de la variable, aunque no necesariamente de manera exhaustiva. Reconoce que el número *n* de ensayos influye en la distribución, pero al calcular probabilidades asociadas, utiliza simplemente la regla de Laplace o se basa en sus creencias subjetivas.

3. *Nivel cuantitativo informal.* Se reconoce el carácter combinatorio de la situación, es decir, la importancia del orden de los resultados. Se calculan probabilidades mediante diagrama en árbol o utilizando regla del producto.
4. *Nivel numérico.* Se reconoce la variable aleatoria, sus valores posibles y sus parámetros. Se utiliza la fórmula binomial para calcular las probabilidades al resolver los problemas.

Mayén y Salazar (2014) comparan el razonamiento de 77 estudiantes de Bachillerato, con y sin instrucción sobre la distribución binomial, utilizando la jerarquía de razonamiento propuesta por Sánchez y Landín (2011). Para ello, aplican un cuestionario con 10 problemas. Los estudiantes sin instrucción se mantienen casi exclusivamente en los primeros niveles, progresando a niveles más avanzados para aplicar las combinaciones y regla del producto al finalizar la enseñanza, llegando en algunos casos al nivel superior de razonamiento. El progreso se logra cuando se comprenden simultáneamente las combinaciones y la regla del producto. Igualmente, el uso del diagrama en árbol permite organizar la solución del problema.

Los trabajos anteriores se interesan por la competencia de los estudiantes de calcular problemas binomiales. Una aproximación diferente es la de Shaughnessy, Ciancetta y Canada (2004), quienes utilizan distribuciones binomiales con 272 estudiantes (10-19 años), pidiéndoles dar el número de sucesos de un cierto tipo en 10 y 100 repeticiones de un experimento. La mayoría de los estudiantes sobreestimaron la variabilidad de la distribución, independientemente del número de experimentos, mientras que algunos le otorgaron muy poca variabilidad. Los autores identifican tres niveles progresivos en el razonamiento de los estudiantes: 1) el *nivel de razonamiento aditivo* (el más frecuente), que consiste en considerar diferentes muestras de valores de la variable como subconjuntos disjuntos; 2) el *nivel de razonamiento proporcional*, en el que se utilizan proporciones al realizar estimaciones y se comprende el valor esperado de la distribución; y 3) el *nivel de* *razonamiento distribucional* (el menos frecuente) donde se integran las ideas de valor esperado y de variabilidad.

Un trabajo relacionado es el de Gómez, Batanero y Contreras (2014), en una investigación con futuros profesores de Educación Primaria. Los autores se interesan por evaluar la comprensión intuitiva del valor esperado y la variabilidad en una distribución binomial. Para ello proponen un ítem en que les piden cuatro valores de una distribución binomial $B(100, 0.68)$ analizando la media y el rango de los cuatro valores dados y comparándolo con la distribución de la media y el rango que teóricamente se esperaría. Los resultados muestran que solamente una tercera parte de los sujetos de su estudio tuvieron una intuición simultánea de la media y la variabilidad de la distribución. Además, los autores identifican en los sujetos diferentes sesgos como la equiprobabilidad (Lecoutre, 1992).

En Begué, Batanero y Gea (2019) se propone el ítem trabajado por Gómez et al. (2014) a una muestra de estudiantes de Bachillerato, pidiéndoles, además, que justificasen los valores proporcionados. Las respuestas obtenidas se clasificaron en diferentes tipos de argumentos que ayudaron a detectar la presencia de algunos sesgos de razonamiento en los estudiantes, tales como el sesgo de equiprobabilidad y concepciones erróneas sobre el muestreo. El análisis de las respuestas se fundamenta en los elementos del enfoque ontosemiótico (Godino, 2002), en concreto, en la idea de conflicto semiótico. En esta teoría, se asume que los objetos matemáticos surgen de las prácticas realizadas en la resolución de problemas, que son las que dotan de sentido a los objetos matemáticos (Godino, 2017). Debido a la variedad de objetos que intervienen en dichas prácticas (lenguaje, conceptos, propiedades, procedimientos, argumentos) y al carácter inmaterial de muchos objetos, el trabajo en matemáticas requiere muchos procesos interpretativos. En ocasiones el estudiante interpreta incorrectamente el sentido de una expresión matemática o puede conceder un significado diferente al pretendido por el profesor a uno de los objetos que utiliza. En dicho caso, se habla de conflicto semiótico.

En este trabajo, continuamos el de Begué et al. (2019), que, además de describir los argumentos de los estudiantes en una situación binomial *B*(0,68, 100) identificaron en ellos una serie de conflictos semióticos- En concreto se analizar ahora las respuestas de la misma muestra de estudiantes a una situación binomial *B*(10, 0.7), es decir, donde interviene un pequeño tamaño de muestra. Además del análisis cualitativo de las respuestas estudiaremos también la intuición del valor esperado y de la variabilidad de la variable, siguiendo el método de Gómez et al. (2014).

En lo que sigue se describe la metodología seguida para el análisis de los datos, se presentan y discuten como la discusión de los resultados.

Metodología

La muestra estuvo formada por 127 estudiantes de 2º Bachillerato (17-18 años) que cursaban sus estudios en dos centros escolares diferentes de la comunidad de Aragón (España). En el primero de los centros participaron dos grupos de estudiantes de Ciencias (26 y 15 alumnos por grupo) y un grupo de Ciencias Sociales (19 estudiantes). En el segundo centro se tomaron otros dos grupos de Ciencias (29 y 14 alumnos) y uno de Ciencias Sociales (24 alumnos). A todos estos estudiantes se les propuso la siguiente tarea, que describe un contexto familiar para los estudiantes. La tarea (Figura 1) se completó dentro de la clase de matemática, con colaboración del profesor de los estudiantes y autorización de los directores de los centros.

Tarea. Un jugador de baloncesto suele encestar 70 de cada 100 tiros a una canasta desde la posición de tiros libres. Escribe en la siguiente tabla un resultado que sea probable para cuatro partidos en los que lanza 10 tiros desde el punto de lanzamientos personales.

Partido 1 (10 tiros)	Partido 2 (10 tiros)	Partido 3 (10 tiros)	Partido 4 (10 tiros)
Número de encestes:	Número de encestes	Número de encestes	Número de encestes
Número de fallos:	Número de fallos:	Número de fallos:	Número de fallos:

Figura 1 - Tarea planteada a los estudiantes (Adaptado de Gómez et al. (2014)).

La tarea describe la frecuencia de encestes de un jugador de baloncesto, habiéndose elegido este contexto por ser conocido por ser familiar para el estudiante. La variable aleatoria número de encestes sigue una distribución binomial $B(10, p)$, donde p es desconocida. Si consideramos el dato del enunciado, se estima que el valor esperado es np=7, y su desviación estándar es $\sqrt{np(1-p)}$ =1,45.

Se realizó un doble análisis cuantitativo y cualitativo de las respuestas. En el análisis cuantitativo se calculó el valor medio y el rango de los cuatro valores proporcionados por los estudiantes en sus respuestas. En relación al valor esperado, se asume que el sujeto presenta una buena concepción o intuición del mismo, si el valor medio de las cuatro respuestas que proporciona es cercano a 7. Se considerarán cercanas aquellas respuestas cuyos valores para el número de encestes se localicen dentro del intervalo (5.55, 8.45), redondeando, entre 6 y 9. Los extremos de dicho intervalo se obtienen al determinar la media más/menos la desviación típica, que es el intervalo que contiene el 68% de las observaciones en la distribución normal. Se toma como aceptable el intervalo obtenido al sumar y restar dos veces la desviación típica a la media, es decir, aquellos valores que se localizan entre 4.1 y 9.9, redondeando, entre 4 y 10. Finalmente, otro aspecto que se debe reconocer es la ausencia de equiprobabilidad por la asimetría física del dispositivo.

En cuanto al estudio de la variabilidad, Gómez el al. (2014) consideran que la la variabilidad es adecuada cuando el rango se encuentra entre una y dos desviaciones típicas, aproximadamente entre 3 y 6. Si está entre dos y tres veces la desviación típica (entre 6 y 9 redondeando) se considera alta, pero aceptable y si es mayor excesiva. Si es menor que 3 se considera demasiada concentración y quiere decir que no se comprende la variabilidad muestral.

Para el análisis cualitativo, las argumentaciones escritas de los estudiantes se sometieron a análisis de contenido (Krippendorf, 2013), una técnica de investigación que permite realizar inferencias sobre el contenido de un texto. En nuestro caso, las unidades de análisis son la respuesta de cada estudiante. Seguidamente, se transcribieron a un fichero

Word las respuestas de todos los estudiantes con un código para identificarlas. Para formar las categorías se compararon todas las respuestas de cada ítem entre sí, agrupando las semejantes para formar una primera lista, que fue revisada por los investigadores. Mediante la discusión de casos discordantes y después de sucesivas revisiones se llegó a la lista definitiva de categorías. Finalmente se analizaron los datos obtenidos.

Resultados y discusión

Análisis cuantitativo

En primer lugar, se analiza la distribución del valor medio de las cuatro respuestas dadas por cada estudiante, que se presenta gráficamente en la Figura 2, donde se ha representado mediante una línea roja vertical el valor medio teórico de la distribución binomial $B(10, 0.7)$. Podemos observar una buena intuición de este valor medio en los estudiantes de la muestra, pues dicho valor teórico, que además tiene una frecuencia apreciable, coincide con la moda de la distribución. Además, el valor medio empírico de la distribución de las medias coincide casi con el teórico.

El análisis anterior se completa con la Tabla 2, donde clasificamos las medias de las cuatro respuestas de los estudiantes en diferentes intervalos. Observamos que una amplia mayoría proporciona valores cuya media está en el intervalo de estimación óptimo, y otro grupo realiza una estimación aceptable, por lo que los estudiantes muestran una buena intuición del valor esperado en la distribución binomial. Aun así, encontramos estudiantes con valores medios muy bajos, la mayor parte de los cuales ha mostrado el sesgo de equiprobabilidad (Lecoutre, 1992), al dar valores muy cercanos al 5 (considerando equiprobables los dos resultados) o bien por encima y debajo de cinco de modo que la media es aproximadamente 5. También encontramos casos aislados de estimaciones muy alejadas del promedio, como un alumno que da todos los valores igual a 1 y otro, que obtiene un promedio igual a 9 en sus cuatro valores.

Tangram
Revista de Educação Matemática
e-ISSN: 2595-0967

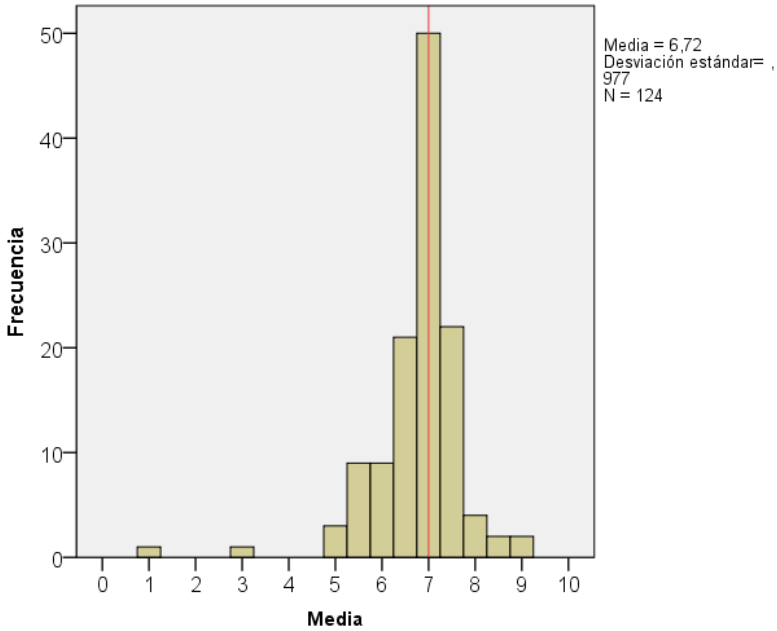

Figura 2 - Distribución de la media de las cuatro respuestas de los estudiantes

Tabla 2: Porcentaje de estudiantes según intervalo en que se sitúa la media de los cuatro valores en su respuesta

Estimación	Intervalo	Porcentaje
Estimación muy baja	< 5.6	11,1
Estimación óptima	6.3–7.7	73
Estimación aceptable (fuera del intervalo 6,3-7,7)	5.6–8.4	11
Estimación muy alta	> 8.4	2,5
No responden		2,4

Seguidamente se presenta en la Figura 3 la distribución del rango las cuatro respuestas dadas por cada estudiante, donde se ha representado, mediante una línea roja vertical, el valor medio teórico de la desviación típica de la distribución binomial $B(10, 0.7)$, que como vemos es también muy próxima a la teórica. La información se completa con la

Tabla 3, en la que se presentan los porcentajes de estudiantes con el rango de las cuatro respuestas en diferentes intervalos.

Se deduce del análisis de la Tabla 3 una amplia mayoría de estudiantes proporciona muestras de variabilidad óptima. Además, si consideramos el porcentaje de los que conceden una variabilidad aceptable se obtiene un porcentaje elevado, el 88,2%. No obstante, se identifican alumnos con alta variabilidad o alta concentración, destacando cuatro alumnos que repiten el valor 7 cuatro veces en su respuesta. En este caso, se concluye que dichos alumnos tendrían una concepción determinista de la situación.

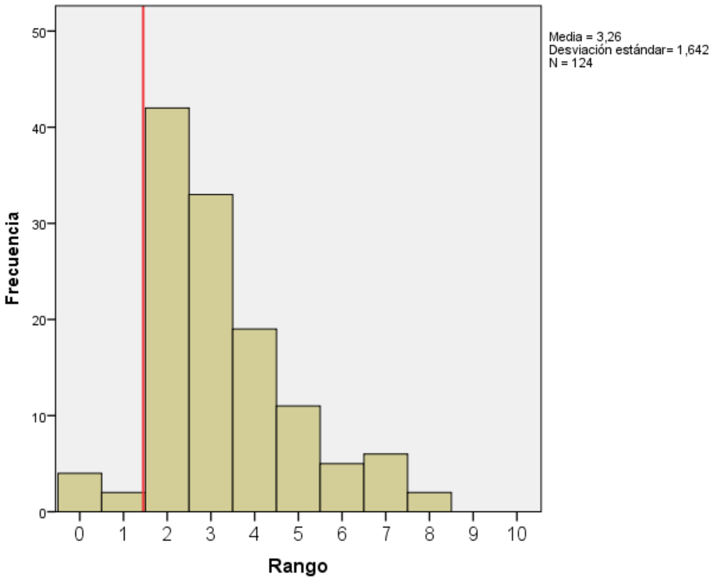

Figura 3 - Distribución del rango las respuestas de los estudiantes

Tabla 3: Porcentaje de alumnos según intervalo en que se sitúa el rango de los valores dados

Variabilidad	Intervalo	Porcentaje
Alta concentración	0	3,1
Estimación óptima	2–4	74
Estimación aceptable	1–6	14,2
Alta variabilidade	7–10	6, 3
No completa		2,4

Análisis cualitativo

Para completar el estudio cuantitativo, a continuación, se caracterizan los distintos tipos de justificaciones encontrados, incluyendo ejemplos de cada una de ellas.

Tipo 1. Aleatoriedad. En estas respuestas, se puede observar que bastantes estudiantes utilizan la expresión *hacer a boleo* u otras similares, como *porque se me ha venido a la cabeza*, para justificar que han escrito al azar los datos que proporcionan. También se ha incluido dentro de esta categoría a los estudiantes que indican que el resultado es imprevisible por ser aleatorio, por lo que aplicarían una concepción de la aleatoriedad como sinónimo de un proceso que es impredecible, concepción descrita por Fischbein, Sainati y Sciolis (1991) y Serrano (1996), en su estudio de la comprensión de la aleatoriedad por los estudiantes.

Esta categoría también se encuentra en aquellas respuestas en las que el estudiante asocia la aleatoriedad con la imposibilidad de medir la probabilidad. Según Serrano (1996), esta concepción de la aleatoriedad denotaría en los estudiantes el enfoque en el resultado, descrito por Konold (1989). Es decir, sujetos que interpretan una pregunta de probabilidad (¿cuál es el resultado más probable?) en forma no probabilística, sino determinista (¿qué resultado ocurrirá?), donde se le pide dar una predicción segura del resultado. Este argumento aparece en todos los ítems. Algunos ejemplos de respuestas en esta categoría son los siguientes:

> A522: Las respuestas son aleatorias, porque un jugador puede ser malo o bueno. (8, 5, 1, 7)
> A495: En este último también existe el azar pero también hay factor de habilidad. (7, 8, 6, 0)

Observamos que estos estudiantes proporcionan valores con un recorrido muy amplio, unos por exceso y otros por defecto, por lo que podría estar razonando de acuerdo a la representatividad; más específicamente de acuerdo a la ley de los pequeños números (Kahneman, Slovic y Tversky, 1982), al tratar de compensar sus resultados.

Tipo 2. Razonamiento basado en aspectos físicos del experimento aleatorio. Esta categoría recoge las justificaciones en las que se describe algún rasgo físico del jugador para argumentar que sea más frecuente la obtención de un resultado. Subyace una interpretación de la probabilidad próxima al significado de propensión debido a Peirce (Batanero y Díaz, 2007) en la que la probabilidad de un generador aleatorio es una tendencia de producir cada uno de sus posibles resultados. Dicha interpretación constituye una alternativa a la interpretación frecuencial, según Rolleri (2002), pues es aplicable a sucesos físicos singulares. La propensión sería una propiedad del sistema físico en ciertas situaciones experimentales.

En esta concepción de la probabilidad, un generador aleatorio, por ejemplo, un dado, tendría una tendencia de producir cada uno de sus posibles resultados. Esta tendencia estaría relacionada tanto con la frecuencia relativa a la larga como con cada resultado singular. Sería una tendencia causal del comportamiento del sistema aleatorio de una cierta manera; en este ítem, sería una propiedad del jugador una propensión a actuar de una cierta manera. Batanero y Díaz (2007) indican que la propensión es diferente de la probabilidad; la propensión sería fuerte si la probabilidad del suceso es alta, pero si es baja (por ejemplo, en un dado una probabilidad 1/6 de obtener un 5) la propensión es fuerte, porque esta probabilidad se producirá con mucha seguridad a la larga. Algunos ejemplos de esta respuesta siguen a continuación:

> A524: Depende de lo bueno o malo que sea un jugador (8, 7, 6, 9)
> A482: El porcentaje de aciertos en 10 tiros libres puede variar debido a otros factores como si tiene un buen día o no pero generalmente en menos tiros menos cansancio y subirá su porcentaje. (7, 9, 8, 9)

Tipo 3. Razonamiento basado en asignación frecuencial de probabilidad. Se ha clasificado en esta categoría a los estudiantes que apoyan su argumento en los datos proporcionados en el enunciado del ítem sobre el experimento. Por tanto, su respuesta puede indicar una comprensión intuitiva de la probabilidad frecuencial, puesto que asignan la probabilidad del suceso de interés, en relación a los datos obtenidos en la muestra dada en el enunciado. Además, asumen explícitamente que la probabilidad de dicho suceso ha de estar cerca del valor dado en el enunciado, por lo que las muestras que genera tienen que ser representativas de dicha probabilidad. Todo ello implica que el estudiante comprende y aplica intuitivamente la ley de los grandes números. Añadimos a continuación algunas respuestas de este tipo donde los estudiantes calculan la probabilidad frecuencial de acierto dada por la frecuencia del enunciado y la relacionan con el número esperado de éxitos en 10 lanzamientos.

> A412: Si encesta el 70 de 100 tiros, equivale a un 70%. Si lanza 10 tiros, entonces el porcentaje son 7,2 tiros encesta de cada 10. Si en total realiza 40 tiros, debería encestar 28. (7, 6, 8, 7)
> A429: Si de cada 100 marca 70, sus tiros tendrán una probabilidad del 70%, 7/10, por lo que los aciertos no pueden variar mucho (7, 6, 8, 9).

Tipo 4. Sesgo de equiprobabilidad implícito. En algunas respuestas se observa que el estudiante asigna la probabilidad del suceso aplicando la regla de Laplace. Por tanto, se asume que el fenómeno aleatorio cumple la condición de equiprobabilidad, por lo que todos los sucesos elementales son equiprobables. Esta respuesta en fenómenos aleatorios que no cumplen la condición de equiprobabilidad, conlleva la presencia del sesgo de equiprobabilidad (Lecoutre, 1992). Entonces, el estudiante resuelve incorrectamente la tarea. En el siguiente ejemplo, el estudiante A430 considera la equiprobabilidad, indicando que el número de encestes de los cuatro valores será cercano a 5 y, en contra de su argumento, los resultados que proporcionan se acercan a la probabilidad dada en el ítem.

> A430: Todo a boleo, no sé cuánto será. Podría salir cualquier cosa, pero es más probable que los números estén cerca de 5. (7, 8, 6, 7).

Tipo 5. Convergencia y variabilidad del muestreo. Son aquellas argumentaciones en las que se justifica que los resultados deben ser parecidos a los dados en el enunciado, lo que supone que el estudiante percibe la relación entre el valor esperado de la variable y el valor

Tangram Revista de Educação Matemática
e-ISSN: 2595-0967

dado en su muestra, y, además, el estudiante también indica que se debe esperar una cierta variabilidad. Por tanto, utiliza las propiedades de convergencia de la frecuencia relativa a la probabilidad y de variabilidad. En consecuencia, el estudiante se encuentra en el nivel más alto de razonamiento entre los descritos por Landín y Sánchez (2010) y Mayén y Salazar (2014). Son muy numerosas estas respuestas y son consistentes con las cuaternas que proporcionan, como vemos en algunos ejemplos:

> A520: Porque en 40 tiros tocaría marcar 28 tiros y luego esos 28 se deben repartir entre 4 días que dan a 7 pero he puesto algún caso con más y menos porque no debe ser una media (7, 8, 6, 7).
> A444: En partidos diferentes con características diferentes (puede tener mejor/peor día, estar o no confiado, estar nervioso, …) los resultados no serán siempre 70% de aciertos pero cuántos más partidos juegue, más se acercará a esa proporción de aciertos. (6, 5, 8, 2)

Tipo 6. Creencias subjetivas. La probabilidad es un terreno en que los estudiantes, e incluso algunos sujetos adultos muestran creencias subjetivas para explicar la ocurrencia de ciertos sucesos en fenómenos aleatorios. Así, Cañizares y Batanero (1997) describen la creencia en la suerte o en la posibilidad de controlar la ocurrencia del suceso. Por su parte, Konold (1989) considera el enfoque en el resultado, donde se interpreta una pregunta de probabilidad en forma no probabilística. Estos argumentos exponen criterios que no tienen que ver con los datos del enunciado, ni con la probabilidad, pero se deducen del conocimiento que los estudiantes tienen del contexto. Por tanto, se interpreta la pregunta probabilística en modo no probabilístico, como en los siguientes ejemplos:

> A470: Porque aquí hay una técnica y aparte del azar, influye la técnica y podrá meter más o menos dependiendo del día del jugador. De si está cómodo en el partido, si está motivado. (5, 8, 8, 7)
> A410: Aumentará su número de encestes progresivamente porque tendrá más práctica conforme vaya jugando y entrenando. (6, 7, 8, 9)

Tabla 4: Frecuencia y porcentaje de estudiantes con diferentes argumentos

Argumentos	Frecuencia	Porcentaje

Tipo 1. Aleatoriedad	9	7,1
Tipo 2. Aspectos físicos del experimento	1	0,8
Tipo 3. Asignación frecuencial de la probabilidad	72	56,7
Tipo 4. Sesgo de equiprobabilidad implícito	2	1,6
Tipo 5. Variabilidad y convergencia	51	40,2
Tipo 6. Creencias subjetivas	22	17,3
Tipo 7. No justifica	29	22,8

La Tabla 4 presenta la frecuencia y porcentaje con que se dieron los diferentes argumentos en los estudiantes de la muestra. Para interpretarla hay que tener en cuenta que un mismo estudiante a veces da más de un argumento, por lo que el total del porcentaje es mayor de 100. En esta tabla se observa que el 56,7% de los estudiantes apoyan su argumento en los valores proporcionados en el enunciado, es decir, realizan una estimación frecuencial de la probabilidad.

Por otro lado, como el número de tiros es pequeño la variabilidad debe ser grande en la variable. De hecho, el 40,2% de los estudiantes aluden a aspectos relacionados con la representatividad y variabilidad, mostrando un razonamiento distribucional sobre la variable binomial. Como se ha indicado anteriormente, existen argumentos en los que se identifica varias categorías. Destacamos que el 32,3% hacen referencia tanto a la identificación de la probabilidad desde el enfoque frecuencial (Tipo 3) como a los conceptos de representatividad y/o variabilidad (Tipo 5). Es decir, muestran un buen conocimiento de la distribución y sus propiedades.

Finalmente destacamos las respuestas (17,3%) que se justifican con criterios subjetivos sin tener en cuenta criterios probabilísticos, y también el alto porcentaje de estudiantes (22,8%) que no es capaz de justificar su respuesta. Este porcentaje contrasta con el de alumnos que no dio respuesta, el cual fue muy pequeño. Por tanto, los estudiantes en su mayor parte resuelven correctamente la tarea, pero son menos los que la llegan a argumentar, siendo algunas de estas argumentaciones incorrectas. En estos argumentos hemos observado los siguientes conflictos semióticos de los estudiantes:

C1. *Confusión entre suceso aleatorio y suceso equiprobable*. En este caso se asigna a la aleatoriedad una propiedad que no le corresponde, que sería la equiprobabilidad. Por tanto, se cae en el sesgo de equiprobabilidad, descrito por Lecoutre (1992) puesto que hay fenómenos aleatorios, como el descrito en la tarea que constan de sucesos no equiprobables. Los estudiantes que dan este argumento asumen que la probabilidad de cada suceso se corresponde con el 50% y suelen dar valores del número de encestes en 10 lanzamientos por encima y debajo de 7, pero tendiendo a que la media se acerque a 5.

> A488: El jugador la mayoría de los partidos marca más del 50% de los tiros. (7, 4, 8,1).

C2. *Suponer que la aleatoriedad indica mucha variabilidad*, es decir, esperar resultados muy diferentes entre sí. Es otra propiedad mal asignada, pues, aunque la aleatoriedad implica variabilidad, la variabilidad depende del tamaño de la muestra y no del hecho de ser un suceso aleatorio. En este caso, la afirmación de que haya variabilidad estaría justificada por el tamaño de la muestra, como muestra el ejemplo siguiente:

> A504: Al realizar 10 tiros solamente, los resultados no tienen por qué parecerse a la estimación sobre 100 tiros y pueden variar mucho. (2, 8, 7, 5).

C3. *Creencia en la ley de los pequeños números* (Tversky y Kaneman, 1974). La probabilidad nos enseña que al repetir un número veces suficiente y de forma independiente un experimento aleatorio, la frecuencia relativa de cada suceso posible tiende a su probabilidad de ocurrencia. Sin embargo, a veces se espera que esta convergencia se produzca incluso en un número pequeño de experimentos. En este caso, los estudiantes tratan de reproducir el valor esperado de la variable en sus cuatro valores y, además, dichos valores tienen poca variabilidad, como el siguiente ejemplo.

A451: Comparando el número de tiros libres que suele encestar, y reduciéndolo a diez, es probable que no varíe mucho. (7, 8, 7, 6)

C4. *Interpretación determinista del enunciado.* Un caso extremo del anterior es cuando el estudiante se limita a repetir cuatro veces el mismo valor del enunciado, como en el siguiente ejemplo:

A510: Por estadística (7, 7, 7, 7)

C5. *Suponer que aleatoriedad es sinónimo de impredecibilidad*, ni siquiera en una serie de resultados. Aunque cada suceso aislado de un fenómeno aleatorio es impredecible (no se sabe si ocurrirá o no), cuando se tiene información del experimento podemos deducir los resultados más y menos probables, es decir, deducir la distribución del conjunto de resultados. Interpretamos que los estudiantes no identifican la probabilidad como una medida del grado de incertidumbre, sino que aleatoriedad y probabilidad son considerados como términos similares. Este razonamiento fue descrito por Konold (1989) como enfoque en el resultado. En el siguiente ejemplo, el estudiante no se decide por los valores más probables, lo que se refleja también en la variabilidad de los valores que propone:

A481: Podrá acertar 10 como fallar 10 (6, 7, 5, 10)

Conclusiones

Los cuatro valores proporcionados por los estudiantes al responder la tarea, en su mayor parte, corresponden a una estimación óptima o al menos razonable del valor esperado en la distribución binomial. Además, también la mayoría de estos valores tienen una variabilidad óptima o aceptable. Concluimos que la mayoría de los estudiantes de la muestra participante alcanza el razonamiento distribucional descrito por Shaughnessy et al. (2004), al contrario de lo obtenido con los estudiantes que participaron en el estudio llevado por dichos autores.

Esta conclusión se refuerza mediante el porcentaje de argumentos en que se refleja un conocimiento sobre el enfoque frecuencial de la probabilidad, a los que se añaden los que razonan de acuerdo a la variabilidad y convergencia. Otros estudiantes son capaces de identificar la asimetría del fenómeno debido a sus conocimientos de otras materias, como la física, puesto que hacen referencia en su justificación a conceptos tales como la fuerza del lanzamiento, el ángulo de tiro o la distancia y, en consecuencia, a la visión de la probabilidad como propensión (Batanero y Díaz, 2007).

El resto estudiantes, tanto en sus estimaciones como en sus argumentos, muestra una comprensión pobre de la aleatoriedad y la probabilidad, y algunos, aunque abordan la tarea, no son capaces de argumentarla, lo que indica que sus intuiciones no se corresponden con una comprensión profunda del fenómeno.

Para finalizar, sugerimos que las intuiciones correctas de los estudiantes debieran ser tenidas en cuentas y reforzadas a lo largo de la instrucción. Igualmente, el profesor debe estar atento a la presencia de las intuiciones incorrectas porque su diagnóstico favorecerá un posterior tratamiento y mejora de la capacidad de argumentación de los estudiantes. De acuerdo con Batanero et al. (2016), creemos que la adquisición del razonamiento probabilístico correcto debe ser un fin en sí mismo y no únicamente como instrumento del estudio de la inferencia. En consecuencia, resulta necesario que los estudiantes analicen tareas relacionadas con contextos cotidianos como la presentada en este trabajo, y que la discusión de las posibles soluciones correctas e incorrectas se use para mejorar su conocimiento y razonamiento probabilístico.

Agradecimientos

Proyecto PID2019-105601GB-I00 (MICIN) y Grupo FQM126 (Junta de Andalucía).

Referencias

e-ISSN: 2595-0967

Batanero, C., Chernoff, E., Engel, J., Lee, H. y Sánchez, E. (2016). *Research on teaching and learning probability*. *ICME-13. Topical Survey series*. New York: Springer.

Batanero, C. y Díaz, C. (2007). Meaning and understanding of mathematics. The case of probability. En J. P. Van Bendegen y K. François (Eds.), *Philosophical Dimmensions in Mathematics Education* (pp. 107-128). Nueva York: Springer.

Begué, N., Batanero, C. y Gea, M. M. (2019). Argumentos de los estudiantes de bachillerato en la generación de muestras de la distribucion binomial. 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* (pp. 1-10). Granada: Universidad de Granada.

Cañizares, M. J. y Batanero, C. (1997). Influencia del razonamiento proporcional y de las creencias subjetivas en la comparación de probabilidades. *UNO. Revista de didáctica de las matematicas, 14*(1), 99-114.

Chaput, B., Girard, J. C. y Henry, M. (2011). Frequentist approach: modelling and simulation in statistics and probability teaching. En C. Batanero, G. Burrill, C. Reading (Eds.), *Teaching statistics in school mathematics. Challenges for teaching and teacher education* (pp. 85-95). New York: Springe.

Fischbein, E., Sainati, M. y Sciolis, M. (1991). Factors affecting probabilistics judgements in children and adolescents. *Educational Studies in Mathematics, 22*(6), 523-549.

Gal, I. (2005). Towards "probability literacy" for all citizens: Building blocks and instructional dilemmas. En G. A. Jones (Ed.), *Exploring probability in school. Challenges for teaching and learning* (pp. 39-63). Dordrecht: Kluwer.

Godino, J. D. (2002). Un enfoque ontológico y semiótico de la cognición matemática. *Recherches en Didactique des Mathématiques, 22*(2-3), 237-284.

Godino, J. D. (2017). Construyendo un sistema modular e inclusivo de herramientas teóricas para la educación matemática. En J. M. Contreras, P. Arteaga, G. R. Cañadas, M.M. Gea, B. Giacomone y M. M. López-Martín (Eds.), *Actas del Segundo Congreso International Virtual sobre el Enfoque Ontosemiótico del Conocimiento y la Instrucción Matemáticos* (pp. 1-20). Granada: Universidad de Granada.

Gómez, E., Batanero, C. y Contreras, C. (2014). Conocimiento matemático de futuros profesores para la enseñanza de la probabilidad desde el enfoque frecuencial. *BOLEMA. Boletim de Educação Matemática, 28*(48), 209-229.

Kahneman, D., Slovic, P. y Tversky, A. (1982*). Judgment under uncertainty: Heuristics and biases*. New York: Cambridge University Press.

Konold, C. (1989). Informal conceptions of probability. *Cognition and Instruction, 6,* 59-98.

Krippendorff, K. (2013). *Content analysis: an introduction to its methodology*. Londres: Sage.

Landín; P. y Sánchez, E. (2010). Niveles de razonamiento probabilístico de estudiantes de bachillerato frente a tareas de distribución binomial. Educaçao Matemática Pesquisa, 12(3), 598-618.

Lecoutre, M. P. (1992). Cognitive models and problem spaces in "purely random" situations. *Educational Studies in Mathematics, 23*(6), 557-568.

Mayén, S. y Salazar, A. (2014). Niveles de razonamiento en la resolución de tareas de tipo binomial. En L. Andrade (Ed.), *Memorias I Encuentro Colombiano de Educación Estocástica* (pp. 88-97). Bogotá: Asociación Colombiana de Educación Estocástica.

MECD. (2015). *Real Decreto 1105/2014, de 26 de diciembre, por el que se establece el currículo básico de la Educación Secundaria Obligatoria y del Bachillerato*. Madrid: Ministerio de Educación, Cultura y Deporte.

Rolleri, J. L. (2002). La probabilidad como grado de posibilidad. *Crítica: Revista Hispanoamericana de Filosofía, 34*(101), 3-26.

Sánchez, E. y Landín, P. (2011). Fiabilidad de una jerarquía para evaluar el razonamiento probabilístico acerca de la distribución binomial. En M. Marín, G. Fernández, L. Blanco y M. Palarea (Eds.), *Investigación en Educación Matemática XV* (pp. 533-542). Ciudad Real: SEIEM.

Serrano, L. (1996). *Significados institucionales y personales de objetos matemáticos ligados a la aproximación frecuencial de la enseñanza de la probabilidad* (Tesis doctoral). Universidad de Granada, Granada.

Shaughnessy, J. M., Ciancetta, M. y Canada, D. (2004). Types of student reasoning on sampling tasks. En M. J. Høines y A. B. Fuglestad (Eds.), *Proceedings of the 28th meeting of the International Group for Psychology and Mathematics Education* (Vol. 4, pp. 177–184). Bergen: Bergen University College Press.

Tversky, A. y Kahneman, D. (1974). Judgement under uncertainity: heurístics and biases. *Science*, *185*, 1124-1131.

Van Dooren, W., De Bock, D., Depaepe, F., Janssens, D. y Verschaffel, L. (2003). The illusion of linearity: expanding the evidence towards probabilistic reasoning. *Educational Studies in Mathematics*, *53*(2), 113-138.

Contribuições dos Autores

1ª autor: conceitualização; curadoria de dados; análise formal; investigação; metodologia; visualização; redação – rascunho original; redação – revisão e edição.

2º autor: conceitualização; curadoria de dados; análise formal; investigação; metodologia; administração do projeto; supervisão; visualização; redação – rascunho original; redação – revisão e edição.

3ª autor: conceitualização; curadoria de dados; análise formal; investigação; metodologia; supervisão; visualização; redação – revisão e edição.

4º autor: conceitualização; curadoria de dados; análise formal; investigação; metodologia; administração do projeto; supervisão; visualização; redação – revisão e edição.