

NIVELES DE RAZONAMIENTO PROBABILÍSTICO DE ESTUDIANTES DE BACHILLERATO SOBRE LA NOCIÓN DE LA DISTRIBUCIÓN BINOMIAL

Jaime Israel García García¹; Elizabeth Hernández Arredondo²

¹Universidad Autónoma de Guerrero, ²Universidad Latina (México)
jaime.garcia.matedu@gmail.com, ehernandez@unila.edu.mx

Resumen

El presente trabajo analiza el desarrollo del razonamiento probabilístico que muestran estudiantes de bachillerato a través de sus respuestas a una tarea referida a una situación inscrita en el tema de distribución binomial. El problema consiste en predecir el comportamiento de sorteos de tipo binomial ($n = 2$, $p = \frac{1}{2}$) con el fin de observar si los estudiantes perciben y consideran dos aspectos elementales: la variabilidad y la distribución de los resultados. Las respuestas se analizaron bajo la taxonomía SOLO; es decir, se organizan en cuatro categorías: preestructural, uniestructural, multiestructural y relacional. En general, los jóvenes muestran un mejoramiento en su razonamiento probabilístico apoyado por el uso de la simulación física y computacional.

Palabras clave: razonamiento probabilístico, distribución binomial, taxonomía SOLO

Abstract

This work analyzes the development of the probabilistic reasoning that high-school students show through their responses to a task involving a situation of binomial distribution. The problem consists in predicting the behavior of binomial-type distribution ($n = 2$, $p = \frac{1}{2}$) to observe whether students perceive and consider two basic aspects: variability and result distribution. The responses were analyzed according to the SOLO taxonomy; that is, they are organized in four categories: pre-structural, mono-structural, multi-structural, and relational categories. Overall, the students show improvement in their probabilistic reasoning supported by the use of physical and computational simulation

Key words: probabilistic reasoning, binomial distribution, SOLO taxonomy

■ Introducción

El desarrollo del razonamiento probabilístico es uno de los objetivos importantes de la educación matemática de los niveles básicos y pre-universitarios (Jones, Langrall y Mooney, 2007); mientras que la binomial es la distribución discreta más importante que se estudia en el curso de probabilidad de bachillerato. No obstante, se han hecho pocas exploraciones sobre problemas que intenten desarrollar el razonamiento probabilístico tomando como punto de partida el concepto de distribución, en particular, la binomial (García y Sánchez, 2013).

Este estudio presenta una exploración sobre el desarrollo del razonamiento probabilístico que muestran estudiantes de bachillerato, que no han sido introducidos al estudio formal de la probabilidad y la estadística, tomando como punto de partida la noción de distribución binomial en su forma más simple, cuando se hacen simulaciones físicas (monedas) y computacionales (Fathom). Este interés surge de una de las once preguntas formuladas por Pfannkuch y Reading (2006) para el desarrollo de la noción de distribución: “¿Cómo se desarrolla el razonamiento acerca de distribuciones desde sus formas o aspectos más simples a unos más complejos?”

■ Fundamentos

La distribución binomial es una de las distribuciones discretas más importantes de probabilidad que se prescribe en los programas de estudio de bachillerato; en gran parte, por sus aplicaciones en la inferencia estadística y por su estrecha relación de convergencia con la distribución normal.

La enseñanza de esta distribución tradicionalmente se enfoca en lograr que los estudiantes sepan obtener la fórmula y aplicarla en la resolución de problemas. Sin embargo, aunque aprendan la fórmula y la apliquen de manera eficiente, se ha mostrado que éstos no son capaces de razonar con, y acerca de, esta distribución (Landín, 2014).

El razonamiento alrededor de la distribución binomial requiere articular varios conceptos de probabilidad, como espacio muestral, variable aleatoria, probabilidad, independencia, combinatoria, aleatoriedad, variabilidad y distribución; siendo estos tres últimos los que se presentan en su forma más elemental en la actividad que realizaron los estudiantes. Es muy difícil que éstos estructuren dichas nociones mediante el enfoque clásico que enfatiza sólo en el aprendizaje de la fórmula, por lo que es necesario buscar formas alternativas de introducir el tema. Desde un enfoque constructivista es importante explorar lo que los estudiantes saben y la manera en que razonan, pues a partir del conocimiento resultante se podrán observar las formas alternativas de la enseñanza del tema.

■ Marco Conceptual

Este estudio optó por establecer un marco conceptual apoyado en las posturas de Jabareen (2009) y Miles y Huberman (1994), quienes declaran que este explica, ya sea en forma gráfica o a través de una narración, las principales cosas que van a ser estudiadas y las supuestas relaciones entre ellas, restringiéndose sólo a los principales conceptos que permiten entender a éste; en particular, la *Distribución binomial*, como el contenido probabilístico cuyo aprendizaje se explora; el *Razonamiento probabilístico*, como la forma en que los estudiantes unen y transforman sus ideas previas sobre probabilidad para construir el concepto de distribución binomial; y finalmente, la *Taxonomía SOLO*, como un método que nos permite analizar los razonamientos de éstos mediante la jerarquización de las respuestas que dan a las tareas.

a) *Distribución binomial*. El caso más simple de una *distribución binomial* (diferente a la Bernoulli) se presenta cuando sus parámetros son $n = 2$ y $p = \frac{1}{2}$; este es el modelo de probabilidad de un gran número de experiencias aleatorias, en especial, la de lanzar dos monedas (o equivalentemente dos veces una moneda) y observar el número de ‘águilas’ [nombre popular de una cara de las monedas mexicanas] que ocurren. La distribución binomial $B(x, 2, \frac{1}{2})$ se representa fácilmente mediante una tabla, como se muestra en la Tabla 1.

Tabla 1. Representación tabular de la distribución $B(x, 2, \frac{1}{2})$

Distribución binomial $B(x, 2, \frac{1}{2})$			
X	0	1	2
P	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{1}{4}$

Cabe destacar que las tareas de este estudio no se enfocan en el manejo adecuado del algoritmo de la distribución binomial; sino en el descubrimiento de relaciones de tipo estructural que subyacen a una situación binomial por parte de los estudiantes.

b) *Razonamiento probabilístico.* Se entiende por razonamiento al proceso de formular juicios o afirmaciones a partir de otras proposiciones ya conocidas u observaciones de un fenómeno. Bakker y Derry (2011) mencionan que las personas forman conceptos como consecuencia de su participación en actividades sociales de dar y pedir razones, es decir, gracias al razonamiento se configuran y consolidan sus conceptos. En consecuencia, es conveniente proponer tareas que propicien el razonamiento matemático de los estudiantes.

En el campo de la probabilidad, es necesario proponer tareas en situaciones de incertidumbre que logren captar el interés de los estudiantes, que lo enfrenten a algunas ‘grandes ideas’ de la probabilidad (Gal, 2005) y que le exijan realizar razonamientos probabilísticos; tales tareas deben involucrar a los estudiantes en procesos de estimación de la propensión de ocurrencia de eventos, y de su predicción y representación, y no restringirse sólo a problemas de cálculo (García, Medina y Sánchez, 2014). Particularmente, miramos el razonamiento probabilístico sobre la noción de distribución binomial como la forma en que los estudiantes unen y transforman sus ideas previas sobre probabilidad para desarrollar dicho concepto.

c) *Taxonomía SOLO.* Una manera de describir el razonamiento acerca de la distribución binomial, desde sus formas o aspectos más simples a otras algo más complejas, es analizando y organizando en una jerarquía las respuestas de los estudiantes a una tarea. Por ejemplo, Jonas, Langrall, Thornton y Mogill (1997) generaron y validaron una jerarquía cuyo objetivo es caracterizar el razonamiento probabilístico de los estudiantes de primaria, tomando como base el modelo de Biggs y Collis.

La taxonomía SOLO (Structure of the Observed Learning Outcome) desarrollada por Biggs y Collis (1991), es un instrumento diseñado para evaluar la calidad de aprendizaje de los estudiantes en una amplia variedad de situaciones escolares, en la mayoría de áreas temáticas, que permite identificar elementos de conocimiento que éstos utilizan como recursos para formar razonamientos; también, ha sido la base para construir jerarquías de comprensión o razonamiento en probabilidad y estadística. En este estudio se utiliza la taxonomía SOLO como una herramienta para analizar y organizar las respuestas de los estudiantes en niveles estructurales; es decir, se categorizan en: preestructural, uniestructural, multiestructural y relacional.

■ Metodología

Participantes. Treinta y siete estudiantes de un grupo de primer semestre (15-16 años) del CCH-UNAM y el profesor titular del grupo. Los estudiantes no habían recibido ningún tipo de enseñanza formal de probabilidad y estadística; además, no recibieron información respecto el propósito del estudio.

Instrumento. Se diseñó una actividad con 13 tareas, llamada “*A la suerte*”. El problema consiste en predecir el comportamiento de sorteos de tipo binomial ($n = 2$, $p = \frac{1}{2}$) con el fin de observar si los estudiantes perciben y consideran aspectos de dos elementos principales, a saber: la variabilidad y la distribución de los resultados (0, 1, 2). Para realizar dichas tareas los estudiantes contaron con el uso de objetos manipulables (monedas) y un software educativo de estadística (Fathom), con el fin de realizar simulaciones referentes al número de ‘águilas’ obtenidas en el lanzamiento de dos monedas. Por las limitaciones de espacio, en este informe sólo se expondrá y comentará una tarea de la actividad, cuya finalidad era predecir si Ana, Beto y Carlos podían tener la misma frecuencia de éxito a largo plazo (ver Tabla 2).

Tabla 2. Situación y tarea de estudio

Situación <i>A la suerte</i>	Tarea de predicción
Ana, Beto y Carlos juegan al lanzamiento de dos monedas para decidir qué programa de la televisión verán. Si no sale ninguna ‘águila’, gana Ana; si sale exactamente una ‘águila’, gana Beto; y si salen dos ‘águilas’, gana Carlos.	¿Cuántas veces crees que gana Ana, Beto, y Carlos si el juego se repitiera 1000 veces? Número de veces que gana Ana (0 águilas) = _____ Número de veces que gana Beto (1 águila) = _____ Número de veces que gana Carlos (2 águilas) = _____

Procedimientos. El estudio consistió en cuatro etapas, en las que los estudiantes: 1) contestaron el cuestionario previo de la actividad para conocer su nivel de razonamiento intuitivo frente a la situación binomial, 2) desarrollaron una actividad guiada que consistía en utilizar objetos manipulables (monedas) para propiciar una familiarización con la situación, 3) desarrollaron otra actividad guiada que consistía en utilizar el software con el fin de simular 1000 veces el resultado de la variable aleatoria binomial de la situación en repetidas ocasiones, y 4) contestaron un cuestionario posterior, con las mismas tareas del previo, para conocer su nivel de razonamiento adquirido después de las actividades de simulación.

■ Análisis y resultados

Las respuestas de los estudiantes se organizan de acuerdo a su complejidad estructural, bajo el supuesto de que entre mayor es ésta, mejor la calidad de la respuesta. Para hacerlo se ha tenido como guía la taxonomía SOLO (Biggs y Collis, 1991). Mediante un análisis comparativo de las respuestas, se

identifican los elementos o componentes que utilizan, es decir, los diferentes conceptos y/o procedimientos que emplean en su resolución. Con base en la identificación de tales componentes, se definen niveles crecientes de complejidad estructural intentando aplicar el esquema expuesto en el marco conceptual. La Tabla 3 presenta los niveles jerárquicos de la taxonomía SOLO, y la descripción de cada uno de ellos, que se determinaron de acuerdo a las respuestas de los estudiantes a la tarea en cuestión.

Tabla 3. Niveles jerárquicos para la tarea de predecir el comportamiento de sorteos de tipo binomial

Nivel	Descripción de tipo de características que presentan las respuestas para su clasificación
Preestructural	<ul style="list-style-type: none"> • Proporcionan frecuencias cuya suma no corresponde al número de sorteos • Proporcionan frecuencias que no reflejan la forma de la distribución ni la variabilidad del fenómeno
Uniestructural	<ul style="list-style-type: none"> • Proporcionan frecuencias que corresponden a los valores esperados • Proporcionan frecuencias que reflejan la forma de la distribución: mayor frecuencia al evento de obtener un águila, y menores frecuencias a los eventos de obtener ninguna águila y dos águilas • Proporciona frecuencias que reflejan la aleatoriedad del fenómeno: frecuencias que se presentan de manera desordenada
Multiestructural	<ul style="list-style-type: none"> • Proporcionan frecuencias que reflejan tanto la forma de la distribución (favorece al evento más probable) como la variabilidad (al proporcionar frecuencias distintas a los valores esperados), sin relacionarlas de manera adecuada al proporcionar frecuencias algo improbables que sucedan
Relacional	<ul style="list-style-type: none"> • Proporcionan frecuencias que reflejan tanto la forma de la distribución (favorece al evento más probable) como la variabilidad (al proporcionar frecuencias distintas a los valores esperados), relacionarlas de manera adecuada al proporcionar frecuencias probables que sucedan, es decir, con una desviación de 50

Notas:

- 1) Las frecuencias esperadas son números de referencia alrededor de los cuales estarán las frecuencias reales obtenidas de la realización efectiva de un gran número de sorteos. En nuestro caso, si consideramos la $B(x, 2, 1/2)$, 1000 sorteos, y pensamos en la frecuencia con la que ocurren cada uno de los resultados: 0, 1 y 2: la frecuencia esperada de 0 es 250; la frecuencia esperada de 1 es 500; y la frecuencia esperada de 2 es 250.
- 2) Mediante una simulación estadística en Fathom, se generan 1000 valores de la variable aleatoria y se cuenta la frecuencia de cada resultado (0, 1, 2) observando sus diferencia con las frecuencias esperadas. Este proceso se repite un gran número de veces (digamos 1000 otra vez) y se observa la distribución de las diferencias. Se toma un valor en el que el 95% de las veces el intervalo, con centro la frecuencia esperada y radio el valor tomado, contenga a las frecuencias reales; por lo que consideramos adecuada una desviación de 50.

En la Tabla 4 se presenta una ejemplificación del tipo de respuestas dadas por los estudiantes de acuerdo al nivel de razonamiento, seguida de una breve justificación sobre esta clasificación.

Tabla 4. Ejemplificación de acuerdo a nivel de razonamiento

Nivel de razonamiento	Respuesta del estudiante	Descripción
Preestructural	Número de veces que gana Ana (0 águilas) = 300 Número de veces que gana Beto (1 águila) = 550 Número de veces que gana Carlos (2 águilas) = 250	Proporciona frecuencias cuya suma no corresponde al número de sorteos.
Uniestructural	Número de veces que gana Ana (0 águilas) = 270 Número de veces que gana Beto (1 águila) = 460 Número de veces que gana Carlos (2 águilas) = 270	Proporciona frecuencias correspondientes a los valores esperados, sin mostrar sensibilidad a la variabilidad.
	Número de veces que gana Ana (0 águilas) = 250 Número de veces que gana Beto (1 águila) = 500 Número de veces que gana Carlos (2 águilas) = 250	Proporciona frecuencias destacando únicamente la forma de la distribución.
	Número de veces que gana Ana (0 águilas) = 500 Número de veces que gana Beto (1 águila) = 400 Número de veces que gana Carlos (2 águilas) = 100	Proporciona frecuencias que se presentan de manera desordenada sin reflejar la forma de la distribución.
Multiestructural	Número de veces que gana Ana (0 águilas) = 300 Número de veces que gana Beto (1 águila) = 600 Número de veces que gana Carlos (2 águilas) = 100	Proporciona frecuencias que reflejan tanto la distribución como la variabilidad; sin embargo, son algo improbables que sucedan.
Relacional	Número de veces que gana Ana (0 águilas) = 200 Número de veces que gana Beto (1 águila) = 550 Número de veces que gana Carlos (2 águilas) = 250	Proporciona frecuencias probables que sucedan considerando la forma de la distribución y la variabilidad.

En la Tabla 5 se muestran las frecuencias de respuesta por nivel de razonamiento a la tarea en el cuestionario previo y posterior.

Tabla 5: Frecuencias de las respuestas de los estudiantes de acuerdo a los niveles de razonamiento

Cuestionario	Nivel de razonamiento				Total
	Preestructural	Uniestructural	Multiestructural	Relacional	
Previo	18	13	3	2	36*
Posterior	6	14	7	9	36*

* Un estudiante no responde la tarea.

Las frecuencias de respuestas de la tarea muestran una mejoría del cuestionario posterior al previo. En el cuestionario previo, la mayoría de las respuestas de los estudiantes se ubicaron en los niveles preestructural y uniestructural, dieciocho respuestas presentan inconsistencia o aspectos irrelevantes a la tarea; mientras que en trece se considera la variabilidad o la distribución de los resultados de manera aislada. Después de las etapas de simulación, se observa una mejora, nueve respuestas se ubicaron en el nivel relacional, presentan una relación adecuada entre la distribución y la variabilidad al mostrar frecuencias probables que sucedan, con una desviación de 50; mientras que siete respuestas se clasificaron en el nivel multiestructural, consideran tanto la distribución (mayor frecuencia de ocurrencia en el valor central que en sus valores extremos) como la variabilidad (frecuencias distintas a los valores esperados), pero no logran relacionarlas de manera adecuada al proporcionar frecuencias algo improbables que sucedan.

Estos razonamientos probabilísticos informales que muestran los estudiantes en sus respuestas, se basan en ideas que generaron durante las actividades de simulación y no emanan de la aplicación de alguna definición, puesto que ninguna se introdujo o se explicó durante la intervención. No es posible afirmar esto con certeza, pero el proponer una distribución con la forma y variabilidad adecuada, ya es un indicio de que el estudiante comienza a entender, de manera informal, la noción de la binomial.

■ Discusión y conclusiones

Durante la realización de las actividades del estudio, los estudiantes no recibieron instrucción formal sobre el concepto de distribución binomial y no se les mencionó como responder las tareas; sólo se les proporcionaron indicaciones de cómo llevar a cabo las simulaciones, y con lo observado, se les solicitó responder nuevamente las preguntas de la situación *A la suerte*. El desarrollo del razonamiento probabilístico permitió observar la presencia de distintos elementos de la distribución binomial que influyen en los niveles jerárquicos: aleatoriedad, variabilidad, distribución.

En el cuestionario previo, la mayoría de los estudiantes espontáneamente no proporcionan de manera adecuada la gráfica de frecuencias de los sorteos de tipo binomial; pero con la ayuda de las simulaciones, en particular con la computacional, logran hacerlo aceptablemente, perciben la aleatoriedad de la situación y llegan a asociarla con su idea de frecuencia esperada, esto permite un razonamiento adecuado de la variabilidad relacionada con la distribución binomial. En consecuencia, las respuestas del cuestionario posterior fueron clasificadas en niveles más altos de razonamiento que las respuestas del previo, éstos propusieron distribuciones que cumplen con la forma de la distribución y la variabilidad, lo que indica un avance general del razonamiento probabilístico de los estudiantes. Enfatizamos este aspecto porque

consideramos que una comprensión adecuada de la noción de distribución binomial nunca debe separarse de la idea de que expresa un fenómeno de incertidumbre, en cuyas manifestaciones (repeticiones de sorteos de acuerdo a la distribución) se debe reflejar la variabilidad.

En conclusión, la noción informal de distribución binomial que logran desarrollar los estudiantes mediante el uso de la simulación, en especial con la computacional, puede considerarse como un complemento a los aspectos relacionados con el cálculo formal. Con estas actividades los estudiantes notaron la forma de la distribución y la variabilidad como rasgos característicos de esta distribución.

■ Referencias bibliográficas

- Bakker, A. y Gravemeijer, K. P. E. (2004). Learning to reason about distribution. En D. Ben-Zvi & J. Garfield (Eds.), *The challenge of developing statistical literacy, reasoning and thinking*. Dordrecht, The Netherlands: Kluwer Academic Press.
- 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). New York: Springer.
- García, J. & Sánchez, E. (2013). Niveles de razonamiento probabilístico de estudiantes de bachillerato frente a una situación básica de variable aleatoria y distribución. En J. M. Contreras, G. R. Cañadas, M. M. Gea y P. Arteaga (Eds.), *Actas de las Jornadas Virtuales en Didáctica de la Estadística, Probabilidad y Combinatoria* (pp. 417-424). Granada, Departamento de Didáctica de la Matemática de la Universidad de Granada. España.
- García, J. I., Medina, M., & Sánchez, E. (2014). Niveles de razonamiento de estudiantes de secundaria y bachillerato en una situación-problema de probabilidad. *Avances de Investigación en Educación Matemática*, 6, 5-23.
- Jabareen, Y. (2009). Building a conceptual framework: Philosophy, definitions, and procedure. *International Journal of Qualitative Methods*, 8(4), 49-62.
- Jones, G. A., Langrall, C. W., Thornton, C. A., & Mogill, A. T. (1997). A framework for assessing and nurturing young children’s thinking in probability. *Educational studies in Mathematics*, 32(2), 101-125.
- Jones, G.A., Langrall, C.W. y Mooney, E.S. (2007). Research in probability. En F. K. Lester, Jr. (Ed.), *Second handbook of research on mathematics teaching and learning* (pp. 909-955). Charlotte, NC, USA: Information Age-NCTM.
- Landín, P. (2014). *Niveles de razonamiento probabilístico de estudiantes de bachillerato sobre problemas binomiales*. Tesis de Doctorado no publicada, Centro de Investigación y de Estudios Avanzados del IPN. México.
- Miles, M. & Huberman, A. (1994). *Qualitative Data Analysis. An expanded sourcebook*. London, UK: Sage Publications.
- Pfannkuch y Reading (2006). Reasoning about distribution: a complex process. *Statistical Education Research Journal*, 5(2), pp. 4-9.