Análisis de tres problemas de probabilidad utilizados en la investigación

Ernesto Alonso Sánchez Sánchez[1]

RESUMEN

Los problemas que suelen utilizarse en la investigación didáctica para explorar las intuiciones y el razonamiento probabilístico de los estudiantes frecuentemente no son reelaborados para servir en la enseñanza del tema. La razón es que las explicaciones de los conceptos subyacentes de los problemas que exploran las intuiciones o razonamientos frecuentemente tienen un alto grado de dificultad. En particular, algunas ideas fundamentales de la probabilidad como Incertidumbre, Aleatoriedad y Variación deberían estar presentes en la enseñanza del tema de todos los niveles, pero su tratamiento genuino en las aulas no está del todo resuelto. No obstante, algunos problemas utilizados en la investigación y el uso de recursos tecnológicos pueden contribuir en su solución; además, se abre una ventana para abordar otros conceptos pertinentes de la estadística de manera embrionaria e informal, como distribución, distribución muestral y significación estadística. En este sentido, analizamos tres problemas extraídos de la literatura con el propósito de ofrecer pautas de cómo utilizarlos en la enseñanza de la probabilidad en los niveles básico y medio. Se concluye que, con ayuda de recursos tecnológicos, es posible utilizar en la enseñanza problemas complejos utilizados en la investigación, de modo que se desarrollen las grandes ideas de incertidumbre, aleatoriedad y variabilidad.

PALABRAS CLAVE

Incertidumbre, Aleatoriedad, Variación, Tecnología, Problemas utilizados en la Investigación.

---

[1] esanchez@cinvestav.mx
Departamento de Matemática Educativa, CINVESTAV
https://orcid.org/0000-0002-8995-7962

Sánchez-Sánchez, E. A. (2023). Análisis de tres problemas de probabilidad utilizados en la investigación. En Juárez Ruiz, E. L., Hernández Rebollar, L. A., & Castañeda, A. (Eds.), *Tendencias en la Educación Matemática 2023* (pp. 91–111). SOMIDEM Editorial.
https://doi.org/10.24844/SOMIDEM/S3/2023/01-05

INTRODUCCIÓN

Desde los trabajos pioneros sobre razonamiento e intuiciones sobre probabilidad de los niños y adolescentes (Piaget e Inhelder, 1975; Fischbein, 1975) se ha avanzado en entender las dificultades, falsas concepciones, creencias y posibilidades de los estudiantes en su aprendizaje de la probabilidad, y de los profesores en su práctica docente (Jones et al. 2007, Langrall, et al. 2017). Una parte central en tales estudios son los instrumentos que utilizan los investigadores para obtener datos duros de las intuiciones, razonamientos o pensamientos de los estudiantes, la mayoría de los cuales son muy creativos y originales, pero no necesariamente utilizables para la enseñanza. Tales problemas contrastan con los que se suelen utilizar en el aula, pues estos se enfocan directamente en el cálculo; dicho enfoque deriva en el desarrollo de habilidades lógicas y aritméticas (álgebra de eventos, proporcionalidad y combinatoria), pero no propicia el desarrollo de otras ideas de probabilidad, como incertidumbre, aleatoriedad, variabilidad y la ley de los grandes números.

Desde que se comenzó a popularizar la tecnología digital, varios autores advirtieron sobre el potencial de la tecnología para propiciar un aprendizaje más completo de la probabilidad, utilizando programación para hacer simulaciones (con la función RAMDOM, que genera números aleatorios), cálculo automático y representación de gráficas, todo esto inspirado en las técnicas del método de Montecarlo (Biehler, 1991; Sobol, 1983). Poco después se comenzó un periodo de desarrollo de software dinámico de estadística amigable para la enseñanza en el que no se requería programación. Como resultado se volvieron muy populares, entre otros, los softwares Probability Explorer, TinkerPlots y Fathom.

Con estos softwares se han sugerido muchos problemas y secuencias para propiciar el aprendizaje de probabilidad de los niños y adolescentes (Stohl & Tarr, 2002, Pratt, 2000); no obstante, no ha habido una recuperación de los problemas que se han utilizado en la investigación para explorar su potencial para la enseñanza con ayuda de las posibilidades del software dinámico. Surge la pregunta de investigación ¿Cómo enfocar problemas clásicos surgidos de la investigación didáctica en probabilidad utilizando tecnología digital? El objetivo del presente estudio es mostrar que los problemas utilizados en la investigación se pueden redimensionar con el uso de la tecnología digital, convirtiéndolos en problemas para desarrollar el razonamiento estadístico y no sólo para explorarlo. En particular, en este trabajo recuperamos tres problemas que exploran las intuiciones y el razonamiento en probabilidad de los niños y adolescentes, y los analizamos utilizando tecnología. Al hacerlo, apuntamos hacia un posible uso en el aula.

Grandes ideas de la probabilidad

Una comprensión razonable de la probabilidad requiere situarla en un contexto que abarque distintos conceptos. Tradicionalmente, se enmarca en la matemática, destacando elementos como eventos (considerados conjuntos), el álgebra de eventos y la propia probabilidad (definida como la razón entre casos favorables y posibles). No obstante, existen marcos que ofrecen complementos y proponen incluir otras nociones.

Gal (2005) propone un marco (Tabla 1) para la competencia probabilística (probability literacy) formado por elementos de conocimiento y disposicionales. A cada elemento de los enlistados los llama bloques de construcción (ladrillos) de la competencia probabilística. En los elementos de conocimiento propone cinco elementos, los cuales aclara que están íntimamente relacionados entre sí de forma compleja y que no es apropiado creer que se presentan de manera aislada. Con relación a las grandes ideas, advierte que no se pueden caracterizar mediante una definición, como se suele hacer con los conceptos matemáticos

Tabla 1

Bloques de construcción para la competencia probabilística.

Elementos de conocimiento: 1. *Grandes ideas*: Variación, Aleatoriedad, Independencia y predicción/incertidumbre 2. *Cálculo de probabilidades*. Formas de encontrar o estimar la probabilidad de un evento 3. *Lenguaje*. Los términos y métodos utilizados para comunicarse acerca de la probabilidad 4. *Contexto*. Comprensión del papel e implicación de asuntos de probabilidad y mensajes en varios contextos y en discursos públicos y personales 5. *Preguntas críticas*. Temas para reflexionar cuando se trata de probabilidades
Elementos disposicionales: 1. Posición crítica 2. Creencias y actitudes 3. Sentimientos personales hacia la incertidumbre y el riesgo

Nota. Tomado de Gal (2005), p. 46.

No abundaremos en este modelo, sólo queremos resaltar que el primer elemento de conocimiento que propone Gal son las grandes ideas. Creemos que todavía hacen falta investigaciones que exploren cómo desarrollar las grandes ideas en el aula. En este trabajo analizaremos tres problemas utilizados en la investigación que, aparte de revelar las intuiciones y razonamiento de los estudiantes acerca de la probabilidad, podrían permitir, con ayuda de la tecnología, desarrollar ideas de incertidumbre, aleatoriedad y variabilidad.

Alarcón (1982)

Jesús Alarcón Bortolussi (Papini) fue un investigador de México del Departamento de Matemática Educativa. Fue experto en didáctica de la matemática, especialmente en geometría, cálculo y probabilidad. Contribuyó en una profunda reforma curricular de México en 1993 en lo que concierne a las Matemáticas. Fue el autor principal de “El Libro para el Maestro” de matemáticas, el cual fue publicado durante muchos años por la Secretaría de Educación Pública de México. Infortunadamente falleció prematuramente en 1997. Su trabajo de doctorado consistió en el diseño, aplicación y análisis de un cuestionario de probabilidad a 300 alumnos de entre 12 y 14 años. El objetivo fue evaluar los juicios probabilísticos de los estudiantes frente a preguntas que realmente contuvieran incertidumbre. En esta sección analizaremos el problema más controvertido de su cuestionario, ver la Figura 1.

Figura 1
Un problema con máxima incertidumbre de Alarcón (1982)

Juan ha elegido una de las bolsas que se presentan abajo. Después, él sacó al azar seis bolas con reemplazo de la bolsa elegida. El resultado de las extracciones fue de 2 bolas blancas y 4 negras. ¿Qué bolsa eligió Juan?

Resultados de la aplicación del problema

El cuestionario se aplicó a dos grupos, uno de 12-13 años y el otro de 13-14 años. La mayoría de las respuestas de ambos grupos se distribuyen en las opciones 2, 3 y 4, siendo viable para algunos que la extracción de 2B4N pudo provenir de la bolsa 2. En efecto, 26% de los estudiantes de 12-13 años eligió la opción 2 (Muy posiblemente la bolsa 1), 46% se inclinó por la opción 3 (No hay razón para preferí una bolsa) y 20% por la opción 4 (Muy posiblemente la bolsa 2); mientras que 21% de los estudiantes de 13-14 años eligió la opción 2, 46% la opción 3 y 23% la opción 4 (Tabla 2).

Es interesante que tanto la opción 2 como la 4 son atractivas para los estudiantes. En ambos grupos, el 46% no encuentran razones para decidir la urna de la que puede provenir el resultado. Es plausible interpretar que el criterio de los que responden correctamente (26% de 7º grado y 21% de 8º grado) es porque perciben que es muy poco probable el resultado a partir de la urna 2 y, entonces, apuestan por que la urna 1 fue la elegida. Este resultado y el opuesto (inciso 4) pueden interpretarse con conceptos de riesgo en el siguiente sentido. La bolsa 1 encarna total incertidumbre, de manera que conlleva mucho riesgo, pues el contenido de la bolsa puede ser cualquiera. En cambio, la bolsa 2 permite calcular la probabilidad de que haya ocurrido el evento 2B4N. Los estudiantes que son adversos al riesgo elegirán la bolsa 2, pues la probabilidad de obtener el evento objetivo al menos es medible, mientras que los que son propensos al riesgo elegirán la bolsa 1, considerando que vale la pena correr el riesgo.

Tabla 2

Resultados del reactivo de la Figura 2

Edad	Seguro la bolsa 1	Muy posible la bolsa 1	No hay razón para elegir	Muy posible la bolsa 2	Seguro la bolsa 2	Total
12 – 13 años	3	12	21	9	1	46
	6%	26%	46%	20%	2%	100%
13 – 14 años	3	12	26	13	2	56
	5%	21%	46%	23%	4%	99%

¿Dónde está la incertidumbre?

El problema didáctico que formuló Alarcón (1982) se basa en el reconocimiento de que se solían utilizar problemas de "pronóstico" en la investigación sobre razonamiento probabilístico de su tiempo. Estos problemas tienen el formato:

> Una urna A tiene *u* bolas blancas y *w* bolas negras; otra urna B tiene *x* bolas blancas y bolas negras. Se va a elegir una urna, de ella se extrae al azar una bola y se observa su color. Si quieres obtener una bola blanca ¿Qué urna elegirías? O ¿Da lo mismo cualquiera de las dos?

Alarcón criticó estos problemas observando que son de naturaleza aritmética, ya que sólo requieren del cálculo y comparación de fracciones, pero supuestamente se debían interpretar como probabilidades. En realidad, las respuestas a tales problemas son exactas y no enfrentan a los estudiantes a la incertidumbre. Es por esto que propuso mejorar los instrumentos de exploración a través de problemas en "situaciones de decisión a posteriori", estos tienen el siguiente esquema:

> Encuentra entre dos bolsas, que contienen bolas blancas y negras, aquella en la cual se obtuvo un resultado dado de uno o varios sorteos. Ver Figura 2.

Conviene preguntarse ¿En qué sentido estos problemas contienen incertidumbre? La respuesta es que, en muchas situaciones particulares no se podrá saber con certeza de qué bolsa provino la bola resultante, por ejemplo, supongamos que una bolsa A contiene 1 bola blanca y una bola negra, una bolsa B contiene una bola blanca y dos negras. El resultado fue bola blanca. ¿De qué urna provino ésta? El cálculo lleva a que es más probable que haya provenido de la bolsa A por el teorema de Bayes:

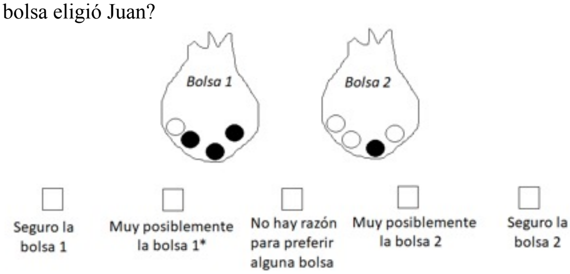
$$\left[P(bolsa\ A \mid bola\ blanca = \frac{\frac{1}{2}}{\frac{1}{2}+\frac{1}{3}} = \frac{3}{5};\ P(bolsa\ B \mid bola\ blanca = \frac{\frac{1}{3}}{\frac{1}{2}+\frac{1}{3}} = \frac{2}{5} \right]$$

Una pequeña reflexión, o con ayuda de una simulación, lleva a la conclusión de que, aunque la mejor respuesta es “Proviene de la bolsa A”, se corre mucho riesgo de fallar. Por otro lado, la respuesta “no hay razón para preferir una bolsa” no es del todo incorrecta, pero no toma en cuenta la pequeña ventaja que ofrece la primera respuesta. En resumen, este problema es cualitativamente diferente a los problemas de pronóstico en los que simplemente se deben calcular probabilidades, pues en los problemas de decisión a posteriori ninguna decisión lleva a una respuesta segura.

Figura 2

Un problema de decisión a posteriori

Juan ha elegido una de las bolsas que se presentan abajo. Después, él sacó al azar seis bolas con reemplazo de la bolsa elegida. El resultado de las extracciones fue de 2 bolas blancas y 4 negras. ¿Qué bolsa eligió Juan?



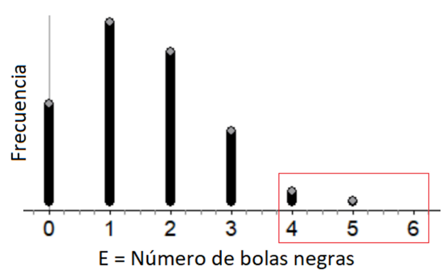
El problema de máxima incertidumbre de Alarcón y las pruebas de significación

Conviene mostrar que el trabajo de Alarcón (1982) está bien enraizado en el razonamiento estadístico. Él no elaboró la relación del paradigma que propuso con el esquema de pruebas de significación, pero sin duda lo utilizó para evaluar la calidad de las respuestas. Volvamos al problema de la figura 1, que hemos juzgado de máxima incertidumbre. En él, cada bolsa representa una población y el evento que se supone que ocurrió (2B4N) es una muestra; un estadístico característico de esta situación es la proporción de bolas negras, pero para facilitar la simulación vamos a considerar el estadístico E como el número de bolas negras. La hipótesis nula (H_0) es que el resultado proviene de la bolsa 2 (porque se quiere rechazar esa hipótesis y es la única que ofrece datos para hacer el cálculo), entonces se calcula la probabilidad de obtener el resultado: {E = 4 u otro más extremo, es decir, E = 4, 5, 6; bajo H_0}. Aunque esta probabilidad requiere de la distribución binomial, se puede estimar con ayuda de una simulación. En la figura 3 presentamos una distribución simulada de la variable E obtenida de repetir mil veces el experimento con el software Fathom. En ella hemos encerrado en un cuadro los puntos que son favorables al evento {E=4, 5, 6}

Una estimación de la probabilidad del evento de que ocurran 4, 5 o 6 bolas negras es de 26/1000, es decir 0.026. Este número es muy pequeño, por lo que tenemos la opción de que ocurrió un milagro o es erróneo suponer . El método de pruebas de significación sugiere que en el caso de que el evento tenga probabilidad menor de 0.05 bajo la hipótesis de que H_0 es verdadera, entonces se rechaza la hipótesis nula H_0. Así, la decisión racional al problema de la figura 2 es elegir la bolsa 1.

Figura 3

Distribución de frecuencias de la variable No. de negras



Measures from Sample of ...

NoNegras	0	192
	1	350
	2	293
	3	139
	4	22
	5	4
Column Summary		1000

S1 = count ( )

La tecnología y los problemas de decisión de Alarcón

Consideraremos los problemas de decisión propuestos por Alarcón (1982) en su forma generalizada, antes de llegar al problema límite en el que una bolsa tiene contenido desconocido. Lo formulamos en términos de urnas y no de bolsas:

> Una urna A tiene *u* bolas blancas y *w* bolas negras; otra urna B tiene *x* bolas blancas y bolas negras. Se elige una urna y de ella se extraen seis bolas y se observa su color. Si el resultado es una muestra con B bolas blancas y N negras ¿De qué urna se obtuvo dicho resultado?

Como hemos dicho, este tipo de problemas está cargado de incertidumbre. Para resolverlo, los estudiantes suelen aplicar una heurística de representatividad (Tversky & Kahneman, 1982), es decir, eligen la urna cuya proporción de bolas blancas es más cercana a la proporción de bolas blancas del resultado. Aunque esta estrategia suele ser apropiada, no ofrece una forma de medir las probabilidades involucradas en la decisión tomada. Pero para calcular tales probabilidades se requiere de técnicas probabilísticas que no están al alcance de los estudiantes de niveles preuniversitarios. Es en este punto en el que surge la hipótesis de que es posible que los estudiantes estimen informalmente tales probabilidades con ayuda de tecnología. Pero queremos enfatizar que, desde nuestro punto de vista, lo importante de utilizar simulación computacional para resolver los problemas es conocer su estructura y familiarizarse con representaciones del comportamiento de eventos inciertos en repeticiones masivas de un experimento, y no tanto la sola estimación de las probabilidades. Nosotros utilizamos Fathom, pero la idea se puede desarrollar también con otros softwares.

Para ilustrar cómo utilizar el software Fathom para simular un problema de decisión de Alarcón tomamos una situación mencionada arriba, pero con un evento-resultado diferente:

> La urna A contiene una bola blanca y otra negra y la urna B contiene una bola blanca y dos negras. Se extraen seis bolas con reemplazo y se obtienen 2 bolas Blancas y 4 bolas Negras (Abreviado 2B4N). ¿De qué urna proviene el resultado?

El resultado 2B4N es más similar al contenido de la urna 2 que al contenido de la urna 1, por lo que la elección de la urna 2 es razonable. ¿Pero qué probabilidad tiene está decisión de atinar a lo que realmente sucedió? ¿Qué tan probable es que el resultado provenga de la urna 1? Como vimos arriba, estos problemas se resuelven con la fórmula de Bayes, pero se puede modelar y simular sin recurrir a ella. Algo importante es que, resolviéndolo informalmente, se va prefigurando el contenido de dicha fórmula. La manera de hacerlo en Fathom es crear dos poblaciones de 1000 resultados provenientes de cada urna. Esto se hace mediante los siguientes pasos.

Paso 1. Se llenan dos columnas en Fathom (llamada Urna 1 y Urna 2) con 1000 datos de los sorteos de las funciones RandomPick (“B”, “N”) para la urna 1, y RandomPick (“B”, “N”, “N”) para la urna 2. Esto quiere decir que en cada columna se registran 1000 resultados de la urna 1 y 1000 de la 2, independientemente unos de los otros.

Paso 2. En el segundo paso se genera una muestra de tamaño 6 proveniente de cada urna y se definen dos medidas que indican el número de bolas blancas de la muestra de la Urna 1 y el número de bolas blancas de la Urna 2; se registran ambos números.

Paso 3. Se repite 1000 veces el paso 2, de manera que se obtengan 1000 parejas ordenadas. Cada pareja representa en su primer componente el número de bolas blancas de una muestra de la Urna 1 (Esta variable la nombramos BUrna 1), y en su segundo componente el número de bolas blancas de una muestra de la Urna 2 (Esta variable la nombramos BUrna 2). La generación de la población de cada urna asegura que las entradas son independientes una de la otra.

Paso 4. Las 1000 parejas ordenadas tienen en cada uno de sus componentes números del 0 al 6. Con la función "Summary" de Fathom se pueden organizar las frecuencias de cada pareja ordenada en una tabla de 2×2, como la tabla de la Figura 4. El número 276 (marcado en la figura) representa a todas las parejas que en su primer componente tienen el número 2, es decir, que favorecen al evento 2B4N de la Urna 1, mientras que 347 corresponden a resultados que al evento 2B4N obtenidos de la Urna 2. Hay entonces 623 (= 276+347) parejas que representan resultados favorables al evento 2B4N, de éstas, 276 provienen de la urna 1, por lo que:

$$P(Urna\ 1 \mid 2N4B) = \frac{276}{623} = 0.44$$

$$P(Urna\ 2 \mid 2N4B) = \frac{347}{623} = 0.56$$

La captura de pantalla de la Figura 5 muestra las diferentes salidas que el software proporciona a algunas de las acciones correspondientes a los pasos que se describieron arriba.

Figura 4

Tabla de doble entrada de 1000 parejas ordenadas

Measures from Sample of Situación

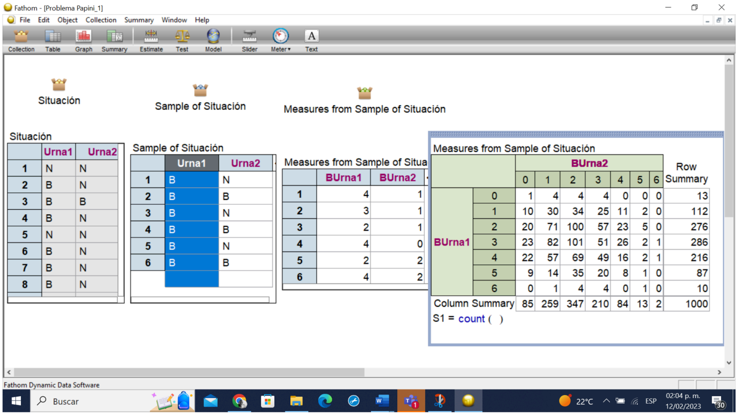
		BUrna2							Row Summary
		0	1	2	3	4	5	6	
BUrna1	0	1	4	4	4	0	0	0	13
	1	10	30	34	25	11	2	0	112
	2	20	71	100	57	23	5	0	276
	3	23	82	101	51	26	2	1	286
	4	22	57	69	49	16	2	1	216
	5	9	14	35	20	8	1	0	87
	6	0	1	4	4	0	1	0	10
Column Summary		85	259	347	210	84	13	2	1000

S1 = count ( )

Hemos expuesto muy brevemente el método para mostrar una posible manera de abordar el problema, aunque, sin duda, puede haber otras formas más ingeniosas. En el espacio que tenemos no ha sido posible destacar el papel de la incertidumbre y cómo se obtienen consecuencias acerca de ella, pero haciendo dos o más simulaciones con el software, se puede experimentar con la incertidumbre; en el trabajo del aula, si se cuenta con software adecuado, el profesor puede encontrar la manera de hacerlo.

Figura 5

Captura de pantalla de los resultados de la simulación



Green (1982a y 1982b)

David Robert Green es un investigador del Reino Unido de la universidad de Loughborough. Es experto en desarrollo curricular, evaluación educativa, probabilidad, estadística y enseñanza y aprendizaje de las matemáticas. Su trabajo sobre las intuiciones del azar y la probabilidad tuvo mucho impacto en los años noventa. Colaboró con Batanero et al. (1998) en un proyecto sobre el significado de la aleatoriedad y sus consecuencias para la enseñanza. En su investigación de doctorado desarrolló el proyecto en el que elaboró y aplicó un cuestionario a cerca de 3000 estudiantes de los grados 7 a 12. Se planteó el objetivo de explorar las intuiciones acerca del azar y la probabilidad de los estudiantes, en particular, sus juicios sobre aleatoriedad. El problema de la figura 6 fue diseñado para saber las intuiciones sobre aleatoriedad que tienen los niños y jóvenes.

Figura 6

Problema de Green (1982a)

El maestro le pidió a Clara y a Susana que arrojaran 150 veces una moneda y que registraran en cada tirada si el resultado es cara o cruz. Cada vez que salga cara se anota "1" y cada vez que salga cruz se anota "0". En seguida aparecen las listas de unos y ceros, anotados por cada una de ellas:

Clara:

010110011001010110110100011100011011010101100100010101001110011010110010110
010110010010111011001101101010010110010101100010011010110011101110101100011

Susana

100111011110100111001001110010001110111111010101011110000001000101001000001

0001100010100000000011001000000001111100001101010010010011111101001100011000

Se sabe que una chica hizo correctamente los 150 lanzamientos, pero la otra hizo trampa, inventando la sucesión. ¿Puedes descubrir quién de ellas hizo realmente el experimento y quién inventó el resultado? ¿Qué chica hizo trampa? Explica tu respuesta.

Resultados de la aplicación del problema

En la tabla 3, los datos revelan que poco más de la tercera parte de los estudiantes de cada nivel responden correctamente que la secuencia de Clara es la que se hizo con trampa. En consecuencia, poco menos de dos terceras partes de los estudiantes de cada nivel responden, erróneamente, que Susana hizo trampa. Una observación importante es que la calidad de las respuestas no mejora con la edad, pues la proporción de respuestas correctas es casi constante en los cinco niveles. Las razones que ofrecieron y sus frecuencias relativas se presentan a continuación:

Tabla 3

Frecuencias de respuestas a la pregunta 20

Grado	Clara	Susana	Sin respuesta
1	38	53	1
2	34	56	2
3	30	59	4
4	40	46	3
5	35	49	5
Promedio	35.4	52.4	3

Argumentos de los estudiantes para justificar su elección:

Relacionadas con la regularidad de los patrones:

1. La secuencia de Clara tiene un patrón demasiado regular (21%)*
2. La secuencia de Susana tiene un patrón irregular (3%)

Relacionada con la diferencia entre el número de "1s" y de "0s"

3. La secuencia de Clara tiene muchos unos; por lo tanto, la proporción está lejos de 0.5 (5%)
4. La secuencia de Susan tiene muchos ceros, por lo tanto, su proporción está lejos de 0.5. (18%)

Relacionada con la cercanía entre el número de "1s" y de "0s"

5. La secuencia de Clara está muy cerca de la proporción 0.5 (3%)
6. La secuencia de Susan está muy cerca de la proporción 0.5 (1%)

Relacionadas con la longitud de las cadenas:

8. La secuencia de Clara tiene cadenas muy cortas (3%)*
9. La secuencia de Susan tiene cadenas muy largas (18%)

Las únicas razones que son correctas son la 1 y la 7; en ésta, sólo 3% utilizan el argumento de que la secuencia de Clara tiene cadenas muy cortas.

¿Qué propiedades tiene la aleatoriedad?

El problema de Green (1982a) pide diferenciar entre una secuencia realmente aleatoria y otra que trata de imitarla; su solución requiere de un profundo sentido de la aleatoriedad y del funcionamiento del razonamiento estadístico. Para resolverlo en el contexto de la estadística requiere recurrir al concepto de pruebas de significación y, por lo tanto, no era fácilmente accesible para los profesores en aquel tiempo. En la actualidad es posible analizar el problema con la misma idea de pruebas de significación con la ayuda de la tecnología, pero desde un enfoque informal. De lo que se trata es de evaluar si el resultado que obtuvo cada una de las chicas es usual bajo la hipótesis de que fueron generados por una distribución de Bernoulli con . Para esto, conviene considerar tres estadísticos: 1) El número de éxitos ("1") de la secuencia. 2) El número de cadenas[1] de la secuencia. 3) La longitud de la cadena más larga.

En la tabla 4 se proporcionan los datos de los tres estadísticos (n_1 es el número de unos en la secuencia)

Tabla 4

Datos básicos de las secuencias de Clara y Susana

Lista de Clara: $n_1 = 78$, $n_0 = 72$	Lista de Susana: $n_1 = 67$, $n_0 = 83$
Número de cadenas = 96	Número de cadenas = 68
Longitud de la cadena más larga: 3	Longitud de la cadena más larga: 9

De lo que se trata es de ubicar los valores de cada estadístico dentro de la distribución correspondiente y decidir si es usual o no. Para aproximar las distribuciones se simula la generación de secuencias de longitud 150 y se calcula el estadístico respectivo. Esto se repite unas 1000 veces y los valores del estadístico se grafican en una distribución. Las figuras 7, 8, y 9 muestran las distribuciones correspondientes a cada estadístico. Se obtuvieron con el software Fathom.

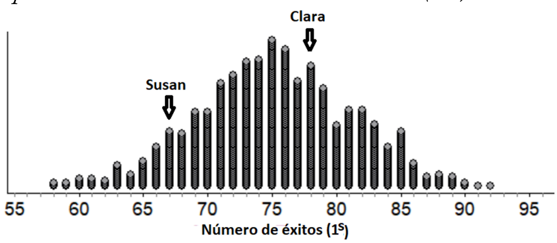
Se han señalado en las figuras el lugar en el que se ubican los datos particulares de la secuencia de Clara y de la de Susana. Se dice que una muestra es inusual cuando está en los extremos y su frecuencia u otra más extrema no excede el 5% de todas las secuencias.

En el caso del estadístico "número de éxitos" (equivalente al número de "1") (Figura 7), tanto el valor de la variable en la secuencia de Susan como en la de Clara están en una zona de resultados usuales, por lo que no se puede descartar la hipótesis de que fueron generadas por la distribución de Bernoulli con $p = \frac{1}{2}$.

[1] Una cadena es una serie de resultados iguales, por ejemplo, la secuencia 0110001011 tiene seis cadenas, a saber: 0, 11, 000, 1, 0, 11. La longitud de una cadena es el número de resultados de la cadena, así, las longitudes de las anteriores cadenas son: 1, 2, 3, 1, 1, 2.

Figura 7

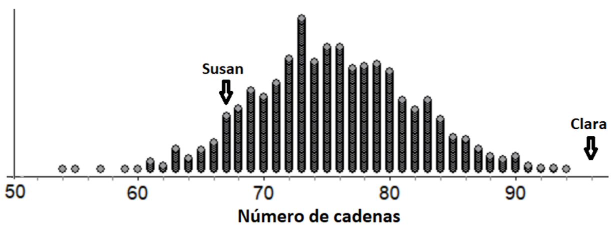
Distribución aproximada del estadístico "Número de unos ("1")"



En el caso del estadístico "número de éxitos" (equivalente al número de "1") (Figura 7), tanto el valor de la variable en la secuencia de Susan como en la de Clara están en una zona de resultados usuales, por lo que no se puede descartar la hipótesis de que fueron generadas por la distribución de Bernoulli con $p = \frac{1}{2}$.

Figura 8

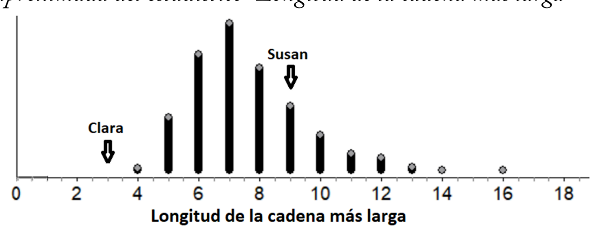
Distribución aproximada del estadístico "Número de cadenas"



La distribución del estadístico "Número de cadenas" (Figura 8), a diferencia del estadístico anterior, muestra que el resultado de Clara (Número de cadenas = 96) es inusual, pues tiene una frecuencia de cero ocurrencias. Este resultado lleva a rechazar la hipótesis de que la secuencia de Clara fue generada con la distribución de Bernoulli con p= ½. En cambio, el valor del estadístico para la secuencia de Susana es usual, pues está dentro de un rango de resultados frecuentes. Esta prueba es suficiente para resolver el problema y decidir que quien hizo trampa fue Clara. Pero también el tercer estadístico ofrece el mismo resultado.

Figura 9

Distribución aproximada del estadístico "Longitud de la cadena más larga"



La figura 9 presenta la distribución de los datos simulados de variable "longitud de la cadena más larga", en ella se observa que la secuencia de Clara es inusual bajo la hipótesis de que la secuencia se genera con la distribución de Bernoulli con p= ½, mientras que el valor del estadístico de Susana está en una región de resultados frecuentes. Esto reafirma el resultado de que Clara hizo trampa.

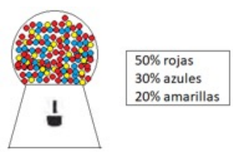
Shaughnessy et al. (1999)

J. Mike Shaughnessy es un investigador retirado de la Universidad de Portland, Estados Unidos. Su investigación se ha enfocado en la comprensión de los estudiantes sobre probabilidad y estadística, para lo que ha sintetizado y aprovechado los trabajos de psicólogos y educadores matemáticos. Ha sido autor de los artículos de reseña de la investigación en probabilidad y estadística de los Handbooks de la NCTM: Grows (1992), Lester (2007) y del compendio Cai (2017). Shaughnessy et al. (1999) elaboraron y administraron un cuestionario de datos y azar para explorar el razonamiento de 324 estudiantes en los grados 4-12 (9-17 años). El objetivo fue explorar el razonamiento de los estudiantes sobre la variabilidad estadística en situaciones tanto de manejo de datos como de azar. El problema de la máquina de gomas de mascar se presenta en la figura 10.

Figura 10

El problema de una máquina expendedora de gomas de mascar

Una máquina expendedora de gomas de mascar contiene mezcladas 50% de gomas rojas, 30% de gomas azules y 20% de gomas amarillas. Pregunta: *Jenny saca 10 gomas de mascar de la máquina ¿Cuál es tu mejor predicción del número de gomas rojas que saldrán?*



Resultados de la aplicación del problema

Se encontraron respuestas variadas que corresponden a estrategias diferentes. Unas son las respuestas basadas en la idea de que cualquier cosa puede pasar, por ejemplo: 0, 9, 4, 7, 10 y 1. Los resultados 0, 1 y 10 rojas son muy poco probables de modo que la secuencia no es un buen ejemplo de lo que puede ocurrir. Otras respuestas están sesgadas hacia mayor número de rojos, probablemente por considerar que la población tiene mayor número de rojos, pero sin considerar la proporción, por ejemplo: 6, 7, 8, 8, 7 y 9. Dentro del grupo de los estudiantes de grado 12 hubo varios cuya respuesta fue 5, 5, 5, 5, 5 y 5; los autores conjeturaron que probablemente estos estudiantes estaban acostumbrados a responder preguntas, como "¿Cuál es la probabilidad de que...?", que requieren respuestas de un solo número en lugar de un rango de resultados posibles. Las mejores respuestas son aquellas que ofrecen valores alrededor del, o igual al, valor esperado, por ejemplo: 3, 5, 3, 6, 7, 5. Estas se evalúan como mejores teniendo en cuenta un criterio de representatividad: varían y están alrededor del valor esperado. Pero hay que tener en cuenta que su probabilidad no es mayor que la secuencia 5, 5, 5, 5, 5, 5. Además, dentro de dichas respuestas hay dos clases que no son muy representativas, a saber: variabilidad muy estrecha: 4, 6, 5, 6, 4, 4 y variabilidad muy amplia: 2, 5, 8, 9, 8, 5. Los resultados expuestos son una descripción de las estrategias que suelen tener los estudiantes y una valoración con un criterio de representatividad, pero no con un criterio de probabilidad; en el análisis del problema se llegará a la respuesta desde un punto de vista estadístico.

¿Dónde está la variabilidad?

El cambio, movimiento o diferencia se presenta en cualquier fenómeno y, por tanto, también en los datos; de ahí que se diga que la variabilidad es omnipresente. Después de que los estadísticos profesionales Moore (1990) y Snee (1990) declararan que la variabilidad es una de las ideas fundamentales de la estadística y que esta afirmación se volviera un conocimiento regular en el ambiente de la estadística, en educación estadística, Shaughnessy (1997) se preguntó en su artículo sobre las oportunidades perdidas por la investigación, ¿dónde está la investigación sobre la variación estadística? Él alertó que las medidas de tendencia central y las gráficas eran los temas favoritos de los investigadores en educación estadística, pero la variabilidad prácticamente no había sido explorada por los investigadores hasta ese momento.

Shaughnessy et al. (1999) llevaron a cabo una investigación motivada por el análisis de las respuestas al problema de las gomas de mascar que hemos mencionado arriba. Este era un ítem de un cuestionario para la Evaluación Nacional del Progreso Educativo (NAEP) de 1996 en Estados Unidos. Lo que llamó la atención de Shaughnessy fue que los estudiantes, e incluso algunos profesores, creían que la respuesta correcta era el valor esperado (5 en este caso). En una muestra de 232 estudiantes, Zawojewski y Shaughnessy (2000) observaron que respondieron con un valor determinado, principalmente el valor esperado 5; excepto un participante respondió dando un rango de posibles números de gomas de mascar. A partir de esta observación idearon una serie de estudios sistemáticos para entender y propiciar el razonamiento sobre variación de los estudiantes.

Veamos desde la perspectiva de un especialista cuál es la solución al problema. Para esto simplificamos el problema suponiendo que la extracción de cada goma de mascar se reemplaza, de modo que no se alteran las proporciones de gomas de cada color en la población original. Para que la respuesta sea una predicción útil y realista es necesario que se cumplan las siguientes condiciones:

- La respuesta debe ser un rango de valores y no un resultado particular
- Dado que la variable es binomial y su distribución es simétrica, el rango con centro en el valor esperado es un evento con mayor probabilidad que otro equivalente no centrado en el valor esperado
- La probabilidad de que un resultado particular esté en el rango de valores debe ser mayor que $\frac{1}{2}$.
- Conviene elegir un rango que:
    1. No sea tan amplio que cubra todos los posibles resultados,
    2. Sea lo suficientemente amplio para que la probabilidad de que ocurra un resultado del rango sea mayor que un valor dado (mayor que $\frac{1}{2}$).

Por ejemplo, la predicción "de las 10 extracciones se obtendrán entre 3 y 7 bolas rojas" tiene una probabilidad de casi 90%, de modo que sólo hay una probabilidad de 0.1 de equivocarse. En la tabla 5 se muestran todos los posibles rangos centrados en el valor esperado (5) y su respectiva probabilidad.

Tabla 5
Probabilidades de que el número de rojas este en el rango indicado

Rango	5	4–6	3–7	2–8	1–9	0–10
Probabilidad	0.2461	0.6562	0.8906	0.9785	0.9980	1

Conviene hacer énfasis en que hacer la predicción "puede salir cualquier número entre 0 y 10 bolas rojas" [rango (0–10)], es equivalente a "cualquier resultado puede ocurrir" , esta respuesta es trivial y no utiliza el conocimiento probabilístico. En cambio, predecir "ocurrirá exactamente 5" es una predicción muy débil, pues se fracasará 3 de cada 4 veces. Las soluciones interesantes son las intermedias, las cuales, no son triviales y tienen una probabilidad mayor que ½ de ocurrir; la elección específica de un rango dependerá de lo que esté en juego en el contexto en el que se realiza el experimento, por ejemplo, si hay ganancias y pérdidas monetarias.

Volviendo al problema original de predecir qué puede ocurrir en la repetición de seis veces el experimento, una solución con 53% de probabilidad de ocurrir es:

$\underline{\text{entre 3 y 7 bolas rojas}}, \underline{\text{entre 3 y 7 bolas rojas}}, \underline{\text{entre 3 y 7 bolas rojas}}, \underline{\text{entre 3 y 7 bolas rojas}}, \underline{\text{entre 3 y 7 bolas rojas}}, \underline{\text{entre 3 y 7 bolas rojas}}.$

La probabilidad se mejora al 89% de probabilidad si se elige el rango de 2–8, es decir:

$\underline{\text{entre 2 y 8 bolas rojas}}, \underline{\text{entre 2 y 8 bolas rojas}}, \underline{\text{entre 2 y 8 bolas rojas}}, \underline{\text{entre 2 y 8 bolas rojas}}, \underline{\text{entre 2 y 8 bolas rojas}}, \underline{\text{entre 2 y 8 bolas rojas}}$

En el análisis del problema hemos utilizado recursos que los estudiantes no tienen o no manejan con fluidez, por lo que es lícito esperar las respuestas ingenuas que encontraron Shaughnessy et al. (1999). Una consecuencia posible es que, para que los estudiantes entiendan y puedan ofrecer una respuesta razonablemente buena, deben tratar este tipo de problemas en el curso de probabilidad una vez que hayan aprendido cómo interpretar y calcular probabilidades binomiales. No obstante, hay otra forma alternativa de que los estudiantes avancen en el problema.

Utilización de la tecnología para atacar el problema

Una vieja sugerencia didáctica para aprender probabilidad es la de hacer experimentos físicos reales y recopilar y graficar datos para descubrir los patrones subyacentes al experimento (Martin et al, 2021 y 2022). Esta técnica es promisoria, excepto porque se deben repetir muchos experimentos antes de que se comiencen a revelar los patrones subyacentes. Pero, si es posible complementar dichas actividades con la ayuda de un software educativo de estadística, entonces los estudiantes podrían, de manera informal, resolver el problema y entender varios conceptos subyacentes a él.

En lugar de una máquina expendedora de gomas de mascar se puede utilizar una urna o algo equivalente, con 50% de bolas rojas, 30% de bolas azules y 20% de bolas amarillas. Para hacer más simple la experiencia con manipulativos y, posteriormente, la simulación correspondiente con el software, conviene aclarar que es importante el supuesto de que las extracciones de cada bola deben ser con reemplazo. Una vez que los estudiantes hayan realizado una actividad con manipulativos, por ejemplo, que hayan sacado varias muestras de tamaño 10 de una urna y hayan anotado el número de bolas rojas, se puede recurrir al software.

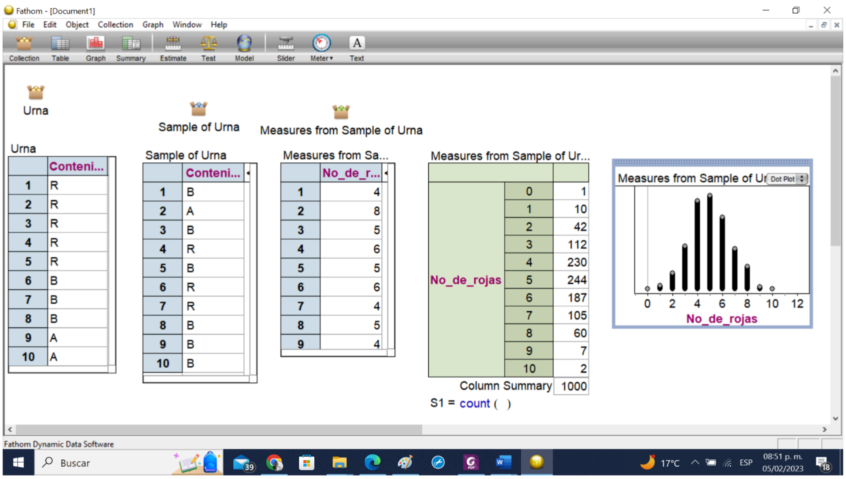
Con el software Fathom la simulación se puede hacer en sólo cinco pasos:

- Se define una población con 5 R's, 3 B's y 2 A's que representan bolas rojas, azules (blue) y amarillas respectivamente.
- Con el comando "sample of urna" obtener una muestra de 10 bolas con reemplazo
- Contar el número de R's en las muestras
- Repetir 1000 veces el proceso anterior para obtener 1000 valores
- Graficar la distribución de los resultados.

La captura de pantalla de la Figura 11 muestra las diferentes salidas que el software proporciona a algunas de las acciones correspondientes a los pasos anteriores. Note que la salida gráfica es una distribución que indica que los valores centrales son muy frecuentes, mientras que los valores en los extremos son muy poco frecuentes. Para hacer una buena predicción se descartan aquellos valores que tienen muy poca probabilidad, por ejemplo, 0, 1, 9 y 10, pues su frecuencia total suma menos del 5% (en la simulación es de 2%). La predicción es entonces su complemento, es decir, el rango de 2 a 8 que tiene una probabilidad de poco más del 95%. Una predicción más precisa, pero con menos probabilidad, se obtiene descartando los resultados, 0, 1, 2, 5, 9 y 10, cuya frecuencia es del 12.2%. Entonces, el rango de 3 a 7 tiene una probabilidad de 87.8%.

Figura 11

Captura de pantalla de los resultados de la simulación



Conclusiones

Los tres problemas que hemos recuperado de la investigación didáctica en probabilidad han sido innovaciones metodológicas para explorar las intuiciones y razonamiento en probabilidad de los estudiantes. Cada uno buscaba ofrecer evidencia de hasta qué punto las intuiciones o razonamiento de los estudiantes son capaces de lidiar con la incertidumbre, en el caso de Alarcón, con la aleatoriedad en el problema Green, y con la variabilidad en la investigación de Shaughnessy et al. (1999). Los resultados obtenidos, en los tres casos, indican que la mayoría de los estudiantes no ha desarrollado las intuiciones o recursos de razonamiento para elegir las respuestas más razonables. Por esto nos preguntamos ¿Cómo transformar los problemas en instrumentos para la enseñanza?

Esperamos que los bosquejos que hemos hecho sobre el uso del software para abordar los problemas hayan mostrado que con pocos recursos matemáticos y simbólicos por parte de los estudiantes se pueden obtener resultados matemática y estadísticamente sofisticados. Esto es posible porque, con el software, el estudiante entra en un micro–mundo en el que se encuentra con formas concretas de varios conceptos y puede manipularlos y obtener información de ellos (Biehler et al. 2013); por ejemplo, destaca que los estudiantes pueden generar, operar y sacar información de distribuciones de frecuencia sin haber pasado por episodios en los que se les enseñe el concepto de distribución.

No estamos obviando que el manejo del software requiere un proceso de aprendizaje, sin embargo, el gasto de tiempo y energía para las generaciones actuales (más acostumbradas que las anteriores al funcionamiento de la tecnología) es menor que el tiempo y energía que requiere el aprendizaje del contenido matemático (en los casos en que dicho aprendizaje es posible en este nivel) para entender los conceptos de probabilidad en cuestión. Tampoco se quiere implicar que el aprendizaje de aspectos de probabilidad con ayuda de tecnología sustituya a un aprendizaje analítico más formal; más bien, lo que se quiere es crear antecedentes intuitivos efectivos que fortalezcan y ayuden a profundizar el aprendizaje formal.

También conviene aclarar que todavía hace falta traducir las propuestas que se hicieron acerca del uso del software en investigación de diseño (Bakker, 2018). Esto quiere decir elaborar lecciones para un nivel escolar determinado en las que se especifiquen los pasos a seguir en el aula y la forma de presentar los problemas, implementar las lecciones y observar el avance en el desempeño y razonamiento de los estudiantes. Finalmente, y lo más importante, mediante el análisis retrospectivo revelar en qué medida se cumplen los supuestos de que la tecnología ayuda, con la materialización de las propuestas aquí sugeridas, a desarrollar las grandes ideas de incertidumbre, aleatoriedad y variabilidad de los estudiantes.

REFERENCIAS

Alarcón, J. (1982). *L'Appréhension des Situations Probabilistes chez de Elevés de 12-14 ans*. Tesis doctoral no publicada. Université Louis Pasteur.

Bakker, A. (2018). Design Research in Education. *A practical guide for early career researchers.* Routledge.

Batanero, C., Green, D. R., & Serrano, L. R. (1998). Randomness, its meanings and educational implications. *International Journal of Mathematical Education in Science and Technology, 29*(1), 113-123. https://doi.org/10.1080/0020739980290111

Biehler, R. (1991), Computers in probability education. En R. Kapadia, & M. Borovcnick (Eds.), *Chance Encounters: Probability in Education* (pp. 169–211). Springer. https://doi.org/10.1007/978-94-011-3532-0_6

Biehler, R., Ben-Zvi, D., Bakker, A., & Makar, K. (2013). Technology for enhancing statistical reasoning at the school level. En M.A. K. Clements, A. J. Bishop, C. Keitel, J. Kilpatrick, & F. K.S. Leung (Eds.). Third International Handbook of Education (pp. 643–689). Springer. https://doi.org/10.1007/978-1-4614-4684-2_21

Cai, J. (Ed.) (2017). *Compendium for Research in Mathematics Education*. National Council of Teachers of Mathematics.

Fischbein, E. (1975). *The Intuitive Sources of Probabilistic Thinking in Children.* Springer. https://doi.org/10.1007/978-94-010-1858-6

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). Springer. https://doi.org/10.1007/0-387-24530-8_3

Green, D. R. (1982a). A survey of probability concepts in 3000 pupils aged 11-16 years. En D. R. Grey, P. Holmes, V. Barnett, & G. M. Constable (Eds.), *Proceedings of the First International Conference on Teaching Statistics* (Vol. 1, pp. 766–783). Teaching Statistics Trust. https://bit.ly/49s0kzp

Green, D. R. (1982b). *A survey of probability concepts in 3000 pupils aged 11-16 years.* [Tesis doctoral, Loughborough University]. Loughborough University Research Repository. https://bit.ly/47pe9wV

Jones, G. A., Langrall, C. W., & Mooney, E. S. (2007). Research on probability. Responding to classroom realities. En F. K. Lester (Ed.), *Second handbook of research on mathematics teaching and learning* (pp. 909–955). National Council of Teachers of Mathematics-Information Age Publishing. https://bit.ly/462FxPW

Langrall, C. W., Makar, K., Nilsson, P., & Shaughnessy, J. M. (2017). Teaching and Learning probability and statistics: An integrated perspective. En J. Cai (Ed.), *Compendium for Research in Mathematics Education* (pp. 490–525). National Council of Teachers of Mathematics.

Lester, F. K. (Ed.) (2007). *Second Handbook of Research on Mathematics Teaching and Learning. National Council of Teachers of Mathematics and Information.* Age Publishing.

Martin, V., Thibault, M., & Homier, M. (2021). Self-reported practices of probability teaching: The use of manipulatives and technological tools. En ICME-14 (Ed.), *Proceedings of the Fourteenth International Congress on Mathematical Education* (pp. 1–8). International Commission on Mathematical Instruction. https://bit.ly/40yBFoT

Martin, V., Thibault, M., & Homier, M. (2022). Convergence and divergence in probability teaching in elementary and secondary school in Québec: A closer look at eight teachers' self-reported practices. *Canadian Journal of Science, Mathematics and Technology Education, 22*(3), 659–678. https://doi.org/10.1007/s42330-022-00241-2

Moore, D. S. (1990). Uncertainty. En L. A. Steen (Ed.), *On the shoulders of Giants. New Approaches to Numeracy* (pp. 95-137). National Academy Press. https://doi.org/10.17226/1532

Piaget, J., & Inhelder, B. (1975). *The Origin of the Idea of Chance in Children.* W.W. Norton & Company

Pratt, D. (2000). Making sense of the total of two dice. *Journal for Research in Mathematics Education, 31*(5), 602–625. https://doi.org/10.2307/749889

Shaughnessy, J. (1997). Missed opportunities in research on the teaching and learning of data and chance. In F. Biddulph & K. Carr (Eds.), *The Twentieth Annual Conference of the Mathematics Education Research Group of Australasia (MERGA)*(pp. 6–22). MERGA. https://bit.ly/3tUEhB9

Shaughnessy, J., Watson, J., Moritz, J., & Reading, C. (1999). School mathematics students' acknowledgment of statistical variation. En C. Maher (Ed.), *Proceedings of the 77th Annual Conference of the National Council of Teachers of Mathematics.* NCTM.

Snee, R. D. (1990). Statistical thinking and its contribution to total quality. *The American Statistician, 44* (2), 116–121. https://doi.org/10.2307/2684144

Sobol, I. M. (1983). *Método de Montecarlo.* Editorial MIR.

Stohl, H., & Tarr, J. E. (2002). Developing notions of inference using probability simulation tools. *Journal of Mathematical Behavior, 21*(3), 319–337. https://doi.org/10.1016/S0732-3123(02)00132-3

Tversky, A., & Kahneman, D. (1982). Judgements of and by representativeness. En D. Kahneman, P. Slovic, & A. Tversky (Eds.), *Judgement under Uncertainty: Heuristic and Biases* (pp. 84–98). Cambridge University Press. https://doi.org/10.1017/CBO9780511809477

Zawojewski, J. S., & Shaughnessy, J. M. (2000). Data and chance. En E. A. Silver, & P. A. Kenney (Eds.), *Results from the seventh mathematics assessment of the National Assessment of Educational Progress* (pp. 235–268). National Council of Teachers of Mathematics.