# Análisis de las Intuiciones y Conocimientos sobre Probabilidad de Estudiantes de Bachillerato

**Sandra Areli Martínez Pérez**
sandra.martinezperez@cinvestav.mx
https://orcid.org/0000-0002-9627-5150
*DME, CINVESTAV, IPN*
CDMX, México.

**Ernesto A Sánchez Sánchez**
esanchez@cinvestav.mx
https://orcid.org/0000-0002-8995-7962
*DME, CINVESTAV, IPN*
CDMX, México.
.

**Recibido:** 28/09/2020 **Aceptado:** 21/02/2021

## Resumen

El objetivo de esta investigación es hacer un diagnóstico de las intuiciones y conocimientos sobre conceptos básicos de probabilidad, en sus enfoques clásico y frecuencial, que exhiben los estudiantes de bachillerato antes de llevar a cabo con ellos un experimento de diseño sobre el enfoque frecuencial. Se aplicó un cuestionario de 10 ítems a 22 estudiantes de bachillerato (17-18 años) que habían estudiado temas de introducción a la probabilidad (hasta las distribuciones) del curso institucional que llevan. El cuestionario contiene tres preguntas para explorar su comprensión de algunos términos de probabilidad, seis problemas en situaciones de urnas y un problema con información incompleta, que explora las ideas espontáneas que surgen cuando intentan relacionar la probabilidad con contextos diferentes a los juegos de azar. Las respuestas se categorizaron para determinar los patrones presentes que permitan ofrecer características de sus conocimientos. Los resultados del análisis revelan conocimientos parciales de los términos experiencia aleatoria, frecuencia relativa y probabilidad que se relacionan más con nociones de sus vivencias personales que con las definiciones técnicas. Aunque calculan probabilidades clásicas y frecuenciales en situaciones simples de urnas, tienen dificultades en la consideración de resultados de extracciones sucesivas. La noción de repetibilidad de una experiencia aleatoria no emerge en algunas situaciones en que sería pertinente y se percibe que se basan en un modelo subjetivo que no requiere la repetibilidad del experimento.
***Palabras clave:*** Conocimientos. Enfoque frecuencial. Probabilidad. Experimento de diseño. Enfoque clásico.

# Análise das Intuições e Conhecimento sobre Probabilidade de Alunos do Ensino Médio

## Resumo

O objetivo desta pesquisa é fazer um diagnóstico das intuições e conhecimentos sobre os conceitos básicos de probabilidade, em suas abordagens clássica e frequencial, apresentados por alunos do ensino médio antes de realizar um experimento projetual sobre a abordagem frequencial com eles. Um questionário de 10 itens foi aplicado a 22 alunos do ensino médio (17-18 anos) que haviam estudado tópicos introdutórios à probabilidade (até as distribuições) do

curso institucional que estão cursando. O questionário contém três perguntas para explorar a compreensão de alguns termos de probabilidade, seis problemas em situações de urna eleitoral e um problema mal definido que explora as ideias espontâneas que surgem quando tentam relacionar a probabilidade a contextos diferentes do jogo. As respostas foram categorizadas para determinar os padrões presentes que permitem oferecer características de seu conhecimento. Os resultados da análise revelam um conhecimento parcial dos termos experiência aleatória, frequência relativa e probabilidade, que estão mais relacionados com noções de suas experiências pessoais do que com definições técnicas. Embora calculem as probabilidades clássicas e de frequência em situações de urna simples, eles têm dificuldade em considerar os resultados de sorteios sucessivos. A noção de repetibilidade de uma experiência aleatória não surge em algumas situações onde seria relevante e é percebida como baseada em um modelo subjetivo que não requer repetibilidade do experimento.
***Palavras chave:*** Conhecimento. Abordagem de frequência. Probabilidade. Experiência de design. Abordagem clássica.

**Analysis of Intuitions and Knowledge about Probability of High School Students**

**Abstract**

The objective of this research is to make a diagnosis of high school students' intuitions and knowledge about basic concepts of probability, in its classical and frequentist approaches, before carrying out a design experiment on the frequentist approach . A 10-item questionnaire was given to 22 high school students (17-18 years old) who had studied introductory topics to probability ( until distributions) in the institutional course they were taking. The questionnaire contains three questions to explore their understanding of some probability terms, six problems in ballot box situations, and one ill-defined problem that explores the spontaneous ideas that arise when the students try to relate probability to contexts different to gambling. The responses were categorized to determine the patterns that help us to offer characteristics of their knowledge. The results of the analysis reveal partial knowledge of the terms random experience, relative frequency and probability that are more related to notions of their personal experiences than to the technical definitions. Although the students calculate classical and frequency probabilities in simple ballot box situations, they have difficulties in considering the results of successive draws. The notion of repeatability of a random experience does not emerge in some situations where it is relevant and we interpret that it is based on a subjective model that does not require the repeatability of the experiment.
***Keywords:*** Knowledge. Frequency approach. Probability. Design experiment. Classical approach.

## Introducción y pregunta de investigación

En el panorama internacional de la investigación sobre aprendizaje y enseñanza de la probabilidad se puede observar una tendencia emergente a considerar un enfoque de modelación como una alternativa de enseñanza de la materia, que puede ofrecer un contenido flexible y apto para las aplicaciones científicas y prácticas (Pfannkuch, et al., 2016). Un enfoque de modelación implica que los estudiantes sean capaces de movilizar intuiciones (Fischbein, 1987) y nociones de los conceptos probabilísticos básicos que les permitan enfocar su atención sobre aquellos aspectos de las situaciones en contexto que son pertinentes para construir el modelo matemático correspondiente. Las situaciones en contexto, diferentes a las de juegos de azar, raramente satisfacen las condiciones para que pueda utilizarse el enfoque clásico de probabilidad (Batanero, Henry, & Parzysz, 2005). Por esta razón, entre los contenidos que se requieren para enseñar la modelación probabilística es fundamental que los estudiantes entiendan a un cierto nivel (aún por determinar) el enfoque frecuencial de probabilidad.

La investigación acerca del aprendizaje y enseñanza de este enfoque se ha emprendido apenas recientemente (Pfannkuch, & Ziedins, 2014; Stohl & Tarr, 2002). Jones, Langall y Mooney (2007, p. 368) en su reseña de investigación de probabilidad señalaban que: "…es evidente que existe un vacío en la investigación asociada con el enfoque frecuencial de la probabilidad". Para diseñar y llevar a cabo una secuencia de enseñanza con énfasis en la modelación conviene hacer un diagnóstico sobre las intuiciones y conocimientos probabilísticos que se requieren para ello, en particular, de aquellos relacionados con el enfoque frecuencial.

El informe que aquí se presenta es parte de una investigación cuyo propósito es llevar a cabo un experimento de diseño para introducir a estudiantes de bachillerato al concepto de probabilidad desde un enfoque frecuencial. En un primer momento se diseñó e implementó una secuencia de enseñanza basada en la resolución de un problema en un contexto de urnas desde un enfoque frecuencial. Del análisis de los datos obtenidos en esta experiencia surgieron dos tipos de consideraciones para la revisión del experimento.

Por un lado, se requerían ajustes y cambios en la secuencia de enseñanza y, por otro, se vio la necesidad de realizar un diagnóstico de las intuiciones y conocimientos de probabilidad de los estudiantes. El presente artículo tiene como objetivo analizar el resultado de dicha evaluación. En particular, en la perspectiva de diseñar lecciones tendientes a que los estudiantes se acerquen a un enfoque frecuencial se consideró importante explorar en tres direcciones: La primera, la manera en que entienden los términos básicos de *experiencia aleatoria*, *frecuencia relativa y probabilidad*. La segunda, investigar la manera en que responden preguntas sencillas en contexto de urnas con repetición de extracciones con algunas que implican pequeñas dosis de incertidumbre. La tercera, explorar sus ideas espontáneas acerca de la relación entre la probabilidad y algunos contextos reales. En consecuencia, nuestra investigación describe las intuiciones y conocimientos que expresan los estudiantes en sus respuestas a un cuestionario que explora estas tres direcciones.

**Antecedentes**

Con relación a los antecedentes de investigación, Dantal (1998) describe las respuestas a tres preguntas realizadas a estudiantes de un bachillerato francés: 1) ¿Qué es una experiencia aleatoria? 2) ¿Qué es un evento en una experiencia aleatoria? 3) ¿Qué es la probabilidad de un evento en una experiencia aleatoria? Aunque hay una variedad de respuestas para cada pregunta Dantal destaca los rasgos de las más frecuentes. Para la primera pregunta, las respuestas con mayor frecuencia (33%,) se centran en la impredecibilidad, por ejemplo: "Es una experiencia en la que no se puede predecir de antemano el resultado" (p. 67). Se puede observar que en general no mencionan como un rasgo definitorio de una experiencia aleatoria la posibilidad de describir todos sus posibles resultados (espacio muestral) ni la propiedad de repetibilidad. Para la segunda pregunta, las respuestas con más frecuencia (25%) identifican evento con resultado: "es uno de los resultados posibles al realizar la experiencia aleatoria" (p. 68). Dantal sugiere que la dificultad más grande de los estudiantes para concebir la noción de evento probabilístico, que se refleja en sus respuestas, es la de creer que un evento pertenece al mundo sensible, es decir, que es un hecho o dato de la experiencia y no un concepto abstracto. Con relación a las respuestas a la tercera pregunta las más frecuentes tratan de reproducir la definición clásica (14%) o se basan en la idea frecuencial (5%). Observa que los estudiantes creen sólo en un enfoque de la probabilidad (clásico o frecuencial), aunque a veces los mezclan; nadie define la probabilidad por sus axiomas (en la enseñanza francesa se incluía este enfoque); tienen muchas dificultades para describir adecuadamente el enfoque frecuencial de probabilidad y entre los que lo intentan, nadie alude a la idea de "un gran número de experiencias repetidas", ni de aproximación de frecuencias relativas a la probabilidad. Concluye que la enseñanza de la probabilidad en el bachillerato es difícil, pues requiere de la clarificación, desde la base, de muchos conceptos.

Jones et al. (2007) hacen una revisión de la investigación educativa en probabilidad y su exposición muestra los contenidos de líneas o tendencias que han sido de interés para los investigadores. En particular, cuatro puntos de su exposición se relacionan con el presente trabajo, a saber, el lenguaje del azar, la aleatoriedad, espacio muestral y la medida de la probabilidad, cuyos resultados se resumen en lo que sigue.

En diversos estudios se ha constatado un uso idiosincrático de términos relacionados con el azar (como eventos imposibles, posibles y ciertos) en proporciones significativas de estudiantes de niveles previos al bachillerato (Fischbein, Nello, & Marino, 1991; Green, 1983, Watson & Moritz, 2003). Amir y Williams (1999) encontraron que la mayoría de los estudiantes en su investigación no sabía explicar lo que entendían por azar más allá de que es algo que sucede. Muchos estudiantes utilizan la expresión 50-50 para indicar la presencia de incertidumbre (Tarr, 2002; Watson, 2005,). Por tanto, parece pertinente que un estudio diagnóstico sobre los conocimientos e intuiciones de probabilidad de los estudiantes contemple examinar la manera en que entienden los términos básicos de la teoría de la probabilidad. Las intuiciones acerca de la aleatoriedad en secuencias repetidas se reducen, en el mejor de los casos, a percibir el desorden (imprevisibilidad y patrones irregulares), sin que los estudiantes desarrollen espontáneamente un sentido de la estabilidad de las frecuencias a la larga, ni tampoco la frecuencia de cadenas largas de un solo resultado que suelen presentarse en secuencias aleatorias (Batanero & Serrano, 1999; Green, 1988,).

Con relación al espacio muestral se ha explorado acerca de la dificultad de describir los espacios muestrales de experimentos compuestos (Batanero, Navarro-Pelayo & Godino, 1997; Fischbein y Grossman, 1997). Otras investigaciones muestran que los estudiantes no siempre consideran el espacio muestral al determinar las probabilidades o examinar las frecuencias de los resultados (Ayres & Way, 2000; Fischbein & Schnarch, 1997; Shaughnessy & Ciancetta, 2002).

Varias investigaciones centradas en un enfoque frecuencial de probabilidad se han ubicado entre el final de la educación primaria y la educación media básica. Nilsson (2014) explora la relación entre probabilidad clásica y frecuencial en la enseñanza de la probabilidad basada en experimentaciones. En su revisión de la literatura sugiere dos categorías para agrupar a dichos estudios:

- *Dirección de mapeo*: Probabilidad teórica → Experimento → Probabilidad frecuencial
- *Dirección de inferencia*: Experimento → Probabilidad frecuencial → Probabilidad teórica

Los estudios que siguen la dirección de mapeo son los de Aspinwall y Tarr (2001), Pratt (2000), Pratt y Noss (2002), Stohl y Tarr (2002) y Nilsson (2009). En ellos los estudiantes utilizan algún recurso tecnológico para simular un generador aleatorio (moneda, dado, urna) y observan las frecuencias con las que ocurre un evento compuesto representadas mediante una gráfica; por ejemplo, en Pratt (2000) se genera la suma de los resultados de dos dados y sus frecuencias se representan en un diagrama de sectores. De esta manera, los estudiantes descubren una distribución no equiprobable y algunas propiedades de las frecuencias relativas, como su estabilidad a la larga. Begué, Batanero, Gea y Beltrán-Pellicer (2017) evaluaron la comprensión del enfoque frecuencial de probabilidad de una muestra de 302 estudiantes de segundo (13 años) y cuarto (15 años) de ESO (Enseñanza Secundaria Obligatoria) y el avance en dicha comprensión del segundo grupo respecto al primero. Como instrumento utilizaron la situación de imaginar que se arrojan 100 chinchetas y se observa que 68 caen con la punta hacia arriba y 32 caen apoyadas sobre la punta y el canto de la cabeza de la chincheta. Se les pide a los estudiantes proponer resultados razonablemente posibles del experimento de arrojar cuatro veces 100 chinchetas (indicar el número de chinchetas que se espera con la punta hacia arriba y el número del complemento con la punta hacia abajo).

Encontraron que aproximadamente 32% de la población propusieron respuestas adecuadas con relación al valor medio y 40% con una variabilidad aceptable. Las respuestas poco razonables se explican por una tendencia a caer en sesgos derivados del sesgo de equiprobabilidad y de la heurística de representatividad. Respecto al avance detectado se registró una leve mejoría, pues la media de las propuestas de los de 2º fue de 56,9 y las de 4º de 58,9; respecto al variabilidad de sus propuestas, se registró que en 2º hubo 81 respuestas dentro de un rango razonable, mientras que para 4º hubo 88. Los autores concluyen que los porcentajes de comprensión de la tarea son aún bajos, que el avance de 2º a 4º curso no es suficiente y que la instrucción en probabilidad debería desarrollarse de manera gradual a lo largo del ciclo de la secundaria.

Por otro lado, en Sánchez y Valdez (2014) se destacó la componente predicción / incertidumbre de las ‘grandes ideas’ de la probabilidad que propuso Gal (2005), componente que prefigura la ley de los grandes números (LGN). En particular, exploraron tareas similares al problema número 10 del presente estudio, en los que debido a que el tamaño de la muestra es pequeño, la incertidumbre es alta. Aunque la versión matemática de la LGN es tema de cursos universitarios de probabilidad, con la ayuda de recursos tecnológicos se han vislumbrado trayectorias para incluir versiones empírico-virtuales de dicha ley en la enseñanza básica. Entre ellos, Lee, Angoti y Tarr (2009) Stohl, Rider y Tarr (2004) y Stohl y Tarr (2002), utilizan el software *Probability Explorer* para entender cómo los estudiantes relacionan las frecuencias relativas con la probabilidad o cómo toman una decisión sobre una hipótesis con base en los datos. Para Ireland y Watson (2009) un aspecto crucial en el proceso de aprendizaje de la LGN en probabilidad es la necesidad de que los estudiantes establezcan una conexión entre las experiencias reales y las experiencias virtuales (simuladas); es decir, el problema es que los estudiantes vean que una simulación es una genuina representación de una situación real y que los resultados vistos en ella se transfieran a ésta.

## Marco conceptual

Concebimos que un marco conceptual es una selección de algunos conceptos interrelacionados que juntos contribuyen a alcanzar un nivel de comprensión de algún aspecto importante del fenómeno estudiado. Esta caracterización se basa en la siguiente, más general, que proponen Miles y Huberman (1994):

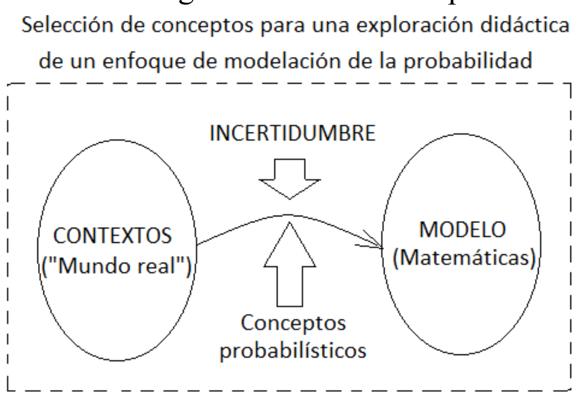
> Un marco conceptual explica, ya sea gráficamente o en forma narrativa, los principales aspectos que van a ser estudiadas –los factores, conceptos o variables clave –y las posibles relaciones entre ellos. Un MC puede ser rudimentario o muy elaborado, muy ligado a una teoría o basado en el sentido común, simplemente descriptivo o conteniendo hipótesis de tipo causal (p. 18).

Para el presente estudio hemos seleccionado explorar los conocimiento e intuiciones sobre nociones de probabilidad de los estudiantes, poniendo atención en aquellos conceptos que podrían favorecer o frenar sus estudios futuros desde un acercamiento de modelación de la probabilidad. En consecuencia, el marco conceptual para el presente estudio está formado por cuatro núcleos conceptuales: *Conceptos probabilísticos*, *enfoque de modelación a la probabilidad*, *contextos e incertidumbr*e (Figura 1). A continuación, resumiremos aspectos de estos conceptos concernientes al presente estudio.

*Enfoque de modelación a la probabilidad*. Consiste en incluir en el proceso de enseñanza-aprendizaje actividades de modelación que permitan entender y resolver problemas de probabilidad de situaciones de los mundos físico, biológico y social. Un modelo es una representación matemática de una situación del "mundo real" que sirve, entre otros propósitos, para entenderla, resolver problemas derivados de ella o tomar decisiones.

*Conceptos probabilísticos.* La modelación se hace posible gracias a la utilización de conceptos básicos de la teoría de la probabilidad pues estos indican aquellos aspectos que deben identificarse y precisarse en las situaciones de los diferentes contextos. Además, los conceptos probabilísticos se vinculan con conceptos matemáticos, lo que también ofrece pautas para la elaboración del modelo matemático de una situación determinada.

Figura 1. Marco Conceptual




FUENTE: Elaborado por los autores (2020)

*Contextos.* En lugar de hacer referencia al "mundo real" se utiliza el término *contexto.* Con esto, se enfatiza que, en el ámbito educativo, la modelación parte de enunciados de situaciones que evocan situaciones reales o realistas, frecuentemente acompañadas con datos, pero no se trata directamente con las situaciones reales. El rasgo común de los contextos que estudia la probabilidad es que se relacionan con la incertidumbre.

*Incertidumbre.* Es difícil que una definición aclare la complejidad del concepto, pero se puede definir la incertidumbre como ausencia o falta de certidumbre. Una clase de incertidumbre es una propiedad de algunos fenómenos del mundo (incertidumbre externa) y otra, subjetiva, es un estado de la mente relacionado con el conocimiento del sujeto acerca del fenómeno (incertidumbre interna) (Kahneman & Tversky, 1982).

En un ambiente educativo, los problemas relacionados con el mundo real son enunciados que describen situaciones y problemas en contextos extra–matemáticos. En probabilidad, los juegos de azar, como las extracciones de bolas en urnas han sido privilegiados, tanto en el origen histórico de la disciplina como en su enseñanza. Una explicación es que las situaciones de juego facilitan la modelación (Chaput, Girard, & Henry, 2011 los llaman modelos pseudo–concretos). No obstante, un enfoque en la modelación de la enseñanza y aprendizaje de la probabilidad se dirige a cubrir fenómenos del mundo físico, biológico y social y no sólo fenómenos de juegos ni otros encubiertos en contextos artificiales. En este sentido, conviene enfatizar que la característica principal de las situaciones que modelan la probabilidad es la incertidumbre. Para esto se ha desarrollado un sistema de conceptos probabilísticos generales que facilitan la tarea, de los cuales mencionamos los siguientes que están en la base del sistema: Experiencia aleatoria, espacio muestral, resultado, evento, repetibilidad de una experiencia aleatoria, frecuencia relativa y, por supuesto, los enfoques clásico y frecuencial de probabilidad.

## Método

Como instrumento para recoger datos sobre nuestro objeto de estudio se elaboró y aplicó un cuestionario; ocho de las preguntas del cuestionario son de respuesta abierta y dos de opción múltiple. En la medida en que la mayoría de los datos que se obtuvieron mediante dicho procedimiento son textos escritos de los estudiantes en los que ellos tratan de comunicar un significado, el estudio se ubica dentro de los métodos cualitativos (Denzin & Lincoln, 2011). De manera más precisa, adoptamos un enfoque de estudio de caso (Flyvbjerg, 2011), donde el caso es el conjunto de respuestas del cuestionario del grupo. El análisis realizado está inspirado en los procedimientos de codificación y categorización que se recomiendan en los primeros pasos de estudios de Teoría Fundamentada (Birks & Mills, 2015), pero cabe aclarar que el objetivo del presente estudio no es producir una teoría como lo recomienda esta metodología general, sino sólo algunas hipótesis que puedan ser útiles para entender mejor la problemática de la enseñanza de la probabilidad desde un enfoque de modelación con el fin de diseñar lecciones para su enseñanza.

### Participantes

Participaron 22 estudiantes de bachillerato (17-18 años) que habían estudiado el primer semestre del curso curricular de probabilidad; en dicho primer semestre se cubre temas de estadística descriptiva y probabilidad justo antes de iniciar el tema de distribuciones. En el momento de la aplicación los participantes llevaban el segundo semestre del curso. Uno de los investigadores era el profesor del grupo y el otro era observador.

### Instrumento.

Se diseñó un cuestionario, que se incluye como Anexo, de 10 preguntas , algunas tomadas de otros estudios, con el fin de evaluar los conocimientos e intuiciones de los estudiantes participantes sobre conceptos de probabilidad. El cuestionario tiene tres tipos de preguntas que agruparemos en tres bloques correspondientes. El primer bloque evalúa conceptos básicos de probabilidad, el segundo aspectos de la relación entre los enfoques clásico y frecuencial; el tercer bloque consiste en situaciones en diferentes contextos, pero con información incompleta; se trata de que argumenten si existe la probabilidad de un evento dado. En lo que sigue se analizan las preguntas en cada uno de los tres bloques y su intención.

*Bloque 1. Caracterización de conceptos básicos.* Contiene las tres preguntas 1 a 3 (Ver Anexo), cuyo objetivo es revelar los conocimientos e intuiciones que los estudiantes ponen en juego al tratar de caracterizar tres conceptos básicos de la probabilidad: experiencia aleatoria, frecuencia relativa y probabilidad. Se siguió el formato del estudio de Dantal (1998) quien formula de manera directa: ¿Qué entiendes por…? Éste pregunta por los conceptos de experiencia aleatoria, evento y probabilidad", nosotros en lugar de preguntar ¿Qué entiendes por evento?, preguntamos por la noción de frecuencia relativa. Este cambio obedece a que queremos centrarnos en los conocimientos que sobre el enfoque frecuencial de probabilidad puedan expresar los estudiantes. No se esperaba que respondieran con la definición precisa, pues tenemos la hipótesis de que un curso como el que llevaron no aseguraría que los estudiantes se apropiaran de tales conceptos a partir de su definición. Más bien esperamos que los estudiantes acomodasen estos nuevos términos a sus nociones del sentido común, ya sea como resultado de su curso previo, o de manera espontánea durante el proceso de resolución de la tarea. Las preguntas dan la oportunidad de que las respuestas ofrezcan indicios del enfoque frecuencial de probabilidad. Así, respuestas avanzadas a las tres preguntas podrían incluir rasgos de este enfoque, por ejemplo, se podría mencionar que una experiencia aleatoria puede repetirse, que la frecuencia relativa consiste en el número de veces que ocurre un evento entre el número de veces que se repite la experiencia y que la probabilidad es el número al que se acercan las frecuencias relativas cuando se llevan a cabo muchas repeticiones.

*Bloque 2. Situaciones que implican la frecuencia relativa.* En este bloque se exponen, las preguntas del 5 al 10. Las preguntas 5, 7 y 10a piden un cálculo simple de probabilidad y se espera que no representen ninguna dificultad para los estudiantes. Sirven de preguntas de control para constatar que pueden llevar a cabo operaciones de probabilidad de nivel básico.

La pregunta 6, requiere que los estudiantes asocien la probabilidad clásica que pueden calcular de la distribución de bolas en la urna con una posible distribución de frecuencias de 1000 extracciones de una bola de la urna. Se espera y se toman como aceptables la elección de la opción b o de la opción c, pues ambas reflejan la probabilidad de las bolas; la opción b porque es la proporción exacta y tiene la probabilidad máxima y la opción c porque está cerca de dicha proporción. Esta leve indeterminación puede producir cierta incertidumbre, pero también el que la opción a, se conciba como posible. No obstante, conviene descartar la opción b como respuesta correcta porque su probabilidad es cercana a cero y también la opción d ya que incluye a la opción a.

La pregunta 8 se distingue de las demás de problemas de bloque 2 porque no utiliza el contexto de urnas y porque no se conocen las probabilidades de base. Ahora sólo se cuenta con la información de los resultados de haber lanzado las 100 chinchetas (Begué, et al., 2017). Se espera que con este resultado se estime la probabilidad de cada posición (0,68 arriba, 0,32 abajo) y con esto se deduzca que la opción más cercana es la opción b. No obstante, puede surgir la consideración de que, en principio, puede también ocurrir un resultado como el de la opción c (aunque es muy poco probable bajo el supuesto 0,68, 0,32). Percibir esta posibilidad puede llevar a algunos a elegir (equivocadamente) la opción d, como la respuesta correcta. En el estudio de Begué, et al. (2017) se formula una pregunta abierta y esto da pie a que muchos estudiantes se basen en la equiprobabilidad que asigna la misma probabilidad a cada una de las posibles caídas de una chincheta. En el caso de la pregunta 8 de nuestro cuestionario, se proponen los resultados de tres experimentos del lanzamiento de 100 chinchetas y una cuarta opción "d) Cualquiera de las anteriores". Por lo tanto, se puede interpretar que quién elige la opción d, cae en un sesgo de equiprobabilidad, pero no referente a las dos posibles caídas de una chincheta, sino al espacio de los posibles resultados del experimento de arrojar 100 chinchetas. Naturalmente, al elegir la opción d, los estudiantes no asignan conscientemente la misma probabilidad a todos los elementos de este espacio, pero la asignación de la misma probabilidad a las tres opciones dadas sería extrapolable a otros subconjuntos de resultados, lo cual, al final llevaría a la asignación de una distribución uniforme sobre el espacio de todos los resultados posibles de arrojar 100 chinchetas. El problema de fondo, como en los resultados de Begué et al., es que para quienes eligen esta opción la información de las frecuencias no resulta relevante.

La pregunta 9 evalúa la influencia de los resultados de 10 repeticiones del experimento sobre las probabilidades de base. Se esperaba que los estudiantes percibieran la independencia de las extracciones y respondieran basados en las probabilidades de base; pero en este caso, la respuesta es indeterminada "No afecta el resultado elegir una u otra urna". No obstante, la presencia de información de los resultados de 10 extracciones puede sugerir que estos datos deben ser tomados en cuenta e influir en desplazar la atención hacia las distribuciones de resultados olvidando las probabilidades de base. En este caso, se crea la ilusión de que con base en los datos pueden elegir la urna que favorece al evento.

La pregunta 10b implica la formulación de una hipótesis sobre el contenido de la urna. La solución óptima es 2 bolas blancas y 3 bolas negras pues en este caso la proporción correspondiente serían 40 bolas blancas y 60 negras. No obstante, algunas distribuciones alternativas podrían, en casos extremos, generar la distribución de datos dada; además, no son elegibles otras alternativas pues la probabilidad de generar tal distribución es ínfima. De cualquier manera, este problema tiene una pequeña dosis de incertidumbre.

*Bloque 3. Pregunta 4 (situaciones en contexto).* La pregunta 4 (con cinco opciones) plantea en cada opción una situación en contexto. En cada uno se define un evento y se pregunta si existe la probabilidad de ese evento. La razón por la que se formuló esta pregunta proviene de dos consideraciones. Una es que en la exposición del enfoque frecuencial que hizo von Mises (1957) se destacan dos propiedades de un colectivo, la existencia del límite de las frecuencias relativas y su unicidad, en el sentido de que las frecuencias relativas de cualquier sub–sucesión elegida por lugar, también converge al mismo límite. La segunda razón, proviene de que la propiedad de existencia de una probabilidad no parece intuitiva en situaciones en contextos diferentes a los juegos de azar, y menos cuando no se tiene la información para hacer algún procedimiento o cálculo para obtenerla. Hay contextos en torno los cuales se suele hablar de probabilidad, en el cuestionario incluimos los de encuestas, pronósticos del clima, número de accidentes, número de matrimonios, y un contexto poco común consistente en pensar un número, pero tiene la ventaja de que el experimento es susceptible de ser concebido fácilmente por los estudiantes. Conviene notar que en tales situaciones no es claro el modelo con el que se generan dichas probabilidades. Si se les cuestiona sobre un evento en estas situaciones, ¿asumen que existe una probabilidad del evento, aunque no sepan cómo se determinaría? La respuesta a esta pregunta puede ser útil cuando se planea enseñar la probabilidad desde un enfoque de modelación. Kahneman (2012, p. 26) señala que cuando a un sujeto se le hace una pregunta que no puede responder con sus recursos inmediatos a la mano, inconscientemente cambia la pregunta por otra más fácil, que sí puede responder. En este sentido, puede ser útil saber cómo interpretan (es decir, cambian) los estudiantes la pregunta.

Con relación a la validez del cuestionario podemos decir que la mayoría de los problemas u otros muy similares ya han sido utilizado en otros estudios (Begué, et al., 2017; Dantal,1998, Sánchez y Valdez 2014), mientras que las situaciones del Bloque 3 (pregunta 4) son propuestas nuevas de nuestra parte. Antes de aplicarlo no se sometió a ningún proceso de validación más allá de la discusión dada entre los autores. No obstante, cabe señalar que, en los contextos de elegir un estudiante del colegio, pronóstico de lluvia, accidentes y matrimonios es frecuente utilizar la probabilidad. El problema 4d, que propone la situación de "piensa un número" está inspirado en un problema que propone Mosteller (1965) y que fue utilizado por Sánchez y Hernández (2010) en una exploración con profesores.

### Procedimiento de análisis.

Las respuestas de los estudiantes a cada pregunta se compararon entre sí para determinar grupos de respuesta con características semejantes; esto permite reducir la gran variedad de respuestas a dos o tres variantes principales. Como consecuencia, se puede hacer una descripción comprensible de los conocimientos e intuiciones que los estudiantes ponen en juego al responder las preguntas.

## Resultados

A continuación, se ofrece una síntesis de las respuestas de los estudiantes a las preguntas del cuestionario. La caracterización de cada categoría de respuesta refleja lo que tienen en común, pero se debe entender que frecuentemente la manera en que se expresan los estudiantes es muy diversa y con expresiones idiosincráticas. La exposición se divide en tres bloques, cada uno corresponde a un subconjunto de preguntas del cuestionario. Se muestran ejemplos de respuestas de los estudiantes, que caracterizamos como E1, E2, etc.

### Bloque 1. Caracterización de conceptos básicos

En el primer bloque del cuestionario se busca explorar la manera en que los estudiantes entienden tres términos básicos de la teoría de la probabilidad: experiencia aleatoria, frecuencia relativa y probabilidad. Ninguno de los estudiantes participantes respondió a la primera pregunta con una caracterización adecuada de una experiencia aleatoria. La mayoría sólo menciona una de sus componentes pertinentes (espacio muestral, impredecibilidad, repetibilidad) y unos pocos ofrecen respuestas fuera de lugar. Más específicamente, nueve estudiantes (41 %) asocian una experiencia aleatoria sobre todo con el azar (cómo en el siguiente ejemplo E4), ocho (36 %) con la impredecibilidad, como ocurre con E1 y cinco (23%) mencionan que es un fenómeno que puede resultar en diferentes resultados (E6). Consideran la aleatoriedad y, de un modo precario, aluden al espacio muestral, pero la componente de repetibilidad necesaria para desarrollar un enfoque frecuencial está totalmente ausente en sus respuestas.

> Una situación en la que no se predice el resultado exacto de la situación (E1).
> Que se escoge o elige de manera al azar (E4).
> Una situación en que son probables diferentes resultados (E6).

En la segunda pregunta, ocho estudiantes (36 %) se acercan a la definición de frecuencia relativa (Por ejemplo, en la respuesta de E1 a continuación), ya sea como el cociente de la "frecuencia absoluta entre el número de datos" o como "porcentaje de la frecuencia absoluta". Dada la ausencia de la propiedad de repetibilidad de una experiencia aleatoria mostrada en las respuestas de la pregunta anterior, y dado que se refieren al cociente de una frecuencia relativa como "número de datos", no es claro si diferencian entre la n probabilidad clásica y la frecuencia relativa. Es decir, es posible que algunos de ellos se refieran con "frecuencia absoluta" a los "casos favorables" de un suceso que seguramente escucharon cuando llevaron el curso de probabilidad y estadística. De las respuestas restantes, en tres (14 %) se responde que la frecuencia absoluta (ver la respuesta de E7) es "la frecuencia del suceso" y en dos (9 %), reconocen que es un cociente de la frecuencia absoluta, pero sin mencionar el denominador (E17). En dos casos no se responde y las 6 respuestas restantes (27 %) ofrecen ideas erróneas al concepto (cómo E13). Nuevamente, un rasgo relevante es que, excepto por la respuesta de E13, ninguna respuesta alude a la repetibilidad de las experiencias.

> Es el resultado que se da entre la frecuencia absoluta de un determinado valor y el número total de datos (E1).
> Es el número de veces que puede ocurrir un evento (E7).
> Cociente de la frecuencia absoluta (E17).
> Una serie de hechos que se repiten constantemente (E13).

Las respuestas a la tercera pregunta muestran la gran dificultad de los estudiantes para asimilar una definición de probabilidad. En efecto, encontramos que la mayoría (73 %) de las respuestas (16) se pueden clasificar como circulares (cómo en E4), pues lo que entienden por probabilidad lo expresan en términos de "posibilidad", que no es más que un sinónimo. En tres casos (14 %) identifican la probabilidad sólo con lo aleatorio o azaroso sin agregar más características (ver E5). En dos respuestas (9 %) se menciona que la probabilidad es "el número de resultados de un evento" y en otra respuesta se dice que es "la frecuencia de posibilidades de un evento". Conviene remarcar que en estas respuestas está ausente el concepto de espacio muestral, fundamental para cualquier definición de probabilidad.

Que existe una posibilidad de que algo sea (E4).
Que un determinado evento se cumpla al azar (E5).

Más en general, el conjunto de respuestas a estas tres preguntas revela una ausencia de sentido teórico (es decir, ponen en juego sus conceptos espontáneos sin considerarlos conceptos científicos) de nuestros estudiantes que les impide valorar la importancia de asimilar las componentes de un modelo de probabilidad. Esto se deriva de dos vertientes; por un lado, la dificultad para acotar la experiencia aleatoria, de tal modo que sea un objeto observable, por otra, la connotación que atribuyen a las palabras azar o aleatorio, evento o probabilidad en el lenguaje común. En efecto, para ellos, una experiencia aleatoria es la que no se puede predecir, la que ocurre al azar o, en el mejor de los casos, algo que puede dar lugar a uno de varios resultados. El mejor acercamiento a la frecuencia relativa en sus respuestas se apoya en la noción de frecuencia absoluta y número de datos, que no conllevan la noción de repetición del experimento. Suelen caracterizar a la probabilidad de manera circular utilizando la palabra posibilidad para describirla; es remarcable que ninguno utilice las palabras "frecuencia relativa" o simplemente "frecuencia" a pesar de que la pregunta previa les permitió evocar dichas palabras.

**Bloque 2. Exploración de algunos conceptos del enfoque clásico y frecuencial**

La pregunta 5 es de comparación de probabilidades y la respuesta requería observar que la distribución de bolas blancas y negras de una urna es proporcional a la distribución de la otra. Trece respondieron (59 %) que se puede elegir cualquier urna señalando la proporcionalidad de las distribuciones (por ejemplo, E5). En ocho respuestas (36 %) se elige la urna B (como E3) porque en ella hay más negras (sesgo de la atención). Otra respuesta es indicar que ambas tienen la misma probabilidad porque “predominan las negras”.

> Es la misma probabilidad ya que en la urna B, están el doble de bolas de la urna A (E5)
> En la urna B porque hay un mayor número de bolas negras que en la urna A (E3).

La pregunta 6 pide elegir entre cuatro opciones, un posible evento resultante de 1000 experimentos de Bernoulli con $p = 0{,}4 \; y \; q = 0{,}6$. Siete estudiantes (32 %) respondieron la opción b, es decir, el resultado proporcional a las probabilidades; seis estudiantes (27 %) eligieron la opción c, resultado aproximado a la proporción, ocho (36 %) optaron por la opción d, que afirma que cualquier evento de los indicados en las opciones puede ocurrir y en un caso no hubo respuesta. Se debe destacar que la opción a no fue elegida en ningún caso. En este problema se han repartido las respuestas en tres opciones diferentes. Una posible explicación de esta diversidad es que los estudiantes perciben incertidumbre en la situación. Por ejemplo, “la opción b pudo haber ocurrido, pero también la opción c”; de ahí que a ocho estudiantes les pareciera más certero la opción c que incluye a los anteriores.

La pregunta 7 pedía calcular frecuencias relativas de una lista de resultados del experimento y 16 estudiantes (73 %) lo hicieron correctamente. Los estudiantes no perciben la incertidumbre en ambas preguntas, pues saben que sólo hay que hacer el cálculo correspondiente.

La pregunta 8 pide elegir entre tres opciones de resultados o una cuarta opción que abarca a las tres anteriores, la que se crea que ocurrió al lanzar 100 veces una chincheta, con base en el conocimiento de las frecuencias que ocurrieron en un lanzamiento previo (68 arriba y 32 abajo). En 21 respuestas (95 %) se eligió la opción c (cualquiera de las anteriores) y en un caso no hubo respuesta. Llama la atención que la opción c (63 arriba y 37 abajo), la cual es la más cercana a las frecuencias conocidas, no haya sido elegida por ningún estudiante. Es posible que la naturaleza del experimento en que no se conoce la probabilidad que permita suponer que cualquier cosa puede ocurrir y que el resultado conocido no informa sobre las probabilidades.

La pregunta 9 pide elegir entre dos urnas idénticas (1 bola negra y 1 bola blanca) la que ofrezca más chances de obtener bola negra, pero también se informa sobre dos secuencias de resultados de 10 extracciones en cada urna. Hubo sólo 7 respuestas (32 %) que dijeron que daba lo mismo elegir una u otra urna, pues son equivalentes. En cambio, diez (45 %) eligieron la urna B porque la muestra que de ella se sacó tiene más bolas negras (efecto de recencia positiva). En dos casos (9 %) se eligió la urna A, “porque se han extraído menos negras”, es decir, estos estudiantes muestran la falacia del jugador o recencia negativa. En dos (9 %) no hubo respuesta.

La pregunta 10a pedía calcular la frecuencia relativa a partir de frecuencias absolutas. Diecinueve estudiantes (86 %) la respondieron satisfactoriamente. La pregunta 10b pide formular una predicción sobre la distribución de bolas blancas y negras en una urna con 5 bolas, dado un resultado de 38 blancas y 62 negras en 100 extracciones con remplazo. La hipótesis más plausible es que la urna contenga 2 bolas blancas y tres negras. En 11 respuestas (50 %) se proponen esta distribución, en tres casos (14 %) se supone igual número de blancas que de negras, sin percibir que el enunciado que informa que la urna tiene 5 bolas. Hay cinco respuestas (23 %) variadas cuya razón de ser no es muy clara, algunas sólo invierten la distribución (3 blancas y 2 negras), en otras se da la frecuencia relativa (38% negras y 62% blancas) y otro estudiante interpreta que 5 es el número de negras y propone un número de blancas para compensar. Un estudiante no responde.

La pregunta 10c pide que “con base en lo anterior” se calcule las probabilidades de los eventos “obtener bola negra” y “obtener bola blanca”. Las respuestas revelaron que la pregunta es ambigua, pues hay dos informaciones “anteriores”, a saber, una es la información del enunciado “62 negras y 38 blancas” y la otra el contenido de la urna que ellos propusieron; esta es la información que se esperaba que utilizaran. Sólo en ocho respuestas (36 %) se hizo el cálculo que corresponde al contenido 2 bolas blancas y 3 negras, a pesar de que 11 (50 %) propusieron dicho contenido. Cuatro (18 %) asignan probabilidades con base en las frecuencias. Ocho (36 %) proponen respuestas no muy claras. Dos (9 %) no respondieron.

**Bloque 3. Situaciones en contexto**

La situación de la pregunta 4a consiste en elegir un estudiante del colegio, indicar si existe la probabilidad del evento “aprobar matemáticas” y explicar por qué. Once respuestas (50 %) se consideraron adecuadas (como la de E9); de ellas, cinco asimilan el problema a una situación de urnas, o muestreo, diciendo que si se elige un estudiante del colegio puede ocurrir que sea uno que haya acreditado matemáticas (por ejemplo, E5); seis (27 %) argumentan pensando en probabilidades individualizadas, en el sentido de que dependen de las características del estudiante elegido (ver E6). El resto responde introduciendo elementos no pertinentes. La respuesta que afirma que no existe la probabilidad da un argumento inconsistente, quizá se basa simplemente en su percepción de que puede ocurrir el evento contrario (respuesta de E20).

> Si se trata de una probabilidad ya que entre todos los estudiantes que hay se elige solamente uno (E9)
> Puesto que se elige a un estudiante al azar, hay varias probabilidades de que haya o no reprobado la asignatura (E5).
> Porque tanto que acredite o no una materia son resultados probables en un estudiante (E6)
> Muchos estudiantes acreditan esa materia, se tendría que realizar más cosas para decidir quién se elegirá (E20)

La situación 4b consiste en preguntar a un estudiante que piense un número y lo diga, el evento es que el número pensado sea un número entero en el rango entre 3 y 10, preguntando si existe la probabilidad del evento Nueve estudiantes (41 %) ofrecen una respuesta sencilla como que “puede haber alguien que piense un número entre 3 y 10” o “los números de 3 a 10 son elegibles”.De ellas, hay dos respuestas más elaboradas que dan razones que se refieren a la tendencia de las personas a pensar dígitos enteros (ver E6). Hay seis respuestas (27 %) que sugierenque no existe la probabilidad y su argumento se basa en la gran amplitud de posibilidades para elegir un número (cómo E3). Estos estudiantes piensan intuitivamente en un espacio muestral infinito y en un modelo laplaciano.

> El estudiante puede decir cualquier número incluyendo del 3-10 que sean enteros (E6).
> No existe probabilidad ya que hay demasiados números y no solo entre el 3 y 10 (E3).

La pregunta 4c trata de la situación de elegir un día del año por venir y decir si existe la probabilidad de que ese día llueva. Dieciséis estudiantes (73 %) responden que sí existe la probabilidad del evento, pero sólo diez ofrecen la explicación simple de que “en cualquier día puede llover” (cómo E6). Los seis restantes añaden elementos que desvirtúan la respuesta (E10). Las cuatro respuestas (18%) que indican que no existe la probabilidad del evento, argumentan en términos de la escasez de lluvias, la variabilidad del tiempo o que no se puede predecir (E18).

> Si, porque existe la probabilidad de que llueva cualquier día del año (E6).
> Si, porque es un día cualquiera y el clima depende de la naturaliza (E10).
> No, puede variar el resultado y es más por el azar (E18).

La situación 4d consiste en observar un crucero de dos calles durante un día y decir si existe la probabilidad de un accidente. Dieciséis respuestas (73 %) responden que sí existe la probabilidad y de estas sólo 12 se basan en la idea de que "siempre es posible que ocurra un accidente" (E2); las cuatro restantes introducen elementos extraños como que el evento es impredecible, que puede ocurrir cada hora, que depende de la circulación de autos (E10). Cuatro estudiantes (18 %) expresan que no entienden la pregunta. Uno dice que no existe la probabilidad y argumenta que "Es difícil saber si habrá o no un accidente".

> Si, porque existe una posibilidad de que pase o no (E2).
> Si, porque pasan muchos autos y es probable que en diferentes variantes ocurra alguna situación (E10).

La situación 4e consiste en contar el número de matrimonios de un año en una ciudad y considerar el evento de que haya entre 1000 y 2000 matrimonios. En once respuestas (50%) dicen que sí existe la probabilidad del evento (E3) y en nueve de ellas explican que en una ciudad grande se puede alcanzar ese número de matrimonios (entre 1000 y 2000); los dos restantes introducen elementos extraños. Siete afirman (32 %) que no existe la probabilidad del evento, sus argumentos consisten en que el evento no se puede predecir o que se necesita más información (E19); en dos casos (9 %) critican el enunciado del problema. Cuatro (18 %) dicen que no se entiende la pregunta.

> Sí hay probabilidad porque hay muchos casamientos en un registro civil (E3)
> Sí, porque el mismo evento te da la probabilidad que hay de los matrimonios de la ciudad (E9).
> No se puede observar porque los eventos no han ocurrido (E19).

**Discusión y conclusiones**

El estudio realizado se orientó a realizar un diagnóstico de las intuiciones y conocimientos de un grupo de estudiantes de bachillerato con el propósito de diseñar un experimento de enseñanza enfocado al enfoque frecuencial de probabilidad. Del análisis de las respuestas de los estudiantes hemos encontrado tres aspectos para reforzar este diseño, a saber, incluir actividades de:

- Conceptualización, para enriquecer el vocabulario probabilístico;
- Articulación entre el enfoque clásico y frecuencial;
- Vinculación del enfoque frecuencial con contextos diferentes a juegos de azar.

Los resultados muestran que los conocimientos de los estudiantes de la muestrasobre los términos básicos de la probabilidad, la relación entre los enfoques clásico y frecuencial y la manera en que estos funcionan en contextos diferentes a los de juegos de azar, son precarios. Sin embargo, las situaciones de urnas les proporcionan un lugar de experimentación donde pueden operar intuitivamente para calcular probabilidades y frecuencias relativas e incluso percibir algunas relaciones pertinentes entre el enfoque clásico y el frecuencial. No obstante, estas intuiciones no se expresan en sus argumentaciones respecto a la existencia de la probabilidad en situaciones en contextos diferentes de juegos de azar. En lo que sigue resumimos algunas de los aspectos que nos permiten realizar estas afirmaciones.

**Tres conceptos básicos de probabilidad**

Los estudiantes examinados están lejos de haber formado los tres conceptos básicos de la probabilidad que hemos examinado. Nuestros estudiantes muestran prenociones (yuxtaposición de ideas, sincretismo o preconceptos) cuando tratan de explicar lo que entienden por 'experiencia aleatoria', 'frecuencia relativa' y 'probabilidad'. Utilizan frecuentemente el término evento para referirse a una experiencia aleatoria y esta solo la asocian con el azar, la impredecibilidad y a veces con la posibilidad de que ocurran diferentes resultados.

En el informe de Dantal (1998) la respuesta dominante a la pregunta ¿Qué es una experiencia aleatoria? fue "es una experiencia en la que no se puede predecir el resultado" (p. 67) con una frecuencia de 33% en promedio. Un porcentaje similar de las respuestas de nuestros sujetos destacan el rasgo de impredecibilidad, que se presentó en 8 respuestas (de 22). Un rasgo poco visible, pero probablemente más generalizado, es que en algunas respuestas se detecta la confusión entre frecuencias absolutas y casos favorables a un evento y no distinguen entre el enfoque frecuencial y el enfoque clásico de la probabilidad. Dantal encuentra un rasgo similar pues dice que "Para ellos [los estudiantes] solo existe un enfoque de probabilidades, ya sea frecuencial o clásico, aunque a veces en las fórmulas que dan encontramos una mezcla de los dos" (p. 69).

Por otro lado, a diferencia del estudio de Dantal en que la mayoría de las respuestas a la pregunta ¿Qué es la probabilidad de un evento de una experiencia aleatoria? se refieren a un procedimiento o una fórmula para calcularla, en este estudio, ninguno de nuestros sujetos respondió así; la mayoría sólo ofrece una respuesta circular utilizando el sinónimo “posibilidad”, que es más coloquial que el término probabilidad. En conclusión, es importante considerar en el experimento que se planea llevar a cabo incluir actividades dirigidas a proporcionar y ampliar el vocabulario de nuestros estudiantes, de manera tal que inicien un proceso de conceptualización de los modelos clásico y frecuencial de probabilidad que vaya más allá del dominio operacional.

**Situaciones de urnas y chinchetas**

En las situaciones de urnas realizan procedimientos en los que interpretan o aplican sus prenociones de ‘experiencia aleatoria’, ‘frecuencia relativa’ y ‘probabilidad’ como si contaran con los conceptos correspondientes. Es decir, calculan la probabilidad de la extracción de una bola de color de una urna, comparan las probabilidades de un mismo evento, pero proveniente de dos urnas con diferente contenido, calculan frecuencias relativas a partir de las frecuencias absolutas y relacionan la probabilidad clásica con las frecuencias de un evento en un gran número de repeticiones. Aunque en las prenociones que manifiestan no incluyen una que corresponda al término ‘repetibilidad’ de una experiencia aleatoria, en la práctica, entienden la idea de extracciones sucesivas con remplazo de bolas en una urna.

Las situaciones de urnas que dan lugar a dificultades son aquellas en las que perciben cierta dosis de incertidumbre y, estas son precisamente las que requieren llevar a cabo una relación entre la probabilidad clásica y las frecuencias relativas. En las respuestas al problema 9 del cuestionario vemos que muchos estudiantes son proclives a creer que pequeñas diferencias son significativas (Shaughnessy, 1992), pues consideran que la información de los resultados de 10 repeticiones determina una tendencia del evento “sale negra” aún frente al conocimiento de que la urna tiene una bola blanca y una negra.

En cambio, en el problema 10 aparentemente tienen una actitud contraria, pues la distribución de la caída de 100 chinchetas “68 blancas y 32 negras” no los lleva a inferir una tendencia y prefieren decir que las tres opciones propuestas son posibles. Al igual que en el estudio de Begué et al. (2017), para muchos estudiantes la información frecuencial no es relevante a la hora de estimar probabilidades y, a falta de más información, recurren a la noción intuitiva de equiprobabilidad. En consecuencia, para un futuro experimento de diseño conviene poner atención en que los estudiantes valoren de manera adecuada la información que proporcionan las frecuencias, superando la falsa concepción de que las frecuencias de muestras pequeñas son informativas para estimar la probabilidad y la tendencia contraria de ignorar la información frecuencial de grandes muestras.

**Contextos**

Cuando se les pregunta si existe la probabilidad de un evento de una experiencia aleatoria en contexto, en general, no tienen dificultad para responder positivamente, pero no son proclives a llevar la situación contextualizada a un modelo clásico o frecuencial de probabilidad. En efecto, no utilizan términos como "espacio muestral", "casos favorables", "casos posibles" del modelo clásico o "repetición de experimentos", "frecuencias absolutas o relativas" del modelo frecuencial.

A excepción del experimento de la selección al azar de un alumno del colegio, que una tercera parte asimila a un modelo de urnas, en las demás preguntas no mencionan ningún término que revele que piensan en algún modelo, ya sea un modelo pseudo–concreto de urnas, ya el enfoque clásico o el frecuencial. Las explicaciones que hemos considerado aceptables en las demás situaciones se reducen a imaginar una realización del evento en la vida cotidiana: "alguien puede pensar en un número entre 3 y 10" "en cualquier día hay posibilidad de que llueva", "cuando menos te lo esperas ocurre un accidente".

Sus razones se basan en su sentido común, en el sentido de que "siempre puede pasar algo". En la situación del evento "un alumno aprueba matemáticas" poco más de la mitad de las respuestas aceptables ofrecen indicios de que lo ven como un modelo de urna "si se elija un estudiante al azar del colegio, puede haber acreditado matemáticas", la otra mitad piensa a la probabilidad como una propiedad personal: "quien estudia tiene probabilidad de acreditar".

Respecto a la validez de la pregunta 4 (contextos), los resultados muestran que los estudiantes no responden de acuerdo con la intención original de la pregunta, aunque sus respuestas revelan aspectos útiles para entender la manera en que razonan. El objetivo era explorar si en cada uno de los contextos dados, los estudiantes creen o no que existe un número que representa la probabilidad del evento dado, pero de la manera que formulamos la pregunta, los estudiantes la interpretaron como si es preguntase si puede ocurrir el evento. De esta manera, no se problematiza sobre la existencia de un número que represente la probabilidad del evento. Como consecuencia para el experimento de diseño, consideramos que una pregunta (o una serie de preguntas) que atienda la misma intención, debe formularse de una manera más apropiada. Queda abierta entonces la pregunta de si los estudiantes conciben que siempre existe un número para representar la probabilidad de eventos dados en contextos diferentes a los juegos de azar, aun cuando no conozcan un método preciso para calcularla.

## Agradecimientos

Este trabajo ha sido financiado por Proyecto-Conacyt 254301 y Proyecto SEP-CINVESTAV 188

## Referencias

Amir, G. S. & Williams, J. S. (1999). Cultural influences on children's probabilistic thinking. *Journal of Mathematical Behavior, 18*(1), 85-107. https://doi.org/10.1016/S0732-3123(99)00018-8.

Aspinwall, L. & Tarr, J. E. (2001). Middle school students' understanding of the role sample size plays in empirical probability. *Journal of Mathematical Behavior, 20*, 229–245. https://doi.org/10.1016/S0732-3123(01)00066-9.

Ayres, P., & Way, J. (2000). Knowing the sample space or not: The effects on decision making. In T. Nakahara & M. Koyama (Eds), *Proceedings of the 24th Conference of the International Group for the Psychology of Mathematics Education* (Vol. 2, pp. 33-40). Hiroshima: PME Group.

Batanero, C. & Serrano, L. (1999). The meaning of randomness for secondary students. *Journal for Research in Mathematics Education, 30,* 558-567. https://doi.org/10.2307/749774.

Batanero, C., Navarro-Pelayo, V., & Godino, J. D. (1997). Effect of the implicit combinatorial model on combinatorial reasoning in secondary school pupils. *Educational Studies in Mathematics, 32*, 181-199. https://doi.org/10.1023/A:1002954428327.

Batanero, C., Henry, M., & Parzysz, B. (2005). The nature of chance and probability. En G. A. Jones (Ed.), *Exploring probability in school. challenges for teaching and learning.* New York: Springer.

Begué, N., Batanero, C., Gea, M.M. y Beltrán-Pellicer, P. (2017). Comprensión del enfoque frecuencial de la probabilidad por estudiantes de Educación Secundaria Obligatoria. En J.M. Muñoz-Escolano, A. Arnal-Bailera, P. Beltrán-Pellicer, M.L. Callejo y J. Carrillo (Eds.), *Investigación en Educación Matemática XXI (pp. 137-146)*. Zaragoza: SEIEM

Birks, M. y Mills, J. (2015). *Grounded theory. A practical guide.* London: SAGE.

Chaput, B., Girard, J.-C, & Henry, M. (2011). Frequentistic approach: modelling and simulation instatistics and probability teaching. En C. Batanero, G. Burril, & C. Reading (Eds.), *Teaching statistics in school mathematics – challenges for teaching and teacher education* (pp. 85-96). New York: Springer

Dantal, B. (1998). Comment les élèves de terminale perçoivent les concepts d'expérience aléatoire, d'événement et de probabilité. En *Enseigner les probabilités au lycée. Ouvertures statistiques, enjeux épistémologiques, questions didactiques et idées d'activités* (pp. 67-69). Grénoble: Commission inter-IREM Statistiques et Probabilités.

Denzin, N. K. & Lincoln, Y. S. (2011). Introduction. The discipline and practice of qualitative research. En N. K. Denzin & Y. S. Lincoln (Eds.), *The Sage handbook of qualitative research* (pp. 1-19). London: SAGE.

Fischbein, E. (1987). *Intuition in science and mathematics.* Dodrecht: Reidel.

Fischbein, E. & Grossman, A. (1997). Schemata and intuitions in combinatorial reasoning. *Educational Studies in Mathematics, 34*, 27-47. https://doi.org/10.1023/A:1002914202652.

Fischbein, E., Nello, M. S., & Marino, M. S. (1991). Factors affecting probabilistic judgements in children in adolescence. *Educational Studies in Mathematics, 22*, 523-549. https://doi.org/10.1007/BF00312714.

Fischbein, E. & Schnarch, D. (1997). The evolution with age of probabilistic, intuitively based misconceptions. *Journal for Research in Mathematics Education, 28*, 96-105. https://doi.org/10.5951/jresematheduc.28.1.0096.

Flyvbjerg, B. (2011). Case study. En N. K. Denzin & Y. S. Lincoln (Eds.), *The Sage handbook of qualitative research* (pp. 301-316). London: SAGE.

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). Nueva York: Springer.

Green, D. R. (1983). A survey of probability concepts in 3000 pupils aged 11-16 years. In D. R. Grey, P. Holmes, V. Barnett, & G. M. Constable (Eds.), *Proceedings of the First International Conference on Teaching Statistics* (pp. 766-783). Sheffield, England: Teaching Statistics Trust.

Green, D. R. (1988). Children's understanding of randomness: Report of a survey of 1600 children aged 7-11 years. En R. Davidson & J. Swift (Eds.), *Proceedings of the Second International Conference on Teaching Statistics* (pp. 287-291). Victoria, BC, Canada: University of Victoria.

Ireland, S. & Watson, J. (2009). Building a connection between experimental and theoretical aspects of probability. *International Electronic Journal of Mathematics Education, 4*(3), 339-370.

Jones, G. A., Langall, C. W., & Mooney, E. S. (2007). Research on probability. Responding to classroom realities. En F. K. Lester Jr. (Ed.). *Second handbook of research on mathematics teaching and learning.* Charlotte, NC: National Council of Teachers of Mathematics & Information Age Publishing.

Kahneman, D. (2012). *Pensar rápido, Pensar despacio.* Villatuerta, España: Debate.

Kahneman, D., & Tversky, A. (1982). Variants of uncertainty. En D. Kahneman, P. Slovic, & A. Tversky (Eds.). *Judgment under uncertainty: Heuristics and biases (509-520)*. Cambridge: Cambridge University

Lee, H. S., Angotti, R. L., & Tarr, J. E. (2009). Making comparisons between observed data and expected outcomes: students' informal hypothesis testing with probability simulation tools. *Statistics Education Research Journal, 9*(1), 68-96

Miles, M. & Huberman, A. (1994). *An expanded sourcebook qualitative data analysis* (2a ed.). Londres: Sage Publications.

Mosteller, F. (1956). *Fifty Challenging Problems in Probability with Solutions.* New York: Dover.

Nilsson, P. (2014). Experimentation in probability teaching and learning. En Chernoff E. J., Sriraman, B. (Eds.). *Probabilistic thinking. Presenting plural perspective (509-532)*. New York: Springer.

Pfannkuch, M. & Ziedins, I. (2014). A modelling perspective on probability. En E. J. Chernoff, & B. Sriraman (Eds.). *Probabilistic thinking. Presenting plural perspective* (pp. 101-116). New York: Springer.

Pfannkuch, M., Budgett, S., Fewster, R., Fitch, M., Pattenwise, S., Wild, C., & Ziedins, I. (2016). Probability modeling and thinking: what can we learn from practice? *Statistics Education Research Journal, 15*(2), 11-37.

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

Sánchez, E., Hernández, A. (2010). A statistical game: the silent cooperation problem. En C. Reading (Ed.), *Data and context in statistics education: Towards an evidence-based society. Proceedings of the Eighth International Conference on Teaching Statistics* (ICOTS8, July, 2010), Ljubljana, Slovenia.

Sánchez, E., & Valdez, J. (2014). Reasoning development of a high school student about probability concept. En K. Makar, B. de Sousa, & R. Gould (Eds.), *Sustainability in statistics education. Proceedings of the Ninth International Conference on Teaching Statistics.* Flagstaff, AR: International Statistical Institute.

SEP (2017). *Aprendizajes clave para la educación integral. Matemáticas, educación secundaria. Plan y programas de estudio.* Ciudad de México: Secretaría de Educación Pública.

SEP (2011). *Plan de estudio 2011. Educación básica*. Ciudad de México: Secretaría de Educación Pública

Shaughnessy J. M. (1992). Research in probability and statistics: Reflections and Directions. En D. A. Grouws (Ed.), *Handbook of research on mathematics teaching and learning* (pp. 465-494). New York: Macmillan.

Shaughnessy, J. M., & Ciancetta, M. (2002). Students' understanding of variability in a probability environment. In B. Philips (Ed.), *Proceeding of the Sixth International Conference on the Teaching of Statistics* [CD-ROM]. Cape Town, South Africa: International Statistical Institute.

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

Stohl, H., Rider, R., & Tarr, J. (2004). *Making connections between empirical and theoretical probability: Students' generation and analysis of data in a technology environment*. http://www.probexplorer.com/Articles/LeeRiderTarrConnectE&T.pdf.

Tarr, J. E. (2002). The confounding effects of "50-50 chance" in making conditional probability judgments. *Focus on Learning Problems in Mathematics, 24*, 35-53.

Von Mises, R. (1957). *Probability, statistics and truth*. New York: Dover Publications, Inc.

Watson, J. M. (2005). The probabilistic reasoning of middle school students. In G. A. Jones (Ed.), *Exploring probability in school: Challenges for teaching and learning (pp. 145-168).* New York: Springer.

Watson, J. M., & Moritz, J. B. (2003). Fairness of dice: A longitudinal study of students' beliefs and strategies for making judgments. *Journal for Research in Mathematics Education, 34*, 270-304. https://doi.org/10.2307/30034785.

## Apéndice. Cuestionario

1. ¿Qué entiendes por una Experiencia aleatoria?
2. ¿Qué entiendes por una Frecuencia relativa?
3. ¿Qué entiendes por Probabilidad?
4. En los siguientes cuadros se describe una experiencia aleatoria y a su derecha se define un evento. Indica paa cada evento si tiene probabilidad y explica tu respuesta.

| Experiencia aleatoria | Evento |
|---|---|
| a. Se va a elegir un(a) estudiante del Colegio | E1: Acreditó matemáticas |
| b. Se le pregunta a un estudiante que piense un número y lo diga. | E2: Responde un número entero entre 3 y 10 |
| c. Se elige un día cualquiera del año por venir. | E3: Llueve ese día |
| d. Se va a observar un crucero de dos calles transitadas durante un día entero | E4: Hay un accidente en el crucero |
| e. Se van a observar los informes del registro civil del año venidero de una ciudad; en particular, el número de matrimonios. | E5: Hay entre 1000 y 2000 matrimonios. |

5. En una urna A hay tres bolas negras y una bola blanca. En la urna B hay 6 bolas negras y 2 bolas blancas. Se saca una bola al azar de cada urna. ¿En qué urna se tiene mayor probabilidad de obtener bola negra? Explica tu respuesta
6. Una urna tiene dos bolas blancas y tres bolas negras. Un experimento consiste en extraer al azar una bola, observar su color y anotarlo. Un equipo hizo 1000 experimentos. En las siguientes opciones se presentan tres posibles frecuencias de bolas blancas y negras. ¿Cuál de ellas crees que sea el resultado que obtuvo el equipo? Explica tu respuesta
   a. Frecuencia de blancas = 611, Frecuencia de negras = 389
   b. Frecuencia de blancas = 400, Frecuencia de negras = 600
   c. Frecuencia de blancas = 385, Frecuencia de negras = 615
   d. Pudieron haber obtenido cualquiera de las anteriores

7. De una urna con bolas blancas y negras se saca una bola al azar, se nota el resultado y se devuelve la bola a la urna. Se llevaron a cabo 50 experimentos y se obtuvieron los siguientes resultados. Calcula las frecuencias relativas de cada resultado $f(B)\ y\ f(N)$, exprésalas en decimales.

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| B | N | N | N | B | B | B | N | N | N |
| N | B | B | N | N | N | N | B | B | N |
| N | N | N | N | N | B | B | N | N | B |
| N | N | N | B | N | N | N | N | N | B |
| N | N | B | B | N | N | N | N | B | N |

8. Un experimento aleatorio consiste en arrojar 100 chinchetas y observar cuántas caen con la punta para arriba (ARRIBA) y cuántas caen apoyadas tanto en la cabeza como en la punta (ABAJO). En una realización del experimento, se arrojaron las 100 chinchetas y se obtuvieron 68 para arriba y 32 para abajo. Después se volvió a realizar el experimento y se anotaron las frecuencias de cada resultado ¿Cuál de las siguientes opciones crees que sean las frecuencias de esta segunda experiencia? Marca la opción de tu elección.
   a. ARRIBA = 36, ABAJO = 64,
   b. ARRIBA = 63, ABAJO = 37,
   c. ARRIBA = 51, ABAJO = 49,
   d. Cualquiera de las anteriores.

9. Dos urnas A y B, tienen cada una 1 bola blanca y una bola negra. Se realizaron diez extracciones de una bola con remplazo de cada urna y se anotaron los resultados en la tabla siguiente:

| # extracciones | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|
| Urna A | B | B | N | B | B | B | N | B | B | N |
| Urna B | N | N | B | N | N | B | B | B | N | N |

   Si vas a hacer una onceava extracción y quieres obtener bola negra (N) ¿cuál urna elegirías? Justifica tu respuesta.

10. Una urna tiene 5 bolas algunas blancas y otras negras. Se hicieron 100 extracciones con remplazo de la urna y se obtuvieron las siguientes frecuencias absolutas: Bolas blancas: 38; Bolas negras: 62
   a. Calcula la frecuencia relativa (en decimales) de cada resultado: f(B), f(N).
   b. ¿Cuántas bolas negras y cuántas blancas crees que tiene la urna?
   c. Con base en lo anterior, calcula la probabilidad de obtener bola blanca y la probabilidad de obtener bola negra: $P(blanca) =;\quad P(negra) =$
   d. ¿Qué puedes concluir de lo anterior?

## Autores

**Sandra Areli Martínez-Pérez**
Estudiante de doctorado (CINVESTAV, IPN). Licenciatura en Física y Matemáticas (ESFM, IPN) Maestrría en Ciencias, con especialidad en Matemática Educativa (CINVESTAV, IPN).
Línea de Investigación: Didáctica de la Probabilidad y la Estadística.
E-mail: sandra.martinezperez@cinvestav.mx

**Ernesto A. Sánchez Sánchez**
Profesor Titular del Departamento de Matemática Educativa, Centro de Investigación y de Estudios Avanzados del Instituto Politécnico Nacional. Licenciatura en Matemáticas (UNAM), Maestría y Doctorado en Ciencias con especialidad en Matemática Educativa (CINVESTAV, IPN). Miembro del Sistema Nacional de Investigadores, Nivel II, México. Miembro de la Academia Mexicana de Ciencias. Línea de Investigación: Didáctica de la probabilidad y la estadística. Email: esanchez@cinvestav.mx.