Hacia una posible aproximación comprensiva de la variable aleatoria

Felipe **Fernández**
Luisa **Andrade**
Javier **Montañez**
Juan Pablo **Beltrán**
Sandra **Zamora**
Universidad Pedagógica Nacional
Colombia
fjfernandez@pedagogica.edu.co

Resumen

De los contenidos que se estudian en los cursos de probabilidad, la variable aleatoria resalta, no sólo por el poco sentido que tiene para los estudiantes, sino también por la ausencia de significación que al respecto los docentes muestran. En un proyecto de investigación dirigido a desarrollar razonamiento estadístico, se implementa una secuencia de tareas enfocada a la construcción de significado en torno a la variable aleatoria. Los resultados posibilitan ver dificultades en torno a los conceptos relacionados que coinciden con dificultades reportadas en la literatura y alertan sobre consideraciones a tener en cuenta para el diseño curricular y la instrucción.
Palabras clave: educación estadística, variable aleatoria, experimentos de enseñanza

En una serie de proyectos de investigación[1] realizados con el propósito de explorar el desarrollo de razonamiento estadístico en estudiantes de colegio y universidad, emergen dificultades notorias relacionadas con contenidos de la probabilidad. En consecuencia, una indagación para profundizar sobre la didáctica en cuestión, revela que de manera generalizada la enseñanza enfocada de la variable aleatoria se reduce a su presentación a través de una definición donde su significado es tácito, y cuya meta primordial se enfoca en el uso instrumental de su distribución de probabilidades, más que en los antecedentes que la modelan y configuran. Es claro entonces que tanto profesores como estudiantes no le dan un sentido a la variable aleatoria ni entienden el por qué de su existencia. Así, esta comprensión limitada de la variable aleatoria se constituye en uno de los obstáculos que no favorece los propósitos que abogan por la mejora de una alfabetización estadística encaminada al desarrollo de razonamiento y pensamiento estadístico desde una perspectiva variacional.

La complejidad que subyace a la conceptualización de la variable aleatoria ha sido explicitada en trabajos como los de Ruiz (2006) y Ruiz, Albert y Batanero (2006), en donde se da cuenta tanto de aspectos disciplinares e históricos de corte epistemológico relevantes para su comprensión, como del estudio de dificultades de los estudiantes cuando se enfrentan a situaciones y problemas que involucran su aplicación.

---

[1] Estos proyectos se han realizado desde el año 2006 por la línea de investigación en Educación Estadística del “Grupo de Didáctica de la Matemática” de la Universidad Pedagógica Nacional de Colombia, el tercero de los cuales se encuentra actualmente en ejecución. Han sido financiados por el Centro de Investigaciones (CIUP) y el Departamento de Matemáticas de la misma universidad.

Desde el campo disciplinar se reconoce el papel dual que juega la variable aleatoria. Por un lado, sobre la variable aleatoria la atención se centra casi exclusivamente en sus valores como base para cuantificar las probabilidades de los eventos posibles del experimento aleatorio, es decir, en la función de distribución de probabilidad que toma como dominio los valores de la variable aleatoria. De otra parte, la variable aleatoria asume un papel de tipo funcional al relacionar eventos asociados a un experimento aleatorio con valores numéricos; parafraseando a Ruiz (2006), éste sería el estrato donde la variable aleatoria actúa como 'modelo matemático' pues vincula un fenómeno aleatorio con una estructura matemática; se define así como función, en donde el dominio es el espacio muestral del experimento y el recorrido es un conjunto numérico. También cabe en esta perspectiva disciplinar advertir las dificultades que conlleva la transición, o la comparación, del carácter discreto al carácter continuo de una variable aleatoria. En el caso discreto es claro que la finitud o la infinitud enumerable de los valores que asume la variable aleatoria permite una asociación más transparente entre los elementos del espacio muestral (que no necesariamente es un conjunto de números) con valores numéricos. Para lidiar con algunas de las complicaciones que surgen al considerar la variable aleatoria continua, en el terreno formal y riguroso de las matemáticas se imponen condiciones a la variable aleatoria, como la de satisfacer que la imagen inversa de un intervalo acotado en la variable sea un suceso medible. Esto parece conducir, a que el espacio muestral (el dominio) de una variable aleatoria continua sea, a la fuerza, un conjunto numérico. Al respecto, un ejemplo que fue objeto de discusión al interior del grupo de investigación, es el experimento aleatorio en donde se seleccionan bombillas para medir su tiempo de vida útil, medida que corresponde a una variable aleatoria continua. Las posibles respuestas a la pregunta "¿cuál es el espacio muestral o el dominio de la variable aleatoria asociada a este experimento aleatorio?", como "las bombillas seleccionadas" o "los intervalos de tiempo de duración de cada una de las bombillas seleccionadas", permiten vislumbrar algunos de las complicaciones subyacentes a la variable aleatoria continua. En realidad, la secuencia de tareas que sirvió de instrumento para obtener los resultados reportados en este artículo sólo se limitó a la variable aleatoria discreta.

En una revisión histórica, Ruiz (2006) explicita varios momentos históricos en la conceptualización de la variable aleatoria que la llevan a identificar diferentes saltos cualitativos en su comprensión. Así, se puede observar como la variable aleatoria en principio se trabajó de manera intuitiva y solo con el pasar de los años, fue adquiriendo importancia su estudio y su respectiva axiomatización por parte de la comunidad científica (ver Kolmogorov, 1956; Levy, 1936; Petrov, 2002; Parzen, 1957; citados en Ruiz, 2006), entre otros. En particular, esa autora destaca dos paradigmas que pueden distinguirse en la evolución histórica del concepto coincidentes con el doble papel mencionado anteriormente: el de "magnitudes aleatorias" que restringe y concibe a la variable aleatoria como un fenómeno aleatorio de resultados numéricos y el de "variables aleatorias" en el que emerge el carácter funcional de la variable aleatoria, y sugiere que la subsistencia actual de estos dos paradigmas aporta elementos para explicar y poner de manifiesto, algunas de las dificultades que deben enfrentar estudiantes y profesores al abordar la enseñanza y aprendizaje de la variable aleatoria.

Tratamiento de la variable aleatoria en algunos textos de estadística y probabilidad

De manera similar a lo ocurrido en el desarrollo histórico de la variable aleatoria, su tratamiento en la enseñanza y el aprendizaje usualmente ocurre de forma tácita, tanto que, muchas veces, ni siquiera los estudiantes se percatan de su existencia.

La revisión de algunos textos universitarios de probabilidad y estadística posibilita catalogarlos en dos clases: los que presentan de manera explícita el concepto de variable aleatoria y los que lo presentan de manera implícita, que según Ortiz (2002), son la mayoría. Cuando se hace referencia expresa a la variable aleatoria, los textos le dedican por lo menos un capítulo al desarrollo de este concepto (ver Walpole, 1992 y Parzen, 1960), pero no contemplan todos los aspectos que le pueden dar sentido y significado, ya que se trabaja la variable aleatoria sin importar como se obtuvieron los valores de ésta, y se aborda algorítmicamente en ejercicios y teoremas sin utilizarla como modelo matemático. Miller (1998, citado en Ruiz, 2006) afirma que en los libros de introducción a la estadística casi no se desarrolla la idea de variable aleatoria y que cuando se hace, frecuentemente es de forma errónea lo que acarrea problemas en temas posteriores. En los libros de texto del otro tipo, se trabaja el concepto sin nombrar la variable aleatoria ni hacer referencia a ella. Por ejemplo en el texto de bioestadística de Schefler (1981), la variable aleatoria ni siquiera es mencionada dentro de los temas propuestos y se propone el estudio de los conceptos de probabilidad sin definirla.

En los textos universitarios que analizaron Oseguera (1994, citado en Ruiz, 2006) y Tauber (2001, citado en Ruiz, 2006) encontraron que no siempre hay una conexión entre el estudio del modelo probabilístico y el análisis de los datos empíricos, pues se olvida el contexto y la forma en que se asignan los valores a las variables aleatorias.

Aproximación metodológica

En este proyecto, como en el anterior, se trabaja con la metodología de los experimentos de enseñanza mediante la cual se implementa una secuencia de tareas elaborada para construir un sentido de la variable aleatoria, que se implementa con estudiantes universitarios de un curso de estadística y probabilidad durante cinco sesiones de clase, donde el profesor del curso es miembro del equipo de investigación.

El experimento de enseñanza que se plantea tiene el fin de analizar el aprendizaje matemático de los estudiantes de forma cercana y constante a medida que ocurre en la situación social de los grupos o de toda la clase, siguiendo la propuesta de Cobb (1999, citado en Jones et al., 2001); igualmente tiene el propósito de mejorar el diseño de la secuencia de tareas al poner a prueba tal diseño y modificar en la clase las conjeturas sobre el aprendizaje de los estudiantes tal como lo sugieren diversos autores (Brown, 1992, Cobb, 2001, Collins, 1999, Suter y Frechtling, 2000, citados en Cobb, McClain y Gravemeijer, 2003). En este proyecto las conjeturas son las predicciones que se han hecho sobre el camino que debe seguir el aprendizaje de los estudiantes, es decir están constituidas por la trayectoria hipotética de aprendizaje que se ha formulado en el sentido propuesto por Simon (1995).[2]

La construcción de las tareas propuestas implicó un proceso de reflexión y análisis antes y durante el desarrollo de la implementación curricular, en donde se discutían las intencionalidades de las tareas planteadas y se conjeturaba acerca de las dificultades que podían emerger en los

[2] Para Simon (1995) la formulación de una trayectoria hipotética de aprendizaje resalta la importancia de tener metas y fundamentos claros para las decisiones del profesor, y se basa en los supuestos de que en el aprendizaje de un individuo puede verse cierta regularidad, la actividad matemática en la clase con frecuencia ocurre de formas previsibles, y en una misma clase varios estudiantes pueden beneficiarse de la misma tarea matemática.

estudiantes. Las tareas siguientes eran entonces modificadas o ajustadas de acuerdo a los resultados encontrados, lo mismo que las conjeturas hechas de acuerdo a las recomendaciones de Confrey (2006, citado en Molina, 2006) y Steffe y Thompson (2000). Adicional a este análisis permanente que se llevó a cabo sobre el aprendizaje de los estudiantes mientras el experimento de enseñanza progresaba, se realizó un análisis retrospectivo sobre lo sucedido de forma global.

La observación del trabajo de los estudiantes durante el transcurso de las clases estuvo a cargo tanto del profesor del curso como de otros tres miembros del equipo de investigación, quienes además de tomar notas de campo grabaron en audio la interacción en varios de los grupos de estudiantes, con el ánimo de describir con precisión la evolución del aprendizaje de los estudiantes en concordancia con las recomendaciones de Confrey y Lachance (2000).

Trayectoria hipotética

Dentro de los aspectos que se consideran como medio para motivar la construcción del concepto de variable aleatoria están: la explicitación de una partición en el espacio muestral de un experimento aleatorio; la relación de los elementos de la partición con los valores de una variable de la situación que es la indicada para resolver la pregunta central; la caracterización de esa relación; la determinación, en distintas situaciones, de la presencia de la variable aleatoria; la definición de una función de probabilidad asociada a la variable aleatoria propuesta; y la utilización de la función de probabilidad de la variable aleatoria como herramienta de análisis para tomar decisiones ante situaciones de incertidumbre.

Poner en evidencia de manera más puntual el carácter funcional de la variable aleatoria, fue uno de las intenciones perseguidas en la secuencia de tareas implementada en el aula de clase. Resultados preliminares de experimentos de enseñanza puestos a prueba y relacionados con la variable aleatoria, mostraron la dificultad de superar la enseñaza enfocada en el paradigma de las magnitudes aleatorias. El reto de poner en práctica una instrucción que haga más patente el paradigma de 'variables aleatorias' implica en consecuencia, la necesidad de proponer situaciones de aprendizaje en donde la definición de la variable aleatoria no sea arbitraria, sino que responda al contexto de forma que su establecimiento se sugiera de manera más natural; en palabras de Ruiz (2006), situaciones en donde la variable aleatoria actúe como 'modelo matemático' en su carácter funcional. La secuencia de tareas propuesta busca, a través de una situación contextualizada en la vida diaria pero ficticia, propiciar una conceptualización de variable aleatoria. La situación se construyó para que involucrara una variable aleatoria que emergiera de manera natural y ayudara a solucionar o responder la pregunta central, a lo cual contribuyeron en gran medida las ideas sobre la situación planteada por Ruiz (2006) en su tesis maestría.

El trazado de la ruta hipotética de aprendizaje se basa en el planteamiento de una situación ficticia en donde el estudiante debe asumir el papel de asesor administrador de los gastos asociados a la realización de una rifa para los empleados de un centro comercial. Uno de los empleados es el elegido en la rifa y es el ganador de un viaje, con todo su grupo familiar, a alguna reconocida ciudad europea de atracción turística. La secuencia de tareas está dividida en seis partes (ver apéndice), a través de las cuales se realiza el proceso de instrucción, que puede llevarse a cabo en unas cuatro o cinco sesiones de clase de noventa minutos.

La primera parte pretende generar la necesidad intuitiva de definir relaciones funcionales entre pares de conjuntos en donde uno de los conjuntos a identificar sea el espacio muestral correspondiente al experimento aleatorio que subyace al contexto sugerido y el otro, un conjunto de valores numéricos correspondiente al número de personas del grupo familiar del empleado del centro comercial, del que depende el número de tiquetes que se requeriría comprar. Además, también se propone un caso en donde es posible definir una relación funcional entre el espacio muestral y un conjunto no numérico (el lugar de destino preferencial del viaje) que no conduce a una definición de variable aleatoria. Al abordar estas primeras tareas, se espera que los estudiantes por un lado, logren claridad respecto a cuál es el experimento aleatorio implicado en la situación, así como cuáles y cuántos son los resultados del experimento que conforman los elementos del espacio muestral asociado; y por otro lado, que identifiquen el número de personas del núcleo familiar y el lugar de destino preferido, como las dos variables más relevantes en la toma de la decisión implicada.

Para la segunda parte de la secuencia de tareas, se presenta el ejemplo hipotético de una base de datos con información sobre las variables consideradas. Se espera que este ejemplo concreto facilite la tarea de precisar las particiones del espacio muestral que se pretende que generen los estudiantes o que, en su defecto, ayude a replantear en una dirección más atinada las particiones que primero se hubiesen sugerido.

Las tareas de la tercera parte retoman la idea de generar particiones del espacio muestral de acuerdo a las variables consideradas; en particular, los primeros cuestionamientos proveen un espacio para la verificación de la concordancia de las repuestas previas, o para la creación de un conflicto con otras maneras de pensar acerca de la partición del espacio muestral, que proporcionen el camino para identificar y representar el dominio y los recorridos de las dos relaciones funcionales en juego. Como uno de estos recorridos es de tipo numérico, se espera que el camino para definir la variable aleatoria quede así preparado y motivado para su definición. La socialización planteada para finalizar esta tercera parte debe abrir el espacio para la formalización de la definición de la variable aleatoria.

En la cuarta parte de este proceso se abre un lugar para la identificación de variables aleatorias al presentar una colección de cinco casos de variación estocástica. En dos de los casos presentados (el primero y el cuarto) se explicita una valoración cuantitativa del fenómeno estocástico en cuestión; por ello, se espera que estos dos casos sean reconocidos como variables aleatorias. En otros dos casos (el tercero y el quinto) se presentan experimentos aleatorios que no tienen una valoración cuantitativa explícita; así, aunque se espera que en estos dos casos no se mencione la existencia de la variable aleatoria, si que se reconozca la posibilidad de enunciar asignaciones numéricas, que hicieran concretable una variable aleatoria; por ejemplo, en el tercer caso, se podría sugerir la asignación del número uno al artículo defectuoso y del cero al artículo no defectuoso.

La quinta parte de la secuencia de tareas, traslada la discusión acerca de la variable aleatoria a la estimación y definición de su función de probabilidad y a la complejidad relacional subyacente a los conjuntos implicados. En otras palabras, se abre espacio para la discusión de las relaciones funcionales que se constituyen entre el espacio muestral, el recorrido de la variable aleatoria y el intervalo [0; 1] de los números reales. Hay que advertir que en tanto objeto matemático, y en un sentido estricto, la función de probabilidad es una relación cuyo dominio es el recorrido de la variable aleatoria, con codominio el intervalo [0; 1].

Para terminar, en la sexta parte de las tareas propuestas se considera el asunto de la recomendación acerca del número de tiquetes que se debería comprar. Se espera que la secuencia de tareas allí propuesta genere discusión en torno a la aplicabilidad del cálculo de probabilidades para el contexto propuesto. Al respecto se pueden reconocer al menos dos posibles aproximaciones argumentativas: por un lado, están los que simplemente identifican el número de integrantes del núcleo familiar con mayor probabilidad y recomiendan, en consecuencia, la compra de exactamente ese número de tiquetes; y por otro lado, están los que desde una mirada más amplia, estiman el valor esperado y la cantidad de probabilidad que se acumula alrededor de este valor; para los que asumen esta última perspectiva, se reconoce que hay un mayor posicionamiento en su análisis, en cuanto al sentido de lo variacional en una distribución de probabilidad.

Resultados

En los resultados de la implementación de la secuencia de tareas se evidencia la utilización intuitiva e implícita de la variable aleatoria sin haber sido definida y como factor definitivo para la toma de decisiones que da respuesta a la pregunta de la situación, en la misma dirección de lo encontrado por Ruiz (2006). No obstante, también sobresalen algunas dificultades presentadas por la mayoría de los estudiantes, que sin referirse directamente a la variable aleatoria como objeto matemático, si apuntan a conceptos necesarios para otorgarle sentido. La primera de ellas tiene que ver con la identificación del experimento aleatorio, ya que su reconocimiento no parece tan evidente (la elección de un empleado para adjudicarle un premio); una razón para esto es la falta de claridad con respecto a lo que se considera un fenómeno aleatorio. Así por ejemplo, los estudiantes señalan que el experimento consiste en "determinar cuántos tiquetes de viaje hay que comprar", "determinar cuáles son las posibilidades de escoger al empleado, la cantidad de familiares y la ciudad que desea conocer".

Aún después de establecer el experimento aleatorio surgen igualmente dificultades para explicitar los posibles resultados del experimento, es decir, el espacio muestral que en esta situación está constituido por los 120 empleados participantes en la rifa; quizás por motivo de que el enunciado de la situación presenta información acerca de varios hechos que se incluyeron al armarla y porque en la pregunta central están presentes las variables relevantes a la situación, los estudiantes se confunden y tienden a conformar el espacio muestral como duplas de combinaciones de los valores de dos variables; por ejemplo hay respuestas como "{*1L, 2L, 3L, … 9L, 1R, 2R, 3R, … 9R, 1P, 2P, 3P, … 9P*}", donde los números corresponden al número de integrantes de la familia del empleado y la letra es la inicial de la ciudad a visitar (Londres, Roma o París).

En la situación propuesta la variable 'número de integrantes de la familia' coincide con una variable aleatoria, es decir, es una variable que se reconoce como relevante, se trabaja pero sin ubicarla como variable aleatoria y se incorpora en el espacio muestral mismo. De alguna manera el espacio muestral y la variable aleatoria se identifican, hecho que es reportado por autores como Ruiz (2006) y Ruiz, Albert y Batanero (2006). Resalta también la necesidad de los estudiantes de introducir una cota superior en los valores de la variable aleatoria, en este caso el 'número de integrantes de la familia', que en la respuesta exhibida se toma como máximo *9* y en otras respuestas como *22*.

Se perciben asimismo dificultades al trabajar las particiones del espacio muestral, no necesariamente en la construcción de la partición misma que es una tarea de fácil realización en las matemáticas escolares, sino en la comprensión de que la variable aleatoria induce una partición y que los valores de la variable aleatoria que se consideran de manera natural, están ligados a los elementos de esa partición. Aunque en este momento de la secuencia de tareas no se ha abordado la variable aleatoria con este nombre ni como una relación funcional, la intención si es que los estudiantes pudieran establecer la relación entre los subconjuntos de una partición del espacio muestral con los valores de la variable que trabajan, es decir la dificultad encontrada concuerda con la reportada por Ruiz (2006, p. 44) cuando señala que hay dificultad para aceptar que el dominio de la variable aleatoria surge de la consideración de una partición del espacio muestral, en donde cada subconjunto de la partición inducida determina un valor de la variable aleatoria.

Es claro que el uso del verbo 'agrupar' en las instrucciones de las tareas no ayudó, pues llevó a la idea de agrupación en intervalos de clase en el sentido en el que se suele trabajar en la estadística descriptiva para resumir los datos, y por consiguiente los estudiantes esbozaron tablas de frecuencias para la variable. Dado que la agrupación de los puntos muestrales en una partición era el paso principal para la conceptualización de la variable aleatoria, se evidencia que la idea de ésta en los estudiantes sigue vinculada principalmente a su empleo para el establecimiento de la distribución de probabilidades. En la representación de la partición emergen de igual forma dificultades, pues lo estudiantes escogen los diagramas sagitales y el plano cartesiano como representantes de las agrupaciones, cuando lo que ilustran estas gráficos son la correspondencia y si acaso la relación de dependencia entre magnitudes o variables.

En el análisis de las cinco situaciones para determinar posibles variables aleatorias, una vez completadas las partes anteriores de la secuencia de tareas y por lo tanto habiendo discutido en el grupo la caracterización de la variable aleatoria, tal y como se esperaba, los estudiantes establecieron fácilmente la variable aleatoria en los casos de lanzamiento de monedas y dados; en los casos de tomar un artículo para determinar si es o no defectuoso y seleccionar al azar un estudiante para determinar la localidad donde vive, algunos estudiantes concluyeron que no era posible establecer una valoración cuantitativa de manera explícita, pero que sí era posible definir una variable aleatoria, si se asignaban números. No obstante, en el caso del generador aleatorio de una letra, previsto como el más problemático, los estudiantes no lograron establecer la variable aleatoria. Por un lado, en el enunciado de la situación hay un elemento numérico en el enunciado, "determinar la frecuencia de aparición de las letras", que sugiere la posibilidad de considerar una variable aleatoria; pero por otro lado, la identificación del espacio muestral que es simplemente el conjunto $\{a, b, c, d, e\}$ genera dificultades y los estudiantes tienden a tomarlo de la forma "n veces" el producto cartesiano del conjunto anterior. Además aun en la posibilidad de pensar el espacio muestral como un producto cartesiano, hay dificultades al definir la variable aleatoria sugerida por el enunciado pues tiene más sentido establecerla como el conteo de una letra específica (por ejemplo la *a*).

Conclusiones

Por un lado, las acciones realizadas para implementar el experimento de enseñanza presentado, así como los resultados encontrados, confirman la presencia de varias dificultades reportadas previamente en la literatura relacionada con la comprensión de la variable aleatoria. Por otro lado, los resultados del trabajo con la variable aleatoria alertan sobre diferentes aspectos a tener en cuenta para superar dificultades encontradas: la identificación y distinción del experimento aleatorio y del espacio muestral asociado; la asignación de probabilidades a los valores de la variable aleatoria, en términos del cardinal de los subconjuntos que conforman la partición del espacio muestral inducida, asignación clásica de probabilidades posible solamente en espacios muestrales finitos y equiprobables; el uso y aplicación de la distribución de probabilidades de la variable aleatoria en situaciones de relevancia fenomenológica para los estudiantes.

En primer lugar, una tarea que a raíz de los resultados obtenidos se posiciona como primordial para el trabajo de los estudiantes con la variable aleatoria, se refiere a crear espacios donde el estudiante pueda identificar cuál de todos los fenómenos incluidos en un enunciado corresponde al experimento aleatorio, en qué consiste dicho experimento aleatorio, por qué es aleatorio, y cuáles pueden ser posibles resultados de ese experimento. En esta última labor se destaca la necesidad de que una vez determinado el experimento aleatorio se reconozcan los resultados que son realmente consonantes con éste. Es importante entonces proponer variadas situaciones con este fin, que difieran de las usuales en la enseñanza para obligar al estudiante a buscar respuesta a las preguntas mencionadas, y le permitan dilucidar de entre todos los hechos expresados en un enunciado, los relevantes para el trabajo.

En segundo lugar, al proponer situaciones en que se quiera prestar atención a la conceptualización de la variable aleatoria desde una perspectiva que enfatice en su carácter de relación funcional entre un espacio muestral y un conjunto numérico –en donde se aborden, entre otros aspectos, la explicitación de las particiones del espacio muestral inducidas por los valores que toma la variable aleatoria– es pertinente tener presente el significado que se atribuya a la idea de 'agrupar'. En particular, el experimento de enseñanza realizado permitió identificar que el trabajo previo de los estudiantes en construir tablas de frecuencias como parte del análisis de datos en estadística descriptiva, genera obstáculos para que los estudiantes agrupen los elementos del espacio muestral en subconjuntos excluyentes y exhaustivos para generar una partición, pues lleva automáticamente a realizar un conteo de los elementos a los que se les puede asignar un mismo valor de la variable aleatoria, es decir, el cardinal de cada subconjunto de la partición. Aún después de concretar la partición por sugerencias del maestro, persiste la idea de que la relación que se busca establecer se da es términos de los valores de la variable aleatoria con los cardinales de los subconjuntos correspondientes de la partición. En consecuencia, una recomendación al proponer tareas de instrucción, es centrar la atención en formular a los estudiantes preguntas que les permitan trabajar con particiones de conjuntos, reconocer que los valores de una variable aleatoria específica inducen una de tales particiones, y que el conjunto de los subconjuntos de la partición, o sea la partición misma, puede considerarse como el dominio de una relación, independientemente de que sea numérico o no. Igualmente, tratar la palabra 'agrupar' y sus sinónimos, en diversas circunstancias y situaciones estadísticas y probabilísticas de forma que su empleo no se restringa al trabajo con tablas de frecuencia.

Por último, al considerar la aproximación a la variable aleatoria desde una perspectiva conceptual amplia, en donde se considera tanto la definición o construcción de la variable aleatoria en su sentido relacional, como de su función de probabilidad, debe emerger el proceso de la asignación de probabilidades desde el punto de vista matemático que se manifiesta en la asignación de probabilidades como una función compuesta por la variable aleatoria en su perspectiva funcional y su correspondiente función de distribución de probabilidades. Esta mirada contribuye a reconocer el por qué de los valores de la variable aleatoria, a adjudicarle una utilidad como función y por lo tanto, a darle sentido, y a determinar más fácilmente las probabilidades asociadas. Además, cabe advertir que la elección de un contexto fenomenológico significativo para los estudiantes al proponer experimentos de enseñanza en torno a la variable aleatoria, se posiciona como un requisito necesario pero no suficiente para concretar una secuencia de instrucción que dé cuenta de estos procesos diferenciales.

Referencias

Cobb, P., McClain, K. y Gravemeijer, K. (2003). Learning about statistical covariation. *Cognition and Instruction, 21* (1), 1–78.

Confrey, J. y Lachance, A. (2000). Transformative teaching experiments through conjecture driven research design. En A. Kelly y R. Lesh (Eds.), *Handbook of research design in mathematics and science education* (pp. 231–265). Mahwah, NJ: Erlbaum.

Jones, G., Thornton, C., Langrall, C., Mooney, E., Wares, A., Jones, M.R., Perry, B., Putt. I.J. y Nisbet, S. (2001). Using students' statistical thinking to inform instruction. *Journal of Mathematical Behavior*, *20*, 109–144.

Molina, M. (2006). Desarrollo de pensamiento relacional y comprensión del signo igual por alumnos de tercero de educación primaria (pp. 261-290). (Tesis doctoral). Departamento de Didáctica de la Matemática, Universidad de Granada, España. (en línea http://0-hera.ugr.es.adrastea.ugr.es/tesisugr/16546167.pdf).

Ortiz, J. (2002) La probabilidad en los libros de texto. (Tesis doctoral). España: Grupo de Educación Estadística de la Universidad de Granada.

Parzen, E. (1960). Modern probability theory and its applications. New York: Wiley.

Ruiz, B. (2006). Un acercamiento cognitivo y epistemológico a la didáctica del concepto de variable aleatoria (Tesis de maestría). México: Instituto Politécnico Nacional.

Ruiz, B., Albert, J. y Batanero, C. (2006). An exploratory study of students' difficulties with random variables. México: ICOTS 7.

Simon, M. (1995). Reconstructing mathematics pedagogy from a constructivist perspective. *Journal for Research in Mathematics Education*, *2* (26), 114-145.

Schefler, W. (1981). *Bioestadística.* México. Fondo Educativo Interamericano.

Steffe, L.P. y Thompson, P.W. (2000). Teaching experiments methodology: Underlying principles and essential elements. En A. Kelly y R. Lesh (Eds.), *Handbook of research design in mathematics and science education* (pp. 267–306). Mahwah, NJ: Erlbaum.

Walpole, R. (1992). *Probabilidad y estadística aplicadas a la ingeniería*. McGraw-Hill

Apéndice. Rifa del Centro Comercial "Centro Zeta"

Parte 1. Selección de un empleado

Para motivar a los trabajadores del centro comercial "Centro Zeta" el departamento de recursos humanos va a seleccionar a uno de sus 120 empleados al azar para adjudicarle un premio.

1. Determine y describa en qué consiste el experimento aleatorio de la situación.
2. Describa los posibles resultados del experimento aleatorio, es decir, explicite cuál es el espacio muestral.

Para participar, es necesario que el empleado registre cuántas personas componen su núcleo familiar y la ciudad que preferiría visitar (Londres; Roma o Paris). La selección pretende premiar a un empleado, y a los integrantes de su núcleo familiar (si no vive sólo) con tiquetes de viaje y gastos de hotel. Antes de realizar la selección del empleado, los organizadores deben tomar una decisión sobre la compra de los tiquetes, el hospedaje y el destino, ya que si compran más tiquetes y reservas de lo necesario, y el núcleo familiar del empleado es pequeño, perderán dinero; y viceversa, si el núcleo familiar del empleado seleccionado está compuesta por muchas personas y no compran suficientes tiquetes, los organizadores no cumplirán con lo ofrecido.

3. Uno de sus compañeros, sugiere que los posibles resultados del experimento son dos: ganar o perder; otro compañero, alega que los resultados son todos los posibles empleados; y uno tercero, afirma que sólo hay un resultado: el ganador. Asuma una posición respecto a la validez de estas afirmaciones y compárela con su respuesta al ítem anterior. Explique.
4. Dada la respuesta anterior, precise cuáles y cuántos son los posibles resultados del experimento, es decir, describa en definitiva cómo es el espacio muestral.
5. Identifique las variables con base en las cuales los organizadores deben tomar la decisión.
6. ¿Cómo se imagina que es la base de datos en donde se registra la información de los empleados y qué tipo de datos contendría un registro típico de ésta? Especifique el tipo de elementos de las filas y columnas de esta base de datos.
7. Con base en la matriz anterior indique para cada una de las variables consideradas, cómo se pueden organizar en subconjuntos del espacio muestral los diferentes resultados del experimento.
8. En relación con los conjuntos anteriores, y antes de tomar una decisión sobre la compra de tiquetes y reservas de hotel, es importante para los organizadores determinar la probabilidad de ciertos eventos. ¿Cuáles podrían ser las probabilidades que se deberían conocer de antemano?
9. ¿Cómo se pueden calcular estas probabilidades?

Parte 2. Análisis de la base de datos de los empleados participantes en la rifa

Datos de los empleados participantes en la rifa

Nº	Nombre del empleado	Personas del núcleo familiar	Lugar de preferencia	Nº	Nombre del empleado	Personas del núcleo familiar	Lugar de preferencia
1	E1	0	Roma	61	E61	2	Roma
2	E2	1	Paris	62	E62	3	Roma
3	E3	9	Roma	63	E63	3	Roma
4	E4	2	Londres	64	E64	2	Paris
5	…	…	…	…	…	…	…
…	…	…	…	…	…	…	…
58	E58	5	Paris	118	E118	4	Londres
59	E59	4	Londres	119	E119	7	Londres
60	E60	4	Paris	120	E120	4	Londres

La matriz de la base de datos entregada presenta los datos recogidos acerca de los 120 empleados. Con base en esta información responda lo siguiente.

1. Determine si las variables que consideró en la Parte 1 como relevantes para tomar la decisión, están incluidas en la matriz. Explique.
2. Especifique si los resultados que sugirió en la Parte 1 como elementos del espacio muestral coinciden con los presentados en la matriz. Explique.
3. Establezca si es necesario replantear la organización de subconjuntos que sugirió en la Parte 1. Explique. (Sugerencia: recuerde que los conjuntos sugeridos deben ser subconjuntos del espacio muestral).
4. Calcule las probabilidades que determinó que se deberían conocer antes de tomar la decisión y explique si proveen información pertinente para tomar la decisión.

Parte 3. Relaciones entre variables

Considere el espacio muestral conformado por los 120 elementos discutidos en la parte anterior.

1. Haga una partición del espacio muestral que refleje la organización de subconjuntos propuestos en la parte anterior. Nota: recuerde que una partición del espacio muestral es una clasificación de sus elementos en subconjuntos cuya unión es el mismo espacio muestral y cuyas intersecciones entre pares de subconjuntos es vacía.
2. Ahora, agrupe los elementos de acuerdo al lugar de destino preferido por cada empleado.
3. Represente las clasificaciones realizadas de dos maneras diferentes. Note que lo que se está pidiendo no es elaborar tablas de frecuencias sino esquemas gráficos o tabulares de las relaciones generadas por las clasificaciones.
4. Para el caso de cada agrupación o partición (la del ítem 1 y luego la del ítem 2), discuta si en ellas es posible reconocer relaciones funcionales entre un dominio específico y un codominio. Argumente su respuesta describiendo en palabras las relaciones.
5. Identifique cuáles de estas relaciones tienen codominio numérico.

Parte 4. Situaciones sobre variables aleatorias

1. En las siguientes situaciones establezca cuales pueden ser variables aleatorias. Descríbalas en palabras y argumente su respuesta.
   - Lanzar tres monedas de manera consecutiva para determinar el número de caras.
   - Activar varias veces un generador aleatorio de una letra, tomada del conjunto {a, b, c, d, e}, para determinar la frecuencia de aparición de las letras.
   - Tomar un artículo de una fábrica para determinar si está o no defectuoso.
   - Lanzar un par de dados para determinar la suma de los resultados.
   - Seleccionar al azar un estudiante de primer semestre de matemáticas para determinar la localidad de Bogotá en donde vive.

Parte 5. Estimación de probabilidades

A partir de la manipulación de la base de datos de los registros de los 120 empleados participantes en la rifa se pueden elaborar las siguientes tablas:

Tabla 1. *Conjuntos de empleados de acuerdo al número de hijos*

Conjuntos de empleados	E1, E24, E28, ... E8, E83, E90	E2, E43, E50, E55, ... E109, E112, E116	E4, E21, E34, E38, E42, E46, ... E102, E107, E111	E5, E6, E7, E15, E16, E17, E18, E19, E23, ... E77, E78, E81, E84, E86, E91, E92	E9, E20, E35, E36, E40, E44 ... E97, E115, E117, E118, E120	E10, E37, E39, E58, ... E108, E110, E114	E11, E45, ... E103, E113	E12, E48, ... , E119	E13, E14, E82, E89	E3, E22
Números de hijos	0	1	2	3	4	5	6	7	8	9
Cantidad	9	13	20	28	20	12	7	5	4	2

Tabla 2. *Conjuntos de empleados, preferencias y frecuencia.*

Conjuntos de empleados	E10, E102, E108, E111, E112, E14, E16, E19, E23, ... E80, E81, E84, E85, E86, E89, E92, E96, E97, E98	E104, E106, E107, E11, E114, E115, E116, E118, ... E79, E88, E9, E90, E91, E95, E99	E1, E100, E101, E103, E105, E109, E110, E113, ... E68, E72, E75, E78, E82, E83, E87, E93, E94
Preferencia	Paris	Londres	Roma
Cantidad	44	36	40

1. Si se elige un empleado al azar, ¿cuál es la probabilidad de que tenga 5 hijos? ¿cuál es la probabilidad de que no tenga hijos? Elabore una tabla para presentar este tipo de probabilidades.
2. En la pregunta 1) precise qué depende de qué: ¿el número de hijos depende de la probabilidad, de los empleados, de la cantidad, o es al revés? En otras palabras, ¿qué relaciones hay entre las variables implicadas en la asignación de probabilidades?
3. Si se elige a un empleado al azar, ¿cuál es la probabilidad de que prefiera visitar Paris y cuál la de que prefiera Londres? Elabore una tabla para presentar este tipo de probabilidades.
4. En la pregunta 3) ¿qué depende de qué: la preferencia depende de la probabilidad, de los empleados, de la cantidad? Es decir, ¿cuáles son las relaciones que hay entre las variables involucradas en la asignación de probabilidades?
5. Determine si las relaciones que se pidió identificar en los puntos 2 y 4 son funcionales y especifique cuáles son los dominios y recorridos de cada caso.

Parte 6. Recomendación

1. Ahora, con base en la información de las Tablas de probabilidad elaboradas en la parte anterior, ¿qué recomendación le daría a los organizadores de la rifa respecto al número de tiquetes a comprar, y el lugar de destino? Explique detalladamente porqué le daría esa recomendación.
2. De acuerdo a su recomendación ¿con qué probabilidad el empleado premiado demandaría exactamente el número de tiquetes que compró la organizadora del evento? En otras palabras, ¿qué tanto puedes asegurar que se demandarán exactamente los boletos que se compraron?
3. De acuerdo a tu recomendación, ¿con qué probabilidad al empleado premiado le alcanzarán los tiquetes y hasta le sobrarán? ¿Con qué probabilidad el comprador premiado no le alcanzarán los boletos para poder llevar a todos sus hijos de viaje?