rev|Psi

Artículo de investigación

Comparación de la dificultad entre un problema de estimación y otro de probabilidad condicional

Horacio Félix Attorresi[1*], Alcira Myriam García Díaz[1] y Héctor Omar Pralong[1]

[1]Facultad de Psicología, Universidad de Buenos Aires (UBA, Argentina)

*Correspondencia: horacioattorresi@gmail.com

Recibido: 10 jun. 2022 | *1ra decisión*: 4 feb. 2023 | *Aceptado*: 19 may. 2023 | *Publicado*: 17 jun. 2023

Revista de Psicología, 23(2), 3–19 | 2024 | ISSN 2422-572X
https://doi.org/10.24215/2422572Xe167

https://revistas.unlp.edu.ar/revpsi
Revista de Psicología es una publicación de la Facultad de Psicología (Universidad Nacional de La Plata, Argentina)

Resumen

Se evaluó la dificultad para resolver problemas de asignación de probabilidad en dos situaciones de aleatoriedad de diferentes características. En la primera, se compararon proporciones de dos muestras de diferente tamaño y se estimaron probabilidades, mientras que en la segunda se presentó un problema de probabilidad condicional. Se categorizaron las respuestas analizando la comprensión cognitiva alcanzada y la existencia de sesgos tal como el heurístico de representatividad. Se encontró para la primera situación un 28% de resoluciones exitosas y un 42% de respuestas en las que se identifica dicho heurístico y la ausencia de resoluciones correctas para el cálculo de probabilidad condicional. Se seleccionó una muestra por accesibilidad y se administraron los reactivos a dos grupos de estudiantes ingresantes a la universidad, sin conocimientos previos acerca de las teorías de la probabilidad, de manera de analizar la capacidad intuitiva para la resolución de problemas de probabilidad.

Palabras clave: probabilidad, estimación, intuición, condicional.

Comparação da dificuldade entre um problema de estimação e outro de probabilidade condicional

Resumo: Avaliou- se a dificuldade para resolver problemas de designação de probabilidade em duas situações de aleatoriedade de diferentes características. Na primeira, compararam-se proporções de duas amostras de diferente tamanho e estimaram-se probabilidades sendo que na segunda apresentou-se um problema de probabilidade condicional. Categorizaram-se as respostas analisando a compreensão cognitiva alcançada e a existência de sesgos assim como o heurístico de representatividade. Encontrou-se para a primeira situação 28% de resoluções bem-sucedidas e 42% de respostas nas quais se identifica tal heurístico e a ausência de resoluções corretas para o cálculo de probabilidade condicional. Selecionou-se uma amostra por acessibilidade e administraram-se os reativos em dois grupos de estudantes calouros na universidade, sem conhecimentos prévios sobre as teorias da probabilidade, de maneira de analisar a capacidade intuitiva para a resolução de problemas de probabilidade.

Palavras-chave: probabilidade, estimação, intuição, condicional.

Comparing difficulty between a problem of estimation and one of conditional probability

Abstract: This work evaluated the difficulty posed by probability problems in two randomness scenarios with different characteristics. In the first one, proportions were compared and probability was estimated in two samples of different sizes, while the second one presented a conditional probability problem. The responses were categorized on the basis of cognitive understanding and potential bias, such as the representativeness heuristic. Results showed 28% of successful responses and 42% of responses involving the heuristic in the first problem, and no correct responses to the conditional probability problem. The sample was selected on accessibility grounds, and the items were administered to two groups of students recently admitted to university with no previous knowledge of probability theories in order to analyze their intuitive ability to tackle probability problems.

Keywords: probability, estimation, intuition, conditional.

Aspectos destacados del trabajo

- Elección de los problemas: uno de estimación y otro de cálculo de probabilidad condicional.
- Comparación de las dificultades entre los dos problemas planteados.
- Elección por accesibilidad de estudiantes ingresantes a la universidad sin formación previa en probabilidades.
- Hubo respuestas correctas en el problema de estimación y ninguna en el de probabilidad condicional.

Se reconoce la importancia de enseñar temas de estadística desde niveles de escolaridad cada vez más tempranos. La transferencia de estos conocimientos desde el nivel terciario y universitario hacia la escuela media se viene llevando a cabo en las últimas décadas.

La vía de entrada del aprendizaje de la estadística ha sido por supuesto la matemática. Sin embargo, la especificidad de la primera amplía con experiencias vinculadas al azar muchos conceptos deterministas de la segunda. Así como lo manifiesta Santaló (1970) nociones tales como las de gráficos, cambios de escala, valores medios, interpolaciones, extrapolaciones, correlaciones, etc., que forman parte del lenguaje de la estadística lo son también de otras ramas de la matemática. Del mismo modo el método de Monte Carlo, la simulación o el empleo de números aleatorios, son conceptos útiles para resolver problemas en estadística. Estos conceptos, así como el de probabilidad que fundamentan distintas áreas de la estadística pueden ser tratados adecuando su enseñanza a los niveles medio e incluso primario. La noción de probabilidad se alcanza, según Piaget e Inhelder (1974) mediante experiencias y situaciones vinculadas a la génesis de las operaciones combinatorias, a través de juegos como los dados, las cartas y la ruleta, entre otros, en los que se iría conformando experimentalmente la idea que subyace a la ley de los grandes números.

La instrucción sobre los heurísticos y sesgos en el razonamiento probabilístico (Kahneman y Tversky, 1972) analizados en estudiantes de magisterio no alcanza para eliminar nociones erróneas (Díaz, 2003). En el ámbito educativo, se observa una insuficiente confrontación de las nociones intuitivas que las personas adquieren en su vida cotidiana (Konold, 1989) lo cual puede llevar a la transmisión de errores, con los perjuicios que esta situación puede producir en la enseñanza (Alvarado et al., 2018).

Las metodologías que utilizan los profesores en la enseñanza de la probabilidad condicional en el nivel medio y preuniversitario se vinculan con actividades que implican más procesos algorítmicos que tareas cognitivas (Batanero et al., 2015). Incluso Elbehary (2020) plantea que éstos poseen niveles bajos de conocimientos didácticos para la enseñanza de temas probabilísticos. Por lo que resulta importante trabajar con contextos de la vida cotidiana para la enseñanza de la probabilidad.

Revista de Psicología, 23(2), 3–19 | 2024 | ISSN 2422-572X
https://doi.org/10.24215/2422572Xe167

Es importante ver cómo inciden las creencias y aprendizajes previos de los estudiantes universitarios en sus estudios sobre temas vinculados al azar y la probabilidad, así como, analizar los criterios que utilizan para decidir sobre la verosimilitud de sucesos (Barragués et al., 2005). Guisasola y Barragués (2002) señalan la existencia de un sesgo determinista que asocia causas y efectos a eventos aleatorios. Cardeñoso (2001) plantea que la enseñanza puede reafirmar estas estrategias y por lo tanto no descartar los sesgos que se producen. Además, el lenguaje y el contexto con el cual se presentan los reactivos y ejemplos, pueden convertirse en dificultades para la construcción del conocimiento (Attorresi et al., 2005, 2007; Serradó et al., 2005).

En el ámbito de la psicología ha sido estudiada la dificultad que presenta llevar a cabo razonamientos probabilísticos adecuados en situaciones de la vida cotidiana. Los trabajos de Kahneman et al. (1982) dan cuenta de estas dificultades Así mismo, Attorresi et al. (2009, 2015) han encontrado problemas en el razonamiento inferencial en adultos, así como en conceptos tales como el de aleatorización, correlación y probabilidad en general.

A medida que nos adentramos en el tratamiento de temas específicos del área de las probabilidades tales como el de probabilidad condicional, las experiencias no resultan mejores. En trabajos de investigación sobre la didáctica de la probabilidad condicional se pueden observar diversas dificultades conceptuales de los sujetos y se sugiere acompañar los problemas de juegos de azar con otros que abarquen circunstancias de incertidumbre que conlleven consecuencias perjudiciales en la vida cotidiana lo cual resultarían ser motivadoras para los estudiantes y pueden generar un marco de aprendizaje adecuado (Batanero y Gea, 2018; Borovcnik, 2015).

Ya sea por el contexto de los problemas presentados, la formulación de los enunciados y su relación o no con el conocimiento del que disponen los alumnos acerca de situaciones reales, distintos autores como Pollatsek et al. (1987) han estudiado dificultades que se presentan en la comprensión de problemas que involucran la probabilidad condicional. En este mismo sentido, Batanero et al. (2012) encontraron estas dificultades en estudiantes de psicología y en futuros profesores de matemática. Los errores más comunes fueron confusión entre conceptos tales como los de probabilidad simple, condicional y conjunta, errores en la partición del espacio muestral, fallas en los procedimientos para calcular probabilidades, en el diseño de diagramas de árbol, falacia en la información de tasas base o proporciones de ocurrencia de hechos, confusión en la especificación de las probabilidades condicionales y sus inversas, así avanzando en los conceptos hasta llegar al establecimiento erróneo de la probabilidad total y de situaciones que involucran al teorema de Bayes (Attorresi et al., 2016; Díaz y De la Fuente, 2007).

Nuestro conocimiento acerca de atributos y características, reunidos en un conjunto de individuos o más generalmente de elementos que forman un colectivo a estudiar (población), está basado en un número más o menos reducido de dichos elementos, parcialidad que se conoce como muestra. La inferencia estadística a través de sus leyes provee el grado de confianza que se puede atribuir a las estimaciones basadas en una muestra, según el tamaño de la misma. Es sabido que el error de estimación del parámetro disminuye a medida que aumenta el tamaño de muestra.

Revista de Psicología, 23(2), 3–19 | 2024 | ISSN 2422-572X
https://doi.org/10.24215/2422572Xe167

La frecuencia de ocurrencia de un cierto evento aleatorio en una muestra de n ensayos se estabiliza alrededor de un número llamado probabilidad, a medida que el tamaño de la muestra es mayor. Esta última propiedad es una de las formas en que puede expresarse la denominada ley de los grandes números. Por otra parte, si se comparan las frecuencias relativas en muestras de un mismo tamaño, se observa que existe variabilidad entre ellas. Si ahora repetimos la comparación entre frecuencias relativas obtenidas en muestras de menor tamaño, encontraremos que hay mayor variabilidad entre estas últimas, es decir, existe más inestabilidad entre las frecuencias relativas calculadas a partir de un número menor de datos. Sobre la importancia del tamaño muestral y siguiendo la situación propuesta por Fischbein y Schnarch (1997), puede pensarse que no es tan infrecuente que nazcan siete varones sobre un total de diez nacimientos, pero es más infrecuente que nazcan 70 varones sobre un total de 100 nacimientos y aún mucho más infrecuente que nazcan 700 varones sobre un total de 1000, aunque en todos los casos mencionados – y solo visto como un cálculo matemático – la frecuencia de ocurrencia de nacimiento de un varón es la misma para los tres tamaños de muestra y tiene como valor 0.7. Con estas cifras se quiere mostrar que, suponiendo aproximadamente igual probabilidad de que nazca un varón o una mujer, es más frecuente que en una serie de nacimientos nazcan aproximadamente la mitad de varones y la mitad de mujeres especialmente si el tamaño de muestra es grande. Así, a pesar de que 7/10 es la misma proporción que 70/100 o que 700/1000 la probabilidad de ocurrencia de siete casos favorables (nacimientos de varones) sobre un total de diez nacimientos es mayor que la de 70 nacimientos de varones sobre un total de 100 nacimientos y así siguiendo. El modelo estadístico que explica estas situaciones es introducido mediante una variable aleatoria cuya asignación de probabilidad es conocida como distribución binomial. A partir de dicho modelo pueden calcularse las probabilidades que se le atribuyen a las diferentes configuraciones de nacimientos de varones. Cuando se asigna un valor numérico a cada resultado del espacio muestral lo que se está formalizando es una variable aleatoria. Esta puede ser continua o discreta, en este último caso existen algunas que son dicotómicas o dicotomizables. Entonces, se puede establecer como éxito a uno de los resultados y fracaso al restante, asignando el valor uno al primero y cero al segundo. Si ahora se asigna una determinada probabilidad de ocurrencia al éxito entonces se obtiene un modelo Bernoulli, cuya construcción dependerá de un solo parámetro, p, que es la probabilidad del éxito ya que la del fracaso será (1-p) por ser el suceso complementario. Un ejemplo de esta situación podría ser el registro del sexo en un nacimiento cuando la elección del éxito es, por ejemplo, varón. Esta situación responde al modelo Bernoulli con parámetro p = 0.5.

Es de interés conocer la probabilidad de obtener una cantidad específica de éxitos cuando se repite un experimento Bernoulli manteniendo la independencia de cada ensayo y la constancia de la probabilidad de éxito en cada uno de ellos. De esta manera se construye un modelo Binomial cuya variable llamaremos *X*, que asigna las probabilidades según la fórmula siguiente:

$$P(X = x) = \binom{n}{x} p^x (1 - p)^{n-x}$$

Revista de Psicología, 23(2), 3–19 | 2024 | ISSN 2422-572X
https://doi.org/10.24215/2422572Xe167

Donde p es la probabilidad de éxito del modelo dicotómico de Bernoulli y n es la cantidad de veces que se repite el experimento. Siendo estos los dos parámetros que definen el modelo binomial ya que son los necesarios para asignar las probabilidades a cada valor de la variable aleatoria X, cuyo recorrido es $x= 0,1,2,\ldots n$

Siguiendo con el ejemplo anterior, una variable binomial puede obtenerse al contabilizar la cantidad de varones que nacen en n nacimientos si mantenemos constante la probabilidad de éxito y consideramos cada uno de ellos como independientes.

No obstante, estas situaciones también pueden analizarse desde sus aspectos más intuitivos, observando cómo se comparan probabilidades de series aleatorias como la de los nacimientos. La estadística afirma que tienen mayor probabilidad de ocurrencia en una serie de ensayos aquellas configuraciones que se acercan al valor esperado para dicha serie. En el ejemplo, el valor esperado es n/2 es decir suponiendo que hay casi igual probabilidad de que nazca una mujer o un varón, tienen mayor probabilidad de ocurrencias las configuraciones con aproximadamente igual número de varones y de mujeres.

En la toma de decisiones en situaciones bajo incertidumbre, las personas suelen desestimar la lógica probabilística, circunstancia que puede generar errores en el pensamiento, los que han sido estudiados por diversos autores como por ejemplo Hope y Kelly (1983) y Kahneman et al. (1982). En su lugar suelen realizar procesos cognitivos de toma de decisión llamados heurísticos, donde se prioriza cierta información desestimando otra, tratando de esa manera, de reducir la complejidad de la información disponible. Existen diferentes heurísticos como el de anclaje, disponibilidad o el de representatividad (Kahneman, 2012). En este último caso se realizan estimaciones de probabilidad a un suceso según como se lo relacione o no con la población de origen. Esta situación puede provocar una insensibilidad de las estimaciones a los tamaños de la muestra, aplicando la ley de los grandes números incorrectamente a muestras “pequeñas”, falacia conocida como la ley de los pequeños números (Tversky y Kahneman, 1971). Esto puede conducir a sesgos donde se sobreestime la potencia de las herramientas estadísticas, como, por ejemplo, subestimar la amplitud de los intervalos para estimación de parámetros, o asumir altas expectativas en experimentos con pequeñas muestras. Estos sesgos se presentan aun en personas formadas en teoría de probabilidades y perjudican el desempeño cotidiano de las personas (Serrano et al., 1998).

Otro aspecto de interés vinculado a la estimación de la probabilidad en contextos significativos resulta de analizar la medida de la ocurrencia de un determinado suceso teniendo información previa acerca de la ocurrencia de otro suceso que puede estar o no asociado con el primero.

Dado un experimento aleatorio se le asocia un conjunto denominado espacio muestral E que contiene a todos los posibles resultados (sucesos elementales) de dicho experimento. Los subconjuntos del espacio muestral se denominan sucesos y en el caso en que el experimento cumpla ciertas condiciones puede definirse sobre los sucesos de ese espacio muestral una asignación de probabilidad. Cuando los sucesos elementales tienen igual probabilidad (equiprobables) se deduce que la probabilidad asignada a un suceso A es el cociente entre el número de elementos del espacio muestral que tiene el suceso considerado (número de casos favorables)

Revista de Psicología, 23(2), 3–19 | 2024 | ISSN 2422-572X
https://doi.org/10.24215/2422572Xe167

sobre el número de elementales totales que tiene el espacio muestral. Este enfoque se denomina de las probabilidades a priori o de Laplace. También existe el llamado frecuencialista o enfoque a posteriori que entiende la probabilidad como los números a los que tienden las frecuencias relativas cuando el número de repeticiones (ensayos) tiende al infinito.

En este marco de la teoría de probabilidades puede señalarse que la ocurrencia de un suceso A sin ninguna otra información, no necesariamente es igual a la ocurrencia de ese suceso A si se sabe que ocurrió también un suceso B. De ahí que se introduzca el concepto de probabilidad condicional de “A dado B” y se representa P(A/B).

Dicho de otro modo, la probabilidad de ocurrencia de un suceso A puede verse afectada por la ocurrencia de otro suceso B, respecto de la que tendría A sin esa ocurrencia de B.

Por ejemplo, si consideramos que la determinación del sexo del recién nacido se ve afectada por el azar y asumiendo aproximadamente igual probabilidad de nacimiento de varón o mujer, para una familia con dos hijos/as una posible representación del espacio muestral sería:

E = {(F,F) ; (F,M) ; (M,F) ; (M,M)} donde, F= Femenino y M= Masculino

Por ejemplo, el par ordenado (F, M) indica que el primer hijo/a es mujer y el segundo/a es varón.

La Probabilidad de cualquiera de las configuraciones del espacio muestral (sucesos elementales) vale 1/4, es decir, todos tienen igual probabilidad, en particular la probabilidad de que ambos hijos/as sean mujeres (F,F).

Sin embargo, la probabilidad de que ambos hijos/as sean mujeres (F,F) si se sabe que uno/a de los hijos/as es una mujer no es 1/4, teniendo esta información adicional el espacio muestral quedaría restringido a

Erestringido = {(F,F) ; (F,M) ; (M,F)}

Por lo cual la probabilidad de que nazcan dos mujeres si se sabe que uno/a de los hijos/as es una mujer es1/3.

Desde un punto de vista formal se define a la probabilidad condicional como:

$$P(A/B) = P(A \cap B) / P(B)$$

El numerador de la expresión es la probabilidad de la intersección de los sucesos A y B y el denominador es la probabilidad de la ocurrencia sólo de B.

Para el ejemplo anterior, teniendo en cuenta el espacio muestral original, la probabilidad de que los/as dos hijos/as sean mujeres sabiendo que uno/a de los hijos/as es una mujer es:

P(A) = P (los/as dos hijos/as sean mujeres) = P((F,F))=1/4

P(B) = P(uno/a de los hijos/as sea una mujer) = P((F,M) o (M,F) o (F,F))

Teniendo en cuenta que los sucesos elementales del espacio muestral son excluyentes y en este caso los consideramos equiprobables, entonces usando lenguaje simbólico, obtenemos:

$$P(B)=P((F,M)) + P((M,F)) +P((F,F)) = 3/4$$

$$P(A \cap B) = P(A) =1/4$$

$$P (A/B) = P(A \cap B) / P(B) = 1/3$$

Para llegar a una resolución del problema se tiene que interpretar correctamente cuál es la condición, es decir, en el problema presentado, tienen que distinguir que la condición "uno de los hijos/as es una niña" no significa que el primer hijo/a es una niña, puede haber nacido niña el primer o el segundo/a hijo/a, para así descartar la situación "el primer hijo/a es una niña" y entonces recurrir trivialmente a que la probabilidad de nacimiento de una niña para el segundo hijo/a es 1/2.

El objetivo de la presente investigación fue evaluar la intuición para resolver problemas de asignación de probabilidad en jóvenes ingresantes a la universidad sin formación en la escuela media, según indicaron los participantes. Se quiso comparar las dificultades de resolución en dos situaciones de aleatoriedad en contextos similares (nacimiento de niños/as) pero de diferentes características.

En una de ellas se explora la habilidad para analizar secuencias aleatorias (sexo en el nacimiento) en las que se debió tener en cuenta la estabilidad de las frecuencias relativas. La otra situación presentada involucra las posibles resoluciones de un problema de cálculo de probabilidad condicional en el mismo contexto aleatorio de los nacimientos, de planteo sencillo en el lenguaje natural pero que requirió de poder determinar todos los posibles resultados del sexo en dos nacimientos en una misma familia.

Se consideró como hipótesis que presenta mayor dificultad la resolución (planteo y cálculo) del problema de probabilidad condicional que la de comparación de probabilidad en el problema de estabilidad de las frecuencias en los nacimientos. Este último resulta más conocido dada la experiencia en la vida cotidiana, sobre todo en los juegos de azar donde en general se analiza la comparación de probabilidades motivado por las apuestas en juego. Las situaciones que involucran probabilidad condicional son menos frecuentes en cuanto a su uso cotidiano, en general se apela a la condicionalidad a lo sumo de manera cualitativa. El aumento en las dificultades para las resoluciones de los problemas presentados puede atribuirse también a que en el problema de probabilidad condicional se requiere del cálculo de un valor numérico mientras que en el otro problema se trata de una comparación de probabilidades.

Métodos

Diseño de actividades

Se adaptaron dos problemas que permitieron indagar estimaciones de probabilidad en situaciones en las que interviene el azar. En el primero, denominado Problema del Sesgo de Proporciones (SP) (Kahneman y Tversky, 1972). Dicha indagación se efectuó al comparar proporciones de dos muestras de diferente tamaño y en las que se debió tener en cuenta para su resolución nociones intuitivas vinculadas a la estabilidad de las frecuencias relativas lo que formalmente se conoce como ley de

Revista de Psicología, 23(2), 3–19 | 2024 | ISSN 2422-572X
https://doi.org/10.24215/2422572Xe167

los grandes números. En el segundo, denominado Problema de Probabilidad Condicional (PC) (Ríos, 1967), se requirió el cálculo de probabilidades en una situación azarosa, conocida una cierta condición de la probabilidad.

A continuación, se muestran ambos problemas.

Problema del Sesgo de Proporciones (SP)

> Se registran los nacimientos diarios en dos centros de salud, A y B. El centro A tiene menor infraestructura que el B. Teniendo en cuenta que aproximadamente el 50% de los bebés que nacen son varones, ¿qué es más probable que ocurra entre estas dos opciones a acontecer en un mismo día?
>
> a) que nazcan 7 o más varones de 10 bebés nacidos en el centro A.
>
> b) que nazcan 70 o más varones de 100 bebés nacidos en el centro B.

Según el modelo de distribución de probabilidad binomial para una probabilidad de éxito de nacimiento de varón de p=.5 y para n= 10 ensayos, la probabilidad de que nazcan siete o más varones en un total de diez nacimientos es de 0,17187.Siguiendo el mismo modelo, pero variando el número de ensayos, la probabilidad de que nazcan setenta o más varones en un total de cien nacimientos es de 0,00004.

Como se ve en el cálculo formal se observa que a pesar de que la proporción de nacimientos de varones es la misma en las dos situaciones expuestas resulta que al aumentar el número de nacimientos (ensayos) disminuye la probabilidad de ocurrencia del evento propuesto. De lo anterior se deduce que la opción correcta (la de mayor probabilidad) es la a).

La estadística afirma que tienen mayor probabilidad de ocurrencia en una serie de ensayos aquellas configuraciones que se acercan al valor esperado para dicha serie. En el problema SP, el valor esperado es n/2 es decir suponiendo que hay casi igual probabilidad de que nazca una mujer o un varón, tienen mayor probabilidad de ocurrencias las configuraciones con aproximadamente igual número de varones y de mujeres.

Problema de Probabilidad Condicional (PC)

> Una familia tiene dos hijos/as. Uno/a de ellos/as es una niña. ¿Cuál es la probabilidad de que el otro/a sea una niña?

Como se ejemplificó en la introducción se definen los eventos:

A: los/as dos hijos/as sean mujeres

B: uno/a de los hijos/as sea una mujer

Por lo que la probabilidad pedida para este problema es:

$$P(A/B) = P(A \cap B) / P(B)$$

$$P(B) = P((F,M)) + P((M,F)) +P((F,F)) = 3/4$$

$$P(A \cap B) = P(A) =1/4$$

$$P(A/B) = P(A \cap B) / P(B) = 1/3$$

Revista de Psicología, 23(2), 3–19 | 2024 | ISSN 2422-572X
https://doi.org/10.24215/2422572Xe167

Participantes

Se administraron el Problema del Sesgo de la Proporción (SP) a 207 jóvenes y el de Probabilidad Condicional (PC) a 116, todos ingresantes a la Universidad de Buenos Aires, sin cursos formales acerca de probabilidad según afirmaron los participantes, provenientes de colegios públicos y privados de nivel socioeconómico medio. Las muestras fueron seleccionadas por accesibilidad y de manera independiente.

Resultados

A continuación, se define la categorización para el Problema SP y se presenta la distribución correspondiente en porcentaje para la muestra de tamaño n= 207 en la Tabla 1.

Categoría	Porcentaje (n=207)
Correctas. Más probable opción a) Respuestas que distinguieron que los apartamientos de la cantidad de varones en n nacimientos respecto del valor esperado es más fácil de encontrar en series de menor tamaño, razón por la cual se inclinaron por la opción a).	28
Aciertos mal justificados. Más probable opción a) Respuestas que optaron por el suceso de la opción a) que, a pesar de ser la opción correcta, fue fundamentada incorrectamente. Se hace referencia a que, al tratarse de una muestra de menor tamaño, era más fácil de ser contrastada mediante réplicas de la experiencia, cuestión que se desvía del problema presentado.	3
Incorrectas. Igual probabilidad Respuestas que, desestimando el tamaño de la muestra, otorgaron, erróneamente, igual probabilidad a ambas opciones, justificándose en la igualdad de proporciones (7/10 es igual a 70/100).	42
Incorrectas. Más probable opción b) (muestra de mayor tamaño) Respuestas que optan por la opción b) confiriéndole mayor confianza dado que presenta mayor tamaño muestral haciendo referencia a que de esta forma disminuye el error de estimación, cuestión, esta última, que no es incorrecta pero que responde incorrectamente al problema planteado.	20
No interpreta Respuestas que no interpretaron el reactivo presentado dando argumentos totalmente alejados de las cuestiones consultadas.	3
No contesta	4
Total	100

Tabla 1. Categorización de las Respuestas del Problema del Sesgo de Proporciones.

Problema SP

A continuación, se muestran algunos ejemplos textuales de respuestas.

Ejemplos de Respuestas Correctas (más probable opción a):

> Estudiante 1SP: (elije opción (a)), porque al haber una menor cantidad de bebés nacidos hay más probabilidad que 7 a más sean varones. En cambio, en una cantidad mayor de bebés nacidos puede haber más niñas.

> Estudiante 2SP: es más probable la (a); a medida que aumenta la cantidad de nacimientos el número tiende a ser 50% varones 50% mujeres.

Revista de Psicología, 23(2), 3–19 | 2024 | ISSN 2422-572X
https://doi.org/10.24215/2422572Xe167

Ejemplos de Respuestas Incorrectas. Igual probabilidad para ambas opciones:

> Estudiante3SP: Son iguales de probables las dos opciones, ya que en ambas hay una probabilidad del 70 % en que los primeros bebés nacidos en el hospital sean varones.
>
> Estudiante4SP: Es lo mismo decir que 7 de cada 10 sean varones que decir que 70 de cada 100 lo sean.
>
> Estudiante 5SP: Cualquiera de las dos puede ser elegida ya que me están haciendo dos veces la misma pregunta.

Ejemplos de Respuestas Incorrectas (más probable opción b):

> Estudiante 6SP: (a) y (b) son lo mismo, en porcentaje. Pero científicamente queda más corroborado que de los 100, 70 hayan sido varones, ya que se tomó los datos de más personas. Cuánto más lanzamientos, experimentos, etc., se hagan, más confiable la probabilidad.
>
> Estudiante 7SP: Es más probable el (b) ya que cuanta más cantidad se tome las estadísticas van a ser más precisas.

Ejemplo de Respuestas que no interpretan el problema:

> Estudiante 8SP: Ninguno es más probable que otro en este caso, cada nacimiento tendrá una causalidad y no creo que respete una estadística.

La categorización para el Problema PC se presenta a continuación en porcentaje para la muestra de tamaño n= 116 en la Tabla 2.

Categoría	Porcentaje (n=116)
Intentos fallidos Tienen en cuenta la información del enunciado, pero no llegan a plantear buenas estrategias.	4
Simplificación del problema Asumen que el 1er hijo/a es niña y el 2do tiene 50% de chances de ser niña dada la equiprobabilidad del sexo en el nacimiento.	64
Idea de predominio numérico de mujeres Piensan que es más probable que sea niña pues hay más mujeres que hombres	6
Idea de compensación Asumen que el 1er hijo/a es niña entonces el segundo/a debería ser niño.	4
No interpretan o evaden la resolución Brindan respuestas que evaden el problema, "es cuestión de genética" o carentes de sentido.	18
No saben o no contestan Manifiestan expresamente no saber qué contestar o no responden	4
Total	100

Tabla 2. Categorización de las Respuestas del Problema de Probabilidad Condicional.

Revista de Psicología, 23(2), 3–19 | 2024 | ISSN 2422-572X
https://doi.org/10.24215/2422572Xe167

Problema PC

A continuación, se muestran algunos ejemplos de respuestas.

Ejemplo de Respuestas que evidencian intentos fallidos de resolución:

> Estudiante 1PC: 50% ya que hay 2 posibilidades de que sea varón o mujer, pero también pienso en la frase subrayada (hace referencia a “Uno/a de ellos/as es una niña”) porque me figura tiene una vuelta de rosca.

Ejemplos de Respuestas que simplifican el problema:

> Estudiante 2PC: Respuesta: 1/2. La probabilidad de que el segundo hijo sea niña no depende del sexo del primer hijo y dado que hay solo dos opciones (hombre o mujer), la probabilidad de que sea niña es del 50 %.

> Estudiante 3PC: Según la biología la posibilidad sería de un 50%, ya que hay sólo 2 sexos y los padres no pueden hacer nada al respecto, a menos que adopten una niña.

Ejemplo de Respuestas que se basan en el predominio numérico de mujeres:

> Estudiante 4PC: estadísticamente en el mundo hay más mujeres que hombres, la probabilidad de que sea niña es mayor.

Ejemplo de Respuestas que se basan en la idea de compensación:

> Estudiante 5PC: creo que esa probabilidad depende genéticamente del padre y la madre, pero supongo que será varón porque ya tienen una nena, por lo tanto, uno pensaría que ahora tendrían un varón, la probabilidad que tenga una nena es menor.

Ejemplos de Respuestas que no interpretan o evaden el problema:

> Estudiante 6PC: La probabilidad que elijo es que los dos hijos de la familia sean gemelas.

> Estudiante 7PC:La probabilidad se puede dar, pero eso no es algo que el hombre en este caso la familia pueda decidir es algo de la naturaleza o de Dios o a lo que se lo quiera atribuir.

> Estudiante 8PC: No habría una probabilidad 100 % de que el otro hijo sea niña ni tampoco niño. La probabilidad no es del 100% ni nula, por ejemplo, voy a poner en mi familia, mi mamá siempre quiso tener 2 niñas y 2 niños, sin embargo, tuvo 4 varones. Son hechos que, aunque algunos tengan probabilidades siempre puede haber algo que no se puede probar.

> Estudiante 9PC: La probabilidad es de un 75%, ya que los gametos masculinos poseen tanto masculino (Y) como femenino (X) y la femenina sólo XX.

Discusión

Se quiso comparar las dificultades de resolución en dos situaciones de aleatoriedad en contextos similares (nacimiento de niños/as) pero de diferentes características.

Revista de Psicología, 23(2), 3–19 | 2024 | ISSN 2422-572X
https://doi.org/10.24215/2422572Xe167

En una de ellas se compararon proporciones de dos muestras de diferente tamaño y se estimaron probabilidades, mientras que en la otra se presentó un problema que requería el cálculo de probabilidad condicional.

Para el problema del Sesgo de las Proporciones (SP) se encontró que algo menos de un tercio de los participantes dieron argumentos correctos ante la situación presentada. El resto de las categorías reunieron argumentaciones incorrectas, aunque en distinto grado. La respuesta más frecuente y que resultó errónea, superando en porcentaje al de las respuestas correctas, fue la asignación de igual probabilidad a las dos opciones del problema. De este modo se prestó atención a la igualdad matemática de proporciones de las muestras y se desatendió la cuestión del tamaño de las mismas, idea central vinculada a la alta variabilidad en pocos ensayos, asociada a la estabilidad de las frecuencias relativas y que por lo tanto condujo a respuestas erróneas.

Otra categoría que resultó significativa en porcentaje fue la que reunió a las preferencias por la opción incorrecta (opción b) con justificaciones basadas en que a mayor tamaño de muestra menor error de estimación, cuestión que desvía de la pregunta a considerar. Esta justificación no condujo a la cuestión central de que el alejamiento de los valores de la variable aleatoria cantidad de varones en n nacimientos respecto de su valor esperado es más probable en muestras de menor tamaño.

Para el problema de probabilidad condicional puede afirmarse que sólo existieron intentos fallidos de plantear la resolución, sin que éstas se aproximaran mínimamente a un planteo intuitivo cercano a la resolución correcta como el de imaginar el conjunto de los posibles resultados de la situación. Un único participante llegó a hacerlo, aunque no pudo restringirlo a la condición brindada y unos pocos mantuvieron la propuesta del problema sin simplificarlo, aunque no llegaron a una buena resolución.

La mayoría respondió que la probabilidad era un medio. El análisis de las respuestas llevó a pensar que se trivializó el problema bajo el supuesto de que el primer hijo/a había nacido mujer, lo cual cambió el enunciado del reactivo. Con esta última modificación la probabilidad del sexo en el segundo nacimiento, que para el problema sería que nazca otra niña, dada la aceptada equiprobabilidad e independencia de esta variable respecto de nacimientos anteriores es correctamente un medio, pero responde a una situación simplificada respecto de la planteada originalmente. Otras resoluciones manifestaron preconceptos que cuestionaron la noción de equiprobabilidad del sexo en el nacimiento. También aparecieron, aunque en menor medida, sesgos de compensación ya encontrados en otros experimentos aleatorios como la falacia del jugador (Attorresi et al., 2007). Un número reducido señaló explícitamente no saber cómo responder. Hubo un porcentaje destacado de respuestas que vincularon la resolución de la situación presentada con conocimientos de genética, probablemente influenciados por la asignatura biología que estaban cursando en ese momento, que independientemente de que fuesen o no correctos evadían el planteo de los sucesos probabilísticos presentados y la medida de su ocurrencia. También aparecieron respuestas totalmente desviadas del planteo del problema, anécdotas familiares de “búsqueda de la parejita” o de “varones” y a nociones que desvinculaban la idea de

Revista de Psicología, 23(2), 3–19 | 2024 | ISSN 2422-572X
https://doi.org/10.24215/2422572Xe167

azar con la de probabilidad, poniéndose de manifiesto que la noción de probabilidad en el habla puede tener distintas significaciones semánticas.

Es posible entonces observar que, si bien las situaciones presentadas en esta investigación son de simple formulación en el lenguaje natural, los participantes tienden a no comprender o simplifican el enunciado del problema y mayoritariamente no aplican las nociones e intuiciones que son necesarias para su adecuada resolución.

La falta de resoluciones exitosas para el problema de la probabilidad condicional confirma la hipótesis propuesta de que este problema presenta mayor dificultad que el del sesgo de las proporciones. En el primero no hubo un planteo que alcanzara una buena resolución sino algún esbozo aproximado de solución que, es de destacar frente a los magros resultados generales, pero que no culminó en la resolución correcta. Ante la imposibilidad de representar la información brindada en el problema de probabilidad condicional enumerando los resultados del experimento y sacando aquel resultado incompatible con la información explicitada, es decir, que hayan nacido dos varones, la estrategia mayoritaria fue simplificar el problema presentado hasta modificarlo en una situación trivial.

Las dificultades de los participantes en este estudio aparecen con bastante regularidad en la enseñanza de temas de probabilidad dictados por los autores de este trabajo en su tarea como docentes cuando se abordan estos temas (Attorresi et al., 2008; Batanero, 2000).

Para controlar y minimizar los sesgos en el aprendizaje de conceptos probabilísticos se sugiere presentar a los alumnos situaciones problemáticas que involucren el azar cuestionando creencias personales a partir de la evidencia experimental y permitiendo la toma de decisiones en situaciones de incertidumbre con una base racional (Alvarado et al., 2018). Así también, una enseñanza que considere la argumentación y el aprendizaje del espacio muestral para encauzar el uso de recursos intuitivos (Vergara et al., 2020).

Financiación

Esta investigación fue financiada con subsidios de la Universidad de Buenos Aires (UBACyT 2018-22, Mod I, Código 20020170100200BA) y de la Agencia Nacional de Promoción Científica y Tecnológica (ANPCyT PICT-2017-3226).

Referencias

Alvarado H., Estrella, S., Retamal, L. y Galindo, M. (2018). Intuiciones probabilísticas en estudiantes de ingeniería: Implicaciones para la enseñanza de la probabilidad. *Revista Latinoamericana de Investigación en Matemática Educativa, 21*(2), 131–156. https://doi.org/10.12802/relime.18.2121

Attorresi, H., García-Díaz, A. M. y Pralong, H. O. (2005). Identificación de situaciones aleatorias en estudiantes universitarios. *Anuario de Investigaciones, 13*(2), 31–41.

Attorresi, H., García-Díaz, A. M. y Pralong, H. O. (2007) Sesgos en la comprensión de dos situaciones típicas de incertidumbre y de azar. *Perspectivas en Psicología, 4*(1), 28–37.

Attorresi, H., García-Díaz, A. M. y Pralong, H. O. (2008). Sesgos en la estimación de probabilidades para dos situaciones secuenciales aleatorias. *Summa Psicológica, 5*(1), 3–12.

Attorresi, H., García-Díaz, A. M. y Pralong, H. O. (2009). Identificación de variables ocultas y su vinculación con el reconocimiento de la aleatoriedad. *Summa Psicológica, 6*(2), 43–54. https://doi.org/10.18774/448x.2009.6.61

Attorresi, H., García-Díaz, A. M. y Pralong, H. O. (2015). La falacia de la conjunción: efectos de la variación de opciones e independencia. *Cogency, 6*(2), 9–27.

Attorresi, H., García-Díaz, A. M. y Pralong, H. O. (2016). The significance as an extention of representativeness heuristic. Two examples. *Revista de Psicología, 18*(1), 19–26.

Barragués, J., Guisasola, J. y Morais, A. (2005). Concepciones de los estudiantes de primer ciclo de universidad sobre estimación de la probabilidad. *Educación Matemática, 17*(1), 55–85.

Batanero, C. (2000). ¿Hacia dónde va la educación estadística? *Blaix, 15*, 2–13.

Batanero, C. y Gea, M. M. (2018). El riesgo como contexto en la enseñanza de la probabilidad condicional. En L. A., Serna, y D. Páges (Eds.), *Acta latinoamericana de matemática educativa* (pp. 125–132). Comité Latinoamericano de Matemática Educativa.

Batanero, C., Contreras M. y Díaz C. (2012). Sesgos en el razonamiento sobre probabilidad condicional e implicaciones para la enseñanza. *Revista Digital Matemática, Educación e Internet, 12*(2), 1–13. https://doi.org/10.18845/rdmei.v12i2.1673

Batanero, C., Gómez, E., Contreras J. y Díaz C. (2015). Conocimiento matemático de profesores de primaria en formación para la enseñanza de la probabilidad: un estudio exploratorio. *Práxis Educativa, 10*(1), 11–34. https://doi.org/10.5212/praxeduc.v.10i1.0001

Borovcnik, M. (2015). Risk and decision making: The "logic" of probability. *The Mathematics Enthusiast, 12*(1), 113–139. https://doi.org/10.54870/1551-3440.1339

Cardeñoso, J. (2001). *Las creencias y conocimientos de los profesores de primaria andaluces sobre la matemática escolar. Modelización de concepciones sobre la aleatoriedad y probabilidad.* Tesis Doctoral. Universidad de Cádiz.

Díaz, C. (2003). *Heurísticas y sesgos en el razonamiento probabilístico. Implicaciones para la enseñanza de la estadística* [presentación en congreso]. XXVII Congreso Nacional de Estadística e Investigación Operativa. Universidad de Lleida (España).

Díaz, C. y de la Fuente, I. (2007). Validación de un cuestionario de razonamiento probabilístico condicional. *Revista Electrónica de Metodología Aplicada, 12*(1), 1–15. https://doi.org/10.17811/rema.12.1.2007.1-15

Elbehary, S. (2020). Discussing the conditional probability from a cognitive psychological perspective. *American Journal of Educational Research, 8*(7), 491–501. https://doi.org/10.12691/education-8-7-7

Fischbein, E. y Schnarch, D. (1997). The evolution with age of probabilistic, intuitively based misconceptions. *Journal for Research in Mathematics Education, 28*(1), 96-105. https://doi.org/10.2307/749665

Guisasola, J. y Barragués, J. (2002). Heurísticas y sesgos de los estudiantes de primer ciclo de universidad en la resolución de problemas de probabilidad. *Enseñanza de las Ciencias, 20*(2), 285–302.

Hope, J. y Kelly, I. (1983). Common difficulties with probabilistic reasoning. *Mathematics Teacher, 76*(8), 565–570. https://doi.org/10.5951/mt.76.8.0565

Kahneman, D. (2012). *Pensar rápido, pensar despacio.* Debate.

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

Kahneman, D. y Tversky, A. (1972). *Subjective probability: A judgment of representativeness.* Elsevier.

Konold, C. (1989). Informal conceptions of probability. *Cognition and Instruction, 6*(1), 59-98. https://doi.org/10.1207/s1532690xci0601_3

Piaget, J. e Inhelder, B. (1974). *La genèse de l'idèe de hasard chez l'enfant.* Press Universitaires de France.

Pollatsek, A., Well, A., Konold C., Hardiman, P. y Cobb, G. (1987). Understanding conditional probability. *Organitation,Behavior and Human Decision Processes, 40*(2), 255–269. https://doi.org/10.1016/0749-5978(87)90015-x

Ríos, S. (1967). *Métodos estadísticos.* McGraw-Hill.

Santaló, L. (1970). *Probabilidad e inferencia estadística.* Organización de los Estados Americanos.

Serradó, A., Cardeñoso, J. y Azcárate, P. (2005). Los obstáculos en el aprendizaje del conocimiento probabilístico: Su incidencia desde los libros de texto. *Statistics Education Research Journal, 4*(2), 59–81.

Serrano, L., Batanero, C., Ortiz J. y Cañizares J. (1998). Heurísticos y sesgos en el razonamiento probabilística de los estudiantes de secundaria. *Educación Matemática, 10*(1), 7-25.

Tversky, A. y Kahneman, D. (1971). Belief in the law of small numbers. *Psychological Bulletin, 76*(2), 105–110. https://doi.org/10.1037/h0031322

Vergara, A., Estrella, S. y Vidal-Szabó, P. (2020). Relaciones entre pensamiento proporcional y pensamiento probabilístico en situaciones de toma de decisiones. *Revista Latinoamericana de Investigación en Matemática Educativa, 23*(1), 7-36. https://doi.org/10.12802/relime.20.2311

Revista de Psicología, 23(2), 3–19 | 2024 | ISSN 2422-572X
https://doi.org/10.24215/2422572Xe167