

Efecto de los formatos y el tipo de información sobre las respuestas al resolver un problema binario de probabilidad condicional

*Gabriel Yáñez Canal**

*Ana Rátiva Hernández***

*Diana Lozano Rodríguez****

RESUMEN

Este trabajo presenta algunos de los resultados de una investigación dirigida a estudiar el efecto de los formatos y el tipo de información sobre las respuestas y estrategias que utilizan los estudiantes para resolver un problema binario de probabilidad condicional. Para esto se realizó un estudio en el que participaron 240 estudiantes de diversas carreras de la Universidad Industrial de Santander (UIS) en Bucaramanga, Colombia.

Los resultados obtenidos mostraron que para problemas en los que se da información condicional hay efecto tanto del formato como del tipo de información, sin embargo, para aquellos que no contienen información condicional el formato no tuvo un efecto significativo.

Palabras clave: Problema binario de probabilidad condicional, formato de presentación, frecuencias naturales, tipo de información.

* Universidad Industrial de Santander (UIS). Dirección electrónica:gyanez@uis.edu.co.

** Universidad Industrial de Santander (UIS). Dirección electrónica:mayuyao_4@hotmail.com.

*** Universidad Industrial de Santander (UIS). Dirección electrónica:dianalozano28@hotmail.com.

PRESENTACIÓN DEL PROBLEMA

La probabilidad condicional, un tema en general complicado para quienes lo conocen, es un concepto fundamental de la probabilidad que muestra el efecto que puede tener un evento sobre la ocurrencia de otro. Su comprensión y aprendizaje encierran muchas dificultades porque su cálculo requiere tomar como nuevo espacio muestral el evento condicionante y calcular las probabilidades de los eventos ajustadas a este nuevo espacio. Este ajuste a un nuevo espacio muestral conlleva la identificación plena de eventos marginales y condicionados, además de eventos conjuntos.

Atendiendo a esto último, Yáñez (2001) presenta una clasificación de los problemas binarios de probabilidad condicional en consonancia con la cantidad de probabilidades condicionales presentes en la parte informativa y un análisis de cada una de las representaciones utilizadas para resolverlos (árbol, tablas y álgebra) indicando la pertinencia de cada una de ellas para resolver los diferentes tipos de problemas. Se habla de problemas binarios porque solo se consideran dos eventos: A y B, junto con sus complementos.

Entre los estudios que abordan el efecto que ciertas estrategias pueden producir para superar estas dificultades, se destaca el trabajo de Gigerenzer & Hoffrage (1995) quienes realizaron una investigación del razonamiento involucrado en los problemas que contienen probabilidades condicionales, más exactamente los que implican el uso del teorema de Bayes, y llegaron a que dicho razonamiento mejora cuando a las personas se les presenta la información en formato de frecuencias naturales. En su investigación Gigerenzer & Hoffrage (1995), tratan un solo tipo de problema (Tipo 7) según la clasificación realizada por Yáñez (2001), en el cual se encuentra una clara correspondencia entre la información del problema y el diagrama de árbol, lo que produce éxito en la solución del mismo. Por esta razón, es de interés analizar si el éxito descrito por Gigerenzer & Hoffrage (1995) con las frecuencias naturales se extiende a problemas de otro tipo.

Motivados por estos antecedentes y ante la importancia del tema, nos propusimos realizar una investigación que diera respuesta a las siguientes preguntas: si se cambia el tipo de problema usado por Gigerenzer & Hoffrage (1995): ¿qué efectos tiene sobre las respuestas el formato de presentación del problema? y estos efectos ¿son simplemente por el formato de presentación y el tipo de problema, o hay otros elementos que influyen? ¿Cuáles? En este trabajo presentamos las respuestas a la primera pregunta.

A continuación describimos la metodología utilizada para resolver estas preguntas, luego se presentan los resultados y se realiza una discusión alrededor de ellos; finalmente se presentan las conclusiones y las referencias citadas a lo largo del trabajo.

Metodología

Los problemas objeto de estudio fueron de los casos 2 y 5, según la clasificación de Yáñez (2001) que se presentan cuando se da como información:

Caso 2: Dos intersecciones y una marginal (nivel 0, pues no presenta condicionales en su información)

Caso 5: Dos intersecciones y una condicional (nivel 1, presenta una condicional en su información)

El objetivo fue cuantificar el efecto del formato y del tipo de información en el éxito de la solución a un problema de probabilidad condicional. La muestra estuvo conformada por 240 estudiantes de Ingeniería y Ciencias de la Universidad Industrial de Santander (UIS) en Bucaramanga, Colombia. Para esta fase se elaboraron 4 problemas, todos del mismo contexto y con una misma pregunta, pero diferente tipo de información y formato de presentación, de la siguiente manera:

Problema 1: Caso 5 en formato de frecuencias

En un salón de clase con 50 estudiantes, 15 de ellos aprobaron Biología y español, 18 estudiantes aprobaron biología y no aprobaron español. Se sabe, además, que de los que no aprobaron español 10 no aprobaron biología. Calcule la probabilidad de que, elegido un estudiante al azar de los que no aprobaron biología, este haya aprobado español.

Problema 2: Casos 5 en formato de probabilidades

En un salón de clase la probabilidad de que al seleccionar aleatoriamente un estudiante este haya aprobado biología y español es de 0.3, y la probabilidad de que haya aprobado biología y no haya aprobado español es de 0.36. Se sabe, además, que si se selecciona aleatoriamente un estudiante entre los que no aprobaron español la probabilidad de que este no haya aprobado biología es 0.357. Calcule la probabilidad de que elegido un estudiante al azar de los que no aprobaron biología este haya aprobado español.

Problema 3: Caso 2 en formato de frecuencias

En un salón de clases con 50 estudiantes, 33 aprobaron biología. 15 aprobaron biología y español. 7 aprobaron español y no aprobaron biología. Calcule la

probabilidad de que elegido un estudiante al azar de los que no aprobaron biología este haya aprobado español.

Problema 4: Caso 2 en formato de probabilidades

En un salón de clases la probabilidad de que al seleccionar un estudiante al azar este haya aprobado biología es 0.66. La probabilidad de que haya aprobado biología y español es 0.3. Se sabe, además, que la probabilidad de que un estudiante haya aprobado español y no haya aprobado biología es de 0.14. Calcule la probabilidad de que elegido un estudiante al azar de los que no aprobaron biología este haya aprobado español.

Cada estudiante debía resolver uno de estos problemas; la asignación del problema a cada estudiante fue aleatoria y el tiempo de aplicación fue de 10 minutos

RESULTADOS Y DISCUSIÓN

Para estudiar el efecto de los formatos y el tipo de información sobre las respuestas de los estudiantes, se decidió hacer pruebas de proporciones con los datos obtenidos al aplicar las pruebas individuales con el fin de realizar análisis objetivos de los datos que condujeran a deducciones válidas respecto al problema de estudio. En la tabla 1 se muestran las proporciones de respuestas correctas obtenidas para cada uno de los problemas clasificados de acuerdo con el caso al cual corresponde y con el formato de presentación.

Una primera observación que se puede hacer a los resultados de la tabla 1 es la baja proporción de respuestas correctas: únicamente 31 estudiantes resolvieron el problema correctamente (13%). Una prueba t permite inferir que *el formato de presentación del problema* no tuvo un efecto significativo (valor $p=0.28$) sobre el éxito al resolver el problema en ninguno de los dos tipos (frecuencias naturales y probabilidades).

Tabla 2. Proporción de respuestas correctas, según tipo de información y formato.

<i>Formato Caso</i>	<i>Frecuencias naturales</i>	<i>Probabilidades</i>	<i>Total</i>
Caso 2	12/60	14/60	26/120
Caso 5	5/60	0/60	5/120
Total	17/120	14/120	31/240

Lo contrario sucedió con el tipo de información (caso 2 y caso 5); es así que se aprecian diferencias entre las proporciones de respuestas correctas

(valor $p < 0.0001$), lo que permite decir que el caso 2 produce una proporción de respuestas correctas mayor que la que produce el caso 5. Este resultado está en la dirección planteada por el mismo Yáñez (2001) cuando afirma que la dificultad de un problema de probabilidad condicional es mayor cuanto mayor sea el número de probabilidades condicionales que contiene en su parte informativa.

Para ver la interacción que se da entre formato y tipo de información, se elaboró el gráfico 1, en el cual se puede ver que los problemas de caso 2 son refractarios al cambio de formato, en tanto que los problemas de caso 5 sí son sensibles a este cambio, siendo las frecuencias naturales mejores para este tipo de información. En conclusión, cuando la información contiene datos con probabilidades condicionales las frecuencias naturales producen mayor proporción de respuestas correctas que las probabilidades.

Haciendo una prueba para comparar las dos proporciones de respuestas en el caso 5, se obtuvo una diferencia significativa (valor $p = 0.0113$), es decir, que los problemas de tipo 5 presentados en formatos de frecuencias naturales producen más respuestas correctas que los de el mismo tipo, pero en formato de probabilidades.

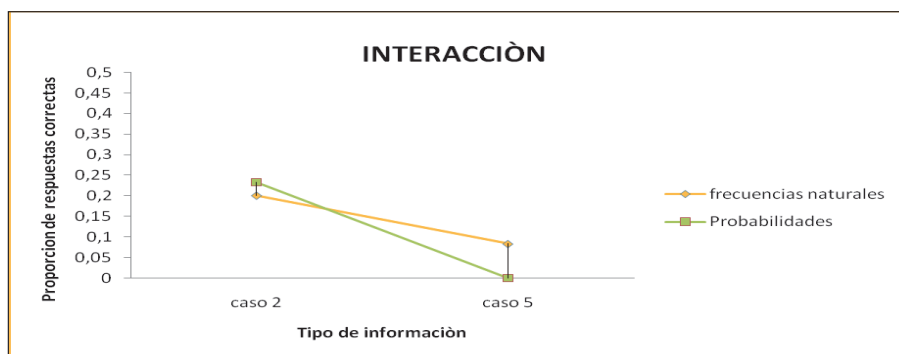


Gráfico 1. Interacción entre formato y tipo de información

CONCLUSIONES

En este estudio se encontró que para los problemas de caso 2 y caso 5 presentados en formatos de frecuencias naturales y probabilidad, el tipo de información tiene un efecto significativo sobre las respuestas, produciendo mayor proporción de respuestas correctas aquellos en los que la información no presenta datos condicionales (caso 2). De esta forma, los logros hallados por Gigerenzer & Hoffrage (1995), se puede afirmar, parecen limitarse so-

lamente al tipo de problemas donde la información contiene probabilidades condicionales, es decir, los efectos que puedan tener las frecuencias naturales (formato de presentación) no son independientes del nivel del problema que se proponga (tipo de información); más aún, este formato solo es mejor cuando existe alguna información sobre probabilidades condicionales.

Finalmente, hay que decir que la discusión alrededor de las frecuencias naturales dada por Gigerenzer & Hoffrage (1995), Martignon & Wassner (2002) y los demás, ahora debe ser ampliada considerando no solo problemas que en su información presenten dos condicionales y una marginal, sino también problemas que den información sobre otros eventos, pues, como se dijo, el tipo de información tiene un efecto mayor que el que puede tener el formato de presentación para problemas que no presentan condicionales.

REFERENCIAS BIBLIOGRÁFICAS

- Gigerenzer, G. & Hoffrage, U. (1995). How to improve Bayesian reasoning without instruction: Frequency format. *Psychological Review*, 102, 684-704.
- Martignon, L. & Wassner, C. (2002). *Teaching Decision Making and Statistical Thinking With Natural Frequencies*. VI International Conference on Teaching Statistics. En B. Phillips (Ed.). Ciudad del Cabo, Sur África, p. 1-4.
- Yáñez, G. (2001). El álgebra, las Tablas y los Árboles en Problemas de Probabilidad Condicional. *Iniciación a la investigación en didáctica de la matemática. Homenaje al profesor Mauricio Castro*. En Gómez, P., y Rico, L. Granada, España, p. 355-37.