

Un estudio longitudinal sobre los efectos de la instrucción en las intuiciones primarias asociadas al razonamiento Bayesiano

GABRIEL YÁÑEZ CANAL

gyanez@uis.edu.co
Universidad Industrial de Santander (Profesor)

CINDY NATHALIA MORGADO HERNÁNDEZ

nathalia.morgado@udes.edu.co
Universidad de Santander (Profesor)

Resumen. Se presentan algunos resultados de una investigación realizada para conocer la transformación del razonamiento bayesiano en estudiantes universitarios en un curso semestral de estadística. Para esto se diseñaron y aplicaron tres pruebas en momentos diferentes durante el semestre. Los resultados muestran que el proceso de enseñanza en ocasiones perjudica las intuiciones primarias y en otras da elementos a los estudiantes para resolver situaciones problemáticas relacionadas con la probabilidad condicional.

Palabras clave: razonamiento bayesiano, teorema de Bayes, datos longitudinales, probabilidad condicional.

1. Introducción

El razonamiento bayesiano hace referencia al cálculo de probabilidades condicionales inversas mediante el teorema de Bayes:

$$P(A_i | B) = \frac{P(B | A_i)P(A_i)}{\sum_{j=1}^n P(B | A_j)P(A_j)}$$

La ocurrencia del evento B está condicionada a la realización de algunos de los eventos A_i que se constituyen en una partición del espacio muestral. La idea es que B ocurrió y se indaga por la probabilidad de que haya ocurrido A_i . $P(B)$ es la probabilidad a posteriori en tanto que $P(A_i)$ es la probabilidad a priori; los valores $P(A_i)$ son las verosimilitudes o probabilidad de que B ocurra cuando ha ocurrido A_i .

El teorema de Bayes y los elementos que lo fundamentan han sido objeto de muchas investigaciones en didáctica de la probabilidad que dan cuenta de algunas de las dificultades que presentan los estudiantes en la comprensión de estos temas. Kahneman y Tversky (1972) mostraron que las personas no tienen en cuenta las probabilidades a priori en el cálculo de la probabilidad inversa. Los estudiantes confunden el evento condicionante con el condicionado, es decir confunden $P(B)$ con la $P(A)$, dando lugar a la falacia de la condicional transpuesta (Falk, 1986). Gras y Totohasina (1995) destacan una concepción temporal que los estudiantes le adjudican a la probabilidad condicional en el sentido de que el evento condicionante B siempre precede en el tiempo al evento A . Ahora, una de las mayores dificultades en el momento de resolver situaciones problemáticas asociadas a la probabilidad condicional es la dificultad que existe para distinguir entre probabilidades condicionales y probabilidades conjuntas (Pollatsek y cols., 1987), así como en la identificación correcta de los sucesos condicionante, condicionado y en la correcta partición del espacio muestral (Díaz y de la Fuente, 2006). En una investigación reciente, Morgado y Yáñez (2013) realizada con estudiantes universitarios que realizaban un curso de estadística para conocer la evolución que sufre el razonamiento bayesiano a lo largo de un semestre, aplicaron tres pruebas en momentos diferentes durante en el semestre y encontraron que si bien los estudiantes responden mejor los problemas bayesianos inmediatamente después de haber estudiado la regla de Bayes, después de un tiempo esta capacidad se reduce a niveles iguales o inferiores a los mostrados en la primera prueba cuando no habían estudiado esos temas.

Motivados por continuar este tipo de estudios realizamos una investigación con la misma estructura del trabajo de Morgado y Yáñez (2013) pero con cuestionarios mucho más reducidos y más homogéneos en términos de los problemas propuestos. Algunos de los resultados obtenidos se presentan en este trabajo. A continuación se describe la metodología implementada para luego dar lugar a la presentación de los resultados y a las conclusiones.

2. Metodología

Se desarrolló un estudio longitudinal con estudiantes universitarios de ingeniería industrial de la Universidad Industrial de Santander que estaban realizando su primer curso de probabilidad y estadística. Se diseñaron tres pruebas con el fin de evaluar el razonamiento bayesiano de los estudiantes, aproximadamente se aplicó una prueba cada cinco semanas. Las pruebas tienen la misma estructura y constan de tres ítems: el primer ítem se resuelve aplicando el teorema de Bayes en un contexto de canales; el segundo ítem, en un contexto de máquinas, se resuelve aplicando el teorema de Bayes; el tercer ítem es un problema de

urna en un contexto de muestreo sin reposición y consta de cuatro incisos que evalúan la probabilidad condicional directa, probabilidad condicional inversa, la regla del producto para eventos dependientes y el teorema de probabilidad total. Es importante aclarar que los ítems de la primera y tercera prueba son idénticos.

La muestra constaba de 76 estudiantes de ingeniería industrial de la Universidad Industrial de Santander, Bucaramanga, Colombia, quienes estaban cursando la asignatura de estadística I en el quinto semestre de su carrera y conformaban dos grupos diferentes con profesores distintos. En los dos grupos, momentos después de estudiar el teorema de Bayes se realizó la aplicación de la segunda prueba.

3. Resultados y discusión

De los 76 estudiantes de ingeniería, 57 de ellos presentaron las tres pruebas. En la Tabla 1 se muestran los porcentajes de éxito obtenidos en cada ítem de acuerdo a su contenido y en cada una de las pruebas.

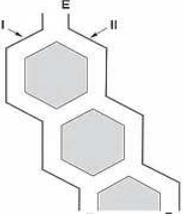
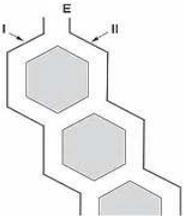
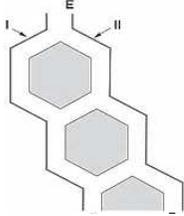
Tabla 1. Porcentaje de respuestas correctas según el contenido de los ítems en cada una de las tres pruebas.

Contexto	Contenido	Primera Prueba	Segunda Prueba	Tercera Prueba
Canales	Razonamiento bayesiano	0%	9%	7%
Fabricación de artículos por dos máquinas	Razonamiento bayesiano	2%	75%	56%
Urna sin reposición	Probabilidad condicional	61%	54%	53%
	Razonamiento bayesiano	14%	2%	4%
	Regla del Producto	11%	40%	44%
	Teorema de Probabilidad Total	4%	11%	23%

Como era de esperarse, el ítem de mejores resultados desde un comienzo fue el que indagaba por una probabilidad condicional directa en un modelo de urna. El hecho de que fueran mejores los resultados de la primera prueba (61%) indica que esta situación tiene un carácter altamente intuitivo para los estudiantes y que esta intuición puede disminuir como consecuencia de la instrucción. En sentido opuesto se comportó el problema de las máquinas donde ya el influjo de la instrucción se vio reflejado en el alto porcentaje de buenas respuestas en la segunda prueba (75%) y, aunque bastante menor, también en la tercera (56%). Ahora bien, las cosas son bien diferentes en el problema de los canales, donde los resultados no fueron nada buenos en ninguna de las tres pruebas: 0%, 9% y 7%

respectivamente, lo que conduce a pensar que el problema no es ni intuitivo ni susceptible de ser comprendido cabalmente en un proceso de enseñanza clásico. Respecto a los otros conceptos evaluados, la regla del producto y el teorema de probabilidad total, la situación fue bien diferente: los porcentajes de éxito aumentaron a lo largo del tiempo. Estos hechos pueden ser debidos al proceso de instrucción que para estos temas se mostró muy efectivo, aspecto que se refleja sobre todo en el teorema de probabilidad total que inicialmente solo obtuvo 4% de resultados positivos en tanto que la misma pregunta tres meses después fue respondida exitosamente por el 23% de los estudiantes.

En el ítem de los canales, concebido para evaluar el razonamiento bayesiano, donde los resultados fueron muy pobres, las respuestas dadas por la mayoría de los estudiantes en la primera y tercera prueba fue $\frac{1}{2}$, identificando el problema con una probabilidad conjunta, donde el segundo evento (salir por el canal R) tiene probabilidad 1 (ver la Figura 1). En la tercera prueba, algunos estudiantes construyeron un diagrama de árbol pero fueron incapaces de utilizarlo para obtener la probabilidad requerida (ver la Figura 3), tal como ya lo habían reportado Morgado y Yáñez (2013).

<p>Figura 1. Respuesta de un estudiante al ítem 1 de la primera prueba.</p>	<p>Figura 2. Respuesta de un estudiante al ítem 1 de la segunda prueba.</p>	<p>Figura 3. Respuesta de un estudiante al ítem 1 de la tercera prueba.</p>
<p>Una bola se suelta por la entrada E. Si sale por R, ¿Cuál es la probabilidad de que haya pasado por el canal 1?</p>  <p>Probabilidad = $\frac{\text{Camino que tomé}}{\text{Total caminos}}$ Probabilidad = $\frac{1}{2} = 0,5$ 50% de probabilidad de que haya pasado por el canal 1.</p>	<p>Una bola se suelta por la entrada E. Si sale por R, ¿Cuál es la probabilidad de que haya pasado por el canal 1?</p>  <p>$P(I) = 0,5$ $P(II) = 0,5$ $P(R I) = 0,5$ $P(R II) = 0,5$ $P(R) = 0,5 + 0,75 + 0,125 = 0,875$ $P(I) = \frac{0,5}{0,875} = 0,5714$</p>	<p>Una bola se suelta por la entrada E. Si sale por R, ¿Cuál es la probabilidad de que haya pasado por el canal 1?</p>  <p>50% I → R 50% 50% II → R 25% 50% II → B 25%</p> <p>$P(I R) = \frac{P(I)P(R I)}{P(I)P(R I) + P(II)P(R II)}$ $P(I R) = 0,5$</p>

En el problema de urna cuando se indaga por la probabilidad de obtener una bola blanca en la segunda extracción, típica aplicación del teorema de probabilidad total, se observa que los estudiantes que no responden acertadamente utilizan la estrategia del “depende” ya reportada por Yáñez (2003) y que refleja la concepción del enfoque del resultado aislado (Konold, 1991).

4. Conclusiones

Esta investigación muestra cómo la instrucción afecta las concepciones de los estudiantes y cuánto se sostiene este efecto en el tiempo. Sin duda, cuando los problemas que implican un razonamiento bayesiano son equivalentes con el diagrama de árbol tal como son los problemas de máquinas, los resultados de la instrucción fueron positivos. Otra cosa muy distinta sucedió cuando los problemas son claramente intuitivos, como es el caso de la probabilidad condicional directa en el modelo de urna: los resultados fueron superiores en el inicio del curso respecto a lo sucedido después de su enseñanza.

Algo muy especial ocurrió con los problemas de canales que tuvieron un resultado deficiente. Parece ser que este tipo de problemas figurativos donde la información está contenida en el gráfico, representan un reto para los estudiantes que no terminan de asimilarlos como situaciones de probabilidad condicional inversa que requieren un razonamiento bayesiano. El estudio del razonamiento bayesiano en este tipo de situaciones amerita mayor investigación.

Desde otro punto de vista se reafirman algunos de los resultados encontrados en Morgado y Yáñez (2013) y reportados por otros investigadores como la asimilación de una probabilidad condicional directa con una probabilidad conjunta, confusión que se manifestó en todas las pruebas realizadas y que prueba que es muy difícil de erradicar. La falta de una adecuada coordinación entre el diagrama de árbol y la regla de Bayes también se evidenció en las dos últimas pruebas: el árbol solo tiene la función de registrar la información dada por el problema y no conduce a su solución. La concepción temporal de la probabilidad condicional (Gras y Totohasina, 1995) también se mantuvo a lo largo del tiempo.

Referencias bibliográficas

- Díaz, C. De la Fuente, I. (2006). Dificultades en la resolución de problemas que involucran el teorema de Bayes, un estudio exploratorio en estudiantes españoles de psicología. *Educación matemática*, 18, 75-94.
- Falk, R. (1986). Conditional probabilities: insights and difficulties. En R. Davidson y J. Swift (Eds.), *Proceedings of the Second International Conference on Teaching Statistics*. University of Victoria. 292 – 297.
- Gras, R. y Totohasina, A. (1995). Chronologie et causalité, conceptions sources d'obstacles épistémologiques à la notion de probabilité conditionnelle. *Recherches en Didactique des Mathématiques*. 15 (1), 49 – 95.
- Kahneman y Tversky (1972) Subjective probability: A judgment of representativeness. *Cognitive Psychology* 3, 430–454.
- Konold, C. (1991). Understanding Students' Beliefs about Probability. En E. Von Glasersfeld (ed.), *Radical Constructivism in Mathematics Education*, Kluwer Academic Publishers, 139-156.
- Pollatsek, A., Well, A. D., Konold, C. y Hardiman, P. (1987). Understanding conditional probabilities. *Organization, Behavior and Human Decision Processes*, 40, 255 – 269.
- Morgado, C., y Yáñez, G. (2013). El cambio en el razonamiento bayesiano de estudiantes universitarios durante un semestre académico. *Revista Científica*, 0(EXTRA), 159-162. Recuperado de: <http://revistas.udistrital.edu.co/ojs/index.php/revcie/article/view/4729>.
- Yáñez, G. (2003). Estudios sobre el papel de la simulación computacional en la comprensión de las secuencias aleatorias, la probabilidad y la probabilidad condicional. Tesis doctoral no publicada. Centro de Investigación y de Estudios Avanzados del Instituto Politécnico Nacional, México.