

# RECURSOS PARA LA ENSEÑANZA DEL RAZONAMIENTO BAYESIANO EN INTERNET

*Carmen Díaz Batanero, [mcdiaz@ugr.es](mailto:mcdiaz@ugr.es)*

*Inmaculada de la Fuente, [edfuente@ugr.es](mailto:edfuente@ugr.es)*

*Facultad de Psicología, Universidad de Granada*

*Congreso Internacional: El Profesorado ante el reto de las Nuevas Tecnologías en la Sociedad del Conocimiento. Departamento de Didáctica y Organización Escolar. Universidad de Granada. Granada, Marzo 2005*

## Summary

Bayesian reasoning is highly relevant for a variety of professionals; however its teaching is not always wide enough to help students overcome the obstacles related to this topic (e.g. base rate fallacy or time-axis fallacy). In this paper we analyse some Internet resources that might facilitate the teaching of Bayesian reasoning, making a special emphasis in its didactic usefulness.

## Resumen

El razonamiento bayesiano es de gran relevancia para todo tipo de profesionales, sin embargo su enseñanza no es en ocasiones lo suficientemente amplia como para superar los obstáculos que presenta este tema (e.g., falacia de las tasas base o falacia del eje temporal). En este trabajo se analizan algunos recursos disponibles en internet para facilitar su enseñanza, haciendo especial énfasis en su utilidad didáctica.

## Introducción

El cálculo de la probabilidad a posteriori  $P(A_i/S)$  de la posible causa  $A_i$  que ha producido un suceso  $S$ , conocidas las probabilidades a priori de las causas  $P(A_i)$  así como de que dichas causas originen el suceso  $P(S/A_i)$  se puede llevar a cabo mediante el teorema de Bayes, que generaliza a toda una rama de la inferencia estadística, conocida como inferencia bayesiana. Este tipo de razonamiento es de gran relevancia para todo tipo de profesionales, como psicólogos, médicos, economistas o políticos, quienes lo encuentran en situaciones de toma de decisión bajo incertidumbre, como el diagnóstico, la evaluación o la investigación empírica. A pesar de esta importancia, encontramos en la actualidad que no se hace suficiente énfasis en la

enseñanza de este tipo de razonamiento, e incluso se encuentran recomendaciones para disminuir los contenidos de probabilidad en los cursos de estadística dirigidos a universitarios. Las razones dadas es que el razonamiento bayesiano es difícil para el alumno, como lo muestran muchas investigaciones en el campo de la psicología.

Pero otras investigaciones recientes sugieren que la enseñanza del razonamiento bayesiano podría facilitarse si se utiliza un formato adecuado de presentación de los problemas y se enseña a los alumnos estrategias de resolución basadas en la representación de diagramas en árbol. En este trabajo sugerimos el uso de recursos disponibles en Internet, que podrían facilitar la adquisición de estas estrategias y la comprensión de los conceptos latentes tras ellas. Todo ello como parte de un trabajo más amplio, en el que pretendemos diseñar estrategias didácticas de introducción de la inferencia bayesiana a alumnos de psicología.

### **Investigaciones sobre razonamiento bayesiano**

La competencia en resolución de problemas bayesianos ha sido investigada, ampliamente por psicólogos (e.g., Tversky y Kahneman, 1982; Falk, 1986). La conclusión es que estos problemas son difíciles y contraintuitivos y que los sujetos no tienen en cuenta las probabilidades a priori en el cálculo de la probabilidad inversa (falacia de las tasas bases). También los educadores matemáticos (e.g., Gras y Totomasina, 1995; Ojeda, 1995) analizan los procedimientos de los estudiantes, sugiriendo la dificultad en construir correctamente un diagrama en árbol o en identificar los datos del problema. Junto a esto se presentan algunas concepciones erróneas, como por ejemplo, pensar que no se puede condicionar un suceso por otro que ha ocurrido con posterioridad a aquél cuya probabilidad calculamos (falacia del eje de tiempos).

Por otro lado, ciertas teorías recientes (ver resumen en Sedlemeier, 1999) sugieren que se puede facilitar la resolución de los problemas bayesianos y llegar a una enseñanza efectiva si los datos del problema se presentan en frecuencias absolutas, en lugar de en porcentajes o por medio de probabilidades. También indican la conveniencia de usar una representación adecuada, como el diagrama en árbol. En nuestro propio trabajo (Díaz, 2004)

hemos encontrado grandes dificultades en el razonamiento bayesiano de alumnos de psicología, incluso después de la enseñanza del tema y con problemas dados en formatos de frecuencia. Todo ello nos indica la necesidad de replantear la enseñanza y nos ha sugerido el uso de algunos recursos disponibles en Internet que analizamos a continuación.

### **Internet y enseñanza de la estadística**

La estadística es una de las materias que más ha sido influida por la tecnología, y en particular por Internet (Galmacci, 1996), a partir de las contribuciones de muchas personas e instituciones en todo el mundo, la mayoría de las cuales no buscan un beneficio económico directo. Destacamos la página de la *International Association for Statistical Education*, <http://www.stat.auckland.ac.nz/~iase/> donde se recogen vínculos a otros servidores, así como publicaciones, incluyendo tesis doctorales, actas de conferencias y revistas electrónicas, como *Statistics Education Research Journal*.

Estas posibilidades, nos llevan a replantear la enseñanza de la estadística (Hawkins, 1997), ya que los libros de texto se empiezan a transformar a ediciones electrónicas y libremente accesibles a la consulta, modificación y sugerencias a través de Internet. Es también sencillo obtener datos de todo tipo para que los estudiantes puedan realizar investigaciones sobre casi cualquier tema, incluso con pocos recursos disponibles. El profesor puede cargar estos conjuntos de datos de la Internet e introducirlos en las calculadoras gráficas de los alumnos que tienen una difusión mucho mayor. De este modo los alumnos pueden trabajar con los datos en casa o exportarlos a otros ordenadores o calculadoras. También pueden combinar diferentes conjuntos de datos en un mismo proyecto o "enviar" a la red sus propias colecciones de datos para que sean usadas por nuevos estudiantes en cualquier rincón del planeta.

Las listas de discusión entre profesores o entre alumnos, la "tutoría" de alumnos a distancia, cuando el trabajo del alumno no permite la comunicación directa con el profesor ya están siendo implementados en muchas escuelas y universidades. La importancia de estos recursos se reflejó en la organización

de una conferencia estadística sobre “*Statistics Education and Internet*” en el año 2003 (<http://www.ph-ludwigsburg.de/iase/> ).

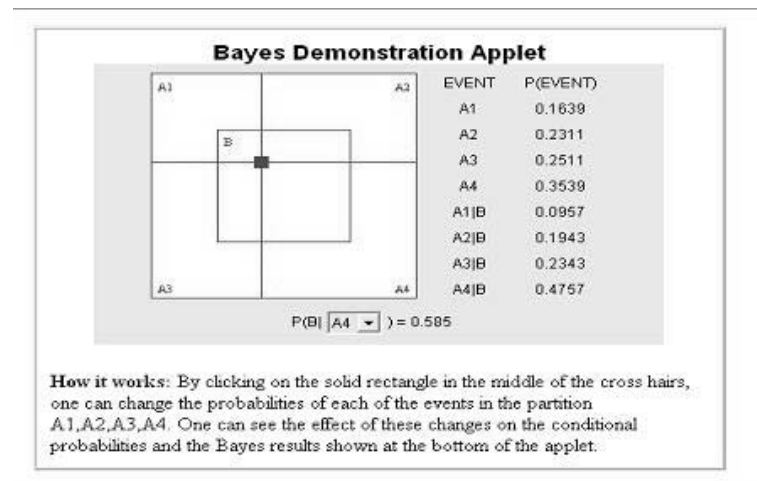
### Recursos en Internet para la enseñanza del razonamiento bayesiano

A continuación describimos algunos applets y recursos disponibles en Internet para facilitar la enseñanza del razonamiento bayesiano, clasificados en varios apartados, describiendo su utilidad para la enseñanza.

### Exploración del teorema de Bayes

En la dirección <http://www.stat.sc.edu/~west/applets/bayesdemo.html> encontramos un applet que permite explorar este teorema. El recurso (Figura 1) muestra un diagrama de rectángulo con cuatro particiones del espacio muestral ( $A_1$ ,  $A_2$ ,  $A_3$  y  $A_4$ ) y un suceso B. Las probabilidades de los sucesos de la partición se pueden modificar moviendo un cursor, modificando también las probabilidades condicionales  $P(B/A_i)$ . Este programa calcula automáticamente las probabilidades  $P(A_i/B)$ , aunque no muestra cómo se hace el cálculo sino tan sólo el resultado.

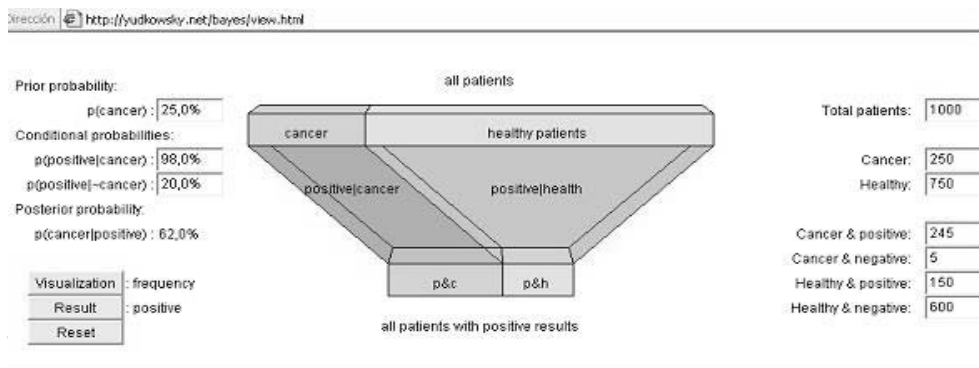
Figura 1. Pantalla del Bayes demonstration Applet



La Figura 2 muestra un ejemplo de la aplicación del teorema de Bayes al diagnóstico (Bayes applet: <http://yudkowsky.net/bayes/view.html>). Permite definir la probabilidad a priori de tener cáncer así como las probabilidades de que un test dé positivo si se tiene cáncer y probabilidad de que dé positivo si no se tiene cáncer. Según se definan estos sucesos, el applet modifica el

diagrama donde se representan las personas con cáncer, personas sanas, personas que dan positivo si tienen cáncer y personas que dan positivo si no lo tienen. El programa calcula la probabilidad de tener cáncer si el test da positivo.

Figura 2. Pantalla del applet de diagnóstico



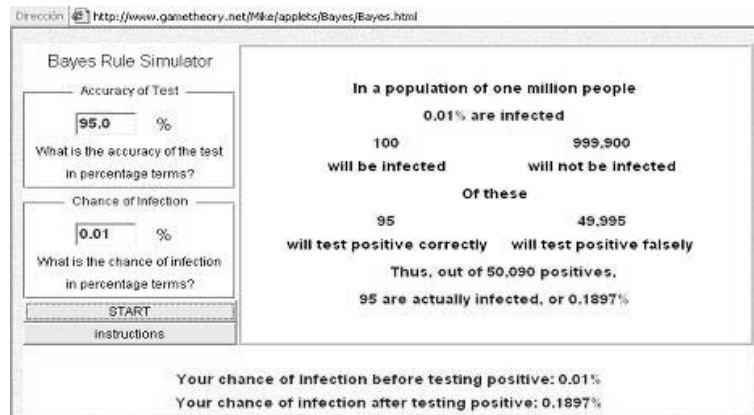
En la parte derecha del applet se incluyen los mismos datos en frecuencias: número de pacientes, personas con cáncer, personas sanas, personas que tienen cáncer y dan positivo, personas que están sanas y dan positivo, etc. Si se introduce un cambio en las probabilidades, el applet calcula los datos en frecuencias y viceversa. Por último, en la parte inferior izquierda del applet hay tres botones: el primero permite visualizar el diagrama en formato de frecuencia o de probabilidades; el segundo permite cambiar el problema para el caso en el que el test dé negativo; y el tercero permite restaurar los valores por defecto.

Un recurso similar (Figura 3) es el Bayes rule simulator: <http://www.gametheory.net/Mike/applets/Bayes/Bayes.html>). Plantea la situación en que alguien es sometido a una prueba para detectar una enfermedad rara y resulta positivo, y pide reflexionar sobre si debería preocuparse esta persona. Hay que definir la precisión de la prueba y la probabilidad de infección.

Cuando presionamos el botón "start", el programa calcula la probabilidad de tener la enfermedad si la prueba da positivo. Un aspecto interesante de este recurso es que representa el procedimiento de cálculo mediante un diagrama de árbol, donde se incluyen los datos en frecuencias y probabilidades. De esta forma el alumno puede ver la secuencia de razonamiento que ha seguido el

programa para llegar a la solución. En la parte inferior, el applet incluye la probabilidad a priori y a posteriori de tener la enfermedad, para poder compararlas.

*Figura 3. Pantalla del Bayes Rule Simulator*



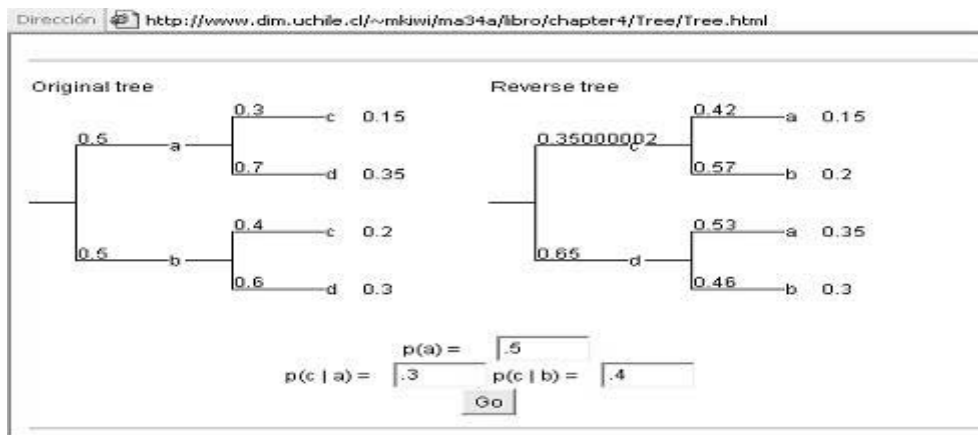
Este mismo autor ha diseñado otro applet muy similar, aunque en otro contexto. En este caso, trata de una empresa que quiere recompensar a los buenos empleados y calcula la probabilidad de que la recompensa haya ido a parar a un mal empleado. Podemos encontrarlo en <http://www.gametheory.net/Mike/applets/Bayes/WhoReward.html>.

### Construcción de árboles de probabilidad

Ya hemos indicado que el diagrama en árbol facilita la resolución de estos problemas, aunque tampoco son sencillos de construir. Bayes tree (<http://www.dim.uchile.cl/~mkiwi/ma34a/libro/chapter4/Tree/Tree.html>) está pensado para enseñar esta tarea. Presenta un diagrama de árbol y unas casillas donde se pueden introducir las probabilidades a priori  $P(A)$  y las probabilidades condicionales  $P(C/A)$  y  $P(C/B)$ . Una vez introducidos los datos, se presiona el botón "go" y el programa calcula las probabilidades conjuntas  $P(A \cap C)$ ,  $P(A \cap D)$ , etc. El applet también incluye el árbol inverso que permite calcular la probabilidad condicional inversa  $P(A/C)$ .

Otro enlace donde se incluye applets para la construcción de diagramas de árbol es: <http://www.stat.sc.edu/~west/applets/tree.html> . Tras introducir los datos en una tabla de contingencia, el programa muestra el diagrama de árbol correspondiente y calcula las probabilidades conjuntas.

Figura 4. Diagrama de árbol representado en el applet



### Software para el cálculo bayesiano

La aplicación de la inferencia bayesiana se ha facilitado bastante con el software disponible. Por ejemplo, un sencillo applet disponible en la página [http://www.amstat.org/publications/jse/secure/v8n1/p\\_discrete.html](http://www.amstat.org/publications/jse/secure/v8n1/p_discrete.html) (Figura 5) que permite hacer la estimación de una proporción mediante inferencia bayesiana, calculando su distribución a posteriori.

Figura 5. Applet para la estimación de una proporción

LEARNING ABOUT A PROPORTION  
BY BAYES RULE

- Enter your prior distribution in the PRIOR column. Note that it may be entered as whole numbers.
- Enter your data (number of successes and failures) in the SUCCESSES and FAILURES boxes.
- Push the UPDATE button to compute the posterior probabilities.
- In the table, the Likelihood column contains the probability of observing the data for each value of p, normalized so that the largest likelihood is equal 10000.

|      | Prior                          | Likelihood           | Product              | Posterior            |
|------|--------------------------------|----------------------|----------------------|----------------------|
| p=0  | <input type="text" value="1"/> | <input type="text"/> | <input type="text"/> | <input type="text"/> |
| p=.1 | <input type="text" value="1"/> | <input type="text"/> | <input type="text"/> | <input type="text"/> |
| p=.2 | <input type="text" value="1"/> | <input type="text"/> | <input type="text"/> | <input type="text"/> |
| p=.3 | <input type="text" value="1"/> | <input type="text"/> | <input type="text"/> | <input type="text"/> |
| p=.4 | <input type="text" value="1"/> | <input type="text"/> | <input type="text"/> | <input type="text"/> |

En la pantalla aparece una tabla de once filas que corresponden a los valores de la proporción ( $p=0$ ,  $p=.1$ ,  $p=.2$ ,..., hasta  $p=1$ ) y cuatro columnas: probabilidad a priori, verosimilitud, producto y probabilidad a posteriori. Tenemos que introducir nuestra distribución a priori (introduciendo la probabilidad para cada proporción en la columna "prior") así como los datos observados (número de éxitos y fracasos). Cuando pulsamos el botón "update", el programa completa la tabla, calculando las verosimilitudes y la distribución a posteriori. Recursos similares se descargan en la dirección <http://members.aol.com/johnp71/bayes.html> (Figura 6) y también en <http://www.dim.uchile.cl/~mkiwi/ma34a/libro/chapter4/Bayes/Bayes.html>. En el primer caso se pueden plantear hasta cinco hipótesis, permite poner una etiqueta a cada hipótesis y cada uno de los sucesos, se establecen las probabilidades a priori, las verosimilitudes y se marca el suceso que ha ocurrido. Cuando pulsamos el botón "compute" te calcula la probabilidad a posteriori.

*Figura 6. Applet para calcular probabilidades a posteriori*

| Hypotheses:                                | Portador | No portador | Hyp 3 | Hyp 4 | Hyp 5 |
|--|----------|-------------|-------|-------|-------|
| Prior Probabilities:                       | 0.5      | 0.5         | 0     | 0     | 0     |
| <input type="checkbox"/> Hemofilia         | 0.5      | 0           | 0     | 0     | 0     |
| <input checked="" type="checkbox"/> Normal | 0.5      | 1           | 0     | 0     | 0     |
| <input type="checkbox"/> Outcome 3:        | 0        | 0           | 0     | 0     | 0     |
| <input type="checkbox"/> Outcome 4:        | 0        | 0           | 0     | 0     | 0     |
| <input type="checkbox"/> Outcome 5:        | 0        | 0           | 0     | 0     | 0     |
| Revised Prob:                              | 0.333    | 0.667       | 0.000 | 0.000 | 0.000 |

En internet, además hay disponibles de forma gratuita versiones de evaluación de algunos programas bayesianos como el programa "Le Bayesian" desarrollado por el equipo de Bruno Lecoutre: <http://www.univ-rouen.fr/LMRS/Persopage/Lecoutre/LeBayesianPlusenglish.htm>. En el enlace podemos encontrar varias pequeñas aplicaciones que permiten el cálculo de distribuciones de probabilidad, probabilidades predictivas y estimación de



proporciones, entre otros. Otro software disponible en internet es el BUGS, que además dispone de una página web muy completa, con manuales y ejemplos:

<http://www.mrc-bsu.cam.ac.uk/bugs/welcome.shtml>.

En otros enlaces, como, por ejemplo, en el servidor <http://espse.ed.psu.edu/espse/hale/507Mat/statlets/free/pdist.htm> y en <http://www.stat.vt.edu/~sundar/java/applets/Distributions.html#BETA> facilitan el cálculo con funciones beta y otras muchas funciones de distribución como la binomial, Poisson, geométrica, Chi-cuadrado, t de Student o normal.

### **Información general**

Por último, en internet podemos encontrar páginas que porcionan tutoriales o artículos con nociones básicas sobre la inferencia bayesiana como en

[http://www.nebulasearch.com/encyclopedia/article/Bayesian\\_inference.html](http://www.nebulasearch.com/encyclopedia/article/Bayesian_inference.html),

<http://www.seh-lelha.org/bayes1.htm> o en las páginas personales de Jim Albert:

<http://www-math.bgsu.edu/~albert/> o Bruno Lecoutre: <http://www.univ-rouen.fr/LMRS/Persopage/Lecoutre/>.

Las personas interesadas en la investigación sobre este tema o en sus aplicaciones prácticas pueden encontrar interesante la página de la Intenational Society for Bayesian Analysis <http://www.bayesian.org/> con enlaces a revistas específicas como Bayesian Analysis (<http://ba.stat.cmu.edu/>) y otros recursos.

### **Conclusiones**

En este artículo hemos descrito algunos ejemplos de las múltiples posibilidades que brinda Internet para la enseñanza de razonamiento bayesiano. Puesto que el número de personas y organizaciones interesadas por el tema es cada día mayor, la lista de lugares de la red que hemos presentado habrá quedado superada ampliamente cuando el artículo sea publicado.

La disponibilidad de recursos libremente accesibles en Internet hace que la cultura y la ciencia se estén democratizando cada vez más, por lo que el aprendizaje se lleva a cabo no sólo en el aula tradicional. El uso de este tipo de recursos aumenta la motivación de los alumnos por el tema, ya que se

presentan los conceptos de una forma más llamativa y permite al alumno adoptar un papel activo en su aprendizaje. Es por ello importante que el profesor tenga en cuenta estos recursos y los incorpore a su enseñanza.

Sin embargo, un recurso didáctico por sí sólo no resuelve todos los problemas. Se plantea, así el reto de diseñar unidades didácticas para la enseñanza del razonamiento bayesiano, que incorpore estos recursos.

**Nota:** El trabajo es parte del Proyecto BSO2002-0334-7 y Beca FPU: AP2003-5130.

### Referencias

- Díaz, C. (2004). *Elaboración de un instrumento de evaluación del razonamiento condicional. Un estudio preliminar*. Trabajo de investigación tutelado. University of Granada.
- Falk, R. (1986). Conditional Probabilities: insights and difficulties. En R. Davidson y J. Swift (Eds.), *Proceedings of the Second International Conference on Teaching Statistics*. (pp. 292 - 297). University of Victoria.
- Galmacci, G. (2001). The impact of Internet on the researchers' training. In C. Batanero (Ed.), *Training researchers in the use of statistics* (pp. 159-169). Granada: International Statistical Institute.
- 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.
- Hawkins, A. (1997). How far have we come? Do we know where we are going? En E. M. En E. M. Tiit (Ed.), *Computational statistics & statistical education* (pp. 100-122). Tartu: International Association for Statistical Education e International Association for Statistical Computing.
- Ojeda, A. M. (1995). Dificultades del alumnado respecto a la probabilidad condicional. *UNO*, 5, 37-55.
- Sedlmeier, P. (1999). *Improving statistical reasoning. Theoretical models and practical implications*. Mahwah, NJ: Erlbaum.
- Tversky, A, y Kahneman, D. (1982). Evidential impact of base rates. En D. Kahneman P. Slovic y A. Tversky (Eds.), *Judgment under uncertainty: Heuristics and biases* pp. 153-160).. New York: Cambridge University Press.