UNIVERSITAT POLITÈCNICA DE VALÈNCIA

Un enfoque Bayesiano de la Probabilidad

Apellidos, nombre	Boigues Planes, Francisco José[1] (fraboipl@mat.upv.es) Estruch Fuster, Vicente D.[1] (vdestruc@mat.upv.es)
Departamento	[1]Matemática Aplicada
Centro	Universitat Politècnica de València

1 Resumen de las ideas clave

Presentamos un enfoque **bayesiano** de la probabilidad donde la probabilidad de un suceso va variando en la medida en que van apareciendo nuevas evidencias relacionadas con el suceso. Partiremos de nociones básicas de la teoría de probabilidades que permitirán construir conceptos típicos del enfoque bayesiano, como probabilidad a priori y probabilidad a posteriori. Además, usaremos procedimientos gráficos para consolidar este enfoque en los esquemas cognitivos de los lectores.

2 Introducción

Una de las nociones fundamentales de la estadística inferencial es la de probabilidad. Es obvio que, al introducir la probabilidad, tenemos que hacer referencia a la Regla de Laplace: casos favorables dividido por casos totales. Pero también, al enfoque frecuencialista (Boigues y Estruch, 2017), que asocia a cada suceso una probabilidad cuyo valor es la frecuencia relativa del suceso, después de experimentar muchas veces, el cual ofrece un enfoque interesante en el campo de la ingeniería. Otro enfoque interesante de la probabilidad es el enfoque bayesiano, donde la probabilidad de un suceso puede variar en función de la aparición de evidencias que modifiquen las circunstancias iniciales. Si queremos que los estudiantes adquieran fundamentos sólidos en el aprendizaje de la probabilidad, tenemos que conseguir que incorporen a su esquema probabilístico las diferentes perspectivas de la probabilidad.

Cuando se estudia el Teorema de Bayes, algunos conceptos relacionados con el mismo, como por ejemplo las nociones de probabilidad condicionada o la probabilidad conjunta, suelen presentar problemas serios de comprensión. Por otra parte, existen trabajos didácticos que indican la existencia de sesgos en el razonamiento con la probabilidad condicional (Batanero, Ortiz y Serrano, 2016). Comprender la probabilidad condicional es esencial para, en su caso, una correcta aplicación del teorema de Bayes en muchos procesos de modelización matemática. Por lo tanto, es importante encontrar dispositivos de enseñanza que puedan ayudar a los estudiantes a comprender mejor estos conceptos (Bárcenas et al. 2017; Díaz, Batanero y Contreras, 2010)

3 Objetivos

Al finalizar este artículo, un estudiante debe ser capaz de asociar probabilidades a sucesos desde una perspectiva bayesiana.

4 Desarrollo

Comencemos repasando las nociones básicas asociadas al enfoque bayesiano (Newbold, Carlson y Thorne, 2010; Estruch, Gregori y Sapena, 2014) acompañadas con un ejemplo fácil de comprender: ***Lanzamos un dado al azar. ¿Cuál es la probabilidad de obtener un seis?*** Si se aplica la Regla de Laplace (ecuación 1) tendremos que la probabilidad de sacar un seis al lanzar un dado será igual al número de casos favorables, que es un caso, dividido por los casos totales o posibles que son 6, es decir, 1/6= 0'1$\hat{6}$≈16'67%, si A es el suceso "*salir un seis*" tenemos que $P(A)$= 1/6, es lo que se denomina, en el enfoque bayesiano, la probabilidad "*a priori*".

$$P(A)= \frac{Casos\ favorables\ a\ A}{Casos\ totales}$$

Ecuación 1. Regla de Laplace

En cambio, si nos dicen que ***en el lanzamiento del dado se sabe que ha salido un número par,*** que, denotaremos como el suceso B, entonces la probabilidad de A condicionada a B sería 1/3, ya que el espacio muestral se ha restringido a {2,4,6}. Dicha probabilidad se suele denotar como P(A|B) = 1/3. Es decir, la probabilidad "a priori" del suceso A ha aumentado ante la evidencia que supone conocer el suceso B. Esta nueva probabilidad suele denominarse probabilidad "a posteriori".

La probabilidad "*a posteriori*", o condicionada, se puede obtener a partir de la igualdad descrita en la ecuación 2.

$$P(A|\mathrm{B})= \frac{p(A \cap B)}{p(B)}$$

Ecuación 2. Probabilidad condicionada

que, aplicada al ejemplo del dado, lleva a otra manera de calcular la probabilidad "*a posteriori*"

$$P(A|B) = \frac{p(A \cap B)}{p(B)} = \frac{1/6}{3/6} = \frac{1}{3}$$

Por otra parte, si despejamos la probabilidad de la intersección en la ecuación 1, obtenemos

$$P(A \cap B) = p(A|B) \cdot p(B)$$

Puesto que los sucesos se comportan como conjuntos, se tiene que $B \cap A = A \cap B$ y, por lo tanto

$$P(A \cap B) = P(B \cap A) = P(B|A) \cdot P(A)$$

Ecuación 3. Probabilidad "conjunta"

A partir de las ecuaciones 2 y 3 se deduce lo que se conoce como Teorema de Bayes

$$P(A|B) = \frac{P(B|A) \cdot P(A)}{P(B)}$$

Ecuación 4. Teorema de Bayes

El teorema de Bayes, se puede generalizar atendiendo a lo que se denomina *partición del espacio muestral*, E, es decir, un conjunto de sucesos $\{A_1, A_2, \ldots, A_n\}$, tales que $\forall i, j \in \{1,2, \ldots, n\},\ i \neq j,\ A_i \cap A_j = \emptyset$ y $\bigcup_{i=1}^{n} A_i = E$. Entonces se tiene que

$$P(A_i|B) = \frac{P(A_i \cap B)}{P(B)} = \frac{P(B|A_i) \cdot P(A_i)}{P(B) \cdot P(B|A_1) \cdot P(A_1) + \cdots + P(B|A_n) \cdot P(A_n)}$$

Ecuación 5. Teorema de Bayes generalizado

4.1 Un enfoque gráfico del teorema de Bayes

Supongamos la siguiente cuestión:

Disponemos de dos monedas. Una de ellas (M1) es normal, es decir, con su correspondiente cara y su cruz y además es no sesgada (al lanzarla, las dos opciones de la moneda tienen las mismas posibilidades de salir). La otra moneda (M2) tendría dos caras. Si consideramos el suceso de elegir una moneda, de las dos, al azar, sabemos que la probabilidad de elegir M1 sería de un 50%. En un enfoque bayesiano la probabilidad, P(M1), sería considerada como una probabilidad a priori. Una vez elegida la moneda al azar, y sin conocer cuál es la moneda elegida, la lanzamos, y obtenemos cara, ¿Podemos mantener que el valor de P(M1) sigue siendo el mismo? o lo que es lo mismos, ¿Ha variado el valor de P(M1)? Observemos que, en este caso, se tiene una evidencia en el experimento (haber obtenidos cara al lanzar la moneda). *En este caso, denotaremos la probabilidad como P(M1|C) con C= {salir cara}, lo que se considera una probabilidad a posteriori.*

Para obtener la probabilidad a posteriori, siguiendo a Erickson (2017), se construye la figura 1, que representa el fenómeno aleatorio descrito. Se trata de un cuadrado/matriz 2x2, donde, además, se consideran dos partes iguales, las columnas. Una de estas partes, la sombreada, representa a M1 y la otra a M2. Además, cada una de estas partes se ha vuelto a dividir en dos partes iguales, las filas, que representan las opciones que pueden ocurrir al lanzar la correspondiente moneda.

Moneda Normal (M1)	Moneda con dos caras (M2)
C	C
X	C

Figura 1. Gráfico síntesis del fenómeno "lanzar una vez la moneda elegida"

Si nos fijamos en el total de caras de la figura 1, aparecen 3, de las cuales solo una (sombreada) corresponde a la moneda normal (M1). Por lo tanto, P(M1|C) = 1/3. Es decir, la probabilidad a posteriori ha disminuido respecto a la inicial o a priori.

4.2 El teorema de Bayes

Veamos cómo resolver a la cuestión planteada desde una perspectiva más formalista a través del teorema de Bayes. En este fenómeno aleatorio, se elige en primer lugar una moneda de un universo que consta de dos posibles resultados, {M1, M2}. Dadas las características de las monedas se tiene que *P*(*C*|*M1*) = ½ y *P*(C|M2)=1. Se lanza la moneda elegida y obtenemos una cara (C) y se nos pide que calculemos p(M1/C). aplicando la ecuación 5 tenemos:

$$p(M1|C) = \frac{p(C|M1)\cdot p(M1)}{p(C|M1)\cdot p(M1) + p(C|M2)\cdot p(M2)} = \frac{\frac{1}{2}\cdot\frac{1}{2}}{\frac{1}{2}\cdot\frac{1}{2}+1.\frac{1}{2}} = \frac{\frac{1}{4}}{\frac{1}{4}+\frac{1}{2}} = \frac{\frac{1}{4}}{\frac{3}{4}} = \frac{1}{3}$$

4.3 Simulando con una hoja de Cálculo

Otra manera de aproximarnos a la probabilidad pedida sería emplear un método de simulación, repitiendo la experiencia aleatoria muchas veces. Con la ayuda de EXCEL, pasamos a realizar la experiencia 400 veces.

En primer lugar, simulemos la elección de una moneda de entre un universo formado por las dos monedas *M1* y *M2*. EXCEL dispone de la función **=ALEATORIO.ENTRE(x;y)** que nos da un valor aleatorio entero entre el **x** y **y**, con la misma posibilidad para cada valor. Por lo tanto, **=ALEATORIO.ENTRE(1;2)**, generará valores 1 y 2, con la misma probabilidad, igual a 0,5. Salir un 1 significaría elegir la moneda M1 y salir un 2 seria elegir la moneda M2.

- *Nos situamos en una celda, A1 y escribimos* **=ALEATORIO.ENTRE(1;2)**
- *Copiar la fórmula anterior hasta la celda A400*

A continuación, lanzamos la moneda elegida. Si la moneda es *M1* puede salir una cara, en cuyo caso asignamos un 0 o una cruz, caso en que asignaremos un 1. En cambio, si la moneda elegida es M2 siempre sale cara al ser lanzada, por lo cual asignaremos un 2.

- *Nos situamos en una celda, B1 y escribimos* **=SI(A1=1;ALEATORIO.ENTRE(0;1);2)**
- *Copiar la fórmula anterior hasta la celda A400*

Contamos las veces que ha salido un 1 y las veces que ha salido un 2 en la columna B, ya que la probabilidad pedida *P*(*M1*/*C*) = Nº de veces que sale 1(Nº de caras de la moneda M1) dividido por la suma de total de caras (Nº de veces de 1 + Nº de veces de 2)

- *Escribir en la celda E1* **Nº de Caras de M1**
- *Escribir en la celda F1* ***=CONTAR.SI(B1:B400; 0)***
- *Escribir en la celda E2* **Nº de Caras de M2**
- *Escribir en la celda F2* ***=CONTAR.SI(B1:B400;2)***
- *Escribir en la celda E3* **p(M1/C):**
- *Escribir en la celda F2* ***=F1/(F1+F2)***

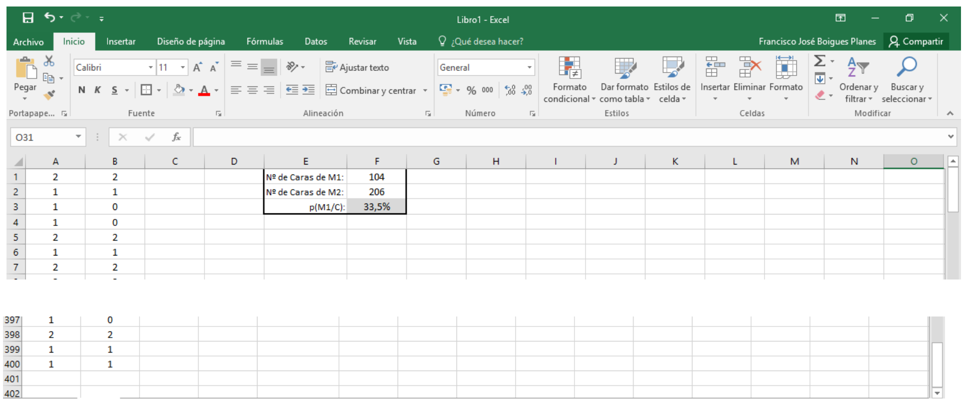

Figura 2. Experimentar 400 veces con una hoja de Cálculo

La probabilidad que obtenemos en la simulación es 0,319, que es un valor próximo a 1/3. Recuérdese que al experimentar de manera aleatoria los resultados de la simulación serán distintos pero muy similares.

4.4 Complicando la cuestión inicial

Supongamos que se ha lanzado dos veces la moneda elegida y nos ha salido cara en las dos ocasiones. ¿Cuál sería la probabilidad de haber elegida la moneda "normal"?

Vamos utilizar el enfoque gráfico para resolver esta cuestión.

Un primer enfoque gráfico sería partir de la figura 1 done teníamos dividido un cuadro en dos partes iguales. Cada parte representaba una moneda. La parte de la moneda 1 la teníamos sombreada. Además cada parte de cada moneda la dividíamos en dos partes iguales que representaba las diversas opciones que podría acontecer al lanzar dicha moneda al azar. Ahora lo que vamos a hacer es dividir cada opción en dos partes iguales y ponemos las opciones que puede acontecer al lanzar por segunda vez la moneda elegida. El gráfico se puede ver en la figura 3.

Moneda Normal (M1)		Moneda con dos caras (M2)	
CC	CX	CC	CC
XC	XX	CC	CC

Figura 3. Grafico síntesis del fenómeno "lanzar dos veces la monea elegida"

Si nos fijamos en el total de opciones con 2 caras contabilizamos 5, de las cuales, solo una está sombreada. Consecuentemente, p(M1|CC)= 1/5

Otro posible enfoque sería considerar las probabilidades a priori en cada lanzamiento. En el primer problema, teníamos que la probabilidad "a priori" de elegir la moneda M1 era de ½ y ante la evidencia de haber sacado una cara al lanzar la moneda elegida, vimos que la probabilidad de elegir *M1* disminuía a 1/3. Vamos a suponer ahora que la probabilidad "a priori" de elegir *M1* es de 1/3 y suponemos que al lanzarla volvemos a obtener cara. ¿Cuál es la probabilidad de haber elegido la moneda normal *(M1)*?

Construyamos ahora un nuevo cuadrado que dividiremos en 2x3=6 partes (rectángulos) iguales. En concreto, lo dividimos verticalmente, en dos partes: una parte (izquierda) que represente a M1|C, con un 1/3 del cuadrado inicial y el resto, 2/3 del cuadrado inicial, a M2|C. Sombrearemos la zona de M1. A continuación, dividiremos el cuadrado, horizontalmente, en 2 partes iguales, que representarán las dos posibilidades que se tienen para cada moneda al ser lanzada (figura 4).

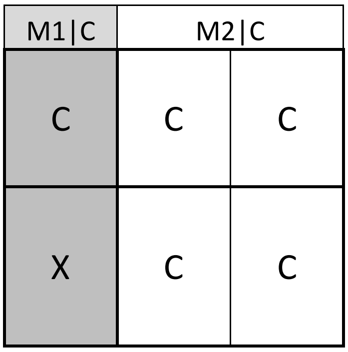

Figura 4. Gráfico del fenómeno lanzar una moneda, con probabilidad a priori de 1/3 para una de ellas.

Fijémonos en la figura 3, de los 5 rectángulos con cara, solo 1 está sombreados. Por tanto, la probabilidad $P(M1\,|\,C \cap C) = 1/5$.

Las probabilidades halladas también se pueden obtener calculando el área que abarca cada opción. Tenemos que el área de las opciones en las que aparece "C", suponiendo que partimos de un cuadrado 1x1, sería 5/6 y el área sombreada de las "C" es de 1/6, por tanto

$$p(M1|C) = \frac{1/6}{5/6} = \frac{1}{5}$$

5 Resumiendo

Veamos los posibles enfoques de resolución de un problema clásico del teorema de Bayes que suele verse en cualquier curso de estadística:

El 60% de los tornillos producidos por una fábrica proceden de la máquina A y el 40% de la maquina B. El porcentaje de tornillos defectuosos producidos por la máquina A es 5% y en cambio por la máquina B es 10%. ¿Cuál es la probabilidad de que, sabiendo que un tornillo es defectuoso, proceda de la máquina A? O preguntada de otra manera, ¿Cuál es el porcentaje de tornillos defectuosos que provienen de la máquina A?

Comencemos aplicando la perspectiva clásica del teorema de Bayes. Definamos los sucesos: A={tornillos producidos por la máquina A}; B={tornillos producidos por la máquina B}; D={ser defectuoso}. De los cuales conocemos las siguientes probabilidades: P(A)= 0'6, p(B)= 0'4, p(D|A)= 0'05, p(D|B)= 0'1. Al ser {A, B} una partición del espacio muestral podemos aplicar el teorema de Bayes generalizado (ecuación 5):

$$p(A|D) = \frac{p(D|A) \cdot p(A)}{p(D|A) \cdot p(A) + p(D|B) \cdot p(B)} = \frac{0'05 \cdot 0'6}{0'05 \cdot 0'6 + 0'1 \cdot 0'4} = \frac{0'03}{0'07} = 0'4667 = 46'67\%$$

Continuemos con el enfoque gráfico discreto. Partimos de un cuadrado de 10x10 que lo dividimos en dos partes: un 60% para identificar la producción de A (que lo sombrearemos) y la otra parte para B. De la parte A son defectuosos (D) el 5% de 60, es decir, 0'05·60= 3, en cambio de B tenemos un 10% de 40, es decir, 0'1·40= 4

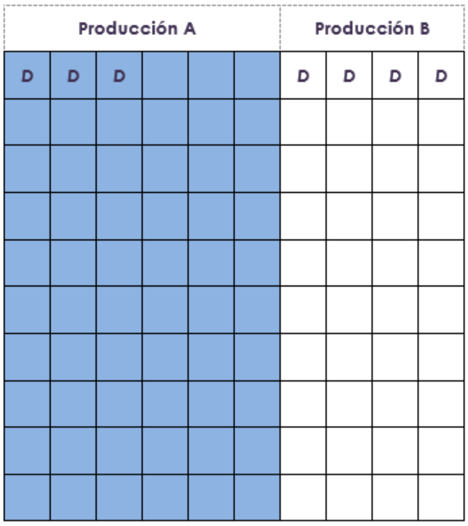

Figura 5. Gráfico discreto del Problema clásico de Bayes

Fijémonos en la figura 5, de los 7 rectángulos con defecto (D), hay 3 sombreados. Por tanto, la probabilidad $P(A|D)$ = 3/7=0'4667=46'67%.

Finalmente veamos el enfoque continuo, mediante el cálculo de áreas (ver gráfica 6): Partimos de un cuadro 1x1 que lo dividimos en dos partes. La parte A cuya área será del 0'6 y la de B 04. El área defectuosa de A es el 5% de 0'6 es decir, 0'05·0'6= 0'03. En cambio, el área defectuosa de B es el 10% de 0'4= 0'1·0'4=0'04.

Figura 6. Gráfico continuo del Problema clásico de Bayes

Tenemos que área total de defectuoso= 0'04+0'03= 0'07, siendo elárea defectuosa en A= 0'03, por tanto

$$p(A|D) = \frac{0'03}{0'07} = 0'4667 = 46'67\%$$

6 Algunas conclusiones

Hemos visto el teorema de Bayes desde la perspectiva formal matemática. A continuación, se ha desarrollado la visión gráfica pero discreta y por último la perspectiva gráfica continua a través de áreas. Integrar los diferentes enfoques bayesianos ayudará al alumnado a ser más eficaz a la hora de resolver problemas, puesto que se podrán comprender mejor los diferentes elementos que acompañan al teorema de Bayes. Las aproximaciones expuestas permiten dotar al alumno con un mayor número de herramientas para un aprendizaje conceptual y práctico más sólidos.

Con el objetivo de comprobar el nivel y la solidez en la comprensión del teorema de Bayes, se propone el siguiente problema de consolidación:

Planteamos el problema con un lanzamiento más, es decir, se ha lanzado tres veces la moneda elegida y nos ha salido cara en las tres ocasiones. ¿Cuál sería la probabilidad de haber elegida la moneda "normal" ($M2$)?

Se deja como ejercicio al lector resolver gráficamente este problema, siguiendo un razonamiento análogo al planteado en el caso anterior. El resultado ha de ser $P(M2|C \cap C \cap C) = 8/9$.

7 Bibliografía

BÁRCENA, M.J.; ARACELI, M.; MARTÍN, A.; TUSELL, F. Y UNZUETA, A. (2017). "Un simulador para asistir en la enseñanza del teorema de Bayes". Editores V. Botti y M.A. Fernández, *Actas congreso INRED* (pp.15-23). Valencia.

BOIGUES, F.J. Y ESTRUCH, V.D. (2017). "Aproximación frecuencialista de la Probabilidad". *Colección articulo docente de la UPV*. http://hdl.handle.net/10251/82991

DÍAZ, C.; BATANERO, C. Y CONTRERAS,J.M. (2010). "Teaching independence and conditional probability". *Boletín de Estadística e Investigación Operativa,* nº. 26 (2), 149-162.

ERICKSON, T. (2017). "Beginning Bayes". *Teaching Estatistics Trust,* 39, 1, 30-35.

ESTRUCH, V.; GREGORI V. Y SAPENA A. (2014) "Lecciones de estadística".. Editorial UPV, Valencia. Ed. 2ª

NEWBOLD, P., CARLSON W.L. Y ; THORNE, B. M. (2010). *Statistics for business and economics.* Pearson Education International. 7th ed