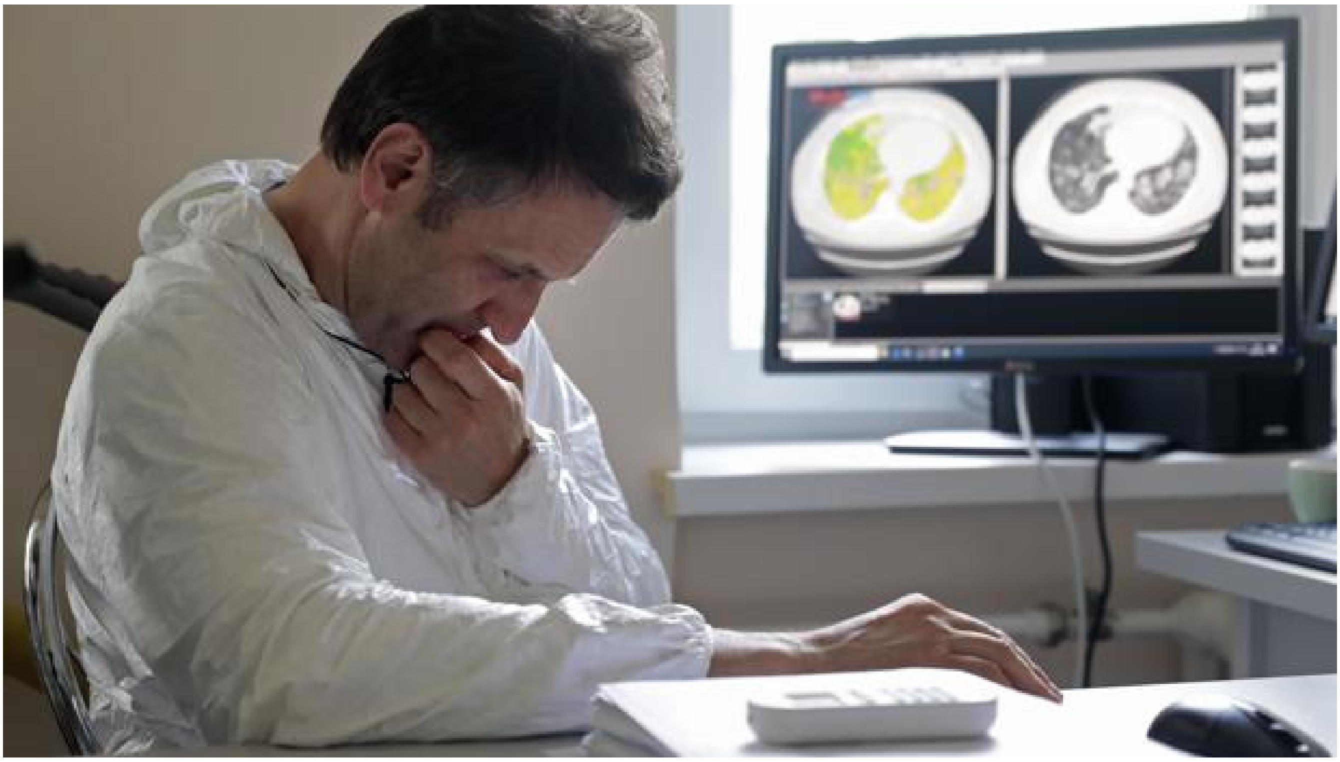
ABC, 15 de Junio de 2020
CIENCIA - El ABCdario de las matemáticas
Alfonso Jesús Población Sáez

Cómo podemos aplicar un famoso teorema a los diagnósticos médicos



EFE

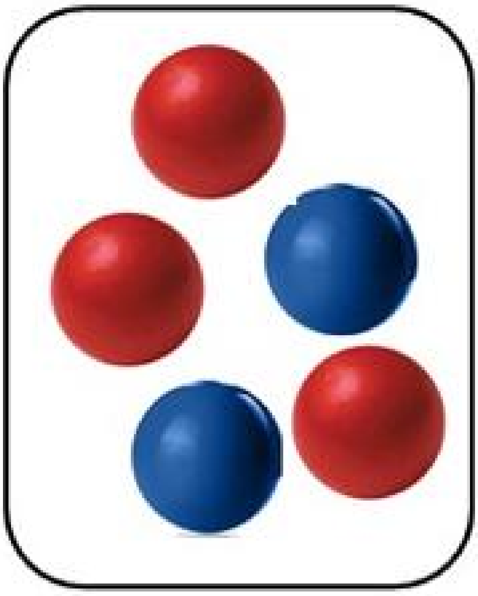
Como seguramente hayan leído más de una vez, el **cálculo de probabilidades** tiene su origen en los intentos de entender y predecir el comportamiento de los juegos de azar (los juegos son tan antiguos como la propia humanidad), y más concretamente los de apuestas. En el siglo XVI dicha preocupación coge cierto auge (se ve que la gente quiere enriquecerse rápido, aunque la mayoría lo que logra en realidad es el efecto contrario, es decir, arruinarse) y se pregunta a los matemáticos.

Sus explicaciones teóricas, en muchas ocasiones, no satisfacen a los jugadores, y menos aún cuando observan que hay ciertas contradicciones con lo que dice la teoría y lo que sucede experimentalmente en las mesas de juego (es célebre el problema del caballero de Méré , que sin duda conocerán). El caso es que no sólo a los jugadores sino también a los matemáticos puros no acaba de convencerles esta nueva faceta de las matemáticas, ese intento de dominar los fenómenos aleatorios, que consideran imposible. Pero, ya entrado el siglo XVII, matemáticos de cierto renombre van proponiendo procedimientos para dotar de rigurosidad a estos asuntos. A pesar de ello, sigue habiendo reticencias (los puristas consideran la matemática como una ciencia exacta, perfecta, donde no caben estimaciones o porcentajes de acierto).

Con estos precedentes, en 1763 se publica de manera póstuma la obra de **Thomas Bayes** (1702 – 1761), donde desarrolla el famoso **Teorema de Bayes** , germen de la **Inferencia Bayesiana** , una importante rama de la Estadística actual. Es un resultado que probablemente (nunca mejor utilizado el adverbio) recuerden más como una fórmula

$$p(A \mid B) = \frac{p(B|A)\, p(A)}{p(B)}$$

que nos permite calcular la probabilidad condicionada, siendo A y B dos sucesos, con B con probabilidad positiva (para que no se anule el denominador). La p(A|B) denota la probabilidad de que suceda A, sabiendo que ha sucedido B. Dicho de otra manera, tener cierta información, ¿modifica la probabilidad de que ocurra un evento? La respuesta es afirmativa, pero hay que saber utilizar bien la expresión anterior porque, en ocasiones, podemos sacar conclusiones equivocadas.



Primero veamos un ejemplo sencillo de cómo puede quedar afectada una probabilidad teniendo cierta información adicional. Supongamos que tenemos una bolsa cerrada y opaca en la que hay tres bolas rojas y dos azules. Sacamos dos bolas. La probabilidad de que la segunda sea roja es 3/5, es decir, 0.6 (¿Sabría el lector por qué?). Pero si sabemos que la primera fue azul, ¿cuál es la probabilidad de que la segunda sea roja? En este caso, como podemos simular todos los casos mediante un sencillo esquema, no haría falta recurrir a la formula anterior (aunque se recomienda al lector que lo intente a ver si le sale lo mismo), comprobando que si la primera fue azul (es decir, quedarían en la bolsa una azul y tres rojas), entonces que la segunda sea roja lo tendríamos con probabilidad de 3/4, es decir, 0.75, la posibilidad de extraer roja en segundo lugar aumenta mucho.

Otro ejemplo típico y muy difundido ya que aparece en series de televisión (*Numb3rs*, episodio 1x13.- *La*

caza del hombre
), películas (
21 Blackjack
, Robert Luketic, EE. UU., 2008) y hasta novelas (
El curioso incidente del perro a medianoche
, Mark Haddon) es el de la llamada
paradoja de Monty Hall
, estrategia habitual de los concursos de televisión (como nuestro célebre
Un, dos, Tres
), en el que el presentador insta al concursante a que decida si mantenerse o cambiar de elección de posible regalo después de haberle mostrado alguna de las opciones iniciales.

Sin embargo, entre los propios matemáticos esta fórmula tampoco estuvo exenta de controversias, distinguiéndose dos grupos: los seguidores de la estadística tradicional que solo admiten cálculos con probabilidades basadas en experimentos repetibles y con confirmación empírica, y los llamados estadísticos bayesianos que aceptan probabilidades "subjetivas", es decir, con información adicional de los sucesos. Un conocido ejemplo de sus objeciones es la llamada **paradoja del cuervo** (o **paradoja de Hempel**, planteada por el filósofo alemán Carl Hempel (1905 – 1997)). Antes de enunciarla, generalicemos la fórmula de Bayes, citada antes para solo dos eventos.

Si tenemos un conjunto de eventos mutuamente excluyentes $\{A_1, A_2, \ldots, A_n\}$ de los cuales necesariamente ocurre uno (o sea, la unión de todos ellos es el espacio muestral completo), y tales que la probabilidad de cada uno de ellos es distinta de cero, y B un suceso cualquiera con probabilidad positiva del que se conocen las probabilidades condicionales $p(B|A_k)$, para todos los k desde 1 a n, entonces, la probabilidad condicionada $p(A_k|B)$ viene dada por la expresión:

$$p(A_k \mid B) = \frac{p(B|A_k)\ p(A_k)}{p(B)}$$

Aunque lo que comento a continuación pueda parecer sumamente técnico (de eso que la gente cree que sólo nos gusta y entienden los matemáticos), es precisamente la idea por la que hay que tener cuidado al aplicar la fórmula de Bayes, y que justifica perfectamente por qué no funciona en el caso de la paradoja del cuervo. Como los A_k hemos indicado que son mutuamente excluyentes y necesariamente ocurre uno, resulta que

$$B = (A_1 \cap B) \cup (A_2 \cap B) \cup \ldots \cup (A_n \cap B)$$

Sólo uno de los sucesos de los paréntesis (al ser una intersección, quiere decir que ocurren a la vez A_k y B) sucede, por lo que la probabilidad total es una suma:

$$p(B) = \sum_{k=1}^{n} p(B|A_k)\, p(A_k)$$

Esta expresión es útil en la práctica porque suele ser más fácil calcular todas las $p(B|A_k)$ que la p(B). Y así se conforma la fórmula de Bayes tal y como la están estudiando estos días los alumnos que se presentan a la EBAU:

$$p(A_k \mid B) = \frac{p(B|A_k)\, p(A_k)}{\sum_{k=1}^{n} p(B|A_k)\, p(A_k)}$$

Seguramente lo entendamos perfectamente con un sencillo ejemplo. Volvamos con los temas médicos (lugar frecuente en el que se utiliza la fórmula de Bayes; seguimos ahondando en que este sector profesional precisa las matemáticas de un modo claro), retomando el hilo de la reseña de la semana pasada. Supongamos que la probabilidad de que un paciente tenga los

síntomas del COVID-19
teniendo la enfermedad es alta, de 0.9 (los números son inventados, no responden a la realidad, sólo los coloco a efectos de entender la fórmula de Bayes y observar su efecto).

Expresado en los términos anteriormente descritos, el evento B será el de tener los síntomas, mientras que A_1 es padecer la enfermedad, y A_2 no tenerla. Los sucesos mutuamente excluyentes en este caso son claros: tener la enfermedad, por un lado, y no tenerla, por otro. Y la unión de ambos es el conjunto de los pacientes en total. El dato anterior sería entonces la probabilidad condicionada $p(B|A_1) = 0.9$. Si el paciente no tiene la enfermedad, la probabilidad de que tenga síntomas será baja, pongamos que 0.15. Conocemos asimismo la prevalencia de la enfermedad, es decir, el porcentaje de la población que ha desarrollado la enfermedad, digamos que ha sido de un 20%. En términos de probabilidades (ya saben, no podemos mezclar churras con merinas, tampoco en matemáticas), ese porcentaje corresponde a $p(A_1) = 0.2$. Por tanto, $p(A_2) = 0.8$.

La probabilidad de que alguien tenga los síntomas, p(B), es la unión de los sucesos tener el COVID y los síntomas, más la probabilidad de que alguien no tenga el COVID, pero sí los síntomas. Entonces

$$p(B) = p(A_1 \cap B) + p(A_2 \cap B) = p(B|A_1)\, p(A_1) + p(B|A_2)\, p(A_2)$$
$$= 0.9 \cdot 0.2 + 0.15 \cdot 0.8 = 0.3$$

Entonces,

$$p(A_1 \mid B) = \frac{0.9 \cdot 0.2}{0.3} = 0.6$$

es decir, de todos los pacientes con síntomas, detectamos correctamente el COVID en el 60% de los casos.

Cuevos, Sherlock Holmes y Bayes

Aunque es muy conocido y difundido, no me resisto a recordar la **paradoja del cuervo**. Tomamos una muestra de un millón de cuervos. Los observamos y establecemos la conclusión de que “Todos los cuervos son negros”. La conclusión parece razonable de acuerdo con el principio de inducción. Sin embargo, de acuerdo a las reglas de la lógica, cuando tenemos una proposición P implica otra Q, es equivalente a decir que, si Q es falsa, entonces P es falsa. Aplicado a la sentencia anterior eso significa que “Todas las cosas que no son negras, no son cuervos”. Quiere esto decir que cualquier objeto verde, rojo, o de cualquier color que no sea negro que encontremos refuerza la idea de que los cuervos son negros. Es extraño porque el conjunto de todas las cosas que no son negras es en comparación muchísimo mayor que el conjunto de los cuervos. Sin embargo, si utilizamos la fórmula de Bayes, observando numerador respecto al denominador, un conjunto de objetos pequeño, apenas influye en el resultado final del otro evento.

En la novela *El signo de los cuatro*, Arthur Conan Doyle habla del arte de la deducción y pone en la boca de su famosa creación, el lógico Sherlock Holmes, la siguiente afirmación: “Cuando han sido descartadas todas las explicaciones imposibles, la que queda, por inverosímil que parezca, tiene que ser la verdadera”. Como el principio de **reducción al absurdo**, tan empleado en las demostraciones matemáticas. Sin embargo, hay algo que también la fórmula de Bayes nos advierte: ¿es posible contemplar todas las posibles situaciones A_k? ¿Son acaso mutuamente excluyentes? Tanto en esta situación como en la paradoja del cuervo los conjuntos son enormes, inabarcables (las posibles explicaciones a un crimen y los objetos no negros) y eso es lo que nos lleva a las paradojas y los absurdos. Igual que una medicación tiene sus especificaciones y sus incompatibilidades, administrar un teorema o una proposición matemática requiere tener muy en cuenta sus hipótesis, porque si no, podemos “cargarnos al paciente”.

Matemática Española (RSME) ... *Real Sociedad*