

El teorema de Bayes con nomogramas

GABRIEL RUIZ-GARZÓN
JAIME RUIZ-ZAPATERO

El objetivo de este artículo es resolver problemas de aplicación del teorema de Bayes utilizando nomogramas. La Nomografía calcula valores de funciones mediante el empleo de tablas gráficas llamadas nomogramas. Esta ciencia fue profusamente utilizada por ingenieros militares durante la Primera Guerra Mundial y olvidada con la irrupción de los ordenadores. El objetivo es introducirla como un recurso didáctico interesante y novedoso en nuestras aulas para mostrar el teorema de Bayes de una forma diferente a la tradicional.

Palabras clave: Recursos didácticos, Nomografía, Teorema de Bayes, Bachillerato y Universidad.

Bayes' Theorem with nomograms

The aim of this article is to solve problems encountered during the application of Bayes' Theorem through the usage of nomograms. Nomography calculates function values by using graphical tables named nomograms. This science was profusely used by military engineers during World War I to be forgotten after the irruption of computers. This paper purposes to reintroduce this tool as a novel an effective way to explain Bayes' Theorem in our classroom.

Keywords: Educational resource, Nomography, Bayes' Theorem, Secondary School and University.

En 1908, el onubense, D. Nicomedes Alcayde y Carvajal (1871-1930), coronel de ingenieros y profesor de la Academia de Artillería, publicó su *Cálculo de Probabilidades*, obra que sería declarada libro de texto en las Academias de Artillería e Ingenieros del Ejército, al igual que en la de Ingenieros y Maquinistas de la Armada y por la que se le concedería, en recompensa por su elaboración, la Cruz de 1.ª clase del Mérito Militar con distintivo blanco. Se publicó, como el propio autor indica en el prólogo, con el ánimo de ayudar a los alumnos que deben tomar apuntes ante la falta de libros de texto adecuados y «al pasar dichos apuntes de unas a otras manos se producen tantos disparates de copia, que llegan a constituir un inconveniente» (Alcayde, 1942).

Alcayde resuelve, utilizando la probabilidad, una serie de problemas de balística que podrían presentarse a un artillero o ingeniero militar en el frente de combate, como es el que sigue:

Ejemplo 1

Si tomamos como origen de coordenadas el centro de impactos de una pieza artillera, podemos señalar dos paralelas a ambos ejes donde se concentren el 50% de impactos de una manera longitudinal y transversal. En las tablas de tiro que suelen acompañar a una arma particular figura que, a la distancia supuesta, el 50% de los impactos caen dentro de una zona de anchura $z=80$ m. Se trataría de determinar el tanto por ciento

de impactos que caerán en una zona centrada en el objetivo de anchura diferente, en este caso de $z' = 16$ m. y de esa manera calcular la probabilidad de que el blanco sea alcanzado.

Matemáticos franceses como Isidore Didion (1798-1878), profesor de la Escuela Militar de Metz, habían efectuado diversos experimentos y conseguido modelizar la cantidad de impactos que caen en una cierta región del plano a través de una distribución normal (Didion, 1858), de ahí que en el libro de Alcayde apareciera junto a una tabla de la distribución normal, una tarjeta denominada *nomograma* que permite la resolución de problemas como el anterior de una manera gráfica.

El nomograma o ábaco permitía al artillero, en el fragor de la batalla, calcular de una manera gráfica sencilla el tanto por ciento de impactos que recibirá el objetivo, sin más que unir, con una regla o hilo, tres puntos de igual número de escalas, como las que mostramos en la figura 2 y que después pasaremos a detallar.

El propio Alcayde publicaría una obra titulada *Elementos de Nomografía* (Alcayde, 1915), donde explicaría los secretos de esta disciplina. La Nomografía se podría definir como el cálculo de valores de funciones mediante el empleo de tablas gráficas, los nomogramas.

En 1912 también inventó una regla llamada «Regla de Alcayde» o de las «Academias Militares», que en 1915 patentó y que fue fabricada por Nes- tler, un constructor alemán. Se trata de una regla

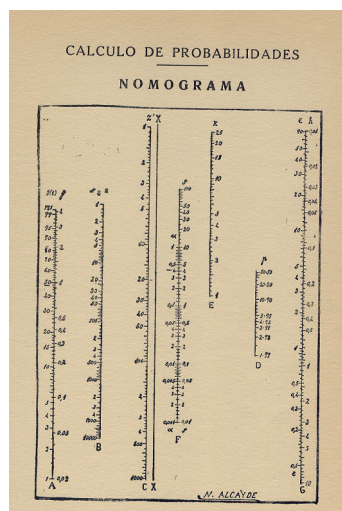


Figura 2. Nomograma de Alcayde

donde figuran los valores de $\sin\alpha \cos\alpha$, $\cos^2\alpha$ y escalas de números (logaritmos de distancias), de gran utilidad para los topógrafos en la nivelación de terrenos y el levantamiento de planos. En 1928, recibiría la medalla de la Legión de Honor de la República francesa por su contribución a la puesta en servicio del túnel de Somport.

Por otra parte, en 1701 había venido al mundo Thomas Bayes. Hijo de Joshua Bayes y Anne Cotton. Su padre fue uno de los primeros seis ministros no-conformistas ordenados públicamente en Londres en 1694, es decir, religiosos protestantes pero que se oponían a la iglesia oficial anglicana. Sus ideas religiosas le privaron de la oportunidad de acceder a universidades de prestigio inglesas como Cambridge u Oxford, que establecían como requisito previo a sus alumnos la aceptación de las normas provenientes de la iglesia anglicana. Uno de sus profesores quizás fuera De Moivre. En 1736, Thomas Bayes publicó, bajo el pseudónimo de John Noon, la obra titulada *Una introducción a la doctrina de las fluxiones, y una defensa de los matemáticos frente a las objeciones del autor del Analista*. Este trabajo era una defensa de Newton frente al obispo Berkeley que había escrito otro titulado *El Analista*, criticando las bases del cálculo diferencial. Bayes fue admitido en la Royal Society en 1742. No se casó y tras heredar la fortuna de su padre, se retiró del ministerio en 1749, muriendo en 1761, estando ac-



Figura 1. Isidore Didion



Figura 3. Tumba de Thomas Bayes en Bunhill Fields (Londres)

tualmente enterrado en el cementerio londinense de Bunhill Fields.

A su muerte, su amigo y también reverendo no-conformista Richard Price (1723-1791), fue requerido por la familia Bayes para examinar los papeles que el matemático inglés había escrito y legado sobre diferentes temas. Tras una revisión concienzuda, que le llevó dos años, Price remitió a la Royal Society la comunicación de Bayes titulada *Essay towards solving a problem in the doctrine of chances*, ya que según sus palabras: «tiene gran mérito y merece ser preservada». En ella aparecerá el renombrado teorema de Bayes en su versión continua, siendo la versión discreta (que es la que se enseña en las aulas universitarias de hoy), obra de Pierre Simon Laplace. Digamos que Price tuvo «gran ojo clínico» ya que pocos artículos han marcado el devenir de la estadística

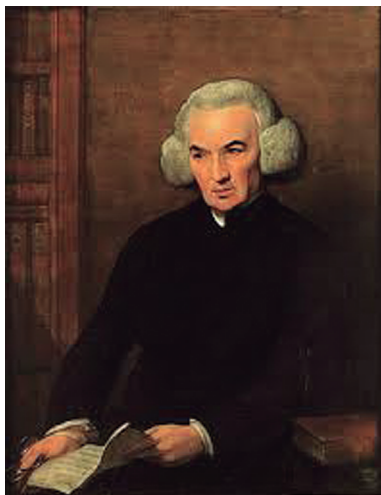


Figura 4. Richard Price

creando toda una rama de las matemáticas, el análisis bayesiano.

Hoy en día, el teorema de Bayes, en particular, es utilizado por los médicos en el desempeño de la práctica clínica al objeto de valorar si una persona presenta una enfermedad una vez se ha obtenido el resultado de una prueba clínica, es decir:

Ejemplo 2

La prevalencia de cierta enfermedad en una población de 5000 personas es del 2%. Con una prueba o test se detectaron que 70 de 100 pacientes enfermos obtuvieron un resultado positivo en el test y en el grupo de control de las 4900 personas que no tenían la enfermedad 76 dieron positivo. Calcular la proporción de pacientes con resultados positivos que tienen la enfermedad.

El objetivo de este artículo es resolver problemas de aplicación del teorema de Bayes utilizando nomogramas. La Nomografía tuvo su apogeo en el primer tercio del siglo XX y fue desplazada por la llegada de los ordenadores, pero todavía se pueden encontrar artículos recientes donde se estudia (Fagan, 1975; Marasco y otros, 2011), ya que permite una valoración rápida del estado del paciente allí donde los ordenadores no llegan, como ocurre aún hoy en día en muchos lugares del Tercer Mundo. Pero no solo es interesante para ingenieros, artilleros o médicos, sino que tiene una aplicación didáctica interesante en nuestras clases de matemáticas. Hace unos años, en los primeros números de esta misma revista *Suma* (Arenzana y otros, 1990), los autores ya proponían nomogramas como un recurso para nuestras clases de bachillerato, calculando raíces de ecuaciones bicuadradas, potencias, divisiones y probabilidades. Nosotros utilizaremos la Nomografía para mostrar el teorema de Bayes desde otra perspectiva.

Para cumplir nuestro objetivo, hemos estructurado este trabajo de la siguiente manera: en la siguiente sección veremos las nociones matemáticas para construir un nomograma de escalas paralelas que nos permita conseguir el resultado de una función como producto de otras dos. A continuación mostraremos el teorema de Bayes en su versión menos conocida, en forma de *odds*

o apuestas en contra, para posteriormente utilizar los nomogramas construidos para resolver problemas probabilísticos, ligados a la aplicación del teorema de Bayes y al cálculo artillero, y acabaremos con unas breves conclusiones.

Nomogramas

Los nomogramas son calculadoras geométricas. Existen nomogramas de diferentes escalas: en línea recta, curva o combinaciones de ambas (Parellada, 1942). Nosotros trataremos los nomogramas de escalas paralelas.

Para que una función de tres variables α , β y γ , de forma general se pueda representar mediante un nomograma de puntos alineados en tres escalas rectilíneas y paralelas, han de cumplirse unas condiciones.

Partiremos de la ecuación

$$f_1(\gamma) \varphi(\alpha) + f_2(\gamma) \psi(\beta) + f(\gamma) = 0 \quad [1]$$

Tracemos dos rectas MM' y NN' paralelas al eje OY , la primera de abscisa arbitraria a y la segunda $-a$. Sea φ una función que se coloca sobre una escala.

Sean α_1 y α_2 dos valores de la variable α y sea L la distancia que separa a $\varphi(\alpha_1)$ y $\varphi(\alpha_2)$ sobre el papel. Definimos la unidad de la escala o módulo de la escala como:

$$U = \frac{L}{\varphi(\alpha_1) - \varphi(\alpha_2)} \quad [2]$$

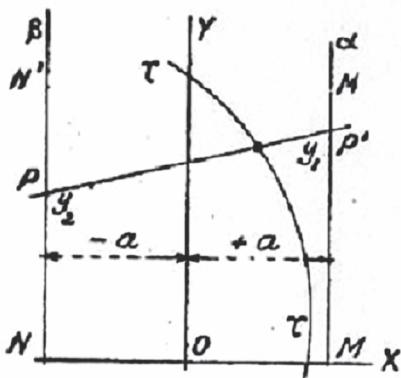


Figura 5. Diagrama de un nomograma de puntos alineados

Luego las coordenadas de todos los puntos de MM' con respecto al origen de la escala serán $(a, U\varphi(\alpha))$. Tomando V como unidad de escala en NN' , todos sus puntos tendrán de coordenadas $(-a, V\psi(\beta))$. Dado un punto cualquiera (x, y) de la recta PP' , la ecuación de dicha recta que corta a las otras dos escalas en los puntos anteriores tendrá por ecuación:

$$\frac{y - U\varphi(\alpha)}{V\psi(\beta) - U\varphi(\alpha)} = \frac{x - a}{-2a}$$

Es decir,

$$U(a+x)\varphi(\alpha) + V(a-x)\psi(\beta) - 2ay = 0 \quad [3]$$

Si los tres puntos $(a, U\varphi(\alpha))$, $(-a, V\psi(\beta))$ y $(x, f(\gamma))$ han de estar sobre una única recta PP' , esas coordenadas han de verificar las ecuaciones [1] y [3] simultáneamente, podremos, por tanto, identificar coeficientes:

$$\begin{aligned} f_1(\gamma) &= U(a+x) \\ f_2(\gamma) &= V(a-x) \\ f(\gamma) &= -2ay \end{aligned}$$

Luego las coordenadas (x, y) de los puntos de la recta PP' que satisfacen simultáneamente las ecuaciones [1] y [3] son:

$$x = a \frac{Vf_1(\gamma) - Uf_2(\gamma)}{Vf_1(\gamma) + Uf_2(\gamma)} \quad [4]$$

$$y = -\frac{f(\gamma)UV}{Vf_1(\gamma) + Uf_2(\gamma)} \quad [5]$$

En general, esos puntos formarán una línea curva pero si $f_1(\gamma)$ y $f_2(\gamma)$ son constantes independientes de γ , la línea determinada por las coordenadas [4] y [5] será una línea paralela al eje OY porque la abscisa será una constante. Y si $f_1(\gamma) = f_2(\gamma)$ resultará que la recta soporte de la escala de γ es el eje de ordenadas OY situado en el centro de las escalas de las variables α y β . Si además consideramos que $f_1(\gamma) = f_2(\gamma) = 1$, la ecuación [1] se reduce a

$$\varphi(\alpha) + \psi(\beta) + f(\gamma) = 0 \quad [6]$$

Y de igual manera, las ecuaciones [4] y [5] que nos dan las coordenadas de los puntos de la recta PP' se reducen a

$$x = a \frac{V - U}{V + U} \quad [7]$$

$$y = f(\gamma)W = f(\gamma)\frac{UV}{U+V} \quad [8]$$

Y si las unidades U y V son iguales para las escalas, todavía se pueden simplificar más las ecuaciones anteriores [7] y [8], quedando entonces como:

$$x = 0$$

$$y = f(\gamma)\frac{V}{2}$$

Con lo que el módulo para la escala de la tercera variable es la mitad del de las otras dos. Pero veamos todo esto con un ejemplo práctico.

Ejemplo 3

Construir un nomograma de escalas paralelas que permita calcular $\alpha\beta=\gamma$, en particular multiplicar tres por cinco, sabiendo que α varía de 0,1 a 10 y β de 2 a 6 y que tenemos un papel de 10 por 20 cm.

Podemos resolver $\alpha\beta=\gamma$ ya que si tomamos logaritmos obtendríamos $\log(\alpha) + \log(\beta) = \log(\gamma)$. Ecuación similar a [6] que sí admite una resolución mediante nomogramas.

Calculamos las unidades de las escalas de las variables α y β al modo de la ecuación [2]:

$$U = \frac{20}{\log(10) - \log(0,1)} = 10$$

$$V = \frac{20}{\log(6) - \log(2)} = 41,91$$

Como la anchura del papel $10 = a + b$ y $a/b = U/V$ podemos calcular la distancia de separación a y b entre las tres rectas:

$$a = \frac{10}{10 + 41,91} 10 = 1,927 \text{ cm}$$

$$b = \frac{41,91}{10 + 41,91} 10 = 8,073 \text{ cm}$$

Y el módulo de la tercera escala vendrá dado por la ecuación:

$$W = \frac{UV}{U+V} = \frac{10(41,91)}{10 + 41,91} = 8,073 \text{ cm}$$

Con lo que podríamos calcular las distancias de algunos puntos al origen de su correspondiente escala (tabla 1).

Construidas las tres escalas sobre cada una de las rectas, las subimos o bajamos hasta que podamos trazar una recta por tres puntos que satisfagan la ecuación para poder calcular los demás.

$U\varphi(\alpha)$	$Wf(\gamma)$	$U\psi(\beta)$
$10 \cdot \log(0,1) = -10$	$8,07 \cdot \log(0,2) = -5,64$	$41,91 \cdot \log(2) = 12,61$
$10 \cdot \log(1) = 0$	$8,07 \cdot \log(1) = 0$	$41,91 \cdot \log(3) = 20$
$10 \cdot \log(2) = 3,01$	$8,07 \cdot \log(10) = 8,07$	$41,91 \cdot \log(4) = 25,25$
$10 \cdot \log(3) = 4,77$	$8,07 \cdot \log(20) = 10,50$	$41,91 \cdot \log(5) = 29,29$
$10 \cdot \log(4) = 6,02$	$8,07 \cdot \log(30) = 11,92$	$41,91 \cdot \log(6) = 32,61$
$10 \cdot \log(5) = 6,99$	$8,07 \cdot \log(40) = 12,93$	
$10 \cdot \log(6) = 7,78$	$8,07 \cdot \log(50) = 13,72$	
$10 \cdot \log(7) = 8,45$	$8,07 \cdot \log(60) = 14,36$	
$10 \cdot \log(8) = 9,03$		
$10 \cdot \log(9) = 9,54$		
$10 \cdot \log(10) = 10$		

Tabla 1

Una vez construidas las tres escalas sobre cada una de las rectas podemos calcular el producto de $\alpha=5$ por $\beta=3$ simplemente observando el punto de corte de la recta que los une en la escala de la recta intermedia, que nos da $\gamma=15$.

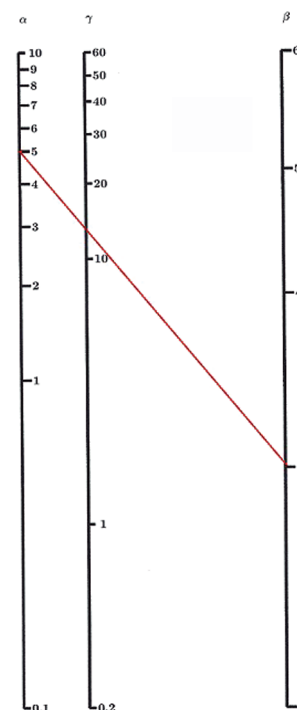


Figura 6. Nomograma para la multiplicación

El teorema de Bayes

En la probabilidad juega un papel predominante el teorema de Bayes a la hora de modificar probabilidades de acuerdo a la información recibida. El teorema de Bayes nos permite, a partir de la

prevalencia de una enfermedad (la proporción de la población que padece la enfermedad en un momento determinado), es decir, una probabilidad a priori establecida y los resultados de un test de laboratorio, calcular cuál es la probabilidad de que un paciente tenga una determinada enfermedad. El teorema nos dice cómo debemos modificar nuestras probabilidades a priori al recibir información adicional.

Partamos de una prueba dicotómica que clasifica a cada paciente como sano o enfermo en función de que el resultado del test sea negativo o positivo, respectivamente. La situación se puede modelizar en una tabla 2×2 (tabla 2).

	Sano	Enfermo	Total
Test –	Verdaderos negativos (TN)	Falsos negativos (FN)	TN + FN
Test +	Falsos positivos (FP)	Verdaderos positivos (TP)	FP + TP
Total	TN + FP	FN + TP	

Tabla 2

Pero los tests o pruebas médicas no son infalibles, sino que cometen errores. Así, se habla de la *sensibilidad* de un test como la proporción de individuos afectados que son dados como positivos correctamente por el test, es decir, es la fracción de verdaderos positivos:

$$p(+ / E) = \frac{TP}{TP + FN}$$

donde el suceso E = «tener la enfermedad» y el suceso $+$ = «dar positivo en el test». La sensibilidad de un test refleja la capacidad del test para detectar la enfermedad cuando está presente. El término tasa de *falsos negativos* hace referencia al complemento de la sensibilidad.

Tasa de falsos negativos = $1 - \text{sensibilidad} =$

$$= p(- / E) = \frac{FN}{FN + TP}$$

Por otro lado, la *especificidad* de un test es la proporción de individuos sobre todos los no afectados que son dados como negativos correctamente por el test, es decir, la capacidad de detectar a los sanos. De ahí que se llame fracción de verdaderos negativos a:

$$p(- / \bar{E}) = \frac{TN}{TN + FP}$$

El término tasa de *falsos positivos* hace referencia al complemento de la especificidad.

Tasa de falsos positivos = $1 - \text{especificidad} =$

$$= p(+ / \bar{E}) = \frac{FP}{TN + FP}$$

Como hemos comentado, una prueba interesa que tenga una alta sensibilidad y especificidad. En una persona con cáncer o tuberculosis que requiera un tratamiento urgente que podría salvar su vida, la sensibilidad de la prueba es muy importante, siendo el gran problema los falsos negativos. En una persona con una enfermedad terminal lo importante sin embargo es la especificidad, ya que tratar a un sano podría tener peores consecuencias que no tratar a un enfermo, aquí el problema serían los falsos positivos.

De manera recurrente, en los medios de comunicación aparecen noticias que alertan del elevado número de falsos positivos, que en estudios recientes sobre mamografías, entre mujeres de 40 y 50 años, se eleva al 61 %, lo que conlleva la realización de pruebas diagnósticas adicionales de elevado coste, sin que la mortalidad por cáncer de mama en mujeres con mamografía y sin mamografía sea diferente.



Figura 7. Noticia sobre falsos positivos

Ejemplo 4

Con los datos del ejemplo 3, calcular la sensibilidad y especificidad de la prueba.

Empezaremos confeccionando la tabla 3.

	Sano	Enfermo	Total
Test –	Verdaderos negativos (TN) 4824	Falsos negativos (FN) 30	4854
Test +	Falsos positivos (FP) 76	Verdaderos positivos (TP) 70	146
Total	TN + FP 4900	FN + TP 100	5000

Tabla 3

La sensibilidad es en nuestro caso:

$$p(+ / E) = \frac{TP}{TP + FN} = \frac{70}{100} = 0,70$$

Y la especificidad es de un 98% como vemos:

$$p(- / \bar{E}) = \frac{TN}{TN + FP} = \frac{4824}{4900} = 0,98$$

Observamos la importancia de la especificidad. Si en el ejemplo anterior modificamos la especificidad al 90% (tabla 4).

	Sano	Enfermo	Total
Test -	4410	30	4440
Test +	490	70	560
Total	4900	100	5000

Tabla 4

Vemos que aumentan considerablemente los falsos positivos a 490 individuos, que casi con seguridad serán tratados con medicación lo que les podría ocasionar problemas, aparte del gasto para el sistema sanitario que ello conllevaría.

Por tanto, la sensibilidad y la especificidad permiten valorar la validez de una prueba diagnóstica, nos informan de la probabilidad de obtener un resultado correcto en función del estado del enfermo. Pero desde un punto de vista clínico al médico le interesa la seguridad de la prueba, es decir, ante un resultado positivo (negativo) de la prueba, cuál es la probabilidad de que el paciente esté realmente enfermo (sano), respectivamente. A esa pregunta se responde utilizando el teorema de Bayes, podemos calcular las probabilidades a posteriori (en función de los resultados del test): el llamado Valor Predictivo Positivo y Negativo. Concretamente:

$$\text{Valor Predictivo Positivo (VPP)} = p(E / +) = \frac{p(+ / E)p(E)}{p(+ / E)p(E) + p(+ / \bar{E})p(\bar{E})} = \frac{TP}{TP + FP}$$

Es decir, el VPP es la proporción de pacientes con resultados positivos que tienen la enfermedad.

O dicho de otra manera, podríamos calcularlo en función de la sensibilidad y especificidad de la siguiente manera:

$$VPP = \frac{\text{Sensibilidad(Prevalencia)}}{\text{Sensibilidad(Prevalencia)} + (1 - \text{Especificidad})(1 - \text{Prevalencia})}$$

También podemos calcular el:

$$\text{Valor Predictivo Negativo (VPN)} = p(\bar{E} / -) = \frac{p(- / \bar{E})p(\bar{E})}{p(- / E)p(E) + p(- / \bar{E})p(\bar{E})} = \frac{TN}{TN + FN}$$

Es decir, la proporción de pacientes con pruebas negativas que han sido diagnosticados correctamente.

O dicho de otra manera:

$$VPN = \frac{\text{Especificidad}(1 - \text{Prevalencia})}{(1 - \text{Sensibilidad})\text{Prevalencia} + \text{Especificidad}(1 - \text{Prevalencia})}$$

Ejemplo 5

Con los datos del ejemplo 2 calcular los VPP y VPN.

Dados los datos tendríamos:

$$VPP = p(E / +) = \frac{TP}{TP + FP} = \frac{70}{146} = 0,4$$

$$VPN = p(\bar{E} / -) = \frac{TN}{TN + FN} = \frac{4824}{4854} = 0,99$$

El valor predictivo positivo sería del 47% y el negativo del 99%.

Los valores predictivos dependen mucho de la prevalencia como vemos en el siguiente ejemplo.

Ejemplo 6

Si la prevalencia de la enfermedad fuera del 10%, calcular los valores predictivos positivo y negativo, VPP y VPN.

En este caso, la especificidad y la sensibilidad no cambian (tabla 5):

	Sano	Enfermo	Total
Test -	4410	150	4560
Test +	90	350	4400
Total	4500	500	5000

Tabla 5

El valor predictivo positivo aumentará:

$$VPP = p(E / +) = \frac{TP}{TP + FP} = \frac{350}{440} = 0,79$$

y el valor predictivo negativo disminuirá:

$$VPPN = p(\bar{E} / -) = \frac{TN}{TN + FN} = \frac{4410}{4560} = 0,96$$

Observamos que si la prevalencia de la enfermedad es alta, un resultado positivo del test tiende a confirmar la enfermedad con un aumento del VPP . Cuando la prevalencia de la enfermedad es baja, un resultado negativo permitirá descartar la enfermedad con mayor seguridad, aumentando el valor predictivo negativo. Otros resultados no nos permitirán confirmar el diagnóstico.

Pero para tratar el teorema de Bayes a través de nomogramas necesitaremos presentarlo a través de las *odds* o apuestas en contra. En el mundo anglosajón, en el mundo de las apuestas de caballos, se habla de expresiones del tipo «6 a 3 en contra», es decir, unas *odds* de 6 a 3, quiere decir que 3 de cada 9 apostantes lo hacen por dicho caballo como vencedor, es decir, la probabilidad de que dicho caballo gane es $p(G) = 3/9$ y $p(\bar{G}) = 6/9$. Podemos ver otras aplicaciones de las *odds* al ámbito del Derecho en (Ruiz-Garzón, 2014).

De igual manera, si denotamos por O (del inglés *odds*) las apuestas expresadas en forma de fracción, tener la enfermedad y no tener la enfermedad, podemos escribir:

$$\text{Odds pretest} = \frac{p(E)}{p(\bar{E})}$$

$$\text{Odds post-test} = \frac{p(E / +)}{p(\bar{E} / +)}$$

Si las *odds pretest* son mayores que 1 entonces el suceso estar enfermo es más probable que no estarlo y al revés, si es menor que 1.

Dado el VPP :

$$p(E / +) = \frac{p(+ / E)p(E)}{p(+ / E)p(E) + p(+ / \bar{E})p(\bar{E})}$$

Si reemplazamos E por su complementario tenemos:

$$p(\bar{E} / +) = \frac{p(+ / \bar{E})p(\bar{E})}{p(+ / \bar{E})p(\bar{E}) + p(+ / E)p(E)}$$

Si dividimos ambas expresiones obtenemos:

$$\begin{aligned} & \frac{p(E / +)}{p(\bar{E} / +)} = \\ &= \frac{p(+ / E)p(E)[p(+ / E)p(E) + p(+ / \bar{E})p(\bar{E})]}{p(+ / \bar{E})p(\bar{E})[p(+ / E)p(E) + p(+ / \bar{E})p(\bar{E})]} \end{aligned}$$

Que simplificado queda:

$$\frac{p(E / +)}{p(\bar{E} / +)} = \frac{p(+ / E)}{p(+ / \bar{E})} \cdot \frac{p(E)}{p(\bar{E})}$$

que es el teorema de Bayes en forma de apuestas (*odds*), que quedaría como:

$$\begin{aligned} & \text{Odds post-test} = \\ &= \frac{\text{sensibilidad}}{(1 - \text{especificidad})} \cdot \text{Odds pretest} \end{aligned}$$

Al primer factor de la derecha se le llama *Razón de Verosimilitud Positiva* ($LR+$), del inglés *Likelihood Ratio* (LR), que definiremos como:

$$LR+ = \frac{\text{sensibilidad}}{(1 - \text{especificidad})} = \frac{\frac{TP}{TP + FN}}{\frac{FP}{FP + TN}}$$

El $LR+$ indica cuánto crece la probabilidad pretest si el test es positivo, es decir, cuánto aumentan las probabilidades de la enfermedad cuando un resultado es positivo. Cuánto más grande es el $LR+$ mayor nivel de discriminación tiene. Por ejemplo, un TAC de abdomen para el diagnóstico de apendicitis suele tener un $LR+$ de 11 y una ecografía para un apendicitis de 23.

Así, el teorema de Bayes se puede enunciar como que las *odds* a posteriori son iguales a las *odds* a priori veces la razón de verosimilitud.

$$\text{Odds post-test} = (LR+) \cdot \text{Odds pretest}$$

Si el resultado del test es negativo, se define la razón de verosimilitud negativa ($LR-$) como:

$$LR- = \frac{(1 - \text{sensibilidad})}{\text{especificidad}} = \frac{\frac{FN}{TP + FN}}{\frac{TN}{FP + TN}}$$

El $LR-$ indica cuánto disminuyen las probabilidades de la enfermedad cuando un resultado es negativo.

Ejemplo 7

Con los datos del ejemplo 2, calcula la razón de verosimilitud positiva y negativa $LR+$ y $LR-$.

La razón de verosimilitud positiva $LR+$ vale:

$$LR+ = \frac{\text{sensibilidad}}{(1 - \text{especificidad})} = \frac{\frac{TP}{TP + FN}}{\frac{FP}{FP + TN}} = \frac{\frac{70}{100}}{\frac{76}{4900}} = \frac{1715}{38} = 45,13$$

Esto indica, que un test positivo es 45 veces más probable en un paciente enfermo que en otro que no lo es.

La razón de verosimilitud negativa $LR-$ es:

$$LR- = \frac{(1 - \text{sensibilidad})}{\text{especificidad}} = \frac{\frac{FN}{TP + FN}}{\frac{TN}{FP + TN}} = \frac{\frac{30}{100}}{\frac{4824}{4900}} = \frac{245}{804} = 0,3047$$

La interpretación del cociente de verosimilitud es que el resultado negativo de la prueba es 0,3 veces más frecuente en paciente con enfermedad con sospecha de padecerla. Es mejor expresar en forma inversa $1:0,3 = 3,3$, y concluir que el test negativo es 3,3 veces más frecuente en los que no tienen la enfermedad que en los que la tienen.

La tabla 6 refleja el grado de utilidad de la prueba diagnóstica en función de los valores de la razón de verosimilitud. Los mejores tests para confirmar el tener una enfermedad serán aquellos que tenga el $LR+$ más alto y para descartar la enfermedad los del $LR-$ más bajos.

Utilidad de la prueba	$LR-$	$LR+$
No es útil	1	1
Poco útil	0,33-1	1-3
Utilidad regular	0,16-0,33	3-6
Prueba útil	0,10-0,16	6-10
Muy útil	<0,1	>10

Tabla 6

Ejemplo 8

Con los datos del ejemplo 2 calcular el VPP en forma de *odds*.

Con los datos ya calculados de $LR+ = 1715/38 = 45,13$, y conocida que la prevalencia es $p(E) = 0,02$, y por tanto los $Odds$ pretest $= 1/49$, luego:

$$Odds \text{ posttest} = (LR+) \cdot Odds \text{ pretest} = \left(\frac{1715}{38}\right) \cdot \frac{1}{49} = \frac{1715}{1862} = 0,92$$

Con lo que:

$$VPP = p(E / +) = \frac{0,92}{1 + 0,92} = 0,47$$

valor que coincide con el previamente hallado en el ejemplo 5.

El teorema de Bayes a través de nomogramas

Gracias a la versión en *odds* del teorema de Bayes como producto:

$$Odds \text{ posttest} = (LR+) \cdot Odds \text{ pretest}$$

lo podemos transformar en una suma tomando logaritmos:

$$\log(Odds \text{ posttest}) = \log(LR+) + \log(Odds \text{ pretest})$$

lo que nos permite construir un nomograma y resolver el ejemplo 2 de aplicación del teorema de Bayes mediante nomogramas.

Ejemplo 9

Con los datos del ejemplo 2 calcular el VPP a través de un nomograma.

Construidas las tres escalas del nomograma, unimos el punto de la primera escala del 2% de la prevalencia con el 45 de la $LR+$ de la segunda, dándonos en la tercera escala el VPP , que en este caso es de aproximadamente el 47% (figura 8).

Volviendo a la resolución del ejemplo 1, Alcaide nos ofrecía un nomograma que permite la resolución gráfica de problemas, sin hacer cálculos numéricos. Así en la escala señalada con la letra A están contenidos en la graduación iz-

quiera los valores, en tanto por ciento, de la función:

$$\theta(t) = \frac{2}{\sqrt{\pi}} \int_0^t e^{-t^2} dt$$

y en la graduación derecha los de la variable que él llama f (*factor de probabilidad*), sustituyendo esta escala, por tanto, a la tabla de la distribución normal. La escala nombrada con la letra B es una escala logarítmica que contiene los números entre 1 y 10 000 y representa las anchuras, en metros, de las zonas del 50% de los impactos, correspondientes a un arma a una distancia determinada. La escala C es otra escala logarítmica de módulo doble que la anterior y su graduación expresará, en metros, las anchuras z' de zonas batidas en tantos por ciento de impactos determinados.

Así, si se quiere determinar el tanto por ciento de impactos que corresponden a una zona centrada de anchura $z' = 16$ m, siendo $z = 80$ m la zona del 50% que a la distancia supuesta corresponde el arma empleada, se coloca una regla sobre las divisiones $z' = 16$ y $z = 80$ de C y B y el trazo en $\theta(t)$ de A nos da la solución de $p = 11\%$. Y al revés, podemos calcular z' en función de z y $\theta(t)$.

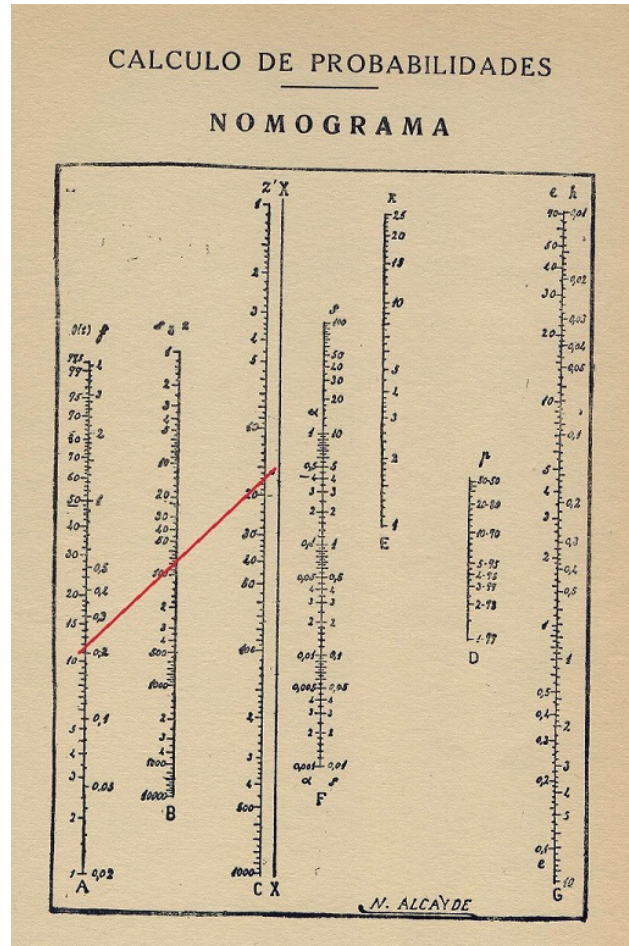


Figura 9. Resolución del ejemplo 1

Epílogo

En resumen, en este trabajo nos hemos acercado al teorema de Bayes de una forma diferente, a través de *odds* y hemos resuelto gráficamente un ejemplo de aplicación del mismo utilizando la técnica de la nomografía, de manera similar a como a comienzos del siglo XX calculaban los artilleros probabilidades. En un reciente artículo (Marasco y otros, 2011), los autores aportan una serie de razones para seguir utilizando nomogramas en lugar de algoritmos implementados en el ordenador. En la tabla 7 hemos seleccionado algunas de ellas.

La Nomografía, una ciencia olvidada que se resiste pasado un siglo a morir y puede sobrevivir como recurso didáctico, depende de nosotros.

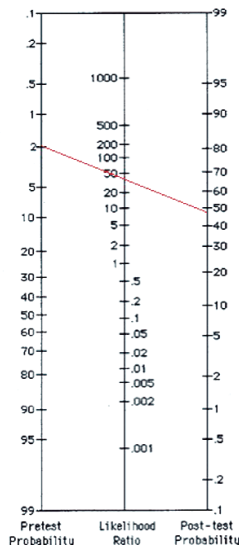


Figura 8. Nomograma para el teorema de Bayes

Requerimientos	Ordenador	Nomograma
De hardware	Ordenador o Tablet o impresora o Smartphone...	Solo lápiz y regla
De software	Windows u otro licenciado	Solo el que llevas en tu cerebro
De infraestructura	Acceso a internet, app...	Solo luz ambiente
De energía	Fuente alimentación o baterías	Solo luz ambiente
Rapidez del resultado	La del procesador Intel de tu ordenador	La de tu destreza dibujando líneas rectas
Fallo más común	Teclear mal los números	¡No encuentro mis gafas!
Uso en el Tercer Mundo	Limitado por el acceso a los ordenadores	El papel y el lápiz es fácil de conseguir

Tabla 7

Referencias bibliográficas

- ALCAYDE, N. (1915), *Elementos de Nomografía*, Dossat Editor, Madrid.
- (1942), *Cálculo de probabilidades*, Dossat Editor, Madrid.
- ARENZANA, V., P. BUERA y L. RODRÍGUEZ (1990), «Los cambios de escala y el cálculo gráfico», *Suma*, n.º 5, 59-64.
- DIDION, I. (1858), *Calcul des probabilités appliqué au tir des projectiles*, Mallet-Bachelier. París.

- FAGAN, T. J. (1975), «Letter: Nomogram for Bayes's Theorem», *New England Journal of Medicine*, vol. 293, n.º 5, 257.
- MARASCO, J., R. DOERFLER y L. ROSCHIER (2011), «Doc, what are my chances?», *The UMAP Journal*, vol. 32, n.º 4, 279-298.
- PARELLADA, A. (1942), *Gráficas y nomogramas*, Dossat Editor, Madrid.
- RUIZ-GARZÓN, G. (2014), *Condenado por la Estadística (Casos reales de aplicación de la Estadística al Derecho)*, Servicio de Publicaciones de la Universidad de Cádiz, Cádiz.

GABRIEL RUIZ-GARZÓN
Universidad de Cádiz (UCA)
<gabriel.ruiz@uca.es>

JAIME RUIZ-ZAPATERO
University College London (UCL)
<jaimeruizzapatero@gmail.com>