Bayes

Recalculando ….

Este teorema es válido en todas las aplicaciones de la teoría de la probabilidad aunque sin embargo, ha existido mucha controversia sobre el tipo de probabilidades que emplea. En esencia, los seguidores de la estadística tradicional también denominada objetivista o frecuencialista (6) sólo admiten probabilidades basadas en experimentos repetibles y que tengan una confirmación empírica mientras que los llamados estadísticos bayesianos permiten y defienden la la utilidad de las probabilidades subjetivas. El teorema, que ha resurgido con gran popularidad desde hace ya algunos años, puede servir entonces para indicar cómo debemos modificar nuestras probabilidades subjetivas cuando recibimos información adicional de un experimento. Este enfoque que propugna la estadística bayesiana está demostrando su utilidad en ciertas estimaciones basadas en el conocimiento subjetivo a priori y permitir revisar esas estimaciones en función de la evidencia , lo que está abriendo nuevas formas de hacer conocimiento.

El problema en medicina

Supóngase una prueba diagnóstica, por ejemplo nivel de glucosa en sangre, en ayunas, para diagnosticar la diabetes. Se considera que la prueba es positiva si se encuentra un nivel por encima de un cierto valor, digamos 120 mg/l.
Para evaluar la prueba, para distintos puntos de corte, se somete a la misma a una serie de individuos diabéticos diagnosticados por otro procedimiento (el patrón de oro o "gold standar") y a una serie de individuos no diabéticos. Los resultados se pueden representar en una tabla de doble entrada.

		Patrón de oro		
		NE	E	
Prueba	-	a	b	r
	+	c	d	s
		t	u	

Si la prueba fuera perfecta $b=c=0$, desgraciadamente nunca ocurre. Se denomina falso-positivo (FP) al cociente c/t, y es una estimación de la probabilidad condicionada $p(+|NE)$; se denomina falso-negativo (FN) al cociente b/u, y es una estimación de la probabilidad condicionada $p(-|E)$. Estos dos valores cuantifican los dos errores que la prueba puede cometer y caracterizan a la misma. Simétricamente, los coeficientes que cuantifican los aciertos son la sensibilidad, $p(+|E)$, y la especificidad $p(-|NE)$.

Cuando la prueba se usa con fines diagnósticos (o de "screening") interesa calcular p(E|+) (o valor Predicitivo Positivo VPP) y/o p(NE|-) (o Valor Predicitivo Negativo VPN), que aplicando el teorema de Bayes y considerando los suceso Enfermedad (E) y No Enfermedad (NE) como una partición de A, tendremos la siguiente expresión:

Como E y NE son una partición, es decir sucesos incompatibles, usando el Teorema de Bayes

$$p(E|+) = \frac{p(+|E)p(E)}{p(+|E)p(E) + p(+|NE)p(NE)}$$

Pero de acuerdo a las definiciones hechas anteriormente es posible entonces expresar que:

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

Se trata de una prueba de laboratorio que tiene un 95 % de sensibilidad y 95 % de especificidad. Pero conocemos a la vez que la prevalencia de la enfermedad en cuestión es de 5 por cada 1000 individuos en la población adulta. Haciendo uso de la fórmula (1) para calcular el Valor Predictivo Positivo (VPP):

$$VPP = \frac{\text{Sensibilidad} \times \text{Prevalencia}}{\text{Sensibilidad} \times \text{Prevalencia} + (1-\text{Especificidad}) \times (1-\text{Prevalencia})} \quad (2)$$

Sensibilidad p(+|E) =95%
Prevalencia p(E)=5/1000
Especificidad p(-|NE)=95%

VPP=
0.95 X 0.005 / 0.95 X 0.005 + 0.05 X 0.995= 0.00475/0.0545= 0.087

Dieron positivos al test pero no estan enfermos

No estan enfermos

Aproximadamente 9 casos de los mil darán positivos estando realmente enfermos

Por ejemplo calculemos el VPP si se tratara de un paciente de 25 años, con dolor precordial, fumador y con una elevación en suero de una enzima útil en el diagnóstico de infarto del miocardio, suponiendo un 1 % de prevalencia para esta edad , y con un 90 % de sensibilidad y especificidad de la prueba un resultado positivo de la prueba se corresponde con un VPP de 8,3 %.

$$VPP = \frac{.90 \times 0.01}{.90 \times 0.01 + (1 - 0.90) \times (1 - 0.01)} = 0.083 \times 100 = 8.3$$

Sin embargo una prueba positiva en un paciente adulto de 68 años con dolor anginoso, fumador, dolor precordial y estimando un 90 % de prevalencia del infarto para este grupo poblacional , entonces el VPP será de 99 % (con el mismo cálculo que en el paciente de 25). Es decir el paciente de 68 años y el joven de 25 tienen ambos positiva la misma prueba diagnóstica, sin embargo el de 68 años tiene una probabilidad mas de 10 veces mayor (99 vs. 8 %) de tener la enfermedad