Tema 1.6 Regla de Bayes.

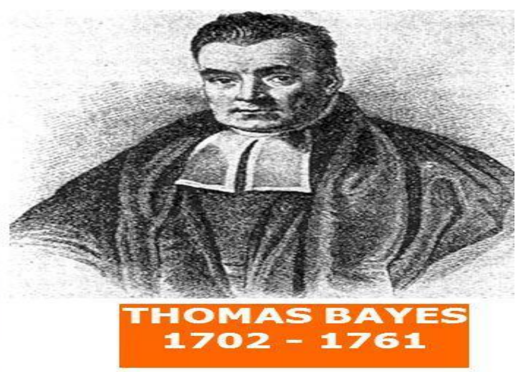
Motivación del tema. Thomas Bayes fue el primero en utilizar la probabilidad inductivamente y establecer una base matemática para la inferencia probabilística (la manera de calcular, a partir de la frecuencia con la que un acontecimiento ocurrió, la probabilidad de que ocurrirá en el futuro). El reverendo Bayes abordó también el problema de las causas a través de los efectos observados. En nuestro contexto $A_1, \cdots, A_n$ son las causas y E es el efecto observado. Queremos saber la probabilidad con que las causas contribuyen a producir el efecto. La genialidad de Bayes es que nos ofreció un método para cuantificar esto:

$$P(E) = \frac{P(A_k)P(E|A_k)}{P(A_1)P(A_1) + P(A_2)P(A_2) + \ldots + P(A_n)P(A_n)}$$

Las probabilidades $P(A_1), \cdots, P(A_n)$ son estimaciones previas o probabilidades a priori, mientras que las probabilidades $P(A_1|E), \cdots, P(A_n|E)$ son correcciones de las anteriores obtenidas en base a la información que nos da el hecho de que ya ocurrió el evento E. También las probabilidades $P(A_1|E), \cdots, P(A_n|E)$ son conocidas como probabilidades inversas, mientras que las probabilidades $P(E|A_1), \cdots, P(E|A_n)$ son directas, en el sentido de que serían las probabilidades de observar el efecto E bajo las hipótesis $A_1, \cdots, A_n$.

$$f(\mu / \bar{x}) = \frac{f(\bar{x} / \mu).f(\mu)}{\sum f(\bar{x})}$$

LEY DE PROBABILIDAD DE LAS CAUSAS
INVERSION DE LA PROBABILIDAD
PRINCIPIO DE LA RAZON INSUFICIENTE
(MODO DE SUBSANAR EL ESTADO DE
IGNORANCIA PREVIA)



Teorema 4.4 Ley de Probabilidad Total. Sea E un evento en un espacio muestral S y sean $A_1, A_2, \ldots, A_n$ eventos ajenos cuya unión es S. Entonces

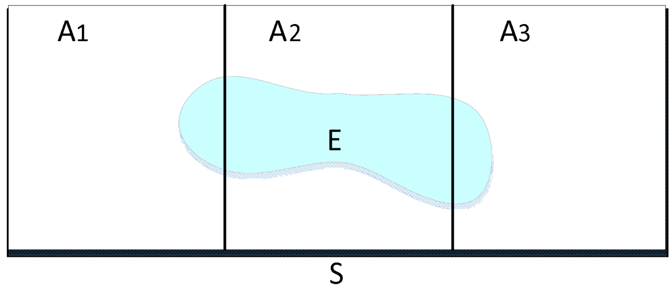
$$P(E) = P(A_1)P(A_1) + P(A_2)P(A_2) + \ldots + P(A_n)P(A_n)$$

Esta fórmula se puede obtener trazando el siguiente diagrama de árbol

$P(A_1) \rightarrow A_1$ $\quad P(E|A_1) \rightarrow E$

$P(A_2)$ A_2 $P(E|A_2)$ E

$P(A_n)$ A_n $P(E|A_n)$ E



Demostración. Como

$$E = (E \cap A_1) \cup \cdots \cup (E \cap A_n)$$

y $E \cap A_1, \cdots, E \cap A_n$ son ajenos entonces

$$P(E) = P(E \cap A_1) + \cdots + P(E \cap A_n)$$

Por la fórmula de probabilidad condicional se obtiene también las siguientes fórmulas para cada sumando

$$P(E \cap A_1) = P(A_1)P(A_1), \cdots, P(E \cap A_n) = P(A_n)P(A_n)$$

Sustituyendo estas fórmulas se obtiene el resultado

Ejemplo 1 Una fábrica utiliza tres máquinas X, Y, Z para producir ciertos artículos. Supongamos que:

1. La máquina X produce el 50% de todos los artículos, de los cuales el 3% son defectuosos.
2. La máquina Y produce el 30% de todos los artículos, de los cuales el 4% son defectuosos.
3. La máquina Z produce el 20% de todos los artículos, de los cuales el 5% son defectuosos.

0.5 X 0.03 D

0.3 Y 0.04 D

0.2 Z 0.05 D

Encuentre la probabilidad P de que el artículo seleccionado aleatoriamente sea defectuoso.

Solución. Sea D el evento en que un artículo está defectuoso. Entonces por la ley de la probabilidad total o sumando los 3 caminos del diagrama de árbol obtenemos

$$P(D) = P(X)P(X) + P(Y)P(Y) + P(Z)P(Z)$$

$$= (0.50)(0.03) + (0.30)(0.04) + (0.20)(0.05) = 0.037 = 3.7\%$$

Teorema 4.5 Fórmula de Bayes. Sea E un evento en un espacio muestral S y sean $A_1, A_2, \ldots, A_n$, eventos ajenos cuya unión es S. Entonces, para k=1,2, … , n

$$P(E) = \frac{P(A_k)P(E|A_k)}{P(A_1)P(A_1)+P(A_2)P(A_2)+...+P(A_n)P(A_n)}$$

Demostración. Basta sustituir la fórmula de probabilidad total en el denominador

$$P(E) = \frac{P(A_k \cap E)}{P(E)} = \frac{P(A_k)P(E|A_k)}{P(E)} = \frac{P(A_k)P(E|A_k)}{P(A_1)P(A_1)+P(A_2)P(A_2)+...+P(A_n)P(A_n)}$$

Esta fórmula se puede recordar dividiendo el camino con A_k entre la suma de todos los caminos que aparecen en el siguiente diagrama de árbol

$$P(A_1) \longrightarrow A_1 \quad P(E|A_1) \longrightarrow E$$

$$P(A_k) \longrightarrow A_k \quad P(E|A_k) \longrightarrow E$$

$$P(A_n) \longrightarrow A_n \quad P(E|A_n) \longrightarrow E$$

Ejemplo 2. Considere la fábrica en el ejemplo 1, suponga que se ha encontrado un artículo defectuoso entre la producción. Encuentre la probabilidad de que este provenga de cada una de las maquinas. Es decir, encuentre $P(D)$, $P(D)$, $P(Z|D)$.

Solución. Recuerde que $P(D) = P(X)P(X) + P(Y)P(Y) + P(Z)P(Z) = 0.037$, por consiguiente, por la fórmula de Bayes.

$$P(D) = \frac{P(X)P(D|X)}{P(D)} = \frac{(0.50)(0.03)}{0.037} = \frac{15}{37} = 40.5\%$$

$$P(D) = \frac{P(Y)P(D|Y)}{P(D)} = \frac{(0.30)(0.04)}{0.037} = \frac{12}{37} = 32.5\%$$

$$P(D) = \frac{P(Z)P(D|Z)}{P(D)} = \frac{(0.20)(0.05)}{0.037} = \frac{10}{37} = 27.0\%$$

P(X)=0.5 → X $P(D|X) = 0.03$ → D

P(Y)=0.3 → Y $P(D|Y) = 0.04$ → D

P(Z)=0.2 → Z $P(D|Z) = 0.05$ → D

Interpretación Estocástica de la Fórmula de Probabilidad Total y de la Fórmula de Bayes. Los problemas que involucran la ley de la probabilidad total y la fórmula de Bayes pueden ser interpretados como un proceso estocástico de dos pasos. La figura 2 da el árbol estocástico, donde el primer paso en el árbol comprende los eventos $A_1, A_2, \ldots, A_n$ que son una particiones de S y el segundo paso comprende el evento E.

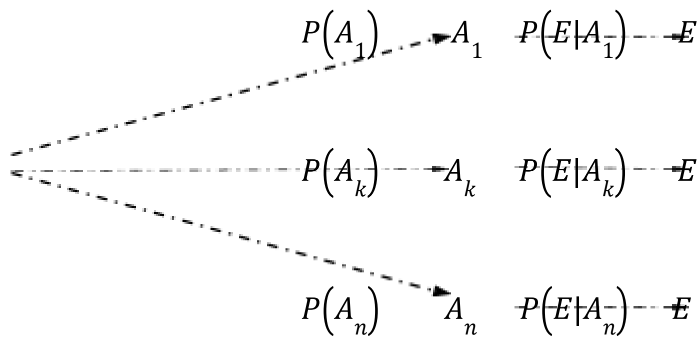


Figura 2

Para obtener la fórmula para $P(E)$, utilizando el diagrama de árbol, seguimos todos los caminos que conducen a E

$$P(E) = P(A_1)P(A_1) + P(A_2)P(A_2) + P(A_3)P(A_3)$$

Para obtener la fórmula de Bayes dividimos el camino que contiene a A_k entre la suma de todos los caminos

$$P(E) = \frac{P(A_k)P(E|A_k)}{P(A_1)P(A_1) + P(A_2)P(A_2) + P(A_3)P(A_3)}$$

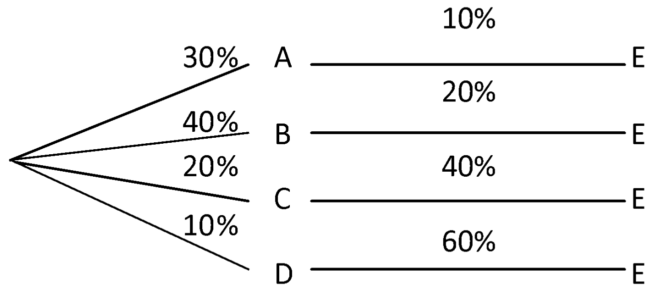
Ejemplo 3. Supongamos que un dormitorio estudiantil en una universidad está formado por:

1. 30% de estudiantes de primer año, de los cuales el 10% poseen un auto
2. 40% de estudiantes de segundo año, de los cuales el 20% poseen un auto
3. 20% de estudiantes de tercer año, de los cuales el 40% poseen un auto
4. 10% de estudiantes de cuarto año, de los cuales el 60% poseen un auto

a) Encuentre la probabilidad de que un estudiante en el dormitorio posea un auto.
b) Si un estudiante posee un auto, encuentre la probabilidad de que el estudiante sea de tercer año.

Solución. Sea A, B, C, D, el conjunto de estudiantes de primero, segundo, tercero y cuarto año y sea E el conjunto de estudiantes que poseen un auto. La figura 3 es un árbol estocástico que describe la información dada

Figura 3



a) Se busca $P(E)$ Por la ley de probabilidad total o al utilizar la figura 3, se obtiene

$$P(E) = (0.30)(0.10) + (0.40)(0.20) + (0.10)(0.60)$$

$$= 0.03 + 0.08 + 0.08 + 0.06 = 0.25 = 25\%$$

b) Se busca $P(C|E)$, Por la fórmula de Bayes

$$P(E) = \frac{P(C)P(E|C)}{P(E)} = \frac{(0.20)(0.40)}{0.25} = \frac{8}{25} = 32\%$$

Ejercicios.

1. En cierta ciudad, el 40% de las personas se consideran conservadoras©, el 35% se consideran liberales (L) y el 25% se consideran independientes (I). Durante una elección, el 45% de los conservadores votaron, el 40% de los liberales votaron y el 60% de los independientes votaron. Suponga que una persona se selecciona aleatoriamente, (a) encuentre la probabilidad de que la persona vote, (b) si la persona votó, encuentre la probabilidad de que la persona sea (i) conservador, (ii) liberal, (iii) independiente.
   Respuesta: (a) 0.47, (b) (i)0.38, (ii) 0.29, (iii) 0.31.
2. En una universidad, la altura del 4% de los hombres y el 1% de las mujeres es mayor de 6 pies. Además, el 60% de los estudiantes son mujeres. Supongamos que la altura de un estudiante seleccionado aleatoriamente es superior a los 6 pies. Encuentre la probabilidad de el estudiante sea mujer.
   Respuesta: 3/11
3. Tres máquinas A,B y C producen el 40%, el 10% y el 50% de los artículos en una fábrica respectivamente. El porcentaje de artículos defectuosos producidos por las máquinas es 2%, 3% y 4% respectivamente. Se selecciona un artículo de la fábrica aleatoriamente (a) encuentre la probabilidad de que el artículo sea defectuoso, (b) si el artículo es defectuoso, encuentre la probabilidad de que el artículo haya sido producido por: (i) la máquina A, (ii) la máquina B, (iii) la máquina C.
   Respuesta: (a) 0.031, (b) (i) 8/31, (ii) 3/31, (iii) 20/31.