Topic: Measuring Uncertainty with Probability

Bayes' Theorem

School of Mathematics and Applied Statistics

UNIVERSITY
OF WOLLONGONG
AUSTRALIA

Recall: Conditional probability

We know

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

and similarly

$$P(B|A) = \frac{\boxed{P(B \cap A)}}{P(A)} \to = \frac{\boxed{P(A \cap B)}}{P(A)}$$

Rearranging

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

and $\times P(A)$

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

So equating these two expressions for $P(A \cap B)$ we can see

$$P(A|B)P(B) = \boxed{P(B|A)}P(A)$$

Then dividing both sides by $P(A)$ we get

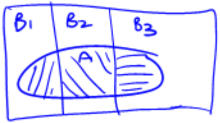
$$P(B|A) = \frac{P(A|B)P(B)}{P(A)} \qquad \text{Eqn. (1)}$$

Recall: Law of Total Probability

Recall: that the collection of sets $\{B_1, \ldots, B_k\}$ is said to be **a partition of S** when

- for some positive integer k,

$$B_1 \cup B_2 \cup \cdots \cup B_k = \bigcup_{i=1}^{k} B_i = S, \qquad i = 1, \ldots k$$



- where $\{B_1, \ldots, B_k\}$ are non-overlapping such that $B_i \cap B_j = \varnothing$ for all $i \neq j$

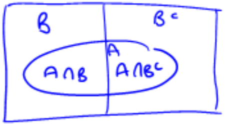
The **Law of Total Probability**:

$$P(A) = \sum_{i=1}^{k} P(A \cap B_i) = \sum_{i=1}^{k} P(A|B_i)P(B_i) \tag{2}$$

Recall: Conditional probability

So let's apply the Law of Total Probability if we have two partitions: B and B^c:

$$\overset{P(A \cap B) \quad + \quad P(A \cap B^c)}{P(A) \quad = \quad P(A|B)P(B) + P(A|B^c)P(B^c)}$$



Then substitute for $P(A)$ into Eqn(1): we obtain **Bayes' Theorem**

$$\underline{P(B|A)} = \frac{P(A|B)P(B)}{P(\underline{A|B})P(B) + P(\underline{A|B^c})P(B^c)} \leftarrow P(A)$$

Thus, if we know $P(A|B)$, we can determine the $P(B|A)$

Bayes' Theorem cont.

This result can be generalised: If the sets B_1, B_2, $\ldots B_k$ constitute a partition of *S*, then **Bayes' Theorem** may be written as

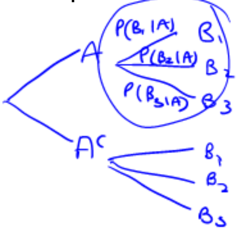
$$P(B_i|A) = \frac{P(A|B_i)P(B_i)}{P(A|B_1)P(B_1) + P(A|B_2)P(B_2) + \ldots + P(A|B_k).P(B_k)}$$

$\leftarrow P(A)$

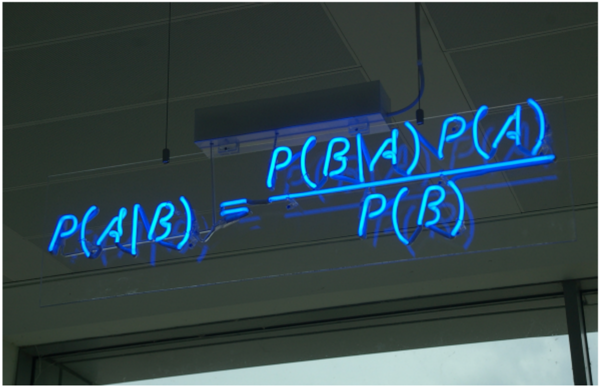
The denominator uses the Law of Total Probability from Eqn2

Notice that the sum of these conditional probabilities of the partition sets, given A, is 1.

$$\sum_{i=1}^{k} P(B_i|A) = 1$$



Bayes' Theorem in Neon Lights!



Bayes' Theorem has many applications in computer science.

Video: Prof Saharmi from Stanford University http://www.youtube.com/watch?v=MSIoBqvTKOY

Bayes' Rule

Bayes' Theorem may be conveniently presented in table

① $P(B_i)$	② $P(A\|B_i)$	③ $P(B_i)P(A\|B_i) = P(B_i \cap A)$	④ $\frac{P(B_i)P(A\|B_i)}{P(A)} = P(B_i\|A)$
x	y	$x \times y$	$\frac{x \times y}{P(A)}$
1		$\sum_i P(B_i)P(A\|B_i) = \underline{P(A)}$	1

- The first two columns usually given information
- The 3rd column is product of first two, and the sum is $P(A)$
- The 4th column is the 3rd divided by $P(A)$
  these are by Bayes' theorem $P(B_i|A)$