Total Probability Theorem and Bayes' Rule

Total Probability Theorem

A set of events $A_1, A_2, \cdots, A_n$ is said to be

- ***mutually exclusive*** (or ***disjoint***) if $A_i \cap A_j = \emptyset, \forall i \neq j$, meaning that *at most one* event can occur (if one occurs then any other cannot occur).
- (collectively) ***exhaustive*** if $A_1 \cup A_2 \cup \cdots \cup A_n = S$, meaning that *at least one* of the events will occur.
- a ***partition*** of sample space S if *one and only one* of the events will occur. Symbolically,

$$\text{partition} = \text{mutually exclusive} + \text{exhaustive}$$

Clearly, the probabilities of the member events of a partition $A_1, A_2, \cdots, A_n$ sum up to unity:

$$P\{A_1\} + \cdots + P\{A_n\} = P\{A_1 \uplus A_2 \uplus \cdots \uplus A_n\} = P\{S\} = 1$$

Consider an event B in S and a partition $A_1, A_2, \cdots, A_n$ of S. Clearly,

$$B = B \cap S = B \cap (A_1 \uplus \cdots \uplus A_n) = (B \cap A_1) \uplus \cdots \uplus (B \cap A_n)$$

Since $(B \cap A_1), \ldots, (B \cap A_n)$ are mutually exclusive, we have

$$\begin{aligned} P\{B\} &= P\{(B \cap A_1) \uplus (B \cap A_2) \uplus \cdots \uplus (B \cap A_n)\} \\ &\overset{\text{Axiom 3}}{=} P\{B \cap A_1\} + P\{B \cap A_2\} + \cdots + P\{B \cap A_n\} \end{aligned}$$

But, for $P\{A_i\} \neq 0$,

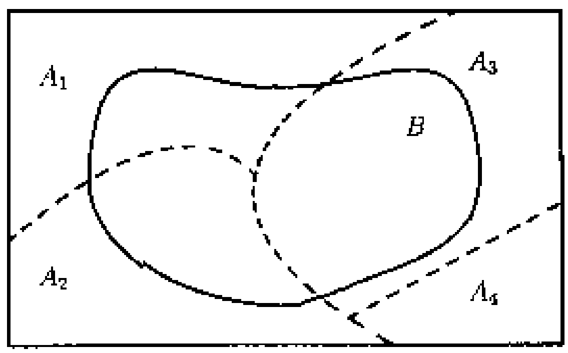
$$P\{B \cap A_i\} \overset{(2.15)}{=} P\{B|A_i\}P\{A_i\}$$

Hence, we have the following result, known as ***total probability theorem***:

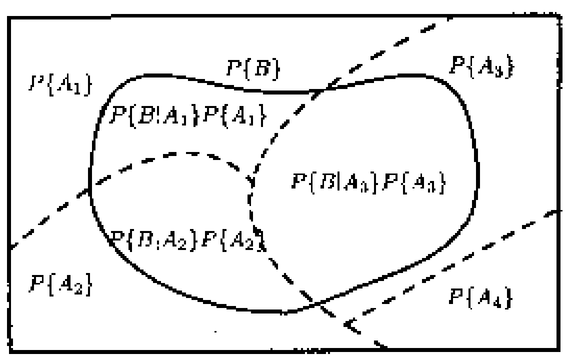
$$\boxed{P\{B\} = P\{B|A_1\}P\{A_1\} + P\{B|A_2\}P\{A_2\} + \cdots + P\{B|A_n\}P\{A_n\}}$$

This theorem is valid for any event B and any partition $A_1, A_2, \cdots, A_n$ of the sample space S. It facilitates greatly the calculation of $P\{B\}$ in many situations because both $P\{B|A_i\}$ and $P\{A_i\}$ may be much easier to calculate than a direct calculation of $P\{B\}$.

Total probability theorem is often useful for the calculation of the unconditional probability of an event $P\{B\}$ knowing various conditional probabilities of the events $P\{B|A_i\}$ and the probabilities of the conditioning events $P\{A_i\}$. Intuitively, it provides a way to find an "effect" from its "causes": It calculates the probability of an "effect" (event B) from the probabilities of all its possible "causes" (events A_i's) and the relationships between these possible "causes" and "effect" ($P\{B|A_i\}$).



(a) sample space partitioning and event B



(b) probability decomposition

Illustration of total probability theorem.