MDI Probability Theory.

Class 4: Total Probability. Bayes rule.

Total probability (Partition theorem)

Let $(\Omega, \mathcal{F}, P)$ be a probability space for a particular random experiment.
We introduce term *partition* — collection $\{B_i\}$ of pairwise disjoint events B_i, such that $\bigcup_i B_i = \Omega$.

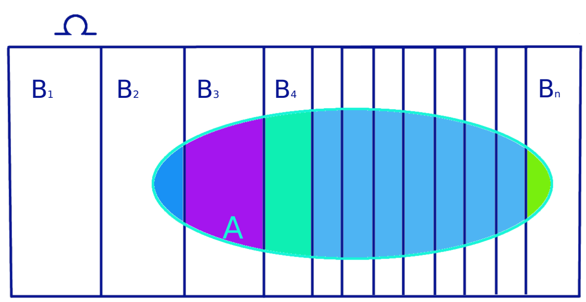
The following result, based on partition concept, is of great importance:

Theorem: If $\{B_i\}$ is a partition of Ω, with $P(B_i) > 0$, $\forall i$, then for any event A from $\mathcal{F}$:

$$\boxed{P(A) = \sum_i P(A|B_i)\,P(B_i)}$$

Indeed, let us reconstruct event A from its intersections with all events B_i:

$$A = \bigcup_i (A \cap B_i)$$



Initial events $\{B_i\}$ are pairwise disjoint, so their intersections $\{(A \cap B_i)\}$ with A are also pairwise disjoint, because each intersection is like restriction for B_i — there are no new elements. Since that, we can apply important property (additivity) of the probability function:

$$P(A) = P\left(\bigcup_i (A \cap B_i)\right) = \sum_i P(A \cap B_i)\,.$$

After that let's apply concept of conditional probability:

$$P(A \cap B_i) = P(A|B_i)P(B_i),$$

and collect everything together:

$$P(A) = \sum_i P(A|B_i)P(B_i).$$

The intuition is that event A may happen as a consequence of some others events $\{B_i\}$ which togethert completely deplete the sample space ($\bigcup_i B_i = \Omega$). In this case we can decompose probability of A to the sum of probabilities of initial B_i's multiplied by conditional probability of A to happen if B_i has happened.

One of the common situations is when partition is represented by some event B and its complement: $\{B, B^c\}$. This means that event A may happen only after either event B or B^c.

Example: Tomorrow there will be either rain or snow but not both; the probability of rain is $\frac{2}{5}$ and the probability of snow is $\frac{3}{5}$. If it rains, the probability that I will be late for my lecture is $\frac{1}{5}$, while the corresponding probability in the event of snow is $\frac{3}{5}$. What is the probability that I will be late?

Let A be the event that I am late and B be the event that it rains. The collection $\{B, B^c\}$ is a partition of the sample space. Then, by total probability principle we have:

$$P(A) = P(A|B)P(B) + P(A|B^c)P(B^c) = \frac{1}{5} \cdot \frac{2}{5} + \frac{3}{5} \cdot \frac{3}{5} = \frac{11}{25}.$$

Bayes rule

Sometimes we can treat event A as some observation, or evidence, and events B_i as some states of nature which precede A. Clearly as B_i's are disjoint, there is only one B_i happening together with A. If probabilities $P(A|B_i)$ are known, we may want to have an answer to the question "How likely that specifically this B_i preceded A?", *i.e.* to find a probability $P(B_i|A)$.

Theorem (**Bayes rule**): Let $\{B_i\}$ is a partition of Ω, with $P(B_i) > 0$, $\forall i$. Then for any event A:

$$\boxed{P(B_j|A) = \frac{P(A|B_j)P(B_j)}{\sum\limits_i P(A|B_i)P(B_i)}}$$

Let us recall definition of conditional probability:

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

We can notice that probability of intersection ($A \cap B$) may be written in two ways:

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

which gives us a formula, how two 'inverted' conditional probabilities are connected:

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

Let us change B by some element of partition B_j, and use the total probability formula for $P(A)$, and we obtain needed result.

Example (**False positives**): A rare but potentially fatal disease has an incidence of 1 in 10^5 in the population at large. There is a diagnostic test, but it is imperfect. If you have the disease, the test is positive with probability $\frac{9}{10}$; if you do not, the test is positive with probability $\frac{1}{20}$. Your test result is positive. What is the probability that you have the disease?