Lecture Notes 3: Probability and Statistics

In the previous lecture note we had discussed about the σ-field and also Borel σ-field. As I had mentioned it was mainly for completeness purposes. Please keep in mind we need these concept to define probability on certain class of subsets as it is NOT always possible to define probability on all class of subsets. Hence, from now on we do not go back to the σ-field any more. It is assumed that there exists a σ-field on which the probability has been defined. Now let us discuss some important properties and concepts of probability.

First we will define what is an event. An event is an element of a σ-field. Therefore, if A is an event, then $P(A)$ is properly defined. Similarly, if A and B are events, then $P(A \cup B)$, $P(A \cap B)$, $P(A')$ etc. are all defined. Suppose A and B are two events, and $P(B) > 0$, then the conditional probability of A given B is defines as

$$P(A|B) = \frac{P(A \cap B)}{P(B)}. \tag{1}$$

It is clear that (1) is properly defined as $P(B) > 0$. Intuitively, it means how the $P(A)$ changes, if you know the event B has taken place. Let us look at the following example:

Example: Suppose you toss a coin three times. Then the all possible outcomes are

$$\Omega = \{HHH, HHT, HTH, HTT, THH, THT, TTH, TTT\}.$$

Suppose the following probability has been defined on the elements of Ω.

$$P(HHH) = \frac{8}{27}, P(HHT) = P(HTH) = P(THH) = \frac{4}{27},$$

$$P(HTT) = P(THT) = P(TTH) = \frac{2}{27}, P(TTT) = \frac{1}{27}.$$

It is clear that once the probability has been defined on all the elements of Ω, it is uniquely defined on all events also. For example if A denotes the event that exactly at most two heads have appeared, then $A' = \{HHH\}$, and $P(A) = 1 - \frac{8}{27} = \frac{19}{27}$. Similarly, if B denotes the event that the first toss is head, then $B = \{HHH, HHT, HTH, HTT\}$ and $P(B) = \frac{18}{27} = \frac{2}{3}$. Now, suppose we want to compute the following probability: we know that the first toss is head, then what is the probability that at most two heads have appeared. Therefore, we want to compute $P(A|B)$. Observe that $A \cap B = \{HHT, HTH, HTT\}$, hence, $P(A \cap B) = \frac{10}{27}$. Therefore,

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{10}{18} = \frac{5}{9}.$$

You can see that $P(A)$ changes, if you know that the event B has taken place. Similarly, if we know that at most two heads have appeared, what is the probability that the first one is head. In this case, we want $P(B|A)$, and it becomes

$$P(B|A) = \frac{P(A \cap B)}{P(A)} = \frac{10}{19}.$$

Now we say two events A and B are independent if

$$P(A \cap B) = P(A)P(B).$$

From the definition of the conditional probability (1), it is immediate that if $P(A) > 0$ and $P(B) > 0$, then

$$P(A \cap B) = P(A)P(B) \quad \Leftrightarrow \quad P(A|B) = P(A) \quad \Leftrightarrow \quad P(B|A) = P(B).$$

Example: A news magazine publishes three columns regularly entitled: "Art" (A), "Books" (B) and "Cinema" (C). Reading habits of a randomly selected reader with respect to these columns are

Read Regularly	A	B	C	$A \cap B$	$A \cap C$	$B \cap C$	$A \cap B \cap C$
Probability	0.14	0.23	0.37	0.08	0.09	0.13	0.05

Therefore, we have

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{0.08}{0.23} = \frac{8}{23}$$

$$\begin{aligned} P(A|B \cup C) &= \frac{P(A \cap (B \cup C))}{P(B \cup C)} = \frac{P(A \cap B) + P(A \cap C) - P(A \cap B \cap C)}{P(B) + P(C) - P(B \cap C)} \\ &= \frac{0.08 + 0.09 - 0.05}{0.23 + 0.37 - 0.13} = \frac{12}{47}. \end{aligned}$$

Find $P(A|A \cup B \cup C)$.

The independence concept can be extended for more than two events also. Suppose A, B and C are three events. They are called independent events if

$$P(A \cap B) = P(A)P(B), \quad P(A \cap C) = P(A)P(C),$$

$$P(B \cap C) = P(B)P(C), \text{and} \;\; P(A \cap B \cap C) = P(A)P(B)P(C).$$

All the four conditions need to be satisfied for the three events to be independently distributed. The three events are called pairwise independent if

$$P(A \cap B) = P(A)P(B), \quad P(A \cap C) = P(A)P(C) \quad P(B \cap C) = P(B)P(C).$$

If the three events are independent then clearly they are pairwise independent, but if the three events are pairwise independent, then they may not be independent. Let us look at the following example.

Example: Let a ball be drawn from an urn containing four balls, numbered 1,2,3,4. Let $E = \{1, 2\}$, $F = \{1, 3\}$ and $G = \{1, 4\}$. If all four outcomes are assumed to be equally likely, then

$$P(E \cap F) = P(E)P(F) = \frac{1}{4}, \quad P(E \cap G) = \frac{1}{4}, \quad P(F \cap G\} = \frac{1}{4},$$

but

$$P(E \cap F \cap G) = \frac{1}{4} \neq \frac{1}{64} = P(E)P(F)P(G).$$

Hence, E, F and G are pairwise independent but they are not independent.

Now we can define the independent events for the general case. Suppose $A_1, A_2, \ldots, A_n$ are n events, they are called independent events if for every subset $\{i_1, i_2, \ldots, i_m\}$ of $\{1, 2, \ldots, n\}$ for $m \leq n$,

$$P(E_{i_1} \cap E_{i_2} \cap \ldots \cap E_{i_m}) = P(E_{i_1})P(E_{i_2}) \ldots P(E_{i_m}).$$

Now we will introduce another important concept known as Bayes theorem. It is one of the most used theorem in probability. It is simple, but it is extremely useful. Suppose $A_1, A_2, \ldots, A_n$ are n disjoint (mutually exclusive) and exhaustive, i.e. $\bigcup_{i=1}^{n} A_i = \Omega$, events. Let B be any other event. Then observe that

$$P(B) = P(B \cap \Omega) = P(B \cap \cup_{i=1}^{n} A_i) = P(\cup_{i=1}^{n} B \cap A_i) = \sum_{i=1}^{n} P(B \cap A_i).$$

Note that the last equality follows because $B \cap A_i$ for $i = 1, \ldots, n$ are mutually exclusive events. Hence,

$$P(A_i|B) = \frac{P(A_i \cap B)}{P(B)} = \frac{P(B|A_i)P(A_i)}{\sum_{k=1}^{n} P(B|A_k)P(A_k)}. \tag{2}$$

The above equality (2) is important because if we know $P(B|A_i)$ and $P(A_i)$, for $i = 1, \ldots, n$, then we can compute $P(A_i|B)$.

Problem: Consider two urns. The first contains two white and seven black balls, and the second contains five white and six black balls. We flip a fair coin and then draw a ball from the first urn or the second urn depending on whether the outcome was heads or tails. What is the conditional probability that the outcome of the toss was heads given that a white ball was selected?

Solution: Let W be the event that a white ball is drawn, and let H be the event that the coin comes up heads. The desired probability $P(H|W)$ may be calculated as follows

$$\begin{aligned} P(H|W) &= \frac{P(H \cap W)}{P(W)} = \frac{P(W|H)P(H)}{P(W|H)P(H) + P(W|T)P(T)} \\ &= \frac{\frac{2}{9} \times \frac{1}{2}}{\frac{2}{9} \times \frac{1}{2} + \frac{5}{11} \times \frac{1}{2}} = \frac{22}{67}. \end{aligned}$$

Problem: In answering a question on a multiple-choice test a student either knows the answer or guesses. Let p be the probability that she knows the answer and $1-p$ the probability that she guesses. Assume that a student who guesses at the answer will be correct with probability $1/m$, where m is the number of multiple-choice alternatives. What is the conditional probability that a student knew the answer to a question given that she answered it correctly?

Solution: Let C and K denote respectively the event that the student answers the question correctly and the event that she actually knows the answer. Now

$$\begin{aligned} P(K|C) &= \frac{P(K \cap C)}{P(C)} = \frac{P(C|K)P(K)}{P(C|K)P(K) + P(C|K')P(K')} \\ &= \frac{p}{p + (1-p)(1/m)} = \frac{mp}{1 + (m-1)p}. \end{aligned}$$