Mathematical Foundations of Computer Science
Lecture Outline
October 8, 2024

Conditional Probability

We now introduce a very important concept of conditional probability. Conditional probability allows us to calculate the probability of an event when some partial information about the result of an experiment is known. As we shall see conditional probability is often a convenient way to calculate probabilities even when no information about the result of an experiment is available.

Suppose we want to calculate the probability of event A given that event B has already occured. We denote this by $\Pr[A|B]$ (read as "the probability of A given B"). Since we know that event B has occured our sample space reduces to the outcomes in B. Is this a valid probability space? No, because the sum of probabilities of the outcomes in B is less than 1. How do we change the probabilities so that this is a valid probability distribution while making sure that the relative probabilities of outcomes in B do not change? We do this by scaling the probability of all sample points in B by $\frac{1}{\Pr[B]}$. Thus for each sample point $\omega \in B$,

$$\Pr[\omega|B] = \frac{\Pr[\omega]}{\Pr[B]}$$

To calculate $\Pr[A|B]$ we just sum up the probabilities of sample points in $A \cap B$. Thus we get

$$\Pr[A|B] = \sum_{\omega \in A\cap B} \Pr[\omega|B] = \sum_{\omega \in A\cap B} \frac{\Pr[\omega]}{\Pr[B]} = \frac{\Pr[A \cap B]}{\Pr[B]}$$

In order to avoid division by 0, we only define $\Pr[A|B]$ when $\Pr[B] > 0$. Conditional probabilities can sometimes get tricky. To avoid pitfalls, it is best to use the above mathematical definition of conditional probability. Note that the R.H.S. of the above equation are unconditional probabilities.

Example. Suppose we flip two fair coins. What is the probability that both tosses give heads given that one of the flips results in heads? What is the probability that both tosses give heads given that the first coin results in heads?

Example. Suppose we flip two fair coins. What is the probability that both tosses give heads given that one of the flips results in heads? What is the probability that both tosses give heads given that the first coin results in heads?

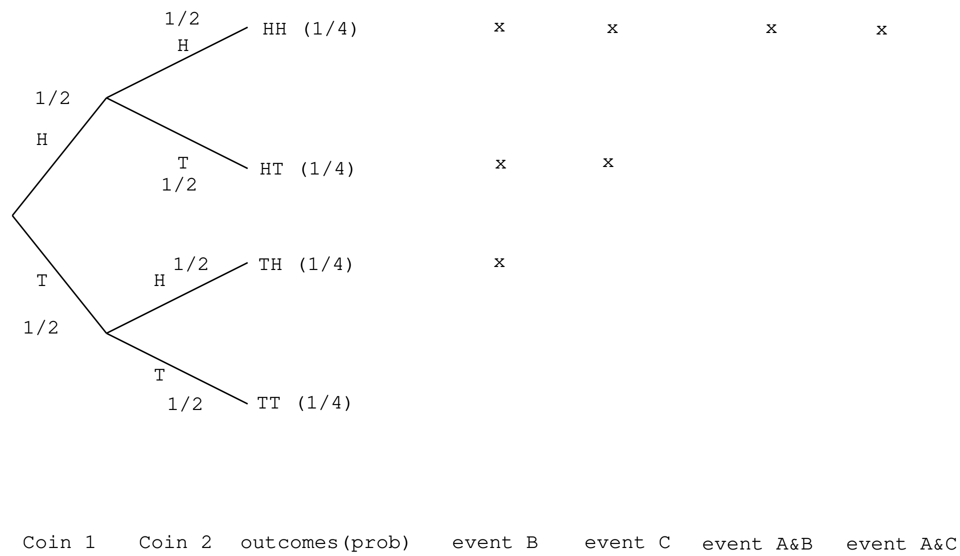

Figure 1: Tree diagram for the experiment in Example 1.

Solution. We consider the following events to answer the question.

A: event that both flips give heads.
B: event that one of the flips gives heads.
C: event that the first coin flip gives heads.

Let's first calculate $\Pr[A|B]$.

$$\Pr[A|B] = \frac{\Pr[A \cap B]}{\Pr[B]} = \frac{\Pr[A]}{\Pr[B]} = \frac{1/4}{3/4} = \frac{1}{3}.$$

Similarly we can calculate $\Pr[A|C]$ as follows.

$$\Pr[A|C] = \frac{\Pr[A \cap C]}{\Pr[C]} = \frac{\Pr[A]}{\Pr[C]} = \frac{1/4}{1/2} = \frac{1}{2}.$$

The above analysis also follows from the tree diagram in Figure 1.

The Multiplication Rule. For any two events A_1 and A_2 we have

$$\Pr[A_1 \cap A_2] = \Pr[A_1] \cdot \Pr[A_2|A_1]$$

The above formula follows from the definition of $\Pr[A_2|A_1]$. This formula can be generalized to n events. We state the generalization without proof.

$$\Pr[A_1 \cap A_2 \cap \cdots \cap A_n] = \Pr[A_1] \cdot \Pr[A_2|A_1] \cdot \Pr[A_3|A_1 \cap A_2] \cdots \Pr[A_n|A_1 \cap A_2 \cap A_3 \cap \cdots \cap A_{n-1}]$$

Example. The probability that a new car battery functions for over 10,000 miles is 0.8, the probability that it functions for over 20,000 miles is 0.4, and the probability that it functions for over 30,000 miles is 0.1. If a new car battery is still working after 10,000 miles, what is the probability that (i) its total life will exceed 20,000 miles, (ii) its additional life will exceed 20,000 miles?

Solution. We will consider the following events to answer the question.

L_{10}: event that the battery lasts for more than 10K miles.
L_{20}: event that the battery lasts for more than 20K miles.
L_{30}: event that the battery lasts for more than 30K miles.

We know that $\Pr[L_{10}] = 0.8$, $\Pr[L_{20}] = 0.4$ and $\Pr[L_{30}] = 0.1$. We are interested in calculating $\Pr[L_{20}|L_{10}]$ and $\Pr[L_{30}|L_{10}]$.

$$\begin{aligned}\Pr[L_{20}|L_{10}] &= \frac{\Pr[L_{20} \cap L_{10}]}{\Pr[L_{10}]} \\ &= \frac{\Pr[L_{20}] \cdot \Pr[L_{10}|L_{20}]}{0.8} \\ &= \frac{0.4 \times 1}{0.8} \\ &= \frac{1}{2}\end{aligned}$$

By doing similar calculations it is easy to verify that $\Pr[L_{30}|L_{10}] = \frac{1}{8}$.

Example. An urn initially contains 5 white balls and 7 black balls. Each time a ball is selected, its color is noted and it is replaced in the urn along with two other balls of the same color. Compute the probability that the first two balls selected are black and the next two white.

Solution. We will consider the following events to answer the question.

B_1: event that the first ball chosen is black.
B_2: event that the second ball chosen is black.
W_3: event that the third ball chosen is white.
W_4: event that the fourth ball chosen is white.

We are interested in calculating $\Pr[B_1 \cap B_2 \cap W_3 \cap W_4]$. Using the Multiplication rule we get,

$$\begin{aligned}\Pr[B_1 \cap B_2 \cap W_3 \cap W_4] &= \Pr[B_1] \cdot \Pr[B_2|B_1] \cdot \Pr[W_3|B_1 \cap B_2] \cdot \Pr[W_4|B_1 \cap B_2 \cap W_3] \\ &= \frac{7}{12} \times \frac{9}{14} \times \frac{5}{16} \times \frac{7}{18} \\ &= \frac{35}{768}\end{aligned}$$

The Total Probability Theorem. Consider events E and F. Consider a sample point $\omega \in E$. Observe that ω belongs to either F or $\overline{F}$. Thus, the set E is a disjoint union of two sets: $E \cap F$ and $E \cap \overline{F}$. Hence we get

$$\begin{aligned}\Pr[E] &= \Pr[E \cap F] + \Pr[E \cap \overline{F}] \\ &= \Pr[F] \times \Pr[E|F] + \Pr[\overline{F}] \times \Pr[E|\overline{F}]\end{aligned}$$

In general, if $A_1, A_2, \ldots, A_n$ form a partition of the sample space and if $\forall i, \Pr[A_i] > 0$, then for any event B in the same probability space, we have

$$\Pr[B] = \sum_{i=1}^{n} \Pr[A_i \cap B] = \sum_{i=1}^{n} \Pr[A_i] \times \Pr[B|A_i]$$

Example. A medical test for a certain condition has arrived in the market. According to the case studies, when the test is performed on an affected person, the test comes up positive 95% of the times and yields a "false negative" 5% of the times. When the test is performed on a person not suffering from the medical condition the test comes up negative in 99% of the cases and yields a "false positive" in 1% of the cases. If 0.5% of the population actually have the condition, what is the probability that the person has the condition given that the test is positive?

Solution. We will consider the following events to answer the question.

C: event that the person tested has the medical condition.
$\overline{C}$: event that the person tested does not have the condition.
P: event that the person tested positive.

We are interested in $\Pr[C|P]$. From the definition of conditional probability and the total probability theorem we get

$$\begin{aligned}
\Pr[C|P] &= \frac{\Pr[C \cap P]}{\Pr[P]} \\
&= \frac{\Pr[C]\Pr[P|C]}{\Pr[P \cap C] + \Pr[P \cap \overline{C}]} \\
&= \frac{\Pr[C]\Pr[P|C]}{\Pr[C]\Pr[P|C] + \Pr[\overline{C}]\Pr[P|\overline{C}]} \\
&= \frac{0.005 \times 0.95}{0.005 \times 0.95 + 0.995 \times 0.01} \\
&= 0.323
\end{aligned}$$

This result means that 32.3% of the people who are tested positive actually suffer from the condition!

Example. A transmitter sends binary bits, 80% 0's and 20% 1's. When a 0 is sent, the receiver will detect it correctly 80% of the time. When a 1 is sent, the receiver will detect it correctly 90% of the time.
(a) What is the probability that a 1 is sent and a 1 is received?
(b) If a 1 is received, what is the probability that a 1 was sent?

Solution. We will consider the following events.

S_0: event that the transmitter sent a 0.
S_1: event that the transmitter sent a 1.
R_1: event that 1 was received.

(a) We are interested in finding $\Pr[S_1 \cap R_1]$.

$$\begin{aligned} \Pr[S_1 \cap R_1] &= \Pr[S_1] \times \Pr[R_1 | S_1] \\ &= 0.2 \times 0.9 \\ &= 0.18 \end{aligned}$$

(b) We are interested in finding $\Pr[S_1 | R_1]$.

$$\begin{aligned} \Pr[S_1 | R_1] &= \frac{\Pr[S_1 \cap R_1]}{\Pr[R_1]} \\ &= \frac{0.18}{\Pr[R_1 \cap S_1] + \Pr[R_1 \cap S_0]} \\ &= \frac{0.18}{0.18 + \Pr[S_0] \times \Pr[R_1 | S_0]} \\ &= \frac{0.18}{0.18 + 0.8 \times 0.2} \\ &= 0.5294 \end{aligned}$$

Independent Events. Two events A and B are *independent* if and only if $\Pr[A \cap B] = \Pr[A] \times \Pr[B]$. This definition also implies that if the conditional probability $\Pr[A|B]$ exists, then A and B are independent events if and only if $\Pr[A|B] = \Pr[A]$.

Events $A_1, A_2, \ldots, A_n$ are *mutually independent* if $\forall i, 1 \leq i \leq n$ A_i does not “depend” on any combination of the other events. More formally, for every subset $I \subseteq \{1, 2, \ldots, n\}$,

$$\Pr[\cap_{i \in I} A_i] = \prod_{i \in I} \Pr[A_i]$$

In other words, to show that $A_1, A_2, \ldots, A_n$ are mutually independent we must show that all of the following hold.

$$\begin{aligned} \Pr[A_i \cap A_j] &= \Pr[A_i] \cdot \Pr[A_j] \quad \forall \text{ distinct } i, j \\ \Pr[A_i \cap A_j \cap A_k] &= \Pr[A_i] \cdot \Pr[A_j] \cdot \Pr[A_k] \quad \forall \text{ distinct } i, j, k \\ \Pr[A_i \cap A_j \cap A_k \cap A_l] &= \Pr[A_i] \cdot \Pr[A_j] \cdot \Pr[A_k] \cdot \Pr[A_l] \quad \forall \text{ distinct } i, j, k, l \\ & \ldots \\ \Pr[A_1 \cap A_2 \cap \cdots \cap A_n] &= \Pr[A_1] \Pr[A_2] \cdots \Pr[A_n] \end{aligned}$$

The above definition implies that if $A_1, A_2, \ldots, A_n$ are mutually independent events then

$$\Pr[A_1 \cap A_2 \cap \ldots \cap A_n] = \Pr[A_1] \times \Pr[A_2] \times \cdots \times \Pr[A_n]$$

However, note that $\Pr[A_1 \cap A_2 \cap \ldots \cap A_n] = \Pr[A_1] \times \Pr[A_2] \times \cdots \times \Pr[A_n]$ is not a sufficient condition for $A_1, A_2, \ldots, A_n$ to be mutually independent.

Do not confuse the concept of disjoint events and independent events. If two events A and B are disjoint and have a non-zero probability of happening then given that one event happens reduces the chances of the other event happening to zero, i.e., $\Pr[A|B] = 0 \neq \Pr[A]$. Thus by definition of independence, events A and B are not independent.

Example. Two cards are sequentially drawn (without replacement) from a well-shuffled deck of 52 cards. Let A be the event that the two cards drawn have the same value (e.g. both 4s) and let B be the event that the first card drawn is an ace. Are these events independent?

Solution. To decide whether the two events are independent we need to check whether $\Pr[A \cap B] = \Pr[A]\Pr[B]$.

$$\begin{aligned}
\Pr[A] &= \frac{3}{51} = \frac{1}{17} \\
\Pr[B] &= \frac{4}{52} = \frac{1}{13} \\
\Pr[A \cap B] &= \frac{\binom{4}{2}}{\binom{52}{2}} \\
&= \frac{1}{221} \\
&= \frac{1}{17} \times \frac{1}{13} \\
&= \Pr[A]\Pr[B]
\end{aligned}$$

Example. Suppose that a fair coin is tossed twice. Let A be the event that a head is obtained on the first toss, B be the event that a head is obtained on the second toss, and C be the event that either two heads or two tails are obtained. (a) Are events A, B, C pairwise independent? (b) Are they mutually independent?

Solution. Note that $\Omega = \{HH, HT, TH, TT\}$. $A = \{HH, HT\}$, $B = \{HH, TH\}$, $C = \{HH, TT\}$, $A \cap B = \{HH\}$, $A \cap C = \{HH\}$, $B \cap C = \{HH\}$, $A \cap B \cap C = \{HH\}$. The probabilities of the relevant events are as follows.

$$\begin{aligned}
\Pr[A] &= 1/2 \\
\Pr[B] &= 1/2 \\
\Pr[C] &= 1/2 \\
\Pr[A \cap B] &= 1/4 = \Pr[A]\Pr[B] \\
\Pr[A \cap C] &= 1/4 = \Pr[A]\Pr[C] \\
\Pr[B \cap C] &= 1/4 = \Pr[B]\Pr[C] \\
\Pr[A \cap B \cap C] &= 1/4 \neq \Pr[A]\Pr[B]\Pr[C]
\end{aligned}$$

Thus we see that A, B, C are pairwise independent but not mutually independent.