2.5 Independence of events

Unit 2: Conditioning

Adapted from Blitzstein-Hwang Chapter 2.

We have now seen several examples where conditioning on one event changes our beliefs about the probability of another event. The situation where events provide no information about each other is called *independence*.

> **Definition 2.5.1 (Independence of two events).**
>
> Events A and B are *independent* if
>
> $$P(A \cap B) = P(A)P(B).$$
>
> If $P(A) > 0$ and $P(B) > 0$, then this is equivalent to
>
> $$P(A|B) = P(A),$$
>
> and also equivalent to $P(B|A) = P(B)$.

In words, two events are independent if we can obtain the probability of their intersection by multiplying their individual probabilities. Alternatively, A and B are independent if learning that B occurred gives us no information that would change our probabilities for A occurring (and vice versa).

Note that independence is a *symmetric relation*: if A is independent of B, then B is independent of A.

> ☣ **Warning 2.5.2.**
>
> Independence is completely different from disjointness. If A and B are disjoint, then $P(A \cap B) = 0$, so disjoint events can be independent only if $P(A) = 0$ or $P(B) = 0$. Knowing that A occurs tells us that B definitely did not occur, so A clearly conveys information about B.

We also often need to talk about independence of three or more events.

Definition 2.5.3 (Independence of three events).

Events A, B, and C are said to be *independent* if all of the following equations hold:

$$\begin{aligned} P(A \cap B) &= P(A)P(B), \\ P(A \cap C) &= P(A)P(C), \\ P(B \cap C) &= P(B)P(C), \\ P(A \cap B \cap C) &= P(A)P(B)P(C). \end{aligned}$$

If the first three conditions hold, we say that A, B, and C are *pairwise independent*. Pairwise independence does *not* imply independence: it is possible that just learning about A or just learning about B is of no use in predicting whether C occurred, but learning that *both* A and B occurred could still be highly relevant for C. Here is a simple example of this distinction.

Example 2.5.4 (Pairwise independence doesn't imply independence).

Consider two fair, independent coin tosses, and let A be the event that the first is Heads, B the event that the second is Heads, and C the event that both tosses have the same result. (A coin has two sides, called Heads and Tails. A coin is called *fair* if the outcomes Heads and Tails are equally likely to occur when the coin is tossed; a coin is called *biased* if it is not fair.) Then A, B, and C are pairwise independent but not independent, since $P(A \cap B \cap C) = 1/4$ while $P(A)P(B)P(C) = 1/8$. The point is that just knowing about A or just knowing about B tells us nothing about C, but knowing what happened with *both* A and B gives us information about C (in fact, in this case it gives us perfect information about C).

We can define independence of any number of events similarly. Intuitively, the idea is that knowing what happened with any particular subset of the events gives us no information about what happened with the events not in that subset.

Conditional independence is defined analogously to independence.

Definition 2.5.5 (Conditional independence).

Events A and B are said to be *conditionally independent* given event E if $P(A \cap B|E) = P(A|E)P(B|E)$.

Example 2.5.6 (Conditional independence doesn't imply independence).

Returning once more to the scenario from Example 2.3.6, suppose we have chosen either a fair coin or a biased coin with probability 3/4 of heads, but we do not know which one we have chosen. We flip the coin a number of times. Conditional on choosing the fair coin, the coin tosses are independent, with each toss having probability 1/2 of heads. Similarly, conditional on choosing the biased coin, the tosses are independent, each with probability 3/4 of heads.

However, the coin tosses are not unconditionally independent, because if we don't know which coin we've chosen, then observing the sequence of tosses gives us information about whether we have the fair coin or the biased coin in our hand. This in turn helps us to predict the outcomes of future tosses from the same coin.

To state this formally, let F be the event that we've chosen the fair coin, and let A_1 and A_2 be the events that the first and second coin tosses land heads. Conditional on F, A_1 and A_2 are independent, but A_1 and A_2 are not unconditionally independent because A_1 provides information about A_2.

Example 2.5.7 (Independence doesn't imply conditional independence).

Suppose that my friends Alice and Bob are the only two people who ever call me. Each day, they decide independently whether to call me: letting A be the event that Alice calls and B be the event that Bob calls, A and B are unconditionally independent. But suppose that I hear the phone ringing now. Conditional on this observation, A and B are no longer independent: if the phone call isn't from Alice, it must be from Bob. In other words, letting R be the event that the phone is ringing, we have $P(B|R) < 1 = P(B|A^c, R)$, so B and A^c are not conditionally independent given R, and likewise for A and B.