2.6 Independent events

Definition 2.4. Two events $A, B \subseteq S$ are **independent** if $P(A \cap B) = P(A)P(B)$.

This just sort of comes out of nowhere, but it has a nice interpretation. Consider two independent events $A, B \subseteq S$. If you observe that B occurs, then the conditional probability of A is

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)} = \frac{P(A)P(B)}{P(B)} = P(A). \tag{2.10}$$

So your probability assessment of the event A was not changed by the news that B occurred. These events "live separate lives," and knowledge of one does not teach you anything about the likelihood of the other.

Example 2.15. Recall the example of rolling two fair, six-sided die. The outcomes are visualized in Table 2.2. Next consider the following events:

$$\begin{aligned}
A &= \text{"Die 1 is less than or equal to 3"} \\
B &= \text{"Die 2 is equal to 1"} \\
C &= \text{"The sum of the rolls is 10"} \\
D &= \text{"The sum of the rolls is 7"} \\
E &= \text{"The rolls are the same (doubles)"} \\
F &= \text{"The sum is even".}
\end{aligned}$$

Since the dice are fair, outcomes are equally likely, so by inspecting Table 2.2 and counting, we know that $P(A) = 18/36 = 1/2$, $P(B) = 6/36$, $P(C) = 3/36$, $P(D) = 6/36$, $P(E) = 6/36$, and $P(F) = 18/36 = 1/2$. Furthermore, we see that $P(A \cap B) = 3/36$, $P(A \cap C) = 0$, $P(A \cap D) = 3/36$, and $P(E \cap F) = 6/36$. As a consequence, we see that

$$\begin{aligned}
P(A \mid B) &= \frac{P(A \cap B)}{P(B)} = \frac{3/36}{6/36} = 1/2 = P(A) \\
P(A \mid C) &= \frac{P(A \cap C)}{P(C)} = \frac{0}{3/36} = 0 \neq P(A) \\
P(A \mid D) &= \frac{P(A \cap D)}{P(D)} = \frac{3/36}{6/36} = 1/2 = P(A) \\
P(E \mid F) &= \frac{P(E \cap F)}{P(F)} = \frac{6/36}{18/36} = 1/3 \neq P(E) \\
P(F \mid E) &= \frac{P(E \cap F)}{P(E)} = \frac{6/36}{6/36} = 1 \neq P(F).
\end{aligned}$$

So A and B are independent events. This makes sense because the dice are rolled separately and their outcomes should not affect one another. So knowing that Die 2 is a 1 should not teach you anything about Die 1. A and C are dependent events, because they are mutually exclusive. If you know that you rolls sum to ten, that guarantees that Die 1 is greater than a three. A and D are independent. If you roll a double, that necessarily means that your sum is also even, so in fact $E \subseteq F$, and these events are quite *dependent*.

Remark 2.11. Something that inevitably confuses folks is the difference between disjoint and independent events. The colloquial meanings of those words make it feel like the concepts would be similar, but they are actually quite different. Recall:

- **Disjoint events**: they cannot occur at the same time, so $A \cap B = \varnothing$;
- **Independent events**: knowledge of one does not teach you anything about the other, so $P(A \mid B) = P(A)$.

Right off the bat, one difference we see is that disjointedness is a property of the sets themselves. Independence is a property of the probability measure and how it factors. So disjoint events remain disjoint regardless what measure is being used, but two events that are independent with respect to one probability measure may not be independent under another.

But here is a more vivid illustration of the difference. Say you had two disjoint events A and B, and $P(A) > 0$. If you compute the conditional probability of one given the other, you get

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)} = \frac{P(\varnothing)}{P(B)} = \frac{0}{P(B)} = 0 \neq P(A).$$

So if A and B are disjoint, and I know B occurs, then this *guarantees* that A does not. Far from teaching us nothing about A, knowing B teaches us everything about it. In this way we see that disjointedness is in fact a rather severe form of *dependence*.

Remark 2.12. Can we visualize independent events with a Venn diagram? No. And the reason again is that independence is not fundamentally a property of the sets themselves, but a property of the measure P that is acting on the sets. Visualizing P is an inherently difficult task, and so independence cannot be easily drawn with our familiar blobs.

Definition 2.5. Two events $A, B \subseteq S$ are **conditionally independent** given a third event C if $P(A \cap B \mid C) = P(A \mid C)P(B \mid C)$.

Equivalently,

$$P(A \mid B \cap C) = P(A \mid C). \tag{2.11}$$

(Check this is the same as the above definition!)

Conditioning on C can *remove* dependence between events. Suppose C is the event that a patient truly has COVID. Consider two diagnostic tests administered to the same patient:

$$A = \{\text{rapid test is positive}\}, \qquad B = \{\text{PCR test is positive}\}.$$

The outcome of the tests depends on two factors: the true disease status and random error.

Clearly, A and B are dependent, since both are related to true disease status. But if we condition on the event of having COVID, then A and B become *independent*, since the only remaining randomness is coming from independent measurement errors of the two tests.

Conditioning can also *create* dependence between two unconditionally independent events A and B.

Example 2.16. Suppose we roll two dice. Let $A = \{\text{1st roll} = H\}$, $B = \{\text{2nd roll} = H\}$, and $C = \{\text{1st roll} = \text{2nd roll}\}$. Then A, B are independent. But

$$P(A \mid B \cap C) = P(\text{1st roll} = H \mid \text{2nd roll} = H, \text{1st roll} = \text{2nd roll}) = 1,$$

whereas $P(A \mid C) = \frac{P(HH)}{P(HH)+P(TT)} = 1/2 \neq 1$. Clearly, knowing the two rolls are the same links events A and B together.

Definition 2.6. Events $A_1, \dots A_k$ are **jointly independent** (or mutually independent) if for every subset $A_{i_1}, \dots, A_{i_j}$ of j of these events,

$$P(A_{i_1} \cap A_{i_2} \cap \cdots \cap A_{i_j}) = P(A_{i_1})P(A_{i_2}) \cdots P(A_{i_j}).$$

.

So for three sets A, B, C, we should have

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

If we only have the equalities in the first line, the sets are called *pairwise* independent. Three events could be pairwise but not jointly independent. This is the case for A, B, C from Example 2.16 (check!).

Example 2.17 (Tossing until a head)**.** Suppose I toss a fair coin until I observe a head.

1. Probability I need exactly three tosses: $P(HHT) = \frac{1}{2} \times \frac{1}{2} \times \frac{1}{2}$ by independence of the tosses.
2. Probability I need more than 20 tosses: same as probability of 20 tails in a row $\implies (\frac{1}{2})^{20}$.
3. Probability I never see a head?

$$1 - P(H) - P(TH) - P(TTH) - P(TTH) - .... = 1 - \sum_{k=1}^{\infty} \left(\frac{1}{2}\right)^k = 1 - 1 = 0.$$