Law of Total Probability

We ended the last lesson with a reminder that $\Pr(E \cap F)$ does not generally equal $\Pr(E) \Pr(F)$. You can only find the joint probability by multiplying marginals if the events are independent. So a rule that you might vaguely remember from basic statistics classes doesn’t generalize when events are related to each other.

But we do have a formula for the probability of an intersection (that is, the joint probability) that always works, whether the events are independent or not. Remember the definition of conditional probability

$$\Pr(E \mid F) = \frac{\Pr(E \cap F)}{\Pr(F)}.$$ 

We can just rearrange it as

$$\Pr(E \cap F) = \Pr(E \mid F) \Pr(F).$$

This rule generalizes to three or more events as well:

$$\Pr(E \cap F \cap G) = \Pr(E \mid F \cap G) \Pr(F \mid G) \Pr(G).$$

Figure 1 illustrates calculation of three-way joint probabilities. As in the previous lesson, we find the probabilities of a joint event on an end node by multiplying the edge probabilities leading to it.

In the previous lesson, we used a tree to visualize the calculation of $\Pr(P \cap L) = \Pr(P \mid L) \Pr(L)$ and $\Pr(P \cap C) = \Pr(P \mid C) \Pr(C)$, and then we added these numbers to get $\Pr(P)$. This is a very common process – we often find the probability of one event by thinking about different ways conditions under which it could occur. Because it’s so useful, we enshrine it in a named formula:

**Law of Total Probability** $\Pr(E) = \Pr(E \mid F) \Pr(F) + \Pr(E \mid F^c) (1 - \Pr(F))$
A natural way to interpret the Law of Total Probability is as a weighted average. The overall likelihood that $E$ occurs is a mixture of the likelihood that $E$ occurs in the $F$-scenario and the not-$F$ scenario. The more likely the $F$-scenario is, the more weight we give it in the mixture.

More generally,

**Law of Total Probability** (long version). If $F_1, F_2 \ldots F_n$ partition $S$,\(^1\) Then \(\Pr(E) = \Pr(E|F_1)P(F_1) + \Pr(E|F_2)P(F_2) \ldots + \Pr(E|F_n)P(F_n)\)

\(^1\) If $F_1, F_2 \ldots F_n$ partition $S$ then $E \cap F_1, E \cap F_2 \ldots E \cap F_n$ will partition $E$. 