

16 Conditional Probability

If we have partial information, this effectively shrinks the available sample space and changes the probabilities. We begin with an example.

Monty Hall show. The setting is a game show in which a prize is hidden behind one of three curtains. Call the curtains X , Y , and Z . You can win the prize by guessing the right curtain.

STEP 1. You choose a curtain.

This leaves two curtains you did not choose, and at least one of them does not hide the prize. Monty Hall opens this one curtain and this way demonstrates there is no prize hidden there. Then he asks whether you would like to reconsider. Would you?

STEP 2A. You stick to your initial choice.

STEP 2B. You change to the other available curtain.

Perhaps surprisingly, Step 2B is the better strategy. As shown in Figure 16, it doubles your chance to win the prize.

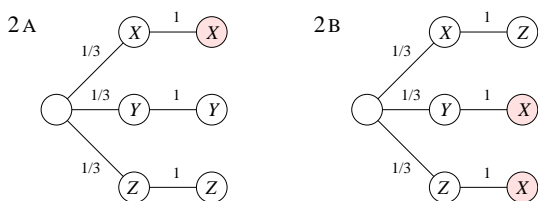


Figure 16: Suppose the prize is behind curtain X . The chance of winning improves from $\frac{1}{3}$ in 2A to $\frac{2}{3}$ in 2B.

Formalization. We are given a sample space, Ω , and consider two events, $A, B \subseteq \Omega$. The *conditional probability* of even A given event B is

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

We illustrate this definition in Figure 17. If we know that the outcome of the experiment is in B , the chance that it is also in A is the fraction of B occupied by $A \cap B$. We say A and B are *independent* if knowing B does not change the probability of A , that is,

$$P(A | B) = P(A).$$

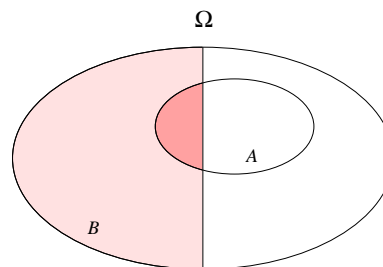


Figure 17: Assuming B , the probability of A is represented by the fraction of the shaded region, B , that is dark shaded, $A \cap B$.

Since $P(A | B) = \frac{P(A \cap B)}{P(B)} = P(A)$, we have

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

We thus see that independence is symmetric. However, it fails to be an equivalence relation because it is neither reflexive nor transitive. Combining the definition of conditional probability with the condition of independence, we get a formula for the probability of two events occurring at the same time.

PRODUCT PRINCIPLE FOR INDEPENDENT PROB. If A and B are independent then $P(A \cap B) = P(A) \cdot P(B)$.

Trial processes. In many situations, a probabilistic experiment is repeated, possibly many times. We call this a *trial process*. It is *independent* if the i -th trial is not influenced by the outcomes of the preceding $i - 1$ trials, that is,

$$P(A_i | A_1 \cap \dots \cap A_{i-1}) = P(A_i),$$

for each i .

An example is picking a coin from an urn that contains one nickel, two dimes, and two quarters. We have an independent trial process if we always return the coin before the next draw. The chance we get a quarter is therefore $\frac{2}{5}$ each time. The chance to pick the quarter three times in a row is therefore $(\frac{2}{5})^3 = \frac{8}{125} = 0.064$. More generally, we have the

INDEPENDENT TRIAL THEOREM. In an independent trial process, the probability of a sequence of outcomes, a_1, a_2, \dots, a_n , is $P(a_1) \cdot P(a_2) \cdot \dots \cdot P(a_n)$.

Trial processes that are not independent are generally more complicated and we need more elaborate tools to

compute the probabilities. A useful such tool is the tree diagram as shown in Figure 18 for the coin picking experiment in which we do not replace the picked coins.

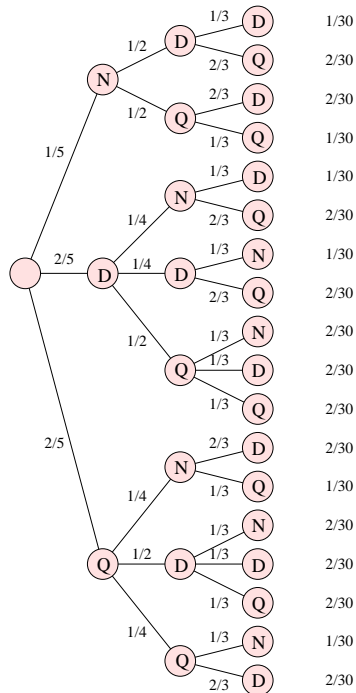


Figure 18: What is the probability of picking the nickel in three trials?

Medical test example. Probabilities can be counterintuitive, even in situations in which they are important. Consider a medical test for a disease, D . The test mostly gives the right answer, but not always. Say its false-negative rate is 1% and its false-positive rate is 2%, that is,

$$\begin{aligned} P(y | D) &= 0.99; \\ P(n | D) &= 0.01; \\ P(y | \neg D) &= 0.02; \\ P(n | \neg D) &= 0.98. \end{aligned}$$

Assume that the chance you have disease D is only one in a thousand, that is, $P(D) = 0.001$. Now you take the test and the outcome is positive. What is the chance that you have the disease? In other words, what is $P(D | y)$? As illustrated in Figure 19,

$$P(D | y) = \frac{P(D \cap y)}{P(y)} = \frac{0.00099}{0.02097} = 0.047 \dots$$

This is clearly a case in which you want to get a second opinion before starting a treatment.

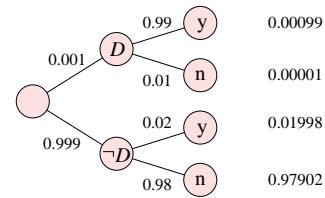


Figure 19: Tree diagram showing the conditional probabilities in the medical test question.

Summary. Today, we learned about conditional probabilities and what it means for two events to be independent. The Product Principle can be used to compute the probability of the intersection of two events, if they are independent. We also learned about trial processes and tree diagrams to analyze them.