Total Probability Theorem

If we have at our disposal a natural way to partition the sample space Ω, then we have yet another way we can use conditional probability to break apart calculations.

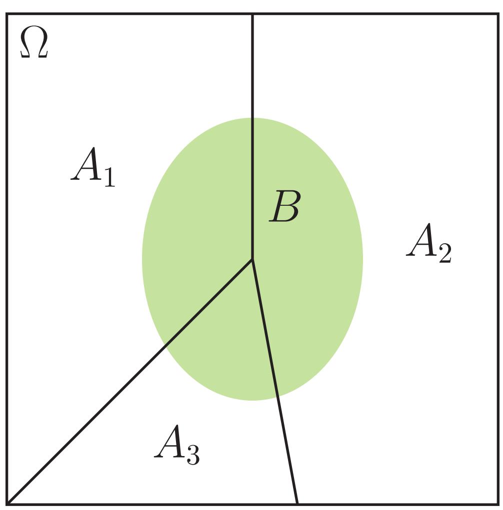
Let $A_1, A_2, \ldots, A_n$ be **disjoint events** that **partition** Ω, meaning:

$$\begin{aligned} A_i \cap A_j &= \emptyset \quad \text{for all } i \neq j \\ A_1 \cup A_2 \cup \cdots \cup A_n &= \Omega. \end{aligned}$$

That is, every outcome is in one and only one of the events A_n. We will also assume that $\mathrm{P}(A_n) > 0$ for $n = 1, \ldots, N$. Then for any event B,

$$\begin{aligned} \mathrm{P}(B) &= \mathrm{P}(A_1 \cap B) + \mathrm{P}(A_2 \cap B) + \cdots + \mathrm{P}(A_n \cap B) \\ &= \mathrm{P}(B|A_1)\,\mathrm{P}(A_1) + \mathrm{P}(B|A_2)\,\mathrm{P}(A_2) + \cdots + \mathrm{P}(B|A_n)\,\mathrm{P}(A_n) \\ &= \sum_{i=1}^{n} \mathrm{P}(B|A_i)\,\mathrm{P}(A_i). \end{aligned}$$

Here is a picture:



Example. No discussion of conditional probability would be complete without mention of the famous "Monty Hall problem". Suppose you are on a game show where there are three doors. Behind one of the doors is a prize (e.g., a car), and behind the other two doors something of little value (usually a goat). You get to pick one of the doors. Clearly, at this stage in the game, your chance of having picked the correct door is 1/3. Where it gets interesting is that next, the host of the show reveals a goat behind one of the two doors you did not pick. You now have the chance to either stick with your original guess, or switch to the other door. Which should you choose?

Let's first define A to be the event that the strategy of sticking with the original guess results in winning the prize. For convenience, let's number the doors, and we will start with the original guess as "Door 1". We can then compute

$$\begin{aligned} \mathrm{P}(A) &= \sum_{i=1}^{3} \mathrm{P}(A|\text{Car behind "Door } i\text{"})\, \mathrm{P}(\text{Car behind "Door } i\text{"}) \\ &= \end{aligned}$$

Next define B to be the event that the strategy of switching doors results in winning the prize. Then we have

$$\begin{aligned} \mathrm{P}(B) &= \sum_{i=1}^{3} \mathrm{P}(B|\text{Car behind "Door } i\text{"})\, \mathrm{P}(\text{Car behind "Door } i\text{"}) \\ &= \end{aligned}$$

Exercise:
Robert is the star quarterback for your favorite football team. His knee is bothering him, and so there is only a 40% chance he plays in the next game. If he plays, the probability that your teams wins is 0.75. If he does not, it is only 0.35. What is the probability that your team wins the game?

Exercise:
Anders and Blake both have coolers; Anders' is filled with 8 sodas and 3 beers, while Blake's is filled with 2 sodas and 11 beers. The coolers look identical, so you just choose one at random and start pulling out drinks.

(a) Suppose you pull out one drink. What is the probability that you pull out a soda?

(b) Now suppose you pull out two drinks. What is the probability you pull out two sodas?

Exercise:
Consider the following game: suppose that you have two decks of (well-shuffled) cards in front of you. You draw a card from the first deck. If the card is red, you are allowed to draw a card from the next deck. You win \$1 for each diamond, \$2 for each heart, \$3 for each club, and \$4 for each spade. If the game costs \$3 to play, what is the probability that you come out ahead (i.e., win \$4 or more)?

Bayes Rule

There is a clever trick that lets us relate $\mathrm{P}(A|B)$ to $\mathrm{P}(B|A)$. It is called *Bayes Rule*, and it is one of the most important results in all of science and engineering.

Let A, B be events with $\mathrm{P}(B) > 0$. Then

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

This relationship is easy to verify. We know that

$$\frac{\mathrm{P}(A \cap B)}{\mathrm{P}(B)} = \mathrm{P}(A|B), \quad \text{and} \quad \frac{\mathrm{P}(A \cap B)}{\mathrm{P}(A)} = \mathrm{P}(B|A),$$

and so

$$\mathrm{P}(A \cap B) = \mathrm{P}(A|B)\,\mathrm{P}(B) = \mathrm{P}(B|A)\,\mathrm{P}(A),$$

and Bayes Rule follows immediately.

One of the main applications of Bayes rule is for inference from indirect observations. The following example is a classic illustration of this.

Example. The incidence rate for a certain disease is 15/100,000. There is a test for the disease which is 95% accurate: if a person has the disease, the test comes back positive with probability 0.95; if a person does not have the disease, the test comes back negative with probability 0.95.

Jim goes to the doctor and is given the test as part of a routine checkup. It comes back positive. What is the probability that Jim has the disease?

To answer the question, define the events

$$A = \{\text{Jim has the disease}\}, \quad B = \{\text{Jim tested positive}\}.$$

We know that

$$\begin{aligned} \mathrm{P}(A) &= 0.00015 \\ \mathrm{P}(B|A) &= 0.95 \\ \mathrm{P}(B^c|A^c) &= 0.95, \quad (\text{and so } \mathrm{P}(B|A^c) = 1 - \mathrm{P}(B^c|A^c) = 0.05). \end{aligned}$$

We can calculate $\mathrm{P}(B)$ using the Total Probability Theorem:

$$\mathrm{P}(B) = \mathrm{P}(A)\,\mathrm{P}(B|A) + \mathrm{P}(A^c)\,\mathrm{P}(B|A^c),$$

where A^c is the complement of A:

$$A^c = \{\text{Jim does not have the disease}\}.$$

We have

$$\mathrm{P}(B) = (0.00015)(0.95) + (0.99985)(0.05) \approx 0.0501,$$

and so

$$\begin{aligned} \mathrm{P}(A|B) &= \frac{\mathrm{P}(B|A)\,\mathrm{P}(A)}{\mathrm{P}(B)} \\ &= \frac{(.00015)(0.95)}{(.00015)(0.95) + (0.99985)(0.05)} \approx 0.0028, \end{aligned}$$

or about 0.28%.

Exercise:
Let's go back to the problem of Robert, the injured star quarterback.

Suppose you could not bear to watch the game, but you find out later that your team won. What is the probability that Robert played?

Methods for Solving Problems

Let's take a step back for the moment and review some of the techniques we've developed for calculating probabilities.

Enumeration

The simplest way is to calculate the probability of an event is often to just enumerate (i.e., write an explicit list of) all possible outcomes of the experiment. If the outcomes are equally likely, then the probability of the event occurring is the ratio of the number of "successes" to the total number of possible outcomes. If the elementary outcomes have different probabilities, then the probability of the event occurring is the sum of the probabilities of the outcomes that correspond to a "success". This is the method we have used the most so far.

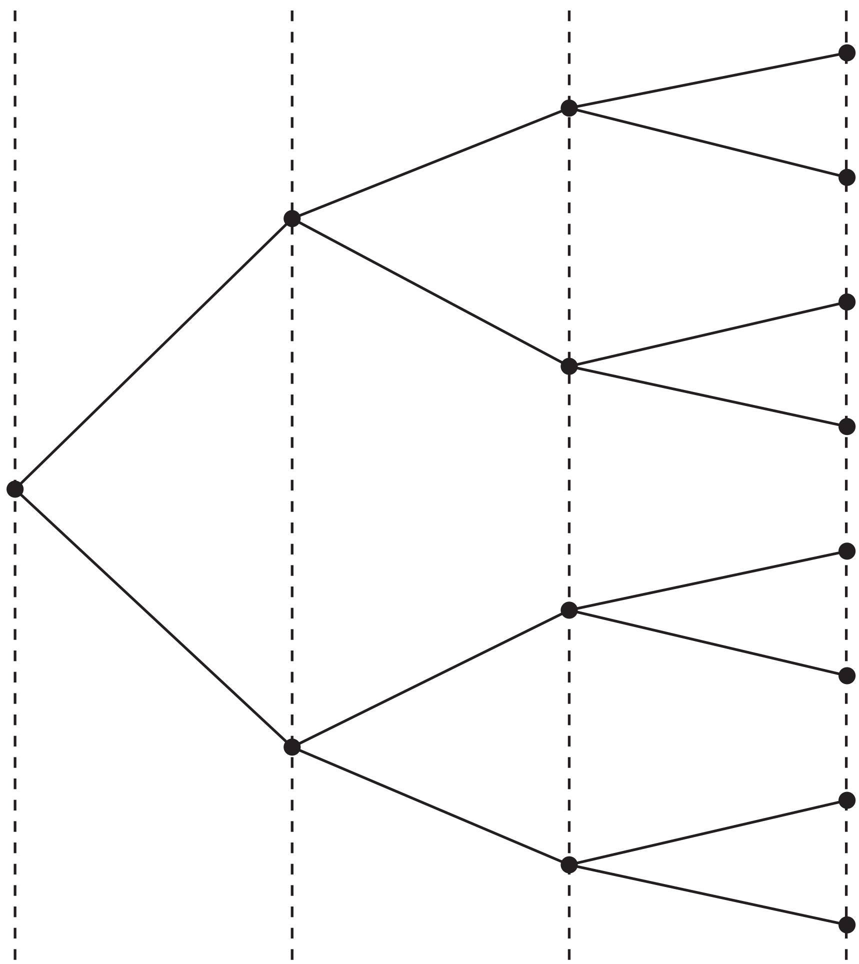
Sequential

When dealing with a sequence of outcomes, it can be useful to break up the calculation of a particular probability into more manageable pieces by following along the "history" of the sequence of outcomes. We have implicitly done this a few times already, but a couple of examples will hopefully clarify how this technique can be used to help organize and sometimes simplify our calculations.

(Recall the football game betting example on page 28 ...)

The next example shows that the sequential approach can also be very handy when the sequence of outcomes are not independent.

Exercise:
We draw 3 cards from an ordinary 52 card deck. What is the probability that we draw at least two hearts? You may find this diagram useful. (Remember the multiplication rule for conditional probability!)



Divide-and-conquer

Another approach (that we discussed at the beginning of this lecture) is the divide-and-conquer approach of splitting up the sample space into more manageable subsets and calculating the conditional probabilities for each subset, combining the results using the total probability theorem. In many ways this approach mirrors the sequential approach, but often applies in scenarios that aren't exactly sequential in nature.

Counting

You may have noticed that all the problems we've been talking about so far have involved experiments with only a small number of outcomes (e.g., experiments with only two outcomes, repeated only two or three times). Strategies where we explicitly enumerate all possible outcomes (or draw a tree showing all possible outcomes) are no longer feasible when the number of outcomes gets much larger than this. To handle cases where the number of outcomes is rather large, we typically have to simplify our calculations carefully counting the number of ways a certain event can occur *without explicitly listing them all*. We've already done this a few times (e.g., we have used the fact that there are 2^{10} possible outcomes if we toss a coin ten times without ever listing all 2^{10} possibilities) but we haven't discussed principled methods for handling more complicated scenarios yet.

There is an entire branch of mathematics — called *combinatorics* — that deals with problems of this type. We will of course only scratch the surface in this course, but we will end up acquiring some powerful tools. This will be the focus of the next lecture.