Lecture Notes

This lecture concerns the probability of one event, given that some other event happens. This is called a *conditional probability.* Trying to solve conditional probability problems by intuition can be very difficult. On the other hand, we can chew through these problems with the same four-step method that we used for the Monty Hall problem:

1. Find the sample space.
2. Define events of interest.
3. Compute outcome probabilities.
4. Compute event probabilities.

The main new feature today is a formula called the *Product Rule*, introduced below.

1 Conditional Probability

Suppose that we pick a random person in the world. Everyone has an equal chance of being picked. Let A be the event that the person is an MIT student, and let B be the event that the person lives in Cambridge. The situation is shown in Figure 1. Clearly, both events A and B have low probability. But what is the probability that a person is an MIT student, *given* that the person lives in Cambridge?

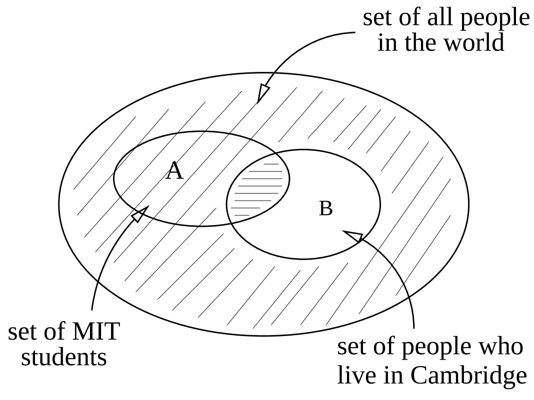

Figure 1: *What is the probability that a random person in the world is an MIT student, given that the person is a Cambridge resident?*

This is a conditional probability question. It can be concisely expressed in a special notation. In general, $\Pr(A \mid B)$ denotes the probability of event A, given event B. In this example, $\Pr(A \mid B)$ is the probability that the person is an MIT student, given that he or she is a Cambridge resident.

How do we compute $\Pr(A \mid B)$? Since we are *given* that the person lives in Cambridge, all outcomes outside of event B are irrelevant; these irrelevant outcomes are diagonally shaded in the figure. Intuitively, $\Pr(A \mid B)$ should be the fraction of Cambridge residents that are also MIT students. That is, the answer should be the probability that the person is in set $A \cap B$ (horizontally shaded) divided by the probability that the person is in set B. This gives:

$$\Pr(A \mid B) = \frac{\Pr(A \cap B)}{\Pr(B)}$$

Rearranging terms gives the following rule, which we will regard as the definition of $\Pr(A \mid B)$.

Rule 1.1 (Product Rule, special case) Let A and B be events.

$$\Pr(A \cap B) = \Pr(B) \cdot \Pr(A \mid B)$$

(We are now using the term "Product Rule" for two separate ideas. One is the definition of $\Pr(A \mid B)$ given above, and the other is the formula for the cardinality of a product of sets. In this lecture, the term always refers to the rule above.)

As an example, what is $\Pr(B \mid B)$? That is, what is the probability of event B, given that event B happens? Intuitively, this ought to be 1! The Product Rule gives exactly this result:

$$\begin{aligned}\Pr(B \mid B) &= \frac{\Pr(B \cap B)}{\Pr(B)} \\ &= \frac{\Pr(B)}{\Pr(B)} \\ &= 1\end{aligned}$$

The Product Rule can be generalized to many events.

Rule 1.2 (Product Rule, general case) Let $A_1, A_2, \ldots, A_n$ be events.

$$\Pr(A_1 \cap A_2 \cap \ldots \cap A_n) = \Pr(A_1)\Pr(A_2 \mid A_1)\Pr(A_3 \mid A_1 \cap A_2)\ldots\Pr(A_n \mid A_1 \cap \ldots \cap A_{n-1})$$

This generalization of the Product Rule can be proven from the special case by induction on n, the number of sets. We will not give a proof here, however.

2 Conditional Probability Examples

This section contains five examples of conditional probability problems. In each case, the solution requires only the Product Rule and the four-step method that we applied to the Monty Hall problem.

2.1 A Two-out-of-Three Series

The MIT computer science department's famed D-league hockey team, The Halting Problem, is playing a 2-out-of-3 series. That is, they play games until one team wins a total of two games. The probability that The Halting Problem wins the first game is $\frac{1}{2}$. For subsequent games, the probability of winning depends on the outcome of the preceding game; the team is energized by victory and demoralized by defeat. Specifically, if The Halting Problem wins a game, then they have a $\frac{2}{3}$ chance of winning the next game. On the other hand, if the team loses, then they have only a $\frac{1}{3}$ chance of winning the following game. What is the probability that The Halting Problem wins the 2-out-of-3 series, given that they win the first game?

This problem involves two types of conditioning. First, we are told that the probability of the team winning a game is $2/3$, *given* that they won the preceding game. Second, we are asked the odds of The Halting Problem winning the series, *given* that they win the first game.

Step 1: Find the Sample Space

The sample space for the hockey series is worked out with a tree diagram in Figure 2. Each internal node has two children, one corresponding to a win for The Halting Problem (labeled W) and one corresponding to a loss (labeled L). The sample space consists of six outcomes, since there are six leaves in the tree diagram.

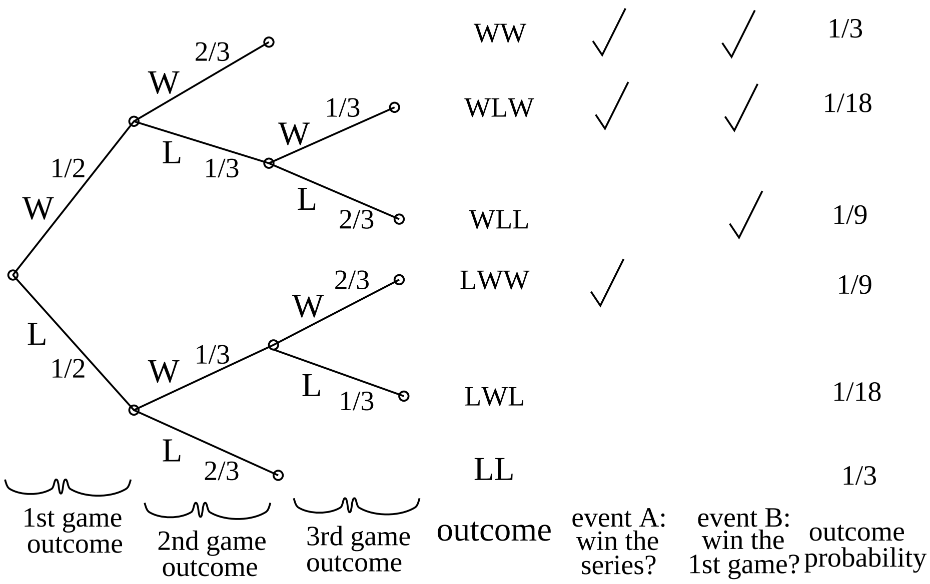

Figure 2: *What is the probability that The Halting Problem wins the 2-out-of-3 series, given that they win the first game?*

Step 2: Define Events of Interest

The goal is to find the probability that The Halting Problem wins the series given that they win the first game. This suggests that we define two events. Let A be the event that The Halting Problem wins the series, and let B be the event that they win the first game. The outcomes in each event are checked in Figure 2. Our problem is then to determine $\Pr(A \mid B)$.

Step 3: Compute Outcome Probabilities

Next, we must assign a probability to each outcome. As in the Monty Hall problem, we initially assign probabilities to edges in the tree diagram. These probabilities are given explicitly in the problem statement. Specifically, The Halting Problem has a $\frac{1}{2}$ chance of winning the first game, so the two edges leaving the root are both assigned probability $\frac{1}{2}$. Other edges are labeled $\frac{1}{3}$ or $\frac{2}{3}$ based on the outcome of the preceding game. We find the probability of an outcome by multiplying all probabilities along the corresponding root-to-leaf path. The results are shown in the figure.

This method of computing outcome probabilities by multiplying edge probabilities was introduced last lecture, but not justified. In fact, the justification is actually the Product Rule! For example, by multiplying edge weights, we conclude that the probability of outcome WW is $\frac{1}{2} \cdot \frac{2}{3} = \frac{1}{3}$. We can justify this rigorously with the Product Rule as follows.

$$
\begin{aligned}
\Pr(WW) &= \Pr(\text{win 1st game} \cap \text{win 2nd game}) \\
&= \underbrace{\Pr(\text{win 1st game}) \cdot \Pr(\text{win 2nd game} \mid \text{win 1st game})}_{\substack{\text{product of edge weights on} \\ \text{root-to-leaf path}}} \\
&= \frac{1}{2} \cdot \frac{2}{3} \\
&= \frac{1}{3}
\end{aligned}
$$

The first equation states that WW is the outcome in which we win the first game and win the second game. The second equation is an application of the Product Rule. In the third step, we substitute probabilities from the problem statement, and the fourth step is simplification. The heart of this calculation is equivalent to multiplying edge weights in the tree diagram!

Here is a second example. By multiplying edge weights in the tree diagram, we conclude that the probability of outcome WLL is $\frac{1}{2} \cdot \frac{1}{3} \cdot \frac{2}{3} = \frac{1}{9}$. We can formally justify this with the Product Rule as follows:

$$\begin{aligned}
\Pr(WLL) &= \Pr(\text{win 1st} \cap \text{lose 2nd} \cap \text{lose 3rd}) \\
&= \underbrace{\Pr(\text{win 1st}) \cdot \Pr(\text{lose 2nd} \mid \text{win 1st}) \Pr(\text{lose 3nd} \mid \text{win 1st } \cap \text{ lose 2nd})}_{\substack{\text{product of edge weights on} \\ \text{root-to-leaf path}}} \\
&= \frac{1}{2} \cdot \frac{1}{3} \cdot \frac{2}{3} \\
&= \frac{1}{9}
\end{aligned}$$

Step 4: Compute Event Probabilities

We can now compute the probability that The Halting Problem wins the tournament given that they win the first game:

$$\begin{aligned}
\Pr(A \mid B) &= \frac{\Pr(A \cap B)}{\Pr(B)} \\
&= \frac{\frac{1}{3} + \frac{1}{18}}{\frac{1}{3} + \frac{1}{18} + \frac{1}{9}} \\
&= \frac{7}{9}
\end{aligned}$$

The first line uses the Product Rule. In the second line, we use the fact that the probability of an event is the sum of the probabilities of the outcomes it contains. The last line is simplification.

The Halting Problem has a $\frac{7}{9}$ chance of winning the tournament, given that they win the first game.

2.2 An *a posteriori* Probability

In the preceding example, we wanted the probability of an event A, given an *earlier* event B. In particular, we wanted the probability that The Halting Problem won the series, given that they won the first game. Harder to think about is the probability of an event A, given a *later* event B. For example, what is the probability that The Halting Problem wins its first game, given that the team wins the series? This is called an *a posteriori* probability.

An *a posteriori* probability question can be interpreted in two ways. By one interpretation, we reason that since we are given the series outcome, the first game is already either won or lost; we do not know which. The issue of who won the first game is a question of fact, not a question of probability. Though this interpretation may have philosophical merit, we will never use it.

We will always prefer a second interpretation. Namely, we suppose that the experiment is run over and over and ask in what fraction of the experiments did event A occur when event B occurred? For example, if we run many hockey series, in what fraction of the series did the Halting Problem win the first game when they won the whole series? Under this interpretation, whether A precedes B in time is irrelevant. In fact, we will solve *a posteriori* problems exactly the same way as other conditional probability problems. The only trick is to avoid being confused by the wording of the problem!

We can now compute the probability that The Halting Problem wins its first game, given that the team wins the series. The sample space is unchanged; see Figure 2. As before, let A be the event that The Halting Problem wins the series, and let B be the event that they win the first game. We already computed the probability of each outcome; all that remains is to compute the probability of event $\Pr(B \mid A)$:

$$\begin{aligned}\Pr(B \mid A) &= \frac{\Pr(B \cap A)}{Pr(A)} \\ &= \frac{\frac{1}{3}+\frac{1}{18}}{\frac{1}{3}+\frac{1}{18}+\frac{1}{9}} \\ &= \frac{7}{9}\end{aligned}$$

The probability of The Halting Problem winning the first game, given that they won the series is $\frac{7}{9}$.

This answer is suspicious! In the preceding section, we showed that $\Pr(A \mid B) = \frac{7}{9}$. Could it be true that $\Pr(A \mid B) = \Pr(B \mid A)$ in general ? We can determine the conditions under which this equality holds as follows:

$$\begin{aligned} & \Pr(A \mid B) &= \quad & \Pr(B \mid A) \\ \Leftrightarrow \quad & \frac{\Pr(A \cap B)}{\Pr(B)} &= \quad & \frac{\Pr(B \cap A)}{\Pr(A)} \\ \Leftrightarrow \quad & \Pr(B) = \Pr(A) & \text{or} \quad & \Pr(A \cap B) = \phi \end{aligned}$$

Evidently, $\Pr(A \mid B) = \Pr(B \mid A)$ only when A and B have equal probability or when the probability of the intersection is zero. This is true for the hockey problem, but only by coincidence. In general, $\Pr(A \mid B)$ and $\Pr(B \mid A)$ are *not* equal!

2.3 A Problem with Two Coins

We have two coins. One coin is fair; that is, comes up heads with probability $\frac{1}{2}$ and tails with probability $\frac{1}{2}$. The other is a trick coin; it has heads on both sides, and so always comes up heads.

Consider the following experiment. We choose a coin at random; in particular, we choose the fair coin with probability $\frac{1}{2}$. If we flip this coin and get heads, then what is the probability that we flipped the fair coin?

This is one of those tricky *a posteriori* problems, since we want the probability of an event (the fair coin was chosen) given the outcome of a later event (heads came up). Intuition may fail us, but the standard four-step method works perfectly well.

Step 1: Find the Sample Space

The sample space is worked out with a the tree diagram in Figure 3.

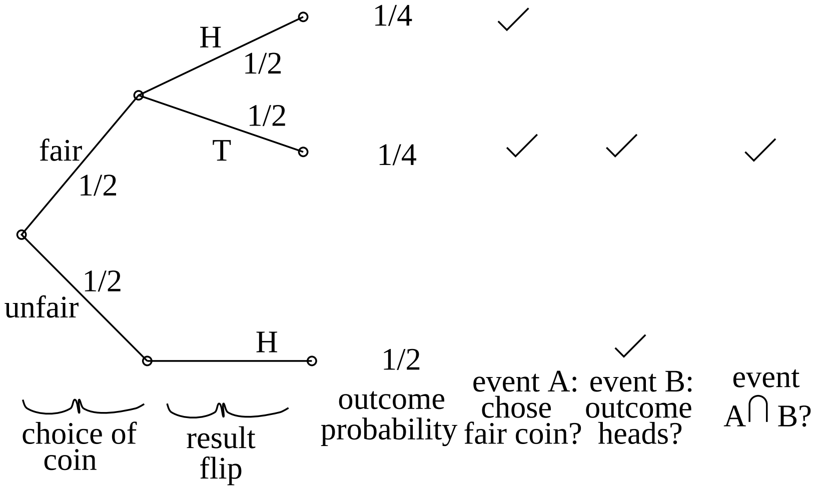

Figure 3: *What is the probability that we flipped the fair coin, given that the result was heads?*

Step 2: Define Events of Interest

Let A be the event that the fair coin was chosen. Let B the event that the result of the flip was heads. The outcomes in each event are marked in the figure. We want to compute $\Pr(A \mid B)$, the probability that the fair coin was chosen, given that the result of the flip was heads.

Step 2: Compute Outcome Probabilities

First, we assign probabilities to edges in the tree diagram. Each coin is chosen with probability $\frac{1}{2}$. If we choose the fair coin, then head and tails each come up with probability $\frac{1}{2}$. If we choose the trick coin, then heads comes up with probability 1. By the Product Rule, the probability of an outcome is the product of the probabilities on the corresponding root-to-leaf path. All of these probabilities are shown in Figure 3.

Step 4: Compute Event Probabilities

$$
\begin{aligned}
\Pr(A \mid B) &= \frac{\Pr(A \cap B)}{\Pr(B)} \\
&= \frac{\frac{1}{4}}{\frac{1}{4} + \frac{1}{2}} \\
&= \frac{1}{3}
\end{aligned}
$$

The first equation uses the Product Rule. In the second line, we use the fact that the probability of an event is the sum of the probabilities of the outcomes it contains. The final line is simplification. The probability that the fair coin was chosen, given that the result of the flip was heads, is $\frac{1}{3}$.

2.4 A Variant of the Two Coins Problem

Here is a variant of the two coins problem. Someone hands us either the fair coin or the trick coin, but we do not know which. We flip the coin 100 times and see heads every time. What can we say about the probability that we flipped the fair coin? Remarkably— nothing!

Suppose we were given the fair coin with probability p. In terms of p, we can find the probability that we flipped the fair coin, given that it comes up heads 100 times.

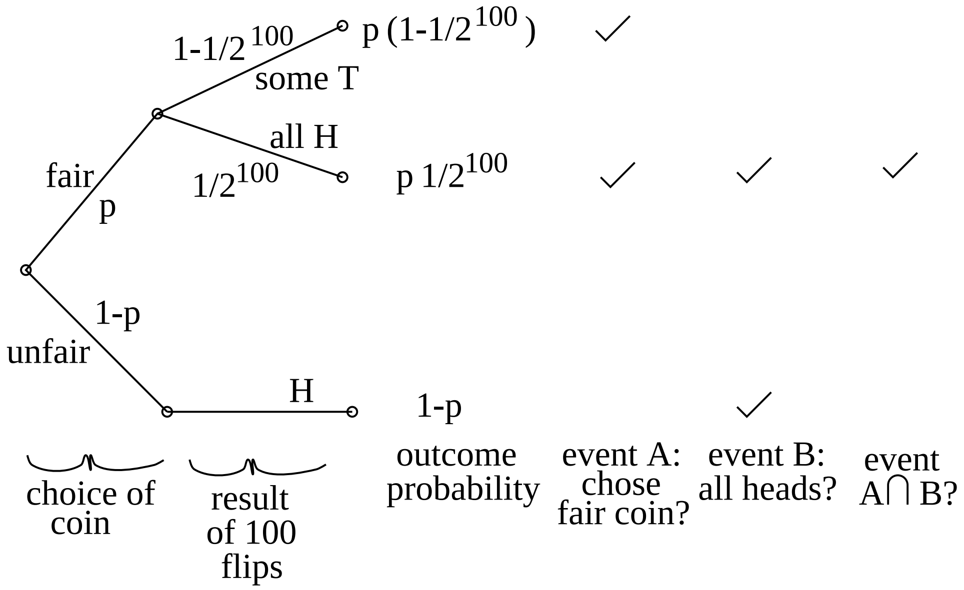

Figure 4: *We flip the coin 100 times and see heads every time. What can we say about the probability that we flipped the fair coin?*

The sample space is worked out in Figure 4. Let A be the event that we were handed the fair coin, and let B be the event that we flipped 100 heads. We want to compute $\Pr(A \mid B)$, the probability that we were handed the fair coin, given that we flipped 100 heads. The outcome probabilities are computed as shown in the figure. We can compute the probability of $\Pr(A \mid B)$ as follows.

$$\begin{aligned}\Pr(A \mid B) &= \frac{\Pr(A \cap B)}{\Pr(B)} \\ &= \frac{p2^{-100}}{1 - p + p2^{-100}} \\ &= \frac{p}{2^{100}(1-p) + p}\end{aligned}$$

Because of the 2^{100} term in the denominator, this expression is very small for reasonable values of p. For example, if p is $\frac{1}{2}$, then we almost certainly have the trick coin.

The problem is that we do not know the probability p with which we were given the fair coin; someone just handed us one or the other! For example, maybe $p = 1 - 2^{-100}$. Then there is only about a $\frac{1}{2}$ chance that we have the fair coin!

In fact, maybe we were handed the fair coin with probability $p = 1$. Then the above expression gives probability 1 that we have the fair coin! This makes sense; if we are given the fair coin with probability 1, then the fact that we flipped 100 heads is irrelevant!

A similar problem arises in polls around election time. A pollster picks a random American and ask his or her party affiliation. Suppose he repeats this experiment several hundred times and 60% of respondents say that they are Democrats. What can be said about the population as a whole? Nothing!

To make the analogy clear, suppose the country contains only two people. There is either one Democrat and one Republican (like the fair coin), or there are two Democrats (like the trick coin). The pollster picks a random citizen 100 times; this is analogous to flipping the coin 100 times. Even if he always picks a Democrat (flips heads), he can not determine the probability that the country is all Democratic!

Later we will see exactly what can be concluded from a poll.

2.5 Medical Testing

There is a deadly disease called X that has infected 10% of the population. There are no symptoms; victims just drop dead one day. Fortunately, there is a test for the disease. The test is not perfect, however:

- If you have the disease, there is a 10% chance that the test will say you do not. (These are called "false negatives".)
- If you do not have disease, there is a 30% chance that the test will say you do. (These are "false positives".)

A random person is tested for the disease. If the test is positive, then what is the probability that the person has the disease?

Step 1: Find the Sample Space

The sample space is found with a tree diagram in Figure 5.

Step 2: Define Events of Interest

Let A be the event that the person has the disease. Let B be the event that the test was positive. The outcomes in each event are marked in Figure 5. We want to find $\Pr(A \mid B)$, the probability that a person has disease X, given that the test was positive.

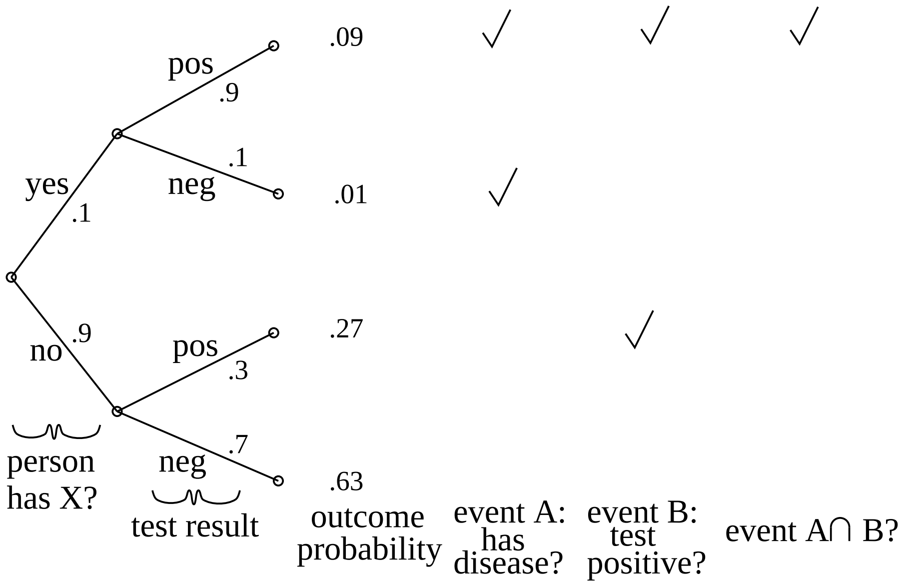

Figure 5: *What is the probability that a person has disease X, given that the test is positive?*

Step 3: Find Outcome Probabilities

First, we assign probabilities to edges. These probabilities are drawn directly from the problem statement. By the Product Rule, the probability of an outcome is the product of the probabilities on the corresponding root-to-leaf path. All probabilities are shown in the figure.

Step 4: Compute Event Probabilities

$$\begin{aligned} \Pr(A \mid B) &= \frac{\Pr(A \cap B)}{\Pr(B)} \\ &= \frac{0.09}{0.09 + 0.27} \\ &= \frac{1}{4} \end{aligned}$$

If you test positive, then there is only a 25% chance that you have the disease!

This answer is initially surprising, but makes sense on reflection. There are two ways you could test positive. First, you could be sick and the test correct. Second, you could be healthy and the test incorrect. The problem is that almost everyone is healthy; therefore, most of the positive results arise from incorrect tests of healthy people!

We can also compute the probability that the test is correct for a random person. This event consists of two outcomes. The person could be sick and the test positive (probability 0.09), or the person could be healthy and the test negative (probability 0.63). Therefore, the test is correct with probability $0.09 + 0.63 = 0.72$. This is a relief; the test is correct almost $\frac{3}{4}$ of the time.

But wait! There is a simple way to make the test correct 90% of the time: always return a negative result! This "test" gives the right answer for all healthy people and the wrong answer only for the 10% that actually have the disease. The best strategy is to completely ignore the test result!

There is a similar paradox in weather forecasting. During winter, almost all days in Boston are wet and overcast. Predicting miserable weather every day may be more accurate than really trying to get it right! This phenomenon is the source of many paradoxes; we will see more in coming weeks.

3 Conditional Probability Pitfalls

The remaining sections illustrate some common blunders involving conditional probability.

3.1 Carnival Dice

There is a gambling game called Carnival Dice. A player picks a number between 1 and 6 and then rolls three fair dice. The player wins if his number comes up on at least one die. The player loses if his number does not appear on any of the dice. What is the probability that the player wins? This problem sounds simple enough that we might try an intuitive lunge for the solution.

False Claim 3.1 The player wins with probability $\frac{1}{2}$.

Proof. Let A_i be the event that the i-th die matches the player's guess.

$$\begin{aligned}
\Pr(win) &= \Pr(A_1 \cup A_2 \cup A_3) \\
&= \Pr(A_1) \cup \Pr(A_2) \cup \Pr(A_3) \\
&= \frac{1}{6} + \frac{1}{6} + \frac{1}{6} \\
&= \frac{1}{2}
\end{aligned}$$

■

The only justification for the second equality is that it looks vaguely reasonable; in fact, equality does not hold. Let's examine the expression $\Pr(A_1 \cup A_2 \cup A_3)$ to see exactly what is happening. Recall that the probability of an event is the sum of the probabilities of the outcomes it contains. Therefore, we could argue as follows:

$$\begin{aligned}
\Pr(A_1 \cup A_2 \cup A_3) &= \sum^{w \in A_1 \cup A_2 \cup A_3} \Pr(w) \\
&= \sum^{w \in A_1} \Pr(w) + \sum^{w \in A_2} \Pr(w) + \sum^{w \in A_3} \Pr(w) \\
&= \Pr(A_1) + \Pr(A_2) + \Pr(A_3)
\end{aligned}$$

The argument is valid provided that the events A_1, A_2, and A_3 are disjoint; that is, there is no outcome in more than one event. If this were not true for some outcome, then a term would be duplicated when we split the one sum into three. Nevertheless, this argument proves the following theorem:

Theorem 3.2 Let $A_1, A_2, \ldots A_n$ be disjoint events.

$$\Pr(A_1 \cup A_2 \cup \ldots \cup A_n) = \Pr(A_1) + \Pr(A_2) + \ldots + \Pr(A_n)$$

We can evaluate the probability of a union of events that are not necessarily disjoint using a theorem analogous to Inclusion-Exclusion. Here is the special case for a union of three events.

Theorem 3.3 Let A_1, A_2, and A_3 be events, not necessarily disjoint.

$$\begin{aligned}\Pr(A_1 \cup A_2 \cup A_3) &= \Pr(A_1) + \Pr(A_2) + \Pr(A_3) \\ &\quad - \Pr(A_1 \cap A_2) - \Pr(A_1 \cap A_3) - \Pr(A_2 \cap A_3) \\ &\quad + \Pr(A_1 \cap A_2 \cap A_3)\end{aligned}$$

The proof is much the same as for Inclusion-Exclusion; details are omitted.

We can use this theorem to compute the chance of winning at Carnival Dice. The probability that one die matches the player's guess is $\frac{1}{6}$. The probability that two dice both match the player's guess is $\frac{1}{36}$. The probability that all three dice match is $\frac{1}{216}$. Plugging these numbers into the preceding theorem gives:

$$\begin{aligned}\Pr(\text{win}) &= \frac{1}{6} + \frac{1}{6} + \frac{1}{6} \\ &\quad - \frac{1}{36} - \frac{1}{36} - \frac{1}{36} \\ &\quad + \frac{1}{216} \\ &\approx 42\%\end{aligned}$$

These are terrible odds in a gambling game; it is much better to play roulette, craps, or blackjack!

3.2 Other Identities

There is a close relationship between computing the cardinality of a set and computing the probability of an event. Theorem 3.3 is an example; the probability of a union of events and the cardinality of a union of sets are computed in very similar ways.

In fact, all of the methods we developed for computing cardinalities carry over to computing probabilities. This is because a probability space is just a weighted set; the sample space is the set and the probability function assigns a weight to each element. Earlier, we were counting the number of items in a set. Now, when we compute the probability of an event, we are just summing the weights of items in a set. We will see many examples of the relationship between probability and counting in tutorial over the next few weeks.

Furthermore, identities carry over to conditional probabilities; that is, any law still holds if all probabilities are conditioned on the same event. For example, the following identity is analogous to the Inclusion-Exclusion formula for two sets, but all probabilities are conditioned on event C.

$$\Pr(A \cup B \mid C) = \Pr(A \mid C) + \Pr(A \mid C) - \Pr(A \cap B \mid C)$$

Be careful not to mix up events before and after the conditioning bar! For example, the following is *not* a consequence of the Sum Rule:

$$\Pr(A \mid B \cup C) = \Pr(A \mid B) + \Pr(A \mid C) \qquad (B \cap C = \phi)$$

A counterexample is shown in Figure 6. In this case, $\Pr(A \mid B) = 1$, $\Pr(A \mid C) = 1$, and $\Pr(A \mid B \cup C) = 1$. However, since $1 \neq 1 + 1$, the equation above does not hold. This is a "false identity". This seems clear now, but the following example will test our resolve!

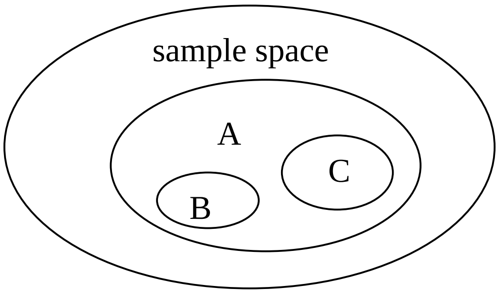

Figure 6: *This figure illustrates a case where the equation* $\Pr(A \mid B \cup C) = \Pr(A \mid B) + \Pr(A \mid C)$ *does not hold.*

3.3 Discrimination Lawsuit

Several years ago there was a sex discrimination lawsuit against Berkeley. A female professor was denied tenure, allegedly because she was a woman. She argued that in every one of Berkeley's 22 departments, the percentage of male applicants accepted was greater than the percentage of female applicants accepted. This sounds very suspicious!

However, Berkeley's lawyers argued that across the whole university the percentage of male applicants accepted was actually *lower* than the percentage of female applicants accepted! This suggests that if there was any sex discrimination, then it was against men! Must one party in the dispute be lying?

Let's simplify the problem and express both arguments in terms of conditional probabilities. Suppose that there are only two departments, EE and CS, and consider the experiment where we pick a random applicant. Define the following events:

- Let A be the event that the applicant is accepted.
- Let F_{EE} the event that the applicant is a female applying to EE.
- Let F_{CS} the event that the applicant is a female applying to CS.
- Let M_{EE} the event that the applicant is a male applying to EE.
- Let M_{CS} the event that the applicant is a male applying to CS.

Assume that all applicants are either male or female, and that no applicant applied to both departments. That is, the events F_{EE}, F_{CS}, M_{EE}, and M_{CS} are all disjoint.

The female plaintiff makes the following argument:

$$\begin{aligned} \Pr(A \mid F_{EE}) &< \Pr(A \mid M_{EE}) \\ \Pr(A \mid F_{CS}) &< \Pr(A \mid M_{CS}) \end{aligned}$$

That is, in both departments, the probability that a woman is accepted is less than the probability that a man is accepted. The university retorts that overall a woman applicant is *more* likely to be accepted than a man:

$$\Pr(A \mid F_{EE} \cup F_{CS}) > \Pr(A \mid M_{EE} \cup M_{CS})$$

It is easy to believe that these two positions are contradictory. In fact, we might even try to prove this by adding the plaintiff's two inequalities and then arguing as follows:

$$\begin{aligned} \Pr(A \mid F_{EE}) + \Pr(A \mid F_{CS}) &< \Pr(A \mid M_{EE}) + \Pr(A \mid M_{CS}) \\ \Rightarrow \qquad \Pr(A \mid F_{EE} \cup F_{CS}) &< \Pr(A \mid M_{EE} \cup M_{CS}) \end{aligned}$$

The second line exactly contradicts the university's position! However, there is a problem with this argument; the second inequality follows from the first only if we accept the "false identity" from the preceding section! Therefore, this argument is invalid; maybe the two parties do not hold contradictory positions after all!

In fact, the table below shows a set of application statistics for which the assertions of both the plaintiff and the university hold:

CS	0 females accepted, 1 applied	0%
	50 males accepted, 100 applied	50%
EE	70 females accepted, 100 applied	70%
	1 male accepted, 1 applied	100%
Overall	70 females accepted, 101 applied	$\approx$ 70%
	51 males accepted, 101 applied	$\approx$ 51%

In this case, a higher percentage of males were accepted in both departments, but overall a higher percentage of females were accepted! Bizarre!

3.4 On-Time Airlines

Newspapers publish on-time statistics for airlines to help travelers choose the best carrier. The on-time rate for an airline is defined as follows:

$$\text{Airline on-time rate} = \frac{\#\text{flights less than 15 minutes late}}{\#\text{flights total}}$$

This seems reasonable, but actually can be completely misleading! Here is some on-time data for two airlines in the late 80's.

	Alaska Air			America West		
Airport	#on-time	#flights	%	#on-time	#flights	%
Los Angeles	500	560	89	700	800	87
Phoenix	220	230	95	4900	5300	92
San Diego	210	230	92	400	450	89
San Francisco	500	600	83	320	450	71
Seattle	1900	2200	86	200	260	77
OVERALL	3330	3020	87	6520	7260	90

This is the same paradox as in the Berkeley lawsuit; America West has a better overall on-time percentage, but Alaska Airlines does a better job at every single airport! The problem is that Alaska Airlines flies proportionally more of its flights to bad weather airports like Seatte; whereas America West is based in fair-weather, low-traffic Phoenix!