CHAPTER 15

Probability Rules!

Pull a bill from your wallet or pocket without looking at it. An outcome of this trial is the bill you select. The sample space is all the bills in circulation: **S** = {\$1 bill, \$2 bill, \$5 bill, \$10 bill, \$20 bill, \$50 bill, \$100 bill}.[1] These are *all* the possible outcomes. (In spite of what you may have seen in bank robbery movies, there are no \$500 or \$1000 bills.)

We can combine the outcomes in different ways to make many different events. For example, the event **A** = {\$1, \$5, \$10} represents selecting a \$1, \$5, or \$10 bill. The event **B** = {a bill that does not have a president on it} is the collection of outcomes (Don't look! Can you name them?): {\$10 (Hamilton), \$100 (Franklin)}. The event **C** = {enough money to pay for a \$12 meal with one bill} is the set of outcomes {\$20, \$50, \$100}.

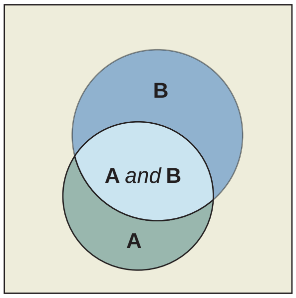
Notice that these outcomes are not equally likely. You'd no doubt be more surprised (and pleased) to pull out a \$100 bill than a \$1 bill—it's not very likely, though. You probably carry many more \$1 than \$100 bills, but without information about the probability of each outcome, we can't calculate the probability of an event.

The probability of the event **C** (getting a bill worth more than \$12) is *not* 3/7. There are 7 possible outcomes, and 3 of them exceed \$12, but they are not *equally likely*. (Remember the probability that your lottery ticket will win rather than lose still isn't 1/2.)

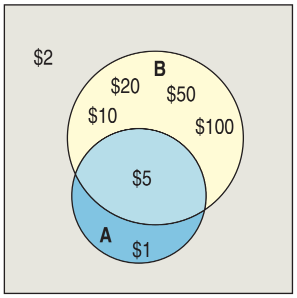
The General Addition Rule

Now look at the bill in your hand. There are images of famous buildings in the center of the backs of all but two bills in circulation. The \$1 bill has the word ONE in the center, and the \$2 bill shows the signing of the Declaration of Independence.

[1] Well, technically, the sample space is all the bills in your pocket. You may be quite sure there isn't a \$100 bill in there, but *we* don't know that, so humor us that it's at least *possible* that any legal bill could be there.



Events **A** *and* **B** *and their intersection.*



*Denominations of bills that are odd (***A***) or that have a building on the reverse side (***B***). The two sets both include the $5 bill, and both exclude the $2 bill.*

What's the probability of randomly selecting **A** = {a bill with an odd-numbered value} *or* **B** = {a bill with a building on the reverse}? We know **A** = {$1, $5} and **B** = {$5, $10, $20, $50, $100}. But *P*(**A** *or* **B**) is not simply the sum *P*(**A**) + *P*(**B**), because the events **A** and **B** are not disjoint. The $5 bill is in both sets. So what can we do? We'll need a new probability rule.

As the diagrams show, we can't use the Addition Rule and add the two probabilities because the events are not disjoint; they overlap. There's an outcome (the $5 bill) in the *intersection* of **A** and **B.** The Venn diagram represents the sample space. Notice that the $2 bill has neither a building nor an odd denomination, so it sits outside both circles.

The $5 bill plays a crucial role here because it is both odd *and* has a building on the reverse. It's in both **A** and **B,** which places it in the *intersection* of the two circles. The reason we can't simply add the probabilities of **A** and **B** is that we'd count the $5 bill twice.

If we did add the two probabilities, we could compensate by *subtracting* out the probability of that $5 bill. So,

$$\begin{aligned}&P(\text{odd number value or building})\\ &\quad= P(\text{odd number value}) + P(\text{building}) - P(\text{odd number value } \textit{and} \text{ building})\\ &\quad= P(\$1, \$5) + P(\$5, \$10, \$20, \$50, \$100) - P(\$5).\end{aligned}$$

This method works in general. We add the probabilities of two events and then subtract out the probability of their intersection. This approach gives us the **General Addition Rule,** which does not require disjoint events:

$$P(\mathbf{A} \cup \mathbf{B}) = P(\mathbf{A}) + P(\mathbf{B}) - P(\mathbf{A} \cap \mathbf{B}).$$

FOR EXAMPLE **Using the General Addition Rule**

A survey of college students found that 56% live in a campus residence hall, 62% participate in a campus meal program, and 42% do both.

Question: What's the probability that a randomly selected student either lives or eats on campus?

Let **L** = {student lives on campus} and **M** = {student has a campus meal plan}.

$$\begin{aligned}P(\text{a student either lives or eats on campus}) &= P(\mathbf{L} \cup \mathbf{M})\\ &= P(\mathbf{L}) + P(\mathbf{M}) - P(\mathbf{L} \cap \mathbf{M})\\ &= 0.56 + 0.62 - 0.42\\ &= 0.76\end{aligned}$$

There's a 76% chance that a randomly selected college student either lives or eats on campus.

Would you like dessert or coffee? Natural language can be ambiguous. In this question, is the answer one of the two alternatives, or simply "yes"? Must you decide between them, or may you have both? That kind of ambiguity can confuse our probabilities.

Suppose we had been asked a different question: What is the probability that the bill we draw has *either* an odd value *or* a building but *not both*? Which bills are we talking about now? The set we're interested in would be {$1, $10, $20, $50, $100}. We don't include the $5 bill in the set because it has both characteristics.

Why isn't this the same answer as before? The problem is that when we say the word "or," we usually mean *either* one *or* both. We don't usually mean the *exclusive* version of "or" as in, "Would you like the steak *or* the vegetarian entrée?" Ordinarily when we ask for the probability that **A** *or* **B** occurs, we mean **A** *or* **B** or both. And we know *that* probability is $P(\mathbf{A}) + P(\mathbf{B}) - P(\mathbf{A}\ and\ \mathbf{B})$. The General Addition Rule subtracts the probability of the outcomes in **A** *and* **B** because we've counted those outcomes *twice*. But they're still there.

If we really mean **A** or **B** but NOT both, we have to get rid of the outcomes in $\{\mathbf{A}\ and\ \mathbf{B}\}$. So $P(\mathbf{A}\ or\ \mathbf{B}\ \text{but}\ not\ \text{both}) = P(\mathbf{A} \cup \mathbf{B}) - P(\mathbf{A} \cap \mathbf{B}) = P(\mathbf{A}) + P(\mathbf{B}) - 2 \times P(\mathbf{A} \cap \mathbf{B})$. Now we've subtracted $P(\mathbf{A} \cap \mathbf{B})$ twice—once because we don't want to double-count these events and a second time because we really didn't want to count them at all.

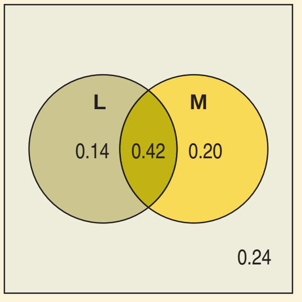
Confused? *Make a picture*. It's almost always easier to think about such situations by looking at a Venn diagram.

FOR EXAMPLE Using Venn diagrams

Recap: We return to our survey of college students: 56% live on campus, 62% have a campus meal program, and 42% do both.

Questions: Based on a Venn diagram, what is the probability that a randomly selected student

a) lives off campus and doesn't have a meal program?
b) lives in a residence hall but doesn't have a meal program?



Let **L** = {student lives on campus} and **M** = {student has a campus meal plan}. In the Venn diagram, the intersection of the circles is $P(\mathbf{L} \cap \mathbf{M}) = 0.42$. Since $P(\mathbf{L}) = 0.56$, $P(\mathbf{L} \cap \mathbf{M}^C) = 0.56 - 0.42 = 0.14$. Also, $P(\mathbf{L}^C \cap \mathbf{M}) = 0.62 - 0.42 = 0.20$. Now, $0.14 + 0.42 + 0.20 = 0.76$, leaving $1 - 0.76 = 0.24$ for the region outside both circles.

Now . . . $P(\text{off campus and no meal program}) = P(\mathbf{L}^C \cap \mathbf{M}^C) = 0.24$

$P(\text{on campus and no meal program}) = P(\mathbf{L} \cap \mathbf{M}^C) = 0.14$

JUST CHECKING

1. Back in Chapter 1 we suggested that you sample some pages of this book at random to see whether they held a graph or other data display. We actually did just that. We drew a representative sample and found the following:

48% of pages had some kind of data display,

27% of pages had an equation, and

7% of pages had both a data display and an equation.

a) Display these results in a Venn diagram.
b) What is the probability that a randomly selected sample page had neither a data display nor an equation?
c) What is the probability that a randomly selected sample page had a data display but no equation?

STEP-BY-STEP EXAMPLE Using the General Addition Rule

Police report that 78% of drivers stopped on suspicion of drunk driving are given a breath test, 36% a blood test, and 22% both tests.

Question: What is the probability that a randomly selected DWI suspect is given

1. a test?
2. a blood test or a breath test, but not both?
3. neither test?

THINK

Plan Define the events we're interested in. There are no conditions to check; the General Addition Rule works for any events!

Plot Make a picture, and use the given probabilities to find the probability for each region.

The blue region represents **A** but not **B**. The green intersection region represents **A** *and* **B**. Note that since $P(\mathbf{A}) = 0.78$ and $P(\mathbf{A} \cap \mathbf{B}) = 0.22$, the probability of **A** but not **B** must be $0.78 - 0.22 = 0.56$.

The yellow region is **B** but not **A**.

The gray region outside both circles represents the outcome neither **A** nor **B**. All the probabilities must total 1, so you can determine the probability of that region by subtraction.

Now, figure out what you want to know. The probabilities can come from the diagram or a formula. Sometimes translating the words to equations is the trickiest step.

Let A = {suspect is given a breath test}.

Let B = {suspect is given a blood test}.

I know that $P(A) = 0.78$

$P(B) = 0.36$

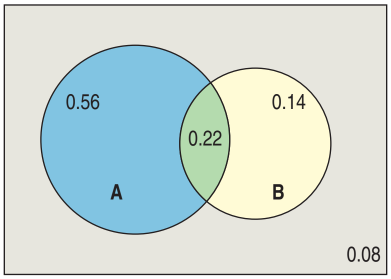
$P(A \cap B) = 0.22$

So $P(A \cap B^C) = 0.78 - 0.22 = 0.56$

$P(B \cap A^C) = 0.36 - 0.22 = 0.14$

$P(A^C \cap B^C) = 1 - (0.56 + 0.22 + 0.14)$

$= 0.08$



Question 1. What is the probability that the suspect is given a test?

SHOW

Mechanics The probability the suspect is given a test is $P(\mathbf{A} \cup \mathbf{B})$. We can use the General Addition Rule, or we can add the probabilities seen in the diagram.

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$
$$= 0.78 + 0.36 - 0.22$$
$$= 0.92$$

OR

$$P(A \cup B) = 0.56 + 0.22 + 0.14 = 0.92$$

TELL

Conclusion Don't forget to interpret your result in context.

92% of all suspects are given a test.

Question 2. What is the probability that the suspect gets either a blood test or a breath test but NOT both?

SHOW

Mechanics We can use the rule, or just add the appropriate probabilities seen in the Venn diagram.

$$P(\mathbf{A} \text{ or } \mathbf{B} \text{ but NOT both}) = P(\mathbf{A} \cup \mathbf{B}) - P(\mathbf{A} \cap \mathbf{B}) = 0.92 - 0.22 = 0.70$$

OR

$$P(\mathbf{A} \text{ or } \mathbf{B} \text{ but NOT both}) = P(\mathbf{A} \cap \mathbf{B}^C) + P(\mathbf{B} \cap \mathbf{A}^C) = 0.56 + 0.14 = 0.70$$

TELL

Conclusion Interpret your result in context.

70% of the suspects get exactly one of the tests.

Question 3. What is the probability that the suspect gets neither test?

SHOW

Mechanics Getting neither test is the complement of getting one or the other. Use the Complement Rule or just notice that "neither test" is represented by the region outside both circles.

$$P(\text{neither test}) = 1 - P(\text{either test}) = 1 - P(\mathbf{A} \cup \mathbf{B}) = 1 - 0.92 = 0.08$$

OR

$$P(\mathbf{A}^C \cap \mathbf{B}^C) = 0.08$$

TELL

Conclusion Interpret your result in context.

Only 8% of the suspects get no test.

It Depends . . .

Two psychologists surveyed 478 children in grades 4, 5, and 6 in elementary schools in Michigan. They stratified their sample, drawing roughly 1/3 from rural, 1/3 from suburban, and 1/3 from urban schools. Among other questions, they asked the students whether their primary goal was to get good grades, to be popular, or to be good at sports. One question of interest was whether boys and girls at this age had similar goals.

Here's a *contingency table* giving counts of the students by their goals and sex:

		Goals			
		Grades	**Popular**	**Sports**	Total
Sex	**Boy**	117	50	60	**227**
	Girl	130	91	30	**251**
	Total	**247**	**141**	**90**	**478**

Table 15.1
The distribution of goals for boys and girls.

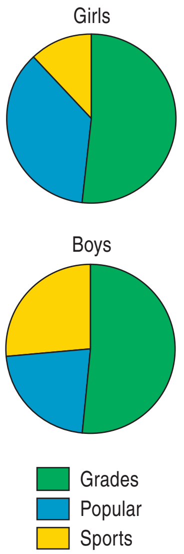


FIGURE 15.1

The distribution of goals for boys and girls.

We looked at contingency tables and graphed *conditional distributions* back in Chapter 3. The pie charts show the *relative frequencies* with which boys and girls named the three goals. It's only a short step from these relative frequencies to probabilities.

Let's focus on this study and make the sample space just the set of these 478 students. If we select a student at random from this study, the probability we select a girl is just the corresponding relative frequency (since we're equally likely to select any of the 478 students). There are 251 girls in the data out of a total of 478, giving a probability of

$$P(\text{girl}) = 251/478 = 0.525$$

The same method works for more complicated events like intersections. For example, what's the probability of selecting a girl whose goal is to be popular? Well, 91 girls named popularity as their goal, so the probability is

$$P(\text{girl} \cap \text{popular}) = 91/478 = 0.190$$

The probability of selecting a student whose goal is to excel at sports is

$$P(\text{sports}) = 90/478 = 0.188$$

What if we are given the information that the selected student is a girl? Would that change the probability that the selected student's goal is sports? You bet it would! The pie charts show that girls are much less likely to say their goal is to excel at sports than are boys. When we restrict our focus to girls, we look only at the girls' row of the table. Of the 251 girls, only 30 of them said their goal was to excel at sports.

We write the probability that a selected student wants to excel at sports *given that we have selected a girl* as

$$P(\text{sports} \mid \text{girl}) = 30/251 = 0.120$$

For boys, we look at the conditional distribution of goals given "boy" shown in the top row of the table. There, of the 227 boys, 60 said their goal was to excel at sports. So, $P(\text{sports} \mid \text{boy}) = 60/227 = 0.264$, more than twice the girls' probability.

A S ***Activity:*** **Birthweights and Smoking.** Does smoking increase the chance of having a baby with low birth weight?

In general, when we want the probability of an event from a *conditional* distribution, we write $P(\mathbf{B} \mid \mathbf{A})$ and pronounce it "the probability of **B** *given* **A.**" A probability that takes into account a given *condition* such as this is called a **conditional probability.**

Let's look at what we did. We worked with the counts, but we could work with the probabilities just as well. There were 30 students who both were girls and had sports as their goal, and there are 251 girls. So we found the probability to be 30/251. To find the probability of the event **B** *given* the event **A,** we restrict our attention to the outcomes in **A.** We then find in what fraction of *those* outcomes **B** also occurred. Formally, we write:

NOTATION ALERT:

$P(\mathbf{B} \mid \mathbf{A})$ is the conditional probability of **B** *given* **A.**

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

Thinking this through, we can see that it's just what we've been doing, but now with probabilities rather than with counts. Look back at the girls for whom sports was the goal. How did we calculate $P(\text{sports} \mid \text{girl})$?

The rule says to use probabilities. It says to find $P(\mathbf{A} \cap \mathbf{B})/P(\mathbf{A})$. The result is the same whether we use counts or probabilities because the total number in the sample cancels out:

$$\frac{P(\text{sports} \cap \text{girl})}{\text{P(girl)}} = \frac{30/478}{251/478} = \frac{30}{251}.$$

A S *Activity:* **Conditional Probability.** Simulation is great for seeing conditional probabilities at work.

To use the formula for conditional probability, we're supposed to insist on one restriction. The formula doesn't work if $P(\mathbf{A})$ is 0. After all, we can't be "given" the fact that **A** was true if the probability of **A** is 0!

Let's take our rule out for a spin. What's the probability that we have selected a girl *given* that the selected student's goal is popularity? Applying the rule, we get

$$P(\text{girl} \mid \text{popular}) = \frac{P(\text{girl} \cap \text{popular})}{P(\text{popular})} = \frac{91/478}{141/478} = \frac{91}{141}.$$

FOR EXAMPLE Finding a conditional probability

Recap: Our survey found that 56% of college students live on campus, 62% have a campus meal program, and 42% do both.

Question: While dining in a campus facility open only to students with meal plans, you meet someone interesting. What is the probability that your new acquaintance lives on campus?

Let $\mathbf{L}$ = {student lives on campus} and $\mathbf{M}$ = {student has a campus meal plan}.

$$P(\text{student lives on campus given that the student has a meal plan}) = P(\mathbf{L} \mid \mathbf{M}) = \frac{P(\mathbf{L} \cap \mathbf{M})}{P(\mathbf{M})} = \frac{0.42}{0.62} \approx 0.677$$

There's a probability of about 0.677 that a student with a meal plan lives on campus.

The General Multiplication Rule

Remember the Multiplication Rule for the probability of **A** *and* **B**? It said

$$P(\mathbf{A} \cap \mathbf{B}) = P(\mathbf{A}) \times P(\mathbf{B}) \text{ when } \mathbf{A} \text{ and } \mathbf{B} \text{ are independent.}$$

Now we can write a more general rule that doesn't require independence. In fact, we've *already* written it down. We just need to rearrange the equation a bit.

The equation in the definition for conditional probability contains the probability of **A** *and* **B.** Rewriting the equation gives

$$P(\mathbf{A} \cap \mathbf{B}) = P(\mathbf{A}) \times P(\mathbf{B} \mid \mathbf{A}).$$

A S *Activity:* **The General Multiplication Rule.** The best way to understand the General Multiplication Rule is with an experiment.

This is a **General Multiplication Rule** for compound events that does not require the events to be independent. Better than that, it even makes sense. The probability that two events, **A** and **B,** *both* occur is the probability that event **A** occurs multiplied by the probability that event **B** *also* occurs—that is, by the probability that event **B** occurs *given* that event **A** occurs.

Of course, there's nothing special about which set we call **A** and which one we call **B.** We should be able to state this the other way around. And indeed we can. It is equally true that

$$P(\mathbf{A} \cap \mathbf{B}) = P(\mathbf{B}) \times P(\mathbf{A} \mid \mathbf{B}).$$

Independence

If we had to pick one idea in this chapter that you should understand and remember, it's the definition and meaning of independence. We'll need this idea in every one of the chapters that follow.

Let's return to the question of just what it means for events to be independent. We've said informally that what we mean by independence is that the outcome of one event does not influence the probability of the other. With our new notation for conditional probabilities, we can write a formal definition: Events **A** and **B** are **independent** whenever

$$P(\mathbf{B} \mid \mathbf{A}) = P(\mathbf{B}).$$

Now we can see that the Multiplication Rule for independent events we saw in Chapter 14 is just a special case of the General Multiplication Rule. The general rule says

$$P(\mathbf{A} \cap \mathbf{B}) = P(\mathbf{A}) \times P(\mathbf{B} \mid \mathbf{A}).$$

whether the events are independent or not. But when events **A** and **B** are independent, we can write $P(\mathbf{B})$ for $P(\mathbf{B} \mid \mathbf{A})$ and we get back our simple rule:

$$P(\mathbf{A} \cap \mathbf{B}) = P(\mathbf{A}) \times P(\mathbf{B}).$$

Activity: **Independence.** Are *Smoking and Low Birthweight* independent?

Sometimes people use this statement as the definition of independent events, but we find the other definition more intuitive. Either way, the idea is that for independent events, the probability of one doesn't change when the other occurs.

Is the probability of having good grades as a goal independent of the sex of the responding student? Looks like it might be. We need to check whether

$$P(\text{grades} \mid \text{girl}) = P(\text{grades})$$

$$\frac{130}{251} = 0.52 \stackrel{?}{=} \frac{247}{478} = 0.52$$

To two decimal place accuracy, it looks like we can consider choosing good grades as a goal to be independent of sex.

On the other hand, $P(\text{sports})$ is 90/478, or about 18.8%, but $P(\text{sports} \mid \text{boy})$ is $60/227 = 26.4\%$. Because these probabilities aren't equal, we can be pretty sure that choosing success in sports as a goal is not independent of the student's sex.

In earlier chapters we said informally that two events were independent if learning that one occurred didn't change what you thought about the other occurring. Now we can be more formal. Events **A** and **B** are independent if (and only if) the probability of **A** is the same when we are given that **B** has occurred. That is, $P(\mathbf{A}) = P(\mathbf{A} \mid \mathbf{B})$.

Although sometimes your intuition is enough, now that we have the formal rule, use it whenever you can.

FOR EXAMPLE *Checking for independence*

Recap: Our survey told us that 56% of college students live on campus, 62% have a campus meal program, and 42% do both.

Question: Are living on campus and having a meal plan independent? Are they disjoint?

Let $\mathbf{L}$ = {student lives on campus} and $\mathbf{M}$ = {student has a campus meal plan}. If these events are independent, then knowing that a student lives on campus doesn't affect the probability that he or she has a meal plan. I'll check to see if $P(\mathbf{M} \mid \mathbf{L}) = P(\mathbf{M})$:

$$P(\mathbf{M} \mid \mathbf{L}) = \frac{P(\mathbf{L} \cap \mathbf{M})}{P(\mathbf{L})} = \frac{0.42}{0.56} = 0.75, \quad \text{but } P(\mathbf{M}) = 0.62.$$

Because $0.75 \neq 0.62$, the events are not independent; students who live on campus are more likely to have meal plans. Living on campus and having a meal plan are not disjoint either; in fact, 42% of college students do both.

Independent ≠ Disjoint

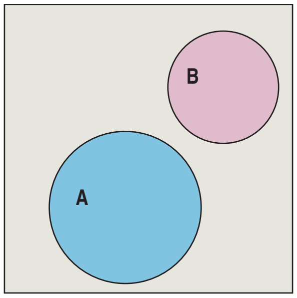


FIGURE 15.2

Because these events are mutually exclusive, learning that **A** *happened tells us that* **B** *didn't. The probability of* **B** *has changed from whatever it was to zero. So the disjoint events* **A** *and* **B** *are not independent.*

Are disjoint events independent? These concepts seem to have similar ideas of separation and distinctness about them, but in fact disjoint events *cannot* be independent.[2] Let's see why. Consider the two disjoint events {you get an A in this course} and {you get a B in this course}. They're disjoint because they have no outcomes in common. Suppose you learn that you *did* get an A in the course. Now what is the probability that you got a B? You can't get both grades, so it must be 0.

Think about what that means. Knowing that the first event (getting an A) occurred changed your probability for the second event (down to 0). So these events aren't independent.

Mutually exclusive events can't be independent. They have no outcomes in common, so if one occurs, the other doesn't. A common error is to treat disjoint events as if they were independent and apply the Multiplication Rule for independent events. Don't make that mistake.

JUST CHECKING

2. The American Association for Public Opinion Research (AAPOR) is an association of about 1600 individuals who share an interest in public opinion and survey research. They report that typically as few as 10% of random phone calls result in a completed interview. Reasons are varied, but some of the most common include no answer, refusal to cooperate, and failure to complete the call.

Which of the following events are independent, which are disjoint, and which are neither independent nor disjoint?

a) **A** = Your telephone number is randomly selected. **B** = You're not at home at dinnertime when they call.
b) **A** = As a selected subject, you complete the interview. **B** = As a selected subject, you refuse to cooperate.
c) **A** = You are not at home when they call at 11 a.m. **B** = You are employed full-time.

Depending on Independence

A S *Video:* **Is There a Hot Hand in Basketball?** Most coaches and fans believe that basketball players sometimes get "hot" and make more of their shots. What do the conditional probabilities say?

A S *Activity:* **Hot Hand Simulation.** Can you tell the difference between real and simulated sequences of basketball shot hits and misses?

It's much easier to think about independent events than to deal with conditional probabilities. It seems that most people's natural intuition for probabilities breaks down when it comes to conditional probabilities. Someone may estimate the probability of a compound event by multiplying the probabilities of its component events together without asking seriously whether those probabilities are independent.

For example, experts have assured us that the probability of a major commercial nuclear plant failure is so small that we should not expect such a failure to occur even in a span of hundreds of years. After only a few decades of commercial nuclear power, however, the world has seen two failures (Chernobyl and Three Mile Island). How could the estimates have been so wrong?

[2] Well, technically two disjoint events *can* be independent, but only if the probability of one of the events is 0. For practical purposes, though, we can ignore this case. After all, as statisticians we don't anticipate having data about things that never happen.

One simple part of the failure calculation is to test a particular valve and determine that valves such as this one fail only once in, say, 100 years of normal use. For a coolant failure to occur, several valves must fail. So we need the compound probability, P(valve 1 fails *and* valve 2 fails *and* . . .). A simple risk assessment might multiply the small probability of one valve failure together as many times as needed.

But if the valves all came from the same manufacturer, a flaw in one might be found in the others. And maybe when the first fails, it puts additional pressure on the next one in line. In either case, the events aren't independent and so we can't simply multiply the probabilities together.

Whenever you see probabilities multiplied together, stop and ask whether you think they are really independent.

Tables and Conditional Probability

One of the easiest ways to think about conditional probabilities is with contingency tables. We did that earlier in the chapter when we began our discussion. But sometimes we're given probabilities without a table. You can often construct a simple table to correspond to the probabilities.

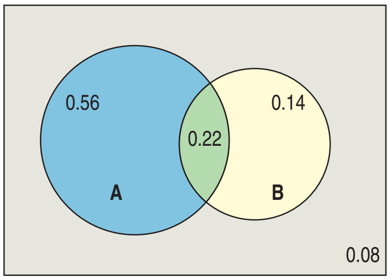
For instance, in the drunk driving example, we were told that 78% of suspect drivers get a breath test, 36% a blood test, and 22% both. That's enough information. Translating percentages to probabilities, what we know looks like this:

		Breath Test		
		Yes	**No**	Total
Blood Test	**Yes**	0.22		**0.36**
	No			
	Total	**0.78**		**1.00**

Notice that the 0.78 and 0.36 are *marginal* probabilities and so they go into the *margins.* The 0.22 is the probability of getting both tests—a breath test *and* a blood test—so that's a *joint* probability. Those belong in the interior of the table.

Because the cells of the table show disjoint events, the probabilities always add to the marginal totals going across rows or down columns. So, filling in the rest of the table is quick:

		Breath Test		
		Yes	**No**	Total
Blood Test	**Yes**	0.22	0.14	**0.36**
	No	0.56	0.08	**0.64**
	Total	**0.78**	**0.22**	**1.00**



Compare this with the Venn diagram. Notice which entries in the table match up with the sets in this diagram. Whether a Venn diagram or a table is better to use will depend on what you are given and the questions you're being asked. Try both.

STEP-BY-STEP EXAMPLE — Are the Events Disjoint? Independent?

Let's take another look at the drunk driving situation. Police report that 78% of drivers are given a breath test, 36% a blood test, and 22% both tests.

Questions: 1. Are giving a DWI suspect a blood test and a breath test mutually exclusive?
2. Are giving the two tests independent?

THINK

Plan Define the events we're interested in.

State the given probabilities.

Let A = {suspect is given a breath test}
Let B = {suspect is given a blood test}.

I know that $P(A) = 0.78$
$P(B) = 0.36$
$P(A \cap B) = 0.22$

Question 1. Are giving a DWI suspect a blood test and a breath test mutually exclusive?

SHOW

Mechanics Disjoint events cannot *both* happen at the same time, so check to see if $P(\mathbf{A} \cap \mathbf{B}) = 0$.

$P(A \cap B) = 0.22$. Since some suspects are given both tests, $P(A \cap B) \neq 0$. The events are not mutually exclusive.

TELL

Conclusion State your conclusion in context.

22% of all suspects get both tests, so a breath test and a blood test are not disjoint events.

Question 2. Are the two tests independent?

THINK

Plan Make a table.

Blood Test	Breath Test: Yes	Breath Test: No	Total
Yes	0.22	0.14	0.36
No	0.56	0.08	0.64
Total	0.78	0.22	1.00

SHOW

Mechanics Does getting a breath test change the probability of getting a blood test? That is, does $P(\mathbf{B} \mid \mathbf{A}) = P(\mathbf{B})$?

Because the two probabilities are *not* the same, the events are not independent.

$$P(B \mid A) = \frac{P(A \cap B)}{P(A)} = \frac{0.22}{0.78} \approx 0.28$$

$P(B) = 0.36$

$P(B \mid A) \neq P(B)$

TELL **Conclusion** Interpret your results in context.	Overall, 36% of the drivers get blood tests, but only 28% of those who get a breath test do. Since suspects who get a breath test are less likely to have a blood test, the two events are not independent.

JUST CHECKING

3. Remember our sample of pages in this book from the earlier Just Checking . . . ?

48% of pages had a data display.

27% of pages had an equation, and

7% of pages had both a data display and an equation.

a) Make a contingency table for the variables *display* and *equation*.
b) What is the probability that a randomly selected sample page with an equation also had a data display?
c) Are having an equation and having a data display disjoint events?
d) Are having an equation and having a data display independent events?

Drawing Without Replacement

Room draw is a process for assigning dormitory rooms to students who live on campus. Sometimes, when students have equal priority, they are randomly assigned to the currently available dorm rooms. When it's time for you and your friend to draw, there are 12 rooms left. Three are in Gold Hall, a very desirable dorm with spacious wood-paneled rooms. Four are in Silver Hall, centrally located but not quite as desirable. And five are in Wood Hall, a new dorm with cramped rooms, located half a mile from the center of campus on the edge of the woods.

You get to draw first, and then your friend will draw. Naturally, you would both like to score rooms in Gold. What are your chances? In particular, what's the chance that you *both* can get rooms in Gold?

When you go first, the chance that *you* will draw one of the Gold rooms is 3/12. Suppose you do. Now, with you clutching your prized room assignment, what chance does your friend have? At this point there are only 11 rooms left and just 2 left in Gold, so your friend's chance is now 2/11.

Using our notation, we write

$$P(\text{friend draws Gold} \mid \text{you draw Gold}) = 2/11.$$

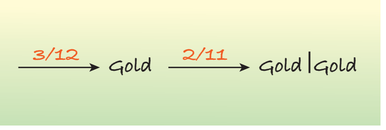
The reason the denominator changes is that we draw these rooms *without replacement*. That is, once one is drawn, it doesn't go back into the pool.

We often sample without replacement. When we draw from a very large population, the change in the denominator is too small to worry about. But when there's a small population to draw from, as in this case, we need to take note and adjust the probabilities.

What are the chances that *both* of you will luck out? Well, now we've calculated the two probabilities we need for the General Multiplication Rule, so we can write:

$$P(\text{you draw Gold} \cap \text{friend draws Gold})$$
$$= P(\text{you draw Gold}) \times P(\text{friend draws Gold} \mid \text{you draw Gold})$$
$$= 3/12 \times 2/11 = 1/22 = 0.045$$

In this instance, it doesn't matter who went first, or even if the rooms were drawn simultaneously. Even if the room draw was accomplished by shuffling cards containing the names of the dormitories and then dealing them out to 12 applicants (rather than by each student drawing a room in turn), we can still *think* of the calculation as having taken place in two steps:



Diagramming conditional probabilities leads to a more general way of helping us think with pictures—one that works for calculating conditional probabilities even when they involve different variables.

Tree Diagrams

For men, binge drinking is defined as having five or more drinks in a row, and for women as having four or more drinks in a row. (The difference is because of the average difference in weight.) According to a study by the Harvard School of Public Health (H. Wechsler, G. W. Dowdall, A. Davenport, and W. DeJong, "Binge Drinking on Campus: Results of a National Study"), 44% of college students engage in binge drinking, 37% drink moderately, and 19% abstain entirely. Another study, published in the *American Journal of Health Behavior,* finds that among binge drinkers aged 21 to 34, 17% have been involved in an alcohol-related automobile accident, while among non-bingers of the same age, only 9% have been involved in such accidents.

What's the probability that a randomly selected college student will be a binge drinker who has had an alcohol-related car accident?

To start, we see that the probability of selecting a binge drinker is about 44%. To find the probability of selecting someone who is both a binge drinker and a driver with an alcohol-related accident, we would need to pull out the General Multiplication Rule and multiply the probability of one of the events by the conditional probability of the other given the first.

Or we *could* make a picture. Which would you prefer?

We thought so.

The kind of picture that helps us think through this kind of reasoning is called a **tree diagram,** because it shows sequences of events, like those we had in room draw, as paths that look like branches of a tree. It is a good idea to make a tree diagram almost any time you plan to use the General Multiplication Rule. The number of different paths we can take can get large, so we usually draw the tree starting from the left and growing vine-like across the page, although sometimes you'll see them drawn from the bottom up or top down.

"Why," said the Dodo, "the best way to explain it is to do it."
—Lewis Carroll

The first branch of our tree separates students according to their drinking habits. We label each branch of the tree with a possible outcome and its corresponding probability.

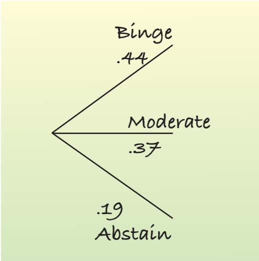


FIGURE 15.3
We can diagram the three outcomes of drinking and indicate their respective probabilities with a simple tree diagram.

Notice that we cover all possible outcomes with the branches. The probabilities add up to one. But we're also interested in car accidents. The probability of having an alcohol-related accident *depends* on one's drinking behavior. Because the probabilities are *conditional,* we draw the alternatives separately on each branch of the tree:

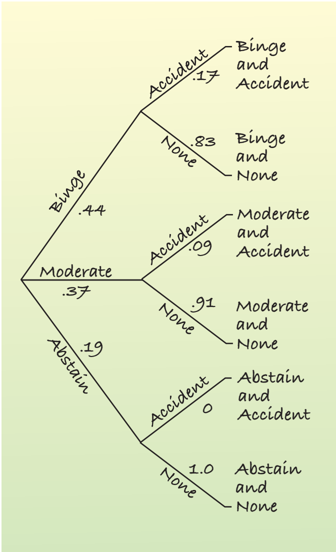


FIGURE 15.4
Extending the tree diagram, we can show both drinking and accident outcomes. The accident probabilities are conditional on the drinking outcomes, and they change depending on which branch we follow. Because we are concerned only with alcohol-related *accidents, the conditional probability P(accident | abstinence) must be 0.*

On each of the second set of branches, we write the possible outcomes associated with having an alcohol-related car accident (having an accident or not) and the associated probability. These probabilities are different because they are *conditional* depending on the student's drinking behavior. (It shouldn't be too surprising that those who binge drink have a higher probability of alcohol-related accidents.) The probabilities add up to one, because given the outcome on the first branch, these outcomes cover all the possibilities. Looking back at the General Multiplication Rule, we can see how the tree depicts the calculation. To find the probability that a randomly selected student will be a binge drinker who has had an alcohol-related car accident, we follow the top branches. The probability of selecting a binger is 0.44. The conditional probability of an accident *given* binge drinking is 0.17. The General Multiplication Rule tells us that to find the *joint* probability of being a binge drinker and having an accident, we multiply these two probabilities together:

$$\begin{aligned} P(\text{binge} \cap \text{accident}) &= P(\text{binge}) \times P(\text{accident} \mid \text{binge}) \\ &= 0.44 \times 0.17 = 0.075 \end{aligned}$$

And we can do the same for each combination of outcomes:

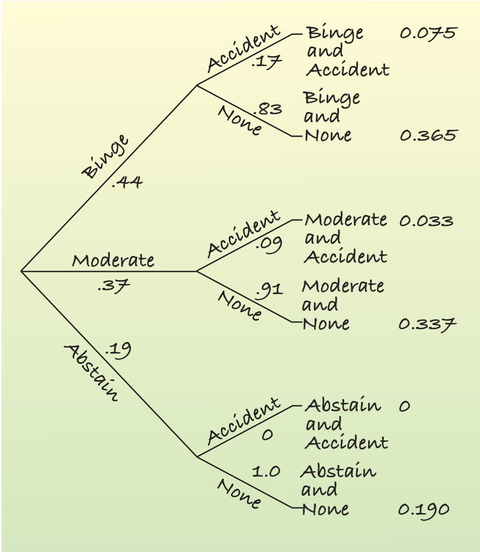


Figure 15.5

We can find the probabilities of compound events by multiplying the probabilities along the branch of the tree that leads to the event, just the way the General Multiplication Rule specifies.

The probability of abstaining and having an alcohol-related accident is, of course, zero.

All the outcomes at the far right are disjoint because at each branch of the tree we chose between disjoint alternatives. And they are *all* the possibilities, so the probabilities on the far right must add up to one. Always check!

Because the final outcomes are disjoint, we can add up their probabilities to get probabilities for compound events. For example, what's the probability that a selected student has had an alcohol-related car accident? We simply find *all* the outcomes on the far right in which an accident has happened. There are three and we can add their probabilities: $0.075 + 0.033 + 0 = 0.108$—almost an 11% chance.

Reversing the Conditioning

If we know a student has had an alcohol-related accident, what's the probability that the student is a binge drinker? That's an interesting question, but we can't just read it from the tree. The tree gives us $P(\text{accident} \mid \text{binge})$, but we want $P(\text{binge} \mid \text{accident})$—conditioning in the other direction. The two probabilities are definitely *not* the same. We have reversed the conditioning.

We may not have the conditional probability we want, but we do know everything we need to know to find it. To find a conditional probability, we need the probability that both events happen divided by the probability that the given event occurs. We have already found the probability of an alcohol-related accident: $0.075 + 0.033 + 0 = 0.108$.

The joint probability that a student is both a binge drinker and someone who's had an alcohol-related accident is found at the top branch: 0.075. We've restricted the *Who* of the problem to the students with alcohol-related accidents, so we divide the two to find the conditional probability:

$$P(\text{binge} \mid \text{accident}) = \frac{P(\text{binge} \cap \text{accident})}{P(\text{accident})}$$

$$= \frac{0.075}{0.108} = 0.694$$

The chance that a student who has an alcohol-related car accident is a binge drinker is more than 69%! As we said, reversing the conditioning is rarely intuitive, but tree diagrams help us keep track of the calculation when there aren't too many alternatives to consider.

STEP-BY-STEP EXAMPLE

Reversing the Conditioning

When the authors were in college, there were only three requirements for graduation that were the same for all students: You had to be able to tread water for 2 minutes, you had to learn a foreign language, and you had to be free of tuberculosis. For the last requirement, all freshmen had to take a TB screening test that consisted of a nurse jabbing what looked like a corncob holder into your forearm. You were then expected to report back in 48 hours to have it checked. If you were healthy and TB-free, your arm was supposed to look as though you'd never had the test.

Sometime during the 48 hours, one of us had a reaction. When he finally saw the nurse, his arm was about 50% bigger than normal and a very unhealthy red. Did he have TB? The nurse had said that the test was about 99% effective, so it seemed that the chances must be pretty high that he had TB. How high do you think the chances were? Go ahead and guess. Guess low.

We'll call **TB** the event of actually having TB and + the event of testing positive. To start a tree, we need to know $P(\mathbf{TB})$, the probability of having TB.[3] We also need to know the conditional probabilities $P(+\,|\,\mathbf{TB})$ and $P(+\,|\,\mathbf{TB^C})$. Diagnostic tests can make two kinds of errors. They can give a positive result for a healthy person (a *false positive*) or a negative result for a sick person (a *false negative*). Being 99% accurate usually means a false-positive rate of 1%. That is, someone who doesn't have the disease has a 1% chance of testing positive anyway. We can write $P(+\,|\,\mathbf{TB^C}) = 0.01$.

Since a false negative is more serious (because a sick person might not get treatment), tests are usually constructed to have a lower false-negative rate. We don't know exactly, but let's assume a 0.1% false-negative rate. So only 0.1% of sick people test negative. We can write $P(-\,|\,\mathbf{TB}) = 0.001$.

THINK

Plan Define the events we're interested in and their probabilities.

Let $\mathbf{TB} = \{\text{having TB}\}$ and $\mathbf{TB^C} = \{\text{no TB}\}$
$+ = \{\text{testing positive}\}$ and
$- = \{\text{testing negative}\}$

Figure out what you want to know in terms of the events. Use the notation of conditional probability to write the event whose probability you want to find.

I know that $P(+\,|\,\mathbf{TB^C}) = 0.01$ and $P(-\,|\,\mathbf{TB}) = 0.001$. I also know that $P(\mathbf{TB}) = 0.00005$.

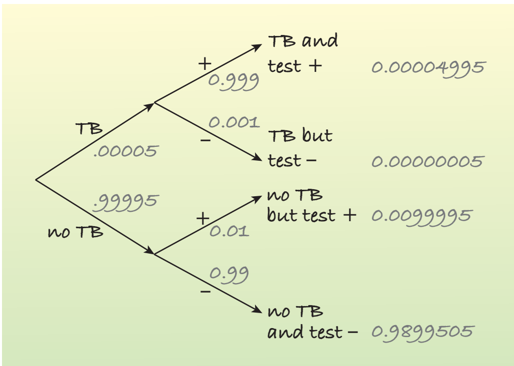
I'm interested in the probability that the author had TB given that he tested positive: $P(\mathbf{TB}\,|\,+)$.

SHOW

Plot Draw the tree diagram. When probabilities are very small like these are, be careful to keep all the significant digits.

To finish the tree we need $P(\mathbf{TB^C})$, $P(-\,|\,\mathbf{TB^C})$, and $P(-\,|\,\mathbf{TB})$. We can find each of these from the Complement Rule:

$$\begin{aligned} P(\mathbf{TB^C}) &= 1 - P(\mathbf{TB}) = 0.99995 \\ P(-\,|\,\mathbf{TB^C}) &= 1 - P(+\,|\,\mathbf{TB^C}) \\ &= 1 - 0.01 = 0.99 \text{ and} \\ P(+\,|\,\mathbf{TB}) &= 1 - P(-\,|\,\mathbf{TB}) \\ &= 1 - 0.01 = 0.999 \end{aligned}$$



[3] This isn't given, so we looked it up. Although TB is a matter of serious concern to public health officials, it is a fairly uncommon disease, with an incidence of about 5 cases per 100,000 in the United States (see *http://www.cdc.gov/tb/default.htm*).

	Mechanics Multiply along the branches to find the probabilities of the four possible outcomes. Check your work by seeing if they total 1.	(Check: $0.00004995 + 0.00000005 + 0.0099995 + 0.98995050 = 1$)
	Add up the probabilities corresponding to the condition of interest—in this case, testing positive. We can add because the tree shows disjoint events.	$P(+) = P(TB \cap +) + P(TB^C \cap +)$ $P = 0.00004995 + 0.0099995$ $= 0.01004945$
	Divide the probability of both events occuring (here, having TB and a positive test) by the probability of satisfying the condition (testing positive).	$P(TB \mid +) = \frac{P(TB \cap +)}{P(+)}$ $= \frac{0.00004995}{0.01004945}$ $= 0.00497$
TELL	**Conclusion** Interpret your result in context.	The chance of having TB after you test positive is less than 0.5%.

When we reverse the order of conditioning, we change the *Who* we are concerned with. With events of low probability, the result can be surprising. That's the reason patients who test positive for HIV, for example, are always told to seek medical counseling. They may have only a small chance of actually being infected. That's why global drug or disease testing can have unexpected consequences if people interpret *testing* positive as *being* positive.

Bayes's Rule

The Reverend Thomas Bayes is credited posthumously with the rule that is the foundation of Bayesian Statistics.

When we have $P(\mathbf{A} \mid \mathbf{B})$ but want the *reverse* probability $P(\mathbf{B} \mid \mathbf{A})$, we need to find $P(\mathbf{A} \cap \mathbf{B})$ and $P(\mathbf{A})$. A tree is often a convenient way of finding these probabilities. It can work even when we have more than two possible events, as we saw in the binge-drinking example. Instead of using the tree, we *could* write the calculation algebraically, showing exactly how we found the quantities that we needed: $P(\mathbf{A} \cap \mathbf{B})$ and $P(\mathbf{A})$. The result is a formula known as Bayes's Rule, after the Reverend Thomas Bayes (1702?–1761), who was credited with the rule after his death, when he could no longer defend himself. Bayes's Rule is quite important in Statistics and is the foundation of an approach to Statistical analysis known as Bayesian Statistics. Although the simple rule deals with two alternative outcomes, the rule can be extended to the situation in which there are more than two branches to the first split of the tree. The principle remains the same (although the math gets more difficult). Bayes's Rule is just a formula[4] for reversing the probability from the conditional probability that you're originally given, the same feat we accomplished with our tree diagram.

[4] Bayes's Rule for two events says that $P(\mathbf{B} \mid \mathbf{A}) = \frac{P(\mathbf{A} \mid \mathbf{B})P(\mathbf{B})}{P(\mathbf{A} \mid \mathbf{B})P(\mathbf{B}) + P(\mathbf{A} \mid \mathbf{B}^C)P(\mathbf{B}^C)}$. Masochists may wish to try it with the TB testing probabilities. (It's easier to just draw the tree, isn't it?)

FOR EXAMPLE Reversing the conditioning

A recent Maryland highway safety study found that in 77% of all accidents the driver was wearing a seatbelt. Accident reports indicated that 92% of those drivers escaped serious injury (defined as hospitalization or death), but only 63% of the non-belted drivers were so fortunate.

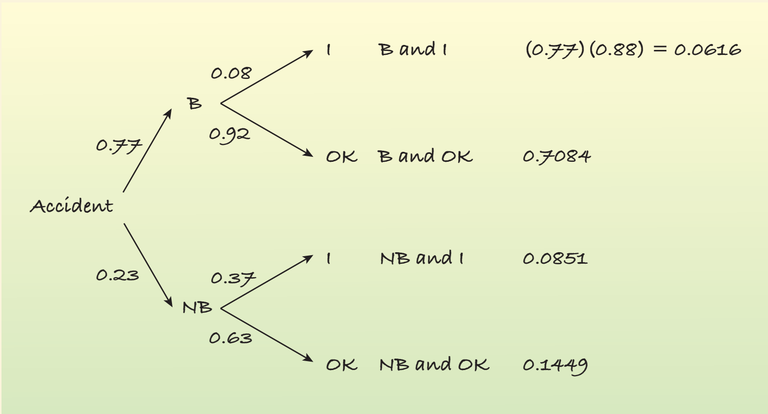
Question: What's the probability that a driver who was seriously injured wasn't wearing a seatbelt?

Let B = the driver was wearing a seatbelt, and NB = no belt.

Let I = serious injury or death, and OK = not seriously injured.

I know $P(B) = 0.77$, so $P(NB) = 1 - 0.77 = 0.23$.

Also, $P(OK \mid B) = 0.92$, so $P(I \mid B) = 0.08$
and $P(OK \mid B) = 0.63$, so $P(I \mid NB) = 0.37$



$$P(NB \mid I) = \frac{P(NB \text{ and } I)}{P(I)} = \frac{0.0851}{0.0616 + 0.0851} = 0.58$$

Even though only 23% of drivers weren't wearing seatbelts, they accounted for 58% of all the deaths and serious injuries.

Just some advice from your friends, the authors: *Please buckle up!* (We want you to finish this course.)

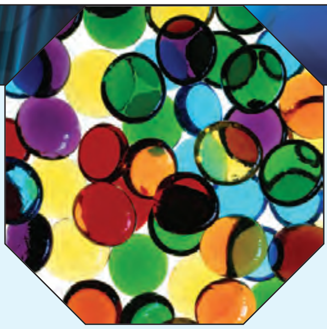
WHAT CAN GO WRONG?

- **Don't use a simple probability rule where a general rule is appropriate.** Don't assume independence without reason to believe it. Don't assume that outcomes are disjoint without checking that they are. Remember that the general rules always apply, even when outcomes are in fact independent or disjoint.
- **Don't find probabilities for samples drawn without replacement as if they had been drawn with replacement.** Remember to adjust the denominator of your probabilities. This warning applies only when we draw from small populations or draw a large fraction of a finite population. When the population is very large relative to the sample size, the adjustments make very little difference, and we ignore them.
- **Don't reverse conditioning naively.** As we have seen, the probability of **A** *given* **B** may not, and, in general does not, resemble the probability of **B** *given* **A.** The true probability may be counterintuitive.
- **Don't confuse "disjoint" with "independent."** Disjoint events *cannot* happen at the same time. When one happens, you know the other did not, so $P(\mathbf{B} \mid \mathbf{A}) = 0$. Independent events *must* be able to happen at the same time. When one happens, you know it has no effect on the other, so $P(\mathbf{B} \mid \mathbf{A}) = P(\mathbf{B})$.

CONNECTIONS

This chapter shows the unintuitive side of probability. If you've been thinking, "My mind doesn't work this way," you're probably right. Humans don't seem to find conditional and compound probabilities natural and often have trouble with them. Even statisticians make mistakes with conditional probability.

Our central connection is to the guiding principle that Statistics is about understanding the world. The events discussed in this chapter are close to the kinds of real-world situations in which understanding probabilities matters. The methods and concepts of this chapter are the tools you need to understand the part of the real world that deals with the outcomes of complex, uncertain events.



WHAT HAVE WE LEARNED?

The last chapter's basic rules of probability are important, but they work only in special cases—when events are disjoint or independent. Now we've learned the more versatile General Addition Rule and General Multiplication Rule. We've also learned about conditional probabilities, and seen that reversing the conditioning can give surprising results.

We've learned the value of Venn diagrams, tables, and tree diagrams to help organize our thinking about probabilities.

Most important, we've learned to think clearly about independence. We've seen how to use conditional probability to determine whether two events are independent and to work with events that are not independent. A sound understanding of independence will be important throughout the rest of this book.

Terms

General Addition Rule
343. For any two events, **A** and **B**, the probability of **A** *or* **B** is

$$P(\mathbf{A} \cup \mathbf{B}) = P(\mathbf{A}) + P(\mathbf{B}) - P(\mathbf{A} \cap \mathbf{B}).$$

Conditional probability
347. $P(\mathbf{B} \mid \mathbf{A}) = \frac{P(\mathbf{A} \cap \mathbf{B})}{P(\mathbf{A})}$

$P(\mathbf{B} \mid \mathbf{A})$ is read "the probability of **B** *given* **A**."

General Multiplication Rule
348. For any two events, **A** and **B**, the probability of **A** and **B** is

$$P(\mathbf{A} \cap \mathbf{B}) = P(\mathbf{A}) \times P(\mathbf{B} \mid \mathbf{A}).$$

Independence (used formally)
349. Events **A** and **B** are independent when $P(\mathbf{B} \mid \mathbf{A}) = P(\mathbf{B})$.

Tree diagram
354. A display of conditional events or probabilities that is helpful in thinking through conditioning.

Skills

THINK

- Understand the concept of conditional probability as redefining the *Who* of concern, according to the information about the event that is *given*.
- Understand the concept of independence.

SHOW

- Know how and when to apply the General Addition Rule.
- Know how to find probabilities for compound events as fractions of counts of occurrences in a two-way table.

- Know how and when to apply the General Multiplication Rule.
- Know how to make and use a tree diagram to understand conditional probabilities and reverse conditioning.

TELL

- Be able to make a clear statement about a conditional probability that makes clear how the condition affects the probability.
- Avoid making statements that assume independence of events when there is no clear evidence that they are in fact independent.

EXERCISES

1. Homes. Real estate ads suggest that 64% of homes for sale have garages, 21% have swimming pools, and 17% have both features. What is the probability that a home for sale has

a) a pool or a garage?
b) neither a pool nor a garage?
c) a pool but no garage?

2. Travel. Suppose the probability that a U.S. resident has traveled to Canada is 0.18, to Mexico is 0.09, and to both countries is 0.04. What's the probability that an American chosen at random has

a) traveled to Canada but not Mexico?
b) traveled to either Canada or Mexico?
c) not traveled to either country?

3. Amenities. A check of dorm rooms on a large college campus revealed that 38% had refrigerators, 52% had TVs, and 21% had both a TV and a refrigerator. What's the probability that a randomly selected dorm room has

a) a TV but no refrigerator?
b) a TV or a refrigerator, but not both?
c) neither a TV nor a refrigerator?

4. Workers. Employment data at a large company reveal that 72% of the workers are married, that 44% are college graduates, and that half of the college grads are married. What's the probability that a randomly chosen worker

a) is neither married nor a college graduate?
b) is married but not a college graduate?
c) is married or a college graduate?

5. Global survey. The marketing research organization GfK Custom Research North America conducts a yearly survey on consumer attitudes worldwide. They collect demographic information on the roughly 1500 respondents from each country that they survey. Here is a table showing the number of people with various levels of education in five countries:

Educational Level by Country

	Post-graduate	College	Some high school	Primary or less	No answer	Total
China	7	315	671	506	3	**1502**
France	69	388	766	309	7	**1539**
India	161	514	622	227	11	**1535**
U.K.	58	207	1240	32	20	**1557**
USA	84	486	896	87	4	**1557**
Total	**379**	**1910**	**4195**	**1161**	**45**	**7690**

If we select someone at random from this survey,

a) what is the probability that the person is from the United States?
b) what is the probability that the person completed his or her education before college?
c) what is the probability that the person is from France *or* did some post-graduate study?
d) what is the probability that the person is from France *and* finished only primary school or less?

6. Birth order. A survey of students in a large Introductory Statistics class asked about their birth order (1 = oldest or only child) and which college of the university they were enrolled in. Here are the results:

Birth Order

College		1 or only	2 or more	Total
	Arts & Sciences	34	23	**57**
	Agriculture	52	41	**93**
	Human Ecology	15	28	**43**
	Other	12	18	**30**
	Total	**113**	**110**	**223**