CHAPTER 4

Probability Judgment

4.1 Introduction

Though the theory of choice explored in Part 1 is helpful for a range of purposes, most real-life decisions are not choices under certainty. When you decide whether to start a company, buy stocks, propose to the love of your life, or have a medical procedure, you will not typically know at the time of making the decision what the outcome of each available act would be. In order to capture what people do, and what they should do, in such situations, we need another theory. Part 2 explores theories of judgment: how people form and change beliefs. In Part 3, we will return to the topic of decision-making.

In this chapter, we explore the theory of probability. There is wide – but far from complete – agreement that this is the correct normative theory of probabilistic judgment, that is, that it correctly captures how we should make probabilistic judgments. Consequently, the theory of probability is widely used in statistics, engineering, finance, public health, and elsewhere. Moreover, the theory can be used as a descriptive theory about how people make judgments, and it can be used as part of a theory about how they make decisions.

Like the theory of rational choice under certainty, probability theory is axiomatic. Thus, we begin by learning a set of axioms – which will be called "rules" – and which you will have to take for granted. Most of the time, this is not hard: once you understand them, the rules may strike you as intuitively plausible. We will also adopt a series of definitions. Having done that, though, everything else can be derived. Thus, we will spend a great deal of time below proving increasingly interesting and powerful principles on the basis of axioms and definitions.

4.2 Fundamentals of probability theory

Here are two classic examples of probability judgment.

Example 4.1 Mrs Jones's children You are visiting your new neighbor, Mrs Jones. Mrs Jones tells you that she has two children, who are playing in their room. Assume that each time somebody has a child, the probability of having a girl is the same as the probability of having a boy (and that whether the mother had a boy or a girl the first time around does not affect the probabilities involved the second time around). Now, Mrs Jones tells you that at least one of the children is a girl. What is the probability that the other child is a girl too?

Example 4.2 The Linda problem Linda is 31 years old, single, outspoken and very bright. She majored in philosophy. As a student, she was deeply concerned with issues of discrimination and social justice and also participated in anti-nuclear demonstrations.

(a) What is the probability that Linda is a bank teller?
(b) What is the probability that Linda is a bank teller and a feminist?

Answers to these questions will be given once we have developed the apparatus required to address them rigorously. For now, I will just note that one reason why this theory is interesting is that people's intuitive probability judgments – and therefore many of their decisions – tend to fail in predictable ways.

Before we start, we need to develop a conceptual apparatus that will permit us to speak more clearly about the subject matter. For example, we want to talk about the different things that can conceivably happen. When you flip a coin, for instance, you can get heads or you can get tails; when you roll a six-sided die, you can get any number between one and six.

Definition 4.3 Definition of "outcome space" *The* ***outcome space*** *is the set of all possible individual outcomes.*

We represent outcome spaces following standard conventions, using curly brackets and commas. To denote the outcome space associated with flipping a coin, we write: {Heads, Tails} or {H, T}. To denote the outcome space associated with rolling a six-sided die, we write: {1, 2, 3, 4, 5, 6}.

Oftentimes, we want to talk about what actually happened or about what may happen. If so, we are talking about actual outcomes, as in "the coin came up tails" and "I might roll snake eyes (two ones)."

Definition 4.4 Definition of "outcome" *An* ***outcome*** *is a subset of the outcome space.*

We write outcomes following the same conventions. Thus, some of the outcomes associated with one roll of a six-sided die include: {1} for one, {6} for six, {1, 2, 3} for a number less than or equal to three, and {2, 4, 6} for an even number. There is one exception: when the outcome only has one member, we may omit the curly brackets and write 6 instead of {6}. Notice that in all these cases, the outcomes are subsets of the outcome space.

Definition 4.5 Definition of "probability" *The* ***probability function*** *is a function Pr(·) that assigns a real number to each outcome. The* ***probability*** *of an outcome A is the number Pr(A) assigned to A by the probability function Pr(·).*

Hence, the probability of rolling an even number when rolling a six-sided die is denoted Pr({2, 4, 6}). The probability of rolling a six is denoted Pr({6}), or relying on the convention introduced above, Pr(6). The probability of getting heads when flipping a coin is denoted Pr({H}) or Pr(H). The probability of an outcome, of course, represents (in some sense) the chance of that outcome happening. Sometimes people talk about odds instead of probabilities. Odds and probabilities are obviously related, but they are not identical. Refer to the text box on page 83 for more about odds.

The next propositions describe the properties of this probability function. They will be referred to as the **rules** or **axioms** of probability.

Axiom 4.6 The range of probabilities *The probability of any outcome A is a number between 0 and 1 inclusive; that is,* $0 \leq \Pr(A) \leq 1$.

Thus, probabilities have to be numbers no smaller than zero and no greater than one. Equivalently, probabilities can be no lower than 0 percent and no greater than 100 percent. You might not know the probability that your internet startup company will survive its first year. But you do know this: the probability is no lower than 0 percent and no greater than 100 percent.

In general, it can be difficult to compute probabilities. People such as engineers and public health officials spend a lot of time trying to determine the probabilities of events such as nuclear disasters and global pandemics. There is one case in which computing probabilities is easy, however, and that is in the case when individual outcomes are equally probable, or **equiprobable**.

Axiom 4.7 The* EQUIPROBABILITY *rule *If there are n equally probable individual outcomes* $\{A_1, A_2, \dots, A_n\}$*, then the probability of any one individual outcome* A_i *is* $1/n$*; that is,* $Pr(A_i) = 1/n$.

Suppose we are asked to compute the probability of getting a four when rolling a fair die. Because all outcomes are equally likely (that is what it means for the die to be fair) and because there are six outcomes, the probability of getting a four is 1/6. So $\Pr(4) = 1/6$. Similarly, the probability of getting heads when flipping a fair coin is 1/2. So $\Pr(H) = 1/2$.

Exercise 4.8 Suppose that you are drawing one card each from two thoroughly shuffled but otherwise normal decks of cards. What is the probability that you draw the same card from the two decks?

You could answer this question by analyzing all $52^2 = 2704$ different outcomes associated with drawing two cards from two decks. The easiest way to think about it, though, is to ask what it would take for the second card to match the first.

Exercise 4.9 The Large Hadron Collider According to some critics, the Large Hadron Collider has a 50 percent chance of destroying the world. In an interview with the *Daily Show*'s John Oliver on April 30, 2009, science teacher Walter Wagner argued: "If you have something that can happen, and something that won't necessarily happen, it's going to either happen or it's going to not happen, and so the best guess is one in two." Why is this not a correct application of the EQUIPROBABILITY rule?

As it happens, we have already developed enough of an apparatus to address Example 4.1. First, we need to identify the outcome space associated with having two children. Writing G for "girl" and B for "boy," and BG for "the oldest child is a boy and the youngest child is a girl," the outcome space is {GG, GB, BG, BB}. Once you learn that at least one of the children is a girl, you know for a fact that it is not the case that both children are boys. That is, you know that BB does not obtain. This means that the outcome space has been reduced to {GG, GB, BG}. In only one of three cases (GG) is the other child a girl also. Because these three outcomes are equally likely, you can apply the EQUIPROBABILITY rule to find that the probability that the other child is a girl equals Pr(GG) = 1/3.

Exercise 4.10 Mrs Jones's children, cont. Instead of telling you that at least one of the children is a girl, Mrs Jones tells you that her oldest child is a girl. Now, what is the probability that the other child is also a girl?

Exercise 4.11 Mr Peters's children Your other neighbor, Mr Peters, has three children. Having just moved to the neighborhood, you do not know whether the children are boys or girls. Let us assume that every time Mr Peters had a child, he was equally likely to have a boy and a girl (and that there are no other possibilities).

(a) What is the relevant outcome space?
(b) Imagine that you learn that at least one of the children is a girl. What is the new outcome space?
(c) Given that you know that at least one of the children is a girl, what is the probability that Mr Peters has three girls?
(d) Imagine that you learn that at least two of the children are girls. What is the new outcome space?
(e) Given that you know that at least two of the children are girls, what is the probability that Mr Peters has three girls?

Exercise 4.12 Three-card swindle Your friend Bill is showing you his new deck of cards. The deck consists of only three cards. The first card is white on both sides. The second card is red on both sides. The third card is white on one side and red on the other. Now Bill shuffles the deck well, occasionally turning individual cards over in the process. Perhaps he puts them all in a hat and shakes the hat for a long time. Then he puts the stacked deck on the table in such a way that you can see the visible face of the top card only.

(a) What is the outcome space? Write "W/R" to denote the outcome where the visible side of the top card is white and the other side is red, and so on.
(b) After shuffling, the visible side of the top card is white. What is the new outcome space?
(c) Given that the visible side of the top card is white, what is the probability that the other side of the top card is red?

This last exercise is called the "three-card swindle," because it can be used to fool people into giving up their money. If you bet ten dollars that the other side is white, you will find that many people are willing to accept the bet. This is so because they (mistakenly) believe that the probability is 50 percent. You might lose. Yet, because you have got the probabilities on your side, on average you will make money. It is not clear that this game deserves the name "swindle" since it involves no deception. Still, because this might be illegal where you live, you did not hear it from me.

Exercise 4.13 Four-card swindle Your other friend Bull has another deck of cards. This deck has four cards: one card is white on both sides; one card is black on both sides; one card is red on both sides; and one card is white on one side and red on the other. Imagine that you shuffle the deck well, including turning individual cards upside down every so often.

(a) What is the outcome space? Write "W/R" to denote the outcome where the visible side of the top card is white and the other side is red, and so on.
(b) Suppose that after shuffling, the visible side of the top card is black. What is the new outcome space?
(c) Given that the visible side of the top card is black, what is the probability that the other side of the card is black as well?
(d) Suppose that after shuffling, the visible side of the top card is red. What is the new outcome space?
(e) Given that the visible side of the top card is red, what is the probability that the other side of the card is white?

We end this section with one more exercise.

Exercise 4.14 The Monty Hall Problem You are on a game show. The host gives you the choice of three doors, all of which are closed. Behind one door there is a car; behind the others are goats. Here is what will happen. First, you will point to a door. Next, the host, who knows what is behind each door and who is doing his best to make sure you do not get the car, will open one of the other two doors (which will have a goat). Finally, you can choose to open either one of the remaining two closed doors; that is, you can keep pointing to the same door, or you can switch. If you do not switch, what is the probability of finding the car?

4.3 Unconditional probability

The theory should also allow us to compute unknown probabilities on the basis of known probabilities. In this section we study four rules that do this.

Axiom 4.15 The OR rule *If two outcomes A and B are mutually exclusive (see below), then the probability of A OR B equals the probability of A plus the probability of B; that is,* $Pr(A \vee B) = Pr(A) + Pr(B)$.

Suppose that you want to know the probability of rolling a one or a two when you roll a fair six-sided die. The OR rule tells you that the answer is $\Pr(1 \vee 2) = \Pr(1) + \Pr(2) = 1/6 + 1/6 = 1/3$. Or suppose that you want to know the probability of flipping heads or tails when flipping a fair coin. The same rule tells you that $\Pr(H \vee T) = \Pr(H) + \Pr(T) = 1/2 + 1/2 = 1$.

Notice that the rule requires that the two outcomes be **mutually exclusive**. What does this mean? Two outcomes A and B are mutually exclusive just in case at most one of them can happen. In the previous two examples, this condition holds. When flipping a coin, H and T are mutually exclusive since at most one of them can occur every time you flip a coin. No coin can land heads and tails at the same time. Similarly, when you roll one die, one and two are mutually exclusive, since at most one of them can occur. Notice that the latter two outcomes are mutually exclusive even though neither one may occur.

Exercise 4.16 Mutual exclusivity Which pairs of outcomes are mutually exclusive? More than one answer may be correct.

(**a**) It is your birthday; you have a test.
(**b**) It rains; night falls.
(**c**) You get Bs in all of your classes; you get a 4.0 GPA.
(**d**) Your new computer is a Mac; your new computer is a PC.
(**e**) You are a remarkable student; you get a good job after graduation.

Exercise 4.17 What is the probability of drawing an ace when drawing one card from a regular (well-shuffled) deck of cards? If you intend to apply the OR rule, do not forget to check that the relevant outcomes are mutually exclusive.

The importance of checking whether two outcomes are mutually exclusive is best emphasized by giving an example. What is the probability of rolling a fair die and getting a number that is either strictly less than six or strictly greater than one? It is quite obvious that you could not fail to roll a number strictly less than six or strictly greater than one, so the probability must be 100 percent. If you tried to take the probability that you roll a number strictly less than six *plus* the probability that you roll a number strictly greater than one, you would end up with a number greater than one, which would be a violation of Axiom 4.6. So there is good reason for the OR rule to require that outcomes be mutually exclusive.

The answer to the question in the previous paragraph follows from the following straightforward rule:

Axiom 4.18 The EVERYTHING rule *The probability of the entire outcome space is equal to one.*

So, $Pr(\{1, 2, 3, 4, 5, 6\}) = 1$ by the EVERYTHING rule. We could also have computed this number by using the OR rule, because $Pr(\{1, 2, 3, 4, 5, 6\}) = Pr(1$ OR 2 OR 3 OR 4 OR 5 OR $6) = Pr(1) + Pr(2) + Pr(3) + Pr(4) + Pr(5) + Pr(6) = 1/6 * 6 = 1$. The OR rule (Axiom 4.15) applies because the six individual outcomes are mutually exclusive; the EQUIPROBABILITY rule (Axiom 4.7) applies because all outcomes are equally probable. What the EVERYTHING rule tells us that we did not already know is that the probability of the entire outcome space equals one whether or not the outcomes are equiprobable. The next rule is easy too.

Axiom 4.19 The NOT rule *The probability that some outcome A will not occur is equal to one minus the probability that it does. That is,* $Pr(\neg A) = 1 - Pr(A)$.

For example, suppose that you want to know the probability of rolling anything but a six when rolling a six-sided die. By the NOT rule, the probability that we roll anything but a six can be computed as $\Pr(\neg 6) = 1 - \Pr(6) = 1 - 1/6 = 5/6$. Given that the outcomes are mutually exclusive, we could have computed this using the OR rule too. (How?) In general, it is good to check that you get the same number when solving the same problem in different ways. If you do not, there is something wrong with your calculations.

Axiom 4.20 The AND rule *If two outcomes A and B are independent (see below), then the probability of A* AND *B equals the probability of A multiplied by the probability of B; that is, $Pr(A \& B) = Pr(A) * Pr(B)$.*

Suppose you flip a fair coin twice. What is the probability of getting two heads? Writing H_1 for heads on the first coin, and so on, by the AND rule, $\Pr(H_1 \& H_2) = \Pr(H_1) * \Pr(H_2) = 1/2 * 1/2 = 1/4$. You could also solve this problem by looking at the outcome space $\{H_1H_2, H_1T_2, T_1H_2, T_1T_2\}$ and using the EQUIPROBABILITY rule. Similarly, it is easy to compute the probability of getting two sixes when rolling a fair die twice: $\Pr(6_1 \& 6_2) = \Pr(6_1) * \Pr(6_2) = 1/6 * 1/6 = 1/36$.

Exercise 4.21 Are you more likely to get two sixes when rolling one fair die twice or when simultaneously rolling two fair dice?

Notice that the AND rule requires that the two outcomes be **independent**. What does this mean? Two outcomes *A* and *B* are independent just in case the fact that one occurs does not affect the probability that the other one does. This condition is satisfied when talking about a coin flipped twice. H_1 and H_2 are independent since the coin has no memory: whether or not the coin lands heads or tails the first time you flip it will not affect the probability of getting heads (or tails) the second time.

Exercise 4.22 Independence What pairs of outcomes are independent? More than one answer may be correct.
(a) You sleep late; you are late for class.
(b) You are a remarkable student; you get a good job after graduation.
(c) You write proper thank-you notes; you get invited back.
(d) The first time you flip a silver dollar you get heads; the second time you flip a silver dollar you get tails.
(e) General Electric stock goes up; General Motors stock goes up.

Exercise 4.23 Luck in love According to a well-known saying: "Lucky in cards, unlucky in love." Is this to say that luck in cards and luck in love are independent or not independent?

The importance of checking whether two outcomes are independent is best emphasized by giving an example. What is the probability of simultaneously getting a two and a three when you roll a fair die once? The answer is not $1/6 * 1/6 = 1/36$, of course, but zero. The outcomes are not independent, so you cannot use the AND rule. This example also tells you is that when two outcomes are mutually exclusive they are not independent. (We will return to the topic of independence in Section 5.2.)

Exercise 4.24 When rolling two fair dice, what is the probability that the number of dots add up to 11? If you intend to use the OR rule, make sure the relevant outcomes are mutually exclusive. If you intend to use the AND rule, make sure the relevant outcomes are independent.

Exercise 4.25 Suppose you draw two cards from a well-shuffled deck of cards *with replacement*, meaning that you put the first card back into the deck (and shuffle the deck once more) before drawing the second card.
(a) What is the probability that you draw the ace of spades twice?
(b) What is the probability that you draw two aces? (Here, you can use your answer to Exercise 4.17.)

We are now in a position to address Example 4.2. Obviously, the theory of probability by itself will not tell you the probability that Linda is a bank teller. But it can tell you something else. Let F mean that Linda is a feminist and B that Linda is a bank teller. Then the probability that she is both a feminist and a bank teller is $Pr(B\&F) = Pr(B) * Pr(F)$. (In order to apply the AND rule here, I am assuming that the outcomes are independent; the general result, however, holds even if they are not.) Because $Pr(F) \leq 1$ by Axiom 4.6, we know that $Pr(B) * Pr(F) \leq Pr(B)$. For, if you multiply a positive number x with a fraction (between zero and one) you will end up with something less than x. So whatever the relevant probabilities involved are, it must be the case that $Pr(B\&F) \leq Pr(B)$; that is, the probability that Linda is a bank teller and a feminist has to be smaller than or equal to the probability that she is a bank teller. Many people will tell you that Linda is more likely to be a bank teller and a feminist than she is to be a bank teller. This mistake is referred to as the **conjunction fallacy**, about which you will hear more in Section 5.3.

We end this section with one more exercise.

Exercise 4.26 For the following questions, assume that you are rolling two fair dice:
(a) What is the probability of getting *two* sixes?
(b) What is the probability of getting *no* sixes?
(c) What is the probability of getting *exactly one* six?
(d) What is the probability of getting *at least one* six?

To compute the answer to (c), note that there are two ways to roll exactly one six. When answering (d), note that there are at least two ways to compute the answer. You can recognize that *rolling at least one six* is the same as *rolling two sixes or rolling exactly one six* and add up the answers to (a) and (c). Or you can recognize that *rolling at least one six* is the same as *not rolling no sixes* and compute the answer using the NOT rule.

Exercise 4.27 In computing the answer to Exercise 4.26(d), you may have been tempted to add the probability of rolling a six on the one die (1/6) to the

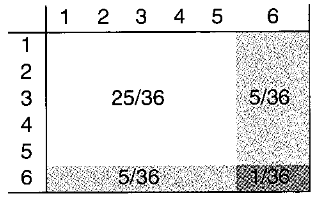

Figure 4.1 The two dice

probability of rolling a six on the other die (1/6) to get the answer 2/6 = 1/3. That, however, would be a mistake. Why?

Odds

Sometimes probabilities are expressed in terms of **odds** rather than probabilities. Imagine that you have an urn containing 2 black and 3 white balls, so that the probability of drawing a black ball is 2/5. One way to get this figure is to divide the number of favorable outcomes (outcomes in which the event of interest obtains) by the total number of outcomes. By contrast, you get the odds of drawing a black ball by dividing the number of favorable outcomes by the number of unfavorable outcomes, so that the odds of drawing a black ball are 2 to 3 or 2:3. Under the same assumptions, the odds of drawing a white ball are 3:2. If there is an equal number of black and white balls in the urn, the odds are 1 to 1 or 1:1. Such odds are also said to be **even**. When people talk about a 50–50 chance, they are obviously talking about even odds, since 50/50 = 1. How do odds relate to probabilities? If you have the probability p and want the odds o, you apply the following formula:

$$o = \frac{p}{1-p}$$

When p equals 2/5, it is easy to confirm that o equals 2/5 divided by 3/5 which is 2/3 or 2:3. If the probability is 1/2, the odds are 1/2 divided by 1/2 which is 1 or 1:1. If you have the odds o and want the probability p, you apply the inverse formula:

$$p = \frac{o}{o+1}$$

When o equals 2:3, you can quickly confirm that p equals 2/3 divided by 5/3 which is 2/5. If the odds are even, the probability is 1 divided by 1+1, which is 1/2. The use of odds instead of probabilities can come across as old-fashioned. But there are areas – for example, some games of chance and some areas of statistics – where odds are consistently used. It is good to know how to interpret them.

If the answers to Exercise 4.26 are not completely obvious already, refer to Figure 4.1. Here, the numbers to the left represent what might happen when you roll the first die and the numbers on top represent what might happen when you roll the second die. Thus, the table has $6 * 6 = 36$ cells representing all the possible outcomes of rolling two dice. The dark gray area represents the possibility that both dice are sixes; because there is only one way to roll two sixes, this area contains but one cell and the answer to (a) is 1/36. The white area represents the possibility that both dice are non-sixes; because there are $5 * 5$ ways to roll two non-sixes, this area contains 25 cells and the answer to (b) is 25/36. The light gray areas represent the possibility that one die is a six and the other one is a non-six; because there are $5 + 5$ ways to attain this outcome, these areas contain ten cells and the answer to (c) is 10/36. You can compute the answer to (d) by counting the $5 + 5 + 1 = 11$ cells in the two light gray and the one dark gray areas and get an answer of 11/36. But a smarter way is to realize that the gray areas cover everything that is not white, which allows you to get the answer by computing $1 - 25/36 = 11/36$. Why this is smarter will be clear in Section 5.3. The figure also illustrates why you cannot compute the probability of getting at least one six by adding the probability of rolling a six on the first die to the probability of rolling a six on the second one. If you were to do that, you would add the number of cells in the bottom row to the number of cells in the right-most column – thereby double-counting the cell on the bottom-right.

4.4 **Conditional probability**

In Exercise 4.25, you computed the probability of drawing two aces when drawing two cards with replacement. Suppose, instead, that you draw two aces *without replacement*, meaning that you put the first card aside after looking at it. What is the probability of drawing two aces without replacement? You know you cannot use Axiom 4.20, since the two outcomes we are interested in (drawing an ace the first time and drawing an ace the second time) are not independent. You can, however, approach the problem in the following way. First, you can ask what the probability is that the first card is an ace. Because there are 52 cards in the deck, and 4 of those are aces, you know that this probability is 4/52. Second, you can ask what the probability is that the second card is an ace, *given that the first card was an ace*. Because there are 51 cards left in the deck, and only 3 of them are aces, this probability is 3/51. Now you can multiply these numbers and get:

$$\frac{4}{52} * \frac{3}{51} = \frac{12}{2652} = \frac{1}{221}.$$

This procedure can be used to calculate the probability of winning certain types of lotteries. According to the Consumer Federation of America, about one in five Americans believe that "the most practical way for them to accumulate several hundred thousand dollars is to win the lottery." The poor, least educated, and oldest are particularly likely to think of the lottery as a smart way to get rich. So it might be useful to ask just how likely or unlikely it is to win common lotteries.

Exercise 4.28 Lotto 6/49 Many states and countries operate lotteries in which the customer picks n of m numbers, in any order, where n is considerably smaller than m. In one version of this lottery, which I will call Lotto 6/49, players circle 6 numbers out of 49 using a ticket like that in Figure 4.2. The order in which numbers are circled does not matter. You win the grand prize if all 6 are correct. What is the probability that you win the Lotto 6/49 any one time you play? Notice that this is similar to picking six consecutive aces out of a deck with 49 cards, if 6 of those cards are aces.

The fact that the probability of winning the lottery is low does not imply that it is necessarily irrational to buy these tickets. (We will return to this topic in Part 3.) Nevertheless, it may be fun to ask some questions about these lotteries.

Problem 4.29 Lotto 6/49, cont. *What does the probability of winning the Lotto 6/49 tell you about the wisdom of buying Lotto tickets? What does it tell you about people who buy these tickets?*

Exercise 4.30 Lotto 6/49, cont. Use the idea of anchoring and adjustment from Section 3.6 to explain why people believe that they have a good chance of winning these lotteries.

Considerations like these clarify why state lottery schemes are sometimes described as a tax on innumeracy.

The probability that something happens given that some other thing happens is called a **conditional probability**. We write the probability that A given C, or the probability of A conditional on C, as follows: $\Pr(A \mid C)$. Conditional probabilities are useful for a variety of purposes. It may be easier to compute conditional probabilities than unconditional probabilities. Knowing the conditional probabilities is oftentimes quite enough to solve the problem at hand.

Notice right away that $\Pr(A \mid C)$ is not the same thing as $\Pr(C \mid A)$. Though these two probabilities may be identical, they need not be. Suppose, for example, that S means that Joe is a smoker, while H means that Joe is human. If so,

LOTTO 6/49						
1	2	3	4	5	6	7
8	9	10	11	12	13	14
15	16	17	18	19	20	21
22	23	24	25	26	27	28
29	30	31	32	33	34	35
36	37	38	39	40	41	42
43	44	45	46	47	48	49

Figure 4.2 Lotto 6/49 ticket

Pr(S | H) is the probability that Joe is a smoker given that Joe is human, which is a number somewhere between zero and one. Meanwhile, Pr(H | S) is the probability that Joe is human given that Joe is a smoker, which is one (or at least close to one). Joe may not be a human being for quite as long if he is a smoker, but that is another matter.

Exercise 4.31 Conditional Probabilities Suppose that H means "The patient has a headache" and T means "The patient has a brain tumor."

(a) How do you interpret the two conditional probabilities Pr(H | T) and Pr(T | H)?
(b) Are the two numbers more or less the same?

It should be clear that the two conditional probabilities in general are different, and that it is important for both doctors and patients to keep them apart. (We will return to this topic in sections 5.4 and 5.6.)

Suppose that you draw one card from a well-shuffled deck, and that you are interested in the probability of drawing the ace of spades given that you draw an ace. Given that you just drew an ace, there are four possibilities: the ace of spades, ace of clubs, ace of hearts, or ace of diamonds. Because only one of the four is the ace of spades, and because all four outcomes are equally likely, this probability is 1/4. You can get the same answer by dividing the probability that you draw the ace of spades by the probability that you draw an ace: 1/52 divided by 4/52, which is 1/4. This is no coincidence, as the formal definition of conditional probability will show.

Definition 4.32 Conditional probability *If A and B are two outcomes,* $Pr(A|B) = Pr(A \& B)/Pr(B)$.

As another example of conditional probability, recall the problem with the two aces. Let A denote "the second card is an ace" and let B denote "the first card is an ace." We know that the probability of drawing two aces without replacement is 1/221. This is $\Pr(A \& B)$. We also know that the probability that the first card is an ace is 1/13. This is $\Pr(B)$. So by definition:

$$\Pr(A|B) = \frac{\Pr(A\&B)}{\Pr(B)} = \frac{1/221}{1/13} = \frac{3}{51}.$$

But we knew this: 3/51 is the probability that the second card is an ace given that the first card was: $\Pr(A \& B)$. So the formula works.

Exercise 4.33 Ace of spades Use Definition 4.32 to compute the probability that you draw an ace of spades conditional on having drawn an ace when you draw one card from a well-shuffled deck. You can imagine a game show host who draws a card at random and announces that the card is an ace, and a contestant who has to guess what kind of ace it is. Given what you know about that card, what is the probability that it is the ace of spades?

Because you cannot divide numbers by zero, things get tricky when some probabilities are zero; here, I will ignore these complications.

One implication of the definition is particularly useful:

Proposition 4.34 The general* AND *rule *Pr(A & B) = Pr(A | B) ∗ Pr(B).*

Proof. Starting off with Definition 4.32, multiply each side of the equation by $\Pr(B)$. □

According to this proposition, the probability of drawing two aces equals the probability of drawing an ace the first time multiplied by the probability of drawing an ace the second time given that you drew an ace the first time. But again, we knew this. In fact, we implicitly relied on this rule when computing the answers to the first exercises in this section. Notice that this rule allows us to compute the probability of *A* AND *B* without requiring that the outcomes be independent. This is why it is called the general AND rule.

Exercise 4.35 The general AND rule Use the general AND rule to compute the probability that you will draw the ace of spades twice when drawing two cards from a deck *without* replacement.

The general AND rule permits us to establish the following result.

Proposition 4.36 *Pr(A | B) ∗ Pr(B) = Pr(B | A) ∗ Pr(A).*

Proof. By Proposition 4.34, $\Pr(A\&B) = \Pr(A \mid B) * \Pr(B)$ but also $\Pr(B\&A) = \Pr(B \mid A) * \Pr(A)$. Because by logic $\Pr(A\&B) = \Pr(B\&A)$, it must be the case that $\Pr(A \mid B) * \Pr(B) = \Pr(B \mid A) * \Pr(A)$. □

Suppose that you draw one card from a well-shuffled deck, and that **A** means that you draw an ace and that ♦ means that you draw a diamond. If so, it follows that Pr(**A** | ♦) ∗ Pr(♦) = Pr(♦ | **A**) ∗ Pr(**A**). You can check that this is true by plugging in the numbers: $1/13 * 13/52 = 1/4 * 4/52 = 1/4$.

This notion of conditional probability allows us to sharpen our definition of independence. We said that two outcomes *A* and *B* are independent if the probability of *A* does not depend on whether *B* occurred. Another way of saying this is to say that $\Pr(A \mid B) = \Pr(A)$. In fact, there are several ways of saying the same thing.

Proposition 4.37 Independence conditions *The following three claims are equivalent:*

(i) $\Pr(A \mid B) = \Pr(A)$
(ii) $\Pr(B \mid A) = \Pr(B)$
(iii) $\Pr(A \& B) = \Pr(A) * \Pr(B)$

Proof. See Exercise 4.38. □

Exercise 4.38 Independence conditions Prove that the three parts of Proposition 4.37 are equivalent. The most convenient way of doing so is to prove (**a**) that (i) implies (ii), (**b**) that (ii) implies (iii), and (**c**) that (iii) implies (i).

Notice that part (iii) is familiar: it is the principle that we know as the AND rule (Axiom 4.20). Thus, the original AND rule follows logically from the general AND rule and the assumption that the two outcomes in question are independent. This is pretty neat.

4.5 Total probability and Bayes's rule

Conditional probabilities can also be used to compute unconditional probabilities. Suppose that you are running a frisbee factory and that you want to know the probability that one of your frisbees is defective. You have two machines producing frisbees: a new one (B) producing 800 frisbees per day and an old one (¬B) producing 200 frisbees per day. Thus, the probability that a randomly selected frisbee from your factory was produced by machine B is $\Pr(B) = 800/(800 + 200) = 0.8$; the probability that it was produced by machine ¬B is $\Pr(\neg B) = 1 - \Pr(B) = 0.2$. Among the frisbees produced by the new machine, one percent are defective (D); among those produced by the old one, two percent are. The probability that a randomly selected frisbee produced by machine B is defective is $\Pr(D|B) = 0.01$; the probability that a randomly selected frisbee produced by machine ¬B is defective is $\Pr(D|\neg B) = 0.02$. It may be useful to draw a tree illustrating the four possibilities (see Figure 4.3).

What is the probability that a randomly selected frisbee from your factory is defective? There are two ways in which a defective frisbee can be produced: by machine B and by machine ¬B. So the probability that a frisbee is defective Pr(D) equals the following probability: that the frisbee is produced by machine B and turns out to be defective or that the frisbee is produced by machine ¬B and turns out to be defective; that is, $\Pr([D\&B] \vee [D\&\neg B])$. These outcomes are obviously mutually exclusive, so the probability equals $\Pr(D\&B) + \Pr(D\&\neg B)$. Applying the general AND rule twice, this equals $\Pr(D|B) * \Pr(B) + \Pr(D|\neg B) * \Pr(\neg B)$. But we have all these numbers, so:

$$\Pr(D) = \Pr(D|B) * \Pr(B) + \Pr(D|\neg B) * \Pr(\neg B) = 0.01 * 0.8 + 0.02 * 0.2 \doteq 0.012.$$

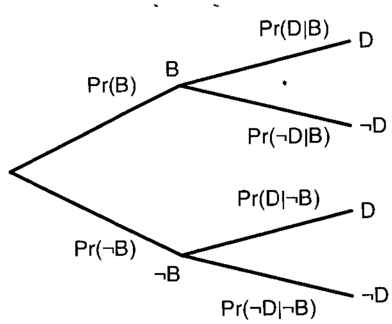

Figure 4.3 The frisbee factory

The probability that a randomly selected frisbee produced by your factory is defective is 1.2 percent. These calculations illustrate **the rule of total probability.**

Proposition 4.39 The rule of total probability

$$Pr(D) = Pr(D|B) * Pr(B) + Pr(D|\neg B) * Pr(\neg B).$$

Proof. By logic, D is the same as $[D \& B] \vee [D \& \neg B]$. So $Pr(D) = Pr([D \& B] \vee [D \& \neg B])$. Because the two outcomes are mutually exclusive, this equals $Pr(D \& B) + Pr(D \& \neg B)$ by the OR rule (Axiom 4.15). Applying the general AND rule (Proposition 4.34) twice, we get $Pr(D) = Pr(D|B) * Pr(B) + Pr(D|\neg B) * Pr(\neg B)$. □

Exercise 4.40 Cancer Use the rule of total probability to solve the following problem. You are a physician meeting with a patient who has just been diagnosed with cancer. You know there are two mutually exclusive types of cancer that the patient could have: type A and type B. The probability that he or she has A is 1/3 and the probability that he or she has B is 2/3. Type A is deadly: four patients out of five diagnosed with type A cancer die (D) within one year. Type B is less dangerous: only one patient out of five diagnosed with type B cancer dies (D) within one year.
(a) Draw a tree representing the four possible outcomes.
(b) Compute the probability that your patient dies within a year.

Exercise 4.41 Scuba diving certification You are scheduled to sit the test required to be a certified scuba diver and very much hope you will pass (P). The test can be easy (E) or not. The probability that it is easy is 60 percent. If it is easy, you estimate that the probability of passing is 90 percent; if it is hard, you estimate that the probability is 50 percent. What is the probability that you pass (P)?

There is another type of question that you may ask as well. Suppose you pick up one of the frisbees produced in your factory and find that it is defective. What is the probability that the defective frisbee was produced by the new machine? Here you are asking for the probability that a frisbee was B conditional on D, that is, $Pr(B|D)$.

We know that there are two ways in which a defective frisbee can be produced. Either it comes from the new machine, which is to say that $D \& B$, or it comes from the old machine, which is to say that $D \& \neg B$. We also know the probabilities that these states will obtain for any given frisbee (not necessarily defective): $Pr(D \& B) = Pr(D|B) * Pr(B) = 0.01 * 0.8 = 0.008$ and $Pr(D \& \neg B) = Pr(D|\neg B) * Pr(\neg B) = 0.02 * 0.2 = 0.004$. We want the probability that a frisbee comes from the new machine given that it is defective, that is, $Pr(B|D)$. By looking at the figures, you can tell that the first probability is twice as large as the second one. What this means is that in two cases out of three, a defective frisbee comes from the new machine.

Formally, Pr (D | B) = 0.008/0.012 = 2/3. This may be surprising, in light of the fact that the new machine has a lower rate of defective frisbees than the old one. But it is explained by the fact that the new machine also produces far more frisbees than the old one.

The calculations you have just performed are an illustration of **Bayes's rule**, or **Bayes's theorem**, which looks more complicated than it is.

Proposition 4.42 Bayes's rule

$$\Pr(B|D) = \frac{\Pr(D|B) * \Pr(B)}{\Pr(D)}$$
$$= \frac{\Pr(D|B) * \Pr(B)}{\Pr(D|B) * \Pr(B) + \Pr(D|\neg B) * \Pr(\neg B)}.$$

Proof. The rule has two forms. The first form can be obtained from Proposition 4.36 by dividing both sides of the equation by $\Pr(D)$. The second form can be obtained from the first by applying the rule of total probability (Proposition 4.39) to the denominator. □

Exercise 4.43 Cancer, cont. Suppose that your patient from Exercise 4.40 dies in less than one year, before you learn whether he or she has type A or type B cancer. Given that the patient died in less than a year, what is the probability he or she had type A cancer?

Exercise 4.44 Scuba diving certification, cont. You passed the scuba diving test! Your friend says: "Not to rain on your parade, but you obviously got the easy test." Given that you passed, what is the probability that you got the easy test?

Bayes's rule is an extraordinarily powerful principle. To show how useful it can be, consider the following problem. If it is not immediately obvious how to attack this problem, it is almost always useful to draw a tree identifying the probabilities.

Exercise 4.45 The dating game You are considering asking L out for a date, but you are a little worried that L may already have started dating somebody else. The probability that L is dating somebody else, you would say, is 1/4. If L is dating somebody else, he/she is unlikely to accept your offer to go on a date: in fact, you think the probability is only 1/6. If L is not dating somebody else, though, you think the probability is much better: about 2/3.

(a) What is the probability that L is dating somebody else but will accept your offer to go on a date anyway?
(b) What is the probability that L is *not* dating somebody else and will accept your offer to go on a date?
(c) What is the probability that L will accept your offer to go on a date?
(d) Suppose L accepts your offer to go on a date. What is the probability that L is dating somebody else, given that L agreed to go on a date?

There are more exercises on Bayes's rule in sections 4.6 and 5.4. See also Exercise 5.34 on page 113.

4.6 **Bayesian updating**

Bayes's rule is often interpreted as describing how we should update our beliefs in light of new evidence. We update beliefs in light of new evidence all the time. In everyday life, we update our belief that a particular presidential candidate will win the election in light of evidence about how well he or she is doing. The evidence here may include poll results, our judgments about his or her performance in presidential debates, and so on. In science, we update our assessment about the plausibility of a hypothesis or theory in light of evidence, which may come from experiments, field studies, or other sources. Consider, for example, how a person's innocent belief that the Earth is flat might be updated in light of the fact that there are horizons, the fact that the Earth casts a circular shadow onto the Moon during a lunar eclipse, and the fact that one can travel around the world. Philosophers of science talk about the **confirmation** of scientific theories, so the theory of how this is done is called **confirmation theory**. Bayes's rule plays a critical role in confirmation theory.

To see how this works, think of the problem of belief updating as follows: what is at stake is whether a given hypothesis is true or false. If the hypothesis is true, there is some probability that the evidence obtains. If the hypothesis is false, there is some other probability that the evidence obtains. The question is how you should change your belief – that is, the probability that you assign to the possibility that the hypothesis is true – in light of the fact that the evidence obtains. Figure 4.4 helps to bring out the structure of the problem.

Let H stand for the **hypothesis** and E for the **evidence**. The probability of H, $\Pr(H)$, is called the **prior probability**: it is the probability that H is true before you learn whether E is true. The probability of H given E, $\Pr(H|E)$, is called the **posterior probability**: it is the probability that H obtains given that the evidence E is true. The question is what the posterior probability should be.

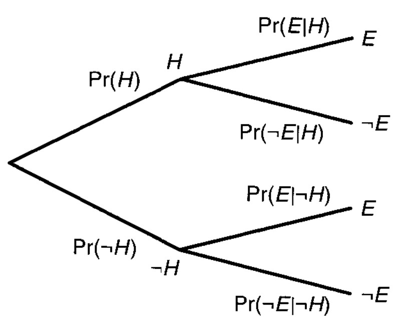

Figure 4.4 Bayesian updating

This question is answered by a simple application of Bayes's rule. Substituting H for B and E for D in Proposition 4.42, we can write Bayes's rule as follows:

$$\Pr(H|E) = \frac{\Pr(E|H) * \Pr(H)}{\Pr(E|H) * \Pr(H) + \Pr(E|\neg H) * \Pr(\neg H)}.$$

The result tells you how to update your belief in the hypothesis H in light of the evidence E. Specifically, Bayes's rule tells you that the probability you assign to H being true should go from $\Pr(H)$ to $\Pr(H|E)$. If you change your beliefs in accordance with Bayes's rule, we say that you engage in **Bayesian updating**.

Suppose that John and Wes are arguing about whether a coin brought to class by a student has two heads or whether it is fair. Imagine that there are no other possibilities. For whatever reason, the student will not let them inspect the coin, but she will allow them to observe the outcome of coin flips. Let H be the hypothesis that the coin has two heads, so that $\neg H$ means that the coin is fair. Let us consider John first. He thinks the coin is unlikely to have two heads: his prior probability, $\Pr(H)$, is only 0.01. Now suppose the student flips the coin, and that it comes up heads. Let E mean "The coin comes up heads." The problem is this: What probability should John assign to H given that E is true?

Given Bayes's rule, computing John's posterior probability $\Pr(H|E)$ is straightforward. We are given $\Pr(H) = 0.01$, and therefore know that $\Pr(\neg H) = 1 - \Pr(H) = 0.99$. From the description of the problem, we also know the conditional probabilities: $\Pr(E|H) = 1$ and $\Pr(E|\neg H) = 0.5$. All that remains is to plug the numbers into the theorem, as follows:

$$\Pr(H|E) = \frac{\Pr(E|H) * \Pr(H)}{\Pr(E|H) * \Pr(H) + \Pr(E|\neg H) * \Pr(\neg H)}$$

$$= \frac{1 * 0.01}{1 * 0.01 + 0.5 * 0.99} \approx 0.02.$$

The fact that John's posterior probability $\Pr(H|E)$ differs from his prior probability $\Pr(H)$ means that he updated his belief in light of the evidence. The observation of heads increased his probability that the coin has two heads, as it should. Notice how the posterior probability reflects both the prior probability and the evidence E.

Now, if John gets access to ever more evidence about the coin, there is no reason why he should not update his belief again. Suppose that the student flips the coin a second time and gets heads again. We can figure out what John's probability should be after observing this second flip by simply treating his old posterior probability as the new prior probability and applying Bayes's rule once more:

$$\Pr(H|E) = \frac{1 * 0.02}{1 * 0.02 + 0.5 * 0.98} \approx 0.04.$$

Notice that his posterior probability increases even more after he learns that the coin came up heads the second time.

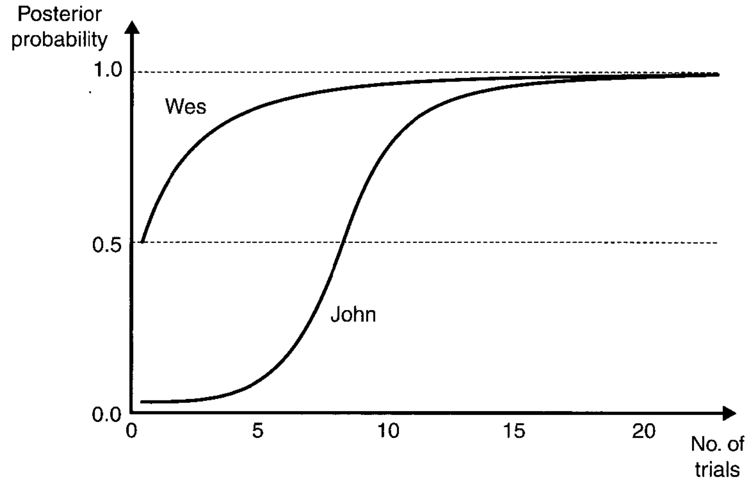

Figure 4.5 John's and Wes's probabilities after repeated trials

Exercise 4.46 Bayesian updating Suppose Wes, before the student starts flipping the coin, assigns a probability of 50 percent to the hypothesis that it has two heads.

(a) What is his posterior probability after the first trial?
(b) After the second?

Figure 4.5 illustrates how John's and Wes's posterior probabilities develop as the evidence comes in. Notice that over time both increase the probability assigned to the hypothesis. Notice, also, that their respective probabilities get closer and closer. As a result, over time (after some 15–20 trials) they are in virtual agreement that the probability of the coin having two heads is almost 100 percent. We will return to questions of rational updating in the next chapter. Until then, one last exercise.

Exercise 4.47 Bayesian updating, cont. Suppose that, on the third trial, instead of flipping heads, the student flips tails. What would John's and Wes's posterior probability be? To solve this problem, let E mean "The coin comes up tails."

4.7 Discussion

In this chapter we have explored the theory of probability. The theory has something of the feel of magic to it. When you are facing these difficult problems where untutored intuitions are conflicting or wide off the mark, all you have to do is to apply the incantations (rules) in the right way and in the right order and the answer pops right up! Anyway, probability theory is critical to a wide range of applications, among other things as the foundations of statistical inference. It is relevant here because it can be interpreted as a theory of judgment, that is, as a theory of how to revise beliefs in light of evidence.

The section on Bayesian updating (Section 4.6) shows how the argument goes. As long as you are committed to the axioms and rules of the probability calculus, we established that you will (as a matter of mathematical necessity) update your probabilities consistently with Bayes's rule. Notice how neat this is. While the theory says nothing about what your prior probabilities are or should be, it does tell you exactly what your posterior probability will or should be after observing the evidence.

How plausible is this theory? Again, we must separate the descriptive question from the normative question. Do people in fact update their beliefs in accordance with Bayes's rule? Should they? The axioms might seem weak and uncontroversial, both from a descriptive and from a normative standpoint. Yet the resulting theory is anything but weak. As in the case of the theory of choice under certainty, we have built a remarkably powerful theory on the basis of a fairly modest number of seemingly weak axioms. It is important to keep in mind that the theory is not as demanding as some people allege. It is not intended to describe the actual cognitive processes you go through when updating your beliefs: the theory does not say that you must apply Bayes's theorem in your head. But it does specify exactly how your posterior probability relates to your prior, given various conditional and unconditional probabilities.

In the next chapter, we will see how the theory fares when confronted with evidence.

ADDITIONAL EXERCISES

Exercise 4.48 SAT test When you take the SAT test, you may think that the correct answers to the various questions would be completely random. In fact, they are not. The authors of the test want the answers to *seem* random, and therefore they make sure that not all correct answers are, say, (d). Consider the following three outcomes. A: The correct answer to question 12 is (d). B: The correct answer to question 13 is (d). C: The correct answer to question 14 is (d). Are outcomes A, B, and C mutually exclusive, independent, both, or neither?

Exercise 4.49 Mr Langford Multiple lawsuits allege that area gambling establishments, on multiple occasions, doctored equipment so as to give Birmingham mayor Larry Langford tens of thousands of dollars in winnings. Langford, already in prison for scores of corruption-related charges, has not denied winning the money; he does deny that the machines were doctored.

We do not know exactly what the probability of winning the jackpot on a machine that has *not* been doctored might be, but we can make some intelligent guesses. Suppose that Langford on three occasions bet \$1 and won \$25,000. For each jackpot, in order to break even, a gambling establishment needs 24,999 people who bet \$1 and do not win. So we might infer that the probability of winning when betting a dollar is somewhere in the neighborhood of 1/25,000. If the establishment wants to make a profit, which it does, the probability would have to be even lower, but let us ignore this fact.

how the argument goes.
s of the probability cal-
nathematical necessity)
rule. Notice how neat
r prior probabilities are
erior probability will or

separate the descriptive
fact update their beliefs
xioms might seem weak
a normative standpoint.
the case of the theory of
powerful theory on the
axioms. It is important to
some people allege. It is
es you go through when
u must apply Bayes's the-
our posterior probability
conditional probabilities.
res when confronted with

T test, you may think
would be completely
est want the answers to
not all correct answers
comes. *A*: The correct
er to question 13 is (d).
outcomes *A*, *B*, and *C*
r?

allege that area gam-
ctored equipment so as
ns of thousands of dol-
or scores of corruption-
oney; he does deny that

y of winning the jackpot
ht be, but we can make
d on three occasions bet
r to break even, a gam-
et $1 and do not win. So
when betting a dollar is
the establishment wants
would have to be even

ADDITIONAL EXERCISES cont.

What is the probability that Langford would win the jackpot three times in a row when playing three times on undoctored machines?

Note that the probability that Langford would win three times in a row given that the machines were not doctored is different from the conditional probability that the machines were not doctored given that Langford won three times in a row.

Exercise 4.50 Economists go to Vegas According to professional lore, economists are not welcome to organize large meetings in Las Vegas. The reason is a sort of sin of omission. What is it economists, unlike most normal people, allegedly do not do when in Vegas?

Exercise 4.51 Gender discrimination Imagine that an editorial board of 20 members is all male.

(a) What is the probability that this would happen by chance alone assuming that the board members are drawn from a pool of 1/2 men and 1/2 women?
(b) Perhaps the pool of qualified individuals is not entire balanced in terms of gender. What is the answer if the pool consists of 2/3 men and 1/3 women?
(c) And what if it is 4/5 men and 1/5 women?

Exercise 4.52 Softball A softball player's batting average is defined as the ratio of hits to at bats. Suppose that a player has a 0.250 batting average and is very consistent, so that the probability of a hit is the same every time she is at bat. During today's game, this player will be at bat exactly three times.

(a) What is the probability that she ends up with three hits?
(b) What is the probability that she ends up with no hits?
(c) What is the probability that she ends up with exactly one hit?
(d) What is the probability that she ends up with at least one hit?

Exercise 4.53 Gov. Schwarzenegger After vetoing a bill from the California State Assembly in 2009, California Governor Arnold Schwarzenegger published a letter (see Figure 4.6). People immediately noticed that the first letter on each line together spelled out a vulgarity. When confronted with this fact, a spokesperson said: "It was just a weird coincidence."

(a) Assuming that a letter has eight lines, and that each of the 26 letters in the alphabet is equally likely to appear at the beginning of each line, what is the probability that this exact message would appear by chance?
(b) It is true that the Governor writes many letters each year, which means that the probability of any one letter spelling out this vulgarity is higher than your answer to (a) would suggest. Suppose that the Governor writes 100 eight-line letters each year. What is the probability that at least one of them will spell out the vulgarity?

ADDITIONAL EXERCISES cont.

To the Members of the California State Assembly:

I am returning Assembly Bill 1176 without
[my s . . .]

For some time now I have lamented the fact tha . . .
unnecessary bills come to me for consideration . . .
care are major issues my Administration has br . . .
kicks the can down the alley.

Yet another legislative year has come and gone . . .
overwhelmingly deserve. In light of this, and . . .
unnecessary to sign this measure at this time.

Sincerely,

Arnold Schwarzenegger

Figure 4.6 Governor Schwarzenegger's letter

Exercise 4.54 Max's bad day Max is about to take a multiple-choice test. The test has ten questions, and each has two possible answers: "true" and "false." Max does not have the faintest idea of what the right answer to any of the questions might be. He decides to pick answers randomly.

(a) What is the probability that Max will get all ten questions right?
(b) What is the probability that Max will get the first question wrong and the other questions right?
(c) What is the probability that Max will get the second question wrong and the other questions right?
(d) What is the probability that Max will get exactly nine questions right?
(e) Max really needs to get an A on this test. In order to get an A, he needs to get nine or more questions right. What is the probability that Max will get an A?

Exercise 4.55 Pregnancy tests You are marketing a new line of pregnancy tests. The test is simple. You flip a fair coin. If the coin comes up heads, you report that the customer is pregnant. If the coin comes up tails, you report that the customer is not pregnant.

(a) Your first customer is a man. What is the probability that the test accurately predicts his pregnancy status?
(b) Your second customer is a woman. What is the probability that the test accurately predicts her pregnancy status?
(c) After you have administered the test ten times, what is the probability that you have not correctly predicted the pregnancy status of any of your customers?

ADDITIONAL EXERCISES cont.

(d) After you have administered the test ten times, what is the probability that you correctly predicted the pregnancy status of *at least one* of your customers?

Notice how high the probability of getting at least one customer right is. This suggests the following scheme for getting rich. Issue ten, or a hundred, or whatever, newsletters offering advice about how to pick stocks. No matter how unlikely each newsletter is to give good advice, if you issue enough of them at least one is very likely to give good advice. Then sell your services based on your wonderful track record, pointing to the successful newsletter as evidence. You would not be the first. We will return to this kind of problem in Section 5.3.

Problem 4.56 Pregnancy tests, cont. *The pregnancy test of Exercise 4.55 is needlessly complicated. Here is another test that is even simpler: just report that the customer is not pregnant. Roughly, what is the probability that you would get the pregnancy status of a randomly selected college student right when using the simplified test?*

FURTHER READING

There are numerous introductions to probability theory. Earman and Salmon (1992) deals with probability theory in the context of the theory of confirmation and is the source of the stories about frisbees and coins (pp. 70–74); it also contains a discussion about the meaning of probability (pp. 74–89). The Consumer Federation of America (2006) discusses people's views about the most practicable way to get rich. The fate of Birmingham mayor Larry Langford is discussed in Tomberlin (2009) and that of California governor Arnold Schwarzenegger in McKinley (2009).