Chapter 16

Information Cascades

16.1 Following the Crowd

When people are connected by a network, it becomes possible for them to influence each other's behavior and decisions. In the next several chapters, we will explore how this basic principle gives rise to a range of social processes in which networks serve to aggregate individual behavior and thus produce population-wide, collective outcomes.

There is a nearly limitless set of situations in which people are influenced by others: in the opinions they hold, the products they buy, the political positions they support, the activities they pursue, the technologies they use, and many other things. What we'd like to do here is to go beyond this observation and consider some of the reasons why such influence occurs. We'll see that there are many settings in which it may in fact be rational for an individual to imitate the choices of others even if the individual's own information suggests an alternative choice.

As a first example, suppose that you are choosing a restaurant in an unfamiliar town, and based on your own research about restaurants you intend to go to restaurant A. However, when you arrive you see that no one is eating in restaurant A while restaurant B next door is nearly full. If you believe that other diners have tastes similar to yours, and that they too have some information about where to eat, it may be rational to join the crowd at B rather than to follow your own information. To see how this is possible, suppose that each diner has obtained independent but imperfect information about which of the two restaurants is better. Then if there are already many diners in restaurant B, the information that you can infer from their choices may be more powerful than your own private information, in which case it would in fact make sense for you to join them regardless of your own private

D. Easley and J. Kleinberg. *Networks, Crowds, and Markets: Reasoning about a Highly Connected World.* Cambridge University Press, 2010. Draft version: June 10, 2010.

information. In this case, we say that *herding*, or an *information cascade*, has occurred. This terminology, as well as this example, comes from the work of Banerjee [40]; the concept was also developed in other work around the same time by Bikhchandani, Hirshleifer, and Welch [59, 412].

Roughly, then, an information cascade has the potential to occur when people make decisions sequentially, with later people watching the actions of earlier people, and from these actions inferring something about what the earlier people know. In our restaurant example, when the first diners to arrive chose restaurant B, they conveyed information to later diners about what they knew. A cascade then develops when people abandon their own information in favor of inferences based on earlier people's actions.

What is interesting here is that individuals in a cascade are imitating the behavior of others, but it is not mindless imitation. Rather, it is the result of drawing rational inferences from limited information. Of course, imitation may also occur due to social pressure to conform, without any underlying informational cause, and it is not always easy to tell these two phenomena apart. Consider for example the following experiment performed by Milgram, Bickman, and Berkowitz in the 1960s [298]. The experimenters had groups of people ranging in size from just one person to as many as fifteen people stand on a street corner and stare up into the sky. They then observed how many passersby stopped and also looked up at the sky. They found that with only one person looking up, very few passersby stopped. If five people were staring up into the sky, then more passersby stopped, but most still ignored them. Finally, with fifteen people looking up, they found that 45% of passersby stopped and also stared up into the sky.

The experimenters interpreted this result as demonstrating a social force for conformity that grows stronger as the group conforming to the activity becomes larger. But another possible explanation — essentially, a possible mechanism giving rise to the conformity observed in this kind of situation — is rooted in the idea of information cascades. It could be that initially the passersby saw no reason to look up (they had no private or public information that suggested it was necessary), but with more and more people looking up, future passersby may have rationally decided that there was good reason to also look up (since perhaps those looking up knew something that the passersby didn't know).

Ultimately, information cascades may be at least part of the explanation for many types of imitation in social settings. Fashions and fads, voting for popular candidates, the self-reinforcing success of books placed highly on best-seller lists, the spread of a technological choice by consumers and by firms, and the localized nature of crime and political movements can all be seen as examples of herding, in which people make decisions based on inferences from what earlier people have done.

Informational effects vs. Direct-Benefit Effects. There is also a fundamentally different class of rational reasons why you might want to imitate what other people are doing. You may want to copy the behavior of others if there is a direct benefit to you from aligning your behavior with their behavior. For example, consider the first fax machines to be sold. A fax machine is useless if no one else owns one, and so in evaluating whether to buy one, it's very important to know whether there are other people who own one as well — not just because their purchase decisions convey information, but because they directly affect the fax machine's value to you as a product. A similar argument can be made for computer operating systems, social networking sites, and other kinds of technology where you directly benefit from choosing an option that has a large user population.

This type of direct-benefit effect is different from the informational effects we discussed previously: here, the actions of others are affecting your payoffs directly, rather than indirectly by changing your information. Many decisions exhibit both information and direct-benefit effects — for example, in the technology-adoption decisions just discussed, you potentially learn from others' decisions in addition to benefitting from compatibility with them. In some cases, the two effects are even in conflict: if you have to wait in a long line to get into a popular restaurant, you are choosing to let the informational benefits of imitating others outweigh the direct inconvenience (from waiting) that this imitation causes you.

In this chapter, we develop some simple models of information cascades; in the next chapter, we do this for direct-benefit effects. One reason to develop minimal, stylized models for these effects is to see whether the stories we've been telling can have a simple basis — and we will see that much of what we've been discussing at an informal level can indeed be represented in very basic models of decision-making by individuals.

16.2 A Simple Herding Experiment

Before delving into the mathematical models for information cascades [40, 59, 412], we start with a simple herding experiment created by Anderson and Holt [14, 15] to illustrate how these models work.

The experiment is designed to capture situations with the basic ingredients from our discussion in the previous section:

(a) There is a decision to be made — for example, whether to adopt a new technology, wear a new style of clothing, eat in a new restaurant, or support a particular political position.

(b) People make the decision sequentially, and each person can observe the choices made by those who acted earlier.

(c) Each person has some private information that helps guide their decision.

(d) A person can't directly observe the private information that other people *know*, but he or she can make inferences about this private information from what they *do*.

We imagine the experiment taking place in a classroom, with a large group of students as participants. The experimenter puts an urn at the front of the room with three marbles hidden in it; she announces that there is a 50% chance that the urn contains two red marbles and one blue marble, and a 50% chance the urn contains two blue marbles and one red marble. In the former case, we will say that it is a "majority-red" urn, and in the latter case, we will say that it is a "majority-blue" urn.[1]

Now, one by one, each student comes to the front of the room and draws a marble from the urn; he looks at the color and then places it back in the urn without showing it to the rest of the class. The student then guesses whether the urn is majority-red or majority-blue and publicly announces this guess to the class. (We assume that at the very end of the experiment, each student who has guessed correctly receives a monetary reward, while students who have guessed incorrectly receive nothing.) The public announcement is the key part of the set-up: the students who have not yet had their turn don't get to see which colors the earlier students draw, but they do get to hear the guesses that are being made. This parallels our original example with the two restaurants: one-by-one, each diner needs to guess which is the better restaurant, and while they don't get to see the reviews read by the earlier diners, they do get to see which restaurant these earlier diners chose.

Let's now consider what we should expect to happen when this experiment is performed. We will assume that all the students reason correctly about what to do when it is their turn to guess, using everything they have heard so far. We will keep the analysis of the experiment informal, and later use a mathematical model to justify it more precisely.

We organize the discussion by considering what happens with each student in order. Things are fairly straightforward for the first two students; they become interesting once we reach the third student.

- *The First Student.* The first student should follow a simple decision rule for making a guess: if he sees a red marble, it is better to guess that the urn is majority-red; and if he sees a blue marble, it is better to guess that the urn is majority-blue. (This is an intuitively natural rule, and — as with the other conclusions we draw here — we will justify it later mathematically using the model we develop in the subsequent sections.) This means the first student's guess conveys perfect information about what he has seen.

[1]It's important that the students believe this statement about probabilities. So you can imagine, if you like, that the experimenter has actually filled two urns with marbles. One has two red marbles and one blue marble, and the other urn contains two blue marbles and one red marble. One of these urns is selected at random, with equal probability on each urn, and this is the urn used in the experiment.

- *The Second Student.* If the second student sees the same color that the first student announced, then her choice is simple: she should guess this color as well.

  Suppose she sees the opposite color — say that she sees red while the first guess was blue. Since the first guess was exactly what the first student saw, the second student can essentially reason as though she got to draw twice from the urn, seeing blue once and red once. In this case, she is indifferent about which guess to make; we will assume in this case that she breaks the tie by guessing the color she saw. Thus, whichever color the second student draws, her guess too conveys perfect information about what she has seen.

- *The Third Student.* Things start to get interesting here. If the first two students have guessed opposite colors, then the third student should just guess the color he sees, since it will effectively break the tie between the first two guesses.

  But suppose the first two guesses have been the same — say they've both been blue — and the third student draws red. Since we've decided that the first two guesses convey perfect information, the third student can reason in this case as though he saw three draws from the urn: two blue, and one red. Given this information, he should guess that the urn is majority-blue, ignoring his own private information (which, taken by itself, suggested that the urn is majority-red).

  More generally, the point is that when the first two guesses are the same, the third student should guess this color as well, *regardless* of which color he draws from the urn. And the rest of class will only hear his guess; they don't get to see which color he's drawn. In this case, an information cascade has begun. The third student makes the same guess as the first two, regardless of which color he draws from the urn, and hence regardless of his own private information.

- *The Fourth Student and Onward.* For purposes of this informal discussion, let's consider just the "interesting" case above, in which the first two guesses were the same — suppose they were both blue. In this case, we've argued that the third student will also announce a guess of blue, regardless of what he actually saw.

  Now consider the situation faced by the fourth student, getting ready to make a guess having heard three guesses of "blue" in a row. She knows that the first two guesses conveyed perfect information about what the first two students saw. She also knows that, given this, the third student was going to guess "blue" no matter what he saw — so his guess conveys no information.

  As a result, the fourth student is in exactly the same situation — from the point of view of making a decision — as the third student. Whatever color she draws, it will be outweighed by the two draws of blue by the first two students, and so she should guess "blue" regardless of what she sees.

This will continue with all the subsequent students: if the first two guesses were "blue," then everyone in order will guess "blue" as well. (Of course, a completely symmetric thing happens if the first two guesses are "red".) An information cascade has taken hold: no one is under the illusion that every single person is drawing a blue marble, but once the first two guesses turn out "blue," the future announced guesses become worthless and so everyone's best strategy is to rely on the limited genuine information they have available.

In the next section, we'll discuss a model of decision-making under uncertainty that justifies the guesses made by the students. More generally, our discussion hasn't considered every possible eventuality (for example, what should you do if you're the sixth student and you've heard the guesses "blue, red, red, blue, blue"?), but our subsequent model will actually predict an outcome for any sequence of guesses.

For now, though, let's think about the particular scenario discussed here — the way in which a cascade takes place as long as the first two guesses are the same. Although the setting is very stylized, it teaches us a number of general principles about information cascades. First, it shows how easily they can occur, given the right structural conditions. It also shows how a bizarre pattern of decisions — each of a large group of students making exactly the same guess — can take place even when all the decision-makers are being completely rational.

Second, it shows that information cascades can lead to non-optimal outcomes. Suppose for example that we have an urn that is majority-red. There is a $\frac{1}{3}$ chance that the first student draws a blue marble, and a $\frac{1}{3}$ chance that the second student draws a blue marble; since these draws are independent, there is a $\frac{1}{3} \cdot \frac{1}{3} = \frac{1}{9}$ chance that both do. In this case, both of the first two guesses will be "blue"; so, as we have just argued, all subsequent guesses will be "blue" — and all of these guesses will be wrong, since the urn is majority-red. This $\frac{1}{9}$ chance of a population-wide error is not ameliorated by having many people participate, since under rational decision-making, everyone will guess blue if the first two guesses are blue, no matter how large the group is.

Third, this experiment illustrates that cascades — despite their potential to produce long runs of conformity — can be fundamentally very fragile. Suppose, for example, that in a class of 100 students, the first two guesses are "blue," and all subsequent guesses are proceeding — as predicted — to be "blue" as well. Now, suppose that students 50 and 51 both draw red marbles, and they each "cheat" by showing their marbles directly to the rest of the class. In this case, the cascade has been broken: when student 52 gets up to make a guess, she has four pieces of genuine information to go on: the colors observed by students 1, 2, 50, and 51. Since two of these colors are blue and two are red, she should make the

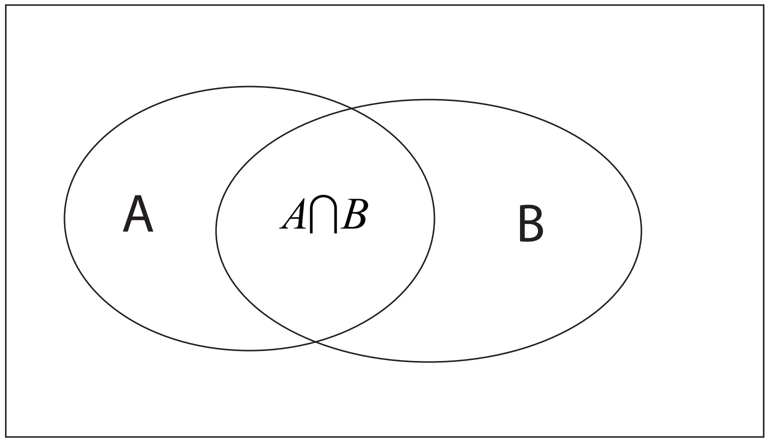

Figure 16.1: Two events A and B in a sample space, and the joint event $A \cap B$.

guess based on her own draw, which will break the tie.

The point is that everyone knew the initial run of 49 "blue" guesses had very little information supporting it, and so it was easy for a fresh infusion of new information to overturn it. This is the essential fragility of information cascades: even after they have persisted for a long time, they can be overturned with comparatively little effort.[2]

This style of experiment has generated a significant amount of subsequent research in its own right, and understanding the extent to which human subjects follow the type of behavior described above under real experimental conditions is a subtle issue [100, 223]. For our purposes, however, the simple description of the experiment is intended to serve mainly as a vivid illustration of some of the basic properties of information cascades in a controlled setting. Having now developed some of these basic properties, we turn to the formulation of a model that lets us reason precisely about the decision-making that takes place during a cascade.

16.3 Bayes' Rule: A Model of Decision-Making Under Uncertainty

If we want to build a mathematical model for how information cascades occur, it will necessarily involve people asking themselves questions like, "What is the probability this is the

[2]It is important to note that not all imitative effects are so easy to overturn. As we will see in the next chapter, for example, imitation based on direct-benefit effects can be very difficult to reverse once it is underway.

better restaurant, given the reviews I've read and the crowds I see in each one?" Or, "What is the probability this urn is majority-red, given the marble I just drew and the guesses I've heard?" In other words, we need a way to determine probabilities of events given information that is observed.

Conditional Probability and Bayes' Rule. We will be computing the probability of various *events*, and using these to reason about decision-making. In the context of the experiment from Section 16.2, an event could be "The urn is majority-blue," or "the first student draws a blue marble." Given any event A, we will denote its probability of occuring by $\Pr[A]$. Whether an event occurs or not is the result of certain random outcomes (which urn was placed at the front of the room, which marble a particular student grabbed when he reached in, and so forth). We therefore imagine a large *sample space*, in which each point in the sample space consists of a particular realization for each of these random outcomes.

Given a sample space, events can be pictured graphically as in Figure 16.1: the unit-area rectangle in the figure represents the sample space of all possible outcomes, and the event A is then a region within this sample space — the set of all outcomes where event A occurs. In the figure, the probability of A corresponds to the area of this region. The relationship between two events can be illustrated graphically as well. In Figure 16.1 we see two events A and B. The area where they overlap corresponds to the joint event when both A and B occur. This event is the intersection of A and B, and it's denoted by $A \cap B$.

If we think about the examples of questions at the start of this section, we see that it is not enough to talk about the probability of an event A; rather, we need to consider the probability of A, *given that some other event B has occurred*. For example, A may be the event that the urn in the experiment from Section 16.2 is majority-blue, and B may be the event that the ball you've drawn is blue. We will refer to this quantity as the *conditional probability of A given B*, and denote it by $\Pr[A \mid B]$. Again, the graphical depiction in Figure 16.1 is useful: to determine the conditional probability of A given B, we assume that we are in the part of the sample space corresponding to B, and we want to know the probability that we are also in A (that is, in $A \cap B$). We can think of this as the fraction of the area of region B occupied by $A \cap B$, and so we define

$$\Pr[A \mid B] = \frac{\Pr[A \cap B]}{\Pr[B]}. \tag{16.1}$$

Similarly, the conditional probability of B given A is

$$\Pr[B \mid A] = \frac{\Pr[B \cap A]}{\Pr[A]} = \frac{\Pr[A \cap B]}{\Pr[A]}, \tag{16.2}$$

where the second equality follows simply because $A \cap B$ and $B \cap A$ are the same set.

Rewriting (16.1) and (16.2), we have

$$\Pr[A \mid B] \cdot \Pr[B] = \Pr[A \cap B] = \Pr[B \mid A] \cdot \Pr[A], \tag{16.3}$$

and therefore, dividing through by $\Pr[B]$,

$$\Pr[A \mid B] = \frac{\Pr[A] \cdot \Pr[B \mid A]}{\Pr[B]}. \tag{16.4}$$

Equation (16.4) is called *Bayes' Rule*. There is also a bit of extra useful terminology surrounding Bayes' Rule. When we want to make explicit that we're interested in the effect of event B on the probability of an event A, we refer to $\Pr[A]$ as the *prior probability* of A, since it reflects our understanding of the probability of A without knowing anything about whether B has occurred. Correspondingly, we refer to $\Pr[A \mid B]$ as the *posterior probability* of A given B, since it reflects our new understanding of the probability of A now that we know B has occurred. The effect of knowing B is thus captured in the change from the prior probability of A to the posterior probability of A, using Equation (16.4).

An Example of Bayes' Rule. As noted above, we will be applying Bayes' Rule in cases where a decision-maker is assessing the probability that a particular choice is the best one, *given* the event that he has received certain private information and/or observed certain other decisions. To get used to Bayes' Rule, we first work through a basic example that illustrates how it is typically applied.

The example involves eyewitness testimony. Suppose that in some city 80% of taxi cabs are black and the remaining 20% are yellow. A witness to a hit-and-run accident involving a taxi states that the cab involved was yellow. Suppose that eyewitness testimony is imperfect in the sense that witnesses sometimes misidentify the colors of cabs. In particular, let's suppose that if a taxi is yellow then a witness will claim it is yellow after the fact 80% of the time; and if it is black, they will claim it is black 80% of the time.

Interpreting eyewitness testimony, therefore, is at some level a question of conditional probability: what is the probability the cab is yellow (or black), given that the witness says it is yellow? Introducing some notation, let *true* denote the true color of the cab, and let *report* denote the reported color of the cab; let Y denote yellow and B denote black. We are looking for the value of $\Pr[\mathit{true} = Y \mid \mathit{report} = Y]$.

The data we have does not directly include the answer to this question, but we can determine the answer using Bayes' Rule. Applying Equation (16.4) with A equal to the event $\mathit{true} = Y$ and B equal to the event $\mathit{report} = Y$, we have

$$\Pr[\mathit{true} = Y \mid \mathit{report} = Y] = \frac{\Pr[\mathit{true} = Y] \cdot \Pr[\mathit{report} = Y \mid \mathit{true} = Y] \cdot}{\Pr[\mathit{report} = Y]}. \tag{16.5}$$

Now, we've been told that $\Pr[\mathit{report} = Y \mid \mathit{true} = Y]$ is 0.8 (this is the accuracy of eyewitness testimony) and that $\Pr[\mathit{true} = Y]$ is 0.2 (this is the frequency of yellow taxi cabs, and hence provides the prior probability of the event $\mathit{true} = Y$). We can also figure out the denominator with a little work, as follows. There are two ways for a witness to report that a cab is yellow:

one is for the cab to actually be yellow, and the other is for it to actually be black. The probability of getting a report of yellow via the former option is

$$\Pr[true = Y] \cdot \Pr[report = Y \mid true = Y] = 0.2 \cdot 0.8 = 0.16,$$

and the probability of getting a report of yellow via the latter option is

$$\Pr[true = B] \cdot \Pr[report = Y \mid true = B] = 0.8 \cdot 0.2 = 0.16.$$

The probability of a report of yellow is the sum of these two probabilities,

$$\begin{aligned}\Pr[report = Y] &= \Pr[true = Y] \cdot \Pr[report = Y \mid true = Y] + \\ &\quad \Pr[true = B] \cdot \Pr[report = Y \mid true = B] \\ &= 0.2 \cdot 0.8 + 0.8 \cdot 0.2 = 0.32.\end{aligned}$$

We can now put everything together via Equation (16.5) so as to get

$$\begin{aligned}\Pr[true = Y \mid report = Y] &= \frac{\Pr[true = Y] \cdot \Pr[report = Y \mid true = Y]}{\Pr[report = Y]} \\ &= \frac{0.2 \cdot 0.8}{0.32} \\ &= 0.5.\end{aligned}$$

So the conclusion is that if the witness says the cab was yellow, it is in fact equally likely to have been yellow or black. Since the frequency of black and yellow cabs makes black substantially more likely in the absence of any other information (0.8 versus 0.2), the witness's report had a substantial effect on our beliefs about the color of the particular cab involved. But the report should not lead us to believe that the cab was in fact more likely to have been yellow than black.[3]

A second example: Spam filtering. As the example with taxi cabs illustrates, Bayes' Rule is a fundamental way to make inferences from observations, and as such it is used in a wide variety of settings. One application where it has been very influential is in e-mail spam detection — automatically filtering unwanted e-mail out of a user's incoming e-mail stream. Bayes' Rule was a crucial conceptual ingredient in the first generation of e-mail spam filters, and it continues to form part of the foundation for many spam filters [187].

We can appreciate the connection between Bayes' Rule and spam filtering through the following example. Suppose that you receive a piece of e-mail whose subject line contains

[3]Kahneman and Tversky have run an experiment with a similar example which shows that people sometimes do not make predictions according to Bayes' Rule [231]. In their experiment, subjects place too much weight on their observations and too little weight on prior probabilities. The effect of errors in predictions on actions, and the subsequent effect on cascades is an interesting topic, but we will not address it here.

the phrase "check this out" (a popular phrase among spammers). Based just on this (and without looking at the sender or the message content), what is the chance the message is spam?

This is already a question about conditional probability: we're asking for the value of

$$\Pr\left[\textit{message is spam} \mid \textit{subject contains ``check this out''}\right].$$

To make this equation and the ones that follow a bit simpler to read, let's abbreviate *message is spam* to just *spam*, and abbreviate *subject contains "check this out"* to just *"check this out"*; so we want the value of

$$\Pr\left[\textit{spam} \mid \textit{``check this out''}\right].$$

To determine this value, we need to know some facts about your e-mail and the general use of the phrase "check this out" in subject lines. Suppose that 40% of all your e-mail is spam and the remaining 60% is e-mail you want to receive. Also, suppose that 1% of all spam messages contain the phrase "check this out" in their subject lines, while 0.4% of all non-spam messages contain this phrase. Writing these in terms of probabilities, it says that $\Pr[\textit{spam}] = 0.4$; this is the prior probability that an incoming message is spam (without conditioning on events based on the message itself). Also, we have

$$\Pr\left[\textit{``check this out''} \mid \textit{spam}\right] = .01$$

and

$$\Pr\left[\textit{``check this out''} \mid \textit{not spam}\right] = .004$$

We're now in a situation completely analogous to the calculations involving eyewitness testimony: we can use Bayes' Rule to write

$$\Pr\left[\textit{spam} \mid \textit{``check this out''}\right] = \frac{\Pr[\textit{spam}] \cdot \Pr\left[\textit{``check this out''} \mid \textit{spam}\right]}{\Pr\left[\textit{``check this out''}\right]}.$$

Based on what we know, we can determine that the numerator is $.4 \cdot .01 = .004$. For the denominator, as in the taxicab example, we note that there are two ways for a message to contain "check this out" — either by being spam or by not being spam. As in that calculation,

$$\begin{aligned} \Pr\left[\textit{``check this out''}\right] &= \Pr[\textit{spam}] \cdot \Pr\left[\textit{``check this out''} \mid \textit{spam}\right] + \\ &\quad \Pr[\textit{not spam}] \cdot \Pr\left[\textit{``check this out''} \mid \textit{not spam}\right] \\ &= .4 \cdot .01 + .6 \cdot .004 = .0064. \end{aligned}$$

Dividing numerator by denominator, we get our answer:

$$\Pr\left[\textit{spam} \mid \textit{``check this out''}\right] = \frac{.004}{.0064} = \frac{5}{8} = .625.$$

In other words, although spam (in this example) forms less than half of your incoming e-mail, a message whose subject line contains the phrase "check this out" is — in the absence of any other information — more likely to be spam than not.

We can therefore view the presence of this phrase in the subject line as a weak "signal" about the message, providing us with evidence about whether it's spam. In practice, spam filters built on Bayes' Rule look for a wide range of different signals in each message — the words in the message body, the words in the subject, properties of the sender (do you know them? what kind of an e-mail address are they using?), properties of the mail program used to compose the message, and other features. Each of these provides its own estimate for whether the message is spam or not, and spam filters then combine these estimates to arrive at an overall guess about whether the message is spam. For example, if we also knew that the message above came from someone you send mail to every day, then presumably this competing signal — strongly indicating that the message is not spam — should outweigh the presence of the phrase "check this out" in the subject.

16.4 Bayes' Rule in the Herding Experiment

Let's now use Bayes' Rule to justify the reasoning that the students used in the simple herding experiment from Section 16.2. First, notice that each student's decision is intrinsically based on determining a conditional probability: each student is trying to estimate the conditional probability that the urn is majority-blue or majority-red, given what she has seen and heard. To maximize her chance of winning the monetary reward for guessing correctly, she should guess majority-blue if

$$\Pr\left[\textit{majority-blue} \mid \textit{what she has seen and heard}\right] > \frac{1}{2}$$

and guess majority-red otherwise. If the two conditional probabilities are both exactly 0.5, then it doesn't matter what she guesses.

We know the following facts from the set-up of the experiment, before anyone has drawn any marbles. First, the prior probabilities of majority-blue and majority-red are each $\frac{1}{2}$:

$$\Pr\left[\textit{majority-blue}\right] = \Pr\left[\textit{majority-red}\right] = \frac{1}{2}.$$

Also, based on the composition of the two kinds of urns,

$$\Pr\left[\textit{blue} \mid \textit{majority-blue}\right] = \Pr\left[\textit{red} \mid \textit{majority-red}\right] = \frac{2}{3}.$$

Now, following the scenario from Section 16.2, let's suppose that the first student draws a blue marble. He therefore wants to determine $\Pr\left[\textit{majority-blue} \mid \textit{blue}\right]$, and just as in the examples from Section 16.3, he can use Bayes' Rule to calculate

$$\Pr\left[\textit{majority-blue} \mid \textit{blue}\right] = \frac{\Pr\left[\textit{majority-blue}\right] \cdot \Pr\left[\textit{blue} \mid \textit{majority-blue}\right]}{\Pr\left[\textit{blue}\right]}. \tag{16.6}$$

The numerator is $\frac{1}{2} \cdot \frac{2}{3} = \frac{1}{3}$. For the denominator, we reason just as in Section 16.3 by noting that there are two possible ways to get a blue marble — if the urn is majority-blue, or if it is majority-red:

$$\begin{aligned}\Pr[\mathit{blue}] &= \Pr[\mathit{majority\text{-}blue}] \cdot \Pr[\mathit{blue} \mid \mathit{majority\text{-}blue}] + \\ &\quad \Pr[\mathit{majority\text{-}red}] \cdot \Pr[\mathit{blue} \mid \mathit{majority\text{-}red}] \\ &= \frac{1}{2} \cdot \frac{2}{3} + \frac{1}{2} \cdot \frac{1}{3} = \frac{1}{2}.\end{aligned}$$

The answer $\Pr[\mathit{blue}] = \frac{1}{2}$ makes sense, given that the roles of blue and red in this experiment are completely symmetric.

Dividing numerator by denominator, we get

$$\Pr[\mathit{majority\text{-}blue} \mid \mathit{blue}] = \frac{1/3}{1/2} = \frac{2}{3}.$$

Since this conditional probability is greater than $\frac{1}{2}$, we get the intuitive result that the first student should guess majority-blue when he sees a blue marble. Note that in addition to providing the basis for the guess, Bayes' Rule also provides a probability, namely $\frac{2}{3}$, that the guess will be correct.

The calculation is very similar for the second student, and we skip this here to move on to the calculation for the third student, where a cascade begins to form. Let's suppose, as in the scenario from Section 16.2, that the first two students have announced guesses of blue, and the third student draws a red marble. As we discussed there, the first two guesses convey genuine information, so the third student knows that there have been three draws from the urn, consisting of the sequence of colors blue, blue, and red. What he wants to know is

$$\Pr[\mathit{majority\text{-}blue} \mid \mathit{blue, blue, red}]$$

so as to make a guess about the urn. Using Bayes' Rule we get

$$\Pr[\mathit{majority\text{-}blue} \mid \mathit{blue, blue, red}] = \frac{\Pr[\mathit{majority\text{-}blue}] \cdot \Pr[\mathit{blue, blue, red} \mid \mathit{majority\text{-}blue}]}{\Pr[\mathit{blue, blue, red}]}. \tag{16.7}$$

Since the draws from the urn are independent, the probability $\Pr[\mathit{blue, blue, red} \mid \mathit{majority\text{-}blue}]$ is determined by multiplying the probabilities of the three respective draws together:

$$\Pr[\mathit{blue, blue, red} \mid \mathit{majority\text{-}blue}] = \frac{2}{3} \cdot \frac{2}{3} \cdot \frac{1}{3} = \frac{4}{27}.$$

To determine $\Pr[\mathit{blue, blue, red}]$, as usual we consider the two different ways this sequence could have happened — if the urn is majority-blue, or if it is majority-red:

$$\begin{aligned}\Pr[\mathit{blue, blue, red}] &= \Pr[\mathit{majority\text{-}blue}] \cdot \Pr[\mathit{blue, blue, red} \mid \mathit{majority\text{-}blue}] + \\ &\quad \Pr[\mathit{majority\text{-}red}] \cdot \Pr[\mathit{blue, blue, red} \mid \mathit{majority\text{-}red}] \\ &= \frac{1}{2} \cdot \frac{2}{3} \cdot \frac{2}{3} \cdot \frac{1}{3} + \frac{1}{2} \cdot \frac{1}{3} \cdot \frac{1}{3} \cdot \frac{2}{3} = \frac{6}{54} = \frac{1}{9}.\end{aligned}$$

Plugging all this back into Equation (16.7), we get

$$\Pr\left[\textit{majority-blue} \mid \textit{blue, blue, red}\right] = \frac{\frac{4}{27}\cdot\frac{1}{2}}{\frac{1}{9}} = \frac{2}{3}.$$

Therefore, the third student should guess majority-blue (from which he will have a $\frac{2}{3}$ chance of being correct) — this confirms our intuitive observation in Section 16.2 that the student should ignore what he sees (red) in favor of the two guesses he's already heard (both blue).

Finally, once these three draws from the urn have taken place, all future students will have the same information as the third student, and so they will all perform the same calculation, resulting in an information cascade of blue guesses.

16.5 A Simple, General Cascade Model

Let's return to the motivation for the herding experiment in Section 16.2: the experiment served as a stylized metaphor for any situation in which people make decisions sequentially, basing these decisions on a combination of their own private information and observations of what earlier people have done. We now formulate a model that covers such situations in general. We will see that Bayes' Rule predicts in this general model that cascades will form, with probability tending to 1 as the number of people goes to infinity.

Formulating the Model. Consider a group of people (numbered $1, 2, 3, \ldots$) who will sequentially make decisions — that is, individual 1 will decide first, then individual 2 will decide, and so on. We will describe the decision as a choice between *accepting* or *rejecting* some option: this could be a decision about whether to adopt a new technology, wear a new fashion, eat in a new restaurant, commit a crime, vote for a particular political candidate, or choose one route to a common destination rather than an alternative route.

First model ingredient: States of the world. At the start of everything, before any individual has made a decision, we assume that the world is randomly placed into one of two possible *states*: it is either placed in a state in which the option is actually a good idea, or a state in which the option is actually a bad idea. We imagine that the state of the world is determined by some initial random event that the individuals can't observe, but they will try to use what they observe to make inferences about this state. For example, the world is either in a state where the new restaurant is good or a state where it is bad; the individuals in the model know that it was randomly placed in one of these two states, and they're trying to figure out which.

We write the two possible states of the world as G, representing the state where the option is a good idea, and B, representing the state where the option is a bad idea. We suppose that each individual knows the following fact: the initial random event that placed