Chapter 2
Scientific inference

Scientists often tell us things about the world that we would not otherwise have believed. For example, biologists tell us that we are closely related to chimpanzees, geologists tell us that Africa and South America used to be joined together, and cosmologists tell us that the universe is expanding. But how did scientists reach these unlikely sounding conclusions? After all, no one has ever seen one species evolve from other, or a single continent split into two, or the universe getting bigger. The answer, of course, is that scientists arrived at these beliefs by a process of *reasoning* or *inference*. But it would be nice to know more about this process. What exactly is the nature of scientific inference?

Deduction and induction

Logicians make an important distinction between *deductive* and *inductive* inference, or deduction and induction for short. An example of a deductive inference is the following:

> All Frenchmen like red wine
> Pierre is a Frenchman
> ______________________
> Therefore, Pierre likes red wine

The two statements above the line are called the *premises* of the inference, while the statement below the line is called the *conclusion*. This is a deductive inference because it has the following property: *if* the premises are true, then the conclusion must be true too. If it's true that all Frenchmen like red wine, and that Pierre is a Frenchman, it follows that Pierre does indeed like red wine. This is sometimes expressed by saying that the premises of the inference *entail* the conclusion. Of course the premises of this inference are almost certainly not true—there are bound to be Frenchmen who do not like red wine. But that is not the point. What makes the inference deductive is the existence of an appropriate relation between premises and conclusion, namely that the truth of the premises guarantees the truth of the conclusion.

Not all inferences are deductive. Consider the following example:

> The first five eggs in the box were good.
> All the eggs have the same best-before date stamped on them.
>
> ---
>
> Therefore, the sixth egg will be good too.

This looks like a perfectly sensible piece of reasoning. But nonetheless it is not deductive, for the premises do not entail the conclusion. Even if the first five eggs were good, and all the eggs do have the same date stamp, it is quite conceivable that the sixth egg will be rotten. That is, it is logically possible for the premises of this inference to be true and yet the conclusion false, so the inference is not deductive. Instead it is known as an *inductive* inference. In a typical inductive inference, we move from premises about objects that we have examined to conclusions about objects of the same sort that we haven't examined—in this example, eggs.

Deductive inference is safer than its inductive cousin. When we reason deductively, we can be certain that if we start with true premises we will end up with a true conclusion. By contrast, inductive reasoning is quite capable of taking us from true premises to a false conclusion. Despite this defect, we seem to rely on inductive reasoning throughout our lives. For example, when you turn on your computer in the morning, you are confident it will not explode in your face. Why? Because you turn on your computer every morning, and it has never exploded up to now. But the inference from 'up until now, my computer has not exploded when I turned it on' to 'my computer will not explode this time' is inductive, not deductive. It is logically possible that your computer will explode this time, even though it has never done so before.

Do scientists use inductive reasoning too? The answer seems to be yes. Consider the condition known as Down's syndrome (DS). Geneticists tell us that people with DS have three copies of chromosome 21 instead of the usual two. How do they know this? The answer, of course, is that they examined a large number of people with DS and found that each had an additional copy of chromosome 21. They then reasoned inductively to the conclusion that *all* people with DS, including those they hadn't examined, have an additional copy. This inference is inductive not deductive. For it is possible, though unlikely, that the sample examined was unrepresentative. This example is not an isolated one. In effect, scientists reason inductively whenever they move from limited data to a more general conclusion, which they do all the time.

The central role of induction in science is sometimes obscured by how we talk. For example, you might read a newspaper report which says that scientists have found 'experimental proof' that genetically modified maize is safe to eat. What this means is that the scientists have tested the maize on a large number of people and none have come to any harm. But strictly speaking this doesn't *prove* that the maize is safe, in the sense in which mathematicians can prove Pythagoras' theorem, say. For the inference from 'the maize didn't harm any of the people on whom it was tested' to 'the maize will not harm anyone' is inductive, not deductive. The newspaper report should really have said that scientists have found good *evidence* that the maize is safe for humans. The word 'proof' should strictly only be used when we are dealing with deductive inferences. In this strict sense of the word, scientific hypotheses can rarely if ever be proved true by the data.

Most philosophers think it's obvious that science relies heavily on induction, indeed so obvious that it hardly needs arguing for. But remarkably, this was denied by the philosopher Karl Popper, whom we met in the last chapter. Popper claimed that scientists only need to use deductive inferences. This would be nice if it were true, for deductive inferences are safer than inductive ones, as we have seen.

Popper's basic argument was this. Although a scientific theory (or hypothesis) can never be proved true by a finite amount of data, it can be proved false, or refuted. Suppose a scientist is testing the hypothesis that all pieces of metal conduct electricity. Even if every piece of metal they examine conducts electricity, this doesn't prove that the hypothesis is true, for reasons that we've seen. But if the scientist finds even one piece of metal that fails to conduct electricity, this conclusively refutes the theory. For the inference from 'this piece of metal does not conduct electricity' to 'it is false that all pieces of metal conduct electricity' is a deductive inference—the premise entails the conclusion. So if a scientist were trying to refute their theory, rather than establish its truth, their goal could be accomplished without the use of induction.

The weakness of Popper's argument is obvious. For the goal of science is not solely to refute theories, but also to determine which theories are true (or probably true). When a scientist collects experimental data, their aim *might* be to show that a particular theory—their arch-rival's theory perhaps—is false. But much more likely, they are trying to convince people that their own theory is true. And in order to do that, they will have to resort to inductive reasoning of some sort. So Popper's attempt to show that science can get by without induction does not succeed.

Hume's problem

Although inductive reasoning is not logically watertight, it seems like a sensible way of forming beliefs about the world. Surely the fact that the sun has risen every day in the past gives us good reason to believe that it will rise tomorrow? If you came across someone who professed to be entirely agnostic about whether the sun will rise tomorrow or not, you would regard them as very strange indeed, if not irrational.

But what justifies this faith we place in induction? How should we go about persuading someone who refuses to reason inductively that they are wrong? The 18th-century Scottish philosopher David Hume (1711–76) gave a simple but radical answer to this question. He argued that the use of induction cannot be rationally justified at all. Hume admitted that we use induction all the time, in everyday life and in science, but insisted that this was a matter of brute animal habit. If challenged to provide a good reason for using induction, we can give no satisfactory answer, he thought.

How did Hume arrive at this startling conclusion? He began by noting that whenever we make inductive inferences, we seem to presuppose what he called the 'uniformity of nature'. To see what Hume meant by this, recall our examples. We had the inference from 'the first five eggs in the box were good' to 'the sixth egg will be good'; from 'the Down's syndrome patients examined had an extra chromosome' to 'all those with Down's syndrome have an extra chromosome'; and from 'my computer has never exploded until now' to 'my computer will not explode today'. In each case, our reasoning seems to depend on the assumption that objects we haven't examined will be similar, in relevant respects, to objects of the same sort that we have examined. That assumption is what Hume means by the uniformity of nature.

But how do we know that the uniformity assumption is true? Can we perhaps prove its truth somehow? No, says Hume, we cannot. For it is easy to *imagine* a world where nature is not uniform but changes its course randomly from day to day. In such a world, computers might sometimes explode for no reason, water might sometimes intoxicate us without warning, and billiard balls might sometimes stop dead on colliding. Since such a non-uniform world is conceivable, it follows that we cannot prove that the uniformity assumption is true. For if we could, then the non-uniform universe would be a logical impossibility.

Even if we cannot prove the uniformity assumption, we might nonetheless hope to find good empirical evidence for its truth. After all, the assumption has always held good up to now, so surely this is evidence that it is true? But this begs the question, says Hume! Grant that nature has behaved largely uniformly up to now. We cannot appeal to this fact to argue that nature will continue to be uniform, says Hume, because this assumes that what has happened in the past is a reliable guide to what will happen in the future—which *is* the uniformity of nature assumption. If we try to argue for the uniformity assumption on empirical grounds, we end up reasoning in a circle.

The force of Hume's point can be appreciated by imagining how you would persuade someone who doesn't trust inductive reasoning that they should. You might say: 'look, inductive reasoning has worked pretty well up until now. By using induction scientists have split the atom, landed on the moon, and invented lasers. Whereas people who haven't used induction have died nasty deaths. They have eaten arsenic believing it would nourish them, and jumped off tall buildings believing they would fly. Therefore it will clearly pay you to reason inductively.' But this wouldn't convince the doubter. For to argue that induction is trustworthy because it has worked well up to now is to reason inductively. Such an argument would carry no weight with someone who doesn't *already* trust induction. That is Hume's fundamental point.

This intriguing argument has exerted a powerful influence on the philosophy of science. (Popper's attempt to show that science need only use deduction was motivated by his belief that Hume had shown the unjustifiability of induction.) The influence of Hume's argument is not hard to understand. For normally we think of science as the very paradigm of rational enquiry. We place great faith in what scientists tell us about the world. But science relies on induction, and Hume's argument seems to show that induction cannot be rationally justified. If Hume is right, the foundations on which science is built do not look as solid as we might have hoped. This puzzling state of affairs is known as *Hume's problem of induction*.

Philosophers have responded to Hume's problem in literally dozens of ways; this is still an active area of research today. One response says that to seek a 'justification of induction', or to bemoan the lack of one, is ultimately incoherent. Peter Strawson, an Oxford philosopher from the 1950s, defended this view with the following analogy. If someone worried whether a particular action was legal, they could consult the lawbooks and see what they say. But suppose someone worried about whether the law itself was legal. This is an odd worry indeed. For the law is the standard against which the legality of other things is judged, and it makes little sense to enquire whether the standard itself is legal. The same applies to induction, Strawson argued. Induction is one of the standards we use to decide whether someone's beliefs about the world are justified. So it makes little sense to ask whether induction itself is justified.

Has Strawson really succeeded in defusing Hume's problem? Some philosophers say yes, others say no. But most agree that it is very hard to see how there *could* be a satisfactory justification of induction. (Frank Ramsey, a famous Cambridge philosopher, wrote in 1919 that to ask for a justification of induction was 'to cry for the moon'.) Whether this is something that should worry us, or shake our faith in science, is a difficult question that you should ponder for yourself.

Inference to the best explanation

The inductive inferences we've examined so far have all had essentially the same structure. In each case, the premise has had the form 'all examined *Fs* have been *G*', and the conclusion the form 'other *Fs* are also *G*'. In short, these inferences take us from examined to unexamined instances of a given kind.

Such inferences are widely used in everyday life and in science, as we have seen. However, there is another common type of non-deductive inference which doesn't fit this simple pattern. Consider the following example:

> The cheese in the larder has disappeared, apart from a few crumbs.
> Scratching noises were heard coming from the larder last night.
> ___
> Therefore, the cheese was eaten by a mouse.

It is obvious that this inference is non-deductive: the premises do not entail the conclusion. For the cheese could have been stolen by the maid, who cleverly left a few crumbs to make it look like the handiwork of a mouse; and the scratching noises could have been caused by the boiler overheating. Nonetheless, the inference is clearly a reasonable one. For the hypothesis that a mouse ate the cheese seems to provide a *better explanation* of the data than the 'maid and boiler' hypothesis. After all, maids do not normally steal cheese, and modern boilers rarely overheat. Whereas mice do eat cheese when they get the chance, and do make scratching sounds. So although we cannot be certain that the mouse hypothesis is true, on balance it looks plausible.

Reasoning of this sort is known as 'inference to the best explanation', or IBE for short. Certain terminological confusions surround the relation between IBE and induction. Some philosophers describe IBE as a *type* of inductive inference; in effect, they use 'inductive inference' to mean 'any inference which is not deductive'. Others *contrast* IBE with induction, as we have done. On this way of cutting the pie, 'induction' is reserved for inferences from examined to unexamined instances of a given kind; IBE and induction are then two different types of non-deductive inference. Nothing hangs on which choice of terminology we favour, so long as we stick to it consistently.

Scientists frequently use IBE. For example, Darwin argued for his theory of evolution by calling attention to various facts about the living world which are hard to explain if we assume that current species have been separately created, but which make perfect sense if current species have descended from common ancestors, as his theory held. For example, there are close anatomical similarities between the legs of horses and zebras. How do we explain this, if God created horses and zebras separately? Presumably he could have made their legs as different as he pleased. But if horses and zebras have descended from a common ancestor, this provides an obvious explanation of their anatomical similarity. Darwin argued that the ability of his theory to explain such facts constituted strong evidence for its truth. 'It can hardly be supposed', he wrote, 'that a false theory would explain, in so satisfactory a manner as does the theory of natural selection, the several large classes of fact above specified.'

Another example of IBE is Einstein's famous work on Brownian motion—the zig-zag motion of microscopic particles suspended in a liquid or gas. A number of attempted explanations of Brownian motion were advanced in the 19th century. One theory attributed the motion to electrical attraction between particles, another to agitation from external surroundings, and another to convection currents in the fluid. The correct explanation is based on the kinetic theory of matter, which says that liquids and gases are made up of atoms or molecules in motion. The suspended particles collide with the surrounding molecules, causing their erratic movements. This theory was proposed in the late 19th century but not widely accepted, not least because many scientists didn't believe that atoms and molecules were real entities. But in 1905, Einstein provided an ingenious mathematical treatment of Brownian motion, making a number of predictions that were later confirmed experimentally. After Einstein's work, the kinetic theory was quickly agreed to provide a better explanation of Brownian motion than the alternatives, and scepticism about the existence of atoms and molecules subsided.

The basic idea behind IBE—reasoning from one's data to a theory or hypothesis that explains the data—is straightforward. But how do we decide which of the competing hypotheses provides the 'best explanation' of the data? What criteria determine this? One popular answer is that a good explanation should be simple, or parsimonious. Consider again the cheese-in-the-larder example. There are two pieces of data that need explaining: the missing cheese and the scratching noises. The mouse hypothesis postulates just one cause—a mouse—to explain both pieces of data. But the maid-and-boiler hypothesis must postulate two causes—a dishonest maid and an overheating boiler—to explain the same data. So the mouse hypothesis is more parsimonious, hence better. The Darwin example is similar. Darwin's theory could explain a diverse range of facts about the living world, not just anatomical similarities between species. Each of these facts could in principle be explained in other ways, but the theory of evolution explained all the facts in one go—that is what made it the best explanation of the data.

The idea that simplicity or parsimony is the mark of a good explanation is quite appealing, and helps flesh out the abstract idea of IBE. But if scientists use simplicity as a guide to inference, this raises a deep question. Do we have reason to think that the universe is simple rather than complex? Preferring a theory which explains the data in terms of the fewest number of causes seems sensible. But are there any objective grounds for thinking that such a theory is more likely to be true than a less simple rival? Or is simplicity something that scientists value because it makes their theories easier to formulate and to understand? Philosophers of science do not agree on the answer to this difficult question.

Causal inference

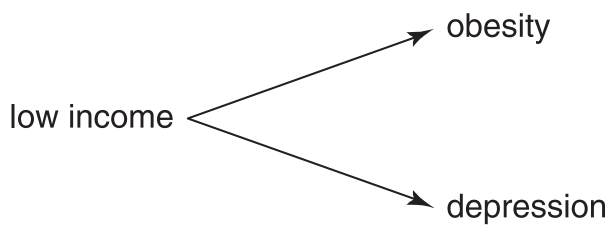
A key goal of science is to discover the causes of natural phenomena. Often this quest is successful. For example, climate change scientists know that burning fossil fuels causes global warming; chemists know that heating a liquid causes it to become a gas; and epidemiologists know that the MMR vaccine does not cause autism. Since causal connections are not directly observable (as David Hume famously argued), scientific knowledge of this sort must be the result of inference. But how exactly does causal inference work?

It is helpful to distinguish two cases: inferring the cause of a particular event versus inferring a general causal principle. To illustrate the distinction, consider the contrast between 'a meteorite strike caused the extinction of the dinosaurs' and 'smoking causes lung cancer'. The former is a singular statement about the cause of a particular historical event, the latter a general statement about the cause of a certain *sort* of event (getting lung cancer). In both cases a process of inference has led scientists to believe the statements in question, but the inferences work in somewhat different ways. Here we focus on inferences of the second sort, i.e. to general causal principles.

Suppose a medical researcher wishes to test the hypothesis that obesity causes depression. How should they proceed? A natural first step is to see whether the two attributes are correlated. To assess this, they could examine a large sample of obese people, and see whether the incidence of depression is higher in this group than in the general population. If it is, then unless there is some reason to think the sample unrepresentative, it is reasonable to infer (by ordinary induction) that obesity and depression are correlated in the overall population.

Would such a correlation show that obesity causes depression? Not necessarily. First-year science students are routinely taught that correlation does not imply causation, and with good reason. For there are other possible explanations of the correlation. The direction of causation could be the other way round, i.e. being depressed might cause people to eat more, hence to become obese. Or there might be no causal influence of obesity on depression nor vice versa, but the two conditions are joint effects of a common cause. For example, perhaps low income raises the chance of obesity and also raises the chance of depression via a separate causal pathway (see Figure 3). If so, we would expect obesity and depression to be correlated in the population. This 'common cause' scenario is a major reason why causation cannot always be reliably inferred from correlational data.

How could we test the hypothesis that low income causes both obesity and depression? The obvious thing to do is to find a sample of individuals *all with the same income level*, and examine



3. Causal graph depicting the hypothesis that low income is a common cause of both obesity and depression.

whether obesity and depression are correlated within the sample. If we do this for a number of different income levels, and find that within each income-homogeneous sample the correlation disappears, this is strong evidence in favour of the common cause hypothesis. For it shows that once income is taken into account, obesity is no longer associated with depression. Conversely, if a strong obesity–depression correlation exists even among individuals with the same income level, this is evidence against the common cause hypothesis. In statistical jargon, this procedure is known as 'controlling for' the variable income.

The underlying logic here is similar to that of the controlled experiment, a mainstay of modern science. Suppose an entomologist wishes to test the hypothesis that rearing insect larvae at higher temperatures leads to reduced adult body size. To test this, the entomologist gets a large number of insect larvae, rears some at a cool and others at a warm temperature, then measures the size of the resulting adults. For this to be an effective test of the causal hypothesis, it is important that all factors other than temperature be held constant between the two groups, so far as possible. For example, the larvae should all be from the same species, the same sex, and be fed the same food. So the entomologist must design their experiment carefully, controlling for all variables that could potentially affect adult body size. Only then can a difference in adult body size between the two groups safely be attributed to the temperature difference.

It is sometimes argued that controlled experiments are the only reliable way of making causal inferences in science. Proponents of this view argue that purely observational data, without any experimental intervention, cannot give us knowledge of causality. However this is a controversial thesis. For while controlled experimentation is certainly a good way of probing nature's secrets, the technique of statistical control can often accomplish something quite similar. In recent years, statisticians and computer scientists have developed powerful techniques for making causal inferences from observational data. Whether there is a fundamental methodological difference between experimental and observational data, vis-à-vis the reliability of the causal inferences that can be drawn from them, is a matter of continuing debate.

In modern biomedical science, a particular sort of controlled experiment is often given particular prominence. This is the *randomized controlled trial* (RCT), originally devised by R. A. Fisher in the 1930s, and often used to test the effectiveness of a new drug. In a typical RCT, patients with a particular medical condition, e.g. severe migraine, are divided into two groups. Those in the treatment group receive the drug, while those in the control group do not. The researchers then compare the two groups on the outcome of interest, e.g. relief of migraine symptoms. If those in the treatment group do significantly better than the control group, this is presumptive evidence that the drug works. The key feature of an RCT is that the initial division of the patients into two groups must be done at *random*. Fisher and his modern followers argue that this is necessary to sustain a valid causal inference.

Why is randomization so important? Because it helps to eliminate the effect of confounding factors on the outcome of interest. Typically the outcome will be affected by many factors, e.g. age, diet, and exercise. Unless all of these factors are known, the researcher cannot explicitly control for them. However by randomly allocating patients to the treatment and control groups, this problem can be largely circumvented. Even if factors other than the drug do affect the outcome, randomization ensures that any such factors are unlikely to be over-represented in either the treatment or the control group. So if there is a significant difference in outcome between treatment and control groups, this is very likely due to the drug. Of course this does not strictly *prove* that the drug was causally responsible, but it constitutes strong evidence.

In medicine, the RCT is usually regarded as the ‘gold standard’ for assessing causality. Indeed proponents of the movement known as ‘evidence-based medicine’ often argue that *only* an RCT can tell us when a particular treatment is causally effective. However this position is arguably too strong (and the appropriation of the word ‘evidence’ to refer only to RCTs is misleading). In many areas of science, RCTs are not feasible, for either practical or ethical reasons, and yet causal inferences are routinely made. Furthermore, much of the causal knowledge we have in everyday life we gained without RCTs. Young children know that putting their hand in the fire causes a painful burning sensation; no randomized trial was needed to establish this. While RCTs are certainly important, and should be done when feasible, it is not true that they are the only way of discovering causality.

Probability and scientific inference

Given that inductive reasoning cannot give us certainty, it is natural to hope that the concept of probability will help us understand how it works. Even if a scientist’s evidence does not prove that their hypothesis is true, surely it can render it highly probable? Before exploring this idea we need to attend briefly to the concept of probability itself.

Probability has both an objective and a subjective guise. In its objective guise, probability refers to how often things in the world happen, or tend to happen. For example, if you are told that the probability of an Englishwoman living to age 90 is one in ten, you would understand this as meaning that one-tenth of all Englishwomen attain that age. Similarly, a natural understanding of the statement ‘the probability that the coin will land heads is a half’ is that in a long sequence of coin flips, the proportion of heads would be very close to a half. Understood this way, statements about probability are objectively true or false, independently of what anyone believes.

In its subjective guise, probability is a measure of rational degree of belief. Suppose a scientist tells you that the probability of finding life on Mars is extremely low. Does this mean that life is found on only a small proportion of all the celestial bodies? Surely not. For one thing, no one knows how many celestial bodies there are, nor how many of them contain life. So a different notion of probability is at work here. Now since there either is life on Mars or there isn't, talk of probability in this context must presumably reflect our ignorance of the state of the world, rather than describing an objective feature of the world itself. So it is natural to take the scientist's statement to mean that in the light of all the evidence, the rational degree of belief to have in the hypothesis that there is life on Mars is very low.

The idea that the rational degree of belief to have in a scientific hypothesis, given the evidence, may be viewed as a type of probability suggests a natural picture of how scientific inference works. Suppose a scientist is considering a particular hypothesis H. In the light of the evidence to date, the scientist has a certain degree of belief in H, denoted $P(H)$, which is a number between zero and one. (Another name for $P(H)$ is the scientist's 'credence' in H.) Some new evidence then comes to light, e.g. from experiment or observation. In the light of this new evidence, the scientist updates their credence in H to $P_{new}(H)$. If the new evidence supports the theory, then $P_{new}(H)$ will be greater than $P(H)$, i.e. the scientist will have become more confident that H is true.

A toy example will help flesh this out. Suppose a playing card has been drawn from a well-shuffled pack and is concealed from your view. Let H be the hypothesis that the card is the queen of hearts. What is the value of $P(H)$, i.e. your initial rational credence in H? Presumably it is 1/52. For there are fifty-two cards in the pack and they are all equally likely to be chosen. Suppose you then learn that the chosen card is definitely a heart. Call this piece of information e. In the light of e, what is the value of $P_{new}(H)$, i.e. your updated credence in H given the new evidence? Clearly, $P_{new}(H)$ should equal 1/13—for there are thirteen hearts in the pack and you know that the concealed card is one of them. So learning e has increased your credence in H from 1/52 to 1/13.

This is all fairly obvious, but what is the general rule for updating your credence in the light of new information? The answer is called 'conditionalization'. To grasp this rule we need the concept of conditional probability. In the card example, $P(H)$ is your initial credence in hypothesis H. Your initial credence in H conditional on the assumption that e is true is denoted $P(H/e)$. (Read this as 'the probability of H given e'.) What is the value of $P(H/e)$? The answer is 1/13. For on the assumption that e is true, i.e. that the card drawn is a heart, your credence in the hypothesis H equals 1/13. When you learn that e *is* actually true, your new credence in H, i.e. $P_{new}(H)$, should then be set equal to your initial credence in H conditional on e, according to the rule of conditionalization.

Rule of conditionalization

> Upon learning evidence e, $P_{new}(H)$ should equal $P(H/e)$.

To better understand the rule of conditionalization, note that the conditional probability $P(H/e)$ is by definition equal to the ratio $P(H \text{ and } e)/P(e)$. In the card example, $P(H \text{ and } e)$ denotes your initial credence that both H and e are true. But since in this case H logically entails e—for if the card is the queen of hearts then it must be a heart—it follows that $P(H \text{ and } e)$ is simply equal to $P(H)$, i.e. 1/52. What about $P(e)$? This is your initial credence that the chosen card is a heart. Since exactly one quarter of the cards in the deck are hearts, and you regard all the cards as equally likely to be the chosen one, it follows that $P(e)$ is ¼. Applying the definition of $P(H/e)$, this tells us that $P(H/e)$ equals 1/52 divided by ¼, which is 1/13—the same answer as we computed previously.

The rule of conditionalization may sound complicated, but like many logical rules we often obey it without thinking. In the card example, it is intuitively obvious that learning *e* should increase your rational credence in *H* from 1/52 to 1/13, and in practice this is exactly what most people would do. In doing so, they are implicitly obeying the rule of conditionalization even if they have never heard of it. In addition to its implicit uses, the conditionalization rule is often used explicitly by scientists, for example in certain sorts of statistical reasoning. The branch of statistics known as Bayesian statistics makes extensive use of updating by conditionalization. (The name 'Bayesian' refers to the 17th-century English clergyman Thomas Bayes, an early pioneer of probability theory, who discovered the conditionalization rule.)

Some philosophers of science wish to use updating by conditionalization as a general model for scientific inference, applicable even to inferences that are not explicitly probabilistic. The idea is that any rational scientist can be thought of as having an initial credence in their theory or hypothesis, which they then update in the light of new evidence by following the rule of conditionalization. Even if the scientist's conscious reasoning process looks nothing like this, it is a useful idealization according to these philosophers.

This 'Bayesian' view of scientific inference is quite attractive, as it sheds light on certain aspects of the scientific method. Consider the fact that when a scientific theory makes a testable prediction that turns out to be true, this is usually taken as evidence in favour of the theory. In Chapter 1 we had the example of Einstein's theory of general relativity predicting that starlight would be deflected by the sun's gravitational field; when this prediction was confirmed it increased scientists' confidence in Einstein's theory. But why should a successful prediction enhance a scientist's confidence in a theory, given that there will always be other possible explanations that can't be ruled out? Is this simply a brute fact about how scientists reason, or does it have a deeper explanation?

Bayesians argue that it does indeed have a deeper explanation. Suppose that a theory T entails a testable statement e. The scientist initially has credence $P(T)$ that T is true and $P(e)$ that e is true. We assume that both $P(T)$ and $P(e)$ take non-extreme values, i.e. are not zero or one. Suppose the scientist then learns that e is definitely true. If they follow the rule of conditionalization, their new credence in theory T, i.e. $P_{new}(T)$, must then be greater than $P(T)$ as a matter of logic. In other words, upon learning that their theory has made a true prediction, a scientist will necessarily increase their confidence in the theory so long as they obey the conditionalization rule. So the fact that successful predictions typically lead scientists to become more confident of their theories has a neat explanation, on the Bayesian view of scientific inference.

However the Bayesian view has its limitations. Much interesting scientific inference involves inventing theories or hypotheses that have never been thought of before. The great scientific advances made by Copernicus, Newton, and Darwin were all of this sort. Each of these scientists came up with a new theory which their predecessors had never entertained. The reasoning that led them to these theories cannot plausibly be regarded as Bayesian. For conditionalization describes how a scientist's rational credence in a theory should change when they get new evidence; this presumes that the theory has already been thought of. So scientific inferences that go from data to completely new theory cannot be understood in terms of conditionalization.

Another limitation of the Bayesian view concerns the source of the initial credences, prior to updating on the new evidence. In the card example, your initial rational credence that the chosen card was the queen of hearts was easy to determine, because there are fifty-two cards in a deck each with an equal chance of being chosen. But many scientific hypotheses are not like this. Consider the hypothesis that global warming will exceed four degrees by the year 2100. What should a scientist's initial credence in this hypothesis, before getting any relevant evidence, be? There is no obvious answer to this question. Some Bayesian philosophers of science reply that initial credences are purely subjective, i.e. they simply represent a scientist's 'best guess' about the hypothesis, so any initial credence is as good as any other. On this version of the Bayesian view, there is an objectively rational way for a scientist to *change* their credences when they get new evidence, i.e. conditionalization, but no objective constraint on what their initial credences should be.

This intrusion of a subjective dimension is regarded as unwelcome by many philosophers, leading them to conclude that the Bayesian view cannot be the whole story about scientific inference. Also, it shows that there cannot be a Bayesian 'solution' to Hume's problem of induction. The idea that we can somehow escape Hume's problem by invoking probability is an old one. Even if the sun's having risen every day in the past doesn't prove that it will rise tomorrow, surely it makes it highly probable? Whether this response to Hume ultimately works is a complex matter, but we can say the following. If the only objective constraints concern how we should change our credences, but what our initial credences should be is entirely subjective, then individuals with very bizarre opinions about the world will count as perfectly rational. So a probabilistic escape from Hume's problem will not fall out of the Bayesian view of scientific inference.

Further reading

Chapter 1: What is science?

A good discussion of the scientific revolution is Steven Shapin, *The Scientific Revolution* (University of Chicago Press, 1998). Detailed treatment of topics in the history of science can be found in J. L. Heilbron (ed.), *The Oxford Companion to the History of Modern Science* (Oxford University Press, 2003). There are many good introductions to philosophy of science, including Alexander Rosenberg, *The Philosophy of Science* (Routledge, 2011) and Peter Godfrey-Smith, *Theory and Reality* (University of Chicago Press, 2003). An excellent collection of papers on general philosophy of science, with extensive commentaries by the editors, is Martin Curd, J. A. Cover, and Christopher Pincock (eds), *Philosophy of Science: The Central Issues* (W. W. Norton, 2012). Popper's attempt to demarcate science from pseudo-science can be found in his *Conjectures and Refutations* (Routledge, 1963). A good discussion of Popper's demarcation criterion is in Donald Gillies, *Philosophy of Science in the 20th Century* (Blackwell, 1993). A good introduction to Popper's philosophy is Stephen Thornton's article 'Karl Popper', in Edward N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy*, <http://plato.stanford.edu/archives/sum2014/entries/popper/>.

Chapter 2: Scientific inference

A clear discussion of induction and scientific inference is Wesley Salmon, *The Foundations of Scientific Inference* (University of Pittsburgh Press, 1967). David Hume's reflections on induction can be found in Book IV, Section 4 of his *Enquiry Concerning Human Understanding*, ed. L. A. Selby-Bigge (Clarendon Press, 1966). A detailed treatment of inference to the best explanation is Peter Lipton, *Inference to the Best Explanation* (Routledge, 2004). The literature on causal inference spans philosophy, statistics, and computer science. An ambitious work on this topic is Peter Spirtes, Clark Glymour, and Richard Scheines, *Causation, Prediction and Search* (MIT Press, 2001). On randomized controlled trials, see John Worrall, 'Why there is no cause to randomize', *British Journal for the Philosophy of Science* 58 (2007), 451–88, and Nancy Cartwright, 'What are randomized controlled trials good for?', *Philosophical Studies* 147 (2010), 59–70. A good treatment of probability and induction is Ian Hacking, *An Introduction to Probability and Inductive Logic* (Cambridge University Press, 2001). The Bayesian approach to scientific inference is expounded by Colin Howson and Peter Urbach, *Scientific Reasoning: The Bayesian Approach* (Open Court, 2006).

Chapter 3: Explanation in science

Hempel's original presentation of the covering law model is in *Aspects of Scientific Explanation* (Free Press, 1965). A useful account of the debate instigated by Hempel's work is Wesley Salmon, *Four Decades of Scientific Explanation* (University of Minnesota Press, 1989). A detailed recent treatment of scientific explanation, with an extensive bibliography, is James Woodward's article 'Scientific explanation', in Edward N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy* (Winter 2014 edition), <http://plato.stanford.edu/archives/win2014/entries/scientific-explanation/>. The suggestion that consciousness can never be explained scientifically is found in Colin McGinn, *Problems of Consciousness* (Blackwell, 1991). The idea that multiple realization accounts for the autonomy of the higher-level sciences is developed by Jerry Fodor, 'Special Sciences', *Synthese* 28 (1974), 97–115. Further discussion of reductionism is found in section 8 of Martin Curd, J. A. Cover, and Christopher Pincock (eds), *Philosophy of Science* (W. W. Norton, 2012).

Chapter 4: Realism and anti-realism

A detailed analysis of scientific realism, with an extensive bibliography, is Anjan Chakravartty's article 'Scientific realism', in Edward N. Zalta