10 Induction and Probability

Overview

Suppose we settle the dispute between realism and instrumentalism. The problem still remains of how exactly observation and evidence, the collection of data, etc., actually enable us to choose among scientific theories. On the one hand, that they do so has been taken for granted across several centuries of science and its philosophy. On the other hand, no one has fully explained how they do so, and in this century the challenges facing the explanation of exactly how evidence controls theory have increased.

A brief review of the history of British empiricism sets the agenda for an account of how science produces knowledge justified by experience, and introduces the problem of induction raised by David Hume in the eighteenth century. If we cannot solve the problem of induction, we may be able to show that it is a pseudo-problem. Even if we can't do that, scientists are not about to down tools and wait for a resolution of this matter. What is more, they may insist that they know how to proceed inductively without any help from philosophers. All we really need, many scientists would insist, is a theorem of the probability calculus derived in the eighteenth century by Thomas Bayes, a contemporary of David Hume. Some philosophers will concur with this judgment. So we need to understand this theorem and the interpretational issues that are raised by its use in experimental and observational reasoning.

The Problem of Induction

As we noted in Chapter 7, the scientific revolution began in central Europe with Copernicus, Brahe and Kepler, shifted to Galileo's Italy, moved to Descartes' France and came to Newton in Cambridge, England. The scientific revolution was also a philosophical revolution, and for reasons we noted. In the seventeenth century science was "natural philosophy," and figures that history would consign exclusively to one or the other of these two fields—philosophy or science—contributed to both. Thus Newton wrote a good deal of philosophy of science, and Descartes made contributions to physics. But it was the British empiricists who made a self-conscious attempt to examine whether the theory of knowledge espoused by these scientists would vindicate the methods which Newton, Boyle, Harvey, and other experimental scientists employed to expand the frontiers of human knowledge so vastly in their time.

Over a period from the late seventeenth century to the late eighteenth century, John Locke, George Berkeley, and David Hume sought to specify the nature, extent and justification of knowledge as founded on sensory experience and to consider whether it would certify the scientific discoveries of their time as knowledge and insulate them against skepticism. Their results were mixed, as Kant was eager to point out. But nothing would shake their confidence or that of most scientists, in empiricism as the right epistemology.

Locke sought to develop empiricism about knowledge, famously holding against rationalists like Descartes, that there are no innate ideas. "Nothing is in the mind that was not first in the senses." But Locke was resolutely a scientific realist about the theoretical entities which seventeenth-century science was uncovering. As noted in Chapter 2, he embraced the view that matter was composed of indiscernible atoms, or "corpuscles," and distinguished between material substance and its properties ("primary qualities") on the one hand, and the sensory qualities of color, texture, smell or taste (the so-called secondary qualities), which substances cause in us. The real properties of matter, according to Locke, are just the ones that Newtonian mechanics tells us it has—mass, extension in space, velocity, etc. The sensory qualities of things are ideas in our heads that the things cause. It is by reasoning back from sensory effects to physical causes that we acquire knowledge of the world, which gets systematized by science.

That Locke's realism and his empiricism inevitably gives rise to skepticism, is not something Locke recognized. It was a philosopher of the next generation, George Berkeley, who appreciated that empiricism makes doubtful our beliefs about things we do not directly observe. How could Locke lay claim to the certain knowledge of the existence of matter or its features, if he could only be aware of sensory qualities, which by their very nature, exist only in the mind? We cannot compare sensory features like color or texture to their causes to see whether these causes are colorless or not, for we have no access to these things, so we cannot compare them. And to the argument that we can imagine something to be colorless, but we cannot imagine a material object to lack extension or mass, Berkeley retorted that sensory properties and non-sensory ones are on a par in this respect: try to imagine something without color. If you think of it as transparent, then you are adding in the background color and that's cheating. Similarly for the other allegedly subjective qualities that things cause us to experience.

In Berkeley's view, without empiricism we cannot make sense of the meaningfulness of language. Berkeley pretty much adopted the theory of language as naming sensory qualities that was sketched in Chapter 8. Given the thesis that words name sensory ideas, realism—the thesis that science discovers truths about things we cannot have sensory experience of, becomes false, for the words that name these things must be meaningless. In place of realism Berkeley advocated a strong form of instrumentalism and took great pains to construct an interpretation of seventeenth- and eighteenth-century science, including Newtonian mechanics, as a body of heuristic devices, calculating rules, and convenient fictions that we employ to organize our experiences. Doing this, Berkeley thought, saves science from skepticism. It did not occur to Berkeley that another alternative to the combination of empiricism and instrumentalism is rationalism and realism. And the reason is that by the eighteenth century, the role of experiment in science had been so securely established that no alternative to empiricism seemed remotely plausible as an epistemology for science. Even rationalism, as we noted in Chapter 1, argues only that some scientific knowledge has a non-empirical justification.

It was David Hume's intention to apply what he took to be the empirical methods of scientific inquiry to philosophy. Like Locke and Berkeley he sought to show how knowledge, and especially scientific knowledge, honors the strictures of empiricism. Unable to adopt Berkeley's radical instrumentalism, Hume sought to explain why we adopt a realistic interpretation of science and ordinary beliefs, without taking sides between realism and instrumentalism. But Hume's pursuit of the program of empiricism led him to face a problem different from that raised by the conflict of realism and empiricism. This is the problem of induction: Given our current sensory experience, how can we justify inferences from them and from our records of the past, to the future and to the sorts of scientific laws and theories we seek?

Hume's argument is often reconstructed as follows: there are two and only two ways to justify a conclusion: deductive argument, in which the conclusion follows logically from the premises, and inductive argument, in which the premises support the conclusion but do not guarantee it. A deductive argument is colloquially described as one in which the premises "contain" the conclusion, whereas an inductive argument is often described as one that moves from the particular to the general, as when we infer from observation of 100 white swans to the conclusion that all swans are white. Now, if we are challenged to justify the claim that inductive arguments—arguments from the particular to the general, or from the past to the future—will be reliable in the future, we can do so only by employing a deductive argument or an inductive argument. The trouble with any deductive argument to this conclusion is that at least one of the premises will itself require the reliability of induction. For example, consider the deductive argument below:

1. If a practice has been reliable in the past, it will be reliable in the future.
2. In the past inductive arguments have been reliable.

Therefore,

3. Inductive arguments will be reliable in the future.

This argument is deductively valid, but its first premise requires justification and the only satisfactory justification for the premise would be the reliability of induction, which is what the argument is supposed to establish. Any deductive argument for the reliability of induction will include at least one question-begging premise.

This leaves only inductive arguments to justify induction. But clearly, no inductive argument for induction will support its reliability, for such arguments too are question-begging. As we have had occasion to note before, like all such question-begging arguments, an inductive argument for the reliability of induction is like underwriting your promise to pay back a loan by promising that you will keep your promises. If your reliability as a promise keeper is what is in question, offering a second promise to assure the first one is pointless. Hume's argument has for 250 years been treated as an argument for skepticism about empirical science. For it suggests that all conclusions about scientific laws, and all prediction science makes about future events, are at bottom unwarranted, owning to their reliance on induction. And it's not just inferences from the specific to the general or from the past to the future. There are other forms of argument that are clearly inductive without taking either of these forms, including arguments by analogy, and inferences to the best explanation as employed to infer the existence of unobservable entities throughout the sciences. All ampliative forms of argument, in which the conclusions are intended to make claims that transcend those of the premises, will be inductive and open to Hume's challenge. Many ampliative inferences employ or exploit deduction. But they are inductive nevertheless. For example, hypothetico–deductive reasoning involves the deduction of observational consequences from a hypothesis and their direct testing is still inductive. If confirmed, such deductive consequences are said to confirm the hypothesis they are deduced from. The whole inference is clearly inductive: the conclusion, that the credence of the general hypothesis is strengthened by the narrower observational evidence, goes beyond the evidence.

Hume's challenge is theoretical. He noted that as a person who acts in the world, he was satisfied that inductive arguments were reasonable; what he thought the argument shows is that we have not yet found the right justification for induction, not that there is no justification for it.

Hume's problem of induction was surprisingly invisible for the first 150 years after he formulated it. The greatest empiricist epistemologist and philosopher of science of the nineteenth century, John Stuart Mill, completely failed to recognize it despite devoting much attention to induction as the core method of science. According to Mill, inferences from a relatively small number of cases to general laws were how science proceeds. Mill famously articulated several rules of experimental design that still guide scientists today in making such inferences. Contemporary double-blind controlled experiments now commonplace in medical science owe a great deal to the rules Mill set out and the arguments he gave for them.

But that the practice of inductive inference requires independent justification as a whole was something Mill did not appreciate. Mill believed with some justification that inductive inferences were grounded on a commitment to the uniformity of nature: that the future will be like the past. If we can be justified in believing this principle then at least some inductive inferences will be warranted. But what sort of argument can be advanced for the uniformity of nature? A deductive argument with a factual conclusion that the future will be like the past will have to include among its premises a factual claim at least as strong, and this will then require justification, and so on in an infinite regress. An inductive argument for the uniformity of nature will proceed along the following lines: In the recent past, its near future was like the more distant past, in the more distant past, its near future was like the even more distant past, and so on. Therefore, hereafter the future will be like the recent past, the more distant past and the very distant past. But this form of argument is itself inductive and so begs the question. We set out to establish the reliability of inductive inference and do so by an inductive inference. This has all the reliability of an attempt to assure someone that I keep my promises by promising him that I do so!

During the period in which logical positivists were confident that the principles of mathematical and symbolic logic were definitions and consequences of them, attempts were made to solve Hume's problem in a similar way. Philosophers like Rudolph Carnap and Carl G. Hempel sought to frame rules of inductive inference that could be justified, like the laws of mathematical logic, on the basis of definitions and their implications. Like the deductive–nomological model which was proposed to rationally reconstruct the concept of scientific explanation, their aim was to provide a "confirmation theory" that would formalize and explicate the notion of inductive inference and solve Hume's problem too. The strategy was to show that inductive argument turns out to be deductive argument that employs special rules that confer justification on their conclusions without guaranteeing their truth (unlike the laws of deductive logic which did so). These rules would reflect the axioms and theorems of probability theory, a set of logical truths or definitions. In order for these rules to systematize inductive inferences, the statements scientists use to describe the data or evidence to which the rules are applied had to be given a rigid logical structure and a wholly observational vocabulary. This could not reasonably accommodate the actual patterns of scientific inference. But in addition, the entire enterprise of developing a purely formal or logical theory of probability only revealed the problem to be even more serious than Hume had recognized, as we will see in Chapter 11.

Other philosophers sought to show that the problem of induction was a pseudo-problem, a classical example of bewitchment of our understanding by language. Thus it has been repeatedly argued that to employ inductive principles in order to frame expectations about the future is just what common sense and most people mean by being reasonable. If employing inductive inference is, by definition, a necessary condition for acting in a reasonable manner, then it is senseless to demand a justification for it. Or at least it makes no more sense to ask that induction be shown to be reasonable than it makes to ask that being reasonable be shown to be reasonable. Thus, a proper understanding of what it means to be reasonable when framing beliefs about the unobserved solves the problem of induction, or rather shows it to be a pseudo-problem, one reflecting mistakes about language. What mistake? One candidate is the tendency mistakenly to apply deductive standards to induction and then to complain when they can't be met. Validity is a feature of proper deductive arguments: these arguments are always truth-preserving. Since inductive arguments are by their nature not truth-preserving (not intended or expected to be); it is easy but mistaken to describe them as invalid and then demand a justification for them. The mistake is even to apply the valid/invalid distinction to such arguments and then to demand a substitute for validity.

Few philosophers of science could take seriously this way of dismissing Hume's problem. They insisted that the callow mistake identified by those who seek to dissolve the problem of induction is not one they make. The problem of induction is very clearly that of showing inductive inferences to be generally reliable, not universally valid. And this problem can be framed in such a way as to honor the thought that being inductive is being reasonable. To ask whether being reasonable, i.e. using inductive methods, is a reliable method of getting through life is perfectly intelligible. The question of whether being reasonable is reliable is one that we all want to answer affirmatively. Hume in effect invites us to do so in a non-question-begging way.

One way to respond to Hume that recognizes this way of putting his problem was due to the logical positivist philosopher, Hans Reichenbach (he preferred the label "logical empiricist"). He sought to show that if any method of predicting the future works, then induction must work. Suppose we wish to establish whether the Oracle at Delphi is an accurate predictive device. The only way to do so is to subject the Oracle to a set of tests: ask for a series of predictions and determine whether they are verified. If they are, the Oracle can be accepted as an accurate predictor. If not, then the future accuracy of the Oracle is not to be relied upon. But notice that the form of this argument is inductive: if any method works (in the past), only induction can tell us that it does (in the future). Whence we secure the justification of induction. This argument faces two difficulties. First, at most it proves that if any method works, induction works. But this is a far cry from the conclusion we want: that there is any method that does in fact work. Second, the argument will not sway the devotee of the Oracle. Oracle-believers will have no reason to accept our argument. They will ask the Oracle whether induction works, and will accept its pronouncement. No attempt to convince Oracle-believers that induction supports either their method of telling the future or any other can carry any weight with them. The argument that if any method works, induction works, is question-begging too.

Statistics and Probability to the Rescue?

At some point the problems of induction will lead some scientists and philosophers to lose patience with the philosopher of science. Why worry about justifying induction? Why not get on with the serious but perhaps more soluble problem of empirical confirmation? We may grant the fallibility of science, the impossibility of establishing the truth or falsity of scientific laws once and for all. Yet we may still explain how observation, data collection and experiment test scientific theory by turning to statistical theory and the notion of probability.

It turns out that doing so is not as simple a matter as it seems. To begin with the notions of probability and of empirical or inductive evidence don't really line up together as neatly as we might wish.

To begin with there is the problem of whether the fact that some data raise the probability of a hypothesis makes that data evidence for it at all. This may sound like a question trivially easy to answer, but it isn't. Define p (h, b) as the probability of hypothesis h, given background information b, and p (h, e and b) as the probability of h given the background information b, and some experimental observations e. Suppose we adopt the principle that

e is positive evidence for hypothesis h if and only if p(h, e and b) > p(h, b)

So, when data increases the probability of a hypothesis, it constitutes favorable evidence for it.

So, in this case, e is "new" data that counts as evidence for h if it raises the probability of h (given the background information required to test h). For example, the probability that the butler did it, h, given that the gun found at the body was not his, b and the new evidence that the gun carried his fingerprints, is higher than the hypothesis that the butler did it, given the gun found at the body, and no evidence about fingerprints. It is the fingerprints that raise the probability of h. That's why the prints are evidence that the butler did it.

It is easy to construct counterexamples to this definition of positive evidence that show that increasing probability is by itself neither necessary nor sufficient for some statement about observations to confirm a hypothesis. Here are two:

This book's publication increases the probability that it will be turned into a blockbuster film starring Keira Knightley. After all were it never to have been published the chances of its being made into a film would be even smaller than they are. But surely, the actual publication of this book is not positive evidence for the hypothesis that this book will be turned into a blockbuster film starring Keira Knightley. It is certainly not clear that some fact which just raises the probability of a hypothesis thereby constitutes positive evidence for it. A similar conclusion can be derived from the following counterexample, which invokes lotteries, a useful notion when exploring issues about probability. Consider a fair lottery with 1,000 tickets, 10 of which are purchased by Andy and 1 of which is purchased by Betty. h is the hypothesis that Betty wins the lottery. e is the observation that all tickets except those of Andy and Betty are destroyed before the drawing. e certainly increases the probability of h from 0.001 to 0.1. But it is not clear that e is positive evidence that h is true. In fact, it seems more reasonable to say that e is positive evidence that h is untrue, that Andy will win. For the probability that he wins has gone from 0.01 to 0.9. Another lottery case suggests that raising probability is not necessary for being positive evidence; indeed a piece of positive evidence may lower the probability of the hypothesis it confirms. Suppose in our lottery Andy has purchased 999 tickets out of 1,000 sold on Monday. Suppose e is the evidence that by Tuesday 1,001 tickets have been sold of which Andy purchased 999. This e lowers the probability that Andy will win the lottery from 0.999 to 0.998 ... But surely e is still evidence that Andy will win after all.

One way to deal with these two counterexamples is simply to require that e is positive evidence for h if e makes h's probability high, say above 0.5. Then in the first case, since the evidence doesn't raise the probability of Betty's winning anywhere near 0.5, and in the first case the evidence does not lower the probability of Andy's winning much below 0.999, these cases don't undermine the definition of positive evidence when so revised. But of course, it is easy to construct a counterexample to this new definition of positive evidence as evidence that makes the hypothesis highly probable. Here is a famous case: h is the hypotheses that Andy is not pregnant, while e is the statement that Andy eats Weetabix breakfast cereal. Since the probably of h is extremely high, P(h, e)—the probability of h, given e—is also extremely high. Yet e is certainly no evidence for h. Of course we have neglected the background information, b, built into the definition. Surely if we add the background information that no man has ever become pregnant, then P(h, e & b)—the probability of h, given e and b—will be the same as P(h, e), and thus dispose of the counterexample. But if b is the statement that no man has ever become pregnant, and e is the statement that Andy ate Weetabix, and h is the statement that Andy is not pregnant, then p(h, e & b) will be very high, indeed about as close to 1 as a probability can get. So, even though e is not by itself positive evidence for h, e plus b is, just because b is positive evidence for h. We cannot exclude e as positive evidence, when e plus b is evidence, just because it is a conjunct that by itself has no impact on the probability of h, because sometimes positive evidence only does raise the probability of a hypothesis when it is combined with other data. Of course we want to say that in this case, e could be eliminated without reducing the probability of h, e is probabilistically irrelevant and that's why it is not positive evidence. But providing a litmus test for probabilistic irrelevance is no easy task. It may be as difficult as defining positive instance. In any case, we have an introduction here to the difficulties of expounding the notion of evidence in terms of the concept of probability.

Philosophers of science who insist that probability theory suffices to enable us to understand how data test hypotheses will respond to these problems that they reflect the mis-fit between probability and our common sense notions of evidence. Our ordinary concepts are qualitative, imprecise, and not the result of a careful study of their implications. Probability is a quantitative mathematical notion with secure logical foundations. That enables us to make distinctions ordinary notions cannot draw, and to explain these distinctions. Recall the logical empiricists who sought rational reconstructions or explications of concepts like explanation that provide necessary and sufficient conditions in place of the imprecision and vagueness of ordinary language. Likewise, many contemporary students of the problem of confirmation seek a more precise substitute for the ordinary notion of evidence in the quantifiable notion of probability; for them counterexamples such as the ones adduced above simply reflect the fact that the two concepts are not identical. They are no reason not to substitute "probability" for "evidence" in our inquiry about how data test theory. Some of these philosophers go further and argue that there is no such thing as evidence confirming or disconfirming a hypothesis by itself. Hypothesis testing in science is always a comparative affair: it only makes sense to say hypothesis h1 is more or less well confirmed by the evidence than is hypothesis h2, not that h1 is confirmed by e in any absolute sense.

These philosophers hold that the mathematical theory of probability holds the key to understanding the confirmation of scientific theory. And this theory is extremely simple. It embodies only three very obvious assumptions:

1. Probabilities are measured in numbers from 0 to 1.
2. The probability of a necessary truth (like "4 is an even number") is 1.
3. If hypothesis h and j are incompatible, then $p(h \text{ or } j) = p(h) + p(j)$.

It's easy to illustrate these axioms with a deck of normal playing cards. The probability of any one card being drawn from a complete deck is between 0 and 1. In fact it's 1/52. The probability that a card will be red or black (the only two possibilities) is 1 (it's a certainty), and if drawing an ace of hearts is incompatible with drawing a jack of spades, then the probability of drawing one of them is 1/52 + 1/52, or 1/26, about 0.038461 ...

From these simple and straightforward assumptions (plus some definitions) the rest of the mathematical theory of probability can be derived by logical deduction alone. In particular, from these three axioms of the theory of probability, we can derive a theorem, first proved by a British theologian and amateur mathematician in the eighteenth century, Thomas Bayes, which has bulked large in contemporary discussions of confirmation. Before introducing this theorem, we need to define one more notion, the conditional probability of any one statement, assuming the truth of another statement. The conditional probability of a hypothesis, h, on a description of data, e, written p(h/e) is defined as the ratio of the probability of the truth of both h and e to the probability of the truth of e alone:

$$p(h/e) = df\, p(h \text{ and } e)$$

$$p(e)$$

Roughly "the conditional probability of h on e" measures the proportion of the probability that e is true which "contains" the probability that h is also true. Adapting an expository idea of Martin Curd and Jan Cover, we can illuminate this definition with a few diagrams. Suppose we are shooting darts at a board on which two overlapping circles are drawn in the shape of a Venn diagram (Figure 10.1):

If a dart lands inside circle e, what is the probability that it will also land inside circle h, i.e. the probability of landing in h, on the condition that it lands in e, the conditional probability, p(h/e)? That depends on two things: the area of overlap between circle e and circle h (the intersection e & h), relative to the area of e, and the size of e compared to the size of h. To see this, compare the two following diagrams. In Figure 10.2, e is very large compared to the size of h, so the chance that a dart thrown inside e also lands in h is

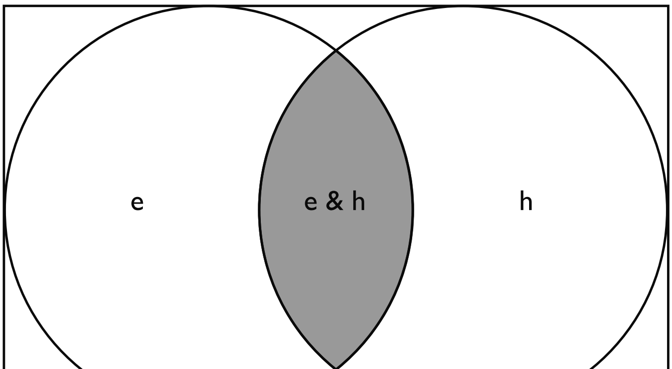

Figure 10.1 Circles e and h are the same size, and between them take up most of the rectangle, to suggest that the probability of a dart hitting one of them (and not the other) is large and about the same.

low. But it would be higher if more of h were inside e. On the other hand, the chance that a dart that lands in h also lands in e is much higher, and increases as the proportion of h inside e grows.

By contrast, consider Figure 10.3. Here e is small and h is large. In this case the chance of a dart which lands in e also landing in h is higher than in the previous case, and becomes even higher the more of e is inside h. Again,

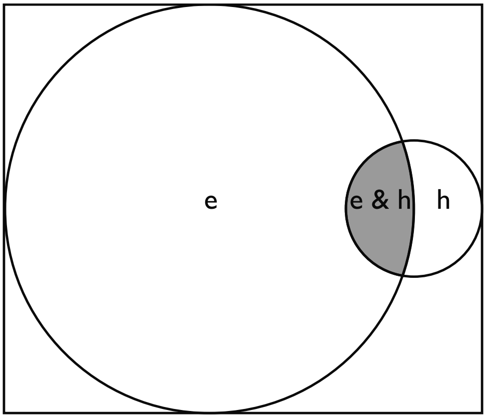

Figure 10.2 Circle e is much larger than circle h, so the probability of the dart hitting e is much higher than the probability of dart hitting h. The shaded intersection e & h is much smaller than e, and a relatively large proportion of h. Thus p(h/e) is low, and p(e/h) is much higher than p(h/e).

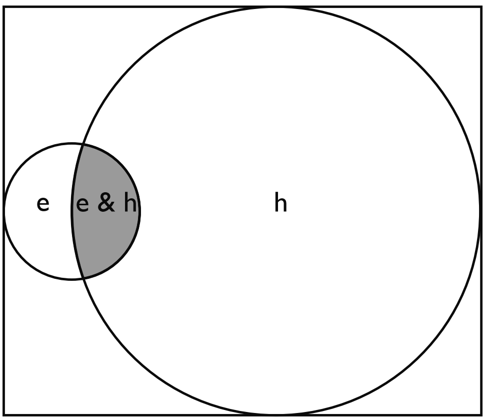

Figure 10.3 Circle h is much larger than circle e, so the probability of the dart hitting h is much higher than the probability of dart hitting e. The shaded intersection e & h is much smaller than h, and is a relatively large proportion of e. Thus p(h/e) is high, and p(e/h) is much lower than p(h/e).

the conditional probability of e on h is of course much lower, the smaller the h circle is and the less it overlaps.

The definition of conditional probability incorporates these two factors on which conditional probability depends. The numerator reflects the size of the overlap of e and h relative to the sizes of e and h, and the denominator measures that size in units of e's size.

Now if h is a hypothesis and e is a report of data, Bayes' theorem allows us to calculate the conditional probability of h on e, p(h/e). In other words, Bayes' theorem give us a mathematical formula for calculating how much more or less probable a bit of evidence, e, makes any hypothesis, h. The formula is as follows:

$$p(h/e) = p(e/h) \times p(h)$$

$$p(e)$$

Bayes' theorem tells us that once we acquire some data, e, we can calculate how the data e change the probability of h, raising it or lowering, provided we already have three other numbers:

> p(e/h)—the probability that e is true assuming that h is true (as noted above, not to be confused with p(h/e), the probability that h is true, given e, which is what we are calculating). This number reflects the degree to which our hypotheses leads us to expect the data we have gathered. If the data is just what the hypothesis predicts then of course p(e/h) is very high. If the data is nothing like what the hypothesis predicts p(e/h) is low.
>
> p(h)—the probability of the hypothesis independent of the test to which the data described by e provides. If e reports new experimental data, then p(h) is just the probability the scientist assigned to h before the experiment was conducted.
>
> p(e)—the probability that the statement describing the data is true independent of whether h is true or not. Where e is a surprising result which previous scientific theory and evidence (independent of h) does not lead us to expect, p(e) will be low.

To see how easily Bayes' theorem follows from the axioms of probability and our definition of conditional probability, return to any of the dart-board diagrams above. If we can calculate p(e/h) by comparing the relative sizes of the circles and the ratio of their intersections to their sizes, we can also calculate p(h/e) the same way. Of course the figures for each conditional probability will be different (as each of the diagrams illustrates).

By drawing e- and h-circles and intersections of them of different sizes, it is easy to see that the probability of a dart which hits the e-circle also hitting the h-circle, p(h/e) will vary directly as the ratio of the intersection of the two circles to the size of the e-circle, and inversely as the ratio of the sizes of e-circle to the size of the h-circle. And this is exactly what Bayes' theorem says: it makes p(h/e) equal to p(e/h)—the ratio of the intersection of e and h to the size of e—times the fraction p(h)/p(e) which is the ratio of the size of h to the size of e.

Two simple examples may help us see how Bayes' theorem is supposed to work: Consider how data on the observed position of Halley's comet provide a test for Newton's laws. Suppose, given prior observations, that p(e), the probability that Halley's comet will be observed in a particular location of the night sky, is 0.8. This allows for imperfections in the telescope, atmospheric irregularities, all the factors that eventually led astronomers to take many photographs of the stars and planets and to average their positions to make estimates of their expected positions in the sky. p(e/h) is also high, the expected position of Halley's comet in the night sky is very close to what the theory predicts it would be. Let's set p(e/h) at 0.95. Let's assume that prior to the acquisition of e, the new data about Halley's comet, the probability that Newton's laws are true is, say, 0.8. Thus, if Halley's comet appears where expected, p(h/e) = (0.95…) × (0.8)/(0.8) = 0.95. Thus, the evidence as described by e has raised the probability of Newton's laws from 0.8 to 0.95.

But now, suppose we acquire new data about, say, the precession of the perihelion of Mercury—that is, data that show that the elliptical orbit of Mercury around the Sun is itself swinging so that the closest point between Mercury and the Sun keeps shifting. Suppose, as was indeed the case, that the figure turns out to be much higher than Newton's laws (and the auxiliary hypotheses used to apply them) would lead us to expect, so that p(e/h) is low, say 0.3. Since Newton's laws did not lead us to expect this data, the **prior probability** of e must be low, so let's let p(e) be low, say, 0.2; and the prior probability of such unexpected data, given Newton's laws plus auxiliary hypotheses, will also be quite low, say, p (e/h) is 0.1. If p(h) for Newton's laws plus auxiliaries is 0.95, then Bayes' theorem tells us that for the new e, the precession data for Mercury, the p(h/e) = (0.1) × (0.95)/(0.2) = 0.475, a significant drop from 0.95. Naturally, recalling the earlier success of Newton's laws in uncovering the existence of Neptune and Uranus, the initial blame for the drop was placed on the auxiliary hypotheses. Bayes' theorem can even show us why. Though the numbers in our example are made up, in this case, the auxiliary assumptions were eventually vindicated, and the data about the much greater than expected precession of the perihelion of Mercury undermined Newton's theory, and (as another application of Bayes' theorem would show), increased the probability of Einstein's alternative theory of relativity.

Philosophers and many statisticians hold that the reasoning scientists use to test their hypotheses can be reconstructed as inferences in accordance with

Bayes' theorem. These theorists are called Bayesians. Some philosophers and historians of science among them seek to show that the history of acceptance and rejection of theories in science honors Bayes' theorem, thus showing that in fact, theory testing has been on firm footing all along. Other philosophers, and statistical theorists attempt to apply Bayes' theorem actually to determine the probability of scientific hypotheses when the data are hard to get, sometimes unreliable, or only indirectly relevant to the hypothesis under test. For example, they seek to determine the probabilities of various hypotheses about evolutionary events like the splitting of ancestral species from one another, by applying Bayes' theorem to data about differences in the polynucleotide sequences of the genes of currently living species.

How Much Can Bayes' Theorem Really Help?

How much understanding of the nature of empirical testing does Bayesianism really provide? Will it reconcile science's empiricist epistemology with its commitment to unobservable events and processes that explain observable ones? Will it solve Hume's problem of induction? To answer these questions, we must first understand what the probabilities are that all these p's symbolize and where they come from. We need to make sense of p(h), the probability that a certain proposition is true. There are at least two questions to be answered: First, there is the "metaphysical" question of what fact is it about the world, if any, that makes a particular probability value, p(h) for a hypothesis, h, the true or correct one? Second, there is the epistemological question of justifying our estimate of this probability value. The first question may also be understood as a question about the meaning of probability statements, and the second about how they justify inductive conclusions about general theories and future eventualities.

Long before the advent of **Bayesianism** in the philosophy of science the meaning of probability statements was already a vexed question. There are some traditional interpretations of probability we can exclude as unsuitable interpretations for the employment of Bayes' theorem. One such is the interpretation of probability as it is supposed to figure in fair games of chance like roulette or blackjack. In a fair game of roulette the chance of the ball landing in any trap is exactly 1/37 or 1/38 because there are 37 (or in Europe 38) traps into which the ball can land. Assuming it is a fair roulette wheel, the probability of the hypothesis that the ball will land on number 8 is exactly 1/37 or 1/38 and we know this *a priori*—without experience, because we know *a priori* how many possibilities there are and that each is equally probable (again, assuming the roulette wheel is fair, a bit of knowledge we could never have acquired *a priori* anyway!). Now, when it comes to hypotheses that can account for a finite body of data, there is no limit to the number of possibilities and no reason to think that each of them has the same probability. Accordingly, the probabilities of a hypothesis about, say, the number of chromosomes in a human nucleus, will not be determinable *a priori*, by counting up possibilities and dividing 1 by the number of possibilities.

Another interpretation of probabilities involves empirical observations, for example, coin flips. To establish the frequency with which a coin will come up heads, one flips it several times and divides the number of times it comes up heads by the number of times it was flipped. Will this frequency be a good estimate of the probability of heads? It will be when the number of coin flips is large, and the frequencies we calculate for finite numbers of coin flips converge on one value and remain near that value no matter how many times we continue flipping. We can call this value, if there is one, the **long-run relative frequency** of heads. And we treat it as a measure of the probability the coin comes up heads. But is the long-run relative frequency of heads identical to the probability it will come up heads? This sounds like a silly question, until you ask what the connection is between the long-run relative frequency's being, say 0.5 and the chance that the very next toss will be heads. Notice that a long-run relative frequency of 0.5 is compatible with a run of 10, or 100, or 1,000,000 heads in a row, just so long as the total number of tosses is very large, so large that a million is a small number in comparison to the total number of tosses. If this is right, the long-run relative frequency is compatible with any finite run of all heads, or all tails, and of course perfectly compatible with the coin's coming up tails on the next toss. Now, suppose we want to know what the probability is that the coin will come up heads on the next toss. If the probability that the coin will come up heads on the next toss is a property of that particular toss, it is a different thing from the long-run relative frequency of heads (which is perfectly compatible with the next 234,382 tosses all being tails). We need some principle that connects the long run to the next toss. One such principle that gets us from the long-run relative frequency to the probability of the next toss being heads is to assume that coins do in any finite run what they do in the long run. But this principle is just false. A better principle for connecting long-run relative frequencies to the probability of the next occurrence is something like this: If you know the long-run relative frequency then you know how to bet on whether the coin will land heads or tails, and if you take all bets against heads at odds greater than even money, you will win. But notice this is a conclusion about what you should do as a gambler, not a conclusion about what the coin will in fact do. We will come back to this insight.

Could long-run relative frequencies provide the probability values for a hypothesis without a track record? It is hard to see how. Compare a novel hypothesis to a shiny new penny about to be flipped. Long-run relative frequencies data provide some reason to ascribe a probability of 0.5 to the chances of heads on the new penny. Is there a track record of previous hypotheses relevant to the new one? Only if we can compare it to the right class of similar hypotheses the way we can compare new pennies to old ones. But hypotheses are not like pennies. Unlike pennies, they differ from one another in ways we cannot quantify as we would have to were we to grade them for similarity to one another. Even if we could identify the track record of truth and falsify for similar hypotheses formulated over the past history of science, we would have the problems of (a) justifying the inference from a finite actual sequence to a long-run relative frequency, and (b) justifying the inference from a long-run relative frequency to the next case, the new hypothesis. Recall, that in the case of coin flipping, the only connection appears to be that relative frequencies are our best guide to how to lay our bets about the next toss. Perhaps the kind of probability which theory testing invokes is the gambler's kind, what has come to be called "subjective probability." "Subjective" because it reflects facts about the gambler, and what the gambler believes about the past and the future, and "probability" because the bets the gambler makes should honor the axioms of probability.

It is the claim that in scientific testing, the relevant probabilities are subjective probabilities, that is, gambler's odds, which is the distinctive mark of the Bayesian. A Bayesian is someone who holds that at least two of the three probabilities we need to calculate p(h/e) are just a matter of betting odds and that within certain weak constraints, they can take on any values at all. You and I may think that the best betting odds are those that mirror our previous experience of actual frequencies or our estimate of long-run relative frequencies, but this is no part of Bayesianism. The Bayesian holds that in the long run it doesn't matter what values they start with, Bayes' theorem will lead the scientist inexorably to the (available) hypothesis best supported by the evidence. These remarkable claims demand explanation and justification.

Calculating the value of p(e/h) is a matter of giving a number to the probability that e obtains if h is true. This is usually easy to do. If h tells us to expect e, or data close to e, then p(e/h) will be very high. The problem is using Bayes' theorem also requires we calculate input values, so called "prior probabilities," p(h) and p(e). P(h) is especially problematical: after all, if h is a new theory no one has ever thought of, why should there be any particular right answer to the question of with what probability it is true? And assigning a value to p(e), the probability that our data description is correct may involve so many auxiliary assumptions, that even if there is a correct number it is hard to see how we could figure out what it is. The Bayesian asserts that these are not problems. Both values, p(h) and p(e) (and p(e/h) for that matter) are simply degrees of belief, and degrees of belief are simply a matter of what betting odds the scientist would take or decline on whether their beliefs are correct. The higher the odds one takes, the stronger the degree of belief. Here the Bayesian takes a page from economists and others who developed the theory of rational choice under uncertainty. The way to measure a degree of belief is to offer the believer wagers against the truth of his or her belief. Other things being equal, if you are rational, and you are willing to take a bet that h is true at odds of 4:1 then your degree of belief that h is true is 0.8. If you are willing to take a 5:1 then your degree of belief is just under 0.9. Probabilities are identical to degrees of belief. The other things that have to be equal for this way of measuring the strength of your beliefs are (a) that you have enough money so that you are not so averse to the risk of losing that it swamps your attraction to the prospect of winning, (b) that the degrees of belief you assign to your beliefs obey the rules of logic and the three laws of probability above. So long as your degrees of belief, a.k.a. probability assignments, honor these two assumptions, the Bayesian says, the initial values or "prior probabilities" you assign to them can be perfectly arbitrary, in fact may be arbitrary, but it doesn't really matter. In the parlance of the Bayesians, as more and more data come in, the prior probabilities will be "swamped," that is, when we use Bayes' theorem to "update" prior probabilities, i.e. feed new p(e)'s into the latest values for p(e/h) and p(e/h), the successive values of p(h/e) will converge on the correct value, no matter what initial values for these three variables we start with! Prior probabilities are nothing but measures of the individual scientist's purely subjective degree of belief before applying Bayes' theorem. In answer to our metaphysical question about what facts about the world probabilities report, prior probabilities report no facts about the world, or at least none about the world independent of our beliefs. In answer to the epistemological question of what justifies our estimates of probabilities, when it comes to prior probabilities, no more justification is needed or possible than that our estimates obey the axioms of probability.

There is no right answer or wrong answer as to what the prior probabilities of p(h), or p(e) are, so long as the values of these probabilities obey the rules of probability and logical consistency on betting. Logical consistency simply means that one places one's bets—that is, assigns strengths to one's degrees of belief—in such a way that bookies can't use you for a money pump: someone is a "money pump" for book-makers when they can be made to bet in ways that are irrational: for example, betting that team A will beat team B, team B will beat team C, and team C will beat team A. A series of such bets by one person guarantees that the bookies will make money no matter what team wins. One thing you must do to be sure that you are not a money pump, is apportion your bets in accordance with the axioms of probability theory.

Another theorem of the probability theory shows that if we apply Bayes' theorem relentlessly to "up-date" our prior probabilities as new evidence comes in, the value of p(h) all scientists assign will converge on a single value no matter where each scientist begins in his or her original assignment of prior probabilities. So not only are prior probabilities arbitrary but it doesn't matter that they are! Some scientists may assign prior probabilities on considerations like simplicity or economy of assumptions, or similarity to already proven hypotheses, or symmetry of the equations expressing the hypothesis. Other scientists will assign prior probabilities on the basis of superstition, aesthetic preference, number worship, or by pulling a ticket out of a hat. It doesn't matter, so long as they all conditionalize on new evidence via Bayes' theorem.

It is not much of an objection to this account of scientific testing that scientists actually offer good reasons for their methods of assigning of prior probabilities. To begin with, Bayesianism doesn't condemn these reasons, at worst it is silent on them. But if features like the simplicity of a hypothesis or the symmetry of its form do in fact increase its prior probability, this will be because a hypothesis having features like this will, via Bayes' theorem, acquire a higher posterior probability than other hypotheses with which it is competing that lack these features. More important, attempts to underwrite the reasoning of scientists who appeal to considerations like economy, simplicity, symmetry, invariance, or other formal features of hypotheses, by claiming that such features increase the objective probability of a hypothesis, come up against the problem that the only kind of probability that seems to make any sense for scientific testing is Bayesian subjective probability. Furthermore, so understood, some Bayesians hold that probabilities can after all deal with some of the traditional problems of confirmation.

One of the major problems confronting Bayesianism, and perhaps other accounts of how evidence confirms theory, is the "problem of old evidence." It is not uncommon in science for a theory to be strongly confirmed by data already well known long before the hypothesis was formulated. Indeed, as we will see in Chapter 14, this is an important feature of situations in which scientific revolutions take place: Newton's theory was strongly confirmed by its ability to explain the data on which Galileo's and Kepler's theories were based. Einstein's general theory of relativity explained previously recognized but highly unexpected data such as the invariance of the speed of light and the precession of the perihelion of Mercury. In these two cases p(e) = 1, p(e/h) is very high. Plugging these values into Bayes' theorem gives us

$$p(h/e) = \frac{1 \times p(h)}{1} = p(h).$$

In other words, on Bayes' theorem the old evidence does not raise the posterior probability of the hypothesis—in this case Newton's laws, or the special theory of relativity, at all. Bayesians have gone to great lengths to deal with this problem. One stratagem is to "bite the bullet" and argue that old evidence does not in fact confirm a new hypothesis. This approach makes common cause with the well established objection to hypotheses that are designed with an eye to available evidence. Scientists who construct hypotheses by intentional "curve fitting" are rightly criticized and their hypotheses are often denied explanatory power on the grounds that they are *ad hoc*. The trouble with this strategy is that it doesn't so much solve the original Bayesian problem of old evidence as combine it with another problem: how to distinguish cases like the confirmation of Newton's and Einstein's theories by old evidence from cases in which old evidence does not confirm a hypothesis because it was accommodated to the old evidence. The alternative approach to the problem of old evidence is to supplement Bayes' theorem with some rule that gives p(e) a value different from 1. For example, one might try to give p(e) the value it might have had before e was actually observed in the past, or else try to rearrange one's present scientific beliefs by deleting e from them and anything which e makes probable; then go back and assign a value to p(e), which presumably will be lower than 1. This strategy is obviously an extremely difficult one to actually adopt. And it is (subjectively) improbable that any scientist consciously thinks this way.

Many philosophers and scientists who oppose Bayesianism do so not because of the difficulties that are faced by the program of developing it as an account of the actual character of scientific testing. Their problem is with the approach's commitment to subjectivism. The Bayesian claim that no matter what prior probabilities the scientist subjectively assigns to hypotheses, their subjective probabilities will converge on a single value, is not sufficient consolation to opponents. Just for starters, values of p(h) will not converge unless we start with a complete set of hypotheses that are exhaustive and exclusive competitors. This seems never to be the case in science. Moreover, objectors argue, there is no reason given that the value on which all scientists will converge by Bayesian conditionalization is the *right* value for p(h). This objection of course assumes there is such a thing as the right, i.e. the objectively correct probability, and so begs the question against the Bayesian. But it does show that Bayesianism is no solution to Hume's problem of induction, as a few philosophers hoped it might be.

And the same pretty much goes for other interpretations of probability. If sequences of events reveal long-run relative frequencies that converge on some probability value and stay near it forever, then we could rely on them at least for betting odds. But to say that long-run relative frequencies will converge on some value is simply to assert that nature is uniform, that the future will be like the past, and so begs Hume's question. Similarly, hypothesizing probabilistic propensities that operate uniformly across time and space, also begs the question against Hume's argument. In general probabilities are useful only if induction is justified, not vice versa.

There is a more severe problem facing Bayesianism. It is the same problem that we came up against in the discussion of how to reconcile empiricism and explanation in theoretical science. Because empiricism is the doctrine that knowledge is justified by observation, in general, it must attach the highest probability to statements that describe observations, and lower probability to those which make claims about theoretical entities. Since theories explain observations, we may express the relation between theory and observation as (t and t → h)—where t is the theory and t → h reflects the explanatory relation between the theoretical claims of the theory, t and an observational generalization, h, describing the data that the theory leads us to expect. The relation between t and h may be logical deductive, or it may be some more complex relation. But p(h) must never be lower than p(t and t → h), just because the antecedent of the latter is a statement about what cannot be observed whose only consequence for observation is h. Bayesian conditionalization on evidence will never lead us to prefer (t and t → h) to h alone. But this is to say that Bayesianism cannot account for why scientists embrace theories at all, instead of just according high subjective probability to the observational generalizations that follow from them.

Of course, if the explanatory power of a theory were a reason for according it a high prior probability, then scientists' embracing theories would be rational from the Bayesian point of view. But to accord explanatory power such a role in strengthening the degree of belief requires an account of explanation. And not just any account. It cannot for example make do with the D-N model, for the principal virtue of this account of explanation is that it shows that the explanandum phenomenon could be expected with at least high probability. In other words, it grounds explanatory power on strengthening probability, and so cannot serve as an alternative to probability as a source of confidence in our theories. To argue, as seems tempting , that our theories are explanatory in large part because they go beyond and beneath observations to their underlying mechanisms is something the Bayesian cannot do.

Summary

Empiricism is the epistemology that has tried to make sense of the role of observation in the certification of scientific knowledge. Since the seventeenth century, if not before, a tradition of English-speaking philosophers like Hobbes, Locke, Berkeley, Hume and Mill have found inspiration in science's successes for their philosophies, and sought philosophical arguments to ground science's claim. In so doing, these philosophers and their successors set the agenda of the philosophy of science and revealed how complex is the apparently simple and straightforward relation between theory and evidence.

But empiricists have never been uncritical in their assessment of scientific methods and of the epistemological warrant of its claims. We saw some of these problems in connection with the problem of the meaning of theoretical terms and scientific realism in previous chapters. Here we have explored another problem facing empiricism as the official epistemology of science: the problem of induction, which goes back to Hume, and added to the agenda of problems for both empiricists and rationalists.

In the twentieth century the successors of the British empiricists, the logical positivists, or logical empiricists as some of them preferred to be called, sought to combine the empiricist epistemology of their predecessors with advances in logic, probability theory, and statistical inference, to complete the project initiated by Locke, Berkeley and Hume. In particular, philosophers have appealed to Bayes' theorem, a result provided at the time Hume formulated his problem of induction, to help understand how evidence supports hypotheses in science. We have seen that appealing to probability is by no means unproblematic. In fact it raises its own problems along with any it might put to rest. We will encounter more of these problems in the next chapter. The problems empiricism faces in illuminating the epistemology of science continue to mount even as the availability of plausible alternatives declines.

Study Questions

1. Discuss critically: "Lots of scientists pursue science successfully without any regard to epistemology. The idea that science has an 'official one,' and that empiricism is it, is wrong-headed."
2. Why would it be correct to call Locke the father of modern scientific realism and Berkeley the originator of instrumentalism? How would Berkeley respond to the argument for realism as an inference to the best explanation of science's success?
3. What should the relationship be between the ordinary concept of evidence as for example employed in law courts and the scientist's use of the term in the testing of general theories?
4. Defend the claim that there are several different but compatible meanings of the word "probability" in science. Is one of them more fundamental than the others?
5. What do you need to add to Bayes' theorem to solve the problem of induction?

Suggested Readings

Empiricism is often thought to officially begin with John Locke's *Essay on Human Understanding*. George Berkeley's *Principles of Human Knowledge* is brief but powerful. The last third develops an explicitly instrumental conception of science that he contrasts to Locke's realism. Berkeley argued for idealism—the thesis that only what is perceived exists, that the only things we perceive are ideas, that therefore only ideas exist. His argument turns on the very same theory of language that the logical positivists initially embraced: the meaning of every term is given by the sensory idea it names. About Berkeley's work, Hume wrote "it admits no refutation, and carried no conviction" in his *Enquiry Concerning Human Understanding*. In this work he develops the theory of causation discussed in Chapter 2, the theory of language common to empiricists from Berkeley to the logical positivists, and the problem of induction. Bertrand Russell's famous paper, "On Induction," reprinted in Balashov and Rosenberg, brought Hume's argument to center stage in twentieth-century analytical philosophy.

J. S. Mill, *A System of Logic*, carried the empiricist tradition forward in the nineteenth century, and proposed a canon for experimental science still widely employed under the name, Mill's methods of induction. The physicist Ernst Mach, *The Analysis of Sensation*, embraced Berkeley's attack on theory as empirically unfounded against Ludwig Boltzman's atomic theory. This work was greatly influential on Einstein. In the first half of the twentieth century logical empiricists developed a series of important theories of confirmation, R. Carnap, *The Continuum of Inductive Methods*, H. Reichenbach, *Experience and Prediction*. Their younger colleagues and students wrestled with these theories and their problems.

W. Salmon, *Foundations of Scientific Inference*, is a useful introduction to the history of confirmation theory from Hume through the positivists and their successors. D. C. Stove's *Hume, Probability and Induction* attempts to solve the problem of induction probabilistically.

L. Savage, *Foundations of Statistics*, provides a rigorous presentation of Bayesianism, as does R. Jeffrey, *The Logic of Decision*. A philosophically sophisticated presentation is P. Horwich, *Probability and Evidence*. An introduction to Bayesianism is to be found in Salmon's *Foundations of Scientific Inference*. Salmon defends the application of the theorem to cases from the history of science in "Bayes' theorem and the history of science," reprinted in Balashov and Rosenberg. Richard Swinburne, *Bayes' Theorem*, collects several recent papers on the theorem and its upshot. Important papers on Bayes' theorem and scientific change by Salmon are reprinted in Lange, and in Curd and Cover's anthology.

The problem of old evidence, among other issues, has led to dissent from Bayesianism by C. Glymour, *Theory and Evidence*. One of his papers on the subject, "Explanation, Tests, Unity and Necessity," is reprinted in Lange, and "Why I Am Not a Bayesian" is reprinted in Curd and Cover.

Peter Achinstein, *The Concept of Evidence*, anthologizes several papers that reflect the complexities of inference from evidence to theory, and the relationship of the notion of evidence to concepts of probability.